UNIVERSITY OF SKÖVDE

VS-GAMES 2015

7th International Conference on Games and Virtual Worlds for Serious Applications

16–18 September 2015
Skövde, Sweden

IEEE

Catalog Number: CFP1538G-ART (Xplore)
CFP1538G-USB (USB)
ISBN: 978-1-4799-8102-1 (Xplore)
978-1-4799-8101-4 (USB)

2015 2015 IEEE 7th International Conference on Games and Virtual Worlds for Serious Applications (VS-Games)

IEEE Catalog Number: CFP1538G-ART (Xplore)
ISBN 15: 978-1-4799-8102-1 (Xplore)

IEEE Catalog Number: CFP1538G-USB (USB)
ISBN 15: 978-1-4799-8101-4 (USB)

Printed copies of this publication are available from:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com

Picture on the front cover:

The restaurant and view of Skövdes´ Boulogner park in three different techniques.

1910 - Ludvig Ericson. Black-White (large format negative)

2015 – Torbjörn Svensson, Digital color photography (Nikon D700)

2014 – Max Mellhage, computer visualization of Boulogner park (Unreal Engine)
+ hand cursor by Torbjörn Svensson

Montage by Torbjörn Svensson.

Produced by IEEE eXpress Conference Publishing

For information on producing a conference proceedings and receiving an estimate, contact conferencepublishing@ieee.org
http://www.ieee.org/conferencepublishing

VS-Games 2015 Preface:

Welcome to the Proceedings of VS-Games 2015, the seventh International Conference on Games and Virtual Worlds for Serious Applications, held at University of Skövde (Sweden) from the 16th to the 18th of September 2015. We welcome all conference delegates to Skövde, one of the major game development hubs in Sweden. The university has some 500 students, studying all aspects of game development, from technology and programming to design and artwork. University of Skövde offers study programs from undergraduate and graduate level as well as PhD study program in informatics, with a possibility to specialize in computer games. In addition to this, research is carried out in two different research groups, focusing on different aspects of computer game development. Finally, Gothia Science Park offers an incubator program for new studios.

The seventh IEEE International Conference in Games and Virtual Worlds for Serious Applications (VS-Games'15) aims to meet the significant challenges of the cross-disciplinary community that work around non-leisure applications of games and game technologies by bringing the community together to share case studies of practice, to present virtual world infrastructure developments, as well as new frameworks, methodologies and theories relevant to our community. As the field has been growing there are more and more reports (scientific as well as more anecdotal) about the use and usefulness of serious games. As a result there is a clear need to consider new frameworks, theories, methods and design strategies for making serious games and virtual world technologies more effective and useful. Furthermore, we also hope for a closer connection between the entertainment games community and the serious games community. Therefore, we arrange VS-Games 2015 back-to-back with the Swedish Game Conference in Skövde. By doing this, we hope to stimulate communication between representatives from both communities.

We had a total of 48 submissions for VS-Games 2015, out of these 32 were full papers, 11 short papers and 4 posters. We have accepted 21 full papers, 9 short papers and 4 posters to be included in the proceedings. The accepted papers cover a range of topics in serious games and virtual worlds. The conference will feature topics such as affective serious games, tools and tool support for serious games development and virtual worlds, gamification, digital storytelling and games for learning and training. Cases and application areas include healthcare, military, organizational development and citizen engagement.

Organizing a scientific conference such as VS-Games requires a lot of support and effort from a number of people and organisations. First of all, thank you to the steering committee of VS-Games who trusted us with the organization of the conference: Professor Sara de Freitas, Dr Fotis Liarokapis and Dr Genaro Rebolledo Méndez. Many thanks also go to the 52 members from 20 different countries in the International Program Committee, who saw to it that each paper had 3 independent reviews. Thank you for your effort.

Thanks are also extended to the IEEE Computer Society (in particular to Shu-Ching Chen, Chair of Technical Committee on Multimedia Computing) for their technical co-sponsorship of the conference, which is now becoming a tradition for VS-Games. Also Christina Zarrello Project Editor for IEEE eXpress Conference Publishing has been of considerable assistance across a number of months in the lead-up to the conference, helping out with organizing the conference proceedings and their publication by IEEE.

University of Skövde and Sweden Game Arena are acknowledged for being the lead sponsor of the event. Many people assisted with the organization of the event behind the scenes, so thanks to all who helped out with a multitude of administrative and practical tasks so that this event could become a reality.

Finally, we appreciate the value coming from the fact that selected authors will be invited to submit extended versions of their work to Visual Computer and Journal of Computing in Higher Education.

Skövde (Sweden) and Brno (Czech Republic), September 2015

Per Backlund (General Chair)

Henrik Engström (General Co-Chair)

Fotis Liarokapis (Steering Committee Representative)

Committees

General Chairs

- Dr Per Backlund (University of Skövde)
- Dr Henrik Engström (University of Skövde)

Technical Program Committee

- Dr Per Backlund (University of Skövde)
- Dr Henrik Engström (University of Skövde)
- Dr Christos Gatzidis (Bournemouth University)

Publication Chairs

- Dr Fotis Liarokapis (Masaryk University, Czech Republic)
- Dr Kurt Debattista (Warwick University, UK)

Steering Committee

- Sara de Freitas (Murdoch University, Australia)

- Fotis Liarokapis (Masaryk University, Czech Republic)
- Genaro Rebolledo Méndez (Universidad Veracruzana, Mexico)

International Program Committee

- Alexiei Dingli , University of Malta
- Ali Arya, Carleton School of Information Technology, Canada
- Anastasios Bakaoukas, University of Northampton, UK
- Anna-Sofia Alklind Taylor, University of Skövde, Sweden
- Athanasios Vourvopoulos, University of Madeira
- Brian Goldiez, University of Central Florida, USA
- Christopher Peters, KTH, Sweden
- Christos Gatzidis , Bournemouth University, UK
- David Panzoli, CUFR Jean-François Champollion, France
- Dimitrios Christopoulos, Foundation of the Hellenic World, Greece
- Eike Anderson, Bournemouth University, UK
- Feng Tian, Bournemouth University, UK
- Fotis Liarokapis, Masaryk University, Czech Republic
- Francisco José Gallego Durán , University of Alicante, Spain
- Frank Dignum, Utrecht University, The Netherlands
- Geoffrey Hamon, Howest,University College West Flanders, Belgium
- Georgios Papaioannou , Athens University of Economics and Business, Greece
- Henrik Engström, University of Skövde, Sweden
- Igor Mayer, Delft University of Technology, The Netherlands
- Ioannis Paliokas, Democritus University of Thrace, Greece
- Iro Voulgari, University of Patras, Greece
- Jana Rambusch, University of Skövde, Sweden
- Jian Chang , Bournemouth University, UK
- Jiří Chmelík, Masaryk University, Czech Republic
- Katerina Mania, Technical University of Crete, Greece

- Kostas Karpouzis, National Technical University of Athens, Greece
- Krzysztof Walczak, Poznan University of Economics, Poland
- Kurt Debattista, Warwick University, UK
- KyungSik Kim, Hoseo University, Rep. of Korea
- Leonel Morgado, Universidade Aberta, Portugal
- Marc Conrad, University of Bedfordshire, UK
- Marc McMahon, Edith Cowan University, Australia
- Marcello Carrozzino, PERCRO - TeCIP Institute, Italy
- Mark Billinghurst, University of Canterbury , New Zealand
- Maria Saridaki, University of Athens, Greece
- Matthew Montebello ,University of Malta
- Maurice Hendrix, University of Northampton, UK
- Mika, Luimula, Turku University of Applied Sciences, Finland
- Minhua Ma, University of Huddersfield, UK
- Nelson Zagalo, University of Minho, Portugal
- Olga De Troyer, Vrije Universiteit Brussel, Belgium
- Panagiotis Petridis, Aston University, UK
- Per Backlund, University of Skövde, Sweden
- Sandro Spina, University of Malta
- Sara de Freitas, Murdoch University, Australia
- Sascha Fagel, zoobe message entertainment GmbH, Germany
- Selma Rizvic, University of Sarajevo, Bosnia and Herzegovina
- Sofia Pescarin, Virtual Heritage Lab, Italy
- Stefan Göbel, Technische Universität Darmstadt, Germany
- Vanessa Camilleri, University of Malta
- Vedad Hulusić, Sarajevo School of Science and Technology, Bosnia and Herzegovina
- Vittorio Scarano, University of Salerno, Italy

Table of Contents

Full Papers

Short Papers

Poster Papers

A study on the collective perceived representation of a real urban area through the usage of an Engaging Framework, based on a 3D Virtual Environment and OpenStreetMap Data.

Marco Alessi, Stefania Castelli, Valentina Chetta, Enza Giangreco, Stefano Pino, Davide Storelli
Engineering-Ingegneria informatica S.p.A.
Lecce, Italy
marco.alessi@eng.it enza.giangreco@eng.it

Angelo Corallo, Laura Fortunato, Andrea Gentile
University of Salento
Lecce, Italy
angelo.corallo@unisalento.it

Abstract— *The need for urban regeneration does not come only by structural requirements, but also by socio-cultural needs. What we are going to propose is the urban regeneration as a way to perceive, in a different way, the surrounding spaces allowing users to receive and provide a wide range of information on the urban environment. Each space of a city has a variety of intrinsic meanings provided by human groups interacting with each other everyday. The purpose is collecting the hidden information thanks to citizens' contribution. The objective is the involvement of citizens as "builders of sense" through a playful attitude as "builders of virtual cities", and using game based on motivation as impetus for the regeneration. Urban regeneration is innovative thanks to a new participatory and cooperative methodology based on the perception of every citizen, and on the collection of players' experiences.*

Keywords—Urban regeneration; 3D application; serious simulation.

I. INTRODUCTION

Urban areas where citizens live are places where different human experiences coexist: tourists, citizens, workers or other types of people, etc. Those spaces are characterized by the presence of different social actors. As stated by Michael Batty, cities are evolving systems whose structure generates from the bottom [1]. For this reason, urban regeneration is a widely experienced but little understood phenomenon [2].

Towns and cities are changing over time, this is inevitable and can be seen as opportunity [2]. This changing is necessary because the operations of the political, economic and social systems constantly generate new demands and offer new opportunities for economic progress and civic improvement. As Mumford argues, in the city, remote forces and influences intermingle with the local ones: the conflicts are no less significant than their harmonies [3].

P. Roberts and H. Sikes define urban regeneration as an outcome of the interplay between many sources of influence and as a response to the opportunities and challenges. Urban regeneration is a comprehensive and integrated vision and action, which leads to solve the urban problems generating improvements in the economic, physical, social and environmental condition of an area that has been subject to change [2].

The challenge we intend to address in this paper begins with the recovery of collective memories and aspirations, which enable the emersion of new characteristics and urban data generated by users' collective contribution. The abstract' boundaries of a real urban environment are defined by people interacting with the city. People interaction, in local area where city is settled, becomes an effective collective good, where scattered elements come together and people give meaning to the whole.

This paper intends to define and analyse the formation of this new relationship between social world, made by actors and practical actions, and the settled environment, made by a variety of symbols belonging to urban structures. The relationship created by the personal collective understanding is explained through the mechanics of a game like experience, leveraged by the use of a 3D application displaying a subset of real spaces. Then we will be able to outline new human processes, which can be grouped by a series of valuable indicators (sex, religion, age, etc.), providing us the opportunity, not only to facilitate urban regeneration, but also to investigate a new Urban Anthropology. This model is the conceptual intersection between the informal city, made by human possible interactions, and the physical city, made by practical constraints. Between those two aspects it is expected to find out a conflict, but at the same time identity and a virtual place are the best approach to measure a human creative attitude inside a more structured environment, where the expression of engaged players give a result.

The approach used here as engaging lever is game like, in order to address users' creativity and feedback, and carry out the research about the regeneration of urban areas through collective creativity. Collaborative phenomena are enhanced

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

addressing the discovery of self-emerging communities in the platform designed within the proposed framework. The intent is to leverage personal and societal perceived representation of spaces, real or emotional ones, and make use of the aggregated data, expressed by the community(ies) discovered in the network of the platform. The data obtained can be used both to regenerate real urban spaces, and collect information in order to make geographic inference based on personal and group perception of urban environments.

The conceptual and technological framework proposed in this paper starts from the distinction between places and non-places of the sociologist Augè [4]. Places are spaces with identity and value, both for the individual and for the community. The non-places are those spaces, which have arisen with the advent of capitalism and globalization. Spaces such as stations, airports, bus stops, shopping malls are considered as the non-places. These are anonymous spaces, not very different in structures, areas of passage fluctuating and alienating [4].

II. PROBLEM

According to the F. Choay, in the history of Europe, the urban morphology of cities has changed over time segmenting different ideal-typical modes, which correspond to the visions of the different social actors [5]. Each mode has different effects on humans, responding to the basic need of security and orientation [6].

Moreover, as stated by Hisham G. Elshimy, over time, urban regeneration has evolved from a simple form of renovation or rehabilitation of obsolete infrastructure to a physical, economical, social and environmental change in a city [7].

There is the need to consider the urban places as relational, open, porous, and hybrid spaces, whose specificity derives not only from structural elements, but by the dynamism of its inhabitants [8].

Therefore, the only possible regeneration would be the one that produces value from the bottom, which gives voice to citizens and residents that enable the community to surpass the conditions of marginality and anonymity [9].

An effective urban regeneration should be based on the participation and co-operation of a wide range of actors and stakeholders including citizens as new actors [10].

The practical problem related to the research questions is: How can the public administration involve all stakeholders, and in particular citizens, in urban regeneration?

To answer this question we have developed a methodological and technological framework able to involve citizens as a "builders of sense", through a playful attitude as "builders of virtual cities", and using game-based motivation as impetus for the regeneration.

III. BACKGROUND

As already indicated in the introduction, our research domain aims to provide a new methodological and technological approach in order to support collective urban regeneration. For this purpose, some important definitions and research areas were analysed.

A. The concept of space and place

Tim Cresswell argues that “place is not just a thing in the world but a way of understanding the world” [11]. For Yi-Fu Tuan “What begins as undifferentiated space becomes place as we get to know it better and endow it with value. The ideas 'space' and 'place' require each other for definition. From the security and stability of place, we are aware of the openness, freedom, and threat of space, and vice versa. Furthermore, if we think of space as that which allows movement, then place is pause; each pause in movement makes it possible for location to be transformed into place” [12].

The political geographer John Agnew (1987) has outlined three fundamental aspects of place: a point on the earth’s surface (Location); the locus of individual and group identity (Locale); and the scale of everyday life (Sense of place) [32].

The place concept is in contrast with the non-place concept theorized by M. Augè (1995). He, in fact, distinguishes between “place encrusted with historical monuments and creative of social life and non-place to which individuals are connected in a uniform manner and where no organic social life is possible” [4].

Reconnect urban regeneration to the concept of non-place becomes important to restore the perception of space as a place with meaning, identity and memory. Each space is the result of signs. In order to turn a space into a place, it is needed to interpret, comprehend and live that particular space. The urban regeneration explained as transformation of space into a place, stimulating the concept of living that is no longer referred to the domestic space.

The German philosopher Heidegger M. defines the living as the persistence in a generally shared public life with a community. It is therefore a condition in which the social aspect and identity is essential. [13]

B. Game Like Experience

Real word objectives can be seen as a sequence of challenges, quests and levels, with a badge awarded in the form of social-recognition or incentives. Since 2009, Volkswagen has released a series of social tests to demonstrate that making funny things, people’s behaviour can change for the better (the fun theory) [14].

Recently it has been affirmed the idea of using mechanics and methods taken from games, in order to enhance the experience of users during their interactions with the applications. Game like experiences were proven an important innovation in not game and leisure contexts: we have experienced the aftermath of Serious Games[15] for training (in a wide range of cases, from soldiers[16] to health and sanitation training[17] and education, and the increasing use of Gamification in non-leisure and non-game contexts[18]. Nowadays the game like approach is being used in full-blown games or common platforms, where the final purpose is not the entertainment but training, users' education, marketing, advertisement, or to reach some important goals, which may be useful or meaningful for the users (i.e. fold.it and scientific research)[1].

The collaborative and social aspect of the gaming experience is well known in analyses of leisure games, where a specific role is assigned to the creation and commitment of in-game communities [34], which can lead also to trust creation phenomena, particularly in MMORPG [35]. The same considerations have started to be performed also for non-leisure gaming experiences [36]. These self-organization phenomena can be specifically addressed and encouraged as explained below. Finally, from the decision maker perspective, it is important to pay attention to those (sub)communities, which are solicited - whenever a novel administrative experiment is performed. This criticality has been since long recognized, for example in development projects [37].

C. Serious simulations and 3D environment

Over the time, 3D Virtual Environments have been used largely to represent urban contexts. It is a fundamental layer over which urban games are developed, for many purposes, as many as the representations of a city may turn useful in a sense of playful or non-leisure attitude. Though 3D urban aesthetics have been changing in time, not always driven by the need or realism, but much more that of effective interaction, in the field of gaming applications, aesthetic realism is set apart from realism of representation [19]. While the virtual realism is important in serious simulations to achieve their purpose [20], other mediums (i.e. for entertainment) require a different type of realism that is not entirely dependent on the mimetic representation [21].

The mimetic representation of the urban context has changed in time, following the evolution of technology and players' tastes, from the first 2D and the later isometric approach, till the latest 3D environment. The first city building game was Maxis's SimCity in 1989, devised by sims legend Will Wright.

Many experiments involving serious games, social goodness, and the relationship between citizens and government offices have been developed using the available technologies and governance offices. Companies moved on to use the state of the art technologies involving the browser-based city building games, creating important applications like runthattown [26], produced by the Australian Bureau of Statistics or a Flash game created by New Zealand energy company Genesis Energy [22] or cityone [28] produced by IBM.

Browser-based and most of the times flash game engines and isometric layout city building have fitted the scope till late 2012, when the simplicity of the isometric visualization could have been overtaken by the extended use of browser-based on advanced technologies, like WebGL. New demands involving interactions and aesthetics lead to investigate new ways of map representations, non-tile based, with a more realistic structure. ViziCity [29] and Osm2World [30] launched their projects in order to give an immersive 3D shape from real urban data for greater mass audiences within a browser-based experience.

In the art of computer game design [24] it is highlighted that the nature of human fantasy can turn an objectively unreal situation into a subjectively real situation, which in essence may indicate that the perceived realism of a virtual scene is a highly important aspect that is set apart from the virtual realism.

Further studies in gaming applications for casual interaction, may outline the trade-off represented by realistic outfits affecting the state of flow, which engage the players in the game, balanced between anxiety of a complex environment and the lack of motivation and boredom given by the excess of simplicity [23]

IV. SOLUTION

A conceptual and technological framework characterizes the solution proposed in this paper in order to involve citizens in an urban regeneration process. The approach described here used an engaging environment in order to address users' creativity, and feedback is game like, similar to the developing of a serious game.

A. Conceptual Framework

Cities are not only spaces, but are socio-political structures, inside which there are norms, values, practices and cultures. From this point of view, the co-production of public space considered as common good is a practice of invention that defines the urban identity.

Therefore, considering urban regeneration, it is not possible to separate the geographical aspect of a place from the anthropological and cultural one, because this split generates the so-called "non-place". Starting from these assumptions, the proposed framework (Fig. 1.) aims to transforms non-place in place through a game like approach, allowing the personalization of urban places and the subsequent regeneration.

The framework is characterized by three layers:

Geographic space is an extension of the terrestrial surface with physical attributes, in which humanity has exercised a work of editing and processing. In other word, the geographical space of a city or territory includes not only the geographical extension, but also roads, palaces, gardens, schools, etc., present in that space and realized over the years. In the proposed framework, the geographical space is the 3D environment that logically reproduces the city or the territory in question and allowing users to navigate in space and visually recognize the places of everyday life.

[1] http://fold.it/portal/

Personal space: This level represents the space experienced by each citizen that consists of memories, experiences, family histories, folktales, etc. The combination of the geographical space with the personal space, allows individuals to customize the places where they lives, giving them a deep meaning. For this reason, we are proposing a tool allowing users to enrich the geographic space with the insertion on map of personal contents related to city places.

Anthropological-cultural space: the environment becomes territory, district and community thanks to the action of man with his culture, values, social structures and economic strategies. Social re-appropriation of "non-places" goes through the creation of a general knowledge, which can create socio-cultural cluster. The interaction and the intersection of all personal spaces can create larger spaces, the anthropological spaces [33], which have the same cultural interests, artistic expressions and crafts, customs and popular beliefs.

Through this framework the process of urban regeneration becomes not a mere exercise of urban beautification, but a deep process of enhancement of the social, cultural and structural places. Urban regeneration becomes **"mnemotop"**[2] dynamic or rather territory/district fixing in its structure elements of the collective cultural memory and innovation, for this dynamic [25].

Fig. 1. *Conceptual Framework*

B. *Technological Framework*

The main purpose for developing the urban simulation in 3D is to create a platform where it is possible to represent the real urban data in a more immersive manner. The gameful approach can leverage the engagement of the end users in a modern way. Since the primary aim is to explore human perception and how players may be involved in the platform, through an application of city building, the platform should provide a good balance between visualization and interaction of data. A gaming application can give us the opportunity to interact with the end users, using a virtual environment, where actions extended and gameplay can establish a stream of information, useful in different ways:

a) Urban Regeneration; city municipalities could use a virtual reality city as an interactive platform over which collect spread proposals involving regeneration of collective areas. In case of scarcity of resources, it is possible to activate players' creativity using the same mechanism already presented in serious/leisure games.

b) Display of local needs; the visual representation, thanks to the spatial coherence of the 3D environment, can be used as a tool to display local needs, expressed by citizens, in a more engaging experience.

c) Visual Representation of Personal Perceived Urban Space; city areas are different from city to city, and those may be perceived as different places by different human groups, based on culture, sex, age, religion and many other factors. The platform may enable a distributed display of personal spaces, useful in order to investigate how spaces are perceived. New services connected to different users' perception may be developed, also, municipalities can take advantage of this personal display in order to represent the same space in more useful ways (i.e. for tourism).

The development of the platform is conducted involving browser-based on graphics WebGL, PHP and Javascript Programming on the client side, database building in a PostgreSQL dbms on the server side, with more services developed on the server machine as bash scripts. At architectural level, the infrastructure supporting the framework will be represented by three layers as following:

a) ***Visual Layer*** *(3D Environment);* the challenging aspect for the visual representation of the urban areas is the optimized display of real geographic features in a virtual space. The approach has been possible using real data received from open source repositories, like Open Street Maps. In this case, the openness of the data in the retrieval of spatial information has been central in the development, and the Open Street Maps Community provides an interesting data source. The limit of the Open Street Map approach is the lack of geographic data in certain areas, where users of the community have been less active. OpenStreetMaps has shown to be a valuable source for archiving and representing geospatial information, thanks to the open source and the vast number of volunteers contributing to this project. Nonetheless, the standard OSM format may contain estimation errors and rapid changes in short periods of time. The geometry for architecture within the virtual city is rendered using WebGL. In particular, it has been made extensive use of Javascript libraries, the most important is the ThreeJS, which has been a valuable resource in order to create a virtual scene and manage the insertion of 3D models inside the scene. The urban representation is divided into two different layers: the first one is available urban data retrieved by open street maps; the second is made by 3D models, usually in Wavefront or Json format. In order to support the construction of the urban Wavefront model it has been created a server process based on bash scripts which uses Osm2world python functions, in order to translate XML data in OSM format to a Wavefront Model.

[2] The *mnemotop is the landscape fixing in its structure elements of the collective cultural memory.*

The use OSM resources may cause lacks of accuracy in building footprints, and the absence of a valid method to value heights [43] [44], apart from for the evaluation of floors. Still those are useful, depending on the scope of the application, making a distinction here between simulations as a mere reproduction of the existing things, and a game like experience, where the level of detail is negotiable in the exchange with the usability of the application, and the sense of authonomy of the end user. Another limit of the use of OSM data, as single resource of geospatial data, is the lack of a model which may represent the elevation of terrain (Digital Elevation Model).

*b) **Content Management Layer** (images, 3D Items, constraints);* the scene initially is created using Open Street Map Data, and then is inserted and visualized in a Virtual Scene using ThreeJS libraries. This is the beginning of the user experience, which can be enhanced leveraging on the features available in the virtual environment. It is important to remember that the contents populating the virtual environment are a very important part of the user experience, especially for motivational and engagement aspects. User features from which the user can choose how to customize virtual world, deal directly with the feeling of Autonomy in the Self Determination Theory [27]. It is useful to have in the platform different types of visual models, with the same informational meaning. Following this approach, for example, different types of building may have the same informational values, but different layouts. Additional information in the virtual reality may be represented by:

- *Urban Models*; different models reproduce the various assets in the urban city customization, as well as the level of personalization reached by the players, in order to populate the 3D map. Different urban models or different classes of models, may represent different urban assets like Education (Schools), street lights (lamp posts), public or social housing (residential buildings), public green (trees and green areas), public mobility (bus stop) and more "Fig 2".

- *Personal Features;* personal features may also be added in the virtual environment, over the 3D initial map. This can enhance the level of affection and personalization bounding the player and the platform. Personal Features are different from urban models because they do not represent a real structure, which can be built in the area, but a service or a personal aspect like images, videos or emotional states recognized as coherent within the area (i.e. cultural heritage, emotions, local food specialties, beautiful landscapes and more). A further investigation has to be made in order to provide a better visualization of personal features in an urban virtual space.

- *Constraints;* constraints are present in every game. They have always been an important part of the gameplay, because they provide realism on one hand and challenge on the other. Constraints may come in two ways: resource constraints, and spatial constraints. Resource constraints (e.g. money, time, energy, quantity, etc.) have been presented in the city building games since their appearance, the gaming mechanic is clear, and it is a source of engagement itself, because it affects directly the sense of competence of the city builder [27]. In this case, the use of realistic resources can enable players' creativity in order to build realistic environments, using the available resources in the municipality. The spatial constraints deal with the existing laws in public building. In some areas, some features may be or not inserted. This happens with the use of law constraints and providing player with effective information. In our 3D model those constraints are available but not fed by real low databases. This functionality should be studied in accordance with the governance "Fig 3".

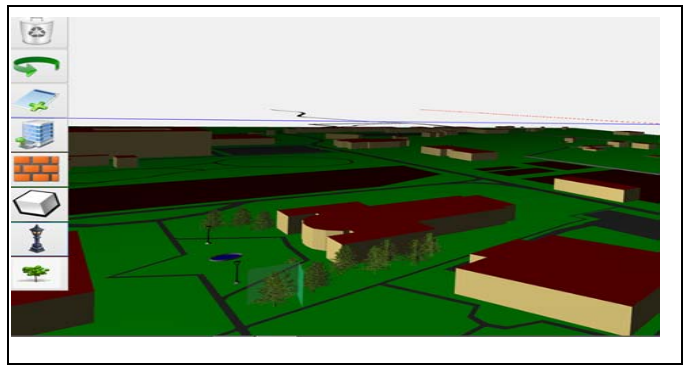


Fig. 2. *Visual Representation of a real area with trees added*

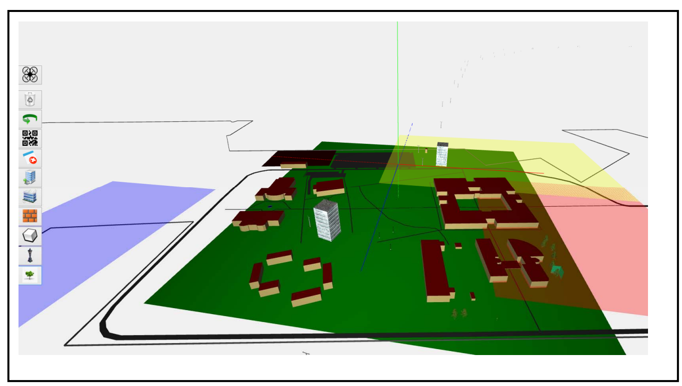


Fig. 3. *Representation of Spacial Constraints (coloured areas)*

*c) **Social Layer** (cooperation, sharing, rating);* Relatedness is an important intrinsic motivator. The social interaction or social participation in an common action is motivating and rewarding at the same time [27]. Both aspects are present in the development of the city building platform. In this case users may interact with other players in order to share the personal views or joining team of players in order to accomplish some social objective. Following social engagement framework proposed by Amy Jo Kim's [31], we can focus on the social actions useful to raise comment, like, greet, share, help, and give. Some of the previous engaging actions are being developed inside the framework.

- *Help*; co-operation can be engaging and a great intrinsic motivator in order to overcome a common challenge. Different players can come together in an urban area exchanging personal interactions with other users in order to display a shared vision of the same physical space. This can be done using an invitation mechanism.
- *Share*; sharing a particular personal map and its representation gives other players the vision of a single player; this may be interesting for different reasons. Enhancing the sense of expression, some players may have, during their customization, leveraging on the fun of exploration some users may find, and the fun of sharing, proper to the cooperative fun in Amy Jo Kim's engagement framework. Sharing can be accomplished by inviting other users in viewing personal virtual spaces, or by visiting those virtual spaces, the players allowed others to visit.
- *Rate*; rating may become a useful resource, showing personal views of physical space, and may produce interesting data on how virtual customizations of real places may be valuable for players and communities. Additional investigations are required in order to embed this functionality within the virtual space or in a different user interface (i.e. separate boards).

C. Social Network Analysis and Community Engagement

The gamification tool designed so far, gathers a considerable amount of data, and this paragraph provides a preliminary assessment of how to use them for two different purposes among others: at user level, to improve the interaction with the community; at management level, to provide a decision support system.

The first aims to increase the community engagement, by improving the collaborative features. In fact, the *help* and *share* features have been introduced, but if the number of users adopting the tool explodes, it becomes increasingly difficult for each user to find people with similar interests/attitudes/focus-area (if they are not already in contact before joining the framework). A few analytics approaches are summarized in Table I, providing solutions to "recommend" other users who may co-design the environment and giving a significant contribution, through the *help* or *share* collaborative design of the environment.

In this way, the aggregated data by any of the methods as shown in Table I, will be provided from an "egocentric" point of view, listing the most important users at the moment of logging in the framework platform.

For the second purpose, the clustering results can be also provided to government bodies as a bird's eye view of the communities' scenario. The peculiar characteristics of each community can be used both to analyse the perception of the environment, as a function of some parameters in the population, as well as to identify "expert roles" for specific areas (specifically referring to communities according to LBC methods).

TABLE I. USER-SIDE CLUSTERING FOR COMMUNITY ENGAGEMENT

Approach name	***Description***	***Comments***
Content-based clustering (CBC)	Data collected during the registration of the users (demographics, interests, …) are used to construct a network. The network is then clusterized according to community detection[38] or clustering algorithms[3].	The intricacy during the registration phase may produce superficial replies in the users. However, CBC can produce highly reliable communities in the presence of accurate descriptions.
Location-based clustering (LBC)	Same idea underlying the CBC approach, but the data used to derive the network are only those related to registered data for the user location, and/or the logfiles tracking the interaction of a user with a certain territory.	Basically reduces the intricacy of the registration phase, compared to the CBC approach.
Interactio n network clustering (INC)	*Rate* data[4] collected by the system are employed to evaluate the interactions among the users, mapped as (weighted) links of a network. This can be clustered as suggested in CBC.	Easier to compute the users'network compared to the CBC, querying the interaction data, especially if data are stored opportunely.
Content-based recommen der system (CRS)	Data collected along with the registration of the users are used to evaluate mutual similarities among them[41]. The users suggested for sharing/helping features will be ranked according to the highest similarity scores.	Robust method, compared to clustering approaches provides each user with a ranking for similar users inside the community of reference. Affected by same "warm-up" considerations as CBC.
Interactio n-based recommen der system (IRS)	Same data used for the INC are here employed to evaluate user-based similarities, in rating other projects.	Basic approach which can be highly efficient whenever the number of users grows to an extent hard to manage with clustering methods [42]. Heavily relies on the frequency of rating activities.

D. Conclusions and future work

The Engineering R&D group has already had field tests, in accordance with the Municipality of Lecce, moving innovative solutions for Public Administration into real settings. We have demonstrated the possibility to engage the public decision makers, into sharing problems and common solutions, at a local and regional level.

---

[3] The choice of the specific algorithm may be driven by considerations about the size of the network. The density envisaged for the final network may suggest to attempt clustering rather than community detection methods [39]. Whenever necessary, hybrid approaches may be invoked [40]

[4] Averaged, wherever multiple "rates" are available for the same couple of users.

A first version of 3D Virtual web based Environment platform has been already developed. It has presented during several meetings between Engineering R&D Group in Lecce and representatives of Lecce Municipality. The Lecce municipality has shown great interest in the proposed platform, giving space in the organization of a possible real scenario. A first setting for the field test seems to be the regeneration of "Quartiere Leuca", during which, group of Engineers and Architects will be involved. This stage will be useful in order to try different methods of data aggregation , and to provide additional insights about the possible development of the platform. As a possible future work, the development of additional functionalities are being considered, such as the inclusion in the system of resource constraints, which are the natural evolution of a urban management game. This additional functionality will require the feeding of adequate data, needed to evaluate the expenditure of resources in the placing of urban items. It is clear that, in order to reach an higher level of detail in the land representation and usage, it should be necessary to include data banks of surface elevation, similar to those provided by Shuttle Radar Mission Topography [45], in addition with building heights, which may be provided in accordance with federal, state and local authorities. An increase of detail may result in an increase of complexity and retardation of the entire process, which have to be avoided, where possible, so that the user experience remains "enjoyable".

REFERENCES

[1] J. M. Batty, “Cities as Complex Systems: Scaling, Interactions, Networks, Dynamics and Urban Morphologies”, UCL Working Papers Series, Paper 131 - Feb 08.

[2] P. Roberts, H. Sykes, Urban Regeneration–A Handbook, British Urban Regeneration Association, 2000.

[3] L. Mumford, The culture of cities, Hacourt Brace and Company, New York, 1938.

[4] M. Augé, Non-Places: Introduction to an Anthropology of Supermodernity, Verso, Londra-New York, 1995.

[5] F. Choay, The modern city: planning in the 19th century, Studio Vista, London, 1969.

[6] H. Lefebvre, La production de l’espace, Editions Anthropos, Paris, 1974.

[7] H. G. ELshimy, “Sustainability of Cites Built Environment by Urban Regeneration concept” Case study of the Historic Cairo, International Scientific Journal Architecture and Engineering.

[8] D. Massey, “A global sense of place”, Space, Place and Gender, University of Minnesota Press, Minneapolis, 1994. http://www.unc.edu/courses/2006spring/geog/021/001/massey.pdf

[9] A. Appandurai, Modernityat large: cultural dimensions of globalization, University of Minnesota press, Minnesota-London, 1996.

[10] F. Friesecke, The Role of Partnerships in Urban Regeneration–Similarities and Differences between Germany and United Kingdom, 2007.

[11] T. Cresswell, Place: a short introduction, Malden, Blackwell, 2004.

[12] Y. Tuan, Space and Place: The Perspective of Experience, University of Minnesota Press, Minneapolis, 1977.

[13] M. Heidegger, Identity and Difference, New York, Harper & Row, 1969.

[14] Volkswagen. (2009), The Funtheory.com, http://www.thefuntheory.com/, accessed June 2013.

[15] C. C. Abt, Serious Game, The Viking Press, New York, 1970.

[16] C. McLeroy (2008), History of Military Gaming, http://www.army.mil/ , accessed September 2014.

[17] F. W. Kron, et al., 2010. Medical student attitudes toward video games and related new media technologies in medical education. bMC Medical education, 10, www.biomedcentral.com/1472-6920/10/50, 28/08/15.

[18] S. Deterding, R. Khaled, E. Nacke, and D. Dixon, “Gamification : Toward a Definition.” Design: 12–15, 1, 2011.

[19] A. R. Galloway, “Social Realism in Gaming”, in The International Journal of Computer Game Research, Volume 4, Issue 1. 2004.

[20] A. Chalmers, “Level of Realism for Serious Games”, Proc. of the IEEE Int’l Conference on Games and Virtual Worlds for Serious Applications (VS-Games 09), IEEE Computer Society, 225-232, 2009.

[21] H. Sommerseth, “Gamic Realism: Player, Perception and Action in Video Game Play”, Situated Play, DiGRA Conference, 2007.

[22] Electrocity, 2015, http://www.electrocity.co.nz/, accessed october 2014.

[23] M. Csikszentmihalyi, Flow: The Psychology of Optimal Experience, Harper & Row, New York, 1990.

[24] C. Crawford, The Art of Computer Game Design, McGraw-Hill/Glencoe, New York, 1982.

[25] J.Assmann, Daskulturelle Gedachtnis: Schrift, Erinnerung und politische Identitat in fruhen Hochkulturen, C. H. Beck, Munchen, pp. 33-34, 1992.

[26] http://runthattown.abs.gov.au/ Accessed April 2015.

[27] Deci, E., & Ryan, R. (Eds.), (2002). Handbook of self-determination research. Rochester, NY: University of Rochester Press.

[28] IBM, City One Game, http://www-01.ibm.com/software/solutions/soa/innov8/cityone/index.jsp, accessed December 2014.

[29] Vizi Cities, http://vizicities.com/, accessed February 2015.

[30] OSM2World, http://osm2world.org/, accessed December 2014.

[31] http://amyjokim.com/2012/09/19/social-engagement-whos-playing-how-do-they-like-to-engage/ Accessed April 2015.

[32] N. J. Clifford, S. H. Holloway, S. P. Rice, G. Valentine, Key concepts in geography, SAGE Publications, London, 2009.

[33] M. Merleau-Ponty, Phenomenology of perception, Humanity Press, New York, 1962.

[34] Steinkuehler, C. A. & Williams, D. Where everybody knows your (screen) name: Online games as ‘third places’. *J. Comput.-Mediat. Commun.* **11,** 885–909 (2006).

[35] Park, S.-B. & Chung, N. Mediating roles of self-presentation desire in online game community commitment and trust behavior of Massive Multiplayer Online Role-Playing Games. *Comput. Hum. Behav.* **27,** 2372–2379 (2011).

[36] Ang, C. S., Zaphiris, P. & Wilson, S. Social interaction in game communities and second language learning. in *The 19th British HCI Group Annual Conference. Edinburgh, UK* (Citeseer, 2005). at <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.140.4158&rep=rep1&type=pdf>

[37] Mansuri, G. & Rao, V. Community-Based and -Driven Development: A Critical Review. *World Bank Res. Obs.* **19,** 1–39 (2004).

[38] Fortunato, S. Community detection in graphs. *Phys. Rep.* **486,** 75–174 (2010).

[39] Schaeffer, S. E. Graph clustering. *Comput. Sci. Rev.* **1,** 27–64 (2007).

[40] Gentile, A. A., Corallo, A., Bisconti, C. & Fortunato, L. in *Trends and Applications in Knowledge Discovery and Data Mining* 549–561 (Springer, 2014). at <http://link.springer.com/chapter/10.1007/978-3-319-13186-3_49>

[41] Lops, P., De Gemmis, M. & Semeraro, G. in *Recommender systems handbook* 73–105 (Springer, 2011). at <http://link.springer.com/chapter/10.1007/978-0-387-85820-3_3>

[42] Herlocker, J. L., Konstan, J. A., Terveen, L. G. & Riedl, J. T. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst. TOIS* **22,** 5–53 (2004).

[43] Over M., Schilling A., Neubauer S., Zipf A., Generating web Based 3D City Models from OpenStreetMap: The Current Situation in Germany. Computers, Environment and Urban Systems. 2010.

[44] Marcus Goetz, Alexander Zipf, OpenStreetMap in 3D – Detailed Insights on the Current Situation in Germany, Proceedings of the AGILE'2012 International Conference on Geographic Information Science, Avignon, April, 24-27, 2012, ISBN: 978-90-816960-0-5

[45] http://srtm.usgs.gov/index.php

Back to Contents

The active instructor: Benefits and barriers to instructor-led serious gaming

Anna-Sofia Alklind Taylor
School of Informatics
University of Skövde
Skövde, Sweden
Email: anna-sofia.alklind.taylor@his.se
Telephone: +46 500–448331

***Abstract*—While there is a wealth of studies on the subject of serious games, the same cannot be said on the issue of teaching with games, especially in game-based learning settings with adult learners. Over the years, most research in this area has been focused on the 'active substance(s)' of games for learning, focusing mainly on characteristics of games, but often failing to take the whole context of game-based learning into consideration, such as the role(s) of the teacher. However, the past two or three years has seen a shift in focus from merely the game as an isolated artefact, to also include more discussions on how games can successfully be integrated into an educational setting, as well as challenges as pitfalls of which instructors need to be aware. This paper aims to outline the contemporary research on instructor-led serious gaming and its implications for the design of serious gaming environments.**

***Keywords*—Serious games; Game-based learning; Instructor-led serious gaming; Instructors; Facilitators**

I. Introduction

What constitutes 'good' practice in serious gaming? As the field is moving towards maturity, scholars have only recently begun to look at the needs and obstacles that instructors face when using games for non-entertainment purposes. This paper explores the use of serious games from an instructor perspective. More specifically, it aims to describe the roles of instructors and the benefits and barriers to instructor-led game-based learning environments. Research within the field of serious games has mostly focused on the learners' perspective, but little attention has been paid to what the instructors do and what challenges that entails.

Computer-aided and game-based instruction was initially seen as something that would, to a degree, replace teachers. For instance, in the 1970's, games were predicted to act as coaches or tutors in the future [1] and there are still efforts in creating algorithms that will perform certain task that are now performed by a human instructor (e.g. [2]–[5]). However, more and more scholars are starting to realise that serious games, as artefacts used for learning and training, cannot fully replace the instructors' tasks, but must rather be designed to support them. Thus, instructors form an important target audience in serious game development – not just as subject matter experts, but also as users and players of the game – with a different set of needs than the learners. Moreover, serious gaming involves more than in-game activities, it also involves actions and events that occur off-game. These activities must also be considered when designing and utilising games for learning and training. A main concern for both educators and scholars of serious games is whether or not games are beneficial for learning and, more to the point, if the quality of learning that they provide is such that they are a good investment for educational organisations. Lately, several scholars have called for a shift of focus from the game itself, to the practicalities and potential barriers for actually using serious games [6]–[11].

This paper aims to describe the current state in research on instructor-led serious gaming and its implications for the design of serious gaming environments. It brings together a comprehensive set of sources which, to varying degrees, discuss instructor roles and the practicalities of planning and carrying out learning activities involving games and game technology.

II. Instructor roles in serious gaming

In serious gaming, instructors can have varying roles depending on their perspective of learning. They also transition between different roles, depending on the current situation, and they may also take on several roles at the same time. Table I summarises instructor roles discussed in this paper.

A. Facilitator

The most common role mentioned in serious game literature is that of facilitator and is here used as an umbrella term. The terms coach, trainer or tutor are sometimes used as synonyms, although coach is mostly used for the more specific in-game facilitator [6], [12]. A facilitator is most often a teacher or instructor, but can also have an administrative employment. In non-educational settings (e.g. games used in healthcare or rehabilitation), the facilitator is any professional (e.g. physical therapist) responsible for the serious gaming activities. Facilitators provide structure and guidance to a learning experience, rather than providing the correct answers in an authoritarian way [13]–[18]. This guidance can be more or less active. For instance, during the introduction or briefing of a game, the facilitator has a key role in making sure that all participants comprehend the notion of serious gaming, which is critical to the learning process [19]. The introduction sets the tone and atmosphere of the game, as well as the style of guidance [19]. An in-game scripted introduction would most likely be unable to capture these subtle nuances as effectively as a human facilitator, especially in dealing with critical participants, false expectations or group tensions [20].

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

TABLE I. SUMMARY OF THE MOST COMMON TYPES OF INSTRUCTOR ROLES. PLEASE NOTE THAT SOME OF THESE ROLES ARE NOT EXCLUSIVE TO INSTRUCTORS. FOR INSTANCE, AN IT SUPPORT TECHNICIAN CAN ALSO PROVIDE TECHNICAL SUPPORT, ESPECIALLY FOR MORE ADVANCED ISSUES.

Instructor role	Description
Facilitator	Provides structure and guidance, motivates and paces the experience, gives feedback and meta-cognitive aid.
Debriefer	Encourages off-game reflection/reflection-on-action by guiding learners in analysing and interpreting their in-game experience and performance, and provides a cool-down event.
Coach or in-game facilitator	Provides guidance and formative feedback as the game is in progress and scaffolds in-game performance (i.e. acts as director, process manager, game master).
Player or participant	An extreme form of in-game facilitator who participates in the game as a player or puckster.
Off-game facilitator	Observes the game's progression from a detached and passive (fly-on-the-wall) position.
Leader	Stays in control of the learning experience without taking an authoritarian position.
Expert	Provides content expertise during the preparation of game-based learning (e.g. scenario authoring) and in assessing learner performance.
Subject matter expert	Provides content expertise during game development.
Champion	Promotes game-based learning practices at the workplace.
Technical support	Provides help with technical issues related to game-based learning practices.

B. Debriefer

The facilitator is also active in another obvious part of a game-based learning experience and that is in the debriefing. The instructor role is then as debriefer, that is, someone who guides reflection-on-action, assesses the performance, and makes sure that everyone goes away emotionally unscathed [17], [18], [21]. Between-game debriefings also allow the facilitator to "influence the subsequent run of the game by allowing participants to reflect on its progression and to formulate actions for improvement together" [19, p. 187]. Thus, the facilitator plays a critical role in the transfer process, by guiding participants towards a reflective practice, in which making connections between the game and the work practices are fostered [22].

C. In-game facilitator or coach

With regard to the facilitators' activities during gameplay, the role is less clear-cut and ranges from active participant in the gaming experience to passive observer [21]. In adult learning, instructors play a pivotal role in facilitating a change in behaviour from a mere acceptable level of performance to one that will excel at work tasks [23, p. 667]:

> Some researchers suggest a facilitator is doing well when participants scarcely notice their presence. However, a facilitator must always remain sufficiently active in the background to ensure arrival at an appropriate end point. Additionally, they must be present in a very focused and observing manner to be aware of participants' decision processes and group dynamics. This attentiveness ensures the facilitator can make appropriate regulatory interventions, which is described as a form of "active inactivity"

The in-game facilitator or coach is able to provide just-in-time information and meaningful feedback [6], [10], [24], provide scaffolding by handling some of the learners tasks [25], act as a director or process manager in the learning experience [26]–[28], and has an insider view of the experience that gives depth and legitimacy to the debriefing [29]. Thus, the instructor can, by careful timing of events, create situations in which reflection-in-action can occur.

An extreme form of in-game facilitator that is rarely described in serious game literature is when the instructor him- or herself takes part in the gameplay. The instructor then becomes a participant alongside the learners, by, for instance, collaborating with the learners [10] or controlling one or several avatars (becomes a 'puckster') [30]. More or less 'disguised' as another player, the instructor can provide guidance without breaking the flow of the game [31], [32]. Of course, from an instructor's point of view, gameplay will be a different experience compared to that of the learners. Instructors take on a role that is related to that of game master, which means they have an almost God-like overview of events and can interfere when the game takes the players in a direction that is counter to the instructional goal of the exercise [6], [33]. In contrast, learners only have a limited view of the virtual environment and the events taking place, which reflects the level of information that their role has access to in reality. For instance, a trainee playing as commander will have access to more information than one playing a role with a lower ranking. Instructors, on the other hand, usually have no restrictions tied to their role, no matter what character they play. This can be especially useful if the learners enter 'gamer mode' [34], which could potentially result in learners using methods that work for winning the game, but would be inappropriate in real life. As game master, the instructor could tweak the game in real-time so that learners using a sound strategy would win the game and others would not. Furthermore, a human facilitator is necessary in order to spot when learners enter gamer mode [9]. Letting the instructor join the fun of gameplay is suggested to be more motivating for instructors and, thus, increase their acceptance of serious gaming [31], [35].

Another aspect of the in-game facilitator is that different instructors vary in style; some are very active during gameplay, while others let the game run without too many interventions. In a study of 59 facilitators, Van Kessel and Datema [19] noticed two aspects in which facilitators mainly differed: focus (content or process) and the extent of interventions (many or few). Each combination has its strengths and weaknesses. For instance, the content-focused facilitator who intervenes a lot is able to show all the possibilities of what is to be learned [19]. Interventions can, however, break the fidelity and flow of the game, which will be frustrating and, in the worse case, confusing to the learners [36]. An instructor who is too controlling may fail to facilitate transfer, by giving the participants the 'correct' answer before they have had the opportunity to reflect upon different solutions themselves, which can thus make the game a frustrating experience [19]. At the other end of the spectrum, the process-focused facilitator who lets the game run its course with no or few interventions will teach the participants to change, but will also miss important opportunities for critical moments of in-game reflection [19]. Consequently, the in-game facilitator must be able to recognise situations where interventions are appropriate and at the same time know when to let the learners explore the game freely. Finding the right balance helps to pace the game, synchronise the participants, and steer them in the right direction when necessary [28], [36].

D. Off-game facilitator

The off-game facilitator, on the other hand, is able to observe the game unfold, from a detached perspective, which can also be useful during debriefing [29], especially as a counterpart to the participants' own views and inferences [37]. Thus, the off-game facilitator is more similar to a-fly-on-the-wall instructor role [16]. It is worth noting, however, that complete detachment from the gameplaying stage, even as a passive observer, is not advisable, since "debriefing is likely to be rather general and abstract, for it is more difficult now to discuss specific experiences of participants or concrete events that did take place. Important to note is that this way to proceed will give participants the impression that their behaviours were highly predictable" [37, p. 81]. Consequently, a combination of both in-game and off-game facilitators can be very powerful [29], [37].

E. Champion or early adopter

Since formal training for serious game instructors is rare, most skilled facilitators have reached their level of expertise through informal channels, such as pure interest or passion for game-based learning [24] or through an organisational infrastructure that encourages instructors to share knowledge and learn from each other's mistakes and successes [13], [15], [38]. This means that existing instructors with interest and competence in serious gaming (so-called champions) should be encouraged to share their knowledge with their less experienced colleagues [24].

F. Technical support

Apart from tasks related to didactics, technical support is another important role or task for the facilitator, which either falls to the instructor or a separate individual or organisation that deals with the more advanced technical issues [39]. Involving digital games, technical problems will always be an issue that might deter instructors from fully embracing serious gaming [40].

G. Subject matter expert

Lastly, some instructors also take on the role of subject matter expert during the development of a particular serious game. As such, they are responsible for making sure that the right content is added to the game and that it is represented correctly [21]. Instructors can also act as subject matter experts as a teaching approach, in which they "teach by example, modeling and encouraging critical thinking as they systematically organize and analyze the subject matter knowledge" [12, p. 221].

III. Benefits and barriers

As should be evident by now, the instructor can and does take on many roles as a facilitator of game-based learning. This has a number of positive implications, such as:

- increased *instructor buy-in* through active involvement in the game production and play [15], [24], [35];
- increased *learner buy-in* and motivation through the presence of the instructor, who legitimises the use of games for serious purposes [24], [41];
- leverage of *emotional aspects* of serious gaming, such as boosting morale [42], decreasing anxiety [24], [42], and establishing rapport between the instructor and the learners [43];
- ensuring that *deliberate practice* [44] is achieved by learners who have yet to become self-monitoring [6];
- avoiding the type of gaming that leads to behaviours unsuitable or even detrimental for specific work practices, such as the behaviours exhibited while in *gamer mode* [34], [45];
- forming or reshaping *communities of practice* in which instructors and learners are all participants in creating shared experiences, and the instructors facilitate learning through advice, coaching and other instructor-learner interactions [27], [28], [46];
- enabling *transfer* of knowledge from the gaming context to the work context by explicit learning and reflective activities [6], [22];
- more *effective learning* through guided discovery and feedback [23], [46]–[48];
- decreasing the need for complex and resource-heavy simulations, since the instructor, as in-game facilitator, adds *complexity, noise, and dynamics* [48].

Most importantly, instructor involvement leads to more high-quality serious gaming [19]. A skilled facilitator is able to make real-time assessments and create an adaptive and dynamic experience that goes beyond the game artefact itself [27], [48]. For instance, Bauman and Wolfenstein [36] claim that instructors can react in real-time to inappropriate in-world behaviour and appearance, such as bullying or contextually unsuitable (avatar) appearance. This is especially useful for aspects of the gaming situation that have not been hard-coded into the game engine or rule set, but important for the learning experience. There are, however, two sides to this coin [49, p. 171]:

> In order to support purely computer-based accurate feedback, the vocabulary of operations and situations in the system has to be specified in advance so that rules can be written. Once deployed, students can only do what the system has been prepared to support. It is considerably harder for instructors, as non-programmers, to modify a computer application when they want to customise it for their courses.

Instructors' lack of programming skills is not the only barrier to instructor-led gaming. Serious game literature mentions the following barriers or potential weaknesses:

- Even if instructors recognise the benefit of being involved in serious game production, they may hesitate in actually

becoming involved, due to *time restraints* and *priorities* towards their teaching duties [8], [21], [50];
- Instructor-led serious gaming implicates *teacher control* instead of learner control, and facilitators who take a directive role, enforce control, and perform a great deal of hand-holding and micromanagement, which is detrimental to learning and the development of self-regulatory skills [13], [36], [51];
- Interferences *break the fidelity or flow* of the game [36] and could cause loss of immersion [52];
- Assessments from human instructors are often *subjective* and comparisons between learners are inaccurate [2];
- Increased *costs* in terms of:
  - loss of return of investment [53],
  - licensing and other expenses [13],
  - labour-intensive and time-consuming individualised instruction or coaching [2], [5],
  - lack of competent facilitators [1], [8], [54], and
  - increased need for technical and pedagogical support for inexperienced instructors [20], [35], [55];
- Practical difficulties in harmonising serious games with the *constraints of the educational setting*, such as time needed for playing a game [8], [20], [50], [56];
- Perceived or real *lack of technology reliability and validity* [8];
- Difficulties in adapting games with *fixed content* [35], [49];
- Difficulties in *following dynamic gameplay in real-time* [1], at least without additional functionalities that support in-game facilitation;
- Difficulties in providing real-time feedback and support when gameplay is *asynchronous and distributed*, such as during a distance course [31];
- Incompatible learning theories with regard to instructor-led serious gaming, such as ill-defined instructor roles [27], or other *organisational obstacles* [38].

The increase in costs is usually considered the most serious barrier to instructor-led game-based learning and is also most often used as the main argument for instructor-less training systems. For example, costs are sometimes cut by reducing instructor hours [15]. However, there "is often an up-front cost—monetary and human resources—to introducing simulation and game-based learning into and across curricula. Once this infrastructure is in place, neglecting the content, pedagogy, or delivery mechanism can be costly" [24, p. 94]. Thus, reducing instructor hours can also reduce training effects, which will lead to mistakes in the workplace. Since monetary costs related to serious gaming usually affect different cost units and training effectiveness is inherently difficult to translate into monetary values, calculating the cost-effectiveness of instructor-led versus instructor-less game-based learning will yield different results, depending on assumptions made about the costs involved [53].

A further concern is that research into serious gaming practices tends to focus on younger learners, while ignoring adult or elderly learner groups [27], [57]. Some of the practical challenges for facilitators, which are applicable in most serious gaming contexts, mentioned in the literature are:

- deciding whether to use serious games and, if so, when to use them [58],
- difficulties persuading other stakeholders to adopt serious gaming within the organisation [59],
- finding or obtaining the time to learn how to practise serious gaming [59],
- identifying which games are available and suitable for the intended learning outcomes [59],
- monetary, staff and technological restrictions [24], [60],
- licensing constraints [54], [61],
- initial efforts in setting up computers and learning the game user interface [56],
- dealing with problems concerning usability and user acceptance [62],
- determining how to practically use games as learning tools [8], [57], [58], [63], and
- receiving formal training specific for teaching with simulations and games [20].

Within the military domain, there is also the trade-off between publicly available games (with realistic scenarios) and the risk of terrorists/enemies misusing them [61].

IV. Possible solutions

There are a number of suggestions to manage these challenges, which can roughly be divided into organisational, technical, and human-computer interaction (HCI) related solutions. The following examples are by no means an exhaustive list of solutions, but rather ideas and suggestions on how to tackle the challenges according to current academic literature. Figure 1 gives an overview of the solutions discussed in this paper.

A. Organisational solutions

Organisational solutions involve creating an infrastructure that facilitates knowledge management at different levels. For instance, formal and informal structures for educating instructors and encouraging them to learn from each other are needed [38], [51]. These efforts should especially focus on understanding the possibilities and limitations of serious games and creating communities of best practices for the sharing of stories [13]. That is, individuals must feel that the act of knowledge sharing will be beneficial for them and reciprocated by other members of their community. Difficulties in knowledge sharing are a well-known issue in knowledge management, but it is still not clear what organisations can do to encourage knowledge sharing [64]. Furthermore, it is important that instructors engage in research, in order to exploit the full potential of game-based learning [24], [50]. In essence, organisational solutions involve making use of the inherent enthusiasm of champions to spread knowledge and increase buy-in, without stifling their enthusiasm through ill-considered reward structures [24]. Another type of organisational solution involves support structures for both pedagogical and technical issues that the instructors might need help with [15], [24].

B. Technical solutions

A great concern for educators is the reliability and validity of the educational system [8]. Technical solutions gear towards either replacing parts of the human instructor's task with a tutoring system or pedagogical agent, or augment the tasks with technical solutions, such as sensors, logging tools and automatic assessment (see e.g. [65]). For instance, many serious games utilise NPCs to deliver content, tasks, or act as mentors [17], [32]. As such, these pedagogical agents can provide scaffolding and drive the narrative forward [32], [66]. There is also work on the enhancement of virtual agents and human behaviour modelling that tries to alleviate the need for human controllers (or pucksters) in serious gaming or simulation. For instance, a single human puckster can control a group of virtual agents, instead of controlling every individual avatar (see e.g. [30]). Similarly, Sycara and Lewis [67] examine multi-agent systems where the virtual agents support human teams in carrying out their tasks. The challenge for these projects is to create agents that model human behaviour as closely as possible, that is, to make human-avatar interactions natural to the point of being almost indistinguishable from human-human interaction. Another example is AutoTutor, a system that employs natural language to support students learning such subjects as Newtonian physics, computer literacy and scientific reasoning. According to the web site (autotutor.org), the system acts as a dialogue partner with the learner and encourages students "to articulate lengthy answers that exhibit deep reasoning, rather than to recite small bits of shallow knowledge". However, Chatham [54, p. 244] notes that NPCs "are today mostly vending machines and jousting dummies" and not very believable as real humans, when interacting with them for more than a few utterances. His statement remains by and large true today [68]. Most adaptive modelling in games is based on ontologies of game entities (objects, NPCs, story elements) that are combined into modules that fit into the narrative [4]. This means that tutoring systems and other adaptive features are limited to well-defined or quantifiable aspects of the gameplay and narrative, that is, those aspects that can be reduced to algorithms. Although these techniques have proven useful for creating interactive and emergent narratives (see e.g. [66], [69]), we are still far from creating AI-agents that are as adaptive as human instructors, especially when training leadership, decision-making and communication skills [9]. Thus, there is a risk that the dynamics of the exercise are lost when pucksters are replaced with virtual agents. Moreover, human behaviour modelling is an endeavour that many serious game developers avoid, due to developmental costs and technological restrictions among the end users.

Technical solutions that aim to relieve some of the cognitive workload put on human instructors seem to be a more promising approach. Most research efforts have focused on automatic assessment systems that are integrated into the game [5], [50], [70] and many also express the need for more work on debriefing tools [48], [65]. To give an example, Ekanayake et al. [2] created an algorithm that assesses driving behaviour in a driving simulator. The algorithm is not only based on the behavioural evaluation of achievement goals, but also accounts for the player's effort towards achieving those goals, such as the physical pressure on the throttle.

C. HCI related solutions

HCI related solutions are related to technical solutions, but more geared towards higher-order considerations on how the instructor's tasks can be simplified or augmented. Some researchers have examined the challenge of involving instructors in game production and have found a number of solutions. The common denominator between these solutions is scenario authoring tools. For instance, Brennecke [41] suggests that a scenario authoring tool could be designed as a game. In a case study, she tested this by implementing a game-based system for crime scene investigation (CSI) training. The idea was to have a cop-and-robber scenario; the teacher would prepare the training session by playing the game as the antagonist, stealing valuable items in a virtual apartment and leaving clues such as foot- and fingerprints. The students were then given the task of investigating the crime scene and trying to 'outsmart' the teacher. The elegance of this idea is that it encourages instructors with less technical skill to create scenarios, since it involves no other skills than playing the game itself. It might also have a positive motivational effect, especially among instructors who enjoy playing games. To deal with the time pressure issue, scenario development could be carried out by a group of instructors. Stiso et al. [71] outline the development of a Common Instructor Operator System (C-IOS) that supports collaborative and distributed scenario authoring. In order to enable the management of different permissions, that is, who is allowed to alter the scenario at specific times during the development cycle, they created an instructor hierarchy consisting of (i) the lead instructor, (ii) the element instructor, and (iii) the platform instructor. A single instructor can inhabit one or several of these roles, or several instructors can be allocated the same role. The main idea is to distribute the work load over several individuals and to enable scenario authoring by instructor teams that are distributed over remote sites or whose work is asynchronous. Distributed scenario authoring still assumes some level of technical competence among instructors, and therefore another popular research direction is to develop automatic or semi-automatic scenario generation [72]. The goal is to relieve instructors of some of the time consuming work, by offering scenario authoring through the 'touch of a button'. Although these solutions still require some scripting skills, the aim is to generate adaptive scenarios from data about player preferences, skill level, game actions, and so on [73]. If successful, these scenarios will be able to adapt to the individual learner's performance, making scenario authoring fully automatic, yet retaining the dynamic nature of manual authoring.

Usability is, of course, a key issue in HCI related solutions, as well as different interfaces specialised to visualise different aspects of the ongoing gameplay and the learners' performance, such as instructor dashboards, HUDs and other control tools [65]. Automatic assessments also need to present measurements to instructors, in a comprehensive and concise way, for use in both formative feedback (e.g. interventions) and debriefings [5]. For instance, Raybourn [48] highlights the importance of being able to bookmark specific events during gameplay, which the instructor wants to re-play during the debriefing later on. This implies that the system needs to not only include logging functionality, but also access to the logs in real-time, through some kind of instructor interface, where events are clear, distinct and selectable. It also implies that the debriefing tool functionality must be able to re-play recorded events, preferably from different angles [48].

Although many instructor tools can be added as features in a system adjacent to the game (e.g. a separate debriefing module), I would argue that a more elegant solution would be to integrate these features within the serious game itself, if possible. This would make transitions between different instructor roles more fluid and flexible, without the need to

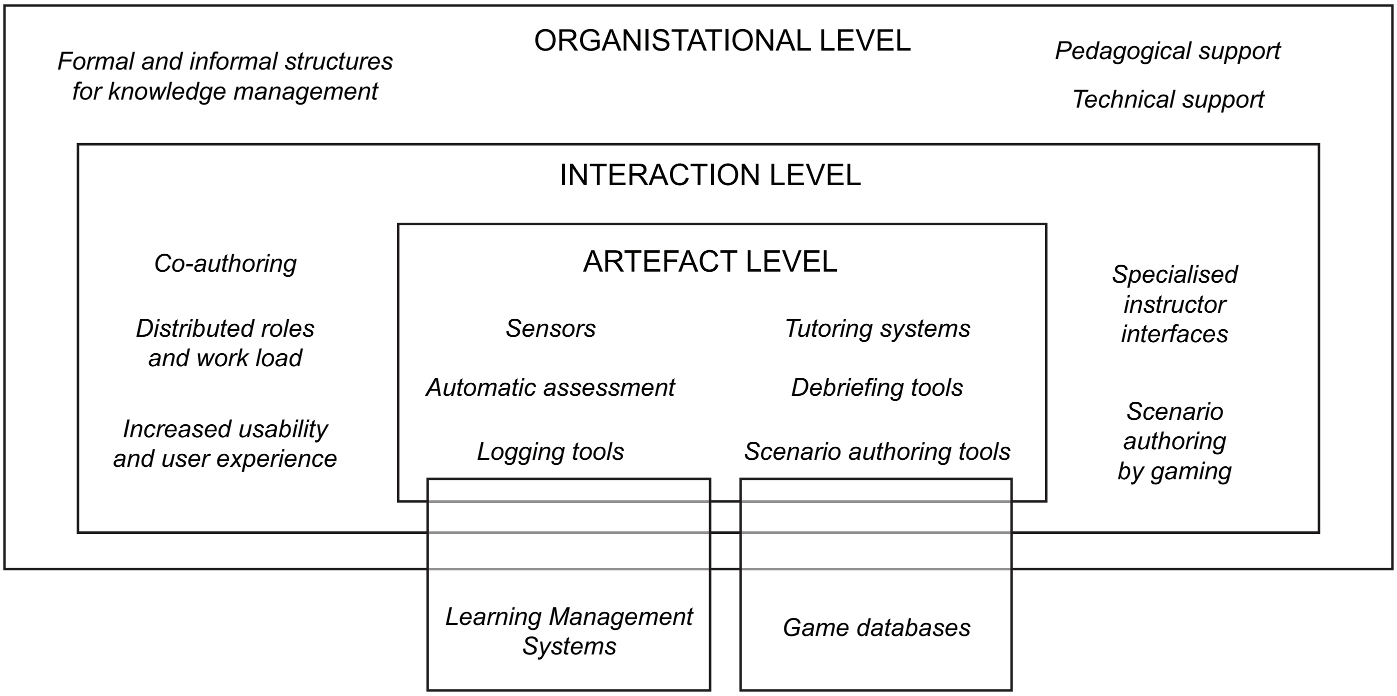


Fig. 1. Overview of possible solutions to tackle the challenges in instructor-led serious gaming

switch program or workstation, and also reduce the costs of educating instructors in different systems.

D. Additional solutions

There are also a few endeavours that are not as easily categorised into organisational, technical or HCI related solutions, but rather span all three. For instance, some efforts have been in the area of integrating serious games into learning management systems, from which games can be introduced, launched and administrated (see e.g. [74]). More recent research is also focused on supporting teachers in the choice of suitable games, by providing databases for sharing the experiences of using specific games (e.g. [75]).

V. Conclusion and future work

In sum, instructors shift between different roles during serious gaming and it is important to not treat them as having only one role. Their main role is that of facilitator, which in turn has different characteristics, depending on the situation and instructional goal. Key roles identified in the literature are debriefer, coach or in-game facilitator, player/participant, off-game facilitator, leader, expert, and technical support. Each of these roles comes with a different set of needs and requirements that must be met in order for serious gaming to run smoothly. Apart from the above mentioned roles, some instructors also take on the roles of subject matter expert and champion, which are broader roles used outside the training situation. Thus, the instructor role extends past the here-and-now serious gaming to situations where experienced instructors introduce novices to the serious gaming practices, or are involved in game production. Furthermore, there are many ways of dealing with the barriers to instructor-led game-based learning, of which I have identified three broad themes: organisational, technical and HCI related solutions.

Instructor-led serious gaming means there is an increased need for instructors with not only high levels of expertise in both pedagogical issues as well as running a simulation, but also able to externalise their knowledge in order to accommodate a culture of cognitive apprenticeship.

This paper makes a case for instructor-led serious gaming, by providing a comprehensive overview of the state of the art in instructor-led serious gaming. It points to the importance of studying the socio-cultural context and off-game activities involved in serious gaming, as well as the importance of designing for user acceptance and user experience, where the instructors form a significant user group. It also highlights the notion of instructors as facilitators who play an important role in helping learners transfer knowledge from the gaming situation to the work context, and the different facilitation techniques that can be used in serious gaming environments.

So far, most work within this area is concentrated on theoretical work and single case studies. However, in order to fully explore the potential and power of serious games, we need more comparative and longitudinal studies. For instance, what instructor roles and needs are associated with specific types of games and educational settings? In order to answer such questions, we need to study not only contexts in which serious gaming is still a novel practice, but also environments where games have been employed for a longer time period.

References

[1] I. Goldstein and B. Carr, "The computer as coach: As athletic paradigm for intellectual education," in *Proc. of the 1977 annual ACM conference*. ACM, 1977, pp. 227–233.

[2] H. Ekanayake, P. Backlund, T. Ziemke, R. Ramberg, and K. P. Hewagamage, "Assessing performance competence in training games," in *Affective Computing and Intelligent Interaction. Proc. of ACII'11, Part II*, ser. Lecture Notes in Computer Science, S. D'Mello, A. Graesser,

B. Schuller, and J.-C. Martin, Eds. Memphis, TN: Springer, 2011, pp. 518–527.

[3] C. Franzwa, Y. Tang, and A. Johnson, "Serious game design: Motivating students through a balance of fun and learning," in *Proc. of VS-Games'13*, 2013, pp. 17–23.

[4] M. D. Kickmeier-Rust and D. Albert, "Emergent design: Serendipity in digital educational games," in *Virtual and Mixed Reality*, ser. Lecture Notes in Computer Science, R. Shumaker, Ed. Berlin: Springer, 2009, vol. 5622, ch. 24, pp. 206–215.

[5] S. M. Stevens-Adams, J. D. Basilico, R. G. Abbott, C. J. Gieseler, and C. Forsythe, "Performance assessment to enhance training effectiveness," in *Proc. of I/ITSEC'10*, 2010.

[6] A.-S. Alklind Taylor, P. Backlund, and L. Niklasson, "The coaching cycle: A coaching-by-gaming approach in serious games," *Simulat. Gaming*, vol. 43, no. 5, pp. 648–672, 2012.

[7] B. Berg Marklund, "Out of context – understanding the practicalities of learning games," in *Proc. of DiGRA'14*. DiGRA, 2014. [Online]. Available: http://www.digra.org/digital-library/publications/out-of-context-understanding-the-practicalities-of-learning-games/

[8] B. Berg Marklund, P. Backlund, and H. Engström, "The practicalities of educational games: Challenges of taking games into formal educational settings," in *Proc. of VS-Games'14*. IEEE Press, 2014.

[9] A. Frank, "The instructor role during educational wargaming," in *Proc. of ISAGA'14*, W. C. Kriz, T. Eiselen, and W. Manahl, Eds., 2014, pp. 153–66.

[10] R. Hämäläinen and K. Oksanen, "Challenge of supporting vocational learning: Empowering collaboration in a scripted 3D game – how does teachers' real-time orchestration make a difference?" *Comput. Educ.*, vol. 59, no. 2, pp. 281–293, 2012.

[11] L. Vos, "Simulation games in business and marketing education: How educators assess student learning from simulations," *Int. J. Educ. Manag.*, vol. 13, no. 1, pp. 57–74, 2015.

[12] A. Y. Kolb, D. A. Kolb, A. Passarelli, and G. Sharma, "On becoming an experiential educator: The educator role profile," *Simulat. Gaming*, vol. 45, no. 2, pp. 204–234, 2014.

[13] H. Axe and H. Routledge, "Practical applications of serious games in education," in *Handbook of research on improving learning and motivation through educational games: Multidisciplinary approaches*, P. Felicia, Ed. Hershey, PA: IGI Global, 2011, ch. 44, pp. 961–973.

[14] B. Berg Marklund, P. Backlund, C.-J. Dahlin, H. Engström, and U. Wilhelmsson, "A game-based approach to support social presence and awareness in distributed project-based learning," *IJGBL*, vol. 4, no. 1, pp. 1–20, 2014.

[15] W. R. Bickley, R. J. Pleban, F. Diedrich, J. Sidman, R. Semmens, and A. Geyer, "Army institutional training: Current status and future research," U.S. Army Research Institute for the Behavioral and Social Sciences, Research Report, 2010. [Online]. Available: http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA516971

[16] G. J. Hofstede, L. de Caluwé, and V. A. M. Peters, "Why simulation games work—in search of the active substance: A synthesis," *Simulat. Gaming*, vol. 41, no. 6, pp. 824–843, 2010.

[17] J. R. Kirkley, T. M. Duffy, S. E. Kirkley, and D. L. H. Kremer, "Implications of constructivism for the design and use of serious games," in *Computer games and instruction*, S. Tobias and J. D. Fletcher, Eds. Charlotte, NC: Information Age Publishing, 2011, ch. 14, pp. 371–394.

[18] L. Ukens, "Learning games: Hands-on participant-centered activities," in *The handbook of experiential learning*, M. Silberman, Ed. San Francisco: Pfeiffer, 2007, ch. 7, pp. 124–137.

[19] M. van Kessel and H. Datema, "Facilitators: Quality, style and attitude," in *Why do games work? In search of the active substance*, L. de Caluwé, G. J. Hofstede, and V. Peters, Eds. Deventer: Kluwer, 2008, ch. 16, pp. 183–189.

[20] G. Der Sahakian, G. Alinier, G. Savoldelli, D. Oriot, M. Jaffrelot, and F. Lecomte, "Setting conditions for productive debriefing," *Simulat. Gaming*, 2015. [Online]. Available: http://dx.doi.org/10.1177/1046878115576105

[21] W. H. Tan, S. Neill, and S. Johnston-Wilder, "How do professionals' attitudes differ between what game-based learning could ideally achieve and what is usually achieved," *IJGBL*, vol. 2, no. 1, pp. 1–15, 2012.

[22] S. Bronack, R. Sanders, A. Cheney, R. Riedl, J. Tashner, and N. Matzen, "Presence pedagogy: Teaching and learning in a 3D virtual immersive world," *IJTLHE*, vol. 20, no. 1, pp. 59–69, 2008.

[23] W. C. Kriz, "A systemic-constructivist approach to the facilitation and debriefing of simulations and games," *Simulat. Gaming*, vol. 41, no. 5, pp. 663–680, 2010.

[24] E. B. Bauman, "Preparing faculty and students for game-based and virtual learning spaces," in *Game-based teaching and simulation in nursing and healthcare*, E. B. Bauman, Ed. New York, NY: Springer, 2013, ch. 4, pp. 77–98.

[25] J. M. Beaubien and D. P. Baker, "The use of simulation for training teamwork skills in health care: how low can you go?" *Qual. Saf. Health Care*, vol. 13, no. suppl 1, pp. i51–i56, 2004.

[26] M. Checa-Romero, "Developing skills in digital contexts: Video games and films as learning tools at primary school," *Games Cult.*, 2015. [Online]. Available: http://dx.doi.org/10.1177/1555412015569248

[27] A. Frank, "Balancing three different foci in the design of serious games: Engagement, training objective and context," in *Proc. of DiGRA'07: Situated Play*, B. Akira, Ed., 2007, pp. 567–574. [Online]. Available: http://www.digra.org/wp-content/uploads/digital-library/07312.29037.pdf

[28] M. G. Wagner and T. Wernbacher, "Iterative didactic design of serious games," in *Proc. of FDG'13*, 2013, pp. 346–351. [Online]. Available: http://www.fdg2013.org/program/papers/paper45_wagner_wernbacher.pdf

[29] L. de Caluwé, "The active substance from the perspective of change," in *Why do games work? In search of the active substance*, L. de Caluwé, G. J. Hofstede, and V. Peters, Eds. Deventer: Kluwer, 2008, ch. 7, pp. 79–89.

[30] J. Colonna-Romano, W. Stacy, M. Weston, T. Roberts, M. Becker, S. Fox, and G. Paull, "Virtual puckster – behavior generation for army small team training and mission rehearsal," in *Proc. of BRIMS'09*, 2009, pp. 153–154.

[31] A.-S. Alklind Taylor and P. Backlund, "Letting the students create and the teacher play: Expanding the roles in serious gaming," in *Proc. of MindTrek'11*. ACM, 2011, pp. 63–70.

[32] D. Charsky, "From edutainment to serious games: A change in the use of game characteristics," *Games Cult.*, vol. 5, no. 2, pp. 177–198, 2010.

[33] M. Villalta, I. Gajardo, M. Nussbaum, J. J. Andreu, A. Echeverría, and J. L. Plass, "Design guidelines for classroom multiplayer presential games (CMPG)," *Comput. Educ.*, vol. 57, no. 3, pp. 2039–2053, 2011.

[34] A. Frank, "Gaming the game: A study of the gamer mode in educational wargaming," *Simulat. Gaming*, vol. 43, no. 1, pp. 118–132, 2012.

[35] A. Brennecke and H. Schumann, "A general framework for digital game-based training systems," in *Proc. of IADIS GET'09*, 2009, pp. 51–58. [Online]. Available: http://www.iadisportal.org/digital-library/a-general-framework-for-digital-game-based-training-systems

[36] E. B. Bauman and M. Wolfenstein, "Using virtual and game-based learning to prepare for actual practice," in *Game-based teaching and simulation in nursing and healthcare*, E. B. Bauman, Ed. New York, NY: Springer, 2013, ch. 5, pp. 99–125.

[37] V. A. M. Peters and G. A. N. Vissers, "A simple classification model for debriefing simulation games," *Simulat. Gaming*, vol. 35, no. 1, pp. 70–84, 2004.

[38] T. D. Henriksen and S. Löfvall, "Integrating learning games in higher education – from technical to organisational obstacles," in *Proc. of ISAGA'13)*, ser. KTH Transport Science publication series, S. A. Meijer and R. Smeds, Eds., 2013.

[39] R. Sandford, M. Ulicsak, K. Facer, and T. Rudd, "Teaching with games: using commercial off-the-shelf computer games in formal education," Futurelab, Project Report, 2007. [Online]. Available: http://www2.futurelab.org.uk/resources/documents/project_reports/teaching_with_games/TWG_report.pdf

[40] S. Egenfeldt-Nielsen, "Practical barriers in using educational computer games," in *Beyond fun: Serious games and media*, D. Davidson, Ed. Pittsburgh: ECT Press, 2008, p. 30–36.

[41] A. Brennecke, "A general framework for digital game-based training systems," Doctoral thesis, University of Rostock, 2009. [Online].

Available: http://rosdok.uni-rostock.de/file/rosdok_derivate_000000003979/Dissertation_Brennecke_2009.pdf

[42] G. J. Hofstede, "One game does not fit all cultures," in *Why do games work? In search of the active substance*, L. de Caluwé, G. J. Hofstede, and V. Peters, Eds. Deventer: Kluwer, 2008, ch. 6, pp. 69–77.

[43] R. D. Duke, "One perspective on the 'active substance' of gaming simulations," in *Why do games work? In search of the active substance*, L. de Caluwé, G. J. Hofstede, and V. Peters, Eds. Deventer: Kluwer, 2008, ch. 1, pp. 27–30.

[44] K. A. Ericsson, "Enhancing the development of professional performance: Implications from the study of deliberate practice," in *Development of professional expertise: Toward measurement of expert performance and design of optimal learning environments*, K. A. Ericsson, Ed. New York: Cambridge University Press, 2009, ch. 18, pp. 405–431.

[45] A. Frank, "Achieving game goals at all costs? – the effect of reward structures on tactics employed in educational military wargaming," in *Proc. of ISAGA'13*, ser. KTH Transport Science publication series, S. A. Meijer and R. Smeds, Eds., 2013.

[46] H. Leemkuil and T. de Jong, "Instructional support in games," in *Computer games and instruction*, S. Tobias and J. D. Fletcher, Eds. Charlotte, NC: Information Age Publishing, 2011, ch. 13, pp. 353–369.

[47] R. Moreno and R. E. Mayer, "Role of guidance, reflection, and interactivity in an agent-based multimedia game," *J. Educ. Psychol.*, vol. 97, no. 1, pp. 117–128, 2005.

[48] E. M. Raybourn, "Applying simulation experience design methods to creating serious game-based adaptive training systems," *Interact. Comput.*, vol. 19, no. 2, pp. 206–214, 2007.

[49] L. Qiu and C. K. Riesbeck, "An incremental model for developing computer-based learning environments for problem-based learning," in *Proc. of IEEE Int. Conf. on Advanced Learning Technologies*, 2004, pp. 171–175.

[50] E. Tsekleves, J. Cosmas, and A. Aggoun, "Benefits, barriers and guideline recommendations for the implementation of serious games in education for stakeholders and policymakers," *Brit. J. Educ. Technol.*, 2014. [Online]. Available: http://dx.doi.org/10.1111/bjet.12223

[51] A. Walker and B. E. Shelton, "Problem-based educational games: Connections, prescriptions, and assessment," *J. Interact. Learn. Research*, vol. 19, no. 4, pp. 663–684, 2008.

[52] I. Ghergulescu and C. H. Muntean, "Motivation monitoring and assessment extension for input-process-outcome game model," *IJGBL*, vol. 4, no. 2, pp. 15–35, 2014.

[53] J. D. Fletcher, "Cost analysis in assessing games for learning," in *Computer games and instruction*, S. Tobias and J. D. Fletcher, Eds. Charlotte, NC: Information Age Publishing, 2011, ch. 16, pp. 417–434.

[54] R. E. Chatham, "Toward a second training revolution: Promise and pitfalls of digital experiential learning," in *Development of professional expertise: Toward measurement of expert performance and design of optimal learning environments*, K. A. Ericsson, Ed. New York: Cambridge University Press, 2009, ch. 10, pp. 215–246.

[55] G. A. Gunter, R. F. Kenny, and E. H. Vick, "Taking educational games seriously: using the RETAIN model to design endogenous fantasy into standalone educational games," *Educ. Tech. Res.*, vol. 56, no. 5–6, pp. 511–537, 2008.

[56] S. Egenfeldt-Nielsen, "Overview of research on the educational use of video games," *Digital Kompetanse*, vol. 1, pp. 184–213, 2006.

[57] F. Ke, "A qualitative meta-analysis of computer games as learning tools," in *Handbook of Research on Effective Electronic Gaming in Education*, R. E. Ferdig, Ed. Hershey, PA: IGI Global, 2009, ch. 1, pp. 1–32.

[58] R. L. Appelman and J. H. Wilson, "Games and simulations for training: From group activities to virtual reality," in *Handbook of human performance technology*, 3rd ed., J. A. Pershing, Ed. San Francisco: Pfeiffer, 2006, ch. 17, pp. 414–436.

[59] J. Kirriemuir and A. Mcfarlane, "Literature review in games and learning," Futurelab, Tech. Rep. hal-00190453, version 1, 2004. [Online]. Available: http://telearn.archives-ouvertes.fr/docs/00/19/04/53/PDF/kirriemuir-j-2004-r8.pdf

[60] P. Backlund and M. Hendrix, "Educational games – are they worth the effort? a literature survey of the effectiveness of serious games," in *Proc. of VS-Games'13*, 2013, pp. 69–76.

[61] M. Ford, M. Barlow, and E. Lewis, "An initial analysis of the military potential of cots games," in *Proc. of SimTecT'03*, 2003.

[62] A. Padrós, M. Romero, and U. Mireia, "Developing serious games: from face-to-face to a computer-based modality," elearningeuropa.info, 2011. [Online]. Available: http://elearningpapers.eu/en/article/Developing-Serious-Games%3A-from-Face-to-Face-to-a-Computer-based-Modality

[63] P. G. Schrader and M. McCreery, "Are all games the same?" in *Assessment in game-based learning: Foundations, innovations, and perspectives*, D. Ifenthaler, D. Eseryel, and X. Ge, Eds. New York, NY: Springer, 2012, ch. 2, pp. 11–28.

[64] M.-J. J. Lin, S.-W. Hung, and C.-J. Chen, "Fostering the determinants of knowledge sharing in professional virtual communities," *Comput. Hum. Behav.*, vol. 25, no. 4, pp. 929–939, 2009.

[65] A.-S. Alklind Taylor and P. Backlund, "Making the implicit explicit: Game-based training practices from an instructor perspective," in *Proc. of ECGBL'12*, P. Felicia, Ed. ACI, 2012, pp. 1–10.

[66] F. Bellotti, R. Berta, and A. De Gloria, "Designing effective serious games: Opportunities and challenges for research," *iJET*, vol. 5, pp. 22–35, 2010.

[67] K. Sycara and M. Lewis, "Integrating intelligent agents into human teams," in *Team cognition: Understanding the factors that drive process and performance*, E. Salas and S. M. Fiore, Eds. Washington, DC: APA, 2004, ch. 10, pp. 203–231.

[68] J. Brusk, "Step towards creating socially competent game characters," Doctoral Dissertation, University of Gothenburg, 2014. [Online]. Available: http://hdl.handle.net/2077/34774

[69] M. D. Kickmeier-Rust, E. Mattheiss, C. Steiner, and D. Albert, "A psycho-pedagogical framework for multi-adaptive educational games," *IJGBL*, vol. 1, no. 1, pp. 45–58, 2011.

[70] F. Bellotti, B. Kapralos, K. Lee, P. Moreno-Ger, and R. Berta, "Assessment in and of serious games: An overview," *Adv. Hum. Comput. Interact.*, vol. 2013, 2013, article ID: 136864.

[71] M. E. Stiso, J. M. Owens, J. E. Fowlkes, S. Eitelman, A. Hafich, M. M. Walwanis Nelson, and D. G. Smith, "Framework for a common instructor operator system, part 1: Enabling scenario development," in *Proc. of BRIMS'04*. Simulation Interoperability Standards Organization (SISO), 2004, pp. 102–111.

[72] G. A. Martin, C. E. Hughes, S. Schatz, and D. Nicholson, "The use of functional L-systems for scenario generation in serious games," in *Proc. of PCGames'10*. ACM, 2010.

[73] R. Lopes and R. Bidarra, "Adaptivity challenges in games and simulations: A survey," *IEEE Trans. Comput. Intell. AI in Games*, vol. 3, no. 2, pp. 85–99, 2011.

[74] J. Torrente, P. Moreno-Ger, I. Martínez-Ortiz, and B. Fernandez-Manjon, "Integration and deployment of educational games in e-learning environments: The learning object model meets educational gaming," *Educ. Technol. Soc.*, vol. 12, no. 4, pp. 359–371, 2009.

[75] M. Hendrix, P. Backlund, M. Lebram, and H. Lundqvist, "Sharing experiences with serious games: The EduGameLab rating tool for parents and teachers," in *Proc. of VS-Games*, 2013, pp. 138–141.

A Game-based Training Approach to Enhance Human Hand Motor Learning and Control Abilities

Ali Asadipour[1], Kurt Debattista and Alan Chalmers
Visualisation Group, WMG
University of Warwick
UK
[1]Email: A.Asadipour@warwick.ac.uk

***Abstract*—This work presents a serious game designed to improve the performance in users' control abilities applied to pressure sensitivity. In particular, the aim of this work is part of a larger goal of providing medical students with further opportunities of palpation experiences and assistance as part of their education. Typically medical students are limited by the number of volunteers they can practice on and the amount of time they can interact with more experienced practitioners to further develop fundamental palpation skills. Correct palpation skills are crucial as they inform the diagnosis in a large number of healthcare fields and a skill required by most healthcare professionals. The ability to be able to enhance the educational process of healthcare professionals' palpation skills could lead to a more holistic student experience. This work presents a serious game in which one aspect of palpation, hand control ability through the correct application of pressure to a patient, is the target for user improvement. A serious game modelled on the infinite runner genre was designed to be controlled via an input device developed in-house with off-the-shelf components that translates real-world pressure to in-game movement. The game was tested in a participant trial involving a game-playing group ($n = 15$) and a control group ($n = 15$) and a significant improvement in a blind-folded pressure test was observed for the game-playing group. User feedback via a questionnaire also showed a positive response of the game.**

I. Introduction

The efficacy of serious games as a training approach has become widely accepted and as a consequence serious games are beginning to be used in a wide variety of domains. Immersion, pleasure and competition are key characteristics of games that enhance user engagement in training activities. Apart from research studies even commercial companies are now interested in taking part in the development and deployment of educational games [1].

One of the main domains demonstrating the fruitfulness of serious games is healthcare [2]. A number of reasons have lead to the success of serious games for healthcare. Ethical restrictions in medical training mean that certain procedures cannot be frequently tested and explored by practitioners. Similarly, a lack of available patients during training entails limited familiarity with the application of procedures across a varying number of ages, genders and body types. Furthermore, the demands for patients' growing thirst for health information has led to health professionals providing novel digital-based interventions.

Current challenges in medical education are particularly difficult when the possibilities of experiencing training in that activity are rare. Game-based solutions can provide an enhanced experience to the existing educational process for such cases [3, 4]. One of the most common medical processes which are carried out by most medical practitioners in a very large variety of conditions and for a diverse number of applications is palpation [5, 6]. Palpation plays an important role in the initial examination of patients and is a crucial initial diagnosis [7] based significantly on haptic sensory feedback. While visual and acoustic digital healthcare practices are becoming more common due to the ubiquitous nature of video and audio displays and their use in both entertainment and real-world applications, the lack of readily available haptic devices has meant very little progress has been made in the areas of motion and pressure sensitivity control within the healthcare domains.

This work is part of a larger project aimed at training medical students to become more proficient at palpation as part of their training process. Training in palpation is of crucial importance as during their training medical students are restricted to the number of hours they can spend with their tutors gaining hands-on experience and are also restricted on the number of body types and participants that they can engage with [8]. A goal of automating the process and providing palpation-based simulators will enhance current practices by allowing the students to practice by themselves or with each other while being guided via a digital tutor. This paper identifies and focuses on one aspect of an automated palpation framework; training of pressure sensitivity using the index finger, one of the crucial characteristics of palpation training, is provided via the use of a serious game in which the player controls an on-screen character via the use of an input device which is sensitive to pressure. Learning to apply the correct amount of pressure plays a significant role in providing the correct diagnosis and also in patient comfort; a too light touch may miss out on important physiological phenomena and a too heavy touch may cause significant patient discomfort further compounding potential diagnosis issues. While a number of novel input technologies beyond the traditional have recently begun to be applied to serious games [9, 10, 11, 12] no serious game, to the best of our knowledge, has targeted the correct application of pressure as its main goal. A study based on two groups composed of general public participants, one group that played the game and a control group demonstrates that there is reason to believe that such a serious game can help improve pressure sensitivity in individuals.

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

The following section presents background and related work. Descriptive information of the three technologies which are used in this study is discussed in Section III. The experimental design and results are discussed in IV. Finally future potentials for extending this study and conclusions are presented in V.

II. Background and Related Work

In general, applications of serious games in healthcare can be classified by their target audience [13] as follows:

- Medical education
- Patient intervention
- Public involvement

Games dedicated to medical education are those that help medical professionals to improve their skills while performing certain tasks. Patient intervention are targeted at patients rather than the medical professionals. Such patient-oriented training games help enhance individuals' knowledge about their condition and to also help improve their engagement in their treatment process. Public involvement applications are directed at the general public and are focused on raising awareness of public health issues and providing motivation for potential behavioural change. A large number of healthcare-related serious games have been developed [2] and we provide a small set of examples in the following.

In terms of medical education and awareness Graafland et al. [14] presented a survey on medical education and surgical skills across 25 publications which comprised 30 games. They explained that games developed for the purpose of such serious applications required the use of further validation before being deployed as there was a lack of robust evaluation for the surveyed games. Dunwell et al. [15] presented a serious game to help create awareness of healthcare-associated infections within wards. They provide feedback and findings on the deployment of the game across 13 hospital wards in the United Kingdom. The serious game we present in this paper could also be considered as part of the medical education sub-category of serious games.

An example of a patient intervention game was the Re-Mission game for cancer patients. Positive behavioral changes were reported on pediatric patients who were diagnosed with cancer by playing this video game [16]. Another patient intervention example was provided by Carmeli et al. [17]. The authors presented a serious game for the improvement of motor, sensory and cognitive performance in rehabilitation of stroke patients (with upper limb impairments) to conduct everyday functional tasks efficiently. An interactive tool was used as an input device for the game to measure range of motion and finger and wrist speed. Results demonstrated an improvement in movements for users.

An example of increasing public involvement was presented by Boulos et al. [18] environment to raise public awareness about sexual health. An online 3D virtual world such as Second Life (SL) with social networking capabilities has been surveyed to highlight positive impact on its audience. Brown et al. [19] also presented a serious game dealing with sex education in which an intervention mapping approach was used in the development and the design of the game. Scarle et al. [10] presented a serious game that had two main goals focusing on public involvement. Firstly, it was targeted at raising the awareness of poor eating habits at primary school children that has been becoming one of the main causes of the rising obesity epidemic and, secondly, through the application of motion controls that enabled the participant to reduce the amount of on-screen time in which the game player was physically inactive.

Figure 1. ParsGlove is an innovative wearable interface which is designed and developed to capture the human hands ergonomics and to provide this information as an interactive input both for the game and the application.

III. Serious Game and Input Device

This section presents the overall framework that has been used for this work. Three key technologies including a serious game were developed to improve pressure sensitivity learning. The input device is a glove developed in-house and can read pressure the amount of pressure applied accurately in Newtons. An application, DigiScale, was developed to help facilitate the input procedure via a user friendly interface. Finally, a serious game was implemented to facilitate the learning of pressure sensitivity for the user.

A. Input Device

A wearable measurement interface, ParsGlove (see Figure 1) has been developed under formal R&D discussions with with medical professionals. It is designed to capture the ergonomics of the human hand during dexterous interaction with the environment. It was essential in our main goals to use the full capabilities of the glove to capture applied pressures, orientations and location parameters although for this work the focus is only on the pressure input.

To provide more freedom the glove is equipped with Bluetooth connectivity and is composed of ultralight materials which help reduce weight and avoid fatigue when the glove is worn for long periods. Twelve force sensitive resistors are mounted in places which were defined by medical professionals. Sensors were also calibrated with a force gauge device to accurately map digital values to actual force.

B. Application

DigiScale is a Graphical User Interface (GUI) specifically developed for this experiment to deliver visual feedback for the force exerted by the tip of users' index fingers. DigiScale has two sections: an information panel providing visual information directly to the user and a toolkit panel for the research investigator. Figure 2 shows a screenshot of DigiScale demonstrating a target force set to 2 Newtons and a sampling timer having collected 3 seconds of data out of a total of 10 seconds (with 7 seconds remaining as shown in the figure).

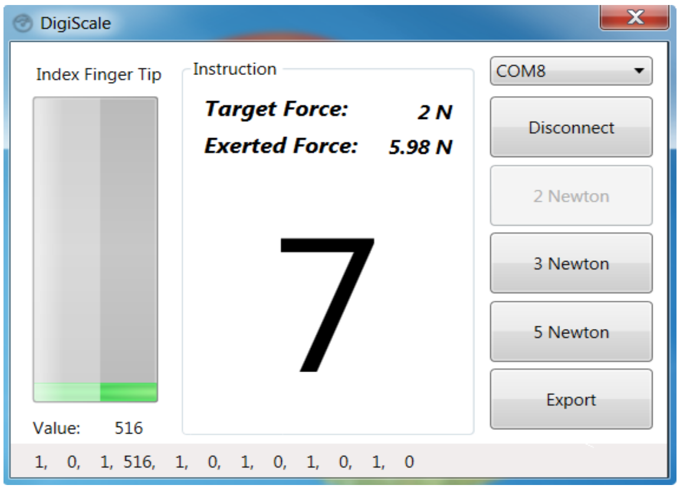


Figure 2. The DigiScale application is designed to deliver visual feedback on the exerted force by the user and to collect samples form the user performance.

The information panel is designed to provide visual feedback of applied forces by the user, the target force level in each task and a countdown timer. The toolkit panel is designed to provide connectivity options to pair ParsGlove with DigiScale, three buttons to set the target force levels in each task and finally an export button to persist recorded samples at the end of each session. The application has capabilities to read continuous data streams over the Bluetooth standard which permits communication with the glove. The application could be run on any normal PC without any specific hardware or software requirements. The only requirement for this application is a Bluetooth receiver.

C. Game

To provide game-based training to help improve performance, a 2D game was developed in the Unity 3D game engine. The game assets were adapted from free assets made available from the Unity store. The game was designed to interface with the ParsGlove playing the role of the game controller. Figure 3 illustrates a number of screenshots of the game in action. Figure 4 shows the game as it is being developed.

The goal of the game is to help users improve their pressure sensitivity by controlling a flying bird that soars higher based on the amount of pressure applied by the index finger of the user. The pressure applied is the only form of input. The gameplay is inspired by the infinite runner genre of games; although the play sessions have been limited in the interest of time. The objective of the game is for the bird to collect coins that appear randomly at three possible heights within the environment. The coins are randomly generated in different quantities from 5 to 15. Auditory feedback is also provided on successful coin collection.

The heights chosen correspond to the application of three levels of force. The three force levels are 2, 3 and 5 Newtons coinciding with very light to medium pressure. These forces were established through a pre-study with medical professionals in a process discussed further in Section IV-A. It is crucial for a medical student to control his hand in a dexterous manner to perform different abdominal palpation tasks. Hence, a very light amount of force such as 2 Newtons could be extremely challenging for a novice. Design of different tasks in game-based and application-based approaches were established on these guidelines.

A collision detection function is implemented to detect if the bird avatar hits a coin. The box collider used for coin to bird intersection has a buffer equivalent to ± 0.25 Newtons along the height dimension. Players should collect more coins in order to achieve a better score at the finish line. A slight increase in the flying speed during the game as well as random generation of coins makes it more challenging for the player and has the goal of keeping the game interesting. An information panel is placed on the top right corner of the game screen to show which the next coin level is. This is demonstrated in Newtons for the user to form an association with the amount of pressure to be applied and the numeric value of the force. The information panel also provides a final score at the finish line. The screen height is normalised to represent 0 to 10 Newtons from bottom to top. Although, the player could reach higher levels of force by pressing harder, the limitation of 10 Newtons was chosen to meet safety regulations.

IV. Experiment

In order to evaluate whether the game discussed in the previous section improves pressure sensitivity a participant-focused experiment was run. This experiment intends to explore if visual and auditory feedback of applied forces in the form of a game-based training approach could improve motor learning and control abilities on non-medical participants.

A. Method

A between participants design was chosen for the experiment. Participants were divided into two groups: Group A ($n = 15$) to be trained by the serious game and group B (n=15) a control group. All participants had been asked to apply force from their index finger tip while sat (in the stand up position kinesthetic help from shoulder may produce variation in the exertion of force) on a table as rigid surface. The exerted force values were sampled for each target force level for 10 seconds with 10 millisecond intervals resulting in 100 samples per each target force level. The three forces of $2, 3, 5$ Newtons that were the target goals where the coins were set in the game (as discussed earlier) were based on a pre-study in which medical tutors' pressure while palpating patients was captured. These studies involved the use of four medical professionals examining five different participants acting as patients composed of both genders and three body types. The data capture consisted

Back to Contents

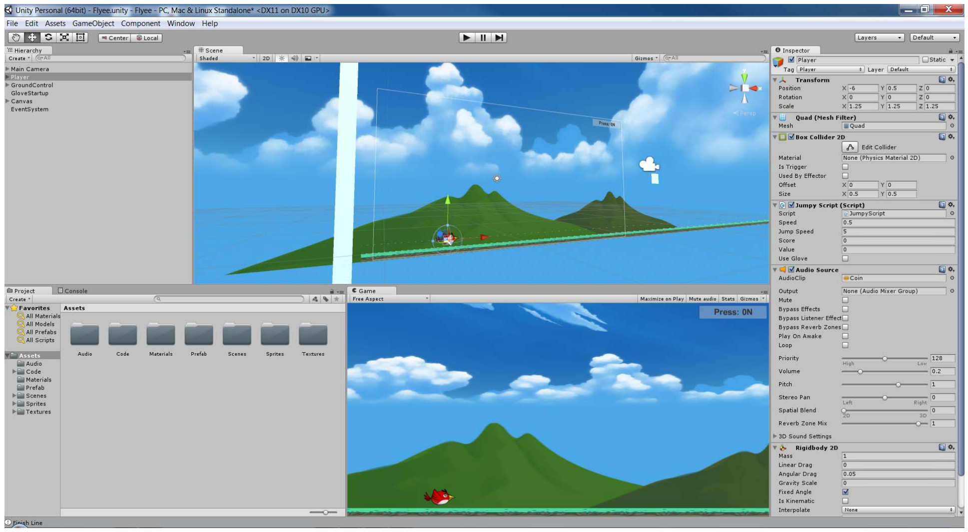


Figure 4. The game being developed within the Unity environment.

of all the medical professionals completing three palpation tasks (liver edge, deep and superficial) for all the patients. All data was captured and analysed and the goals of $2, 3, 5$ Newtons were identified based on the mean force across the medical professionals in each task across the body types.

Table I shows an overview of the experimental design and training methods for each group. The familiarisation phase allowed the participants to acquaint themselves with the equipment and saw the actual value they were pressing on DigiScale. In the final test (no visual feedback) the participants could not see how much pressure they were applying on the display and had to rely only on their pressure sensitivity training. The difference between the target value and the recorded value (in Newtons) for the no visual test was used as the dependent variable. The null hypothesis H0 in this experiment was that there is no difference between the two groups in the accuracy of the exerted target force for the no visual feedback test session. The software used for the familiarisation DigiScale was significantly different from the environment found in the game to avoid any bias of familiarisation that may have led to the game playing group to have an unfair advantage during the testing phase.

Table I. EXPERIMENTAL DESIGN

	Group A	**Group B**
Training	Visual Feedback (GUI)	Visual Feedback (Game)
Familiarisation	Visual Feedback (GUI)	Visual Feedback (GUI)
Test	No Visual Feedback	No Visual Feedback

B. Materials

The primary materials used correspond to the three technologies discussed in Section III. DigiScale was used to convert the raw sensor value from the glove to force in Newtons and to provide visual feedback on the exerted force by the user for the Training and Familiarisation phases. Two TFT displays were used in duplicate mode to provide visual feedback for each participant and to let the research investigator monitor the experiment's progress. An ultra thin powder coated polyvinyl glove was used to meet hygiene requirements prior to provide the measurement glove to participants.

C. Participants

Thirty participants took part in this experiment in two groups of fifteen with seven females and one left handed participant. Participants all had normal or corrected to normal vision. Participants were members of staff or students contacted via internal university email. A Participant Information Leaflet and related ethics documentation were attached to the invitation email before the experiment day to debrief the participants about details prior to the experiment. Participation in this experiment was entirely voluntarily with the right to withdraw at any point.

D. Procedures

Each participant had been debriefed about the experimental steps by research investigators and via email prior to data collection. Each participant was asked to wear a powder coated ultra thin polyvinyl glove and confirm if the sensor on the index finger tip was positioned correctly.

Group A had 5 minutes training with DigiScale application. Group B had the same duration of training with the game in three rounds with one minute intervals between them. The reason for repeating the training for three attempts for the game group was to provide the player with variety within the game environment as aspects of the game are randomised.

In the familiarisation phase, which occurred soon after the training session, participants were asked to attempt to meet target force levels with the aid of a dedicated display that provided visual feedback via the DigiScale application. In the final no visual feedback test, the display was switched off and results collected for each participant. There was a one minute interval between training and tests to avoid human fatigue. For

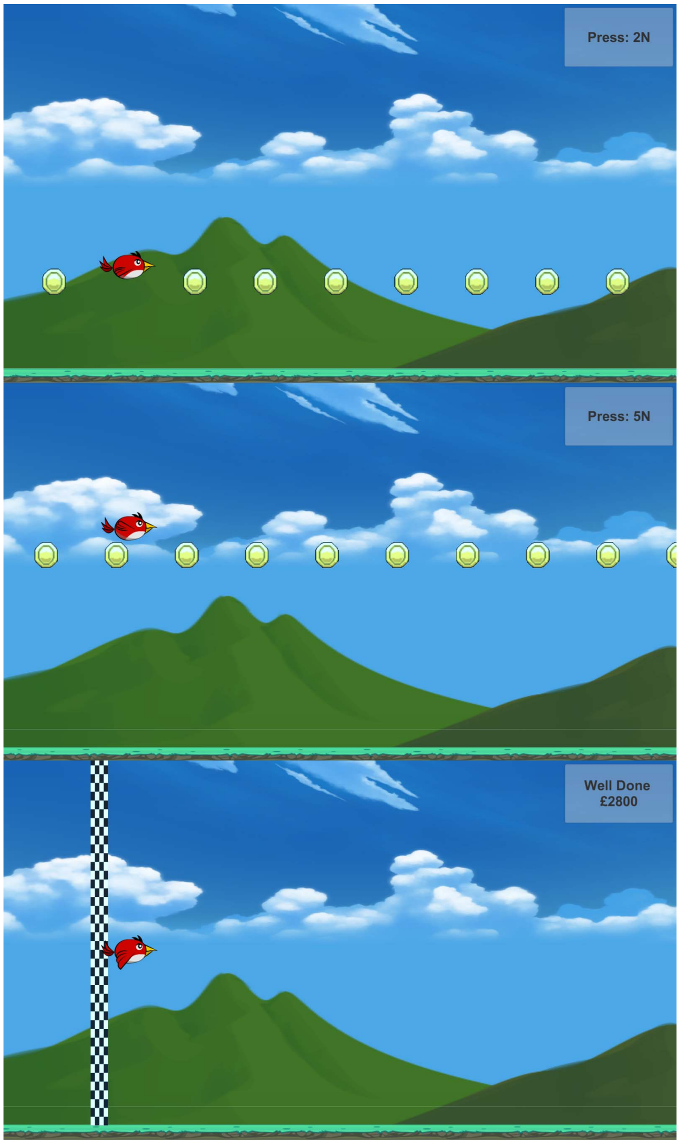


Figure 3. A game-based training approach is proposed by this experiment to help users improve their pressure sensitivity by controlling a flying bird that soars higher based on the amount of pressure applied by the index finger of the user. The final figure demonstrates the score that is displayed at the end of a run. Each captured coin is worth 100.

each target force the following objectives need to be achieved by each participant in each test.

- To reach the given target force level
- To maintain that target force level for 10 seconds.

E. Results

Results for each target force was obtained via the difference in target and recorded force for each observation. The mean exerted force (μ_i) for each target force level (f_i) is calculated from collected samples for each participant. The absolute difference from the target force is calculated as:

$$\delta_i = |\mu_i - f_i|$$

The mean of the delta values for all three target force levels f_i (2, 3 and 5 Newtons) were computed as a final result for each participant. A non-parametric Mann-Whitney test has been selected to analyse the results due to the non-parametric nature of the data.

The accuracy in the exerted target force for the no-visual test for participants in group B, who were trained using our game approach ($Mdn = 0.86$), differed significantly from the participants in group A, who trained using the application only approach ($Mdn = 1.56$), $U = 61$, $z = -2.137$, $p < .05$, $r = -0.36$ and thus H0 is rejected.

This result may highlight the potential role of game-based training on cognitive and control motor learning abilities. One possible reason for this achievement is an improvement in the understanding of the approximate force and sensitivity for the required pressure instilled while playing the game.

Another potential advantage of game-based training is the competition factor characteristic of games. It was observed during the experiment that participants in group B were keen to beat their previous best score in each round which may have led to better focus and concentration on the requested test.

F. Qualitative Feedback

In order to form an understanding of whether the game was considered an enjoyable experience, and whether it was well designed and engaging a number of questions were asked to the group that played the game. An electronic questionnaire was sent via email to each participant in the game group (n=15) to collect their reflective feedback on their experience when playing the game. A total of six questions were asked to rank key features of the game from 1 to 5 (e.g for first question the answer is made from Not at all, Slightly, Moderately, Very, Extremely "Enjoyable"). Fourteen out of fifteen participants replied and a mean scores for each question are reported in Table II.

Table II. QUESTIONNAIR

Questions	Score
Did you enjoy playing this game?	3.85
Did you engage with this game?	3.85
How would you rank this game in terms of design?	3.85
Did you feel any improvement in controlling your force level each time you have played this game?	3.77
Would you play the game again in future?	3.46
Would you recommend this game to a friend?	3.54

Using rounded mean scores for a general evaluation the game can be deemed to be very enjoyable to play, very engaging, well designed, and with the ability to provide a perceived increase in motor ability. Participants also considered that they were likely to play the game again given the opportunity and would very much recommend it to a friend. On the whole, based on this feedback, the game design seems to have been for the most part successful.

V. Conclusions and Future Work

This work has presented a serious game that attempts to teach participants the correct application of pressure by controlling a virtual character on screen via a pressure sensitive input device that rewards players with accurate and controlled input. The results demonstrate that those players that played the game performed significantly better than a control group in a subsequent no-visual task within a very different environment from the game itself. Moreover, questionnaire responses indicate that the game is enjoyable and engaging. It is important to indicate that while this game has appeared to have been successful it is part of a larger framework that is required in order to make automated or assisted palpation training successful.

Future work will look into enhancing this experience by using all the input sensors on the glove, and the capability of the system to capture location and orientation data. This can be then used for the development of a serious simulator or serious game. Furthermore, palpation is not the only application that requires pressure sensitivity and modifications to the main game to adapt to the range of sensitivities of various applications can aid pressure sensitivity training in other fields eq. training for musical instruments.

Acknowledgment

Grateful acknowledgement is made to members of staff and students in WMG for their participation and support. We would like to thank Unity and Creepy Cat for making available the 2D Game Starter Assets used to prepare the game used in this publication. Debattista is partially supported by a Royal Society Industrial Fellowship.

References

[1] D. Wortley, "Immersive technology strategies," *Simulation & Gaming*, vol. 44, no. 2-3, pp. 452–465, 2013.

[2] S. Arnab, I. Dunwell, and K. Debattista, *Serious Games for Healthcare: Applications and Implications: Applications and Implications*. IGI Global, 2012.

[3] A. Sliney and D. Murphy, "Jdoc: A serious game for medical learning," in *Advances in Computer-Human Interaction, 2008 First International Conference on*, Feb 2008, pp. 131–136.

[4] T. Susi, M. Johannesson, and P. Backlund, "Serious games: An overview," 2007.

[5] J. Macleod, G. Douglas, E. F. Nicol, and C. E. Robertson, *Macleod's clinical examination*. Elsevier Health Sciences, 2009.

[6] V. Patel and J. Morrissey, *Practical and professional clinical skills*. Oxford University Press, 2011.

[7] J. Goodwin, "The importance of clinical skills," *BMJ*, vol. 310, no. 6990, pp. 1281–1282, 1995.

[8] R. J. Duvivier, K. van Geel, J. van Dalen, A. J. Scherpbier, and C. P. van der Vleuten, "Learning physical examination skills outside timetabled training sessions: what happens and why?" *Advances in health sciences education*, vol. 17, no. 3, pp. 339–355, 2012.

[9] M. Gotsis, "Games, virtual reality, and the pursuit of happiness," *Computer Graphics and Applications, IEEE*, vol. 29, no. 5, pp. 14–19, Sept 2009.

[10] S. Scarle, I. Dunwell, T. Bashford-Rogers, E. Selmanovic, K. Debattista, A. Chalmers, J. Powell, and W. Robertson, "Complete motion control of a serious game against obesity in children," in *Games and Virtual Worlds for Serious Applications (VS-GAMES), 2011 Third International Conference on*, May 2011, pp. 178–179.

[11] C. Schonauer, T. Pintaric, H. Kaufmann, S. Jansen Kosterink, and M. Vollenbroek-Hutten, "Chronic pain rehabilitation with a serious game using multimodal input," in *Virtual Rehabilitation (ICVR), 2011 International Conference on*, June 2011, pp. 1–8.

[12] S. Saini, D. Rambli, S. Sulaiman, M. Zakaria, and S. Shukri, "A low-cost game framework for a home-based stroke rehabilitation system," in *Computer Information Science (ICCIS), 2012 International Conference on*, vol. 1, June 2012, pp. 55–60.

[13] M. N. Kamel Boulos. (2011) Healthcybermap. [Online]. Available: http://www.healthcybermap.org/

[14] M. Graafland, J. M. Schraagen, and M. P. Schijven, "Systematic review of serious games for medical education and surgical skills training," *British Journal of Surgery*, vol. 99, no. 10, pp. 1322–1330, 2012. [Online]. Available: http://dx.doi.org/10.1002/bjs.8819

[15] I. Dunwell and S. Jarvis, "A serious game for on-the-ward infection control awareness training: Ward off infection," *Serious Games for Healthcare: Applications and Implications: Applications and Implications*, p. 233, 2012.

[16] P. M. Kato, S. W. Cole, A. S. Bradlyn, and B. H. Pollock, "A video game improves behavioral outcomes in adolescents and young adults with cancer: a randomized trial," *Pediatrics*, vol. 122, no. 2, pp. e305–e317, 2008.

[17] E. Carmeli, J.-J. Vatine, S. Peleg, G. Bartur, and E. Elbo, "Upper limb rehabilitation using augmented feedback: Impairment focused augmented feedback with handtutor," in *Virtual Rehabilitation International Conference, 2009*, June 2009, pp. 220–220.

[18] M. N. Kamel Boulos, L. Hetherington, and S. Wheeler, "Second life: an overview of the potential of 3-d virtual worlds in medical and health education," *Health Information & Libraries Journal*, vol. 24, no. 4, pp. 233–245, 2007. [Online]. Available: http://dx.doi.org/10.1111/j.1471-1842.2007.00733.x

[19] K. E. Brown, J. Bayley, and K. Newby, "Serious game for relationships and sex education: Application of an intervention," *Serious Games for Healthcare: Applications and Implications: Applications and Implications*, p. 135, 2012.

Back to Contents

Classification of Affective Data to Evaluate the Level Design in a Role-playing Videogame

Fabrizio Balducci, Costantino Grana, Rita Cucchiara
Dipartimento di Ingegneria "Enzo Ferrari"
Università degli Studi di Modena e Reggio Emilia
Via Vivarelli 10, Modena (MO) 41125, Italy
Email: {name.surname}@unimore.it

Abstract—**This paper presents a novel approach to evaluate game level design strategies, applied to role playing games. Following a set of well defined guidelines, two game levels were designed for Neverwinter Nights 2 to manipulate particular emotions like boredom or flow, and tested by 13 subjects wearing a brain computer interface helmet. A set of features was extracted from the affective data logs and used to classify different parts of the gaming sessions, to verify the correspondence of the original level aims and the effective results on people emotions. The very interesting correlations observed, suggest that the technique is extensible to other similar evaluation tasks.**

I. Introduction

The concept of playing is strictly linked to our lives as one of the first activities that a human experiences, discovering his surrounding world: one method to learn rules or how to safely execute a task is to see it like a gaming activity, gaining experience for future goals [1].

Frasca [2], referring to Csíkszentmihályi's studies [3], defines the modern concept of *ludology* like the "science of the game" which uses research methods and theories from a wide range of scientific communities such as HCI and psychophysiology, with the aim to improve the methodologies to study both players and (video)games: this should help to understand the processes involved in the design of an optimal *player experience*, i.e. the set of feelings and opinions that come from an activity which has to capture the player's senses offering amusement and engagement.

In recent decades the videogame has become a modern entertainment medium with expensive budget, which can rival with cinema, music, comics and literature; considering that every person is different because of factors like age, education, country and personal history it becomes crucial attempting to define and manipulate complex concepts like "the fun" considering also the different player's feelings and preferences. Nacke [4] consider the *affective ludology* referring to investigations of affective player-game interaction to understand emotional and cognitive player experiences: it should investigate cognition, emotion, and goal-oriented behavior of players from a scientific perspective by establishing rigorous methodologies (e.g. psychological player testing or physiological response analysis of players).

Mandryk and Inkpen [5] have successfully demonstrated how electroencephalography can evidence human emotions and cognitive activity during gaming. A brain-computer interface (BCI) is a communication systems that measures brain activity with sensors positioned on the subject's scalp: the brain is a dense network of neurons, nerves cells which communicate using electrical signals obtained from chemical reactions so it's possible observe and identify scalp's electrical potentials variations and retrieving information from the associated brain waves; this is a safely process because EEG is a passive measuring device and less invasive than other methods.

This paper describes a set of guidelines for the affective level design and manipulation of emotions in a Role Playing Game. We investigate how these can be applied to practical design rules and then make use of a BCI to experimentally test the users response. In this way we try to extract objective data from subjective emotions, and use these to uniquely connote the tasks and the areas developed for the different game levels.

In Section II recent related work, which use videogames to study player's behaviors are reported. Section III proposes the formal guidelines and their practical application into game level design, while Section IV briefly illustrates the technical tools used. Section V describes the experimental setup, while Section VI shows and comments the results coming from the data analysis. Finally conclusions and suggestions for future work are drawn in Section VII.

II. Related work

Nacke *et al.* [6], [4] were the first to formalize game level design guidelines and to employ BCI to evaluate them, using the first-person shooter (FPS) *Half-Life*; also in Nacke and Lindley [7] the opportunity to deal with the events and tasks of a game level is considered.

Plotnikov *et al.* [8] investigate the brain waves for the concepts of boredom and flow using the *Tetris* videogame; in the work of Nijholt *et al.* [9] there is an extensive survey on current BCI research about HCI studies with the viewpoint of games and game design, featuring suggestions and proposals.

Herrewijn *et al.* [10] use surveys to analyze the factor of "Immersion" in the videogame *Fallout:New Vegas* while Reuderink *et al.* [11] voluntarily cause frustration to the gamers with an adapted clone of *Pacman* also analyzing their EEG data; in Canossa *et al.* [12] there is a dissertation about metrics to discover the frustration factors during a gaming session using internal game data. The EEG data are also

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

used in Lee and Tan [13] to classify the tasks depending on their cognitive and mnemonic effort with the videogame *Halo*, while Smith *et al.* [14] focus the analysis on the interfaces, to improve their usability with complex tasks. Salminen and Ravaja [15] use the videogame *007:NightFire* to inquire about the virtual violence.

A BCI headset can be used not only to gather neurological data, but also as a real-time input device: in Pour *et al.* [16] the headset is used to control the interaction in a clone of the videogame *Breakout*.

III. Level design for affective ludology

The design of a game and its levels depends on the *gameplay*, i.e. the set of "internal rules" that characterize the specific game dynamics and so denote particularly every game genre (FPS, platform, puzzle,...). According to the definition provided by Brown and Cairns [17] there are three phases of participation (engagement, engrossment, total immersion) in which the concept of immersion and flow are very close. Nacke *et al.* [6] provided a set of affective design guidelines and developed three game levels with the distinction between Immersion and Flow by dividing the concept of environment from that of combat which constitutes the focus of the FPS gameplay.

When considering RPG gameplay features, the focus moves from just action and shooting challenges to other different aspects:

- the interaction using dialogues
- the presence of allies in a group called "party"
- the developing of moral choices and riddles

The gameplay of an RPG is less immediate than FPS because it's based on statistics, bird-flight view, objects inventory and character's skills development, with a great importance given to the environment's exploration to reach the "sense of wonder". For these reasons immersion can be treated as a necessary condition to reach the flow emotional state ($Immersion \subset Flow$). In order to stress this evolution of the participation, only two game levels have to be developed for Neverwinter Nights 2: one will focus on boredom, while the other will stress flow.

A. Boredom game level

Fisher [18] defines boredom like a transitory unpleasant affective state with lack of concentration and difficulty during a task, Csíkszentmihályi [19] furthermore denotes it like a state in which player's skills are greater than required. Formal design guidelines:

Guideline 1: Assuming a spatial progression point p to indicate game level progression and given a set E of n types of enemy e_a with constant strength a smaller than player strength $b > a$, then the player challenge function $f_{ch}(p)$ for an encounter k with enemies type e_a at progression point p should remain constant $\forall e_n \in E$.

Guideline 2: Given n textures $t \in T = \{t_1, t_2, ..., t_n\}$ in a game level L, then L becomes less interesting if contains less elements of T so that $L = \{t_k, ..., t_m\} \subset T$, $m < n$; the same should be for all other structural and visual features like sounds, animations, light and visual effects, trees, enemy types, items, 3D models, weapons.

Fig. 1. The tasks "text dialogue" and "single fight" in the boredom game level.

Guideline 3: For each progression unit p the player should be constantly rewarded with $n = p$ items with i_h that refreshes player health value and i_a for the ammunition value so that the full player ammunition supply value $A = \sum_{a=0}^{n} i_a$ and the full player health value $H = \sum_{h=0}^{n} i_h$ are often reached.

Guideline 4: No real winning condition C_k so that $C = \emptyset$ and the victory goal $V = \{C_1, C_2, ..., C_n\} = \emptyset$ doesn't give any reward r_v to the player.

Guideline 5: Given n characters a with a set S of skills belonging to a specific class t which can be in $C =${warrior, wizard, thief, animal, unskilled}, if a_t can join the player's party G becoming his ally $a_t \in G = \{a_{t1}, a_{t2}, ..., a_{tn}\}$ in a game level L, then L becomes less interesting if contains much less elements of G so that $L = \{a_{tk}, ..., a_{tm}\} \subset G$, $tm \ll tn \vee \exists a_t$ $t =$*unskilled*.

Guideline 6: Given n generic dialogues $d \in D = \{d_1, d_2, ..., d_n\}$ that a player can have with the NPCs of the game level L, then L becomes less interesting if contains much less elements of D so that $L = \{d_k, ..., d_m\} \subset D$ $m \ll n$; given also n questions $q \in Q = \{q_1, q_2, ..., q_n\}$ that the player can choose during an articulated dialog $d_a \in D$, assuming that the relative answer to q_k is a_k, then L becomes less interesting if the dialogues contain much less elements of Q so that for the most part of D is $d_a = \{q_k, ..., q_m\} \subset Q$, $m \ll n$.

Referring to the previous formal guidelines, three areas have been designed by editing the *2311_tunnels* level, contained in the *2300_Crossroad_Keep_Adv* package of the original game, because it has an environment with linear paths, the prevalence of gray and blue colors, few structures and NPCs (Fig. 1).

The challenge offered by this game level is minimal with constant difficulty: there are only two types of weapons and enemies (with low level skills of resistance and attack) and player's health and weapon's munitions are always at the maximum, since every enemy releases an health potion at his death and every chest releases munitions and weapons; the story plot is really basic with a very simple task to accomplish, without victory conditions or gratifications.

The structural and visual features are poor with only indoor environments to emphasize the linearity and repetitiveness of the areas (the path in the third area is the reverse of the one in the first area); there are no visual effects or music and sounds for battle and environment, and every NPC has the same 3D model. The allies are fundamentally unnecessary and tedious, consisting in an unskilled not interactive human character, without Artificial Intelligence, which only follows the player. Dialogues are performed only in "text version" without animations or camera changes: they are unbranched and useless, filled with out-of-context sentences and little variety of questions and answers that don't add anything to the story plot; furthermore some of them create false expectations in the player.

B. Flow game level

In the field of (video)games the notion of flow is closely related to that of *fun* even if doesn't coincide with it: Nakamura and Csíkszentmihályi [20] observed the behaviors of chess players and sport players noting that their enjoyment derived by the mere fact to accomplish their activity independently by other types of rewards; the flow is intended like an emotional peak with total involvement and mental attention in which the skills fully meet the challenge of a task and the player is fully absorbed by it. Formal design guidelines:

Guideline 1: Given a set of indoor level parts $I \neq \emptyset$ and a set of outdoor level parts $O \neq \emptyset$, the game level L should be a set union of outdoor and indoor level parts $L = \{I, O\}$.

Guideline 2: Assuming a spatial progression point p to indicate a game level progression and given a set E of n type of enemy e_a with strength a, then the player challenge function $f_{ch}(p)$ for an encounter k with an enemy e_a at progression point p has to progressively increase $\forall e_a \in E$ (Ermi and Mäyrä [21]).

Guideline 3: Given n sets $X_k \neq \emptyset$ of different structures in a game level L like textures X_t, lighting effects X_l and sounds X_s, then L becomes more atmospheric and fosters imagination if $L = \{X_t, X_l, X_s..., X_n\}$; the same should be for all other structural and visual features like animations, fire effects, trees, enemy types, items, 3D models, weapons.

Guideline 4: For a player progression unit p, after a set progression interval p_i for n intervals in a game level $L = \{p_1, p_2, ..., p_n\}$, a reward type r_p from a set of n rewards $R = \{r_1, r_2, ..., r_n\}$ should be given to the player such as ammunition, health packs, money, new weapons and spells.

Guideline 5: There's at least one winning condition C_k of the victory goal $V = \{C_1, C_2, ..., C_n\}$, $V \neq \emptyset$; the achievement of V must lead to one or more rewards r_v from a set of n rewards $R = \{r_1, r_2, ..., r_n\}$ in order to gratify the player's efforts.

Guideline 6: Given n characters a with a set S of skills belonging to a specific class $t \in C$ ={warrior, wizard, thief, animal, unskilled}, if a_t can join the player's party G becoming his ally $a_t \in G = \{a_{t1}, a_{t2}, ..., a_{tn}\}$ in a game level L, then L becomes more interesting if $L = \{a_{tk}, ..., a_{tm}\} \subseteq G$ with $G \neq \emptyset \wedge \forall a_t \in G \; t \neq unskilled$.

Guideline 7: Given n generic dialogues $d \in D = \{d_1, d_2, ..., d_n\}$ that a player can have with the NPCs of the game level L, has to be $L = \{d_k, ..., d_m\} \subseteq D$; given also n questions $q \in Q = \{q_1, q_2, ..., q_n\}$ that the player can choose during an articulated dialog $d_a \in D$, assuming that the relative answer to q_k is a_k then the game level L becomes more interesting if for the most part of D is $d_a = \{q_k, ..., q_m\} \subseteq Q$ and d_a is a *narrative or ludic game level structure* useful for reaching a condition of victory V or to achieve useful information and items.

Fig. 2. The tasks "animated dialogue" and "group fight" in the flow game level.

Based on these guidelines, four areas have been designed by editing the *3010_highcliff*, *3000_castle_never*, *3063_merchant* and *3032_th_canyon* levels contained in the *3000_Neverwinter_A3* package of the original game. Linking the areas, the original enemies, paths, quests and dialogues have been replaced to create a new coherent story plot; three original dubbed dialogues are preserved (Fig. 2).

The game level must be characterized by the balance between difficulty and player's skills, in fact if the challenge is higher than the ability the activity becomes overwhelming and generates anxiety, while if the challenge is lower than the player's ability, it provokes boredom. In order to continue to experience flow, a person must continue to learn new skills [22]: the challenges steadily increase due to the variety of enemy types provided by every area (better statistics, weapons, and skills). The enemies can also attack in groups with ambushes, furthermore the previous encountered enemies grow and become harder to defeat.

Environment structures are diversified with indoor and outdoor areas, each of which has a specific setting (cliff, castle, haunted town, volcanic caves) and some of them offer optional paths. The structural and visual features are rich of 3D models, items, effects (lights, animated water and trees), particular music and sounds with alternation between day and night and weather conditions.

Rewards are balanced with the player's skills and the challenge offered, in fact only few enemies (the harder to defeat) release money or health potions at their death, and the chests which release magic objects, ammo and weapons are located only before or after the more challenging combat sections; there are wearable objects which increase player's skills which can also grow using the experience points that the character gains. Within the different types of weapons and equipment, also spells are included.

The story plot is rich, with moral and emotional entanglements and well-defined and satisfactory goals (a rescue, an enrollment, a discovery) and related rewards; some tasks are also linked to the player's choices and influence the prosecution of the story. There is a party of skilled allies, useful and diversified by their particular skills (a wizard, a warrior, a thief) with different ways to enlist them (automatic, payment, dialogue). All the allies in the party are interactive with a proactive Artificial Intelligence also additional non-interactive NPCs (guards and peasants) may help the player to perform some tasks.

Dialogues are useful and articulated: they use animations and shots and provide useful information for the plot (memories of events, opinions, riddles, narrative elements). Moreover, the conclusion of the quest (positive or negative) depends on the moral choices (greed, compassion) made in some of them; some dialogues have professional voice dubbing which increases player's emotional involvement.

IV. The technologies

The Emotiv EPOC Headset is a cheap neuro-signal acquisition and processing wireless headset with 14 wet sensors and 2 reference sensors which is capable of detecting brain signals: the Control Panel provides an interface which allows monitoring in real-time the headset's signals implementing the *Affectiv Suite* for visualizing in charts the five detected emotions (Engagement, Frustration, Meditation, Excitement and LongTermExcitement).

The choice of using Neverwinter Nights 2 for the game level design was made considering several factors:

- the *RPG* gameplay which offers variety of story plot, characters and events
- a complete editor (*Electron*) used even by the authors which provides visual tools to create dialogues, NPCs (Non-Player Characters) with Artificial Intelligence, environments, triggers, objects and visual effects
- the powerful C-like scripting language *NWScript* which allows to manage internal game's data and variables
- the possibility to customize the game interface with XML

For the experimental phase there are two types of data to extract during the gaming session:

- internal: environmental values, character's values, map coordinates
- external: player's affective data which come from the headset

The data sources must be automatically stored and synchronized with a timestamp to facilitate their analysis. The open source C++ tool NWNX[23] allows the implementation of a *plugin* which manages the interchange of data between the game, the screen and the "outside world": developed for the online gaming, it uses the server version of the game and permits the calling of the plugin's code from a NWScript and its integration with the code libraries of the headset's SDK .

Screen captures of the gaming sessions can be also useful for the experiments so an *XML UIPane* widget will enable NWNX printing in real-time on the game interface the same timestamp used to synchronize the internal data.

V. Experimental study

Laboratory experimentation involved 13 volunteers male academic students tested in a range hour between 10.00 AM and 07.00 PM on weekdays. A single session was about one hour and half in which the subject had to play the two game levels wearing the headset, and every gaming session was recorded using a screen capture software.

At the beginning of a session a brief explanation about the modalities of the study was provided, followed by the compilation of an informed consent and of a pre-questionnaire about subject's gaming preferences; every subject played using the same game character (a neutral half-elf warrior with all the character statistics set to an average value).

Next a step-by-step game tutorial consisting of two areas with brief tasks was presented, with the aim of learning commands and game rules familiarizing with the user interface. Some of the offered tasks were: "move the character", "rotate and zoom the screen view", "have a chat with someone", "take the control of an ally", "open the inventory or a chest and take an item", "launch a spell", "fight", "rest to recover health".

After the tutorial, the headset was put on subject's scalp with a phase of signal tuning and the starting of the data gathering software. Following the first gaming session the subject compiled a post-questionnaire about his gaming experience. The next game level was then played with the compilation of a new post-questionnaire and finally the headset was removed from the subject.

Different data sources were available for every gaming session:

- an MPEG file with the gaming session screen capture
- the pre and post questionnaires
- two log files with data coming from the game and from the headset, synchronized by a timestamp

Every entry of the log for the *events data* (Fig.3-1) contains the timestamp and the tag of the meaningful event occurred during the gaming session captured using internal NWNScript game triggers.

Every entry of the log for the *affective data* (Fig.3-2) contains the timestamp, the name of the area in which the player was with its coordinates (X and Y values) and the punctual values of the emotions that the headset recognized: these ones are in a range between 0 and 1 and classified by the internal algorithms of the Emotiv EPOC.

A schematization of the emotions available to understand the affective data is the Circumplex model of affect (Fig. 4) by Russell *et al.* [24], [25]: in this two-dimensional spatial model every human affect could be defined on a circumference combining the bipolar dimensions of *valence* and *arousal*; the first suggests if an affect is pleasant or unpleasant (X axis), the second suggests the activation or the deactivation of an emotion (Y axis).

The *Engagement* denotes an increment of the beta brainwave and a reduction of the alfa ones [26]: by the Emotiv EPOC headset it's associated to the mood states of participation and attention and as an indicator of fun so if the aim

Back to Contents

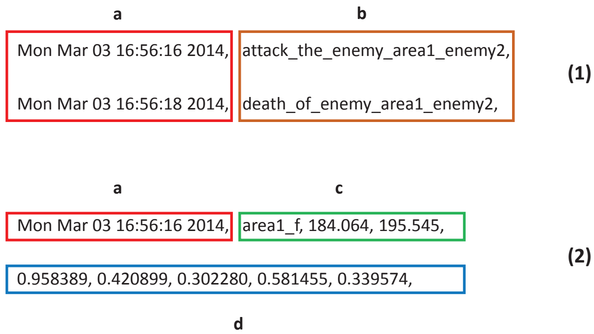


Fig. 3. (1): part of the events log with a sequence of entries featuring timestamps (box "a") and the related events occurred during the gaming session (box "b"). (2): an entry line (divided in two) of the affective log featuring the timestamp (box "a"), the character's map coordinates (box "c") and the punctual values for the five recognized emotions (box "d").

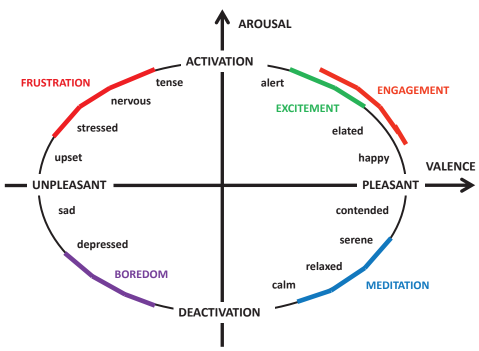


Fig. 4. An interpretation of the circumplex model of affect: Excitement is pleasant and activated so it is located in the upper-right quadrant like Engagement which is its extension; Meditation is pleasant but not activated, so it is located in the bottom-right quadrant. Frustration and Boredom are unpleasant, but the first is activated so it is located in the upper-left quadrant, while the second is in the bottom-left.

was to induce the flow state, most of the tasks and events must show an increase of it for the player's interest perceived; because *Boredom* is a deactivated emotion it can be seen as the opposite of engagement so if a game level was designed to induce it, most of the tasks and events should show a decrease of engagement.

The *Excitement* is a positive indicator linked to a short-time emotional peak so it may be unfit as an indicator for flow or boredom: it is better to consider the *LongTermExcitement* like a global value which expresses how stable over time the excitement is. As observed for engagement, if the aim was to induce boredom most of the tasks and events should show a decrease while the opposite should be true for the flow state. Engagement and Excitement are closely related to the concept of flow due to the mutual interaction between interest and involvement while facing the game tasks.

Inducing and analyzing *Frustration* and *Meditation* in game tasks is difficult because they have a negative valence and are influenced by subjectivity, for example different players can perceive the same amount of frustration as a challenge or as a reason to quit the game.

TABLE I
SUBJECT'S EVENTS COUNTS BY AREA: BOREDOM GAME LEVEL

	a1	a2	a3	Tot.
subj.1	4	5	3	12
subj.2	5	6	4	15
subj.3	5	6	4	15
subj.4	5	6	4	15
subj.5	5	6	4	15
subj.6	5	6	4	15
subj.7	5	6	4	15
subj.8	5	7	4	16
subj.9	5	5	3	13
subj.10	5	6	4	15
subj.11	5	6	3	14
subj.12	5	8	4	17
subj.13	5	6	4	15
Tot.	64	79	49	192

VI. EVALUATION OF THE EEG DATA

A game level can be considered as a sequence of challenges and situations and more specifically as the sequence of tasks and events which the player faces during his gaming session; for the evaluation it was decided to consider the tasks and the events of the areas defined in the design phase: this approach allows the interpretation of data that greatly changes over time due to the duration of a gaming session, focusing on factors like repetitiveness, plot development and modifications in the environment.

The significant tasks and events counted for the subjects are those defined in the design phase extracted from the log files: *dubbed dialogue*, *riddle dialogue*, *dialogue*, *attack an enemy or a group of them* (fight), *death of an enemy or his entire group* (fight end), *chest opened* (with its items gathered); there are two further events which are annotated using the recoded video: *character growth* (skills upgrade) and the *stealing task* with the thief character.

Because the affective data was recorded 10 times per second it is important to identify and extract the sequential data entries between the beginning and the end of a task using the timestamps, with the aid of the video recording and the events log to identify the time edges.

Considering a player session, the subject may face very different task and events (TableI and II): the boredom game level presents an average of 14.7 events while the flow level offered about the double of them (27.9) due to the different structure and development guidelines; some type of event are not present in the boredom level, like dubbed and riddle dialogues or the skills upgrade and the stealing action.

As seen in the design Section, tasks and events can be very different even if they are of the same type, in fact they depend on the kind of emotion which they want to manipulate and finally by the game level (and area) in which they are located.

TABLE II
SUBJECT'S EVENTS COUNTS BY AREA: FLOW GAME LEVEL

	a1	a2	a3	a4	a5	Tot.
subj.1	6	9	13	4	0	32
subj.2	6	8	6	5	0	25
subj.3	6	10	13	5	0	34
subj.4	7	9	7	4	1	28
subj.5	6	10	8	3	1	28
subj.6	7	8	6	5	0	26
subj.7	7	9	7	1	0	24
subj.8	7	9	12	0	0	28
subj.9	7	9	7	5	0	28
subj.10	6	10	14	4	0	34
subj.11	6	9	6	4	1	26
subj.12	5	10	12	5	0	32
subj.13	6	8	7	0	0	21
Tot.	82	118	118	45	3	366

In order to evaluate the game level design we chose to perform a supervised machine learning binary classification with the clauses of *boredom* and *flow*: in this way it is possible to investigate if the brain-recorded data is characteristic enough to differentiate among the two set of development guidelines.

The data for training is a set of points (vectors) x_i along with their categories y_i; for some dimension d, the $x_i \in R^d$ and the $y_i = \pm 1$. The equation of a hyperplane is

$$< w, x > +b = 0 \tag{1}$$

where $w \in R^d, < w, x >$ is the inner (dot) product of w and x, and b is real.

The following problem defines the best separating hyperplane: find w and b that minimize $||w||$ such that for all data points (x_i, y_i)

$$y_i(< w, x_i > +b) \geq 1 \tag{2}$$

The support vectors are the x_i on the boundary, those for which $y_i(< w, x_i > +b) = 1$.

The primary data features associated to a game level task are the emotive data retrieved by the headset, but considering the entire duration of a task measures are required which summarize the values in the log: for every affective emotion E, each of which consists of a set of n log entries related to the task, we considered:

- the values of the *angular coefficient* β and the *intercept* α from a linear regression line calculated considering *time* (T) as a positive variable which constantly increases

$$E_i = \alpha + \beta T_i + \mu_i, i : 1..n \tag{3}$$

  where μ_i is the statistical error

TABLE III
BINARY CLASSIFICATION RESULTS FOR THE TWO GAME LEVELS

	boredom	flow
original events	366 (ext.)	366
classified events	423	309
correctly classified	298	241
accuracy	81.42%	65.85%
miss rate	18.58%	34.15%
false positive rate	29.55%	22.00%
generalization error	0.2896	
bias	0.1656	

- the *Pearson product-moment correlation coefficient* considering *time* (T) in order to estimate its relationship with an *affective emotion* (E)

$$\rho_{T,E} = \frac{cov(T, E)}{\sigma_T \sigma_E}, -1 \leq \rho_{T,E} \leq +1. \tag{4}$$

  In particular the *Pearson correlation coefficient* r_e calculated for a sample of n affective values in the task's time interval considered:

$$r_e = \frac{\sum_{i=1}^{n}(T_i - \bar{T})(E_i - \bar{E})}{\sqrt{\sum_{i=1}^{n}(T_i - \bar{T})^2}\sqrt{\sum_{i=1}^{n}(E_i - \bar{E})^2}}, \quad -1 \leq r_e \leq +1 \tag{5}$$

  if the outcome of Eq. 5 is a strongly significant value ($r_e > 0.7$ or $r_e < -0.7$), then we can assume evidence of a local correlation for the selected emotion; a positive coefficient detects a direct correlation while a negative detects an inverse correlation).
- the arithmetic *mean* $\overline{E}$
- the *variance* σ_E^2

Considering the 13 subjects we have a 588x25 matrix M in which there are 588 labeled events (192 b of the boredom level and 366 f of the flow) with 25 numeric features (five for every emotion); since the set is unbalanced, the number of boredom cases was increased by randomly sampling 174 cases from the original cases.

The feature data matrix becomes 732x25 and it is used as the matrix of predictors to train a SVM (support vector machine) classifier: while by default the training uses a 10-fold cross validation a further training has been done using the option "leaveout" which trains the system n times iteratively using $n - 1$ cases and testing it with the one left out; after the new training session the system has been tested with the same matrix used to train it with the results in the Table III. The evaluation of all this measures by the classifier will suggest if a task was well-designed to induce some trend variations in the manipulation of the emotion also considering the natural presence of noise in the data.

The *accuracy* of the classification is in both game levels up to 65% suggesting a satisfactory success for the design phase: the presence of an hyperplane which splits the events

TABLE IV
MOOD-CONGRUENT EVENTS BY AREA: BOREDOM GAME LEVEL

	total number (ext.)	correctly classified	rate
area1	125	115	92.00%
area2	155	124	80.00%
area3	86	59	68.60%
Tot.	366	298	

in the geometric space of the feature highlights a strong characterization between the measures of the tasks of the same type with the 65.85% for the flow game level tasks and the 81% for the boredom (even considering the extension of its training cases). The false positive classification is under 30% for both the game levels with a greater value for the boredom (29.55%).

A. *Classification for game level areas*

Another useful aspect to evaluate the design of the game level is to consider to which game level area the correctly classified tasks belong, comparing their number with the original; the results for the areas of the boredom game level are in Table IV while for the flow in Table V. For the three areas of the boredom game level the accuracy is up to 65% with a peak for the area 1 (92%) which suggest that this area and the area 2 (80%) were well-designed by the guidelines because they have events and tasks highly characterized by their recognized emotive values: while the first two areas are the same with repetitiveness of path and visual structures the area 3 introduces some small variations which however are at least sufficient to reduce the boredom perceived by the subjects.

For the three areas of the flow game level the accuracy is variable: over the 80% for area 4 (88.89%) and area 3 and smaller for area 2 (63.56%) and area 1 (31.71%): the progression of the values suggests the growth of the subject's engagement and long-term excitement due to the development of the plot story and the challenges faced with increasing new allies and skills.

The fourth area is the shortest but results to be the best-designed due to the presence of duels, dubbed dialogues, a moral choice and a great enemy like a dragon while the third area (which is the largest) has a moral choice and an optional path in which recruiting the thief character (to perform the stealing task). The fifth area was the final stage containing only one dialogue with a 100% rating but since was reached only by 3 subjects it is not very significant for the evaluation.

B. *Classification for task and event types*

Using the log files it is possible to label every task faced by a subject: it can be useful to inquire about the types of task and events and if there are types classified better than others; the results for the boredom game level are in Table VI while for the flow are in Table VII. In the boredom game level, *dialogues* are the best recognized type (85.44%) followed by the *chest opened* with its items gathered by the players (83.50%). The classification rate is over 70%, evidence of a correct development phase and sound guidelines to increase the poorness of dialogues (performed only in texual mode) and the repetitiveness of the chest's items founding.

Also in the flow game level the dialogues are the most recognized: the *dubbed* (93.94%) and the *riddle* (84.61%) have been developed specifically for the gameplay of the flow (as the *skills upgrade* phase also up to 70%) and so their classification success justifies the goodness of the guidelines proposed in the design phase; the most frequent types are *dialogue* (articulated and performed using animations and shoots) and *single fight*, with a rate satisfactorily up to 64% as *fight vs. a group* which are less numerous.

The difference between the two type of fights (single duel or against a group of enemies) is about 0.50% so there is no justification for their differentiation in the guidelines; the *chest opened* task is the only one under 50% of classification and,

TABLE V
MOOD-CONGRUENT EVENTS BY AREA: FLOW GAME LEVEL

	total number	correctly classified	rate
area1	82	26	31.71%
area2	118	75	63.56%
area3	118	97	82.20%
area4	45	40	88.89%
area5	3	3	100%
Tot.	366	241	

TABLE VI
MOOD-CONGRUENT EVENTS BY TYPE: BOREDOM GAME LEVEL

	total number (ext.)	correctly classified	rate
dialogue	158	135	85.44%
single fight	111	82	73.87%
chest opened	97	81	83.50%
Tot.	366	298	

TABLE VII
MOOD-CONGRUENT EVENTS BY TYPE: FLOW GAME LEVEL

	total number	correctly classified	rate
dubbed dialogue	33	31	93.94%
riddle dialogue	13	11	84.61%
dialogue	110	72	65.45%
single fight	107	69	64.48%
fight vs. a group	25	16	64%
chest opened	63	30	47.62%
skills upgrade	11	8	72.73%
stealing action	4	4	100%
Tot.	366	241	

comparing it with the value for the boredom level (-35.88%), it becomes evident that the repetitiveness is more efficient than the variety of the retrieved items. The *stealing action* task was optional and performed only by 4 subjects, but it was characterized very well resulting in a rating of 100%.

VII. Conclusions

The results from experimentation are very encouraging about the usefulness of the proposed guidelines and underlines the goodness of the design phase of the two game level both for their structure (the areas) and for the implementation and conception of the tasks and events; the dialogues, the fulcrum of the RPG gameplay, prove the ability to manipulate the player's emotions positively (in the flow game level) and negatively (in the boredom game level). The factor of repetitiveness and progression (of skills and challenges), respectively at the base of the guidelines of boredom and flow, become evident considering the results about the areas classification.

In future works the analysis of the raw data provided by the headset's sensor will be very useful, in order to extract the trends of the single brain waves and compare them with the affective classification detected from the internal algorithms used there; also the different questionnaires require analysis.

The introduction of some forms of DDA (dynamic difficulty adjustment) could benefit from the results of this work: with a real-time classification of the emotions it could be possible to change the trend of an emotion during a gaming session.

The screen capture of the gaming session can be (automatically) analyzed to detect the dynamics of gameplay like strategies and behavior patterns of the gamers; also the AI of the allies can became dynamic and benefit from a mechanism of machine learning based on this patterns. Finally, given the extreme variety of videogame genres, it can be useful to test the guidelines with other games presenting a different gameplay (survival, stealth, ...) or different modalities of affective interaction (on-line with real people as ally and enemies) like MMORPGs.

References

[1] J. P. Gee, "Learning and games," *The ecology of games: Connecting youth, games, and learning*, vol. 3, pp. 21–40, 2008.

[2] G. Frasca, "Ludologists love stories, too: notes from a debate that never took place." in *DIGRA Conf.*, 2003.

[3] M. Csikszentmihalyi, "Does being human matter? on some interpretive problems of comparative ludology," *Behavioral and Brain Sciences*, vol. 5, pp. 160–160, 3 1982. [Online]. Available: http://journals.cambridge.org/article_S0140525X00010980

[4] L. Nacke, "Affective ludology: Scientific measurement of user experience in interactive entertainment," Ph.D. dissertation, Blekinge Institute of Technology, Karlskrona, Sweden, 2009, http://phd.acagamic.com, (ISBN) 978-91-7295-169-3.

[5] R. L. Mandryk and K. M. Inkpen, "Physiological indicators for the evaluation of co-located collaborative play," in *Proceedings of the 2004 ACM conference on Computer supported cooperative work*. ACM, 2004, pp. 102–111.

[6] L. Nacke, S. Stellmach, and C. Lindley, "Electroencephalographic assessment of player experience: A pilot study in affective ludology," *Simulation and Gaming*, vol. 42, no. 5, pp. 632–655, 2011, cited By 5. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-81755163956&partnerID=40&md5=d5b1a5426fb88f962d4a2feacb322ed3

[7] L. Nacke and C. A. Lindley, "Flow and immersion in first-person shooters: Measuring the player's gameplay experience," in *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, ser. Future Play '08. New York, NY, USA: ACM, 2008, pp. 81–88. [Online]. Available: http://doi.acm.org/10.1145/1496984.1496998

[8] A. Plotnikov, N. Stakheika, C. Schatten, F. Belotti, D. Pranantha, R. Berta, and A. De Gloria, "Measuring enjoyment in games through electroencephalogram (eeg) signal analysis," in *Proceedings of the 6th European Conference on Games-Based Learning (ECGBL 2012)*, 2012, pp. 393–400.

[9] A. Nijholt, B. Reuderink, and D. Oude Bos, "Turning shortcomings into challenges: Brain-computer interfaces for games," in *Intelligent Technologies for Interactive Entertainment*, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, A. Nijholt, D. Reidsma, and H. Hondorp, Eds. Springer Berlin Heidelberg, 2009, vol. 9, pp. 153–168. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-02315-6_15

[10] L. Herrewijn, K. Poels, and G. Calleja, "The relationship between player involvement and immersion: An experimental investigation." in *FDG*, 2013, pp. 364–367.

[11] B. Reuderink, A. Nijholt, and M. Poel, "Affective pacman: A frustrating game for brain-computer interface experiments," in *Intelligent Technologies for Interactive Entertainment*, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, A. Nijholt, D. Reidsma, and H. Hondorp, Eds. Springer Berlin Heidelberg, 2009, vol. 9, pp. 221–227. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-02315-6_23

[12] A. Canossa, A. Drachen, and J. R. M. Sørensen, "Arrrgghh!!!: Blending quantitative and qualitative methods to detect player frustration," in *Proceedings of the 6th International Conference on Foundations of Digital Games*, ser. FDG '11. New York, NY, USA: ACM, 2011, pp. 61–68. [Online]. Available: http://doi.acm.org/10.1145/2159365.2159374

[13] J. C. Lee and D. S. Tan, "Using a low-cost electroencephalograph for task classification in hci research," in *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '06. New York, NY, USA: ACM, 2006, pp. 81–90. [Online]. Available: http://doi.acm.org/10.1145/1166253.1166268

[14] M. E. Smith, A. Gevins, H. Brown, A. Karnik, and R. Du, "Monitoring task loading with multivariate eeg measures during complex forms of human-computer interaction," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 43, no. 3, pp. 366–380, 2001.

[15] M. Salminen and N. Ravaja, "Increased oscillatory theta activation evoked by violent digital game events," *Neuroscience letters*, vol. 435, no. 1, pp. 69–72, 2008.

[16] P. A. Pour, T. Gulrez, O. AlZoubi, G. Gargiulo, and R. A. Calvo, "Brain-computer interface: Next generation thought controlled distributed video game development platform," in *Computational Intelligence and Games, 2008. CIG'08. IEEE Symposium On*. IEEE, 2008, pp. 251–257.

[17] E. Brown and P. Cairns, "A grounded investigation of game immersion," in *CHI'04 extended abstracts on Human factors in computing systems*. ACM, 2004, pp. 1297–1300.

[18] C. D. Fisher, "Boredom at work: A neglected concept," *Human Relations*, vol. 46, no. 3, pp. 395–417, 1993.

[19] M. Csikszentmihalyi, *Beyond boredom and anxiety*. Jossey-Bass, 2000.

[20] J. Nakamura and M. Csikszentmihalyi, "The concept of flow," *Handbook of positive psychology*, pp. 89–105, 2002.

[21] L. Ermi and F. Mäyrä, "Fundamental components of the gameplay experience: Analysing immersion," *Worlds in play: International perspectives on digital games research*, vol. 37, 2005.

[22] D. O. Broin, "Using a criteria-based user model for facilitating flow in serious games," in *Games and Virtual Worlds for Serious Applications (VS-GAMES), 2011 Third International Conference on*. IEEE, 2011, pp. 63–69.

[23] *Neverwinter Nights eXtender v.4*, http://www.nwnx.org/.

[24] J. A. Russell, "A circumplex model of affect," *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.

[25] J. Posner, J. A. Russell, and B. S. Peterson, "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology," *Development and psychopathology*, vol. 17, no. 03, pp. 715–734, 2005.

[26] *Emotiv EPOC Software Development Kit and User Manual for Release 1.0.0.5, par. 3.4.3*, wiki.emotiv.com/tiki-index.php?page=Documentation.

LiverDefense: Using A Tower Defense Game As A Customisable Research Tool

Julia Brich, Katja Rogers, Julian Frommel, Martin Weidhaas, Adrian Brückner,
Sarah Mirabile, Tamara Dorn, Valentin Riemer, Claudia Schrader and Michael Weber

Ulm University, Germany
Email: firstname.lastname@uni-ulm.de
http://www.uni-ulm.de/in

***Abstract*—In this paper, we present LiverDefense - an educational Tower Defense game illustrating the basic functions of the human liver. LiverDefense serves as an easy-to-use evaluation tool for psychological game-related research since it can be customised with regard to its degree of difficulty. Furthermore, Likert scale questionnaires can be displayed throughout the game and relevant game events can be logged. All customisation is performed via human-readable XML files, abolishing the need for programming proficiency. LiverDefense has been successfully used in a user study focused on exploring the effect of perceived control over gameplay on players' emotions. We report on the analysis of this study with regard to enjoyment and frustration and the resulting insights on LiverDefense's usefulness as a customisable research tool.**

I. Introduction

Serious games as introduced by Clark C. Abt in 1968 [1] are games that are not primarily focused on providing a form of entertainment [2]. Instead, many of them aim to utilise the human drive to play as a way to teach facts, instill knowledge or train skills as demonstrated in various game studies [3], [4]. Prensky [5] found that contemporary young learners respond well to more interactive methods like serious games as they tend to find traditional learning media boring and insufficient for their needs. Mangold [6] states that virtual media are potentially useful tools to address the so-called *millenial generation*'s issues of engagement and (inter-)active learning. However, educational games need to follow certain design rules in order to be engaging as well as efficient with regard to teaching. Malone, for instance, developed a framework to promote intrinsic motivation in instructional games [7]. He suggests a number of aspects important to this goal, among them the following: Clear and meaningful goals, performance feedback, varying difficulty, support of fantasy, and adequate complexity to promote curiosity. While simply adhering to these design aspects is no guarantee for creating a motivating serious game [8], it can at least be a valuable step in the right direction.

Tower Defense games are a very successful genre of video games that employ these aspects already in a non-serious field of application. In this genre, the player's objective is to defend a given asset, e.g. a castle. The asset is assaulted by enemies, usually along a given path. In order to prevent them from reaching the castle, the user may place defensive units, so-called towers, along the path. The player takes damage for every enemy unit that reaches the end of the path. The goal is to survive as many enemy waves as possible while these become more and more menacing. Due to cost and placement issues, strategic planning is required. Furthermore, each type of enemy may require its own type of tower to be dealt with. For more information on the genre, refer to Avery et al. [9]. They conducted a comprehensive analysis of the Tower Defense genre, its history, and its applicability for computational intelligence research.

Since Tower Defense games inherently comply with Malone's goals for successful educational games as stated above, the genre harbours a lot of potential to serve as basis for effective serious games. To explore this possibility, we developed LiverDefense, a Tower Defense game in the biological-medical context. We intend LiverDefense to serve as an easily customisable platform for psychological studies in the field of game research. Thus, we decided to focus on customisability with regard to game difficulty, and providing researchers with the opportunity to comfortably add, edit, and display questionnaires during the game.

We also report on a study where LiverDefense's customisation capabilities were successfully used to research the effect of varying degrees of difficulty on player's perceived control and emotions during play.

The rest of the paper is organised as follows. We provide a review of related literature relevant to our approach. Subsequently, we present the concept and implementation details of our game LiverDefense and highlight the customisable features. Finally, we report on a first study that successfully used LiverDefense as an evaluation tool.

II. Related Work

Sawyer [10] reports on the general applicability of games in the health sector and presents a number of possible application areas. Wattanasoontorn et al. [11] reviewed 108 academic and commercial serious games for health and developed four criteria for classification: main purpose, targeted health care phase, intended target group, and game technology. They performed a detailed classification of all 108 games and found a high degree of diversity in the field. McCallum [12] reports in more detail on the research challenges for serious health games, such as measuring effects due to varying game experience for players, the necessary quality of health games to be engaging, and the effect of extrinsic versus intrinsic motivators. Papastergiou [13] reviewed thirty-four articles with regard to the usefulness of video games in health education and found them to have many potential benefits, among them the

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

ability to engage students. Her findings indicate however that empirical evidence pertaining to the educational effectiveness of health-related video games is still sparse. Baranowski et al. [14] conducted a survey on twenty-seven articles regarding health-related behavior change and found that video games can indeed have a positive influence. However, they conclude that more research is needed on the optimal use of game elements such as interactivity to promote beneficial changes.

With regard to the application of Tower Defense games in the health sector, we found the following projects particularly interesting. Thompson developed Food Fight, a diabetes-related tower defense game [15]. The game aims to teach players the necessary balance of a healthy diet and steady exercise with regard to diabetes management. Various foods embody the incoming enemies, while the player has to place *diet* and *exercise* towers. Motivation is used as the game's currency. In a novel approach, diet towers do not destroy enemies. Instead, they symbolise the dietary choices of the player and influence the types of incoming food. For instance, placing a carbohydrate diet tower results in less sugar-rich foods, i.e. enemies, entering the playing field. However, the healthier the diet, the more motivation it costs to uphold. Motivation can only be regained by using exercise towers to destroy the approaching enemies. To promote beneficial self-care behaviour, players are encouraged to keep their virtual blood sugar at a stable and healthy level. Hence, points are awarded for keeping a horizontal blood sugar meter close to the optimal middle value.

Bassilious et al. [16] focused on diabetes and the Tower Defense genre as well. They developed the game *Power Defense* to teach diabetes type 1 patients the numeracy skills needed for insulin management. The focus of the game lies on experiential learning. They chose to use primarily implicit teaching methods by employing the metaphor of an unstable power plant to symbolise the body's blood sugar level. Players learn the necessary calculations by deciding in real time how much additional energy (e.g. food) may enter the facility to prevent power failure and how much coolant (e.g. insulin) is necessary to avoid a blowup. In addition, explicit methods like diabetes-related quizzes are used to convey and manifest knowledge. Bassilious performed an evaluation of PowerDefense with 44 adolescent participants [17], but did not find any positive effects regarding diabetes management. He does however list numerous limitations of the study himself which might have prevented him from discovering the desired effects.

Clements et al. [18] developed a game that allows players to experience the workings of the human immune system first-hand. They chose the genre of Tower Defense as they found its mechanics most suitable to illustrate the underlying principles of immune response. The goal of the game was to prevent the infection of healthy cells. To achieve this, players could specialize generic immune system cells to defend against pathogens entering the blood stream. Clements et al. chose a two-dimensional approach and cartoon-like animations to award the game a fanciful atmosphere while still retaining biological accuracy as much as possible. This game was designed according to Malone's recommendations [7] for use in an actual teaching environment; however, Clements et al. do not report on any findings pertaining to the reception and efficiency of their approach.

Avery et al. researched Tower Defense games with regard to their usefulness for computational intelligence research [9]. They found the genre particularly suitable to serve as a test-bed for game-related research due to its easily implemented and easy-to-understand mechanics, its applicability to various domains, and its inherently challenging and addictive nature.

Regarding a game's difficulty, Clements et al. emphasise the need for adjustability thereof to keep the game interesting and engaging for all manner of players [18]. Charles and Black, who conducted research related to player-centered game design, found that the difficulty of a game influences its perceived entertainment value [19] and is thus worth looking into concerning its motivational and emotional impact. Based on this, van Lankveld et al. researched the affective impact of various difficulty settings [20] based on Rauterberg's incongruity theory [21]. This theory is similar to Csikszentmihalyi's concept of "flow" [22] in that it states that the difference between the context complexity (e.g. the difficulty of the game) and the system complexity (e.g. the skills and knowledge of the player) influences the player's affective state: If the game is too difficult, the player will be frustrated; if the game is too simple, the player will be bored. Only the right amount of challenge ("slightly positive incongruity") will bring the player pleasure and keep them in the flow. Van Lankveld et al. were able to support these hypothesised effects to some extent. Please refer to [20] for more details.

While there are many frameworks that define evaluation criteria for serious games [23], [24], they mostly focus on analysing games outside of the playing experience. Liu and Ding [25] have developed an evaluation component for in-game analysis, however it is integrated into their own game engine and can only be modified via script programming.

The above-mentioned works show the applicability of Tower Defense games in the biological-medical context, and as a research tool. With LiverDefense, we intend to provide an easily customisable tool from this domain which can be employed by researchers without programming knowledge.

III. Design Considerations

Due to the inherent conceptual similarities between liver metabolism (neutralizing harmful or waste materials) and the principles of Tower Defense (neutralize assaulting enemies), we chose the basic functions of the human liver as a content domain for the learning matter of the game.

In accordance with Thompson [15], Bassilious et al. [16], and Clements et al. [18], we employed a real-time instead of a turn-based approach to keep in touch with the biological reality and provide an engaging challenge for the player.

As the related work indicates that difficulty is an important aspect in creating successful games, we chose this as a point of focus for LiverDefense. The potential to influence affective states makes game difficulty an interesting element of research for psychological studies. However, many researchers in this field are not proficient programmers. For this reason, LiverDefense is intended to allow for the manual customisation of difficulty in order to induce affective states such as frustration, boredom, and pleasure as needed by the researcher. Since questionnaires are a common tool in psychological research, LiverDefense should be able to add and display them in the game, and provide capabilities for easy editing.

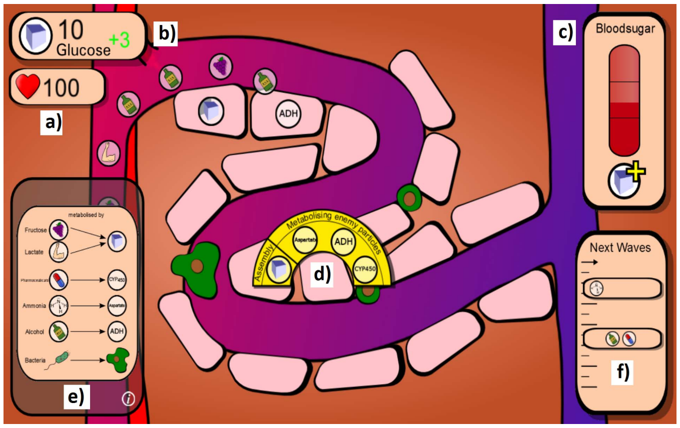


Fig. 1. Screenshot of the LiverDefense playing field and interface, showing a) current health points, b) current glucose balance, c) current blood sugar level, d) the build menu to specialise liver cells, e) the info box for mouse-over help, and f) the next wave preview

Fig. 2. Screenshot of a displayed Likert scale questionnaire (translated from German); the blood sugar management is disabled for the time of the inquiry.

IV. LiverDefense

A. Biology Basics and Learning Matter

LiverDefense aims to convey basic human liver functionalities to players with little or no previous knowledge in this area. The human liver detoxifies biological waste materials like alcohol, ammonia, and pharmaceuticals in the bloodstream by the use of specific enzymes (e.g. alcohol dehydrogenase/ADH, aspartate transaminase and cytochrome P450) [26]–[29]. The liver also has the ability to drain glucose from the bloodstream and store it in the form of glycogen [26]. Furthermore, bacteria are destroyed inside the liver tissue by a layer of macrophagic Kupffer cells lining the bloodstream [30].

B. Necessary Simplifications

The biological context is complex and requires a lot of prerequisite knowledge to be fully grasped. However, to provide a smooth introduction and enhance the playability with regard to a Tower Defense game, we decided to simplify several aspects. Instead of confronting the player with an entire liver system to manage, one singular bloodstream was chosen as basis for the game's playing field to reduce the cognitive demand. We also reduced the number of waste materials metabolised by the liver to alcohol, ammonia, and pharmaceuticals. These materials capitalise on a certain familiarity on part of the player and are representative of the liver's functions. The large and complex urea cycle, which deals with the metabolism of nitrogen, is thereby reduced to a single representative reaction, i.e. neutralising ammonia. In this manner, the player is not overwhelmed and fluid gameplay is ensured. Glycogen synthesis, too, is simplified. Glycogen itself is not introduced in the game mechanics, so that glucose functions as the game's sole currency for income as well as expenses. Instead of confronting the player with the entire glycogen synthesis, we decided to opt for an approach where glucose is simply built from fructose and lactate.

Actual liver cells are capable of several concurring functions. In the spirit of Tower Defense game mechanics, however, we employed generic cells that can be specialized for one specific function to serve as "towers". Cells may thus occupy only one function at any given time.

Another simplification concerns the distribution and behaviour of Kupffer cells. Contrary to the actual situation wherein Kupffer cells line the entire bloodstream within the liver tissue, we decided to provide the player with at most three distinct Kupffer cells. This improves the clarity of the playing field design. Additionally, this approach makes it possible to treat Kupffer cells as towers as well.

With regard to blood sugar management, we chose to present generic negative effects for both exceeding and undershooting the ideal level. Insulin is not mentioned nor is more insight provided into the real-life effects of dangerously high or low blood sugar so that the player will not be diverted from the main task of protecting the liver.

C. Game Elements & Mechanics

The application of the Tower Defense genre to the simplified learning matter presented above results in a set of game elements and mechanics which are described in the following. The developed game interface can be seen in Figure 1. In the case of LiverDefense, the "castle" that the player is given to protect with their "towers" is the liver, and, as a result, the entire body's health. To reflect this in the game mechanic, the player is given 100 health points (see Fig.1, a). Once all health points are lost, the player is presented with a death screen.

As stated above, a singular bloodstream constitutes the path along which the enemy units advance to attack the liver. In accordance with the biological context, these enemy units are alcohol, ammonia, pharmaceuticals, and bacteria. All of these vary in damage potential, and approach in waves of varying composition. Interspersed with these are fructose and lactate particles which can be used to build glucose. Glucose serves as the in-game currency (see Fig.1, b) and is needed to specialize liver cells. As an additional challenge, the user also has to spend glucose to keep the ever-decreasing blood sugar level within acceptable minimal and maximal bounds (see Fig.1, c). To gain glucose and defend against adversarial particles, the player has to specialize the generic liver cells (*hepatocytes*) bordering on the bloodstream. Generic liver cells are specialized into one of the following cell types: glucose hepatocyte (*assemble glucose*), aspartate hepatocyte (*neutralises ammonia*), ADH hepatocyte (*neutralises alcohol*), and CYP450 hepatocyte (*neutralises pharmaceuticals*). A liver cell can be

```xml
<liverDefense>
    <settings>
        <initialGlucose>25</initialGlucose>
        <timeBetweenWaves>15</timeBetweenWaves>
        <timeBetweenParticles>1.5</timeBetweenParticles>

        <particle>
            <damagePotentialAlcohol>1</damagePotentialAlcohol>
            <!-- ... -->
        </particle>

        <hepatocyte>
            <costAdh>5</costAdh>
            <!-- ... -->
            <durationTime>3</durationTime>
            <regenerationTime>10</regenerationTime>
        </hepatocyte>
        <!-- ... -->
    </settings>
    <waves>
        <wave>
            <displayQuestionnaire>true</displayQuestionnaire>
            <elements>
                <element>Lactate</element>
                <element>Alcohol</element>
                <element>Alcohol</element>
                <element>Alcohol</element>
                <element>Fructose</element>
                <!-- -->
            </elements>
        </wave>
        <!-- -->
    </waves>
</liverDefense>
```

Fig. 3. An abbreviated difficulty settings XML file specifiying the various game element parameters (translated from German)

specialized by assigning one of the functions via the build menu (see Fig.1, d).

In order to defend against bacteria, Kupffer cells are necessary. Therefore, the player is provided with three dedicated building spots. Building Kupffer cells is more expensive than specializing cells; however, they are not subject to the same limitations. Specialized liver cells can only ever deal with one of their assigned waste materials at once. This is realised via a "metabolism duration". The cell will only concern itself with the next enemy after this time has passed. To make things more interesting and give the player the potential for more flexibility, specialized cells will also revert back to their generic state after a certain number of enemy encounters. After a short "regeneration time", the player can specialize the cell again according to their strategic considerations. This mechanism is a simplified representation of the death and regeneration of real liver tissue.

An introductory tutorial preceeding the actual game is used to teach the player the basic gameplay and, as such, the basic functions of the human liver. The tutorial design briefly introduces each piece of information (e.g. "*ADH hepatocytes metabolise alcohol.*"), and then instructs the player to perform the corresponding function (in this case: "*Build an ADH hepatocyte to defend against incoming alcohol particles.*"). Players can then take their time to perform the task; the tutorial only progresses after they have finished. All information on game element characteristics (e.g. function, damage potential, building costs, building hotkey etc.) can later be retrieved via mouse-over pop ups on the respective elements during gameplay. An infographic regarding which hepatocytes metabolise which particles can be retrieved via mouse-over on the info box shown in Fig.1, e.

Further assisstance is provided via a wave preview window (see Fig.1, f). Here, players can see which enemy particle

```xml
<questionnaire>
    <item1>Currently, I feel joy.</item1>
    <item2>Currently, I feel bored.</item2>
    <item3>Currently, I feel frustrated.</item3>
    <item4>Currently, I feel angry.</item4>
    <item5>Currently, I feel in control of the game.</item5>
</questionnaire>
```

Fig. 4. A LiverDefense XML file specifiying the employed questionnaire items (translated from German)

types will be included in the next wave and when it will arrive. Neither explicit numbers of enemy particles nor the occurence and numbers of potential fructose/lactate particles are revealed. This helps players with their strategic decisions without telling them what to plan for exactly. In between waves, players have a short time of respite to regroup and reinforce their defenses for the next assault. Since the player cannot collect more glucose in these intervals, the decrease of the blood sugar level is stopped until the next wave arrives.

D. Customisable Features

As stated before, being able to customise the difficulty of an interactive experience is potentially beneficial for psychological studies. LiverDefense employs an XML approach to support the editing of a number of relevant parameters. General parameters (such as initial glucose balance, time between particles, time between waves, damage potential of particles, and drop rate of the blood sugar level), as well as cell-specific parameters (e.g. building costs, metabolism duration, and regeneration time) may be specified as XML parameters of the given elements. The number and composition of waves can be specified as well. With this, it is not only possible to influence the current run of the game; settings can also be saved as complete XML files and be reused or shared. An abridged sample XML file is provided in Figure 3.

In order to support the use of common research methods, LiverDefense enables researchers to add, edit, and display 7-point Likert scale [31] questionnaires. The questionnaire items can be easily specified in an XML file as well, and edited or exchanged as needed. The questionnaires can be displayed throughout the duration of the game, i.e., after every wave. To support the display of questionnaires during waves seemed unnecessary since it would disrupt the flow of the game and potentially annoy participants independent of actual gameplay. Due to layout constraints, each questionnaire frame in LiverDefense is limited to five items. A screenshot of a questionnaire presented in the game can be seen in Figure 2. See Figure 4 for the corresponding XML file. To further supplement questionnaire-based evaluations with game-related data, LiverDefense also provides a CSV-formatted log file which can be imported easily into tools like Microsoft Excel [32] or IBM SPSS [33]. It contains the following values: current timestamp, current action (beginning of new wave or questionnaire display), current questionnaire answers (99 if no questionnaire was displayed, otherwise the chosen Likert scale value), current health points, current glucose balance, death counter, info box counter (e.g. how often players referred to the in-game help), and, once at the beginning, the level of difficulty as indicated by the name of the corresponding XML file. An example can be seen in Figure 5.

Time	Action	Emotion item 1	Emotion item 2	Emotion item 3	Emotion item 4	Emotion item 5	Lives	Glucose	Deaths	Info counter	Difficulty = normal
4/1/2015 2:37:50 PM	Emotiontask 0	5	1	0	0	6	100	25	0	0	
4/1/2015 2:37:53 PM	Spawning wave 0	99	99	99	99	99	100	25	0	0	
4/1/2015 2:38:20 PM	Spawning wave 1	99	99	99	99	99	100	9	0	0	
4/1/2015 2:38:50 PM	Spawning wave 2	99	99	99	99	99	100	9	0	0	
4/1/2015 2:39:23 PM	Spawning wave 3	99	99	99	99	99	98	5	0	0	
4/1/2015 2:39:56 PM	Spawning wave 4	99	99	99	99	99	98	10	0	0	
4/1/2015 2:41:18 PM	Emotiontask 1	5	0	0	0	6	95	4	0	0	
4/1/2015 2:41:30 PM	Spawning wave 5	99	99	99	99	99	95	4	0	0	
4/1/2015 2:42:11 PM	Spawning wave 6	99	99	99	99	99	95	9	0	0	
4/1/2015 2:42:51 PM	Spawning wave 7	99	99	99	99	99	95	18	0	0	
4/1/2015 2:44:18 PM	Emotiontask 2	1	5	2	0	0	39	40	0	0	
4/1/2015 2:44:30 PM	Spawning wave 8	99	99	99	99	99	39	40	0	0	
4/1/2015 2:45:08 PM	Player died	99	99	99	99	99	0	50	1	0	
4/1/2015 2:45:19 PM	Spawning wave 9	99	99	99	99	99	100	75	1	0	
4/1/2015 2:47:10 PM	Emotiontask 3	2	0	0	0	6	94	40	1	1	

Fig. 5. An exemplary LiverDefense log file, pertaining to the study design as detailed in Section V (Evaluation); imported into Excel.

Log files are named according to the player name / participant ID that is entered at the start of the game. They are never overwritten, instead all data pertaining to the same ID is accumulated in one file and saved for subsequent analysis.

E. Comparison with Malone's Goals

A comparison of Liver Defense with Malone's goals for successful educational games [7] yields the following results.

Clear and meaningful goals are provided courtesy of the easily understandable game mechanics of the Tower Defense genre. Due to the implicit character of the employed experience-based teaching method, the player is presented solely with the game-related goal of surviving the enemy attacks. Aside from the relative single-mindedness of the defense task, clear and meaningful goals are further emphasised by the distinct display of health points (see Fig. 1, a). The goal is meaningful in that players are given an asset to protect with which they can identify to some extent, i.e. the human liver. The identification is further supplemented by the death screen which says "Your liver is being regenerated. You will be revived in x seconds.".

The display of health points additionally functions as constant *performance feedback* for the player. Whenever enemies reach the end of the path and damage is taken, the display flashes in alarming red. While this does not mean instant defeat, it indicates that the player needs to revise his or her strategy. The same concept is applied to the blood sugar level (see Fig. 1, c), which starts flashing red whenever it falls below or exceeds the optimal bracket.

LiverDefense supports *varying degrees of difficulty* in that several parameters such as initial glucose balance, time between enemy particles, enemy wave composition, and many more can be adjusted.

Fantasy in the game is supported by presenting the player with a playing field that mimicks the inside of the human body on a scale not normally accessible. The fantastic character of the experience is further aided by the employment of cartoon-like graphics. However, in accordance with Clements et al. [18], a substantial degree of realism is retained to ensure basic biological accuracy.

LiverDefense's *complexity* is comprised of various elements. Time is one factor, as well as the necessary development of strategies to achieve the game's goal. *Curiosity and exploration* are promoted via the various tower and enemy types, i.e. liver cells and adverse particles respectively. Players are pointed to an illustration detailing all tower-enemy mappings (see Fig.1, e) in the tutorial, however the tutorial exercises only cover part of that information actively. Thus, players are prompted to explore the remaining combinations during gameplay.

As shown, LiverDefense adheres to the basic principles of Malone's theory. While not being conclusive, this is at least indicative of LiverDefense incorporating the necessary features to provide an engaging experience for learners.

LiverDefense was used successfully as a research tool in a psychological study. The setup and results of the study are as follows.

V. Evaluation

With regard to the tremendous popularity of developing and using games for educational purposes, most explanations regarding why games determine students' emotions in terms of their positive as well as negative valence, refer to their interactive nature originating from the control provided by the game in connection with students' perceived control over gameplay [5]–[7], [34]–[36]. Being in control over gameplay is regarded as enjoyable [37], while a lack of control appears to support frustration [36], [38].

However, it has not yet been fully investigated in what manner external game events (e.g. variation in the degree of difficulty) influence the perception of control, and how the different control perceptions interact with emotions during gameplay. Investigating these emotions in more detail is important as they are deemed to influence engagement and learning in educational settings [39]–[41]. While positive emotions should facilitate performance, negative emotions should be detrimental thereto.
Thus, the aim of this study was to explore whether the extent of perceived control over gameplay (manipulated by varying the degree of difficulty) affects positive and negative emotions while learning with LiverDefense. While a broader range of emotions was investigated in this study, we will focus on the analysis of enjoyment and frustration in the testing conditions, as these serve as examples for positive and negative emotions. In the following, we report on the method and outcome of our study.

TABLE I. MEANS (STANDARD DEVIATIONS) OF PERCEIVED CONTROL AND EMOTIONS, REPRESENTED PER CONDITION

	High control condition (n = 39) (easy difficulty)	*Moderate control condition* (n = 41) (normal difficulty)	*Low control condition (n = 41)* (hard difficulty)
Perceived Control	5.22 (1.41)	4.36 (1.17)	4.02 (1.25)
Enjoyment	4.26 (1.53)	4.29 (1.34)	4.10 (1.21)
Frustration	1.84 (1.21)	2.81 (1.35)	3.35 (1.44)

A. Method

1) Participants: The sample of this study consists of bachelor and master students in the fields of psychology, computer science, economy, biology, and medicine from Ulm University (Germany). Students were recruited via an email that was sent to all students of the university, and an announcement on the university's Facebook page. In sum, one hundred and twenty-one students participated in this experiment. Their age range varied between 18 and 38, with the mean age equal to 22.89 years (*SD* = 3.01). 54.8% of the participants were female and 45.2% were male. Regarding frequency of play, 26.8% reported that they play games every day, and a further 30.1% said they played games up to once a week.

2) Material: LiverDefense was employed as a research tool to elaborate the link between varying levels of difficulty and learners' perception of control, as well as their emotions during play. The highly specific learning matter, e.g. the basic functions of the human liver, made sure that we could acquire a large pool of participants with equally minimal knowledge in the field. Via extensive pre-testing, three difficulty XML files (easy, normal, and hard) were prepared to induce *high*, *moderate*, and *low perceived control* in the intended target group. Easy difficulty made sure the player could survive waves most of the time, even without strategic planning. Normal difficulty required some strategy, yet players would generally still be able to survive the required amount of time. Hard difficulty, however, bombarded the player with a quick succession of enemy particles and shortened the supply of available glucose significantly via the demands of a rapidly sinking blood sugar level. This condition ensured that players would frequently be overrun by enemy particles. The corresponding parameter details for the three difficulty settings were as follows:

Easy:
- initial glucose balance: 35
- time between particles: 2.0
- blood sugar decrease per second: 0.2

Normal:
- initial glucose balance: 25
- time between particles: 1.5
- blood sugar decrease per second: 0.35

Hard:
- initial glucose balance: 20
- time between particles: 1.0
- blood sugar decrease per second: 0.5

Additionally, from easy to hard there were more enemy particles and comparatively less resources to assemble glucose. Time between waves and particle damage were however the same for all three settings. To guarantee equal playing time for all participants, we employed a respawn mechanism not usually found in Tower Defense games. For this purpose, upon losing all health points during a wave, the player was presented with a death screen, after which they could continue to fight the next wave with a set of new resources.
Questionnaires, detailed in the next paragraph, were presented to the participants after the end of each three minute round. Overall, the participants played three rounds each.

3) Measures:
Perceived control: To examine whether the manipulation of control via LiverDefense's three difficulty settings was effective, students' perceived control was investigated once before the game, and at the end of each of the three rounds. Students were requested to indicate how they had experienced being in control of gameplay on a 7-point Likert scale ranging from 1 (= very little) to 7 (= very much).

Emotions: For measuring enjoyment and frustration, students responded on how much of each of the emotions they experienced before and at three points during gameplay; the responses were given on a 7-point Likert scale ranging from 1 (= very little) to 7 (= very much).

4) Procedure: The experiment was conducted in a university computer lab. All participants were randomly assigned to one of the three game versions (easy, normal, and hard difficulty), with each version representing one of the manipulation conditions of control in the experiment: *high control condition* (n = 39), *moderate control condition* (n = 41), and *low control condition* (n = 41). There was no significant difference between the three conditions regarding participant gender (male = 18, female = 21 for the *high control condition*; male = 20, female = 21 for the *moderate control condition*; and male = 18, female = 23 for the *low control condition*).
In the beginning, all participants were introduced to the game elements and mechanics by playing the introductory tutorial previously described. After playing the tutorial, participants completed a questionnaire with respect to their perceived sense of control and emotions; these baseline ratings were used as covariates. Afterwards, participants were invited to play three rounds of LiverDefense, constituting approximately ten minutes and ten to eleven waves depending on the difficulty setting. After each round, perceived control and emotions were measured as described above. During gameplay, participants worked on an individual basis on a computer. They could interact via keyboard and mouse.

B. Results

1) Differences in Perceived Control: First, we investigated whether the manipulation of the game regarding the players' sense of control took effect as intended. For this, we averaged the ratings of control perception over the three rounds of play. A one-way ANCOVA was carried out to discover the differences between conditions with the experienced sense of being in control as the dependent variable, the game conditions as the independent variable, and the sense of control measured before starting the game as the covariate. The means for all three game conditions are summarized in Table I.

The covariate – the perception of control after the tutorial and before starting the three rounds of gameplay – was significantly related to the perception of control during gameplay: $F(1,115) = 50.70$, $p < .001$, $\eta^2 = .30$. There was also a significant effect of the game conditions on the perception of control after controlling for the covariate: $F(2,115) = 12.26$, $p < .001$, $\eta^2 = .17$.

Bonferroni corrected *post hoc* comparisons revealed that participants in the *high control condition* displayed a higher sense of control than participants in the *moderate* ($p = .002$) and *low control condition* ($p < .001$). Furthermore, there was a significant difference between participants in the *moderate* and *low control condition* ($p = .047$), indicating a higher degree of control perception for the moderate condition. These results demonstrate that the manipulation of the perception of control via the three difficulty settings *easy*, *normal* and *hard* was successful.

2) Differences in Emotions: Next, we explored how participants in the different game conditions differed in their experienced extent of enjoyment and frustration. For both, a one-way ANCOVA using each emotion as the dependent variable, game conditions as the independent variable, and the measured emotions after playing the tutorial and before gameplay as the covariate was performed. Due to missing data the degrees of freedom differed between single procedures. Results showed that the covariate was significantly related to enjoyment ($F(1,115) = 114.13$, $p < .001$, $\eta^2 = .49$) and to frustration ($F(1,115) = 59.97$, $p < .001$, $\eta^2 = .34$). Also, the game condition had a significant effect on participants' enjoyment ($F(2,115) = 3.56$, $p = .03$, $\eta^2 = .05$), and frustration ($F(2,115) = 21.55$, $p < .001$, $\eta^2 = .27$).

For enjoyment, Bonferroni corrected *post hoc* comparisons demonstrate that participants in the *high control condition* reported higher enjoyment, and differed significantly from participants in the *low control condition* ($p = .03$). There was no significant difference between *high* and *moderate*, or *moderate* and *low control condition* ($p = .24$). As expected, frustration was rated lowest in the *high control condition* and was significantly different compared to *low* ($p < .001$) and *moderate control condition* ($p < .001$). Furthermore, significant differences also appeared for the comparison between *moderate* and *low control condition* ($p = .03$).

VI. Discussion

Analysis of the evaluation data shows that the intended variances in game difficulty, as regulated via the XML settings files *easy*, *normal* and *hard*, indeed led to varying levels of *high*, *moderate* and *low perceived control* in the participants.

Analysis regarding enjoyment and frustration indicates that LiverDefense is effective in inducing these emotion via its various difficulty settings in accordance with the theories of incongruity [21] and flow [22]. The lack of significance in cases where both enjoyment and the *moderate control condition* are concerned can be explained by the strategic demands of the game. The peculiarity of a participant's strategic abilities could, in the moderate condition, swing gameplay either way: Independent of the difficulty setting, particularly good or bad decisions could then make the gameplay situation easier or harder; and thus, more similar to one of the other conditions.

Overall, using LiverDefense as a research tool to induce various degrees of control via game difficulty and to collect player data was successful. However, the following limitations of using LiverDefense as an evaluation tool were discovered during this study. To create settings with varying levels of difficulty, extensive testing was required to explore suitable particle parameter values and wave configurations. In order for the settings to have any real-world relevance, they have to be tested by members of the intended target population outside the development team. This does however in turn reduce the pool of potential study participants. There are study contexts for which this can be a problem.

Furthermore, interface layout restrictions imposed restrictions on what kind of questionnaires could be used. Currently, LiverDefense only supports the display of five questionnaire items with answers on a 7-point Likert scale each per frame.

VII. Conclusion & Future Work

LiverDefense is an educational Tower Defense game aimed at teaching players the basic functions of the human liver while adhering to Malone's goals for successful serious games. LiverDefense can be adjusted in its difficulty and display 7-point Likert scale questionnaires throughout the game. Customisations regarding both the difficulty and the employed questionnaires can be done via easily-understandable, human-readable XML files. LiverDefense further supports research studies by saving relevant game events in log files distinguished by player ID. As such, LiverDefense is a useful evaluation tool and can be adapted by non-programmers to fit their needs.

LiverDefense has been successfully used in a psychological study concerning itself with the effects of perceived control on player emotions. Study results have shown that LiverDefense's difficulty can be effectively adjusted. Analysis indicates that this is potentially efficacious in inducing positive as well as negative emotions like enjoyment and frustration.

The next steps will focus on the remaining analysis of the conducted study, namely the effect of perceived control on the emotions boredom and anger, as well as learning outcome. This will provide further insights into the usefulness of games with adjustable difficulty as evaluation tools in serious games research.

To make LiverDefense applicable for more research scenarios, we will look into the possibility of supporting the queXML format for questionnaires [42]. This will free researchers from the limitation of 7-point Likert scale questionnaires and facilitate the integration of LiverDefense with other tools such as the web-based survey application LimeSurvey [43].

To address the underlying problem of creating suitable settings files however, more research needs to focus on which parameters of the employed game elements and which compositions of enemy waves have what kind of effect on the player's level of perceived control over gameplay. Instead of trial-and-error testing of various settings, this would provide researchers with design guidelines regarding game difficulty.

ACKNOWLEDGMENTS

This work was conducted within the project "Serious Games - Skill Advancement Through Adaptive Systems" being funded by the Carl Zeiss Foundation and as part of the project "EffIS - Efficient and Interactive Studying" (FKZ: 16OH21032) funded since 2014 by the BMBF (German Federal Ministry of Education and Research) within the program "Aufstieg durch Bildung: Offene Hochschulen".

REFERENCES

[1] C. C. Abt, *Serious games*. University Press of America, 1987.

[2] D. R. Michael and S. L. Chen, *Serious games: Games that educate, train, and inform*. Muska & Lipman/Premier-Trade, 2005.

[3] M.-C. Li and C.-C. Tsai, "Game-based learning in science education: a review of relevant research," *Journal of Science Education and Technology*, vol. 22, no. 6, pp. 877–898, 2013.

[4] P. Wouters, E. D. Van der Spek, and H. Van Oostendorp, "Current practices in serious game research: A review from a learning outcomes perspective," *Games-based learning advancements for multisensory human computer interfaces: techniques and effective practices*, pp. 232–255, 2009.

[5] M. Prensky, "Computer games and learning: Digital game-based learning," *Handbook of computer game studies*, vol. 18, pp. 97–122, 2005.

[6] K. Mangold, "Educating a new generation: Teaching baby boomer faculty about millennial students," *Nurse Educator*, vol. 32, no. 1, pp. 21–23, 2007.

[7] T. W. Malone, "Toward a theory of intrinsically motivating instruction," *Cognitive science*, vol. 5, no. 4, pp. 333–369, 1981.

[8] M. D. Dickey, "Engaging by design: How engagement strategies in popular computer and video games can inform instructional design," *Educational Technology Research and Development*, vol. 53, no. 2, pp. 67–83, 2005.

[9] P. Avery, J. Togelius, E. Alistar, and R. P. Van Leeuwen, "Computational intelligence and tower defence games," in *Evolutionary Computation (CEC), 2011 IEEE Congress on*. IEEE, 2011, pp. 1084–1091.

[10] B. Sawyer, "From cells to cell processors: The integration of health and video games," *Computer Graphics and Applications, IEEE*, vol. 28, no. 6, pp. 83–85, Nov 2008.

[11] V. Wattanasoontorn, I. Boada, R. García, and M. Sbert, "Serious games for health," *Entertainment Computing*, vol. 4, no. 4, pp. 231–247, 2013.

[12] S. McCallum, "Gamification and serious games for personalized health," *Studies in Health Technology and Informatics*, vol. 177, pp. 85–96, 2012.

[13] M. Papastergiou, "Exploring the potential of computer and video games for health and physical education: A literature review," *Computers & Education*, vol. 53, no. 3, pp. 603–622, 2009.

[14] T. Baranowski, R. Buday, D. I. Thompson, and J. Baranowski, "Playing for real: video games and stories for health-related behavior change," *American journal of preventive medicine*, vol. 34, no. 1, pp. 74–82, 2008.

[15] J. Thompson, "Food fight," Tech. Rep., 2010.

[16] E. Bassilious, A. DeChamplain, I. McCabe, M. Stephan, B. Kapralos, F. Mahmud, and A. Dubrowski, "Power defense: A video game for improving diabetes numeracy," in *Games Innovation Conference (IGIC), 2011 IEEE International*. IEEE, 2011, pp. 124–125.

[17] E. Bassilious, "Numeracy in adolescents with type 1 diabetes: Assessment and gaming intervention–a pilot project," Ph.D. dissertation, University of Toronto, 2013.

[18] P. Clements, J. Pesner, and J. Shepherd, "The teaching of immunology using educational: gaming paradigms," in *Proceedings of the 47th Annual Southeast Regional Conference*. ACM, 2009, p. 21.

[19] D. Charles and M. Black, "Dynamic player modeling: A framework for player-centered digital games," in *Proc. of the International Conference on Computer Games: Artificial Intelligence, Design and Education*, 2004, pp. 29–35.

[20] G. Van Lankveld, P. Spronck, H. J. Van Den Herik, and M. Rauterberg, "Incongruity-based adaptive game balancing," in *Advances in computer games*. Springer, 2010, pp. 208–220.

[21] M. Rauterberg, "About a framework for information and information processing of learning systems." in *ISCO*, 1995, pp. 54–69.

[22] M. Csikszentmihalyi, *Flow: The psychology of optimal experience*. HarperPerennial New York, 1991, vol. 41.

[23] K. Mitgutsch and N. Alvarado, "Purposeful by design?: a serious game design assessment framework," in *Proceedings of the International Conference on the foundations of digital games*. ACM, 2012, pp. 121–128.

[24] L. Nacke, A. Drachen, and S. Göbel, "Methods for evaluating gameplay experience in a serious gaming context."

[25] S. Liu and W. Ding, "An approach to evaluation component design in building serious game," in *Learning by Playing. Game-based Education System Design and Development*. Springer, 2009, pp. 141–148.

[26] H. Liehr, *Leber, Galle, Bauchspeicheldrüse: Wirksame Hilfe bei Beschwerden (German)*. Trias, 2002.

[27] J. F. Riemann, W. Fischbach, P. Galle, and J. Mössner, *Gastroenterologie in Klinik und Praxis: Das komplette Referenzwerk für Klinik und Praxis (German)*. Thieme, 2007, (German).

[28] C. Bode and J. Bode, *Ernährungsmedizin: nach dem Curriculum Ernährungsmedizin der Bundesärztekammer (German)*. Georg Thieme Verlag, 2004, ch. Protektive Wirkungen und Missbrauch von Alkohol, pp. 516–538.

[29] S. Steiner-Welz, *Die wichtigsten Körperfunktionen der Menschen (German)*, R. Welz, Ed. Reinhard Welz Vermittlerverlag Mannheim, 2005.

[30] K. Friese, I. Mylonas, and A. Schulze, *Infektionserkrankungen der Schwangeren und des Neugeborenen (German)*. Springer, 2013.

[31] R. Likert, "A technique for the measurement of attitudes." *Archives of psychology*, 1932.

[32] Microsoft Corporation, "Microsoft Excel - Create order," https://products.office.com/excel.

[33] IBM, "SPSS - Predictive analytics software and solutions," www.ibm.com/spss.

[34] M. DiPietro, R. E. Ferdig, J. Boyer, and E. W. Black, "Towards a framework for understanding electronic educational gaming," *Journal of Educational Multimedia and Hypermedia*, vol. 16, no. 3, pp. 225–248, 2007.

[35] K. Lucas and J. L. Sherry, "Sex differences in video game play: A communication-based explanation," *Communication Research*, vol. 31, no. 5, pp. 499–523, 2004.

[36] P. Vorderer, C. Klimmt, and U. Ritterfeld, "Enjoyment: At the heart of media entertainment," *Communication theory*, vol. 14, no. 4, pp. 388–408, 2004.

[37] R. Pekrun, "The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice," *Educational psychology review*, vol. 18, no. 4, pp. 315–341, 2006.

[38] T. Grodal, "Video games and the pleasures of control," *Media entertainment: The psychology of its appeal*, pp. 197–213, 2000.

[39] D. K. Meyer and J. C. Turner, "Re-conceptualizing emotion and motivation to learn in classroom contexts," *Educational Psychology Review*, vol. 18, no. 4, pp. 377–390, 2006. [Online]. Available: http://dx.doi.org/10.1007/s10648-006-9032-1

[40] P. A. Schutz and R. Pekrun, "Introduction to emotion in education," *Emotion in education*, pp. 3–10, 2007.

[41] E. Um, J. L. Plass, E. O. Hayward, B. D. Homer *et al.*, "Emotional design in multimedia learning." *Journal of Educational Psychology*, vol. 104, no. 2, p. 485, 2012.

[42] queXML - An Open Source, XML, multi-mode questionnaire description toolkit. [Online]. Available: http://quexml.sourceforge.net/

[43] Lime Survey - the open source survey application. https://www.limesurvey.org. [Online]. Available: https://www.limesurvey.org

Back to Contents

Generation of Variable Human Faces from 3D Scan Dataset

Igor Chalás, Zuzana Ferková, Katarína Furmanová, Jiří Sochor and Barbora Kozlíková
Department of Computer Graphics and Design
Faculty of Informatics, Masaryk University
Brno, Czech Republic
{chalas, xferkova, xfurman, sochor, kozlikova}@fi.muni.cz

***Abstract*—Generating human faces is an important task in many research and application fields, including the gaming industry. When the scene contains many characters, it becomes impracticable to create all individual characters manually. On the other hand, the requirement for the different appearance of faces of individuals in a crowd is now more in demand. In this paper we propose our solution to the automatic generation of human faces. Our solution synthesizes facial parts coming from scans of real human faces; a process that is completely automatic. However, the user has the possibility to further adjust the composite by designing replacements, leading to a desired appearance. The final composite can be exported and attached to a given avatar. As the perception of the variability of resulting composites is in our focus as well, we also present results of the user study which was designed for this purpose. The study aims to reveal the minimal changes of the facial parts which are necessary to make in order to perceive a given facial model as different from the original one.**

***Index Terms*—facial composite, 3D face model, face perception, crowd simulation, character modelling**

I. Introduction

The human face is one of the most important features in making people distinguishable. When meeting an unknown person, it is the face that attracts our attention the most. This real life situation is projected to the virtual world as well. When generating a virtual human, one of the crucial parts is the generated face. The quality of the resulting face then determines the final feasibility of the real appearance of the virtual human.

Generated human faces play an important role in many fields. Criminologists use 2D or 3D images of human faces to synthesize the criminals appearance according to the description of a witness. Anthropologists perform a variety of tasks on facial models, stretching from the facial reconstruction of ancient persons to the simulation of aging, or studying facial changes during pregnancy or disease. Among these vast fields, where the human face is in the center of interest, belongs also the gaming industry. One of the essential research topics in gaming is the automatic creation of human crowds where researchers have to focus on several issues, such as the randomness of movement and the treatment of collisions. It is often hard to produce a great variety of characters. Due to the complexity of manual character modeling, it is very time consuming for a 3D modeler to manually create a large amount of heterogeneous agents. There are several research studies dealing with the automated generation of human crowds [1], [2], [3], [4], [5], [6]. Most of them, such as [7], concentrate on the variation of clothes and their color. Such approach enables to create a decent variability in large crowds with respect to their perception. However, in many cases this technique does not lead to sufficient results, especially when some individuals in the crowd are walking towards the camera. Then the uniform face of all humans is unacceptable.

In this paper we present our solution to the problem of uniform faces in the crowd by introducing a technique for automatic or semi-automatic generation of human faces. These faces are synthesized from different parts of real human faces which were scanned and further processed in order to get individual facial parts, such as eyes, lips, or nose. Our technique provides the users with the ability to create a desired variety of 3D models of human faces. The resulting models can be further exported to file formats commonly used by the most current and frequently used game engines, such as Unity or Unreal Engine. Our process of creating faces can be completely automatic. However, the generated faces can be further adjusted, e.g., by manual selection and replacement of individual facial features.

Our technique uses the latest technology of the 3D scanning of human faces, which enables the creation of a vast database of real facial models. These models can be further processed and used with respect to legal issues. In our case, the models are split into several parts with respect to facial features defined by anthropologists. These parts are further used for the synthesis of new artificial faces.

The variability of generated faces and its perception is a crucial part of the whole process of simulating a crowd. Thus, our paper focuses also on the user study aiming to reveal the differences in perception of the generated facial models. More specifically, the goal of the study is to reveal the impact of the number of changes of individual facial parts on the perception of the variability.

II. Background

In this section we discuss the existing solutions for problems related to the topic of this paper. First we present techniques for crowd simulation and then we present current methods for generation of human faces and bodies.

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

During game development, several scenarios exist where populating the scene with computer controlled agents is beneficial. This is related to the scene atmosphere as well as to the general gameplay. These scenarios contain scenes where the gameplay takes place in highly populated cities, as can be seen in *Assassin's Creed: Unity*. Another example are multi massive online games, such as *World of Warcraft*, or campsites with a significant number of refugees, as in *Dragon Age: Inquisition*. However, simulating large crowds in games is usually problematic, as it is sensitive on both performance and the amount of created assets. Due to these difficulties, several strategies for efficient crowd rendering were proposed. Many solutions are based on impostors, i.e., 2D precomputed images, such as [8], [9], [10].

With the ability to generate large crowds in games, the developers are facing the challenge of generating the uniqueness of individuals in the crowd in a limited time. In consequence the players should not feel as they are facing an army of clones. Creating each member of the crowd manually would be very inefficient and time consuming. So it is inevitable that some of the models will be cloned. However, there is an extensive research conducted in the area of human perception of crowds and aspects affecting the player's ability to recognize clones within the stream of agents. Tecchia et al. [1] showed how changes in color of the texture of individual agents affect the ability to recognize the resulting models as clones. Ulicny and Thalmann [2] distinguished between entities in a crowd by altering their movement. Other studies [4], [5], [3] researched ways to distinguish between agents by changing their behavior, adding different variations of cloths and adding decals to agents' outfit. Furthermore, McDonnell et al. [7] and later Thalmann et al. [6] discussed the usage of different parameters, namely the color variety of texture, height, and shape variety of agents and accessories and their effect on human perception. Research in the area of face perception [11] creates conditions for the development of methods based on the facial diversity.

Currently most of the faces in computer games are still created manually by artists. This process is reasonable when creating the main characters in a game. The same process is used when the appearance of the character has to resemble certain characteristics imposed by, for example, game designers, such as specific scars, body deformations, or other unique signs. The manual post-processing is also required when the character should resemble a real person. This is the case of many sport simulations, such as *FIFA 15*, where the scanning of footballers was used to acquire the most precise resemblance. However, when it comes to creating many agents whose appearance is not specified so precisely, the manual modeling becomes highly ineffective.

To avoid manual modeling of large numbers of heads of agents, several solutions already exist. They automatically allow the recreation of human faces, while at the same time, offer a high variety of results. Pighin et al. [12] proposed a method of recreating the model of a human head from multi-view photographs. With the increasing availability of devices which allows us to quickly acquire several angles of a human face, this method is preferred for the purposes of our research as well. One approach suggested by Blanz and Vetter [13] enables to use single or multi-view photographs to reconstruct the 3D head model using the morphable face model. This is reached by transforming the shape and texture of the example set into a vector space representation. Another technique proposed by Kahler et al. [14] is based on fitting of a template head model to input range scan data. It results in creating a head model with anatomical structure which the users are able to animate. Maejima and Morishima [15] tried to overcome the necessity of manual adjustment, such as the model and camera posing or rendering parameters and marker settings, which were required by the previous approaches. They propose a model that is able to fully automatically create a head model. The core of the method is based on the automatic detection of feature points.

Most of the mentioned approaches rely on capturing the real-world data. However, this process possesses several obstacles and such data can rarely be used directly. Such data often contain noise or incomplete information, such as holes in the mesh. This latter is known as the mesh completion problem. Over the past years several solutions to this problem were proposed. Kahler et al. [14] fitted the scanned data to a generic model to fill in the holes created during the data acquisition. Allen et al. [16] also used a template to fill in the holes in the mesh. Their proposed solution allows also the manual definition of areas which are commonly considered to be problematic during the data acquisition, such as ears and fingers. In these areas the template model is favored over the scanned data. Kraevoy and Sheffer [17] proposed another method for the template-based mesh completion supporting both local and global approach.

Other group of existing applications focuses on enabling the user to create their own human head in a reasonable time. Some of them are parametric-based, such as *FaceGen* [18]. It focuses on human faces and allows for adjusting several parameters, such as age, race, and gender. Another well-known software used for the parametric creation of heads as well as of the entire body, is *MakeHuman* [19]. It offers the user a wide variety of adjustments.

III. Methodology

Our approach aims to quickly customize the face of a 3D character using preprocessed facial features. This will provide the user with a considerable variety of the output 3D models. Using our system, users can quickly add their own datasets of facial features and create heads designed specifically for their application. This includes also cases where, e.g., some stylized graphic styles are required.

We implemented the functionality for generating human faces from real data atop of the FIDENTIS Analyst [20]. More specifically, we improved its Composite module, which was originally designed to manually create human 3D facial composites for forensic identification purposes.

The process of creating 3D facial composite consists of several steps (see Figure 1). The preparation phase can be fully automatic or semi-automatic. The individual phases, described in detail in the following subsections, cover the decomposition of 3D scanned models of real human faces, followed by the selection of facial parts, their fitting on a prototype head, and final sewing into one connected mesh.

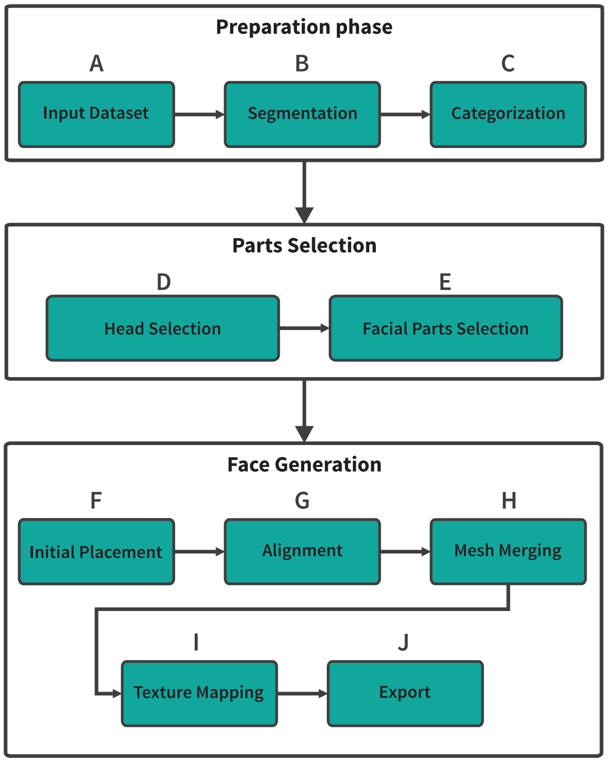


Fig. 1. The process of generating 3D facial composite.

A. Input Dataset

Real 3D facial images, retrieved from the FIDENTIS 3D Face Database (F3D-FD) initially published in [21], were used as the input data for our solution. From the list of the existing somatoscopic features we selected those which influence the appearance of the human face the most. Among these features belong eyes, eyebrows, nose, lips, and chin. Subsequently they are used for the definition of individual parts of the input mesh which are segmented in the following step. In order to cover the variability of human faces available in the F3D-FD database we defined 30 somatoscopic features containing 128 categories. For example, the ’Nose profile’ feature contains four basic categories – concave, straight, convex, and wavy (see Figure 2). Then the representative 3D models containing these predefined features were selected from the F3D-FD database. These models serve as an input to the segmentation phase.

B. Segmentation of Facial Features

In this phase, the 3D facial models acquired from the F3D-FD database are segmented to obtain only the desired facial features. The process of segmentation consists of several steps. First, the original mesh is aligned with the generic 3D facial model in the Frankfort horizontal plane, which represents the anatomical position of a human skull. The Iterative Closest Point (ICP) algorithm is used for the registration [22]. Second, we apply the anisotropic filtering [23] which further reduces the number of artifacts present in the original data, such as the orange peel effect. The level of filtering is highly dependent on the type of the segmented feature. For example, when segmenting a nose, we can use a high filtering value, but when segmenting an eye, no filtering is allowed.

Fig. 2. Example of one feature extracted from the F3D-FD database. From left to right: Nose profile feature and its four categories – concave, straight, convex, and wavy.

On such preprocessed model, the computation of a curvature map is launched. Again, according to the type of the segmented feature, an appropriate type of curvature is used. For the curvature computation we used the definition by [24]. Basically we use four types of curvatures—minimal, maximal, mean and Gaussian. Subsequently we set a threshold (with respect to the segmented feature) which selects parts of the model which can be the best candidates for the desired segment. Here we can obtain more independent mesh segments and we have to select the one representing the desired facial feature. The selection is based on unique characteristics of each feature. For example, the nose is the largest continuous segment after applying selected operators. After selecting the final segment we finalize the contour of the segment boundary. For this purpose, the closing operator from the mathematical morphology is applied to the mesh according to [25]. Figure 3 illustrates the process of the segmentation. In addition, each segmented part is further processed in order to determine landmarks on it. Here we use the approach described in [26].

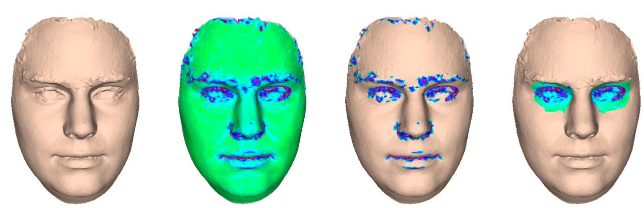


Fig. 3. From left to right: The process of eye segmentation – initial model, maximal curvature, maximum entropy threshold and closing operator.

C. Categorization of Facial Parts

In this step we tag each facial feature with respect to several characteristics, such as gender, age, and nationality of the scanned volunteer. This allows us to further categorize the facial features and select only certain types of scans. This gives the game developer more control over the end results of face generation, especially in cases where the created avatars have to satisfy a certain gender distribution or have a specific ethnic background. The required information about the characteristics of the input data is included directly in the metadata stored in the F3D-FD database.

Once the representative of characteristic facial features is selected and preprocessed, we load them into the FIDENTIS Analyst software. In its Composite module, these features are divided into two categories. The first category consists of paired facial features, such as eyes and eyebrows. The second category then contains single facial features, such as nose, mouth, chin, and head.

D. Head Selection

All steps of the whole process of face generation which were discussed until now belong to the preparation phase. Now we will focus on the description of steps which are necessary to perform each time when the user creates a new composite model. The first step is to select a prototypic head onto which the facial parts will be positioned. This selection can be automatic or semi-automatic. The automatic approach selects one of the predefined prototypic heads, whereas the semi-automatic approach enables to load the head mesh from a given file format. Such loaded mesh has to be further manually equipped with landmarks. Then it is prepared for the next phases of the whole process.

E. Facial Parts Selection

Similarly to the head selection, of the facial parts selection can be also fully automatic or semi-automatic. The automatic version selects a representative of each facial part randomly. The semi-automatic process enables to define several criteria which reduce the amount of available facial parts. Among these criteria belong, e.g., the age limit or sex. From these restricted sets the representatives are again selected randomly.

F. Initial Placement

In this stage, the process of positioning the selected facial parts onto the selected prototypic head is launched. For each facial feature we define landmarks on this feature and the corresponding landmarks on the model of the head. We use three landmarks for each single facial feature and six for each paired facial feature. The sum of distances of corresponding landmarks is minimized, which allows us to quickly acquire the initial translation, rotation and scaling of the facial parts.

G. Alignment

For the placement of the facial parts we created a set of prototypic heads which reflect on several properties of human faces, such as their height and width. However, the initial placement onto the prototypic head gives us only a rough approximation of the position of the facial features.

To gain more precise position of the facial features on the prototypic head, we employ the alignment algorithms. Our models are stored in a point-based format, allowing us to perform the alignment by using the ICP algorithm [22]. More particularly, we use its SICP derivation [27] which also offers the computation of uniform scaling parameter for a given point cloud. Both used algorithms attempt to minimize the differences between two point clouds in each iteration until either a given number of iterations is reached, or the transformations converge to a local minimum. Instead of trying to align all points of the facial features to the entire prototypic head, we only use two corresponding point sets to compute the transformation parameters. The first set is defined on a facial part and its corresponding second set is detected on the prototypic head. Using the algorithm presented in [22], we are able to extract the translation vector and the quaternion corresponding to the rotation in a given iteration. Then, using the algorithm proposed in [27], we compute the uniform scaling parameter and apply the transformations directly to the facial feature mesh, as well as to the set of points we defined for the facial feature. The resulting mesh can be seen in Figure 4.

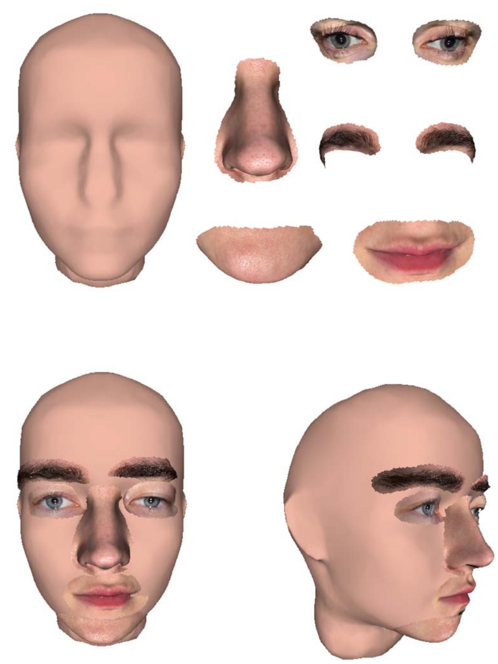


Fig. 4. Generated face before the Mesh Merging phase, showing added facial parts as separated objects. The facial parts are positioned onto the prototypic head.

H. Mesh Merging

Once the alignment is performed for each facial part, we get correctly translated, rotated, and scaled facial parts with respect to a given prototypic head. However, each facial part is still an individual object which is not connected to the rest of the model. To create a single connected mesh consisting of the individual parts we perform the mesh merging algorithm. This algorithm consists of the following three steps.

1) Contour Projection: In the first step of the mesh merging procedure we start with finding and projecting of contours of the facial parts onto the mesh of the prototypic head. We represent the projected contour as an active contour structure, also know as a snake [28]. More specifically, we use the Doubly Connected Edge List (DCEL) of snaxels, i.e., the adjacent points in the snake, so that each pair of consequent snaxels matches with an existing outer edge of the facial part. We project the contour onto the prototypic head mesh using bi-directional ray-tracing. This approach was selected because it is possible that the segments of individual facial parts can be occluded by the head mesh. In such cases we compute the distance between the projected vertex and the intersections of the rays with the head mesh as our target projection. As the algorithm operates in two directions, we compute the distance in each direction and take the smaller from the two resulting values. Now we describe in detail two crucial steps – the process of updating the DCEL and the projection of the snake contour onto the head mesh.

a) Insertion of New Vertices into DCEL: When merging the facial part with the prototypic head mesh, we need to add new vertices to the head mesh. Their position is based on the computed projection through which we interconnect the head mesh with the added facial part. There are three possible cases where a vertex may be projected and we handle each of them separately.

In the first case, a vertex from the facial part is projected to a vertex on the head mesh. This is the simplest case where we only swap the vertex from the head mesh with the vertex from the facial part. Such situation does not require any changes in topology.

In the second case, the projected vertex from the facial part lies on a half edge on the head mesh. In this case we have to split the half edge as well as its corresponding twin edge and replace each of them by a pair of new edges. We also have to replace each face adjacent to the original half edge and its twin edge with the pair of new triangular faces. This preserves the consistency of our topology. Adding new faces also requires to recompute the normal vectors of these faces and the texture coordinates.

The third case occurs when the projected vertex lies inside a face. Then we replace this face by a triangle fan where the projected vertex forms its central point. The normal vectors and texture coordinates are recomputed using the barycentric coordinates of the replaced face.

b) Snake Projection onto the Head Mesh: At this point the snake representing the contour of a facial part is still not merged with the head mesh. So it may not respect the topology of the head mesh and the requirements for mesh triangulation. To fix this we begin with selecting the first snaxel and inserting it to its projected position on the head mesh using the above described contour projection. The other snaxels are iteratively processed in the following way. We create a vector from the current snaxel to the subsequent one. We find a suitable candidate face for projection by searching through the neighboring faces of the current snaxel. Then we project the created vector onto the selected candidate. If the projected vector intersects with either a half edge or a triangle vertex of the face, we need to insert new vertex to the position of intersection. Then we create the corresponding snaxel. In a case when the vector is projected onto the face but it does not intersect with any half edge and any vertex of the corresponding triangle, we insert new vertex at the end position of the vector in the face. Once all the snaxels are processed, the DCEL structure contains vertices matching with the edge contour of the facial part.

2) Hole Cutting: In this step we determine the parts of the head mesh which will be replaced by the corresponding facial parts. In other words, we have to create holes in the head mesh whose size and shape correspond to the facial parts. Using a flooding algorithm we find all polygons which lie inside the borders of the projected contour to create these holes in the head mesh. Then we mark the polygons inside the holes and delete them. They will be subsequently replaced by the meshes of the facial parts.

3) Merging and Sewing: In this stage we use the mean value coordinate [29] to align the mesh of a facial part to the corresponding hole created in the previous phase. To create a single triangular mesh, we merge the points on the contour of the facial part with points forming the hole on the head mesh. It is also necessary to recompute the DCEL structure to reflect the changes in the topology and to include the information about materials used in the facial part. Figure 5 shows an example of the sewing phase result for the mouth facial part.

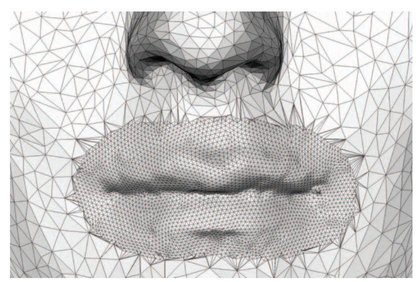


Fig. 5. Resulting polygonal mesh after the contour projection and sewing of the mouth facial part.

I. Texture Mapping

The final step of the whole procedure consists of texture projection. Our solution is currently very simple. Each facial part keeps its original texture. The prototypic head is also textured by a generic skin texture. Along the sewed parts, i.e., the projected contours, the adjacent textures are blurred using the Gaussian kernel.

J. Export

After the process of the procedural generation of the face is finished, we allow for exporting the facial mesh to the *.obj* file format. This format is supported by many modeling software tools where the user can further edit the resulting mesh. The *.obj* format is also quite commonly used in several popular game engines.

IV. Results and User Study

The outputs of our proposed technique for creating artificial 3D facial models from the real datasets were tested with respect to their variability. More specifically, we designed a case study which aimed to reveal the minimal changes of the facial parts which cause that the resulting face is perceived as different from the original one.

Our proposed test was based on the experiment done by McDonnell et al. [7]. The goal of the case study was to analyze peoples ability to identify a single pair of clones, changed by our algorithm, among other completely different faces. The pair of clones which the participant should detect and mark, differed in 0, 1, 2, or 3 facial features.

For this testing phase we implemented a special standalone application. The participants were first given the instructions how to operate the test. Then, for each test, three rows of four models were shown on the screen with two out of twelve slots containing the same character in each test screen. The participants were asked to click on the pair of same or very similar characters as quickly as possible. For each test the maximal time of 120 seconds was allowed for the participants to make their choice. In the tests, we were randomly changing the number of changed facial parts of the same character. So in some tests the characters were completely identical and in others, 1, 2, or 3 changes were made. An example of such testing screen is in Figure 6.

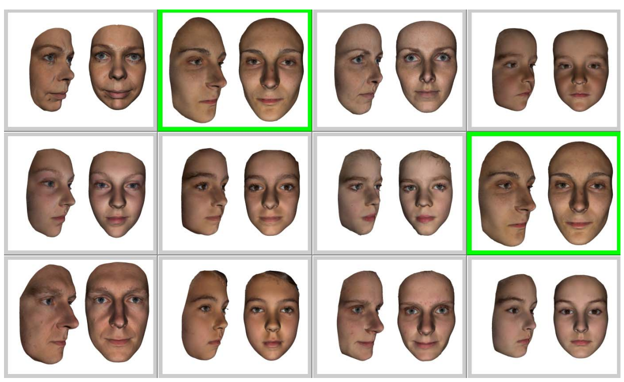


Fig. 6. Example of exactly the same face model in the test application. When user identifies the identical models, these are highlighted by green color.

Our testing was performed by 12 participants (6 males and 6 females). The mean reaction time for the identification of the exact and modified models was measured. The results were evaluated using the Mann-Whitney test [30]. It revealed that, for the facial area model, the change of the facial model by our method significantly increases the time of its identification (the statistical significance—$P = 0.000634$). The mean reaction time for the identification of the exact characters, i.e., without any facial change, was 13.5 ± 5.5 seconds. For characters with changed facial parts, this time value was 51.0 ± 30.9 seconds (see Figure 7).

This preliminary study gives us an interesting result that our algorithm seems to be useful for masking the appearance of cloned facial models, even with small changes of the facial features. Moreover, the results indicate that the more facial features are modulated, the longer time for the identification of the similar model is necessary. However, when changing more facial features, the difference is not statistically significant. Even the change of one facial part leads to the substantial increase of time required for the identification.

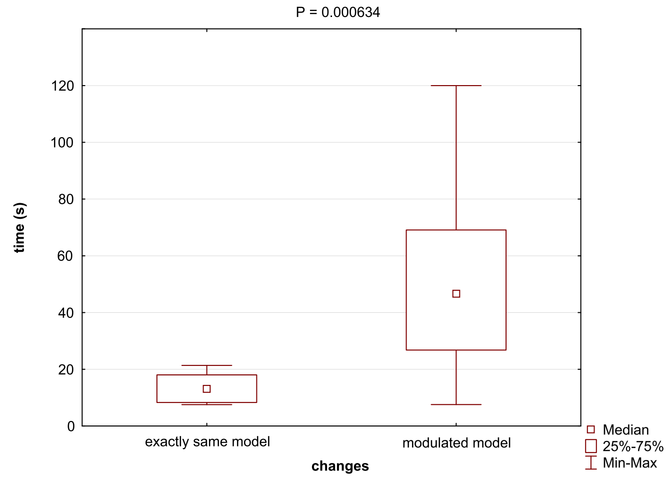


Fig. 7. Distribution of the resulting time spent on the identification of the same (left part of the graph) and modified (right part) human faces.

Figure 8 shows an example of the visual results achieved by a user defined head template where the different number of facial features was replaced.

A. Performance

The application was tested for performance on a machine with the Intel Core i7-3612QM 2.1 GHz CPU, 8 GB of RAM and NVIDIA GeForce GT 630M graphics card. The input data consisted of the segmented models from the F3D-FD database. Each of the input facial features contained about 10,000 polygons. Thus we created their low-poly variations with about 1000 polygons for additional testing. The measured results of the computation time, including the comparison of the generation and mesh completion phase, can be found in Table I. It should be noted that the majority of the computed time is consumed by the mesh completion phase (contour projection and sewing).

TABLE I
Computation Time and Its Dependency on the Number of Generated Facial Models.

	Number of Face Models	Generation	Mesh Completion	Total
Low poly	1 100	0.673 s 18.843 s	1.459 s 1 min 43 s	2.132 s 2 min 2 s
High poly	1 100	0.738 s 21.760 s	48.703 s 54 min 29 s	49.441 s 54 min 51 s

Back to Contents

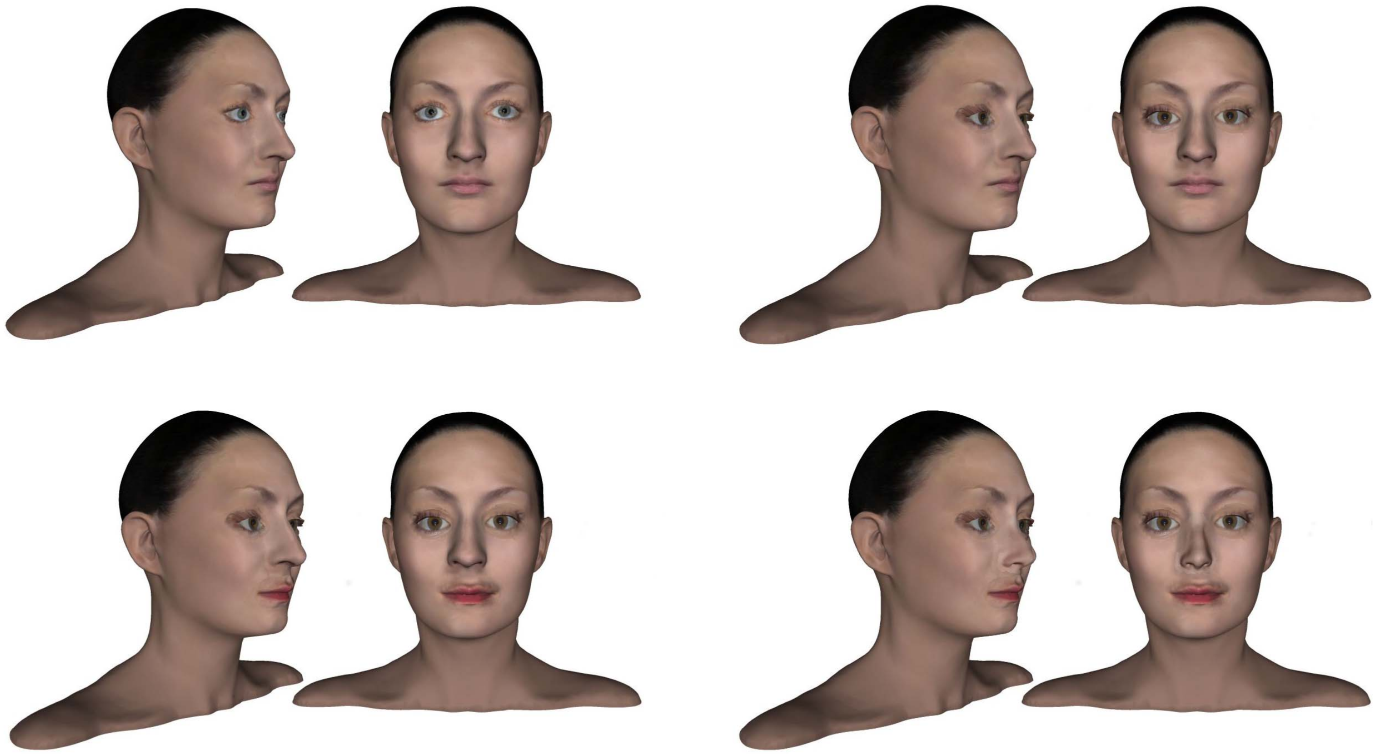


Fig. 8. Results of our technique for a user defined head template. From left to right, from top to bottom: none, one, two, and three facial features changed in the original 3D model.

V. Conclusion and Future Work

In this paper we presented our novel approach to the automatic generation of 3D facial models. Our approach takes as an input a facial mesh of a real person stored in the F3D-FD database or an arbitrarily created model in a given file format. Such model is subsequently segmented and its individual facial parts are then used for automatic or semi-automatic creation of artificial facial models. The resulting facial models can be exported and attached to avatars in various games, mainly those containing large crowds.

As the differences between faces in the crowd are crucial for the evaluation of the variability of the generated crowd, we proposed and conducted a user study focusing on this issue. The aim of the study was to reveal the minimal required changes of individual facial parts that make the face different from the original one. The user study showed significant difference in times required for the identification of the correspondence between two faces. While the identical heads were recognized relatively quickly, the time necessary for the recognition of pairs the faces varying in few facial parts was considerably larger. These results suggest that our algorithm can be successfully used for generating unique faces for avatars. The presented algorithm was integrated into the FIDENTIS Analyst software tool which is freely available at www.fidentis.cz.

A. *Future Work*

One of the problems of the current algorithm is that it excludes two other facial parts, i.e., the forehead and the ears. So the challenge for the future is to incorporate these two facial parts into the pipeline of the automatic generation of human heads. Currently, our algorithm also only remaps the texture coordinates, keeping several texture files – one for each facial part and the prototypic head. In future, we would like to incorporate changes, which would create a single texture file. The texture mapping procedure should be also improved to reach more realistic looking and unified skin. We would also like to extend the algorithm to provide more support for 3D scans of human faces, which usually contain only the front part of the face, by allowing to procedurally generate the back part of the head. We believe this functionality could greatly improve the time necessary for manual editing of the scanned heads. With respect to the idea of recreating the entire human head, it would be also interesting to accessorize the mesh in order to add more variability, e.g., by adding a hat or a model of hair.

Acknowledgements

This work was supported by the Masaryk University projects MUNI/A/1213/2014 and MUNI/33/08/2015.

Back to Contents

References

[1] F. Tecchia, C. Loscos, and Y. Chrysanthou, "Image/based crowd rendering." *IEEE Computer Graphics and Applications*, vol. 22, no. 2, pp. 36–43, 2002.

[2] B. Ulicny and D. Thalmann, "Towards interactive real-time crowd behavior simulation," *Computer Graphics Forum*, vol. 21, pp. 767–775, 2002.

[3] D. Gosselin, P. Sander, and J. Mitchell, "Drawing a crowd," *ShaderX3*, pp. 505–517, 2005.

[4] D. Thalmann, C. O'Sullivan, B. Yersin, J. Maim, and R. McDonnell, "Populating virtual environments with crowds," *Eurographics Tutorials*, vol. 1, pp. 21–124, 2007.

[5] S. Dobbyn, R. McDonnell, L. Kavan, S. Collins, and C. O'Sullivan, "Clothing the masses: real-time clothed crowds with variation," *Eurographics 2006 Short Papers*, pp. 103–106, 2006.

[6] D. Thalmann, H. Grillon, J. Maim, and B. Yersin, "Scalable solutions for simulating, animating, and rendering real-time crowds of diverse virtual humans," in *Modeling, Simulation and Visual Analysis of Crowds*, ser. The International Series in Video Computing, S. Ali, K. Nishino, D. Manocha, and M. Shah, Eds. Springer New York, 2013, vol. 11, pp. 123–145.

[7] R. McDonnell, M. Larkin, S. Dobbyn, S. Collins, and C. O'Sullivan, "Clone attack! perception of crowd variety," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 26:1–26:8, Aug. 2008.

[8] A. Aubel, R. Boulic, and D. Thalmann, "Real-time display of virtual humans: Levels of details and impostors," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, pp. 207–217, 2000.

[9] S. Dobbyn, J. Hamill, K. O'Conor, and C. O'Sullivan, "Geopostors: A real-time geometry / impostor crowd rendering system," in *IN SI3D 05: PROCEEDINGS OF THE 2005 SYMPOSIUM ON INTERACTIVE 3D GRAPHICS AND GAMES*. ACM Press, 2005, pp. 95–102.

[10] J. Maim, B. Yersin, J. Pettre, and D. Thalmann, "Yaq: An architecture for real-time navigation and rendering of varied crowds," *IEEE Computer Graphics and Applications*, vol. 29, no. 4, pp. 44–53, 2009.

[11] V. Bruce and A. Young, *Face Perception*. Psychology Press, 2012.

[12] F. H. Pighin, J. Hecker, D. Lischinski, R. Szeliski, and D. Salesin, "Synthesizing realistic facial expressions from photographs," in *SIGGRAPH*, 1998, pp. 75–84.

[13] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3d faces," in *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '99. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1999, pp. 187–194.

[14] K. Kähler, J. Haber, H. Yamauchi, and H.-P. Seidel, "Head shop: Generating animated head models with anatomical structure," in *Proceedings of the 2002 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, ser. SCA '02. New York, NY, USA: ACM, 2002, pp. 55–63.

[15] A. Maejima and S. Morishima, "Human head modeling based on fast-automatic mesh completion," in *ACM SIGGRAPH ASIA 2009 Posters*, ser. SIGGRAPH ASIA '09. New York, NY, USA: ACM, 2009, pp. 53:1–53:1.

[16] B. Allen, B. Curless, and Z. Popović, "The space of human body shapes: Reconstruction and parameterization from range scans," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 587–594, Jul. 2003.

[17] V. Kraevoy and A. Sheffer, "Template-based mesh completion," in *Proceedings of the Third Eurographics Symposium on Geometry Processing*, ser. SGP '05. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2005.

[18] S. Inversions, "Facegen," http://www.facegen.com/, accessed: 2015-04-14.

[19] M. Bastioni, S. Re, and S. Misra, "Ideas and methods for modeling 3d human figures: The principal algorithms used by makehuman and their implementation in a new approach to parametric modeling," in *Proceedings of the 1st Bangalore Annual Compute Conference*, ser. COMPUTE '08. New York, NY, USA: ACM, 2008, pp. 10:1–10:6.

[20] I. Chalás, P. Urbanová, Z. Kotulanová, M. Jandová, M. Králík, B. Kozlíková, and J. Sochor, "Forensic 3D facial identication software (FIDENTIS)," in *Proceedings of the 20th World Meeting of the International Association of Forensic Sciences*, 2014.

[21] Z. Kotulanová, I. Chalás, and P. Urbanová, "3D virtual model database of human faces: Applications in anthropology and forensic sciences," in *Mikulov Anthropology Meeting. The Dolní Věstonice Studies 20*. Brno: Academy of Sciences of the Czech Republic, 2014, pp. 177–180.

[22] P. J. Besl and N. D. McKay, "Method for registration of 3-D shapes," in *Robotics-DL tentative*. International Society for Optics and Photonics, 1992, pp. 586–606.

[23] K. Hildebrandt and K. Polthier, "Anisotropic filtering of non-linear surface features," in *Computer Graphics Forum*, vol. 23, no. 3. Wiley Online Library, 2004, pp. 391–400.

[24] M. Meyer, M. Desbrun, P. Schröder, and A. H. Barr, "Discrete differential-geometry operators for triangulated 2-manifolds," in *Visualization and mathematics III*. Springer, 2003, pp. 35–57.

[25] N. Loménie and G. Stamon, "Morphological mesh filtering and α-objects," *Pattern Recognition Letters*, vol. 29, no. 10, pp. 1571–1579, 2008.

[26] M. Galvánek, K. Furmanová, I. Chalás, and J. Sochor, "Automated facial landmark detection, comparison and visualization," in *Proceedings of the 31st Spring conference on Computer Graphics*, ser. SCCG '15, 2015.

[27] T. Zinsser, J. Schmidt, and H. Niemann, "Point set registration with integrated scale estimation," in *International Conference on Pattern Recognition and Image Processing*, 2005, pp. 116–119.

[28] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[29] Z. Deng, G. Chen, F. Wang, and F. Zhou, "Mesh merging with mean value coordinates," in *Proceedings of the 2012 Fourth International Conference on Digital Home*, ser. ICDH '12. Washington, DC, USA: IEEE Computer Society, 2012, pp. 278–282.

[30] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *The annals of mathematical statistics*, pp. 50–60, 1947.

Back to Contents

Experience Surveillance Suite for Unity 3D

Alexiei Dingli, Andreas Giordimaina
Department of Intelligent Computer Systems
University of Malta
Msida, Malta

Héctor P. Martínez
Institute of Digital Games
University of Malta
Msida, Malta

***Abstract* - Monitoring the emotional state of players in games can get quite complex, taking into consideration that the game context affects the player and that a game may contain various emotional features. Furthermore, since the experience of playing a game occurs unconsciously, methods such as think aloud may interrupt the playing experience. Other methods include fitting cables and electrodes to the player in order to monitor measurements such as heart rate. Although such devices can offer significant results, they are not commonly found and may cause discomfort. In this project we propose a webcam-based heart rate monitoring method that can be used to predict the player's emotional state. The first objective was to analyze the heart rate changes with respect to the players' emotional state. The evaluation resulted in positive results, where the heart rate showed correlation with the following emotional states; frustration, fun, challenge and boredom. The second objective was to create a webcam-based method to monitor the heart rate. This was performed by extracting the RGB channels from the face region and then retrieving the underlying components using a dimensionality reduction method. Although the results obtained from the webcam-based method were not ideal, this was expected taking into consideration that the method was tested under realistic scenarios. The last objective was to predict the player's emotional state using the heart rate obtained from the webcam-based method. The accuracy of the prediction was up to 76%, which exceeds the aim of the project. Finally, by using the evaluation results it was possible to define a set of approaches on how this project can be extended by future researchers.**

Keywords—image processing; webcam-based, video-based; heart rate monitoring; non-contact; adaptive gaming; affective computing; user experience; human-computer interaction

I. INTRODUCTION AND BACKGROUND

In the past, the main focus of game developers was based on increasing the quality of games. This led to drastic improvements in game engines, for example, graphics, performance and game physics. Recently, more attention is being given to monitoring the player's emotional states, which can then be utilized to adapt the game content to the player's respective abilities [1]. Such games are known as *affective games* [2]. The objective of adaptive games is to create a loop between the user and the game, known as the *affective loop*. For example, if the user is currently feeling frustrated, then the game should automatically identify it and reduce the difficulty level. This would gradually help to reduce the player's frustration. In return, the game will then start increasing the difficulty level again.

Affective games are based on a theory, known as the *flow theory*, created by [3]. This theory defines multiple concepts of what makes activities, such as playing, enjoyable. The idea is that an activity should provide challenges to the people, but not exceeding their abilities. This theory was later adapted by [4] for digital games. Furthermore, [5] performed a research based on the flow theory in order to analyze the player experience. The research highlighted the benefits of adaptive games and how to keep the player engaged in a game, as shown in figure 1. Furthermore, it also highlighted the importance of not disturb the user while playing in order not to interrupt the experience.

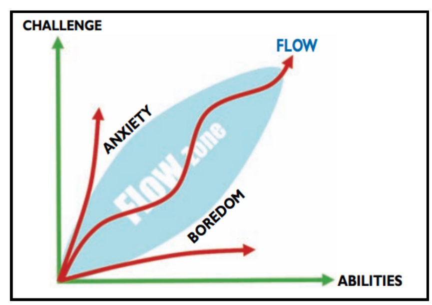


Figure 1. Balance between Challenge and player's ability (Source: Chen, 2007).

Monitoring the player experience can get quite complex, taking into consideration that gameplay data may not be enough. Additional devices may be used, for example fitting cables and electrodes to the player in order to monitor measurements such as facial expressions, heart rate and speech. Although such devices can offer significant results, they are not commonly found and may cause discomfort to the players [6]. Our motivation for this research is to utilize measurements that traditionally required complex devices, but instead extracted using webcam-based methods. For example, there are multiple systems that are able to identify the facial expressions using a webcam. Such systems have been researched for many years and minor changes in facial expressions are still difficult to capture [7]. Alternatively, webcam-based methods for heart rate monitoring have recently been receiving an increasing amount of attention and major advancements have been made in a short time.

The remainder of this paper is organized as follows: section two states the aim and objectives of the research. Then, section three provides an overview of the design used for the implementation. Next, section four gives further implementation details. Finally, section five describes the experiment setup used and evaluates the objectives of the research.

II. Aims and Objectives

Previously webcam-based methods for heart rate monitoring have only been used for medical purposes. This led to the following scientific question:

> *"Can webcam-based heart rate monitoring methods be used to predict the player's emotional state in a real game scenario for the majority of the time?"*

This question combines the research performed in the medical field for monitoring the heart rate using webcam-based methods with the research performed in digital games for analyzing the players' emotional state changes. The answer for the scientific question depends on the following three objectives:

1. Analyzing heart rate changes with respect to the player experience
2. Building and evaluating the webcam-based method
3. Predicting the player's emotional state using the heart rate

In order to answer the scientific question being researched in this project, all of these objectives need to be completed. Furthermore, in order to provide a positive answer for the question, the accuracy of the prediction needs to be above 50%. If the aim of the project is not met, then this research can be used as guidance for future researchers by identifying any difficulties encountered and the possible solutions.

III. Design

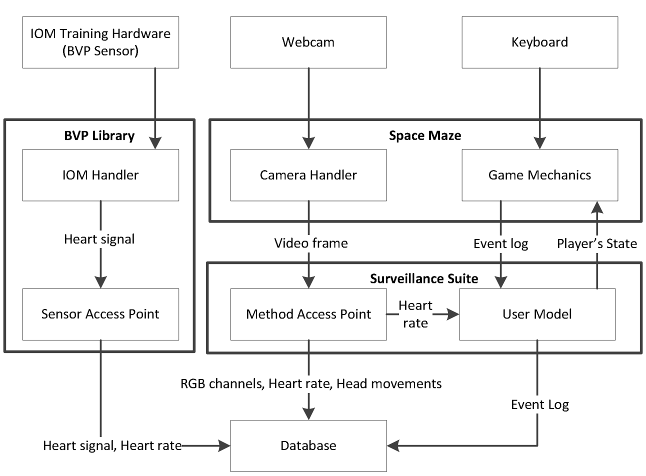


Figure 2. System overview.

The design of the system is divided into three components, as shown in figure 2. The first component is the *Surveillance Suite*, which is the main component of the system. It contains the webcam-based monitoring method and the player's emotional state prediction. The second component is the *BVP library*, which is used to handle the blood volume pulse (BVP) device. The aim of this component is to accurately monitor the heart rate and then compare the data with that obtained from the webcam-based method. The third component is *Space Maze*, which is the game being used in this project for the experiments.

IV. Implementation

In this section, the three components of the system will be explained into further detail. The first component to be described is *Surveillance Suite*, which is divided into two parts; *Method Access Point* and *User Model*. Then an overview about the *BVP library* and *Space Maze* will be given.

Method Access Point

The structure for the webcam-based method is depicted in figure 3. The input for the method is a stream of images received from the webcam. For every image, the face region is detected using the Open Computer Vision (OpenCV) library, which is commonly used by other researchers [8] [9] [10]. Then the Red, Green and Blue (RGB) channels are extracted by taking the average from all the pixels in the face region. Then, taking into consideration that the RGB channels may contain external noise, a dimensionality reduction method is used to extract the underlying components. The method being used in this project is Principal Component Analysis (PCA), which is a linear method. Other methods are available, but [11] showed that the PCA is able to achieve sufficient results in real-time. Next, one of the three components that the PCA returns is chosen. In this research, the third component was used since it was the most accurate when compared to the other two components. Furthermore, the PCA was computed using a 30 seconds moving window [10].

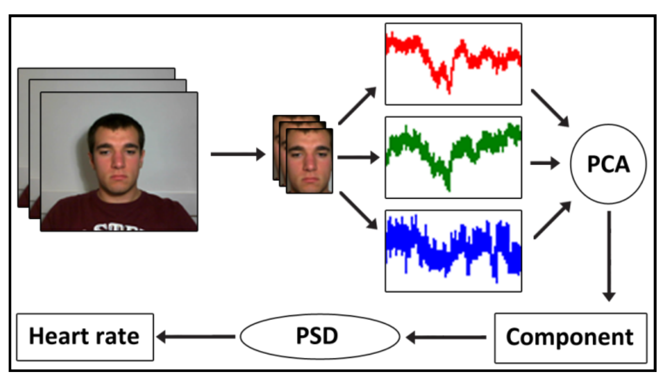


Figure 3. Overview of the heart monitoring method.

After the component is chosen, Power Spectral Density (PSD) is used to extract the heart rate, which is computed using the Fast Fourier Transform (FFT) [11]. The idea is to identify how the power of a signal is distributed over the frequency domain. The frequency range used is that of [0.75, 3], which corresponds to the heart rate’s minimum and maximum beats per minute of 45 and 180 respectively [12] [9].

User Model

The user model is used to represent the player's emotional state. Given a set of measurements, known as input modalities or features, the user model predicts the player's emotional state using mapping. Although the more features collected, the better the accuracy of the prediction tend to be, some features may be irrelevant when predicting some emotional states. Furthermore, if these irrelevant features are not removed, then they may negatively affect the accuracy of the user model [13]. The elimination of irrelevant features is known as *feature selection*, which is performed separately for every emotional state.

The process for feature selection is divided into two parts; searching for a solution and evaluating the solution [14]. For this project, the search algorithm being utilised is *Sequential Forward Selection* (SFS). On the other hand, the solution evaluation is performed using Single-Layer Perceptrons (SLP). After retrieving the sub-set features for every emotional state, the user model is created by inputting the features subsets into a Multi-layer Perceptrons (MLP), in order to support nonlinear relations. Both the SLP and MLP are trained using neuro-evolution, more specifically, genetic algorithms.

BVP library

The BVP device being used is known as IOM Training Hardware, which was provided by the Institute of Digital Games. The device is used to monitor the heart rate of the participants and then compare the results with those obtained from the webcam-based method. To our knowledge, the device does not have API available to connect with Unity 3D, and therefore had to be created manually. This was performed by handling the device as a Human Interface Device (HID) and retrieving a stream of strings as input. Then, a peak detection algorithm is applied on the input in order to retrieve the heart rate.

Space Maze

Space Maze is a prey/predator game created in Unity 3D, which was also provided by the Institute of Digital Games. The game is based on *Maze-ball*, which is another game previously used by other researchers [15] [16]. The player’s objective in the game is to collect all three pellets and then exiting the maze while trying to avoid hitting the enemies in less than 120 seconds. The game is played using only one hand (arrow keys), and therefore makes it possible to attach the BVP device to the other hand. Furthermore, it contains two separate levels, which can be played in any order.

Space Maze was initially created by a previous research in order to analyze the affect of camera control [17]. Therefore, it contains a dynamic camera controller, which is based on the Constraint Satisfaction Problem (CSP) framework. The camera angle is changed according to the following three variables; distance, height and frame coherence, as shown in figure 4.

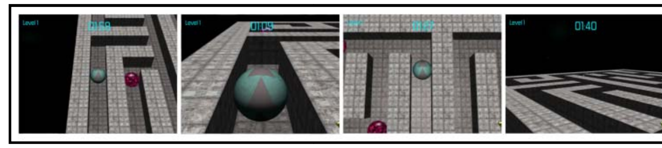


Figure 4. Different camera positions which changes automatically.

V. Results and Evaluation

In this section, the experiment results and evaluation of the research objectives are discussed. The experiments were performed at the Institute of Digital Games in a laboratory setting. In total, 28 Maltese people (25 male and 3 female) participated for the experiments, invited from an event held by the Institute of Digital Games, known as the *Game nights*. The age varied between 18 and 32, with the average age being 23.

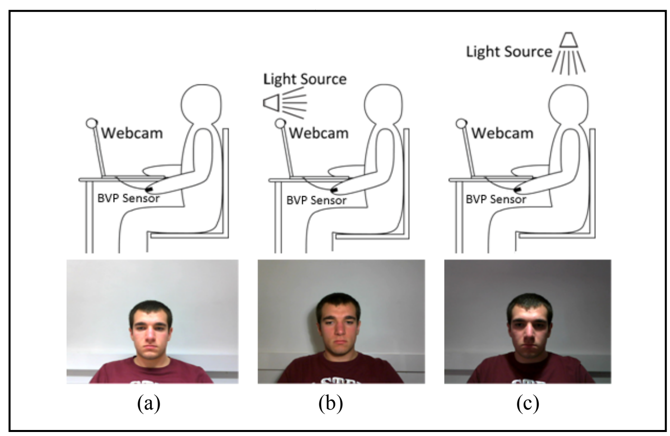


Figure 5. Lighting conditions being taken into consideration for the experiments, where in (a)the light source is the ambient light, in (b) the light source is placed in front of the participants and in (c) the light source is placed above the participant.

The experiment consisted of participants playing the game *Space Maze* under different lighting scenarios. The reason for using multiple lighting scenarios is to provide the accuracy of the system in the real world. The experiment setup is shown in figure 5. The participants were asked to read the consent form and then play two levels of the game, which represent easy and hard, under all three scenarios. Then, after every scenario, the participants were asked to fill in a questionnaire. The emotional states being taken into consideration are; frustration, boredom, fun, and challenge.

The questionnaire used for the experiments was a 4-alternative forced choice (AFC) preference-based questionnaire. Another common questionnaire is the Game Experience Questionnaire (GEQ), which uses a 5-point Likert scale [18]. The reason for choosing the former questionnaire is that it tends to be more consistent when compared to the latter [15]. The idea is that different people may have different definitions for the score “3” in the question “How fun a game is on a scale from 1 to 5”. On the other hand, if a participant is asked which game was more fun between Game A and Game B, then the question is less ambiguous.

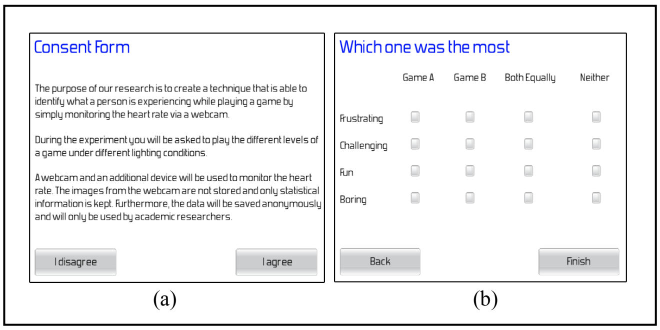


Figure 6. (a) is the consent form, while (b) is the questionnaire

Taking into consideration that every participant played a total of six games and answered 3 questionnaires, a total of 168 game instances and 84 questionnaires were obtained from the experiments. Next, this data is used to evaluate the project’s objective.

Analyzing the players' experience

In this section we are going to analyze the first objective of the project, which is to analyze the heart rate changes with respect to the player experience. This was performed by calculating the correlation coefficients between the emotional states and the heart rate measurements, using equation 1. The equation is computed using game pairs that showed *clear preference*, that is, the game pairs where the questionnaire result was not answered as "Both equally" or "Neither". After calculating the correlation coefficients, the p-value was calculated in order to identify the most significant calculations.

$$c(z) = \sum_{i=1}^{N_s} \frac{z_i}{N_s} \qquad (1)$$

Where $\boldsymbol{N_s}$ is the number of pairs where participants indicated clear preference, and $\boldsymbol{z_i = 1}$ if the participant preferred the game with a higher value of the variable being analyzed, else $\boldsymbol{z_i = -1}$. From the data gathered, $\boldsymbol{N_s}$ is set to 49, 52, 39 and 18 with respect to the emotional states frustration, challenge, fun and boredom.

From the heart rate results, the following measurements were extracted; average heart rate, maximum heart rate, minimum heart rate, difference between maximum and minimum heart rate, standard deviation of heart rate, average inter-beat interval, maximum inter-beat interval and minimum inter-beat interval. Additional features were extracted from the game data, which are; average distance between the character and the closest enemy, times hit by an enemy and game duration.

For the frustration emotional state, the maximum heart rate tends to be higher, while the average inter-beat interval is also higher. Furthermore, frustration was also common in longer games and when the character was repeatedly hit by the enemies.

Next, the boredom emotional state is analyzed. Although some correlation was identified between the heart rate changes and the boredom emotion, none of them were significant enough. This may be due to the lack of clear preference for the boredom emotion, where only 18 pairs of games were available to calculate the correlation.

For the fun emotional state, there are three features related to heart rate that offer correlation with the emotion, which are; minimum heart rate, maximum heart rate and the difference between them. It appears that the maximum heart rate is lower than usual, while the minimum heart rate and minimum inter-beat interval are higher. These results are also supported by the high correlation with the difference between the minimum and maximum heart rates, where they tend to be closer while the participants are enjoying the game. The correlation with the minimum heart rate was also noticed by [16]. For the average distance from enemies, times hit by the enemy and game durations; the results oppose those obtained for the frustration emotion. That is, the participants enjoyed the game when they did not hit the enemies too frequently and managed to complete the level.

For the challenge emotional state, a significant correlation was noted with standard deviation heart rate and with the difference between maximum and minimum heart rates. This shows that during the challenge emotional state, the heart rate tends to perform drastic changes. Furthermore, the maximum heart rate tends to be higher and the participants appear to perform more head movements. The final correlation occurred with the times hit by the enemy and the average distance from the enemy. Although the latter two features offer significant correlations, they are not as significant when compared to the frustration emotional state. That is, although being repeatedly hit by the enemy can be considered challenging, if it occurs too regularly then it becomes frustrating.

Heart rate monitoring

In this section, the second objective of the project is going to be analyzed, which is, building and evaluating the webcam-based method used to monitor the heart rate. The evaluation is performed by comparing the heart rate obtained from the BVP sensor with that obtained from the webcam-based method. The method is being tested using 168 game instances, with the average of 91 seconds long; therefore there is over 4 hours of footage being utilized.

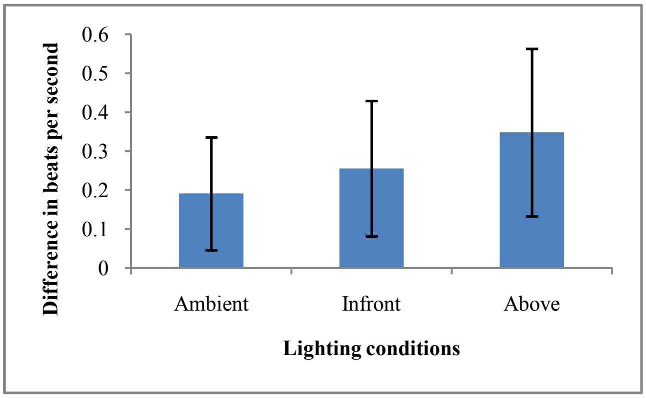


Figure 7. The overall accuracies under different lighting conditions.

As shown in figure 7, the highest accuracy was achieved when the ambient light was used, with the accuracy of 0.19 beats per second difference from the BVP device. Although this accuracy is not ideal, these results were achieved under realistic scenarios and therefore cannot be compared to previous researches. As shown in figure 7, the accuracy was heavily impacted by the lighting scenarios, where the accuracy for the *infront* and *above* scenarios decrease to 0.25 and 0.35 beats per second respectively. In previous research, the participants were asked to not perform head movements. Taking into consideration that such behaviour is unnatural while playing, participants were allowed to performed head movements. After identifying the correlation between the accuracy and the head movements performed, it resulted that the accuracy was heavily impacted by head movements.

User Modeling

In this section, the third objective is going to be evaluated, which is to predict the player's emotional state for the majority of the participants. Furthermore, although the accuracy of the webcam-based method is not ideal, it may still be sufficient for the purpose of predicting the player's emotional state. Therefore, the user model was built using game events and the heart rate obtained from the webcam-based method. For the evaluation, the user model was created using different approaches, as shown in table 1. When comparing SLP with MLP (SF), it shows the advantage of using nonlinear methods over linear methods when creating user models. Furthermore, when comparing MLP (SF) with MLP (All), it shows the impact feature selection has on the accuracy of the user model.

	SLP	MLP (SF)	MLP (All)
Frustration	73%	80%	69%
Bored	71%	78%	83%
Fun	65%	77%	70%
Challenge	63%	69%	53%
Average	68%	76%	68%

Table 1. The accuracies for four different user models. SLP is the user model built during feature selection. MLP (SF) is given the selected features as inputs. MLP (All) is given all the features.

Observations

In this section, the experiment setups and an evaluation on the collected data were discussed. In order to determine whether the aim of the project was met, all the objectives must be completed. For the first objective, it was shown that there is a strong correlation between heart rate changes and player experience, and therefore the heart rate was concluded to be an important asset when predicting the player's emotional state. For the second objective, the method was created but although achieving decent results, these were not ideal. That being said, the webcam-based results were used when creating the user model, and for the user model the accuracy achieved was that of 76%. Therefore, this shows that the webcam-based method accuracy was sufficient for building the user model. Finally for the third objective, the user model was able to predict the emotional state up to 76%, which surpasses the aim of the project. Therefore, in this research it was shown that webcam-based methods are sufficient to predict the player's emotional state in a real game scenarios.

CONCLUSIONS AND FUTURE WORK

Although the aim of the project was reached, there are some components that can be further improved in order to achieve better accuracy. The most vulnerable part of the Surveillance Suite is the webcam-based method, which is heavily impacted by the external noise. This includes handling head movements, speech and different lighting conditions.

Although in this project, talking while playing the game was considered as a negative concept, speech can also be used as an input modality. Therefore, when speech occurs there is a trade-off between decreasing the accuracy of the webcam-based method and increasing the accuracy of the speech input modality.

The toughest challenge for this system is to achieve more accurate results under unfavourable lighting conditions. The first approach that can be taken into consideration is to create a technique that identifies the participant's face regions that are most exposed to the light and use those regions. When the light source is very low, the monitor's brightness should also be taken into consideration in order to identify if the brightness is sufficient to be used as a light source.

Acknowledgements

The author would like to thank that Institute of Digital Games for providing the resources used during the experiments and for their feedback and advice throughout the research.

Bibliography

[1] L. E. Nacke, A. Drachen, K. Kuikkaniemi, J. Niesenhaus, H. J. Korhonen, W. M. v. d. Hoogen, K. Poels, W. A. IJsselsteijn and Y. A. W. d. Kort, “Playability and Player Experience Research,” London, 2009.

[2] E. Hudlicka, “Affective game engines: motivation and requirements,” New York, 2009.

[3] M. Csikszentmihalyi, Flow: The Psychology of Optimal Experience, 1990.

[4] P. Sweetser and P. Wyeth, “GameFlow: a model for evaluating player enjoyment in games,” *CIE Computers in Entertainment (CIE),* vol. 3, no. 3, pp. 3-3, 2005.

[5] J. Chen, “Flow in Games (and Everything Else),” *Communications of the ACM,* vol. 50, no. 4, pp. 31-34, 2007.

[6] N. Shaker, S. Asteriadis, G. N. Yannakakis and K. Karpouzis, “A game-based corpus for analysing the interplay between game context and player experience,” *Proceedings of the 4th international conference on Affective computing and intelligent interaction,* vol. 2, pp. 547-556, 2011.

[7] E. Ganglbauer, J. Schrammel, S. Deutsch and M. Tscheligi, “Applying Psychophysiological Methods for Measuring User Experience Possibilities, Challenges and Feasibility,” in *Possibilities, Challenges, and Feasibility. Proc. User Experience Evaluation Methods in Product Development workshop*, 2009.

[8] M.-Z. Poh, D. J. McDuff and R. W. Picard, “Non-contact, automated cardiac pulse measurements using video imaging and blind source seperation.,” *Optical Society of America,* vol. 18, no. 10, 2010.

[9] L. Wei, Y. Tian, Y. Wang, T. Ebrahimi and T. Huang, “Automatic Webcam-based Human heart rate measurements using laplacian eigenmap,” Heidelberg, 2012.

[10] S. Kwon, H. Kim and K. S. Park, “Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone,” San Diego, CA, 2012.

[11] M. Lewandowska, J. Ruminski, T. Kocejko and J. Nowak, “Measuring Pulse Rate with a Webcam - a Non-contact Method for Evaluating Cardiac Activit,” Szczecin, Poland, 2011.

[12] T. Pursche, J. Krajewski and R. Moeller, “Video-based heart rate measurement from human faces,” Las Vegas, NV, 2012.

[13] C. Pedersen, J. Togelius and G. Yannakakis, “Modeling Player Experience for Content Creation,” *Computational Intelligence and AI in Games,* vol. 2, no. 1, pp. 54 - 67, 2010.

[14] P. Burell and G. N. Yannakakis, “Towards Adaptive Virtual Camera Control in Computer Games,” Bremen, Germany, 2011.

[15] G. N. Yannakakis and J. Hallam, “Ranking vs. Preference: a comparative study of self-reporting,” in *Proceedings of the 4th international conference on Affective computing and intelligent interaction*, Heidelberg, 2011.

[16] H. Martinez, A. Jhala and G. Yannakakis, “Analyzing the Impact of Camera Viewpoint on Player Psychophysiology,” Amsterdam, 2009.

[17] Y. Knight, H. Martínez and G. Yannakakis, “Space Maze: Experience-Driven Game Camera Control,” Crete, Greece, 2013.

[18] R. L. Mandryk, K. M. Inkpenb and T. W. Calverta, “Using psychophysiological techniques to measure user experience with entertainment technologies,” *Behaviour & Information Technology,* vol. 25, no. 2, pp. 141-158, 2006.

Project Sanitarium: Gaming TB

A serious game for a serious problem

Dr. Iain Donald, Dr. Karen Meyer, Mr John Brengman
School of Arts, Media and Computer Games
Abertay University
Dundee, UK
i.donald@abertay.ac.uk, k.meyer@abertay.ac.uk, 1004767@live.abertay.ac.uk

Professor Stephen Gillespie, Dr. Ruth Bowness
School of Medicine
University of St Andrews
St. Andrews, Fife
shg3@st-andrews.ac.uk, rec9@st-andrews.ac.uk

***Abstract*— Collaborative projects between Industry and Academia provide excellent opportunities for learning. Throughout the academic year 2014-2015 undergraduates from the School of Arts, Media and Computer Games at Abertay University worked with academics from the Infection Group at the University of St Andrews and industry partners Microsoft and DeltaDNA. The result was a serious game prototype that utilized game design techniques and technology to demystify and educate players about the diagnosis and treatment of one of the world's oldest and deadliest diseases, Tuberculosis (TB). Project Sanitarium is a game incorporating a mathematical model that is based on data from real-world drug trials. This paper discusses the project design and development, demonstrating how the project builds on the successful collaborative pedagogical model developed by academic staff at Abertay University. The aim of the model is to provide undergraduates with workplace simulation, wider industry collaboration and access to academic expertise to solve challenging and complex problems.**

Keywords—games with purpose; games for change; serious games; games education; educational games

I. Introduction

Project Sanitarium is a game created by undergraduate students at Abertay University in Dundee, Scotland. It is a 'serious game' designed to use real world medical data regarding the treatment of Tuberculosis (TB) and combine it with proven game mechanics to deliver a simulation that could entertain players, mimic clinical trials and provide a powerful tool for increasing awareness of one of the world's oldest diseases. The project was the result of the successful collaboration of Abertay's School of Arts Media and Computer Games pedagogical model for workplace simulation, the medical expertise of the University of St Andrews Infection Group and the technical integration of Microsoft's Windows 8 platform, Azure technologies and DeltaDNA's games analytics platform.

Historically TB has been one of the world's most deadly diseases. Today it is second only to HIV/AIDS as the greatest killer worldwide due to a single infectious agent, and patients have the same chance of surviving Ebola without treatment as surviving TB with treatment. [1][2] Whilst considerable progress has been made in treating the disease there remains much that still needs to be researched and understood. In recent years drug resistance has become a particular concern. It was building on the clinical research of Professor Stephen Gillespie of the University of St Andrews that the opportunity for the project came about. Prof. Gillespie and the Infection Group were keen to develop a simple visual tool to present data from clinical trials. In August 2014 a creative brief was developed with the aim to create a game that would:

- Mimic the disease developing in a person
- Mimic how treatment might cure the disease
- Explore how new treatments might work in the future

Key to the brief was the need to visually simplify the multiple complex parameters for people developing new treatments for the disease, and to take what is known and present it in a way that could engage players in order to help them understand more about how treatment works.

The brief identified several possible audiences for the game and considered the commercial success of games such as *Pandemic 2* and *Plague, Inc.* as examples of people's fascination with disease in a gaming context. The main audiences for the game were identified as:

- Biology, health sciences and medical students
- Global health and development advocates
- TB Drug developers
- Anyone who wants to understand more about the world's oldest disease

The Abertay team was tasked with delivering a playable prototype. The prototype aimed to explore potential options for patient design, enhancing the virulence of the bug, and designing and testing better treatments. Critically the brief emphasized that the game should be fun. While there was some desire that the game should reflect the combat between the bug and the host and provide a quasi-realistic feel with images of the lungs, bacteria and human immune cells, the Abertay team was to decide upon the best direction in terms of game aesthetics, mechanics and delivery.

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

II. BACKGROUND

A. Project Context

Tuberculosis is a disease that kills 1.5 million people each year or more than one person every 20 seconds. [3] Widely ignored by the mainstream media the global threat of TB has been growing, fuelled by global poverty, the AIDS epidemic, and drug resistance. Currently, the recommended treatment therapy for TB lasts for 6 months, and many patients cannot complete the treatment because of toxicity. Put simply, there is a need for faster-acting tools to treat TB in all forms, wherever it may strike around the world. A shorter, less toxic regimen would enable more patients to complete the therapy while also reducing the risk of drug resistance occurring.

New treatments are currently tested in lengthy and expensive clinical trials, but a mathematical model enables researchers to test the effects of varying parameters (including those associated with drug efficacy) and explore 'virtual clinical trials', more cheaply. The team at the University of St Andrews used results from a series of global clinical trials to develop a mathematical model for the treatment of tuberculosis; however there remained an issue in being able to display complex research and data to a wide variety of researchers, academics, funding bodies and medical professionals. [4][5][6] The result was the collaborative project with Abertay University to help demonstrate the power of the mathematical model whilst maintaining an accurate yet engaging tool for visualizing the model.

B. The AMG Abertay Model

An inter-disciplinary approach has always been at the core of teaching game development at Abertay University. The School of Arts, Media and Computer Games (AMG) has successfully developed a pedagogical model based around Etienne Wenger's Communities of Practice and integrated it with the team based and collaborative models widely used throughout the creative industries. [7] The pedagogical model is now core for the curriculum in all the School's undergraduate programs and is centered upon placing all third-year AMG undergraduates into small inter-disciplinary teams to work on an industry and client led creative brief. Each year approximately 20-25 student led teams take on board briefs that challenge them to develop a prototype that can be professionally presented at the end of the academic year. Over the years clients have ranged from large multinational corporations to small indie game developers, with almost every conceivable combination in between.

C. The Project Brief

For the Academic Year 2014-15 the students were given the choice of over 20 clients and 30 different briefs. Two of the clients were Microsoft and the Infection Group at the University of St Andrews. Microsoft set a number of very open briefs that encouraged the use and development of their technology, including the Kinect Version 2, the Windows 8 Phone ecosystem and Azure, Microsoft's cloud platform. The Infection Group had more specific aims and were looking for a team to help them visualize the data from recent clinical trials in a more engaging and entertaining way. [8] One student team proposed combining several briefs to create a game that would focus around the Infection Group creative challenge and examine the treatment of Tuberculosis. This was done by utilizing the Microsoft systems for the Windows 8 Phone and harnessing Azure to handle the data.

D. The Project Team

The Team consisted of 10 undergraduates split into the roles of Producer, Game Designer, 3 Programmers, 3 Artists and 2 Audio Engineers. These directly reflected the makeup of Abertay's undergraduate programs with participants from the following degree courses [9]:

- BSc (Hons) Computer Games Technology
- BSc (Hons) Computer Game Applications Development
- BA (Hons) Computer Arts
- BA (Hons) Creative Sound Production
- BA (Hons) Game Design & Production Management
- BA (Hons) Visual Communication & Media Design

The team were assisted by their module tutors, Dr. Iain Donald and Dr. Robin Sloan, and assigned an academic mentor, Dr. Karen Meyer, to assist with the mathematical modelling. The team also had feedback at regular interviews from both Microsoft and the University of St Andrews Infection Group.

E. Project Structure

Although the module is core for all undergraduate programs it only accounts for a quarter of each student's academic workload. In total the number of hours equates to each person working full-time for approximately 6 weeks. Given the time constraints and the additional complications of working in a University environment (no dedicated PC's or workspace, a flat team hierarchy, difficulty in bringing in additional resources and the restrictiveness of the University calendar) the team resolved to work with a hybrid Waterfall and Agile model. This balanced model was chosen to best deliver features within the confines of the University two-semester structure. [10] The approach allowed the team the benefits of a structured system to communicate with the clients directly, and create identifiable milestones for the clients, and themselves. However it also allowed the space to respond to feedback, via iterative 'sprints' in order to fine-tune the game and make any adjustments as needed within the game development process. From the outset the team resolved that they were creating a serious game using serious technology and that delivery to different clients across industry and academia required clear, strong models for planning and delivery. The project plan developed into two distinct phases that aligned with the University semesters.

In the first phase of the project, the team utilized the Waterfall method and divided the tasks into distinct milestones, with the goal of an established design by the end of the first semester, and a rudimentary digital prototype that could be tested for fun and value. The milestones were broken down as follows:

- M1: Analyzing the problem
- M2: Organizing the Plan
- M3: Creating the Design
- M4: Paper Prototyping
- M5: First Coded Prototypes
- M6: First Play Test

The team worked towards the goals on the basis that there would be indicative feedback that could be iterated on, rather than major changes. The team would then move on to a more agile approach for semester two, with a clear focus on iterative weekly sprints and milestones:

- M7: Game Framework
- M8: Digital Wireframe Prototype
- M9: Three Mini-games
- M10: Create Game Trailer
- M11: Core Game Loop Using Mathematical Model
- M12: Integrating Azure Cloud Features
- M13: Integrating Live Tiles and Push Notifications
- M14: Integrating Analytics
- M15: Play Testing and Bug fixes

The goal of the weekly sprints was to motivate the team and demonstrate progress, while the milestones had the overall aim of having that development component or segment fully completed and testable.

III. RESEARCH & DESIGN

The project had several creative and technical challenges, and from the outset required the students and academics to move into areas that they were not familiar with. Each team member helped to research the overall project area but generally focused on their specific disciplines. The key areas for research were examining tuberculosis, other serious games, understanding the mathematical model, examining core gameplay concepts and defining a clear game aesthetic, visually and aurally.

A. Tuberculosis Research

The team came to the project without any in-depth knowledge and had to learn as much as was possible about the history, diagnosis and treatment of Tuberculosis as was feasible in order to inform the game design. The wealth of material available ensured that the team read medical journals, press articles and documents from various sources to inform themselves. For clear and concise understanding the United Kingdom's National Health Service website provided the basic knowledge in direct language. [11] Two other articles emerged as core for the team to utilize. The first of these was the paper published by the Clinical Trial team in the *New England Journal of Medicine*, which discussed a recent 4-month drug trial treatment regimen. [12] The second was a study on using games to combat Tuberculosis published in the *International Journal of Science and Research*. [13] For students of games and the creative industries the material is quite a departure from the usual research. Moving onto more familiar territory the team researched a number of key games.

B. Game Research

The initial research, beyond understanding the work from the University of St Andrews and Microsoft technology, was to examine other games that the team decided would be similar either in genre or problem-solving. Specifically they examined games that dealt with health and disease, such as *Foldit*, *Fraxinus*, *Play to Cure: Genes in Space*, *Pandemic 2*, *Plague Inc.* These were each played and analyzed in order to inform the design ideas and the user experience.

- *Foldit* is a game that sees the player folding protein strings to try and predict the best possible structure for them. Using this information new protein chains can be designed in an effort to combat diseases and potentially even cure them. Created by various departments and labs from the University of Washington, the game is one of the first examples of crowdsourcing scientific solutions through gaming. The key is in using the player base for their intuitive, 3D-puzzle solving skill that computers simply cannot match. [14]
- *Fraxinus* is a game made to combat a specific species of fungus that was causing Ash Die Back in ash trees across Europe. By creating and matching patterns the player uses real genetic data to try and find why some trees are immune to this fungus. *Fraxinus* demonstrated that it was possible to make a fun, entertaining game based on real scientific data. In particular the gene matching puzzle showed that complex data could be presented in simple form and gameplay. [15]
- *Play to Cure: Genes in Space* was funded by Cancer Research UK and in a similar vein uses the collective force of players to analyze real genetic data. Ostensibly the game sees the player piloting a spaceship in order to collect Element Alpha. Underneath the surface the game finds the optimal route to pick up the most Element Alpha, and in doing so plots a course through genuine 'DNA microarray' data. Crunching the data helps scientists identify the DNA faults that could lead to cancer. [16]
- *Pandemic 2* and *Plague Inc.* are very similar games that put the player in the position of designing and developing a virus to destroy the world's population. In both games the player can evolve and morph the virus to make it more drug-resistant and crucially more deadly. Both of these games influenced the design of Project Sanitarium as the team looked to develop a more positive game and this led to the decision for the player to take on the role of a doctor working on a global scale. The emphasis could therefore be placed on tackling the disease as it surfaces around the globe with

various parameters coming into effect based on local conditions. [17][18]

Each game brought insights into design approaches and the challenge that the team faced in particular with ensuring that the game did not detract from the underlying mathematical model.

C. Mathematical Modelling

The core element of making the game a 'serious game' was the interpretation and integration of the mathematical model for the treatment of tuberculosis (TB) that was developed by Dr. Ruth Bowness and Professor Stephen Gillespie from the University of St Andrews.

The team at the University of St Andrews used results from a series of clinical trials to develop the model. Half of the data were used to constrain the parameters within the model. The other half were used to test the model's results and demonstrate its accuracy. The mathematical model is built into the game to determine a patient's outcome based on the player's choices and performance within mini-games.

The model assumes that two biologically different populations of Mycobacterium Tuberculosis exist within a patient, which they designate "active" and "dormant", and that the bacteria may transfer between states. A pair of coupled first-order ordinary differential equations defines the evolution of these populations. The differential equation that describes the active cells contains a growth term, which represents the replication rate of the actively multiplying bacteria, a lethality rate that captures the drug effect on the bacterial population, and transfer parameters to allow transfer between populations. The differential equation describing the dormant cells equivalently has a lethality rate and transfer parameters, but here there is no replication rate as negligible growth is assumed in this population. Parameter values in the model are based on previous literature, laboratory experiments and available clinical trial data. It was the parameters that make up these equations that the Abertay team used to influence the game design. The equations themselves are soon to be published in a medical journal and hence are not displayed here.

The mathematical model allows different parameters to be altered by TB researchers in order to simulate different future treatments. The game design was required to consider these parameters in order to ensure gameplay affected the model. This included values such as the variation in the number of active/dormant bacteria and the calculated total. Other parameters such as the growth rate of the active population of bacteria, the assumed transfer rates of the bacteria from dormant to active and vice versa, and the death rates of the active and dormant bacteria, were considered for design input. The challenge for the Abertay team was in designing a game that would fit with this mathematical model, and could be used to impact upon the gameplay without taking liberties with the accuracy of the model. To that end the mathematical model strongly influenced the decision to focus on mini-games for the gameplay and frame these with an overarching story for the overall design. The team was fortunate enough to have an academic mentor experienced with the model, Dr. Karen Meyer, based at Abertay University. This enabled the team to iterate on the mini-game design quickly and effectively to ensure that the model was accurately replicated within the gameplay.

D. Game Aesthetics

While the design was being developed, the art and audio team members focused on exploring the aesthetics. The artists began by examining medical imaging to get an understanding of the colors that are used in medical applications and to help get a feel for the overall visual style, before expanding their research to include games that utilize color well. This was much broader than the serious games studied to influence the game design and varied from the DICE's *Mirrors Edge* which uses colors to guide the player through the level, through to the Indie hit *Papers Please* for how information is displayed. [19][20] Initial User Interface Designs were influenced by the clean design of *Deus Ex*. [21] That futuristic feel was further enhanced by research into cinematic blockbusters such as *Iron Man* to the new *Star Trek* reboots. [22][23] The research was met with success and the initial mock-ups created as exploratory designs became the foundation of the look of the game. The audio team worked alongside the art team to ensure that the futuristic feel was delivered consistently through the content, deciding upon a clean electronic ambient soundtrack. The audio was also influenced by Hollywood: *Iron Man, Gravity* and *Interstellar* each encouraged the creation of music that would match the pace of the gameplay without overwhelming or creating audio fatigue in the player. [24][25] However it was the vision to bring in professional voice-over talent, Tara Platt, which helped to establish a clear audio identity. [26] The steady pacing and clear scripting took the early prototypes to a different level and can be identified as the point that people started to take notice of the game.

IV. GAME DEVELOPMENT

The initial ideas were developed through mock-ups and rough prototypes. These were then presented to the clients via a series of meetings and presentations. The level of engagement with the clients led to considerably better feedback and stronger iteration on the core game. Each discipline was involved in the process and the ability of team members to identify issues, problem-solve and re-iterate on all aspects was key to developing a strong prototype.

A. Design Overview

The core game is a simulation based upon the mathematical model, although the model itself is completely hidden from the player. The model is used within the game to give an accurate portrayal of how TB will develop and grow within each patient's body. The game, as a reflection of the model, considers variables such as underlying health concerns, lifestyle choices, geographic locations, socio-economic considerations, or the impact of poor education. Variety in play is created by the interaction with the TB bacteria, which is used to provide a more unique experience for each player.

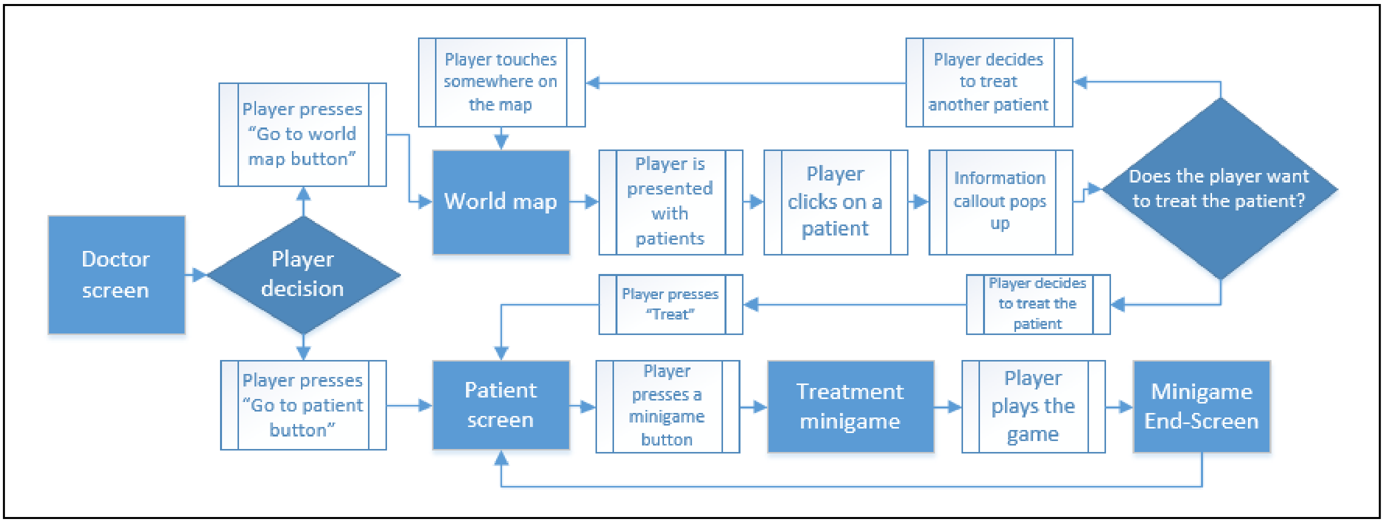


Fig. 1 Main Game Loop

Ostensibly the player takes on the role of a medical doctor throughout their career, from graduating medical school through to retirement. As a doctor the player navigates a variety of mini-games each, asking them to help in the diagnosis and treatment of patients. The more successful the player is, the more complex cases they get to diagnose and treat. The further the player advances, the more they travel to hotspots throughout the world, thereby encountering different parameters and different underlying variables. The player needs to take into consideration the patient history, such as the possibility of the patient being HIV positive, whether they smoke, drink, are overweight, or conversely suffer malnutrition. If the patient is uneducated, then there is a risk that they could stop treatments at the first sign of feeling better, thinking they are cured. This could result in relapse. The game also considers the economic difficulty the patients face in receiving treatment because they don't have the resources to be properly cared for. From a design perspective, having multiple parameters ensured that the game that evolved had depth, but it was often harrowing to realize that these were not fictional or fantastical situations. These were real scenarios faced by medical professionals throughout the world.

The recognition of the number of varying parameters and the importance of understanding how players would react to these led the team to incorporate game analytics software, provided by DeltaDNA. The inclusion of analytics enabled the team to playtest more effectively but added an additional technical challenge to what could have been a complex technology setup. However, the team found that focusing on a single technology ecosystem, provided by Microsoft's Windows 8 platform, enabled iteration to be quick and effective. The design of the main game loop then effectively hid much of the underlying complexity from the player.

B. Main Game Loop

The core of the game centers round the player looping through four event screens: Doctor, World Map, Patient and Diagnosis/Treatment. See Fig 1.

The Doctor screen provides all the player-specific information regarding their status within the game world. This includes personal and medical information, plus game events that may impact upon player actions. From there the player can decide to enter either the World Map or the Patient screens. In the World Map, various cases or geographical hotspots can be selected. On selecting a location, the player is presented with patient information and can make the decision to treat the patient or return to the map if they don't feel confident about that particular case. Underneath the surface of this the game uses the mathematical model to determine a patients' profile (weight, smoker, non-smoker, geography, etc.).

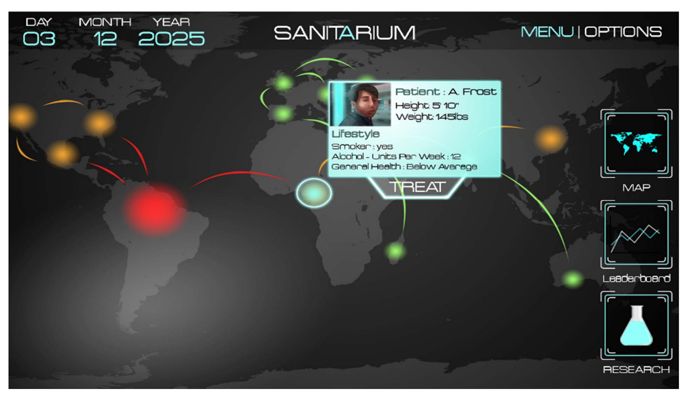


Fig. 2 In-game screenshots of the World Map screen

The patients' profile will then affect the difficulty level of the diagnosis mini-games. Once the player does select to treat the patient they move to the Patient screen. Here, they select a mini-game from a variety of options and progress to play that game. The player succeeds by correctly diagnosing then managing to cure each patient, adding to their overall score. Once the player reaches the end of their career cycle (i.e. retirement), the score is totaled and displayed upon online leaderboards. Since the game features a rogue-like element, each play through will present different challenges and patients. This in turn will mean that players can play through the game multiple times, each time encountering different scenarios and outcomes. Most importantly, difficulty will be altered based upon the player's success each time around.

C. *Mini-Games*

In the game world, the procedures of patient diagnosis and treatment are represented and implemented as mini-games, each representing a diagnosis test and a corresponding treatment. The outcome of playing each determines the condition of the patient and the possible result of their treatment. All of the mini-games factor into the final score for each patient, and whether the player is able to cure them. The design aim here was to provide a quasi-realistic game environment for the player and to provide some gameplay variety. The prototype developed an initial three mini-games, with the intention of expanding this if the project went forward. For the purposes of the prototype development, the mini-game design represents a simplified three-step process in diagnosis and treatment (the first step being a smear test, the second, an X-ray, and third the prescription of some form of drug treatment). The intention was to provide these initial games at a high level of polish and to add more mini-games if time and resources allowed.

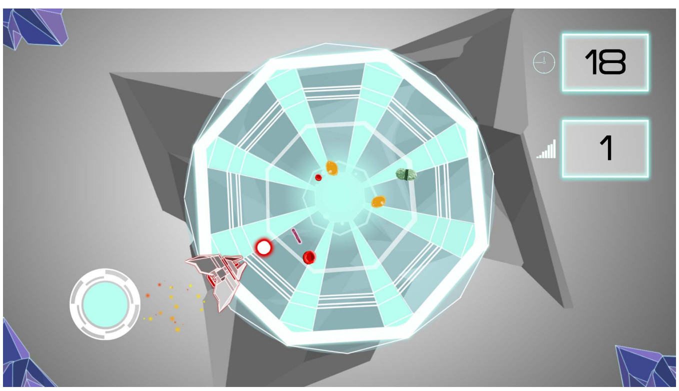


Fig. 3. Mini-Game 1: Petri Dish

The first mini-game developed represented a Petri dish, where the player is required to identify and kill TB bacilli (See Fig. 3). The initial total number of bacteria is determined based on a player's performance in the Petri dish mini-game: a poorer performance in the game results in the patient being infected with TB for longer before detection, and hence a higher initial number of bacteria.

In the mini-game the player is presented with a petri dish with various cells around the scene. They then rotate around the petri dish and count bacteria by shooting them within a specific time frame. This represents a sputum test where a doctor will take a saliva swab then count the bacteria from the swab in the petri dish. The design also identified that there are 'good bacteria' and for every good bacteria the player hits they lose one second on the timer. It builds upon a game mechanic that combined arcade classics such *Whac-A-Mole* and *Space Invaders* before evolving through play testing to the current version. [27] The player's final score is then used to determine how long it has taken to detect TB within the patient's body, and hence the initial number of bacteria.

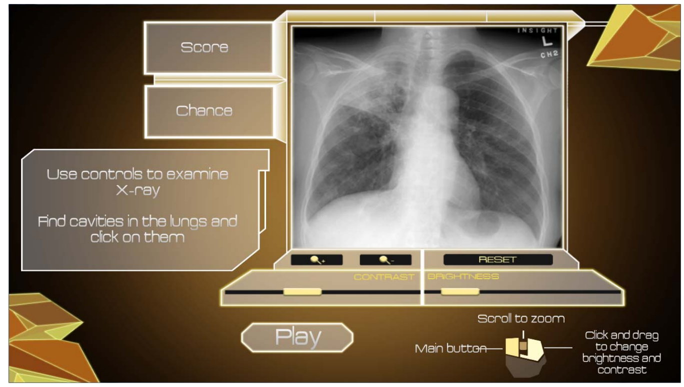


Fig. 4. Mini-Game 2: X-ray

The second diagnosis mini-game is the X-ray game, and proved challenging due to the difficulty in accurately representing real-world diagnosis. The aim of this was to get the player to identify problems within the patient's lung. They are presented with the lungs in the form of an X-ray scan (see Fig. 4). The player has a limited number of guesses within a defined time frame to spot and identify the cavities within the lung X-ray. Although simplistic, the game is fundamentally more realistic and educational than the other mini-games. Reading X-rays is challenging and requires considerable training for clinical staff. The design of the game is based on real software called OsiriX, which medical professionals use to examine X-rays. [28] This allows the player to adjust the brightness and contrast of the image to facilitate the identification of cavities. The results again feed back into the underlying mathematical model and change the patients' status based on their final score: the better the player performs within the game, the easier treatment becomes.

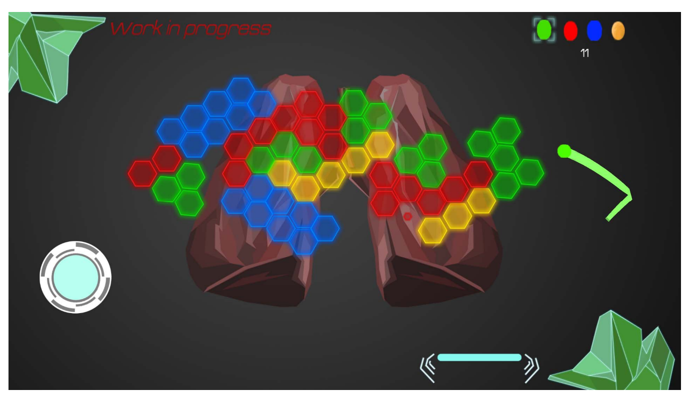


Fig. 5 Mini-Game 3: Treatment Game

The final mini-game for the prototype represented treatment (see Fig. 5). To that end a *Breakout* style game was created to mimic various drug regimens. [29] The mini-game used data from the previous two games to determine the initial layout and difficulty. The player controls a paddle and must reflect the pill (the ball) into the blocks which represent the TB bacteria. The player has one ball that can change into four different colors, each color representing one of the different drugs in the treatment. Each colored ball has a countdown timer that starts and stops when that color is selected or deselected. This mimics the time sensitive nature of the effectiveness of each drug as in the real world. The colored balls are used to destroy the corresponding colored blocks. The mini-game score is determined by how many blocks the player destroys and how many are left intact. Various iterations were designed, developed and play tested. The core game play always focused on the player keeping the ball as centered on the infected area as possible. Multiple restart and power-up opportunities are provided, and again player success has a direct impact upon the virtual patient's health. Essentially success in the mini-games increases the chances of progression in the overall game. Although not implemented for the prototype, each successful treatment results in additional time and monetary resource. Effective management of the resources and making difficult moral decisions then affect the overall game progression.

V. OUTREACH

The criteria for this project to be judged a success were multi-faceted. From the client perspective of the University of St Andrews' Infection Group, success was delivering a game about treating tuberculosis that is engaging and entertaining whilst remaining true (in regard to point of infection, diagnosis, treatment, possible cure, death or relapse) and that incorporated the mathematical model where input variables can be plugged in and modeled with appropriate results. For Microsoft and DeltaDNA it was the technical ease that their platforms can be developed for. For the Abertay students and staff the initial goal of a successful workplace simulation and a completed prototype grew more ambitious as the project developed.

A. *TB Awareness*

One of the key goals of the team was to increase awareness that the fight against TB remains constant. To that end, through contacts from the University of St Andrews, the team began building connections with organizations such as the TB Alliance. There remains much to be done in providing a fully functioning game and in developing the message about TB awareness through gaming. However the initial prototype and trailer demonstrate the potential that the game has for increasing awareness. Reflecting the commercial reality of Game Development, the team found that they were not alone in trying to achieve this and that they were not first to market, with the release of a similar game on World Tuberculosis Day, 24 March 2015. *Tuberspot* is a serious game developed by researchers at the Polytechnic University of Madrid. [30] Although different in many aspects it naturally has some cross-over, for example in *Tuberspot*, players help analyze real digitized sputum samples and it too has a vision to develop for the mobile platform. Specifically *Tuberspot* is looking to develop a microscopy system on a mobile phone to allow telediagnosis. The fact that other researchers are looking at similar game developments is healthy both for TB awareness and the use of serious games. The games remain significantly different and one of the next goals will be to explore potential collaborations.

B. *Game Competitions*

As the project progressed and was met with enthusiasm from those viewing the project and engaging with the game, suitable games contests were identified to enter the prototype into, such as Microsoft's Imagine Cup and the Serious Play Awards. The completed prototype earned the team a place in the UK Finals of the Imagine Cup. It seemed the perfect fit, as Project Sanitarium was a prototype that encompassed all three categories – Games, World Citizenship (Social Impact) and Innovation. In the end the team entered through the Games category and although they did not win, they learned new skills in regard to pitching a 'Serious Game' in an entrepreneurial setting and realizing how their message could be heard through all the noise that is produced in this age of information. It was the recognition of the social impact that gaming can have that remains one of the teams' most interesting challenges and as the development has come to a natural hiatus, they have looked to ensure that the message and recognition that TB remains one of the world's most deadly diseases remains in the forefront of their work. The prototype has since won a Gold award at the 2015 International Serious Play Awards, a fantastic achievement. [31]

VI. CONCLUSION

Project Sanitarium demonstrates the potential that serious games can deliver for an undergraduate level, without specific project funding. For the undergraduate team to take on the challenges of developing and delivering a game that addresses the fundamental aspects of diagnosing and treating tuberculosis in the modern day, whilst remaining true to an underlying mathematical model, and still provide an experience entertaining enough to secure a place in the UK finals of the Imagine Cup and win Gold in the Serious Play Awards is a success for all of the team to be proud of. The project further demonstrates that serious games research and development provide unparalleled opportunities for student learning. Taking students outside of their comfort zone in terms of games-subject matter but still providing the experience of commercial game development and working with industry and academic clients of the caliber of Microsoft and the University of St Andrews will hopefully bode well for their careers. For the clients, Microsoft has a great example of their strong technical ecosystem and a team that is roundly convinced of how easy the system can be utilized. The University of St Andrews Infection Group has a playable, functioning and professional prototype to demonstrate the potential of mathematical model. In project terms, Project Sanitarium moves forward to its next challenge, one that resonates with both academia and industry, and that is to secure funding for further development.

ACKNOWLEDGMENT

The authors would like to thank the Abertay students, Microsoft UK's Technical Evangelist Lee Stott, Chris Wright and the team at DeltaDNA, and the entire Infection Group at the University of St Andrews for all their help and support in making the project possible.

REFERENCES

[1] World Health Organisation, "Tuberculosis Fact Sheet", (no date). [online] Available: http://www.who.int/mediacentre/factsheets/fs104/en/ (Accessed March 30th 2015).

[2] M. Wang and R. Oshi, "Working Paper: Benefits and Costs of the Tuberculosis Targets for the Post-2015 Development Agenda," (2014). [online] Copenhagen Consensus Center, Available: http://www.copenhagenconsensus.com/sites/default/files/health_viewpoint_tb_-_wang.pdf (Accessed March 30th 2015).

[3] TB Alliance, "Our Mission" (no date). [online] Available: http://www.tballiance.org/about/mission.php (Accessed March 30th 2015).

[4] S.H. Gillespie SH, R. D. Gosling, and B. M. Charalambous, "A reiterative method for calculating the early bactericidal activity of antituberculosis drugs." *American Journal of Respiratory and Critical Care Medicine,* Vol. 166, No. 1 (2002) pp. 31–5.

[5] I. Honeyborne, T.D. McHugh, P. P. J. Phillips, et al. "Molecular Bacterial Load Assay, a Culture-Free Biomarker for Rapid and Accurate Quantification of Sputum Mycobacterium tuberculosis Bacillary Load during Treatment." *Journal of Clinical Microbiology* Vol. 49 (2011) pp. 3905–11.

[6] S. H. Gillespie, A. M. Crook, T. D. McHugh, *et al.* "Four-month moxifloxacin-based regimens for drug-sensitive tuberculosis," *New England Journal of Medcine*, Vol. 371 (2014) pp. 1577–87.

[7] E. Wenger, Communities of Practice: Learning, meaning, and identity. New York: Cambridge University Press. 1998.

[8] The Review on Antimicrobial Resistance. 2014. "Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations," [online] Available from: http://www.jpiamr.eu/wp-content/uploads/2014/12/AMR-Review-Paper-Tackling-a-crisis-for-the-health-and-wealth-of-nations_1-2.pdf (Accessed 30th March 2015)

[9] Abertay University. 2015. AMG Course List. [online] Available from: http://www.abertay.ac.uk/studentlife/schools/amg/courses/ (Accessed 30th March 2015)

[10] B. Boehm and R. Turner, Balancing Agility and Discipline: A Guide for the Perplexed, Boston, MA: Addison-Wesley Longman Publishing Co., Inc. 2003.

[11] NHS, "Tuberculosis (TB)," (no date). [online] Available: http://www.nhs.uk/Conditions/Tuberculosis/Pages/Introduction.aspx (Accessed March 30th 2015).

[12] S. H. Gillespie, A. M. Crook, T. D. McHugh, *et al.* "Four-month moxifloxacin-based regimens for drug-sensitive tuberculosis," *New England Journal of Medcine*, Vol. 371 (2014) pp. 1577–87.

[13] A. Singh. "Games to Combat Tuberculosis," *International Journal of Science and Research*, Vol. 3:1, January 2014, pp. 361-367

[14] Foldit: Solve Puzzles for Science, (no date). [online] Available: https://fold.it/portal/ (Accessed 30th March 2015)

[15] Team Cooper, "Fraxinus," (no date). [online] Available: http://teamcooper.co.uk/work/fraxinus/ (Accessed 30th March 2015)

[16] Guerrilla Tea, "Play to Cure," (no date). [online] Available: http://www.guerillatea.com/play-to-cure-genes-in-space/ (Accessed 30th March 2015)

[17] Dark Realm Studios, "Pandemic 2," (no date). [online] Available: http://www.darkrealmstudios.com/ (Accessed 30th March 2015)

[18] Ndemic Creations, "Plague Inc.," (no date). [online] Available: http://www.ndemiccreations.com/en/22-plague-inc (Accessed 30th March 2015)

[19] Deadendthrills, "Building the Impossible: The Art of Mirror's Edge," (no date). [online] Available: http://deadendthrills.com/building-the-impossible-mirrors-edge/ (Accessed 30th March 2015)

[20] First Person Scholar, "The Art of Papers, Please," 23 October 2013. [online] Available: http://www.firstpersonscholar.com/the-art-of-papers-please/ (Accessed 30th March 2015)

[21] Behance.net, "Deus Ex Human Revolution - User Interface," (no date). [online] Available: https://www.behance.net/gallery/2465641/Deus-Ex-Human-Revolution-User-Interface (Accessed 30th March 2015)

[22] Pushing-Pixels.org, "At the intersection of the physical and digital – interview with Michael Goldman," 24 June 2014. [online] Available: http://www.pushing-pixels.org/2014/06/24/at-the-intersection-of-the-physical-and-digital-interview-with-michael-goldman.html (Accessed 30th March 2015)

[23] V. Frei, "Interview - Star Trek Into Darkness: Roger Guyett," 27 June 2013. [online] Available: http://www.artofvfx.com/?p=4623 (Accessed 30th March 2015)

[24] C. Rosen, "Gravity' Composer Steven Price On The Film's Incredible Soundtrack" 4 October 2013. The Huffington Post. [online] Available: http://www.huffingtonpost.com/2013/10/04/gravity-steven-price_n_4039455.html (Accessed 30th March 2015)

[25] B. Child, "Interstellar's sound 'right for an experimental film', says Nolan" 17 November 2014. The Guardian. [online] Available: http://www.theguardian.com/film/2014/nov/17/interstellar-sound-christopher-nolan (Accessed 30th March 2015)

[26] IMDb, "Tara Platt," (no date). [online] Available: http://www.imdb.com/name/nm1297428/ (Accessed 30th March 2015)

[27] J. Sellers. Arcade Fever: The Fan's Guide to thE golden Age of Video Games. Philadelphia: Running Press, 2001.

[28] OsiriX, (no date). [online]. Available http://www.osirix-viewer.com/index.html (accessed 16th June 2015)

[29] S. Kent, Ultimate History of Video Games. New York: Three Rivers Press, 2001.

[30] Tuberspot, (no date). [online] Available http://tuberspot.org/en/ (Accessed 30th March 2015)

[31] 2015 Serious Play Awards, (no date). [online]. Available http://www.seriousplayconference.com/awards/ (Accessed 29th June 2015)

Back to Contents

Towards a Construction and Validation of a Serious Game Product Quality Model

Lilia García-Mundo
Institute of Technologies and Information Systems
University of Castilla-La Mancha
Ciudad Real, Spain
LiliaCarmen.Garcia@alu.uclm.es

Marcela Genero, Mario Piattini
Institute of Technologies and Information Systems
University of Castilla-La Mancha
Ciudad Real, Spain
{Marcela.Genero, Mario.Piattini}@uclm.es

***Abstract*— A Serious Game (SG) is a game for purposes other than mere entertainment. SGs are currently in widespread use and their popularity has begun to increase steadily. The number of users of these systems is also growing day-by-day, signifying that their social impact is very high; it is precisely for this reason that SG quality evaluation is of the utmost importance. The principal objective of our long-term research, initiated one year and a half ago, is therefore to define and validate a quality model adapted specifically to SGs that has been agreed on by experts and is useful in practice, in order to allow SG designers and developers to ensure, evaluate and improve the quality of the SGs they build from the early stages of its development. The main goal of this paper is to present the construction processes of a preliminary version of the SG Quality Model (QSGame-Model) adapted from the ISO/IEC 25010 standard, which is a product quality model that can be applied to any kind of SG and will be validated in the near future.**

Keywords*— *Serious Game; Quality Model; ISO/IEC 25010

I. Introduction

Serious Games (SGs) are games for purposes other than mere entertainment, which means that they have a serious purpose not only as regards education but also training, advertising or simulation [1]. SGs use have many benefits, some of the main ones being the following: 1) SGs allow learners to experience situations that it would otherwise be impossible to come across in real life owing to aspects related to costs, resources, time, security, etc. [1]; 2) There is evidence that SGs support the acquisition of knowledge, that they are more effective than traditional instructional methods as regards training cognitive skills, and that they are promising as regards their use in the learning of fine-grid motor skills that require excellent hand-eye coordination [2]; 3) SGs enable the employment potential of staff to be enhanced, while simultaneously improving their technical capabilities. They also make it possible to catch up with and keep abreast of technological development, in addition to fostering local development and strengthening regional cohesion [3].

SGs are a fast-emerging area of opportunity, in addition to being a rapidly-growing market [4]. In 2012, worldwide revenues for game-based learning (a type of SG) alone amounted to 1.5 billion dollars. With a global growth rate of 8 % a year, it is forecasted that by 2017 worldwide revenue will reach 2.3 billion dollars [5].

SGs are vitally important at present, as they may be a means to achieve relevant goals from both a personal and an institutional point of view. They may be used in fields as diverse as defense, education, scientific exploration, health care, emergency management, city planning, engineering, religion, and politics. What is more, the number of users of these systems is growing each day, signifying that their social impact is very high. It is for this reason that the quality of SGs is so critical; they are not just another variety of software (in which it is already assumed that quality is important), but may have a major impact on many areas of society and on a huge amount of users; it is therefore our duty as researchers and computer professionals to ensure their quality. The principal objective of our long-term research, initiated one year and a half ago, is therefore to define and validate a quality model adapted specifically to SGs that has been agreed on by experts and is useful in practice, in order to allow SG designers and developers to evaluate and improve the quality of the SGs they build. We planned to achieve this main goal by following the research plan shown in Fig. 1.

As a first Step of this research plan, we conducted a systematic mapping study (SMS) in order to discover the current state-of-the-art on research into SG Quality [6] following the guidelines proposed by [7]. An SMS is a well-known research methodology designed to provide a wide

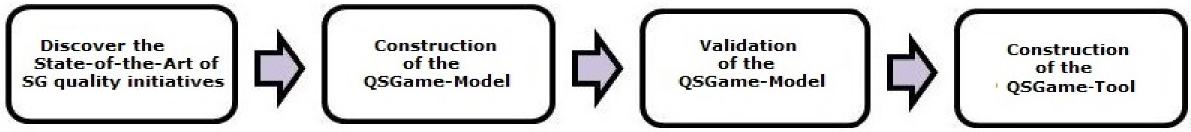


Fig. 1. Research plan

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

overview of a particular research topic in a systematic and rigorous manner [7] [8] [9]. The main characteristics of the SMS performed are shown in TABLE I.

The results of this SMS indicate that researchers are principally concerned with demonstrating or confirming whether an SG has accomplished the purpose for which it was created, along with being capable of providing enjoyment and entertainment in the game. However, we were unable to find an agreed on quality model that considers all the characteristics, sub-characteristics, attributes and measures that are applicable to any kind of SG.

Based on the SMS findings, in a second Step, we proposed an SG product Quality Model (QSGame-Model) by adapting and extending the current standard on software product quality the ISO/IEC 25010 standard [10]. QSGame-Model considers attributes that can facilitate the player's flow experience and can contribute to achieving the serious purpose of the game. In addition, we believe these attributes have an influence when the game is in use, thus allowing a better player experience [11]. The QSGame-Model is a model for product quality, and we believe that it will be useful for designing SGs that satisfy desirable requirements that contribute to achieving better SG usability and playability. The QSGame-Model was built using the methodology proposed by Franch and Carvallo [12] which drives the construction of domain-specific quality models. The main goal of the third step is the empirical validation of the QSGame-Model with the objective of attaining two main goals: 1) to contrast the opinions about the model's relevance and comprehension of the model provided by experts, which will allow us to refine it and 2) to validate the usefulness of the model in practice. Finally, in the last step, once the quality model has been validated it will be automated by means of the QSGame-tool.

The main goal of the current paper is to present the construction process of a preliminary version of the QSGame-Model that will be validated in the near future.

The remainder of this document is organized as follows. Section II briefly introduces the ISO/IEC 25010 standard and presents the related work. Section III thoroughly describes the steps followed to build the QSGame-Model. An outline of the validation that we plan to perform in the near future is described in Section IV. Finally, our main conclusions and ideas for future work will be presented in Section V.

TABLE I. MAIN CHARACTERISTICS OF THE SMS ON SG QUALITY

SMS characteristic	Description
Objective	To collect the existing literature on SG quality in a systematic and rigorous manner.
Inclusion criteria	• Papers that present any kind of research as regards evaluating/assessing/measuring/ testing the quality of SGs. • Journals, conferences and workshop papers. • Papers written in English. • Papers published until April 2013 (inclusive).
Search sources	SCOPUS, Science@Direct, Wiley InterScience, IEEE Digital Library, ACM Digital Library, Springer
Number of papers analyzed	112

II. RELATED WORK

The quality model we are proposing in this work has been developed by taking the ISO/IEC 25010 [10] standard as a starting point and adapting it to the domain of SGs. This section presents an overview of the ISO/IEC 25010 [10] standard, and research works in which quality models based on the extension and adaptation of standards to specific domains have been proposed.

Software quality is defined as the "degree to which a software product satisfies stated and implied needs when used under specified conditions" [10]. Evaluating the quality of software has been tackled for several years and is done using the standards related to software quality as a basis. The reasons for using a standard are: 1) to avoid conflicts and inconsistencies as regards the vocabulary used; and 2) to start with a widely accepted set of quality characteristics that has been agreed on by consensus. The main purpose of the ISO/IEC 25010 standard is to specify and assess the quality of software products [10]. The ISO/IEC 25010 quality model is composed of the Product Quality Model and the Quality in Use Model. Both models are useful as regards evaluating the quality of a software product which is determined from two perspectives: Product Quality, by measuring internal properties (such as software specification, architectural design, among others), or by measuring external properties (typically by measuring the behavior of the code when executed); and Quality in Use by measuring quality in use properties (when the product is in real or simulated use). The Product Quality Model classifies product quality properties in eight characteristics and thirty one quality sub-characteristics (Fig. 2); while the Quality in Use Model describes five quality and nine sub-characteristics, [10]. The main idea behind this standard is the definition of a quality model and its use as a framework for software evaluation. A quality model is defined by means of the general characteristics of software, which are further refined into sub-characteristics, which are in turn decomposed into attributes, thus yielding a multilevel hierarchy. The bottom of the hierarchy contains measurable software attributes whose values are computed by using a particular measure. These measures must be thoroughly and precisely defined within the quality model. The output of a quality evaluation of a software product is therefore a set of measurement values that have to be interpreted in order to provide developers and designers with feedback regarding the quality of the software products.

The ISO/IEC 25010 quality standard is generic; the characteristics defined by it are relevant to all software products and are not exclusively related to code or executable software, but also to analysis and design artefacts. Owing to this generic nature, the standard fixes some high-level quality concepts which can be tailored to specific domains [12].

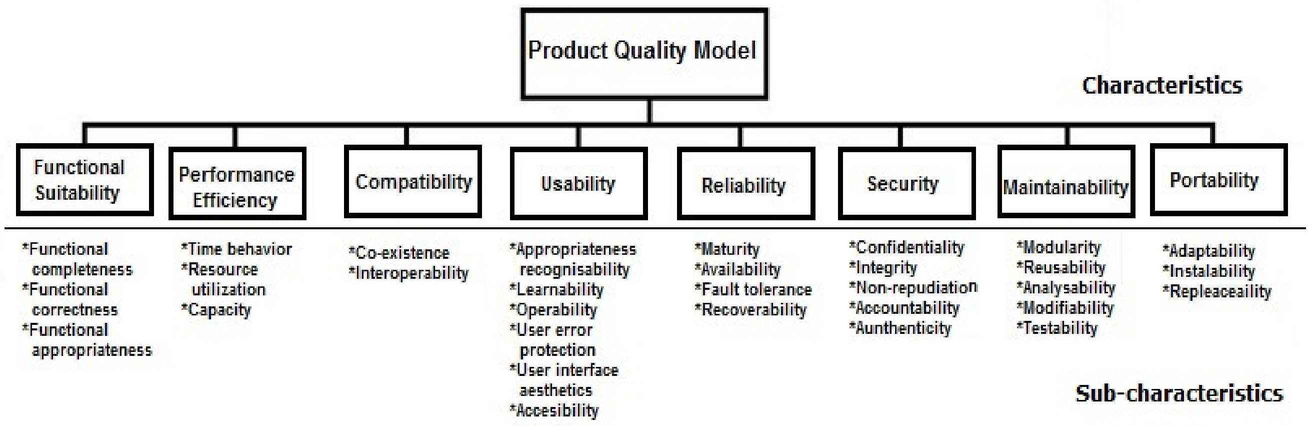


Fig. 2. Product Quality Model of ISO/IEC 25010

Several models have been developed by taking this standard as a starting point and adapting it to specific domains. Some examples are:

1. Radulovic, García-Castro, and Gómez-Pérez [13] presented a product quality model for semantic technologies called SemQuaRE. This quality model is based on the SQuaRE standard and describes a set of quality characteristics specific to semantic technologies and the quality measures that can be used for their measurement.
2. Herrera, Moraga, Caballero, and Calero [14] adapted a quality in use model to assess the level of quality in use of Web portals. The quality model used the ISO/IEC 25010 standard as a base. Some of the sub-characteristics defined in the standard were adapted to the contexts of web portals and other sub-characteristics were not included because they could be considered as not being sufficiently relevant for web portal usage.
3. González, Montero, Padilla, and Gutiérrez [15] presented a quality in use model with which to evaluate the player experience during the use of video games. This quality model is based on the ISO/IEC 9126-4 standard [16] and characterizes the player's experience through the definition of factors and measures.
4. Carvallo, Franch, and Quer [17] built a product quality model based on the ISO/IEC 9126-1 [18] quality standard and adapted it to a particular COTS domain: that of mail servers. This model can be used in two different contexts related to COTS procurement: precise formulation of quality requirements and description of COTS.

These examples have been taken into account for the construction of QSGame-Model, which is the product quality model proposed for SG Quality presented in this paper.

III. PROPOSAL OF THE QSGAME-MODEL

We have defined a product quality model that is specific to SGs, and which is denominated as QSGame-Model. Fig. 3 shows QSGame-Model. The squares with a white background in this figure represent the characteristics or sub-characteristics, to which no changes were made, i.e., they are the same as those in the standard. Squares with a dark background represent the sub-characteristics to which modifications were made. These modifications are related to the attributes and measures that were added to QSGames-Model. Top-down approaches are frequently used to adapt quality models to a specific domain [14][15][17][19]. These approaches start from general characteristics to concrete measures. To adapt our model we consider it appropriate to use the top-down approach methodology proposed by Frach and Carvallo [12]. The methodology proposes six sequential steps but they can also be intertwined or repeated. As a preliminary Step in the methodology, the domain of interest has to be examined and described. We then followed the six Steps shown below in order to identify the quality characteristics, sub-characteristics and attributes, according to the model:

1. Determining quality sub-characteristics. This step deals with the decomposition of characteristics into the sub-characteristics that appear in the standard (add new sub-characteristics specific to the domain, refine the definition of existing ones, or even eliminate some). Taking the top of the hierarchy from the ISO/IEC 25010 standard as a starting point, we adopted the sub-characteristics and made minor modifications to them. The principal changes consisted of adapting all the definitions of the sub-characteristics to the context of SGs; in some cases, we adapted some sub-characteristics and made only minimal changes to them, such as the sub-characteristic learnability. In the ISO/IEC 25010 standard [10], learnability is defined as " the degree to which a product or system can be used by specified users to achieve specified goals of learning to use the product or system with effectiveness, efficiency, freedom from risk and satisfaction in a specified context of use", while our

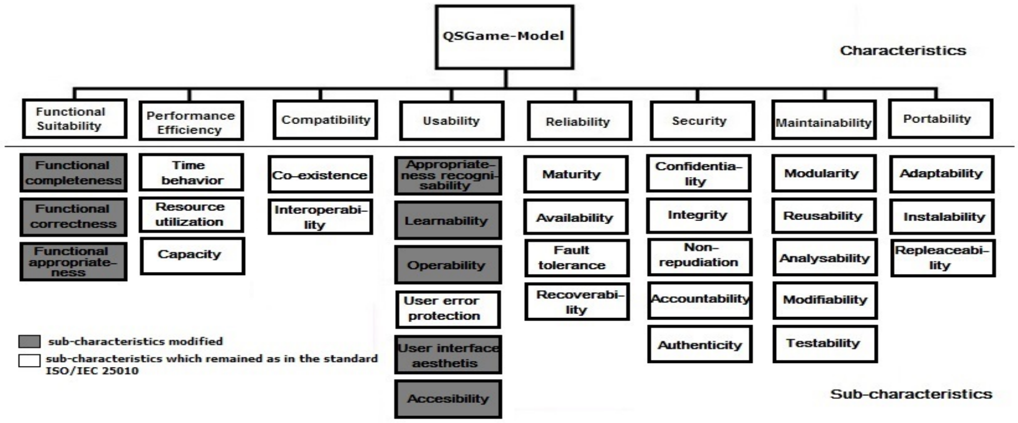


Fig. 3. QSGame-Model

definition is "the degree to which an SG and its mechanics are easy to understand and master by SG users, with effectiveness, efficiency, efficiency and satisfaction in an SG context". We redefined this sub-characteristic in order to consider the learning of the game mechanics as a part of the SG learning process.

2. Defining a hierarchy of sub-characteristics. It is possible to decompose sub-characteristics with regard to certain factors, thus yielding a hierarchy. We adopted all the sub-characteristics that are defined on the second level of the hierarchy as defined in the ISO/IEC 25010 standard.

3. Decomposing sub-characteristics into attributes. Since quality sub-characteristics provide a comprehensible abstract view of the quality model, it is necessary to decompose these abstract concepts into more concrete ones (attributes). An attribute keeps track of a particular observable feature of the software in the domain. Based on the elements of the SGs considered by researchers according to the state-of-the-art on SG quality [6], our main contribution is the addition of specific SG attributes which are not considered in the standard product quality model [10]. We believe that these added attributes could facilitate the player's flow experience [20] while simultaneously contributing to achieving the serious purpose of the game. We also believe that added product attributes, have an influence when the game is in use, thus allowing e a better player experience [11] (that is, quality in use). We considered the addition of several attributes to the Functional Suitability sub-characteristics: Functional completeness, Functional correctness and Functional appropriateness (See Fig. 4); and to the Usability sub-characteristics: Appropriateness recognizability, Learnability, Operability, User interface aesthetics, and Accessibility. We added attributes only in these two characteristics because we believe the elements that facilitate the flow experience (objectives and clear rules, feedback, balance between challenges and skills and concentration) are directly related to them. Furthermore, the results of the SMS on SG quality [6], showed that these two characteristics of the product quality model were those most frequently addressed by researchers. We did not make any significant changes to the rest of the sub-characteristics from the product quality model standard: Performance efficiency, Compatibility, Reliability, Security, Maintainability, and Portability sub-characteristics. As an example, TABLE II describes the attributes added to the Functional appropriateness sub-characteristic, which is a sub-characteristic of the Functional Suitability characteristic.

4. Decomposing derived attributes into basic ones. When an attribute cannot be directly measurable (derived attribute), it should be decomposed until it is completely expressed in terms of basic attributes. Some attributes of the QSGame-Model have been divided into basic attributes. One example of this is the attribute "likeability appearance" of the user interface, which refers to the qualities that a likeable user interface should have. These qualities are, among others, balance, symmetry, regularity, etc.

5. Stating relationships between quality characteristics, sub-characteristics and attributes. If a more exhaustive model is to be obtained, it is important to explicitly state the relationships between quality characteristics,

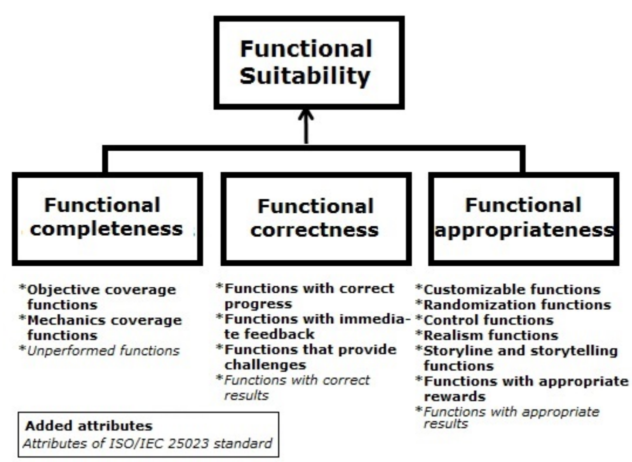


Fig. 4. Functional suitability: sub-characteristics and attributes

TABLE II. ATTRIBUTES ADDED TO FUNCTIONAL APPROPRIATENESS SUB-CHARACTERISTIC.

Functional appropriateness Degree to which the SG functions facilitate the accomplishment of specified SG tasks and objectives.	
Attribute name	**Attribute description**
Customizable functions	Functions that allow the user to establish particular preferences, e.g. identified as a character of a particular sex or given looks, etc.?
Randomization functions	Functions that have actions or tasks produced in a random order.
Control functions	Functions that allow the user to choose different controls in order to operate the game.
Realism functions	Functions that allow virtual worlds to be closer to the real world.
Storyline and storytelling functions	Functions that have tasks and activities that are relevant to the storyline and storytelling being conducted throughout the game.
Functions with appropriate rewards	Functions that offer appropriate rewards, in relation to the challenge achieved.

sub-characteristics and attributes (which we shall refer to as quality entities). Given two quality entities A and B, the types of relationships are:

- Collaboration (+). Growing A implies growing B.
- Damage (-). Growing A implies decreasing B.
- Dependency (D). Some values of A require that B must fulfill certain conditions.

TABLE III (whose design was taken from [17]) shows an excerpt of the relationships established. Attributes in rows contribute to attributes in columns. For example, as more functions are implemented by considering *Mechanics coverage* (i.e., each objective established provides a challenge, and each challenge achieved offers a reward), there will be more functions with clear game mechanics. On the other hand, functions that allow players to set challenges (*Functions that provide challenges*), depend on whether the functions in the game have established *clear game mechanics*.

6. Determining measures for attributes.

We defined measures for each of the attributes that are directly measurable. The measurement description and the measurement function are included for each measure. TABLE IV shows only those measures that were added to the Functional appropriateness sub-characteristic for reasons of space. For each of the measures added or modified, we defined their name, a description of them, and the sources from which these measures were obtained. These sources are the primary studies of the SMS on SG quality [6] and/or [11]. For example, for the functional correctness sub-characteristic, an added measure is functional progress. This measure was added because in SGs it is important to show the player accurate results of his progress during the game. For reasons of space constraints, we cannot include the complete model in this paper, but the complete names of the measures (added or modified), the description of the measures and the sources from which they were obtained is accessible at http://alarcos.esi.uclm.es/SeriousGamesProductQualityModel/.

Most attributes of the QSGame-Model proposed can be evaluated with integer or floating values of a particular unit (for example the ratio of functions which provide a correct score to signal progress or advancement in the game); and very few of the attributes can be represented using logical single values such as yes or no (for example functional customization measure if the SG allows the user to establish particular preferences).

It is important to mention that when using a quality model it is not always necessary to evaluate all the quality characteristics of the model. What we evaluate will depend, among other things, on the type of application, the application needs that are to be covered, the quality requirements, etc. This means that we shall propose the quality models mentioned, consisting of characteristics, sub-characteristics, attributes and measures. When using them, the evaluators must determine what the most relevant quality characteristics requiring assessment are. The evaluators could use the quality

TABLE III. EXCERPT OF ATTRIBUTES RELATIONSHIPS

CHARACTERISTIC			**Usability**	
Functional Suitability	**SUB-CHARACTERISTICS**		**Learnability**	**Operability**
		Attributes	Functions with clear game mechanics	Functions with real controls
	Completeness	Objective coverage functions	+	
		Mechanics coverage functions	+	
	Correctness	Functions that provide challenges	D	
	Appropriateness	Realism functions		+

TABLE IV. MEASURES ADDED TO THE ATTRIBUTES OF THE FUNCTIONAL APPROPRIATENESS SUB-CHARACTERISTIC.

Functional appropriateness Degree to which the SG functions facilitate the accomplishment of specified SG tasks and objectives.		
Measure name	**Measure description**	**Measurement function**
functional customization	Does the SG allow the user to establish particular preferences, e.g. identified as a character of a particular sex or given looks, etc.?	X = yes or not If the game allows the user to establish particular preferences, X value will be "1", otherwise X value will be "0" X [0 or 1]; 1 is better
functional randomization	What proportion of the implemented functions has actions or tasks produced in random order?	X = A / B A = number of implemented functions which have actions or tasks produced in random order B = number of total functions of SG X [0,1]; the closer to 1 the better
functional control	Does the game functions allow the user to choose different controls for operate the game?	X = yes or no If the game allows the user to choose different controls for operate the game, X value will be "1", otherwise X value will be "0" X [0 or 1]; 1 is better
functional realism	What amount of the implemented functions allows that virtual world to be closest to the real world?	X = A / B A = number of implemented functions which allows the virtual world to be closest to the real world B = number of total functions of SG X [0,1]; the closer to 1 the better
functional storyline and storytelling	What amount of the implemented functions has tasks and activities relevant to the storyline and storytelling being conducted throughout the game?	X = A / B A = number of implemented functions which has tasks and activities relevant to the storyline and storytelling B = number of total functions of SG X [0,1]; the closer to 1 the better
appropriateness of reward	What proportion of the implemented functions offered appropriate rewards, in relation to the challenge achieved?	X = A / B A = number of the implemented functions which offer appropriate rewards B = number of total functions of SG X [0,1]; the closer to 1 the better

evaluation as a basis to provide recommendations on how to improve SG quality.

IV. VALIDATION OF THE QSGAME-MODEL

In order to validate the quality model proposed in this research, we decided to a use a research method called "Technical Action Research (TAR)". The TAR method, proposes starting from an artifact, in our case the QSGame-Model, and then seeking ways in which to iteratively and incrementally validate the artifact in different cycles, until the artifact is used in real environments to solve real problems in industry [21].

The objective of the validation following TAR is: To obtain a specific model for SGs that has been agreed on by experts and that is simultaneously useful in practice. QSGame-Model must be useful for designers and developers from two perspectives: 1) in supporting the specification of an SG requirement, thus ensuring a higher quality SG from the early stages of its development; and 2) in evaluating the quality of an SG once it has been built, thus helping improve it if necessary.

The following subsections provide a brief introduction to the validations we plan to carry out.

A. Refinement of the QSGame-Model

As mentioned previously, standard quality models such as that proposed in [10] are relevant to all software products. If the models are to be useful for specific domains they must therefore be adapted in order to address the specific characteristics of each software product. When a quality model is adapted, it will be necessary to refine the model with the objective of verifying that all the characteristics that have been included in it are valid in the specific context to which it was adapted.

In order to refine the model we planned to perform a survey carried out with a group of expert SG developers; and an example of the application of the model with a group of subjects.

1) Survey with a group of expert SG developers

We plan to carry out a survey which will be administered to a group of experts related to the SG development area. The process of applying the survey will be carried out using the principles of survey research [22]. The main objective of the survey is: to obtain feedback from SG developers regarding the understanding and the importance for them of each of the attributes of QSGame-Model. The target population of the survey will be SG and video game designers and/or developers.

We are now gathering representative contacts in order to distribute the survey.

The survey is structured in two blocks of questions:

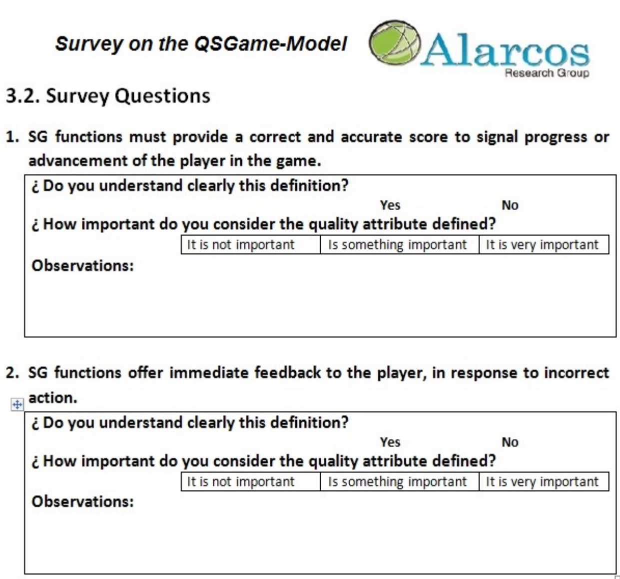
- Background and expertise: The aim of this block of 5 closed questions is to help us contextualize the responses to the survey. The information in these questions is related to the respondent's gender and experience as a software developer in general, as a video game developer and as an SG developer.
- Assessment of the QSGame-Model attributes: This block contains 36 closed questions in which we ask the respondents how important each of the QSGame-Model attributes in the proposed model is for them; and an open question to ask about any other aspects that are important to the SG developer which have not been considered in the survey. For each closed question (36) the respondent is asked (if deemed necessary) to add a comment about the attribute defined. The response format of the closed questions has been standardized in order to reduce the time needed to fill in the survey. An excerpt of the survey is shown in Fig. 5.

The survey was designed using a cross-sectional survey design and is being applied by means of self-administered questionnaires via the Internet [22]. In Genero, Cruz-Lemus, and Piattini [21], it is mentioned that the design of cross-sectional surveys and the use of self-administered questionnaires is that which is most frequently used in the field of software engineering.

The construction process of the survey was based on similar surveys that have been used to refine quality models proposed in domains other than SG, such as those employed in [23] and [24]. Before placing the survey on-line we plan to conduct a pilot study with experts who are professors on a "Video game development course" at our Computer Science School. The main goal of this pilot study is to know the opinion of experts about the quality model and to reduce any ambiguities.

2) An example of application of the model by a group of subjects

Once the model has been refined on the basis of the information collected after executing the survey, we plan to provide a group of subjects with the QSGame-Model and ask them to apply it in order to evaluate the quality of an SG. This exercise will allow us to verify the accuracy of the refined QSGame-Model, i.e., whether all subjects obtain the same value for the measures and whether the subjects comprehend all the attributes and measures.


Survey on the QSGame-Model

3.2. Survey Questions

1. SG functions must provide a correct and accurate score to signal progress or advancement of the player in the game.

¿ Do you understand clearly this definition? Yes No

¿ How important do you consider the quality attribute defined?

It is not important	Is something important	It is very important

Observations:

2. SG functions offer immediate feedback to the player, in response to incorrect action.

¿ Do you understand clearly this definition? Yes No

¿ How important do you consider the quality attribute defined?

It is not important	Is something important	It is very important

Observations:

Fig. 5. Excerpt of survey on QSGame-Model

B. *Validation of the usefulness of QSGame-Model in practice*

We plan to carry out an experiment and replications of QSGame-Model in order to gather empirical evidence about its usefulness, i.e., to assess whether the existence of a quality model permits the construction of better quality SGs. This will be done by following the guidelines provided in [25]. The main characteristics of the experiment are:

1. To ask the (carefully selected) group of subjects to develop the same SG after being provided with the requirements specification of the game.
2. The subjects will be divided into two groups; Group1 will have to develop the SG by using the quality model, while Group 2 will not have the quality model.
3. Finally the experimenters will evaluate the quality of the SG developed using the QSGame-Model defined, and statistically test whether the existence or otherwise of the quality model really allows better quality SGs to be built.

V. Conclusions and Future Work

The results of SMS on SG quality carried out previously [6] revealed that there is no general quality model that can be applied to any kind of SG. This motivated us to focus our long-term research on the construction and validation of a product quality model that is specific to SGs.

The main contributions of this paper are:

- The QSGame-Model, which is a preliminary proposal for a product quality model for SGs, adapted from the ISO/IEC 25010 standard [10] which defines all the characteristics, sub-characteristics, attributes and measures related to SG quality. The QSGame-Model was built using the top-down methodology proposed by Franch and Carvallo [12], which has been used to build quality models that are applicable in others domains such as those of [17] and [19], among others. Based on the findings obtained by means of an SMS [6], this model considers attributes that can facilitate the player's flow experience and can contribute to achieving the serious goal of the game. In addition, we believe these attributes have an influence when the game is in use, thus permitting a better player experience [11] (that is, quality in use).

We are aware that QSGame-Model contains attributes that will not be relevant for every game, and the selection of the most appropriate along with how to score them is pending as future work.

- An outline of how we plan to validate the model in the near future, through a combination of the TAR method and experiments, with the goal of obtaining an agreed product quality Model for SGs that will be useful as regards allowing SG developers to ensure, evaluate and improve the quality of the SGs they build from the early stages of the development.

ACKNOWLEDGMENT

This work has been funded by the following projects: GEODAS-BC (Ministerio de Economía y Competitividad and Fondo Europeo de Desarrollo Regional FEDER, TIN2012-37493-C03-01) and IMPACTUM (Consejería de Educación, Ciencia y Cultura de la Junta de Comunidades de Castilla La Mancha, y Fondo Europeo de Desarrollo Regional FEDER, PEII11-0330-4414).

We would also like to thank the Instituto Tecnológico de Ciudad Victoria and PRODEP for granting the first author the scholarship that made it possible to complete the research work presented in this paper.

REFERENCES

[1] T. Susi, M. Johannesson, and P. Backlund, “Serious Games – An Overview ,” Technical Report HS- IKI -TR-07-001, School of Humanities and Informatics University of Skövde Sweden, 2007.

[2] P. Wouters, E. Van der Spek, and H. Van Oostendorp, Current Practices in Serious Game Research: A Review from a Learning Outcomes Perspective, Games-Based Learning Advancements for Multi-Sensory Human Computer Interfaces: Techniques and Effective Practices. IGI Global, Hershey PA USA, 2009, pp. 232-250.

[3] LUDUS, “How can one Benefit from Serious Games,”, 2010. Retrieved November 19, 2013, from: http://www.ludus-project.eu/sgbenefits.html.

[4] David R. Michael and Sandra L. Chen, Serious Games: Games That Educate, Train, and Inform. Muska & Lipman/Premier, 2006.

[5] Ambient Insight Research, “The 2012-2017 Worldwide Game-based Learning and Simulation-Based Markets,”, 2013. Retrieved February 10th, 2014, from: http://www.ambientinsight.com/Resources/Documents/AmbientInsight_SeriousPlay2013_WW_GameBasedLearning_Market.pdf.

[6] J. A. Vargas, L. Garcia-Mundo, M. Genero, M. Piattini, “A Systematic Mapping Study on Serious Game Quality,” 18th International Conference on Evaluation and Assessment in Software Engineering (EASE´14), May 2014.

[7] B. Kitchenham and S. Charters, “Guidelines for performing Systematic Literature Reviews in Software Engineering,” Technical Report EBSE-2007-01, Software Engineering Group of Keele University Durham UK, 2007.

[8] B. Kitchenham, D. Budgen and P. Brereton, “Using mapping studies as the basis for further research – A participant-observer case study,” Information and Software Technology, vol. 53 (6), pp. 638-651, 2011. doi: http://dx.doi.org/10.1016/j.infsof.2010.12.011.

[9] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, “Systematic Mapping Studies in Software Engineering,”, Conference on Evaluation and Assessment in Software Engineering EASE’08, June 2008.

[10] ISO/IEC, ISO/IEC IS 25010: Systems and Software Engineering - Systems and Software Quality Requirements and Evaluation (SQuaRE) - System and Software Quality Models, ISO (International Organization for Standarization), 2011.

[11] J. L. González, “Jugabilidad. Caracterización de la experiencia del jugador en videojuegos,” PhD Thesis, Universidad de Granada, 2010.

[12] X. Franch, and J.P. Carvallo, "Using quality models in software package selection," IEEE Software, vol. 20(1), pp. 34-41, 2003.

[13] F. Radulovic, R. García-Castro, and A. Gómez-Pérez, “SemQuaRE—An extension of the SQuaRE quality model for the evaluation of semantic technologies” Computer Standards & Interfaces, vol. 38, pp. 101-112, 2015.

[14] M. Herrera, M. A. Moraga, I. Caballero, and C. Calero, "Quality in use model for web portals (QiUWeP)," 10th International Conference on Web Engineering, ICWE 2010, pp. 91-101, July 2010.

[15] J. L. González, F. Montero, N. Padilla, and F. L. Gutiérrez, “Playability as Extension of Quality in Use in Video Games,” 2nd International Workshop on the Interplay between Usability Evaluation and Software Development (I-USED), August 2009.

[16] ISO/IEC 9126-4, ISO/IEC-9126-4 Software Engineering – Product Quality – Quality in use metrics, 2004.

[17] J. Carvallo, X. Franch, and C. Quer, “Defining a quality model for mail servers,” Proceedings of the 2nd International Conference on COTS-based Software Systems, (ICCBSS 2003), pp. 51-61, January 2003.

[18] ISO/IEC 9126-1, Software Engineering—Product Quality—Part 1: Quality Model, 2001.

[19] P. Botella, X. Burgués, J. P. Carvallo, X. Franch, J. A. Pastor, and C. Quer, “Towards a quality model for the selection of ERP systems,” Component-based software quality, vol. 2693, pp. 225-245, 2003..

[20] M. Csikszentmihalyi, Flow: The psychology of optimal experience. New York: Harper Perennial, 1991.

[21] M. Genero, J.A. Cruz-Lemus, and M. Piattini, Métodos de Investigación en Ingeniería del Software. RaMa. Shull, Forrest, 2014.

[22] B. A. Kitchenham, and S. L. Pfleeger, Personal opinion surveys, Guide to Advanced Empirical Software Engineering. Springer London, 2008, pp. 63-92.

[23] M. A. Caro, Modelo de Calidad de Datos para Portales Web. PhD Thesis, Universidad de Castilla-La Mancha, 2007.

[24] M. F. Bertoa, Medidas para la Usabilidad de Componentes. PhD Thesis, Universidad de Málaga, 2006.

[25] C. Wohlin, P. Runeson, M. Höst, M.C. Ohlsson, B. Regnell, and A. Wesslén, Experimentation in Software Engineering: An Introduction. Norwell, MA: Kluwer Academic Publishers, 2000.

Back to Contents

Augmenting Technology Trees: Automation and Tool Support

Teemu J. Heinimäki and Tapio Elomaa
Department of Mathematics
Tampere University of Technology
P.O. Box 553, FI-33101 Tampere, Finland
Email: {teemu.heinimaki, tapio.elomaa}@tut.fi

***Abstract*—In this work we tackle the rigidity of predefined technology trees common in many digital games. Technology trees have been a well-understood concept in the industry for approximately two decades. Unfortunately, so far they have mainly been considered as fixed structures. Some attempts of adding ostensible temporal variability to them have been made before by the means of simulating the effect. Our approach, on the other hand, aims to create a stepping stone towards true runtime technology generation for (both serious and leisure) games in order to improve them. We present potentially useful ideas, constructs, and methods to achieve this goal. Initial observations on our test implementation are also presented.**

I. Introduction

Technology trees, also known as *tech trees* (or *skill trees*, *talent trees*, etc.) are the predominant way of providing development alternatives within digital games. These structures guide and limit possible player-made choices and the overall flow of the game. Moreover, they provide short-term goals for players. Choices offered to the players are known to be essential in meaningful play [1].

In computer role-playing games, usually the protagonist (player character) is improving its skills, piling up money, attaining arms, gaining in knowledge, or developing in some other way as the game progresses. In strategy games, development typically concerns a whole tribe, a nation, a species, or even a larger entity, like a galactic alliance. Concerning the apparent distinction between these development graph structure ("tree") types in different genres, it is basically just a question about whose operation and development does a particular tree guide. Therefore we find it justified to use the term technology tree for all such development-guiding structures.

A tech tree explicates what are the requirements for advancing in technical capability, engineering abilities, scientific knowledge, military power, financial wealth, or such. The tree models and keeps track on the overall development status using distinct *technologies* (also known as *techs*). Technologies can be developed individually, and the tree dictates their relationships and requirements for starting the tech development processes.

The downside of usual tech trees is their rigidity. They are preprogrammed without any ability to adapt to the player and without any variation in time or between different game instances. This paper advances and promotes the possibilities to rectify this issue.

We introduce our approach called CATFAT (Constructing Automatically Technologies For Augmenting Technology trees) to make tech trees more flexible and more interesting by automated content generation. The approach consists of specific kinds of classifications and representations of technologies and a tech tree augmentation method that actually uses them.

The main contributions of this paper are (i) presenting the general principles of CATFAT and (ii) demonstrating the practical usefulness of it, when implemented in suitable software tools. Secondary contributions are (iii) pointing out the common rigidity problem of technology trees and the difficulties related to tackling it and (iv) raising general awareness of the importance of technology trees as a topic for further studies.

We continue by a bit of more background and by explaining our motivation in Section II. A short review to the related work can be found in Section III, after which we introduce our approach in Section IV. Section V illuminates tool support implemented so far, and Section VI describes initial experiments that have been conducted. Section VII discusses different aspects of possibilities to generate technologies based on our observations. Finally, conclusions are drawn in Section VIII.

II. Background, Motivation, and the Problem Domain

Technology trees are essentially data-containing acyclic digraphs that are used to offer players possibilities to develop technologies.[1] Tech trees are used to guide and limit possible choices and the overall flow of the game and to offer goals for players. (The term technology tree can also be used to mean a representation of such a structure, e.g., in a pictorial form.) The idea of a tech tree is demonstrated in Fig. 1. Initially, there are two technologies, *A* and *B*, available to be developed. Technology *D* can be developed after obtaining *B*, but prior to developing *C*, both *A* and *B* are required.[2]

In practice, technology trees are designed manually, and once they have been implemented for a game, they do not change, barring possible corrections and modifications introduced in patches and expansions to the game. Even if a game

[1] We use the word technology quite liberally here. In this paper, the word is used synonymously with any node belonging to any technology tree.

[2] This paper only considers tech trees with such conjunctive prerequisites. In general, however, obtaining all the prerequisite technologies is not always necessary in order to develop a tech, but only some of them may suffice.

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

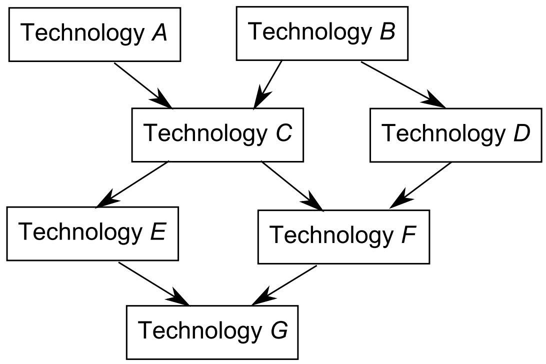


Fig. 1. An illustration of the basic idea of a tech tree.

uses a data-driven approach and defines trees in, say, text files instead of hard-coding them – thus making it easier to alter the game – tech trees are still essentially fixed between the modifications. Making tech trees easy to modify is a good practice, but does not remove the fundamental problem.

Rigidity is a problem because of its negative effects to the replayability and interestingness of a game [2]. Moreover, a fixed tech tree also implies having only a limited amount of technologies to be offered during a game. At some point, there simply may not be any technologies left to be attained.

Unfortunately, the solutions applied in practice are far from perfect and tend to restrict games severely. For instance, the number of turns in a game may be limited in order to force the tree to suffice. It is also possible to offer some generic "future tech" to be researched after all the "real technologies" have been developed, as is done in the Civilization games (*Sid Meier's Civilization* (MicroProse 1991) and its numerous sequels). In Fig. 1, *G* could be such a technology, researchable over and over again. However, this approach may lead to uninteresting gameplay after a critical point, because the central element of making the player to choose the next technology to be developed among several candidates is effectively removed.

There are, of course, lots of cases, in which there is no need to expand the tree to make it larger at the runtime. For instance, in strongly story-driven games the length of a game is determined by the narration, and trees of finite – and often quite modest – sizes can be built to fit the arc of the story. Therefore we acknowledge that good games of certain types can be implemented even with rigid trees.

Nevertheless, especially strategy games featuring large traditional tech trees and so-called sandbox games with considerable freedom to explore the world and proceed in the possible story would benefit from expandability of tech trees. The degree of actually using dynamic tech tree content generation might naturally vary – all the technologies could be generated during runtime, a fixed tree could be augmented only when necessary, or anything in between. In any case, it would be generally appealing to be able to generate techs on runtime and keep discovering specific, meaningful technologies even after exhausting the predefined tree. This is the motivation behind this work.

If one desires to produce better games, also capability to create temporal variance in tech trees is an important aspect to consider. Even if the tree size is fixed, technologies do not have to be. If the set of available technologies changes for each game, one cannot keep playing the game similarly, making the same choices time after time.[3] The whole tech tree does not necessarily have to be replaced; temporal variability in varying degrees can be offered by changing, adding, or removing available individual technologies in a "basic tech tree". Such tech tree–modifying operations can also be performed during the actual gameplay – not only when initializing the game. Combining the idea of temporal variability and runtime content generation means introducing some (pseudo)randomness into the generating process.

Often technology tree structures and technologies in them are fully known to players so that they can easily plan their technological advancement beforehand. If, on the other hand, one is interested in dynamically "inventing" possible areas of research always based on the pre-existing knowledge (the state of technological development) during the game, real runtime generation – instead of revealing preconstructed content – is an intriguing way to proceed, since this way the virtual process of development corresponds to the real one.

A central problem in automated technology generation is natural language: it is rather simple to generate technologies with somewhat arbitrary properties within some guidelines, but typical use of technology trees requires the technologies to have descriptive names (and possibly even more detailed descriptions of some length). Another major issue is also about representation; often technologies have corresponding visible instances, like moving units or buildings, in the game. This paper mainly focuses on the former problem, as it applies also to the abstract technologies without those "physical" instance requirements. Sometimes also graphical images are used in visual technology representations. Images might be generated, for instance, by combining elemental images corresponding to "basic natures" of techs and their property bonuses etc., but this is out of the scope of this paper.

III. Related Work

Procedural content generation (PCG) for digital games in its various forms has been recently studied rigorously. A recent survey [3] gives a decent overview to the field. Our approach can be seen as a PCG method specializing to generate technology trees, and it touches upon at least two PCG subfields identified in the survey: those concerning game systems and game design.

The approach presented in this paper fits in the category of domain specific language approaches to PCG, because the content generation in this case is based on a model defined by a language developed specifically for this purpose. Using the taxonomy of methods in PCG [3], our approach can be seen as one using generative grammars. The basic idea of creating content by "expanding" smaller description is typical to PCG algorithms [4].

[3]Making each game different by changing tech trees for each game instance would add considerably on the replayability value. Naturally this would, however, also annoy some players, so probably the feature should be optional.

Until recently, technology trees have been overlooked as interesting subjects for academic studies. As far as we know, this is the first paper about generating content to technology trees. The fixed structure and the assumption of technology trees or technological paths existing as such – without changes due to outside influence – have been criticized (see, e.g., reference [5]). However, practical remedial suggestions are hard to find. Typically tech trees are treated as totally deterministic constructs. Sometimes determinism is even included in the definition [6].

As the basis for our implementation of the approach put forward in this paper we use a prior software tool, Tech Tree Tool (TTT) [7], [8], created for, e.g., building, manipulating, and analyzing tech trees. The tool facilitates the process of defining the required constructs by offering an easy-to-use graphical user interface (GUI) and generating code automatically.

One possible way to create temporal variability to tech trees is to leave some technologies out of a large predefined technology tree (more or less randomly, often preserving essential "skeleton" technologies). This approach has actually been used in real commercial games – for instance in *Sword of the Stars* (Kerberos Productions 2006). An alternative way – the approach discussed in this paper – proceeds from the opposite direction and tries to generate new technologies (with some random factors involved) and thus expand a skeletal tree.

The problem of managing and directing a game based on user-made decisions has been discussed in previous papers, and frameworks have been presented (see, e.g., reference [9]). Our approach does not alter the game plot as such, but belongs in the wider scope of improving player experience by automation. We also acknowledge the possible benefits to gaming experience obtainable by making runtime changes to the game played based on the observations about the player (see, e.g., reference [10]). For instance, there are numerous ways to adjust the difficulty of a digital game dynamically [11]. At the moment, the suggested method to generate technologies in its simplest form does not explicitly include such functionality. However, it is easy to add considering the modeling and algorithmic choices made.

There have been several publications on generating game mechanics automatically based on user-generated content. For instance, a recent conference paper [12] discusses this topic and introduces a game-generating system. Our approach has a more strict focus on technology trees, but in some sense it also generates mechanics by defining a set of technologies with effects to be chosen at a given time via the tech generation. Our method is also designed to operate during runtime, though it can alternatively be used to help with planning tech trees beforehand when designing and creating a game.

Numerous studies focusing on existing games, their effects, and game design have been carried out during recent years. As nontechnical by nature they are only weakly related. In addition to (merely) entertainment applications, so-called serious games have gained considerable attention. They have various application areas [13], [14]; for example, games can be used in education [15] or in rehabilitation [14]. Our examples in this paper hint towards a leisure setting, but the CATFAT approach can as well be applied in "serious", non-leisure contexts using tech trees.

IV. The Proposed Approach

The main idea of the CATFAT approach is to generate new technologies based on the existing ones in order to make the game more interesting and replayable (see Subsection IV-C). Technologies can be generated either dynamically when running the game, when initializing a game instance, or beforehand when implementing the game.

Properties and categories are attached to technologies, since they guide the technology generation process (see Subsection IV-A) and thus are crucial. To be able to communicate the method clearly, the technologies are also divided into several types (see Subsection IV-B).

A. Categories, Properties, and Non-tech Modifiers

Developing a technology usually manifests itself in a game by enabling the corresponding player to use (or create), e.g., physical abilities, spells, units or other such (possibly abstract) concepts. Let us call these, in the lack of a better term, *products* of their respective enabling technologies. To clarify: technologies are goals for (technological, scientific, equipment-related, or some other kind of) development in a game, and products are the manifestations of the developed technologies.[4]

The CATFAT method attaches *properties* to technologies and their products. These properties reflect, e.g., physical measures, abilities, roles, and effect types in the game world to which the technologies or the products are directly related.[5] The properties are basically name-value pairs having typically numerical or Boolean values, via which they inflict functional effects in the game. There may also be faction-wide or global properties affecting the game more widely.

For instance, each faction in a game could have the property *loyalty* featuring, say, a value between zero and one indicating the general tendency of units to defect to a competitive party in situations offering possibilities to do so. In addition, each unit might have its personal loyalty modificator property, and the actual probability for a defection of a single unit occurring in a given game situation could be calculated based on these two property values.

The relation between properties and technologies is simple: developing technologies modifies property values, and property values define and modify constraints and effects for, e.g., technologies. For instance, achieving some popular form of government (a tech) in a game might boost the global *loyalty* value, and developing an institute for funding research work might increase the personal loyalties of scientist units.

In addition to using properties, another key element in our approach is categorization of techs and products based on an

[4]It is not always necessary to make a distinction, and sometimes the term technology is used to cover also products, but we find this a bit confusing habit and use two distinct words in this paper for clarity.

[5]In computer role-playing games, it is typical to use numerical statistics in order to track character development [16]. Here we are applying the idea more widely: measurable properties can be attached to other entities also, not only characters.

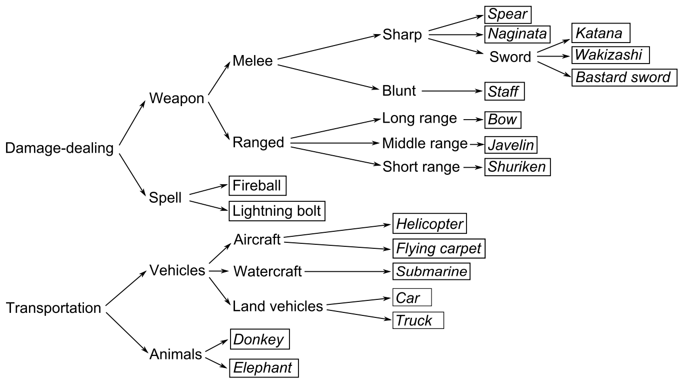


Fig. 2. Example type hierarchies of technologies and products.

Tech Categories:

GENERAL (G)
DAMAGE DEALING (D)
DAMAGE COUNTERING/EVADING (C)
DAMAGE ABSORBING (A)
POWER SOURCE (P)
ENGINES (E)
SENSORS (S)
COMPUTING (M)
HEALING (H)

Non-tech Modifier Sets:

Ammo: {Basic Ammo}
Fuel: {Energy Crystals}
Food: {Worms}
Hull_material: {Basic Alloy}

AC Abbreviations:

C: Cost
S: Size
W: Weight
P: Power Consumption
T: Thrust
N: Range
M: Communication
V: Effect Velocity
U: Accuracy
(K|I)A:
(Kinetic|Incendiary) Attack
(K|I)D:
(Kinetic|Incendiary) Defence

BV Abbreviations:

As corresponding ACs, but prefixed with "B"

Fig. 3. Categories, modifier abbreviations, and NTMs of the example game.

analysis of their different semantic subclasses. Categories built on is-a relationships are often sufficient, but other relations can also be used. In any case, each product and a large portion of technologies should be classified into – possibly several – categories.[6] The categories may form hierarchies. An example hierarchy is shown in Fig. 2: the leaves are technologies and products, and the intermediate nodes are categories. A leaf belongs to all the categories along the path from the root to it. (Separating techs from products here is not a matter of importance, and cannot even be done without extra knowledge of the related game.) Each leaf also forms a category consisting only of the leaf.

When using CATFAT, the current technological state of the gameplay is represented for each player by a set of global properties and properties of technologies and products. It is often convenient to divide properties into two categories: *base values* (BVs) defining initial values and *adjustment coefficients* (ACs) for typical value modifications.

For example, the value for a global AC called “building cost” could be 1.0 at the beginning of the game, and by developing, say, *logistics* the coefficient could be reduced during the game. This reduction would then be reflected – via a simple multiplication with the corresponding base value – to the cost of a specific building operation. For a transport unit, on the other hand, e.g., speed, weight, and cargo capacity could be crucial features to be modeled this way, and fuel type could serve as an example of a non-numeric property that still could be represented with a BV property.

On the other hand, fuel is also a good example of an external non-tech resource affecting features and usability of some technologies. More generally, a technology may be affected in a game by arbitrary non-tech modifiers (NTMs) – variables able to take values in their respective value sets that can also change during a game. For instance, the NTM *fuel* could have a value from the set {gasoline, hydrogen, gunpowder}. Commonly applicable relations between techs and NTMs are “uses” and “has” relations, but other types are also possible.

[6]Strictly speaking at least all the base techs (introduced in Section IV-B, as are also the other technology types mentioned here) should have categories that could be inherited by generated techs. If a modifier tech has categories, they can be inherited too, and any technology that should be affected by property changes for a category should be categorized into it. For simplicity, one can categorize all of the technologies; a generic category can be used, if needed.

B. Technology Types

All the technologies present in a predefined tree (*core tree*) are *core technologies*. That tree will be augmented procedurally with *generated technologies*. Core techs – and their relationships – can be seen as a compact (compressed) presentation of all potential larger tech trees that can be generated – and in the optimal situation, wanted.

Base technologies are selected representants of different fundamental categories among the core technologies. These chosen ones have an important role in the tech generation process: they offer initial tech names and features to be modified in the process.

Modifier technologies are able to modify properties; developing them may affect technologies and products or the gameplay (player-wise or globally), typically by changing AC values. Also BVs can be manipulated. The modifications may also introduce possible technologies to be generated into the tree – *generable techs*. They can manifest themselves as actual generated techs. Generable techs are formed using base technologies or generated techs as the basis and modifying their names and properties according to the rules defined by the modifier tech. Categorizing technologies and their products simplifies forming these rules.

The technology type groups presented here may overlap. Core trees consist of base techs, modifier techs, and possibly other techs not used by CATFAT (all thus being core techs). It is also possible for a single technology to be both, a base tech and a modifier tech simultaneously.

C. Technology Generation

To demonstrate and explain the actual technology generation process, we take an imaginary space strategy game as an example and let building spacecrafts be an essential part of it. Consider a tech tree used to design spaceship parts. The technology categories used are listed in the box of the upper left corner of Fig. 3. In this simple example, there is no need for hierarchies. The box on the right-hand side presents the AC and BV property abbreviations used. All the technologies and products do not necessarily need every one of these properties. The starting NTM sets are listed in the box of the lower left corner of Fig. 3.

The core tree consists of the technologies depicted in Fig. 4 using rectangles and ellipses with solid edges and

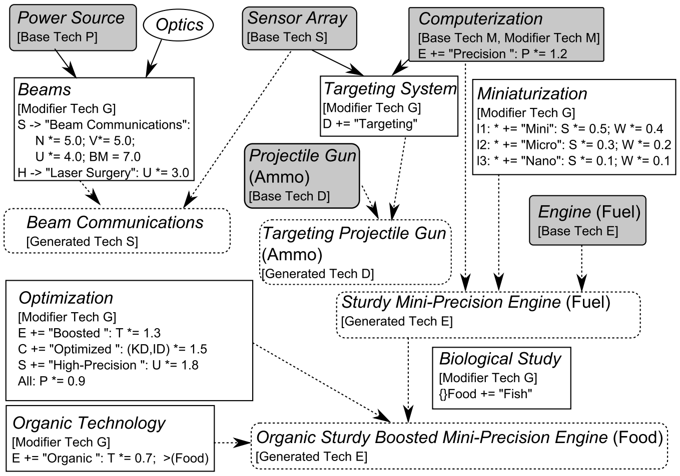


Fig. 4. Example core tree (techs with solid edges and solid arrows) and some generable techs (technologies with dotted edges).

the solid arrows defining the prerequisite relationships. The grey shapes represent base technologies, and the corresponding categories are indicated in brackets. In order to develop a tech with incoming arrows, all the corresponding techs at the start points of the arrows (tails) must first be obtained. Modifier technologies are drawn using rectangles with non-rounded corners. They typically have rules to modify existing tech names and properties category-wise. For instance, *Beams* has a rule for dealing with a technology of the category SENSORS. In the beginning there is only one such a tech available, namely the base tech *Sensor Array*. Based on the rule, the generable technology *Beam Communications* can be produced.

More generally, a base technology or a generable technology (*modified tech*) and a modifier technology with rules compatible with the categories of the modified tech can produce a new generable tech. This is done by creating a copy of the modified tech and then adjusting the copy according to the rules of the modifier technology.[7] In other words, the properties of the produced generable tech are initially set similar to those of the modified tech, and then modified by the rules. These modifications can be, e.g., multiplications, additions, or setting specific property values.

The name of a tech is also determined by the rules generating it. In the case of modifier tech *Beams* and category SENSORS, the name of the base tech is just replaced by a new one. In contrast, all the other modifier techs, except *Biological Study*, concatenate modifying strings with the names of the modified techs in order to form the names for the corresponding generable techs.

There may be several concatenations along the way, and typically the order of modifying strings is important. This must be taken into account when designing the rules. A modifying string of a rule may have, for instance, a numerical value attached indicating its need to occur immediate before – or in the case of postfixes, after – the basic name. The modifier strings can then be ordered based on these values. Other arranging rules can also be applied, if necessary. In Fig. 4, only prefix concatenation is demonstrated without any order data.

By default, the generable techs inherit categories of their parents and can be modified again by different modifier techs. The categories can, however, be controlled by the rules of generation. Either all the parents or only the ones currently not used as modifier techs (as is the case in Fig. 4) may contribute to the categories of their respective children.

Biological Study is here an example of a modifier tech that adds a NTM into a NTM set. Additional sets can also be defined, and existing NTMs and NTM sets can be removed.

Miniaturization demonstrates that a tech can consist of several levels to be developed. This is a convention adopted to make tech tree presentations more compact; instead of defining three different modifier techs in the tree, only one is defined with three levels of development with their corresponding effects. Initially, the actual game should only generate the technology corresponding to the first level. The tech corresponding to the second level ought to be presented only after the first-level tech has been developed, and so forth.

Typically developing technologies in a game costs resources like time, energy, or gold. Defining the prices to be set for generated technologies is basically a challenging task, especially, if there are several resource types in use. If there is only one type to consider, however, the difficulty can be overcome by combining the use of an *evaluator function* and a *balance corrector function*. Evaluator functions have to be implemented case-by-case. The idea is to make such functions evaluate the values of technologies under scrutiny based on their properties (effects in the game) and use these estimates in order to determine the prices for developing the techs. Also, the value estimates of, e.g., the children of the techs can be used in the process.

The purpose of the balance corrector function is to check through the current tree and make necessary modifications to the costs – and possibly also to the effects – in order to prevent generating too powerful paths and keeping the tree in balance in the sense that basically advanced techs should be more expensive than less advanced ones, and different techs of, e.g., the same historical era should have total development costs (including all the prerequisites) of the same magnitude. In the case of several resources to consider, the same principles apply, but additional metadata is needed to keep the assigned costs semantically meaningful.

The basic approach can be improved by using modifier entities that act similarly to modifier technologies, but are not included in the tree as technologies and cannot be developed as conventional techs. Predefined game events or random events could lead to applying modifying rules of these entities. For instance, a player could complete a quest on behalf of an industrial company in a game. As a reward, limited products of that company – specialized items with company-specific characteristics – might be made available to the player as technologies. Negotiating deals with, e.g., nations or species could lead to somewhat similar results. There could also be reverse effects limiting the availability or removing certain techs, launched by other events. Let us call game events with modifier tech–like capabilities *modifier events*.

[7]For simplicity, here we restrict ourselves to the cases, in which each generable tech is based on one modified tech, but more complex schemes can also be allowed.

The general procedure of applying CATFAT in a game with dynamic tech generation based on the actual technological situation is as follows:

```
initially:
  C ← ∅

whenever tech t is developed:
  if t is a modifier tech:
    select the rules of the lowest level of the tech as \
    the active ones and the rest as inactive ones.
    if t has more than one level:
      generate a next level tech, containing the rest of \
      the levels, as a child of t
    for each active rule r of t:
      if r is a general rule not bound to categories:
        apply r
      else:
        for each nonmodifier tech m ∈ r.category:
          generate a tech candidate, c, based on m \
          using rule r
          C ← C ∪ c
  else:
    for each developed modifier tech m:
      for each active rule r of m:
        if t ∈ r.category:
          generate a tech candidate, c, based on t using \
          rule r
          C ← C ∪ c
    for each modifier event e with active generation rules:
      for each active generation rule r of e:
        if t ∈ r.category:
          generate a tech candidate, c, based on t using \
          rule r
          C ← C ∪ c
  Select technologies from C and add each of them as \
  a child of its corresponding modified tech. (All the \
  candidates can be added, or only some of them, \
  based on, e.g., the currently selected focus area of \
  the research work in the game.)
  Remove the selected techs from C.
  run evaluator function
  run balance corrector function

whenever modifier event e takes place:
   for each active rule r of e:
      if r is a general rule not bound to categories:
        apply r
      else:
        for each nonmodifier tech m ∈ r.category:
          generate a tech candidate, c, based on m \
          using rule r
          C ← C ∪ c
   Select technologies from C and add each of them as \
   a child of its corresponding modified tech.
   run evaluator function
   run balance corrector function
```

This basic procedure can be modified as seen fit, but the basic idea is to add technologies to the pool of possible techs as the player progresses in a tree and then augment the tree suitably from that pool. If there are effects reducing possibilities to develop technologies, this must be taken into account, and the pool must be updated correspondingly.

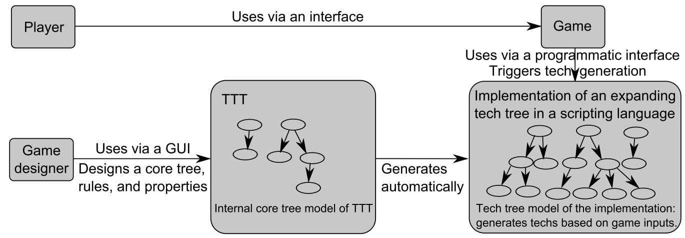


Fig. 5. Relations of different actors and components.

As a part of this method, it is futile to try to define exactly, how the selection of technologies to be added among the candidates should be done, because technology trees and games differ a lot from each other. Therefore, the selection method should be designed or chosen case-by-case.

The selection can be implemented either as an internal tech tree functionality or by the main program using the tree. Our suggestion is to let the trees take care of the whole generation process including the selection, because this way trees can be swapped with each other easily. Unnecessary dependencies between tech trees and their users should be avoided.

V. Tool Support

As a merely abstract approach CATFAT is useless. With suitable software tool support, however, it can be turned to a practical method. As prior work, we have implemented TTT to facilitate the design process, evaluate tech trees, manipulate their properties, and generate code automatically. In order to experiment with CATFAT, TTT was extended to support rule-based technology generation as described in this paper.

So far, TTT only supports disjoint categories, but it is trivial to add support for hierarchies by, e.g., defining and storing hierarchy graphs similarly to tech trees and then defining, which hierarchies a technology tree should use. Categories of technologies can be easily changed via a graphical user interface and default properties for technologies can be set. Naturally, properties of a particular tech can also be easily modified. Generation rules must, for now, be expressed using a syntax resembling the one used in Fig. 4. In the future, graphical tools and mouse gesture shortcuts can be added for even easier rule definition.

TTT can generate a functional technology tree implementation in Lua language based on the design defined via its GUI. Currently, such automatically generated tech tree implementations augment themselves by all the possible additional technologies every time a tech is developed. So far, there is no support for modifier events, but that is a straightforward feature to add, if desired.

Fig. 5 clarifies the role of TTT in the process; a game designer uses it to define the core tree and to create the necessary rules for modifier techs. TTT then generates a functional tech tree implementation in Lua. It is a rather autonomous entity having its internal data models, bookkeeping, and functions for performing necessary operations. It also offers a simple interface, via which the actual game can use the tree and ask the tree to perform different operations concerning itself. Based on input received via this interface and the internal state of the tree, more techs are generated as necessary and the data model is updated.

Moreover, TTT serves as a test environment for tech tree designers; the designer can select technologies to be virtually developed in the design view, and the tool generates techs and adds them to the visual tech tree presentation. The properties of the generated techs can be observed and modifications to the rules can be made accordingly. This speeds up the initial testing, because the tech tree can be “run” outside of the context and constraints of the actual game. Within the tool, there are also different tech tree analysis and measurement functionalities available for evaluating the tree properties.

VI. Experiments

To test the usability of our approach, we created a small core tree – corresponding to the one illustrated in Fig. 4 – for an imaginary game focusing on improving spacecrafts by upgrading, modifying and inventing different components for them. After modeling the core tree using TTT, we simulated proceeding in the tree and let the system generate additional content accordingly.

In Fig. 6, a partial screenshot of the tech manipulating area of TTT user interface is shown. A user can select technologies to be developed using a pointing device (a mouse), and the tool generates more technologies to the tree accordingly. The properties of selected technologies can be seen and manipulated in a separate property window.

We also let TTT to create a functional tech tree implementation in Lua based on the designed core tree, and run several sequences of technology selections via the tech tree–using interface (from a command line) corresponding to the sequences simulated inside TTT. In order to compare these results with the ones obtained by simulations within TTT, the processes were kept deterministic. We could verify that the Lua script generated by TTT was able to create and alter technologies and properties similarly to TTT in its internal simulations.[8] A main game could use the technology properties and general tech tree functionality via the Lua interface, and the tech tree could be expanded at the runtime by the Lua scripts generated by TTT without any effort by the main program.

These are, of course, only initial experiments, and more testing is needed. Because of this fact and the limited space in the paper, explaining the experimental setup in more detail is omitted as negligible; the proper evaluation of the approach deserves its own paper anyway. However, based on the initial tests, it strongly seems that the method is viable and the tool support facilitates creating expandable tech trees considerably. Gathering experiences of using CATFAT and TTT also in the context of real game development is important, but at this point we have to leave this as future work as well.

[8] At the moment, the internal technology development simulation functionality of TTT is implemented in C++, so it made sense to compare the results with the ones given by the Lua implementation generated by TTT. In future, it might be reasonable to make TTT automatically generate a Lua implementation of the core tree under scrutiny for performing any in-tool simulations instead of using its own internal implementation of the algorithm.

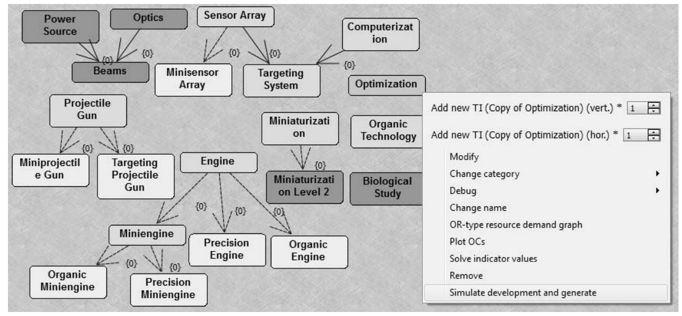


Fig. 6. A partial screen capture of the TTT GUI. Development of *Targeting System* was just simulated, and TTT generated *Targeting Projectile Gun*.

VII. Discussion

The set of technologies that can be generated with the approach presented in this paper is still limited, since additional modifier techs are not generated in the process, and using a single modifier tech several times “in a row” is generally a bad idea. It is also hard to generally guarantee the high quality of each technology that is generated. The need to use natural language restricts possibilities to generate technologies automatically, but with languages like English the approach is still feasible. The quality of the generated content depends on the modeling choices made and the definitions given. Thus, rigorous planning improves the quality.

Because the method is only capable to create content derivable from the given technologies and rules, it may look like only a way to represent data in a compressed form, and in a sense, this point of view is totally valid. However, we argue that we are dealing with automated content generation: all the generation methods need some initialization and use some rules to construct “new” content. A person defining the core tree and the rules may not be able to notice all the implications, and a computer can, based on the provided data, form unforeseen technologies.

Basically, the approach is all about changing the focus from first developing lots of detailed content and then pruning, modifying and unifying it, to designing a compact rule set capable to define the possible achievements. Despite its limitations, our approach has its benefits. In contrast to the more common method to create tech tree variations by removing techs from a large pre-existing technology tree structure, growing a tech tree based on few properties, techs, and rules makes it possible to easily adjust general mechanics and idea manifestations globally. The process is also meant to produce a unified logic throughout the tree, which can be a cumbersome task to carry out with traditional means. Although the core tree has to be built with consideration, applying CATFAT may still also result in savings of time and work.

There are, of course, tech tree types, for which it is hard to generate additional technologies by simply concatenating additional strings to tech names and letting techs inherit most of their properties. In such cases, lots of effort must be put into defining non-generalizable rules for generating single techs. This means returning to the traditional way to define tech trees.

Nevertheless, if the technologies are generated as the game progresses, one cannot – at least in theory and without prior experience or with randomization used in generation – see the possibilities offered in the distant future. Traditionally, tech trees allow one to "build toward a goal" [17]. However, even if the whole tree happens to be predetermined, we consider revealing it all at once in the beginning of the game generally a rather poor idea. Although one can optimize better with such prior knowledge, often this kind of foreseeing makes a game feel unrealistic and highlights the rigidity of the tree.

One could critisize our approach, because it does not necessarily lead to any visual or functional novelties concerning user experience; this work does not try to change technology trees as structures into anything unprecedented. A casual gamer could try a game without even noticing if it was using sophisticated mechanisms to create technologies or not. This argument, however, applies to PCG quite generally and is not specific to our approach; PCG is typically used to enhance conventional games, not to create totally new types of them. Even the casual gamer – not especially interested in tech trees – might, however, notice the difference, if the game was played again, and the tree was formed differently because of some randomization or learning scheme used. If the user actions were taken into account in the dynamic tech tree creation, the overall user experience might be improved from the baseline case of a rigid tech tree, even if the user could not appreciate the mechanisms in work.

In this paper we have presented examples of simple multiplicative property modification only, and so far that is the only modification type supported by our implementation besides setting values explicitly. Such modifications seem to be viable and useful, but naturally, other ways to control properties may be added in order to further improve the approach.

In addition to generating techs during runtime, CATFAT can be used to facilitate traditional tech tree design process: only a small amount of work is required to generate a tech tree suggestion within the desired thematic and modeling constraints. The tech tree for the final product can be prepared based on such generated suggestions. For this kind of use it might be beneficial to produce relatively large and modifiable "generic" core trees for different game genres and settings. With standardized tech tree implementation formats, offering such products might even be convertible into a business model.

VIII. Conclusions

We have presented CATFAT, a general approach for augmenting technology trees by creating additional technologies on runtime. The method reverses the conventional process of creating temporal variance by first defining a large tech tree and then leaving technologies out. CATFAT rather starts with a minimal skeleton tree and a set of rules, and generates techs as they are needed. Designing rules instead of single technologies facilitates keeping the tech tree consistent. The approach can also be used in a more conventional tech tree building process to generate suggestions to start the design work with.

Initial tests indicate that the method can be used in practice, and the resulting tech trees are of adequate quality. However, mostly the method is still on the idea level, and further evaluation and testing is still needed to ensure its general applicability and usefulness. The key challenges are different natures and characteristics of technology trees used within the wide array of digital games, the difficulties of natural language processing, and the visualization of tech-related entities.

As future work, we intend to continue testing and improving the CATFAT method and the TTT software. Generating complex content related to techs – such as unit types with their graphical representations – is a challenge that remains still to be tackled. Interesting results might be obtained, e.g., by using genetic methods. Adding a user behavior–observing manager to control the generation in order to customize the gameplay automatically for individual players might be worth trying. Such functionality could also be incorporated, e.g., into the evaluator and balance corrector functions used.

References

[1] K. Salen and E. Zimmerman, *Rules of Play: Game Design Fundamentals*. MIT Press Ltd, 2003.

[2] A. Solo, "Innovative tech trees in space strategy games," http://www.spacesector.com/blog/2009/07/dynamic-and-specialized-technology-research-in-space-strategy-games/, 2009, accessed March 30, 2015.

[3] M. Hendrikx, S. Meijer, J. Van Der Velden, and A. Iosup, "Procedural content generation for games: A survey," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 9, no. 1, pp. 1:1–1:22, Feb. 2013. [Online]. Available: http://doi.acm.org/10.1145/2422956.2422957

[4] J. Togelius, G. N. Yannakakis, K. O. Stanley, and C. Browne, "Search-based procedural content generation: A taxonomy and survey," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 3, no. 3, pp. 172–186, 2011.

[5] E. Watrall, "Chopping down the tech tree: Perspectives of technological linearity in god games, part one," http://www.gamasutra.com/view/feature/131570/chopping_down_the_tech_tree_.php, 2000, accessed July 19, 2012.

[6] T. Ghys, "Technology trees: Freedom and determinism in historical strategy games," *Game Studies*, vol. 12, no. 1, Sep. 2012.

[7] T. J. Heinimäki and T. Elomaa, "Facilitating technology forestry: Software tool support for creating functional technology trees," in *Proceedings of the Third International Conference on Innovative Computing Technology, INTECH 2013*, Aug. 2013, pp. 510–519.

[8] T. J. Heinimäki and T. Elomaa, "Quality measures for improving technology trees," *International Journal of Computer Games Technology*, vol. 2015, pp. Article ID 975 371, 10 pages, 2015.

[9] D. Thue and V. Bulitko, "Procedural game adaptation: Framing experience management as changing an MDP," in *Intelligent Narrative Technologies: Papers from the 2012 AIIDE Workshop (AAAI Technical Report WS-12-14)*. AAAI Press, Oct. 2012, pp. 44–50.

[10] G. N. Yannakakis and J. Hallam, "Real-time game adaptation for optimizing player satisfaction," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 1, no. 2, pp. 121–133, Jun. 2009.

[11] E. Adams, *Fundamentals of Game Design*, 2nd ed. New Riders Publishing, 2009.

[12] A. Zook and M. O. Riedl, "Automatic game design via mechanic generation," in *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, 2014.

[13] J. Alvarez and D. Djaouti, "An introduction to serious game – Definitions and concepts," in *Proceedings of the Serious Games & Simulation Workshop*, 2011, pp. 10–15.

[14] P. Rego, P. M. Moreira, and L. P. Reis, "Serious games for rehabilitation: A survey and a classification towards a taxonomy," in *the Proceedings of the 2010 5th Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE, June 2010, pp. 47:1–47:6.

[15] K. Squire and H. Jenkins, "Harnessing the power of games in education," *InSight*, vol. 3, no. 1, pp. 5–33, 2003.

[16] N. Hallford and J. Hallford, *Swords & Circuitry: A Designer's Guide to Computer Role-Playing Games*, ser. PRIMA TECH's Game Development Series. Prima Publishing, 2001.

[17] P. Tozour, "Introduction to Bayesian networks and reasoning under uncertainty," *AI Game Programming Wisdom*, vol. 1, pp. 345–357, 2002.

Enhancing User Engagement in Immersive Games through Multisensory Cues

Ernst Kruijff, Alexander Marquardt, Christina Trepkowski, Jonas Schild, André Hinkenjann
Institute of Visual Computing
Bonn-Rhein-Sieg University of Applied Sciences
Sankt Augustin, Germany

***Abstract*—In this article, we report on a user study investigating the effects of multisensory cues on triggering the emotional response in immersive games. Yet, isolating the effect of a specific sensory cue on the emotional state is a difficult feat. The performed experiment is a first of a series that aims at producing usable guidelines that can be applied to reproducing similar emotional responses, as well as the methods to measure the effects. As such, we are interested in methodologies to both design effective stimuli, and assess the quality and effect thereof. We start with identifying main challenges and the followed methodology. Thereafter, we closely analyze the study results to address some of the challenges, and identify where the potential is for improving the induced stimuli (cause) and effect, as well as the analytical methods used to pinpoint the extent of the effect.**

Keywords—virtual reality, immersion, multisensory cues, 3D user interfaces, user engagement

I. Introduction

Immersive display technology is rapidly finding its way into game environments. Graphically advanced devices such as the Oculus RIFT offer convincing graphical output to display compelling game environments. Yet, visuals are only one aspect of producing engaging user experiences. While sound for long has been an integral part of games, the interest in alternative input and output modalities is growing. Over the last decade, a multitude of non-conventional input devices have appeared that allow for alternative and interesting ways of user input [1]. In contrast, not many researchers have focused on the usage of other sensory channels other than audiovisual output. As especially theme parks show [2][3], these modalities can aid in producing compelling experiences. For long, creating and maintaining multisensory immersive installations was an expensive and difficult process [4]. Yet, with current rapid prototyping platforms such as Arduino and Phidgets, it becomes easy to connect different low-cost devices to an installation, be it a low-cost ventilator or a simple smell interface. Stimulated by these technical possibilities, increasingly multisensory installations are being created, both for research and art purposes. While many of these installations engage users, it is difficult to pinpoint which cue actually has which effect. At current, the adaptation of multisensory cues for specific situations – situations we may characterize using emotional response parameters – seems to be a trial and error process, or at least a well-kept secret.

It is here that the current article ties in, as it intends to explore the design and effect of sensory cues on the emotional state of the user. Our main goal is to study the effect of single or multiple cues on specific emotional responses. Analyzing such supposedly interrelated effects is a tough process and requires a flexible yet reliable methodology. This paper aims at developing such a methodology iteratively, through (a) identifying requirements and challenges based on theoretical research on user engagement and multisensory processing and (b) by performing an exploratory study. We will analyze the results of the initial study and perform a reflection to identify the gained knowledge and potential for improvements. These improvements cover both the generation and processing of stimuli as well as measurement and analytical methods, stating open issues that need to be addressed.

II. Terminology and Related Work

Studying the effect of multisensory cues on user engagement touches upon numerous areas. In order to shed light on the different aspects we target in the user study, it is important to both define the terminology and interrelationships, and identify the current state of the art. We can identify two main areas that affect our research: studies on user engagement and experiments with multisensory cues. Within this paper, we relate both areas to the analysis of emotion as at the level of human information processing and behavior. As such, our studies are affected by basic concepts in emotion research [5], [6] and the investigation of awe [7], [8]. Emotion has also played a role in user interface developments, in particular those that relate to affective computing [6], [9], [10]. Yet, as we will discuss in the following subsection on user engagement, most of the work has focused on detecting emotions [11], [12], less on the triggering specific emotional responses. While discussing the topics, we also briefly refer to Table 1, which states the challenges and connected methodologies – we will describe this Table in more detail in section III.

User engagement and emotion. User engagement can be described through a comprehensive range of aspects (C1 in Table 1). User engagement can be regarded as a quality of user experience that depends on the aesthetic appeal, novelty, and usability of the system, the ability of the user to attend to and become involved in the experience, and the user's overall evaluation of the experience. Engagement depends on the

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

depth of participation the user is able to achieve with respect to each experiential attribute [13]. As such, user engagement is thought to span various dimensions that relate to emotional, cognitive and behavioral processes [14], including focused attention, positive affect, aesthetics, endurability, novelty, richness and control, reputation, trust and expectation, and motivation and interests [15]. As such, there is a direct relationship between user engagements and affect or emotion [16], the aspect we mainly focus on in the reported experiment (M1 in Table 1). While people have somewhat agreed on what basic emotions can occur in a human being [5], underlying processes are not well understood. Even more so, measuring any of the aspects contributing to user engagement is not straightforward, though measurement tools exist - tools include the usage of eye tracking or the deployment of questionnaires such as the user engagement scale (UES, [13]). Still, measurements have mostly been done within such domains as web content or advertisement, not virtual environments (C4). Moreover, the reliability and validity of individual measures are questionable, and correlation between approaches is not necessarily being performed: while the analysis of user engagement is likely to be multi-factorial, creating more appropriate methods or approaches (up to standardization) would be important [15], [17] (M1).

Immersive games are increasingly appearing while stimulated by affordable head-worn displays, yet, the study of engagement in these systems still requires further research. In direct relation to interactive systems, user engagement is believed to be also a characteristic of user experience that emphasizes positive aspects of interaction, in particular the fact of being captivated (C2) by a certain technology [15]. Within the domain of virtual environments, the term captivation relates strongly to the terms immersion and presence. While quite some discussion has taken place on how to separate both terms, we adopt the definitions by Slater et al. [18]. They refer to immersion as a quantifiable description of a technology, which includes the extent to which the computer displays are extensive, surrounding, inclusive, vivid and matching. Also, they state that displays are more extensive the more sensory systems that they accommodate. The latter is particularly interesting, as it refers to multisensory stimuli provided through a system (M1). As such it relates directly to sensory richness and fidelity, and vividness, the variety and richness of the sensory information that can be generated by a VR system [19]. There are some studies on user experience, engagement and immersion [20]–[23], while also general concepts in the field of game flow relate as they positively affect user engagement [24], [25]. In immersive environments, presence and immersion has been studied widely [26], [27], while some deployed physiological measurements to study presence [28]. In direct relation, more recently some researchers have started the investigation of the various issues that contribute to fidelity, including [29]–[31], which relates also to the usage of stimuli.

Multisensory stimuli and emotion. Simulating the human body through various input and output modalities has been topic of interest in various related areas. In particular the design of alternative interfaces using sensory and control substitution is a well-known technique [32][33] and has often been applied to improve task performance. Also, the potential of the human body to receive various stimuli and perform actions has been studied to some extent, especially for games [1], [34], [35]. In particular, studies have focused on custom-made devices for the isolated usage of specific secondary cues. These cues include spatial sound [36]–[39], low-frequency sound [37], [40], [41], specific physical cues related to exertion [42], smell [43], [44] gustation [45], [46], and wind [47], [48]. While handling multisensory cues, our study is partially affected by cross-modal effects. These effects have been studied in psychology experiments on the various interplays between cues [49]–[52]. Yet, while the usage of multisensory cues in virtual reality (VR) systems has been explored to some extent at a technological level, the psycho-physiological effects of the interplay between various cues are still not well understood. At a design methodology level, few, and not widely adopted exceptions exist that look into the integration of multisensory cues, including [53], [54]. It is here we intend to focus our studies, aiming at better understanding the effect of multisensory cues on emotional responses, by improving the design and analysis of effective cues. In research the focus is moving away from traditional multimodal techniques in the direction of multi-sensory interfaces that differ at the level of human information processing. When sensory channels are substituted or combined, some implications occur: sensory channels are no longer seen as separate channels but may affect each other [16][24] through integration of sensory signals inside multimodal association areas in the brain [50]. The research on multi-sensory factors still needs to advance in order to fully understand its importance, but some cross-modal effects can already be identified, including bias, enrichment and transfer. With respect to bias, vision mostly plays an important role, since it predominantly alters other modalities [55][56]. Numerous studies have also been performed on other interactions, including [49], [52], [57]–[59]. In our first study we foremost focused on isolated cues instead of cue interplay, to limit the complexity of our analysis. It is important to note that sensory cues might have affected each other in our study. However, at this point we did not study inter-sensory effects, yet tried to isolate the effects of the separate cues on emotion: we assume that the combination of sensory cues has led to enrichment. Still, we have to verify sensory binding in upcoming studies using appropriate methods (M3).

With respect to emotion, little work has been performed on identifying the effect of multisensory cues on emotional responses (C4). This is somewhat surprising: affective computing has been around for long [10], yet little is known on how multisensory cues can trigger emotional responses. Many researchers have rather focused on detecting different kinds of emotions to change interaction there upon [11], [12], [60], [61] – to use the detected emotional state to change the content of the application is not necessarily widespread. Some have experimented with emotional states within the content of adapting educational content [62], while others have investigated biofeedback systems to change the mood of a user [63], [64](M5).

III. Methodology

In summary, our research extends existing work in multisensory stimulation and processing, immersive gaming and underlying validation methodologies. As novel approach in experience design, we propose to influence user engagement in immersive games by applying combinations of cues from different sensory systems. The main approach of creating impact on user engagement is through affecting user emotions. The key questions are how, which emotions and how strongly emotions can be triggered through what quantity and quality of multisensory cues. The identified set of challenges based on related work in the domains of user engagement and multisensory processing lead to our resulting initial methodology as proposed in Table 1.

Table 1: Main challenges and initial methodology of enhancing user engagement in immersive games through multisensory cues

Challenges (C) and methodological approaches (M)	
C1	User engagement is a multifactorial construct that can be difficult to improve and analyze.
M1	Focus on emotional response through multisensory cues, design guidelines for deployment
C2	Captivating users by increasing presence is important.
M2	Improve user presence through an increase of variety, vividness and richness of sensory cues
C3	Sensory channels (vision, audio, tactile, chemical) affect each other, yet analyzing the interplay is challenging.
M3	First isolate effects of separate cues on emotions, then explore inter-sensory effects.
C4	Measuring user engagement and emotional impact in virtual environments is difficult.
M4	Create, evaluate and verify appropriate methods for measuring user engagement and emotional impact.

Our initial experiment assesses emotional responses, trying to address the differences in arousal through specific cues, which we will report upon in this article. Emotion is thought to be one of the three main drivers of user engagement, next to mental states and interaction [14]. As key approach, our studies specifically focus on emotion as main driver for assessing user engagement as we aim at using different cues to cause different moods or responses from a user, such as surprise, fear or happiness. We assume that strong user responses will increase user engagement in immersive games.

IV. Experiment

Research questions

Within a series of experiments, we intend to look in the role of specific combinations of multisensory cues on triggering different kinds of emotional responses. Our interest is stimulated by how we can improve user engagement in computer games. In particular we want to address our *hypothesis* that different combinations of multisensory cues can enhance user engagement and presence by triggering particular emotional responses. As such, we assume that multisensory cues can provide more compelling and immersive experiences in virtual reality systems. Because of the inherent complexity associated with proving this hypothesis, the first experiment was designed as an exploratory study to be able to design more refined studies here after. Doing so, the first study explored two dimensions, namely (a) the extent to which specific cues (stimuli) can potentially induce a specific emotional response, and (b) the investigation of the necessary methods to verify the effect thereof. Hence, the study aimed at providing a better idea of what to improve on the interfaces used to induce the sensations, and to advance our measurement and analytical methods to investigate effects on emotion more closely.

Stimuli and apparatus

The system used for our experiment consisted of the following components. Visual stimuli were presented through a head-mounted display (HMD), the Oculus RIFT[tm] DK2. This HMD provides stereo graphics at a resolution of 960×1080 pixel per eye and a binocular FOV of about 100 degrees. The experiments were programmed in Unity3D[tm] and rendered at 60Hz. Participants were asked to use a Sony Dualshock® 3 gamepad to control movement through the environment. To introduce different sensory stimuli to the user, we customized three games. Each game included two distinct situations in which we triggered different multisensory stimuli (see Table 2 for a summary). These distinct situations should trigger a specific emotional response. We targeted one main emotion per situation, however, expected that other basic emotions could also play a secondary role. Also, we foremost targeted the emotions of disgust, fear, happiness, and surprise, as it was difficult to find appropriate situations in games that could introduce sadness or anger. The latter two emotions we assumed would, among others, need a longer phase of playing and storytelling [23][65], which would be difficult to achieve within the duration of the experiment.

Table 2 summarizes the three games with the six exposed situations, including the designed and perceived emotional triggers, as well as the influence of stimuli. Don't Let Go (Skydome Studios) is a game that plays with the fears of user, introducing various fear-inducing or dangerous events, including a bee swarm, large spiders, or a growling dinosaur. Except the tuning of audio files towards lower frequencies (the sound of the dinosaur), the game content was not modified. The second game, racer (Unity asset VMotor), is a traditional motor racer game that was modified to include am Indiana Jones-style rolling stone simulation and the appearance of a bad weather front. Finally, the Tuscany demo, a standard demo of the Oculus RIFT system offering the exploration of a Mediterranean scenery, was modified to "turn bad" by including a horde of zombies (Unity asset UFPS). In all three games, additional non-visual cues were coupled to the specific situations. We designed these situations inspired by personal experiences from movies and theme parks. However, ultimately the situations are designed by rules-of-thumb, as there is hardly any evidence that supports which combinations of stimuli can trigger specific emotional responses. As such, the gained understanding from this experiment will inform the design of the planned further experiments.

To generate the different non-visual sensory stimuli, we used the following devices. For audio, users were wearing

Sennheiser HD-265 headphones. Furthermore, we custom-designed a large subwoofer (100 Liters) that was used for low frequency feedback (20Hz and up). The latter was combined with a bass-shaker (Ibeam VT-200) mounted underneath the chair the user was seated upon to generate low frequency vibratory sensations in the user's body [40]. In addition, inspired by [41] we made use of a vibrotactile grid of 10 tactors placed at the back of the chair to stimulate the back of the user. The tactors were ordered in three rows of three tactors, and a single tactor at the lower center. As such, we could generate different vibrotactile patterns on the back of the user. In addition we made use of three 12V / 230mm ventilators with a throughput of 275m3/h to provide directional wind feedback in the frontal hemisphere of the user. We also deployed a simple smell-interface made from a 5V smoke generator to vaporize etheric oils that were chosen to match certain emotions. The smell interface was placed before the center ventilator to easily move the smell towards the user. The ventilators, smell interface and tactile grid were controlled within Unity3D by using an Arduino-based connection.

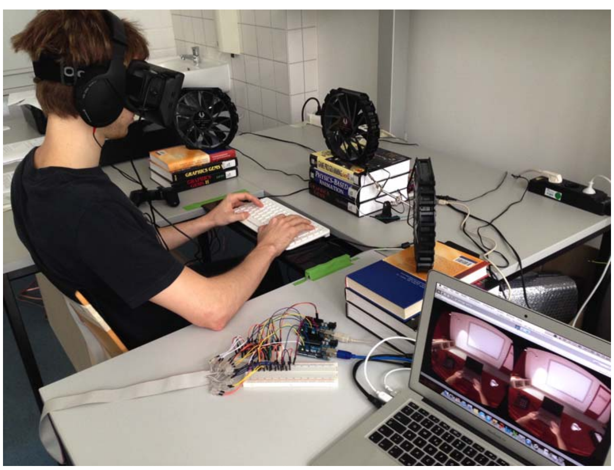


Fig. 1. Experimental setup used for initial experiment

Procedure and design

To gather subjective measures, we designed a questionnaire to address the effect of stimuli on emotional responses, and correlate these with the level of presence and involvement in the game. The questionnaire consisted of five parts. Emotional and stimuli ratings were repeated for each game situation, while presence questions were requested once per game. As a result, we had the following structure: user background (4 questions), stimuli response perception (6 x 7 questions), emotional response (6 x 6 questions), and general comfort (5 questions), and presence based on a modified IPQ rating system (Igroup Presence Questionnaire, 3 x 9 questions, [27]). The rating scale was a 5-point Likert Scale (1-5). In addition, we closely observed users, noting down clear user responses.

V. Results

Demographics and user background. 20 users participated in the experiment. All were male, age 19-33 (M= 24.65, SD 3.54). 45% of the participants played games daily, 45% weekly, and 10% monthly. The medium on which games were played were consoles or PC (80%), some played online (15%) or on a handheld (5%). The total in-game time was around 18 minutes: the racer game took about 5 minutes (SD = 1.19), the Tuscany game around 7 minutes (SD = 2.0), and the spider-game about 6 minutes (SD = 1.10). The time for filling out the questionnaires was between 15 and 20 minutes per participant. Data was analyzed using multiple methods: Spearman correlation, Friedman-tests and repeated measures ANOVA were used as appropriate. Bonferroni correction was applied as needed. To analyse the relationship between emotions and stimuli the primary and secondary emotion were determined according to the highest and second highest mean value in each situation to perform subsequent two-sided correlations with the influence of the stimuli (alpha .05). The analysis of the various situations provided the following results.

Don't Let Go. The *bee swarm* only caused mediocre emotional responses: happiness (M = 2.25, SD = 1.25) and surprise (M = 2.00, SD = 1.17) received somewhat higher ratings, while normal sound showed a significant correlation with happiness (r_s (18) = .53, p < .05). In contrast, the *spider at the back* induced the emotions of surprise (M = 2.95, SD = 1.43) and happiness (M = 2.95, SD = 1.43), while seat vibration showed a significant correlation with surprise (r_s (18) = .72, p < .001). Fear also received a noticeable rating (M = 2.15, SD = .99), and showed a significant correlation with seat vibration r_s (18) = .50, p < .05). Both ratings were mostly in line with our expectations.

Racer. In the racer game, the *rolling stone* situation produced a feeling of surprise (M = 3.15, SD = 1.04), while happiness was also rated rather well (M = 2.3, SD = 1.17). Normal sound (M = 3.4, SD = .68) and seat vibration (M = 3.45, SD = 1.23) affected these emotional responses. Yet, we did not find any significant correlations. The *bad weather* situation produced a rather weak user response. Both surprise (M = 2.3, SD = .92) and happiness (M = 2.25, SD =1.21) were rated somewhat higher. Back vibration showed a significant correlation with surprise (r_s (18) = .53, p < .05) and low-frequency sound with happiness (r_s (18) = .45, p < .05).

Tuscany. Participants were mainly happy in the situation *sea view* (M = 3.2, SD= 1.20). The happier they felt, the higher the influence of wind: wind showed a significant correlation with happiness (r_s (18) = .46, p < .05). Surprisingly, an increase in happiness also produced decreasing ratings for smell, as both were significantly correlated (r_s (18) = -.47, p < .05). In the *zombie* situation, surprise was the strongest emotion (M = 3.4, SD= 1.31). Low-frequency sound (r_s (18) = .60, p < .01), normal sound (r_s (18) = .60, p < .01), back vibration (r_s (18) = .49, p < .05) and seat vibration (r_s (18) = .58, p < .01) showed significant correlations with surprise. The higher the influence of these stimuli was rated, the greater the participants' subjective surprise turned out to be. It was also tested whether the situational context influenced the strength of the primary emotion.

Table 2. Summary of designed versus rated emotions and associated stimuli effect

Game	Situation	Designed primary and secondary emotion	Provided non-visual stimuli (predominant stimuli first)	Rated primary emotion	Rated secondary emotion	Further emotions with score > 2	Correlations main emotions with stimuli
Don't Let Go	Bee swarm	Fear, surprise	Normal sound, low frequency sound	Happiness $M = 2.25$ $SD = 1.25$	Surprise $M = 2{,}00$ $SD = 1.17$		Happiness: normal sound (significant)
	Spider behind back	Fear, disgust	Tactile back, normal sound	Surprise $M = 2.95$ $SD = 1.43$	Happiness $M = 2.5$ $SD = 1.24$	Fear $M = 2{,}15$ $SD = 0.99$	Surprise: Tactile back (significant)
Racer	Rolling stone	Surprise, anger	Low frequency sound, normal sound, seat vibration, smell	Surprise $M = 3.15$ $SD = 1.04$	Happiness $M = 2.3$ $SD = 1.17$		Surprise/Happiness: normal sound, seat vibration (not significant)
	Bad weather	Sadness, surprise	Normal sound, low frequency sound	Surprise $M = 2.3$ $SD = 0.92$	Happiness $M = 2.25$ $SD = 1.20$		Surprise: back vibration, low frequency sound (significant)
Tuscany	Sea view	Happiness, no secondary	Wind, smell, normal sound	Happiness, $M = 3.2$ $SD = 1.20$	Surprise $M = 2.85$ $SD = 1.18$		Happiness: wind, smell (significant)
	Zombie swarm	Surprise, fear	Normal sound, seat vibration	Surprise $M = 3.4$ $SD = 1.31$	Happiness $M = 2.35$ $SD = 1.35$	Fear $M = 2.05$, $SD = 1.10$	Surprise: low frequency and normal sound, back vibration, seat vibration (significant)

In terms of surprise – which was a strong primary emotion in the spider, rolling stone and zombie situation (bad weather excluded) – the difference between surprise ratings of these situations was not significant (χ^2 (2) = 3.57, p = .167). However, Mann-Whitney indicated that the feeling of happiness - which was the primary emotion in Don't Let Go (*Mdn* = 2.5) and sea view (*Mdn* = 3) - was rated significantly stronger for the latter (U= -2.24, $p < .05$). Finally, analysis of the post questionnaire on user comfort revealed no abnormalities. Overall, participants felt comfortable during the experiment (M = 4.3, SD = 0.66) without getting sick (M = 2.05, SD = 1.40) or dizzy (M = 2.25, SD = 1.33). They were quite excited (M = 3.6, SD = 0.94), while their muscles were not tense (M = 2.3, SD = 1.49).

Presence. Table 3 summarizes the scores of presence ratings for each of the games. As the table shows, the games all received mid-range ratings that at first glance seem in line with the emotional responses described before. In both situations of the Don't Let Go game, happiness was not correlated with the IPQ scales. However, the happiness ratings of the racer-game were correlated with the involvement scale for the game (rolling stone, r_s (18) = .52, $p < .05$, r_s (18) = .61, bad weather, $p < .01$). Furthermore surprise in the bad weather situation was significantly correlated with the presence scale (r_s = .47, $p < .05$) and realism (r_s = .55, $p < .05$). For the Tuscany game there was no significant correlation of happiness or surprise with the IPQ. Smell in the spider on the back situation was correlated with the involvement scale (r_s = .45, $p < .05$) as well as back vibration (r_s = .47, $p < .05$) and seat vibration (r_s = .50, $p < .05$). The bee swarm situation and both racer game situations stimuli did not correlate with IPQ scales. In the sea view situation the IPQ presence and realism scale were correlated with low-frequency (presence: r_s = .57, $p < .01$, realism: r_s = .57, $p < .01$) and normal sound (presence: r_s = .72, $p < .001$, realism: r_s = .59, $p < .01$) as well as with back vibration (presence: r_s = .62, p < .01, realism: r_s = .70, p < .01) and seat vibration (presence: r_s = .51, $p < .05$, realism: r_s = .62, $p < .01$). Low-frequency sound also correlated with the involvement (r_s = .61, $p < .01$). Within the zombies situation only normal sound did correlate with realism (r_s = .47, $p < .05$).

Table 3: Presence scores: means and standard deviations.

	Don't Let Go	Racer	Tuscany	Maximum
Presence	7.45 (1.43)	6.0 (2.18)	7.35 (2.13)	10
Involvement	10.05 (2.46)	10.65 (2.89)	11.55 (3.12)	15
Realism	6.75 (1.33)	5.55 (2.21)	6.30 (1.66)	10

In addition, analysis of variance showed significant differences between results on the presence scale (F(2, 38) = 5.63, $p < .01$.), in particular between the Don't Let Go and racer games ($p < .05$), while no significant difference was found with the Tuscany game. Friedman-Test on realism showed also significant differences between the games (χ^2 (2) = 6.72, $p < .05$). Descriptively there was again a difference between Don't Let Go and the racer game, but post-hoc tests were not significant. Involvement showed no significant result (χ^2 (2) = 4.2, p = .122).

VI. Reflection

The analysis of results of the exploratory experiment provides us with some initial indications on the effect of stimuli on emotional responses, supporting but not yet proving our hypothesis: it shows there is still a long way to go in multisensory stimuli design and assessment methodology. In this section we take a closer look at these issues, picking up the main issues we identified in section III. While describing the various topics it should be clearly noted that as a result of both the nature of the experiment (exploratory) and the number of trials we could only describe tendencies.

M1. Focus on emotional response through multisensory cues, design guidelines for deployment.

Within the experiment we successfully introduced cues spanning various sensory systems to trigger different emotional responses. Thereby, the main responses were always happiness and surprise, to a lesser extent also fear. While this mostly matched the designed emotions, it did not always match completely. As such, it will require further tuning of stimuli and responses, thereby also varying various other variables that can affect the response. As we will explain in M2 and M4, we assume cue strength has had an effect in our system. Yet, other variables will likely also have an effect: this includes the difference in strength between ambient and context-related cues, the emotional state before additional stimuli are provided, the number and duration of stimuli and response, and user variety and expectations (M2). Experimentation with various cue combinations and situations will thereby require a change in experimental procedures, but also an improvement of the validation methods (M4).

Furthermore, as a result of M1 to M4, mid-term we hope we can provide guidelines based on statistical evidence. However, while we have shown the validity of the usage of multisensory cues on triggering emotional responses, results of the first exploratory study do not grant the design of guidelines yet.

M2. Improve user presence through an increase of variety, vividness and richness of sensory cues.

While we received reasonable ratings in the IPQ questionnaire – ratings that also reflect user engagement to a certain extent (M1) - users were not highly immersed in game play. Though previous research is not conclusive on the effect of exposure on presence, the shorter duration may have been an issue. However, based on the results we assume that an emotional response – and level of engagement and immersion as a result - does likely not only depend on the kind of cue, but also on its strength (as well as the other variables we identified in M1). What we noticed is – in particular with predominantly male gamers – that cues may not have been strong enough for this particular group of end-users. In direct user feedback gamers often stated they are used to strong visual and auditory cues in games and as such expect even more from an immersive setup, which is in line with previous findings [66]. Even more so, some of the cues – in particular smell – were not highly pronounced, which may have contributed to this user feedback. While higher cue strength may be appropriate for this user group, however, other end-users may react differently. For this reason the next validation will include a far more proliferated user group, varying male and female users with different gaming experience and age. We will test the cue strength assumption with our system iteration, which increases the strength of certain cues considerably: the system will offer the potential to provide very strong wind-feedback, much denser smell stimuli, and back vibration at the upper and lower body. While the latter did not suffer from initial feedback and observations, in particular smell was not always well noticeable in the initial system, a problem that is often noticeable in installations [43].

M3. First isolate effects of separate cues on emotions, then explore inter-sensory effects.

Results indicate we yet have to achieve a perfect match between intended emotional trigger and the actual rating of an emotional response. Until further experimentation, we can identify only likely matches between stimuli and emotional response, yet, there is no solid evidence for clear matches. In the long run, we will target experimenting with a limited number of stimuli, varying its strength and duration, while also designing a wider variety of emotional-loaded situations. To this extent, the context of the stimuli in relation to ambient cues or previously provided higher intensity cues can be identified (M1). Furthermore, in the following experiment, we will add a control group that will only receive audio-visual stimuli (normal audio), to be able to address the effect of the additional secondary cues, as it will help us to further isolate the cue effects. Only once we have a better understanding of the cues in isolation we will start addressing potential cross-modal effects, likely by validation of different cue combinations to induce the same target emotional response.

M4. Create, evaluate and verify appropriate methods for measuring user engagement and emotional impact.

While the procedures used for the first experiment provided us with novel results, we also gained useful insights in how to optimize the procedure for further experiments.

First of all, user observations revealed that users seemed to physically respond more to stimuli than ratings would indicate. For example, we clearly observed many users getting highly surprised when the virtual spider would crawl behind their shoulders, yet, ratings were only mid-range. This leads to the assumption that some users may have had an issue to rate emotions. To gain a better insight, it will be preferential to add other methods of addressing emotional responses, i.e., by correlating other, more objective methods with the user provided subjective ratings. Methods like galvanic skin response, heart rate, and electromyography have been previously used to identify arousal in users [67], [68]. While these methods will not provide an absolute value, indications of arousal will be useful for analytical purposes, and have been shown to be useful in correlation with the measure of presence (M2,[28]). As such, we target the inclusion of galvanic skin response and heart rate in our follow-up studies. For now, we do not intend to adjust stimuli in real-time based on the readings of the biosensors, yet, we keep this possibility open for the future. Secondly, during assessment and analysis, we gained the impression that some of our scales and labeling thereof may be interpreted in multiple ways. As we noted in section III, multiple concepts are intertwined which will require a more detailed way to address the sub-topics and potentially a clear labeling of what is being asked. While in the first experiment we only requested a match between situation and emotional trigger, we are refining the questionnaires to couple emotional ratings directly with stimuli to be able to further isolate the effects (M4). Thereby, we will also specifically address the level of fun users had, differentiating it clearly from happiness, one of dominating emotions in all situations. We currently assume that happiness is often interpreted as "meta-emotion" in direct relation to fun: we expect that when users enjoyed the application they would rate happiness high.

VII. Conclusion

In this article we analyzed how to address the effect of multisensory cues on different emotional responses. We reported on a first user study that forms the initiator of a series of experiments that will focus on overcoming various challenges (cf. sections III and VI): focusing on emotional response (M1), improving spatial presence through stimuli fidelity (M2), first isolating, later missing interrelated cues (M3), and developing appropriate methods for measuring effect (M4). Identification of these challenges as well as the current level of methodological improvements can aid others researchers too, as many more researchers deal with similar topics. Ultimately, the goal of our studies target the creation of both a solid methodology for (a) the design and modification of cues and (b) the assessment of the quality and effect thereof. The results underline the identified challenges: The isolation of cues is difficult and, even though we now have a better understanding of how to improve the assessment, further studies are necessary, for example to address potential cross-modal effects. Nonetheless, we can already provide some take-home messages that may also be of interest to other researchers pursuing similar studies:

Isolating cue effects: there is a strong interplay between different cues that relate to the way humans process multisensory cues. However the effects may be difficult to assess after an experiment (see section III). The assessment is increasingly difficult as multiple processes or concepts - including engagement, user experience, presence, immersion and emotion - intertwine and are difficult to separate. Adequate measurement and analysis methods need to be designed that may correlate various (subjective and objective) measures to isolate the effects of a single cue, or the binding of multiple cues.

Matching user variety and expectations: user variety and expectations will have an effect on the kind and strength of a stimuli, as previous experience in games and real-life will likely effect responses. Yet, it may be difficult to address the effect of a cue without actively assessing the user state as users respond differently to stimuli. As such, while post-experiment analysis can reveal group-specific effects, it may require additional methods to address effects during stimulation.

To this extent, further experiments are planned that take into account the improved understanding of cue design and assessment methodology to further overcome the identified challenges.

References

[1] S. Beckhaus and E. Kruijff, "Unconventional human computer interfaces," in ACM SIGGRAPH 2004 course notes, 2004, vol. Los Angeles.

[2] M. Mine, "Towards virtual reality for the masses," in Proceedings of the workshop on Virtual environments 2003 - EGVE '03, 2003, pp. 11–17.

[3] J. Schell and J. Shochet, "Designing Interactive Theme Park Rides," IEEE Comput. Graph. Appl., vol. 21, no. 4, pp. 11–13, 2001.

[4] C. Park, H. Ko, I.-J. Kim, S. C. Ahn, Y.-M. Kwon, and H.-G. Kim, "The Making of Kyongju VR Theatre," p. 269, Mar. 2002.

[5] R. Morgan and D. Heise, "Structure of Emotions," Soc. Psychol. Q., vol. 51, no. 1, pp. 19–31, 1988.

[6] J. Sykes and S. Brown, "Affective Gaming - Measuring Emotion through the Gamepad," in Proceedings of the 2003 ACM Conference on Human Factors in Computing Systems (CHI 2003), 2003.

[7] M. Shiota, D. Keltner, and A. Mossman, "The nature of awe: Elicitors, appraisals, and effects on self-concept," Cogn. Emot., vol. 21, no. 5, pp. 944–963, 2007.

[8] M. Rudd, K. D. Vohs, and J. Aaker, "Awe expands people's perception of time, alters decision making, and enhances well-being.," Psychol. Sci., vol. 23, no. 10, pp. 1130–6, Oct. 2012.

[9] W. Ark, D. Dryer, and D. Lu, "The Emotion Mouse," in Proceedings of HCI International '99, 1999.

[10] R. Picard, Affective Computing. MITPress, 1997.

[11] R. Calvo and S. D'Mello, "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications," IEEE Trans. Affect. Comput., vol. 1, no. 1, pp. 18–37, Jan. 2010.

[12] S. K. D'mello and J. Kory, "A Review and Meta-Analysis of Multimodal Affect Detection Systems," ACM Comput. Surv., vol. 47, no. 3, pp. 1–36, Feb. 2015.

[13] H. O'Brien and E. Toms, "Examining the generalizability of the User Engagement Scale (UES) in exploratory search," Inf. Process. Manag., vol. 49, no. 5, pp. 1092–1107, Sep. 2013.

[14] S. Attfield, G. Kazai, M. Lalmas, and Piwowarski, "Towards a science of user engagement," in Proceedings of the WSDM Workshop on User Modeling for Web Applications, 2011.

[15] J. Lehmann, M. Lalmas, E. Yom-Tov, and G. Dupret, User Modeling, Adaptation, and Personalization, vol. 7379. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[16] R. Jacques, "Engagement as a Design Concept for Multimedia," Can. J. Educ. Commun., vol. 24, no. 1, pp. 49–59, 1995.

[17] M. Lalmas, H. O'Brien, and E. Tov, "Measuring user engagement, tutorial notes, WWW2013," 2013.

[18] M. Slater, V. Linakis, M. Usoh, and R. Kooper, "Immersion, Presence, and Performance in Virtual Environments: An Experiment with Tri-Dimensional Chess," in Proceedings of ACM Virtual Reality Software and Technology, 1996.

[19] J. Steuer, "Defining Virtual Reality: Dimensions Determining Telepresence," J. Commun., vol. 42, no. 4, 1992.

[20] J. Bell, K. Cooper, G. Kaiser, and S. Sheth, "Proceedings of the Second International Workshop on Games and Software Engineering: Realizing User Engagement with Game Engineering Techniques," 2012.

[21] M. Buchenau and J. F. Suri, "Experience prototyping," in Proceedings of the conference on Designing interactive systems processes, practices, methods, and techniques - DIS '00, 2000, pp. 424–433.

[22] E. Brown and P. Cairns, "A grounded investigation of game immersion," in Extended abstracts of the 2004 conference on Human factors and computing systems - CHI '04, 2004, p. 1297.

[23] S. Dow, "Understanding user engagement in immersive and interactive stories." Georgia Institute of Technology.

[24] M. Csikszentmihalyi, Flow: The Psychology of Optimal Experience. Harper Perennial Modern Classics, 1990.

[25] P. Sweetser and P. Wyeth, "GameFlow," Comput. Entertain., vol. 3, no. 3, p. 3, Jul. 2005.

[26] M. Slater, M. Usoh, and A. Steed, "Depth of Presence in Virtual Environments," Presence Teleoperators Virtual Environ., vol. 3, no. 2, pp. 130–144, 1994.

[27] H. Regenbrecht, T. Schubert, and F. Friedmann, "Measuring the Sense of Presence and its Relations to Fear of Heights in Virtual Environments," Int. J. Hum. Comput. Interact., vol. 10, no. 3, pp. 233–250, 1998.

[28] M. Meehan, B. Insko, M. Whitton, and F. P. Brooks, "Physiological measures of presence in stressful virtual environments," ACM Trans. Graph., vol. 21, no. 3, pp. 645–645–652–652, Jul. 2002.

[29] R McMahan, D. Bowman, D. Zielinski, and R. Brady, "Evaluating display fidelity and interaction fidelity in a virtual reality game.," IEEE Trans. Vis. Comput. Graph., vol. 18, no. 4, pp. 626–33, Apr. 2012.

[30] J. Schild and M. Masuch, "Formalizing the potential of stereoscopic 3D user experience in interactive entertainment," in IS&T/SPIE Electronic Imaging, 2015, p. 93911D.

[31] J. Schild, J. J. LaViola, and M. Masuch, "Altering gameplay behavior using stereoscopic 3D vision-based video game design," in Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14, 2014, pp. 207–216.

[32] C. Lenay, O. Gapenne, S. Hanneton, C. Genouëlle, and C. Marque, "Sensory Substitution: Limits and Perspectives," in Touching for Knowing, Y. Hatwell, A. Streri, and E. Gentaz, Eds. 2003, pp. 275–292.

[33] J. Loomis, "Sensory Replacement and Sensory Substitution: Overview and Prospects for the Future," in Converging technologies for improving human performance, M. Roco and W. Bainbridge, Eds. Kluwer Academic Publishers, 2003.

[34] E. Kruijff, "Human-potential Driven Design of 3D User Interfaces," in Proceedings of the IEEE International Conference on Artificial Reality and Telexistence, 2013.

[35] E. Kruijff, "Unconventional 3D User Interfaces for Virtual Environments," Graz University of Technology, 2007.

[36] V. Pulkki, "Spatial Sound Generation and Perception by Amplitude Panning Techniques," Helsinki University of Technology, 2001.

[37] C. Afonso and S. Beckhaus, "How to not hit a virtual wall," in Proceedings of the 6th Audio Mostly Conference on A Conference on Interaction with Sound - AM '11, 2011, pp. 101–108.

[38] M. Naer, O. Staadt, and M. Gross, "Spatialized Audio Rendering for Immersive Virtual Environments," in Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST 2002), 2002, pp. 65–72.

[39] N. Tsingos, E. Gallo, and G. Drettakis, "Perceptual audio rendering of complex virtual environments," ACM Trans. Graph., vol. 23, no. 3, p. 249, Aug. 2004.

[40] E. Kruijff and A. Pander, "Experiences of using Shockwaves for Haptic Sensations," in Proceedings of 3D user interface workshop, IEEE Virtual Reality Conference (VR2005), 2005.

[41] A. Israr, T. Williams, I. Poupyrev, C. Ioffreda, J. Cox, N. Gouveia, H. Bowles, A. Brakis, B. Knight, and K. Mitchell, "Surround Haptics," in ACM SIGGRAPH 2011 Emerging Technologies on - SIGGRAPH '11, 2011, pp. 1–1.

[42] F. Mueller, F. Vetere, M. Gibbs, D. Edge, S. Agamanolis, J. Sheridan, and J. Heer, "Balancing exertion experiences," in Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12, 2012, p. 1853.

[43] J. Kaye, "Symbolic Olfactory Display," Massachusetts Institute of Technology, Boston, 1999.

[44] Y. Yanagida, H. Kawato, H. Noma, A. Tomono, and N. Tetsutani, "A nose-tracked, personal olfactory display," in In Proceedings of ACM SIGGRAPH Sketches & Applications, 2003.

[45] T. Narumi, S. Nishizaka, T. Kajinami, T. Tanikawa, and M. Hirose, "Augmented reality flavors," in Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11, 2011, p. 93.

[46] N. Ranasinghe, A. Cheok, R. Nakatsu, and E. Y.-L. Do, "Simulating the sensation of taste for immersive experiences," in Proceedings of the 2013 ACM international workshop on Immersive media experiences - ImmersiveMe '13, 2013, pp. 29–34.

[47] M. Feng, R. Lindeman, H. Abdel-Moati, and J. Lindeman, "Haptic ChairIO: A System to Study the Effect of Wind and Floor Vibration Feedback on Spatial Orientation in VEs," in Proceedings of the IEEE Symposium on 3D User Interfaces, 2015.

[48] F. Hülsmann, J. Fröhlich, N. Mattar, and I. Wachsmuth, "Wind and warmth in virtual reality," in Proceedings of the 2014 Virtual Reality International Conference on - VRIC '14, 2014, pp. 1–8.

[49] S. Shimojo and L. Shams, "Sensory Modalities are not Separate Modalities: Plasticity and Interactions," Curr. Opin. Neurobiol., vol. 11, pp. 505–509, 2001.

[50] D. Pai, "Multisensory Interaction: Real and Virtual," in Proceedings of the International Symposium on Robotics Research, 2003.

[51] C. Spence and S. Squire, "Multisensory Integration: Maintaining the Perception of Synchrony," Curr. Biol., vol. 13, pp. 519–521, 2003.

[52] J. Weisenberger and G. Poling, "Multisensory Roughness Perception of Virtual Surfaces: Effects of Correlated Cues," in Proceedings of the 12th International Symposium on Haptic Interfaces for Virtual Environments and Teleoperators Systems (HAPTICS'04), 2004.

[53] R. Lindeman and H. Noma, "A Classification Scheme for Multi-Sensory Augmented Reality," in Proceedings of ACM Virtual Reality Software and Technology (VRST), 2007.

[54] A. Sutcliffe, Multimedia and Virtual Reality: Designing Multisensory User Interfaces. Psychology Press, 2003.

[55] S. Lederman, G. Thorne, and B. Jones, "Perception of Texture by Vision and Touch: Multidimensionality and Intersensory Integration.," J. Exp. Psychol. Hum. Percept. Perform., vol. 12, no. 2, pp. 169–180, 1986.

[56] T. Narumi, T. Kajinami, T. Tanikawa, and M. Hirose, "Meta cookie," in ACM SIGGRAPH 2010 Emerging Technologies on - SIGGRAPH '10, 2010, pp. 1–1.

[57] R. Sekuler, A. Sekuler, and R. Lau, "Sound alters Visual Motion Perception," Nature, vol. 385, 1997.

[58] R. Blake, K. Sobel, and W. James, "Neural Synergy Between Kinetic Vision and Touch," Psychol. Sci., vol. 15, no. 6, 2004.

[59] J.-P. Bresciani, M. Ernst, K. Drewing, G. Bouyer, V. Maury, and A. Kheddar, "Feeling what you Hear: Auditory Signals can modulate Tactile Tap Perception," Exp. Brain Res., vol. 162, pp. 172–180, 2004.

[60] T. Partala, "Affective Information in Human-Computer Interaction," Jun. 2009.

[61] Y. Fu, H. V. Leong, G. Ngai, M. X. Huang, and S. C. F. Chan, "Physiological Mouse: Towards an Emotion-Aware Mouse," in 2014 IEEE 38th International Computer Software and Applications Conference Workshops, 2014, pp. 258–263.

[62] S. Asteriadis, P. Tzouveli, K. Karpouzis, and S. Kollias, "Estimation of behavioral user state based on eye gaze and head pose—application in an e-learning environment," Multimed. Tools Appl., vol. 41, no. 3, pp. 469–493, Oct. 2008.

[63] M. Steffin, "Virtual Reality Biofeedback in Chronic Pain and Psychiatry," eMedicine Journal, 2005.

[64] D. Bersak, G. McDarby, N. Augenblick, P. McDarby, D. McDonell, B. McDonald, and R. Karkun, "Intelligent Biofeedback using an Immersive Competitive Environment," in Proceedings of UbiComp 2001, 2001.

[65] J. Forlizzi and K. Battarbee, "Understanding experience in interactive systems," in Proceedings of the 2004 conference on Designing interactive systems processes, practices, methods, and techniques - DIS '04, 2004, p. 261.

[66] C. Bateman, R. Lowenhaupt, and L. Nacke, "Player Typology in Theory and Practice," in DiGRA '11 - Proceedings of the 2011 DiGRA International Conference: Think Design Play, 2011, vol. 6.

[67] R. Picard and J. Scheirer, "The Galvactivator: A Glove that Senses and Communicates Skin Conductivity," in Proceedings of the 9th International Conference on Human Computer Interaction, 2001.

[68] L. Nacke and C. A. Lindley, "Flow and immersion in first-person shooters," in Proceedings of the 2008 Conference on Future Play Research, Play, Share - Future Play '08, 2008, p. 81.

Fractal Nature - Generating Realistic Terrains for Games

Adrian Cristea
Route 1 Games Ltd
1 Burton Road, S3 8BW
Sheffield, UK
a.cristea@route1games.com

Fotis Liarokapis
Human-Computer Interaction Laboratory
Masaryk University, 60200
Brno, Czech Republic
liarokap@fi.muni.cz

Abstract— This paper presents a software tool – called Fractal Nature - that provides a set of fractal and physical based methods for creating realistic terrains called Fractal Nature. The output of the program can be used for creating content for video games and serious games. The approach for generating the terrain is based on noise filters, such as Gaussian distribution, capable of rendering highly realistic environments. It is demonstrated how a random terrain can change its shape and visual appearance containing artefacts such as steep slopes and smooth riverbeds. Moreover, two interactive erosion systems, hydraulic and thermal, were implemented. An initial evaluation with 12 expert users provided useful feedback for the applicability of the algorithms in video games as well as insights for future improvements.

***Keywords** – serious games, fractal terrain generation, procedural modeling.*

I. INTRODUCTION

Creating procedural generated content is an important issue for the video games industry. One of the most popular approaches is based on fractal algorithms due to a number of reasons. First of all, they can reduce the production cost of generating art assets. Also, they can greatly reduce the download and storage size of the game. Moreover, they can reduce memory usage throughout the graphics processing unit (GPU) making them ideal for video games. Another advantage is that they allow for creativity making the game more fun and entertaining.

On the other hand, the main disadvantage of procedural generation algorithms is that it is not so easy to create very realistic results compared to other methods (i.e. laser scanning techniques). Consequently, although rough approximations of landscapes can be easily controlled, generating a specific location at a landscape is not a trivial task. Other typical problems of these algorithms include modelling specific detail in landscapes such as the generation of realistic caves and overhangs.

During the past three decades a number of rules and algorithms were presented attempting to produce realistic terrains using a variety of methods. Since 1986 [1], there have been a lot of improvements in fractal and physical based methods for representing realistic terrains. Rules have been optimized, combined and used as foundation layers for new, better approximations of naturally occurring events. A small number of the notable improvements the years brought over include: a wide range of filters for generating terrain templates such as mountains, hills and sand dunes, a diverse range of erosion approximations [2], [3] and higher capabilities of rendering extremely detailed terrain [4]. However, until nowadays there has not been an approach that combines everything together providing a solution for creating terrain environments for video games.

This paper tries to tackle these issues by simulating a selection of naturally occurring phenomena into a single environment which is easy to work with. In particular, this work extends previous work [5] by providing a set of techniques that can be used for creating realistic terrains for games such as flight simulators, action and strategy games. All these algorithms are included into a single tool which can be used for creating a variety of realistic terrain landscapes in an easy manner.

Emphasis is given on (a) generating a realistic random terrain and (b) implementing the effects of different types of erosion such as hydraulic and thermal. It is demonstrated how these types of erosion can be simulated in real-time performance. Results demonstrate how a random terrain can change its shape and visual appearance containing artefacts such as steep slopes and smooth riverbeds. An initial evaluation with 12 expert users provided useful feedback for the applicability of the algorithms in video games as well as insights for future improvements.

The rest of the paper is structured as follows. Section II provides background information on different fractal techniques for modelling terrains. Section III describes the architecture of the software tool. Section IV presents the methods used for generating and smoothing the terrain. Sections V and VI illustrate thermal and hydraulic erosion techniques. Section VII presents evaluation results whereas section VIII demonstrates conclusions and future work.

II. BACKGROUND

One of the first attempts to model fractal terrains examined different methods for the generation of fractals based on recursive subdivision [1]. A few years later, researchers addressed two issues in modelling natural terrain using fractal geometry: estimation of fractal

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

Back to Contents

dimension (using the fractal Brownian function approach), and fractal surface reconstruction (by extending Szeliski's regularization [6] with fractal priors method) [7]. Results demonstrated both estimation and reconstruction with noisy range imagery of natural terrain but discontinuities were not taken into account and sensitivity issues existed.

Fractals were also combined with different methods allowing for surface approximations and decomposition. A good example for surface approximations presented a Gibbs sampler for synthesizing constrained fractals with applications to terrain map generation from sparse elevation data [8]. Authors presented a way to synthesize surfaces of any fractal order controlling this way the roughness of the fractal geometry. For surface decomposition, simplification and multi-resolution modelling of a terrain represented as a Triangulated Irregular Network (TIN) was proposed [9]. A TIN was decomposed into areas with uniform morphological properties and the critical net was extracted inputting these values into a multi-resolution terrain model.

An alternative approach combined fractals with Digital Elevation Maps (DEMs) [10]. The model satisfies several kinds of initial constraints given by the user and can either rebuild DEMs from partial elevation datasets, or generate them from scratch. The obtained models do not suffer from steep slope artefacts. Moreover, patches from a sample terrain were used to generate a new terrain emphasizing on large-scale curvilinear features (ridges and valleys) [11]. This technique supported user-controlled terrain synthesis in a wide variety of styles, based upon the visual richness of real-world terrain data. Multifractal terrains were explored by capturing terrain characteristics of real-world data, into five parameters. These parameters were put into a multifractal terrain generation algorithm that produced synthetic terrain with features similar to those in the terrain that was analysed [12]. A later study [13] demonstrated an algorithm that can create rich terrains based on user defined feature lines using the diffusion equation. Results illustrated good visual effects; however the proposed method requires the user to create all the feature lines manually.

In 1989, a simple physical erosion model was proposed which simulates hydraulic and thermal erosion processes to create global stream/valley networks and talus slopes [14]. Rendering was done based on a ray tracing algorithm for general height fields, of which most fractal terrains are a subset. A different approach demonstrated the use of a controllable system that uses intelligent agents to generate terrain elevation heightmaps according to designer-defined constraints allowed the designer to create procedural terrain that has specific properties [15]. Moreover, a self-organization approach, via iterative cellular automata, for generating infinite 2D cave-maps in real-time was proposed [16]. Evaluation of the algorithm in a game level editor showed computational efficiency of the approach in generating playable tunnels however only 2D grids were generated. Finally, modern procedural generation techniques employed in commercial games were surveyed [17] and an extensive discussion of recent work in terrain generation has been recently published [18].

III. Architecture of the System

'Fractal Nature' is designed to act as a procedural terrain editor, with the general functionality and interactivity of any other 3D modelling software. Procedural content is generated in the form of terrain, approximations of naturally occurring events and their effects on the particular ground mesh, while allowing for user-oriented personalization of rendering outputs through the interactive on-screen options and real-time simulations. It is built based on the Unity 3D game engine, version 4.2.2, with the development process running on an Intel Core i5-2410M CPU of 2.30 GHz.

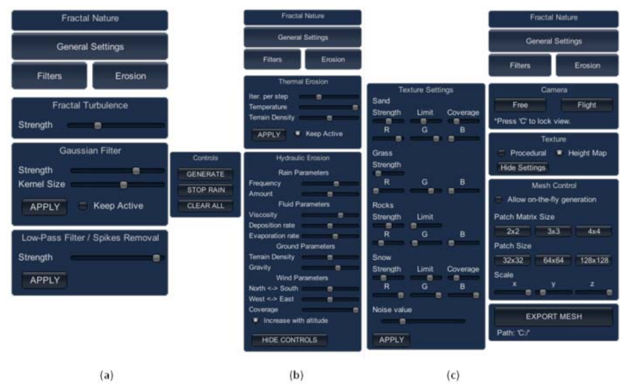


Figure 1 'Fractal Nature' GUI (a) Filters menu (b) Erosion menu, and (c) Mesh control menu

The application provides a Graphical User Interface (GUI) through which a detailed interactivity with the procedural methods is achieved. The main menu buttons, placed at the top of the GUI, are always made visible for the user to interact. An overview of the main capabilities of the application is provided in Figure 1. The GUI also incorporates two visualisation modes, 'Flight Simulator' and 'Free View'. The algorithms were initially explored in the 'Free View' mode, while the final testing was done based on the 'Flight Simulator' game (section VII). Moreover, it provides the option to save the generated mesh as an OBJ file for further processing. This makes it ideal for further editing in 3D modelling software packages or direct usage in video games.

IV. Fractal Turbulence

A complex procedural terrain generator requires certain algorithms that can quickly create different templates as starting points for further modifications. The first step in our approach was to define a terrain based on the well-known approach, the Diamond-Square subdivision [1], [2]. This method is widely used and has been regarded as a convenient approach, although its algorithm sometimes produces unwanted artefacts. It is based on a vertex displacement algorithm that is structurally uniform across the two phases of each iteration. In this first step, the initial modifications required for a smooth, recursive displacement of the vertices are computed. This involves modifying each outward corner's height by applying different values to them. In particular, the sum of the neighbour values is set to zero, the randomness is set to its highest value (returning values in the range [0,1]) and the roughness parameter is given by the user as a randomness scalar. In this second step, the centre of the selected area is heightened. Finally, the roughness parameter is a user-controlled constant which gives the variation of the slopes, statistically uniform across the map.

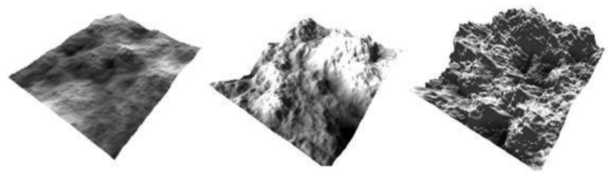


Figure 2 Three gradual values for the 'roughness' parameter and a mesh size of 192 by 192

Figure 2 demonstrates the algorithm's output, considering three gradual values for the 'roughness' parameter and a mesh size of 192x192. To smooth further the terrain and to simulate hills and sand dunes the Gaussian filter was preferred [5]. The core of the filter which is used as a template throughout the iterating process is represented by the Gaussian kernel. The values given to the visual representation of the kernel are the weights of each neighbour that plays a role in the interpolation between the vertices of the mesh. The centre point that is to be displaced is set to be relative to the middle, highest point on the kernel. Each kernel weight is computed based on [19], where the strength variable accounts for the amount of blur to be applied. Finally, each weight is normalized.

The algorithm then applies the changes to the main mesh iterates through each vertex and sums up the height of each of that vertex's neighbours, each height being multiplied by the value of its relative position in the Gaussian kernel. Inside the kernel, the centre point represents the vertex that is to be displaced, having the highest weight value. The final value of the vertex is stored on a separate map so that each mesh vertex will be displaced according to the initial, untouched, heights. When the algorithm ends, the two maps are swapped.

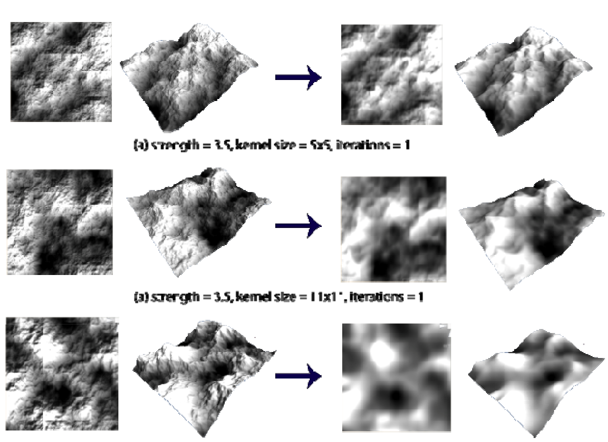


Figure 3 Three turbulent meshes and the effects the Gaussian Filter

Figure 3, illustrates how the algorithm smoothens the meshes as it should and that the level of detail being covered by the blur presents its flexibility. Although the results vary between partially rocky surfaces to smooth hill-like structures, the terrain does not fully resemble a natural environment due to its surface uniformity. To improve this, a low-pass filter iterates through all of the mesh's vertices and for each point, it averages the heights of its Moore neighborhood (Figure 4 (a)).

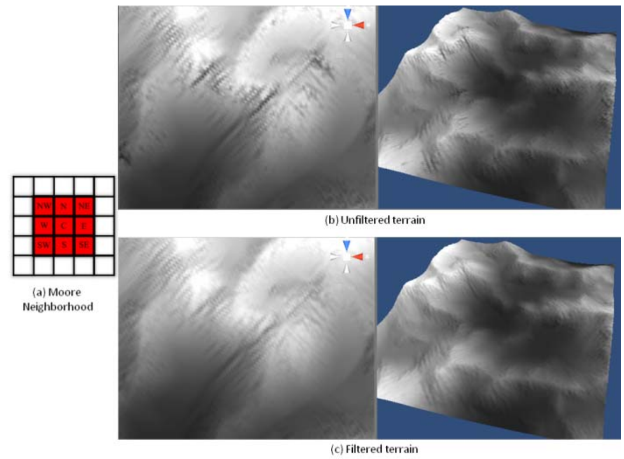


Figure 4 Left to right: (a) Moore neighborhood, (b) Unfiltered terrain, and (c) Filtered terrain

This value is then compared to the actual height of the current point being verified (in this case the vertex at the specific position) and displaced to the nearest value within the allowed threshold. Figure 4 (b) and (c) illustrates the results of the low pass filtering and the subtle change is obvious.

V. THERMAL EROSION

Thermal erosion refers to the process in which surface sediment weakens due to temperature and detaches, falling down the slopes of the terrain until a resting place is reached, where smooth plateaus tend to form. This method has various effects on the terrain, most notable being the appearance of accentuated slopes that now cover larger areas and present natural-looking ridges, as well as flat plateaus where sediment tends to deposit. The algorithm implemented here is based on [14], [32] and consists of two main phases, described following.

First it picks a random location on the map. Then, the random position's Moore neighbourhood is checked for a lowest height out of the eight neighbours, storing it as a direction reference for the sediment that might detach and fall. Secondly, it is decided whether the point is prone to erosion and further modifications are made, as described below. Here, the user can control two parameters: the minimum slope allowed (i.e. the slope at which one recursive process stops before picking another random point on the map) and the erosion process.

Figure 5, presents the output of the thermal erosion algorithm, where the properties of this method are well illustrated: long, lined peaks at high altitudes, steep slopes (that show uniformity over large areas, presenting random ridges across their length due to the amounts of sediment passing over them), eroding and depositing, and finally, smooth plateaus where sediment deposits.

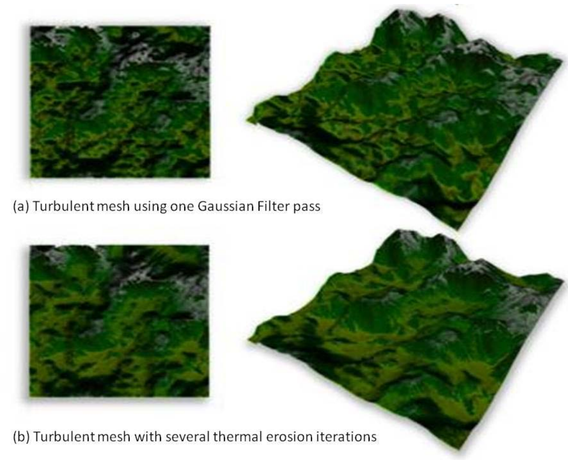


Figure 5 (a) Turbulent mesh plus one Gaussian Filter pass, (b) Same mesh after several thermal erosion iterations

In particular, Figure 5 (a) shows a turbulent mesh which is next filtered using a Gaussian kernel whereas Figure 5, (b) the thermal erosion effects applied in the terrain. Finally, a simulation was performed on an Intel Core i5-2410M CPU with a frequency capacity of 2.30GHz.

Mesh / Iterations per frame (sec/frame)	**20**	**750**	**1500**
128x128	0.07	0.09	0.11
256x256	0.30	0.31	0.32
384x384	0.43	0.45	0.47

Table 1 Simulation of thermal erosion

The average speed at which the software executes when no simulation is running is 0.015 sec/frame. Table 1 describes a comparison of different meshes for the real-time simulation of the thermal erosion model used.

VI. HYDRAULIC EROSION

Hydraulic erosion refers to the natural process in which the motion of fluids, specifically water, produces mechanical weathering over the soil. The source of the water is most typically rain, pouring droplets of liquid at random locations on the ground within a defined area. The water's property as being the best natural solvent takes effect on the terrain, eroding and depositing sediment where fit, according to a number of variables such as the slopes of the terrain, viscosity levels of the fluid or the environmental temperature. This is related to fluid dynamics which shows great flexibility in determining the result of erosion by changing the properties of the fluid in real-time.

Similarly to previous methods [3], [20], [21], [22] the hydraulic erosion's loop stores data onto five layers, continuously exchanging the information that lies at the core of the fluid/erosion simulation process: (a) *water quantities, w*; (b) *outflow flux map,* $f = (f^L, f^R, f^T, f^B)$; (c) *velocity map,* $v = (x, y)$; (d) *dissolved sediment map, s*; and (e) t*errain height map, t*. These layers are being updated in time following the order in which they are presented above, in the context of five main steps.

Firstly, the rainfall causes the water levels to rise up at random positions [20]. This is represented by (a) a switch for the rain simulation; (b) the number controlling the rain droplets per iteration and (c) a value controlling the droplet weight. The outflow flux map *f* is updated by mainly checking the slopes of the terrain, along with the gravity influence over the mass of the fluid, which is given by the virtual pipes' sizes [3]. The velocity map is built upon the outflow flux's values and so, the erosion-deposition process starts. This depends on the water viscosity (i.e. the maximum sediment capacity), the terrain density and the deposition rate. The sediment map is updated according to the velocity map, previously calculated. Finally, a percentage of the water evaporates due to environmental heat.

A. *Rain simulation*

The rain simulation has an intuitive effect: randomly distributing droplets of water across the terrain. This change in water levels populates the water map with values as shown in Figure 6.

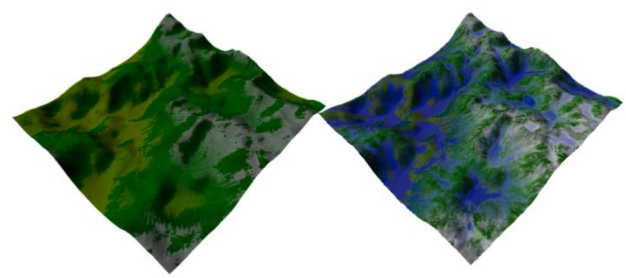


Figure 6 Rain simulation effect

Additionally, the user can control certain aspects of the simulation such as: play or stop the simulation in real-time, controll the random water distributions, and set the individual droplet weight. Each water cell coordinates the amount of fluid it holds through the values of four incoming pipes, representing the inflow flux, and four more pipes pushing the fluid outwards, the outflow flux. These pipes can be interpreted as indices to the mass of the fluid by controlling the hydrostatic pressure between each cell and its Von Neumann neighbourhood (4 neighbours). To minimise performance time, the inflow flux of a cell can be obtained by calling the outflow pipes of its neighbours:

$$outFlow^R_{[x,y]} = inFlow^L_{[x+1,y]}$$

Applying this to all four values of each cell in the outflow flux map *f*, we get a structure resembling (f^L, f^R, f^T, f^B). In this implementation it is assumed that no water can flow outside the mesh. This requires constant out-of-bound checks when calling a cell's neighbouring values and recursive loops. f^L is then calculated for t[x,y]'s left outflow flux as follows:

$$f^L_{[x,y]} = max(\, 0,\ f^L_{[x,y]} + T * A * \frac{G * \Delta h^L_{[x,y]}}{l}\ )$$

where *T* is regarded as the time step, A represents the cross-sectional area of the pipe, G, is the gravity and $\Delta h_{[x,y]}^L$ equals the total height difference between the current vertex and its left neighbour:

$$\Delta h^L_{[x,y]} = (t_{[x,y]} + w_{[x,y]} + s_{[x,y]}) - (t_{[x-1,y]} + w_{[x-1,y]} + s_{[x-1,y]})$$

If the sum of the four outflow flux values exceeds the water amount present at $w_{[x,y]}$, the algorithm will output negative values, so each virtual pipe is multiplied by a factor K, as shown below:

$$f^i_{[x,y]} = f^i_{[x,y]} * K, \quad i = L, R, B, T,$$

where

$$K = \min\left(1, \frac{w_{[x,y]} * l_x * l_y}{(f^R + f^L + f^T + f^B) * \Delta t}\right)$$

where l_x and l_y represent the distances between the two water cells in the X and Y directions. Until now, the terrain map has been fully iterated through one single time and the data gathered concerns the water map w and the outflow flux map f. Next, the program starts the iteration loop once again, this time going through the remaining phases.

B. Water movement

By gathering the inflow flux values and releasing the outflow ones, the net water volume change over one time step can be deducted:

$$\Delta V_{[x,y]} = \max\left(0,\ \mathrm{T} * \left(\sum_{i=L,R,T,B} f^i_{inFlow} - \sum_{i=L,R,T,B} f^i_{outFlow}\right)\right)$$

The volume of the water cell is then updated according to the above $\Delta V_{[x,y]}$:

$$w_{[x,y]} = w_{[x,y]} + \frac{\Delta V_{[x,y]}}{l_x * l_y}$$

To proceed to the erosion-deposition model, the velocity map requires updating. The values in this map hold the amounts of water passing through one cell at a certain time step in a horizontal direction υ = (x, y). The formula presented below computes the X direction value of the current cell's fluid velocity. Similarly, this structure is also applicable for computing the velocity's Y direction.

$$\vec{v}_x = f^R_{(x-1,y)} - f^L_{(x,y)} + f^R_{(x,y)} - f^L_{(x+1,y)}$$

Once we can deduce the speeds and directions of the water from the outflow flux map *f* and the amounts of fluid moved throughout the terrain from the velocity map v, the algorithm renders a realistic water movement over the terrain (Figure 7).

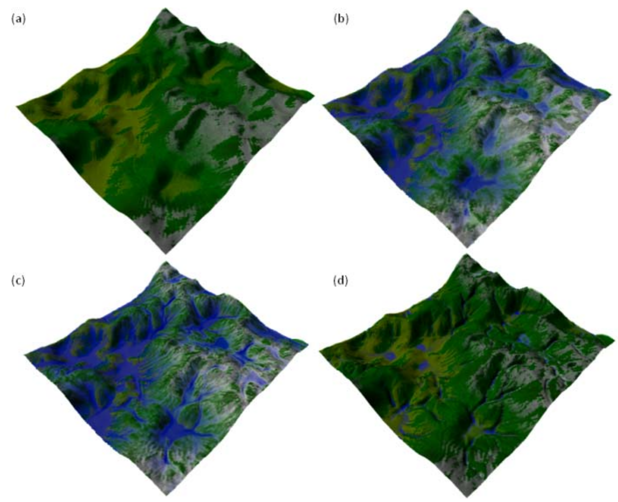


Figure 7 Erosion and deposition steps

C. Erosion and deposition

Next, the method can proceed to the erosion/deposition phase, where the water's solvent property can be exploited. This process heavily relies on the sediment capacity of the fluid which describes the amount of allowed fluid viscosity in a particular cell. The user can manipulate the sediment capacity scalar, Kc, while the algorithm includes the slope amount of the terrain slopeValue and the v velocity:

$$sedCap = Kc * velocity.magnitude * slopeValue[x, y]$$

where *slopeValue* represents the angle at which the normal of the point is orientated relative to the Y axis. This is achieved by subtracting from one length unit the absolute value of the dot product between the two vectors: vertex normal and Y axis. To avoid very small values the lower boundary is set to be 0.1. When accounting for the slope of the terrain, the sediment capacity value gets directly linked to the geometry of the mesh. Thus, it behaves differently in certain circumstances. If the slope is steep, the algorithm will dissolve more sediment, but if the surface is flatter, the terrain will not be heavily affected. The method is also able in producing steep river-like structures, ending with the appearance of canyons.

The next step refers to the approximation of the sediment amount that is to be eroded, Ms, as well as the amount to be deposited, Md, according to the cell's sediment capacity sedCap and actual sediment amount $s_{[x,y]}$.

$$Ms = max\,(0, Ks * (sedCap - s_{[x,y]})$$
$$Md = \max\,(0, Kd * (s_{[x,y]} - sedCap)$$

where Ks is the terrain density and Kd is the deposition scalar. The decision of whether to deposit sediment or dissolve the terrain comes from the comparison of the maximum sediment capacity of the water with the actual sediment value at point [x, y]. By checking these two values, if (sedCap > s[x, y]), the water is capable of displacing more terrain, adding the amount onto the sediment map. Otherwise, the sediment in the water is too 'heavy' for the fluid to displace and so the excess is deposited on the ground and cleared from the sediment map. Next, the current suspended sediment is carried by the velocity field throughout the map. The new value is determined by taking an Euler step backward in time, relative to the velocity X and Y values.

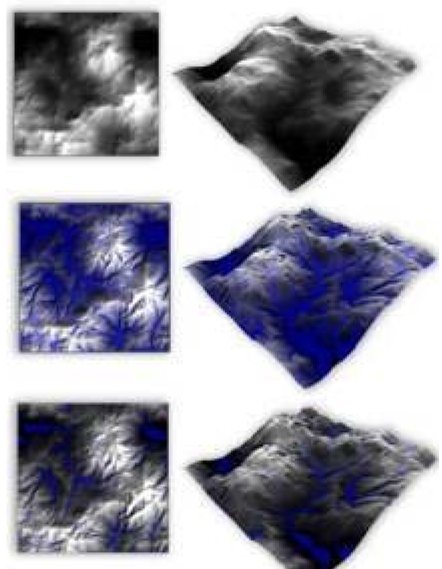


Figure 8 Evaporation results

Finally, part of the water evaporates due to environmental heat (Figure 8). The evaporation equation can be represented as follows, where Ke is the evaporation constant and T is time.

$$w_{[x,y]} = w_{[x,y]} * (1 - Ke * T)$$

Similarly to thermal erosion, a simulation was performed for the hydraulic simulation.

Mesh	Performance Time(sec/frame)
64x64	0.08 sec/frame
128x128	0.17 sec/frame
256x256	0.34 sec/frame
384x384	0.55 sec/frame

Table 2 Simulation of hydraulic erosion

Table 2 describes the real-time simulation of the hydraulic erosion model used for a number of different meshes.

VII. EVALUATION RESULTS

To test the algorithms presented in the previous sections, a simple flight simulator game was created. To make the game more immersive, 3D model of an airplane was imported into the game. For the gaming environment, an infinite terrain was specified. In addition, some other graphics effects were implemented, including basic procedural textures and simple clouds. The procedural textures were generated by weighting each layer according to set rules and lastly combining the colours. Four different types of simple procedural textures were implemented including: sand, grass, rocks and snow. A screenshot illustrating how the game looks like is presented in Figure 9.

Figure 9 Flight simulator game

A. Procedure

To test the functionality of the game as well as the effectiveness and realism of the algorithms a qualitative evaluation with 12 expert users was performed in a laboratory environment. Six expert users had a computer graphics/virtual reality background and are employed as academics in higher education establishments. The rest of them, are creative industry professionals working in game companies.

The testing lasted approximately 30 minutes and the think-aloud methodology [23] was employed. The main features of the GUI were presented followed by the flight simulator game. In addition, a comparison between the generated results from thermal and hydraulic erosion was shown to them as illustrated in Figure 10 and Figure 11 and feedback was recorded. The main purpose of this was to check on whether the randomised simulation results can produce effects which look relatively similar to real landscapes so that they can be applied for generating video games.

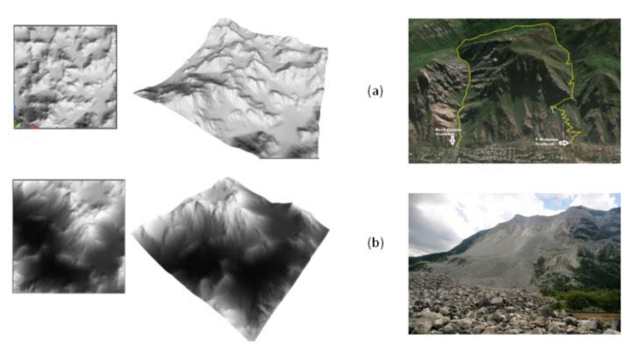


Figure 10 Thermal erosion evaluation (a) Medium temperature values comparison (b) High temperature values comparison

Figure 10 presents the comparison between the simulation results of thermal erosion and photographs collected over the web representing real landscapes. Figure 10 (a) shows how medium values for temperature and terrain density render flat plateaus with considerably more localized rocky segments in between them compared with a Canyon near Utah [22]. Figure 10 (b) illustrates maximum temperature plus a low terrain density results in a large amount of displaced sediment over a turbulent mesh compared with Frank Slide, Canada [25].

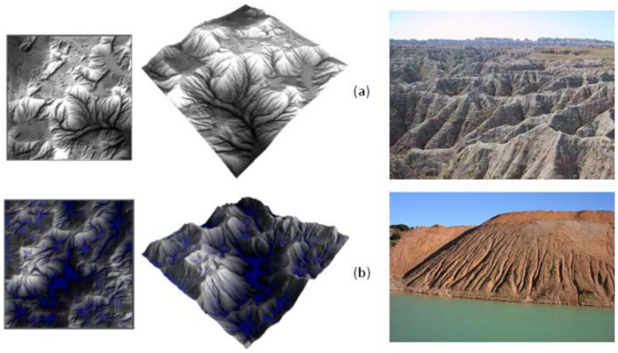


Figure 11 Hydraulic erosion evaluation (a) Low density and heavy erosion comparison (b) Low density and high viscosity comparison

On the other hand, Figure 11 presents the comparison between the simulation results of hydraulic erosion and photographs collected over the web representing real landscapes. Figure 11 (a) illustrates a turbulent terrain with a low density, accompanied by heavy erosion from rain compared with Badlands [24]. Figure 11 (b) shows a low density, thermally eroded turbulent terrain influenced by water with high viscosity capacity compared with Spoil heaps at Cairdshill Quarry [27].

B. Feedback

The received feedback was promising and gave useful insights for further improvements. In particular, the expert users from the computer graphics/virtual reality background provided a number of comments on how to further improve the visual realism of the terrain. They were very critical on the implementation but they liked the idea of having a software tool for speeding up the development process for video games. Their main focus was on using more accurate algorithms for hydraulic erosion such as [28], [29] and [30]. Some proposed making the software open-source so that it can be improved by other researchers and software developers.

One user mentioned that it would be a good idea to compare these methods with Perlin noise generator since by changing the gradient distribution in each octave to an exponential distribution results in realistic and interesting procedurally generated terrain [31]. Two users proposed to port the implementation (or part of it) on GPUs so that it will be significantly faster. Another user pointed out that fluid simulation would be very useful for the representation of sea, rivers and lakes. A final remark was to include more stohastic smoothing algorithms in order to make the landscapes look more realistic.

The games professionals indicated a very positive response and found the level of realism quite satisfying and appropriate for the generation of interactive video games. They concluded that there are strong similarities with real life mountains that have gone through years of erosion by temperature and hydraulic erosion, successfully simulating natural water movement over the terrain. They stated that the proposed approach takes into account a wide range of variables that are able to simulate a high approximation of diverse fluids and the mechanical weathering they exert on the underlying terrain.

Moreover, all of them stated that although the algorithms produce realistic enough results, the generated terrains cannot be used without manual editing into modern games. Professionals from larger studios mentioned that they are using similar procedural approaches but then they perform a lot of editing. On the other hand, independent studios stated that they could use the generated terrain as they are. They suggested incorporating more procedural elements such as different types of trees and plants. They also commented on the simplicity of the texturing techniques. Then it would make more sense to make more comparisons with real-lanscapes and will make the software environment more promosing.

VIII. Conclusions and Future Work

This paper presented a procedural generation terrain tool that includes a set of methods that can be used for simulating naturally occurring phenomena and can be used for creating realistic terrains for games. Emphasis is given on including effects of different types of erosion such as hydraulic and thermal. This aims at simulating how these types of erosion act on the body of the terrain and change its shape to one containing artefacts such as steep slopes and smooth riverbeds. Initial evaluation results with a small number of users provided promising feedback indicating that the algorithms can be used for creating content for interactive games with minor modifications. However, it was suggested that the solution would be much more useful if more functions are available.

Future work aims to reach out for consideration of more accurate erosion algorithms as mentioned in the user evaluation. Real and natural phenomena must be taken into consideration instead of approximating, aiming for highly realistic results coming from fractal geometry. Although the algorithms were applied for the generation of a flight simulator game, they have the potential of being applied for other types of games such as action and strategy games. As a result, a wider scale evaluation study will take place evaluating not only the effectiveness of the algorithms but also their applicability to a wider range of video games.

Acknowledgements

A video that illustrates the functionality of the fractal algorithms is illustrated at: https://www.youtube.com/watch?v=eFVgls1BSnM&feature=youtu.be

References

[1] Miller, G.S.P. “The definition and rendering of terrain maps”, Proc. of the 13th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH), ACM Press, 20(4): 39-48, (1986).

[2] Olsen, J. “Realtime Procedural Terrain Generation”, Technical Report, University of Southern, Denmark, (2004).

[3] Mei, X., Decaudin, P., Hu, B. “Fast Hydraulic Erosion Simulation and Visualization on GPU”, Proc. of the 15th Pacific Conference on Computer Graphics and Applications (Pacific Graphics), IEEE Computer Society, Maui, Hawaii, 47-56, (2007).

[4] Strugar, F. “Continuous distance-dependent level of detail for rendering heightmaps”, Journal of Graphics, GPU, and Game Tools, Taylor & Francis, 14(4): 57-74, (2009).

[5] Noghani, J., Liarokapis, F., Anderson, E.F. “Randomly Generated 3D Environments for Serious Games”, Proc. of the 2nd IEEE Int'l Conference in Games and Virtual Worlds for Serious Applications, IEEE Computer Society, Braga, Portugal, 3-10, (2010).

[6] Szeliski, R., Terzopoulos, D. “From splines to fractals”, Proc. of the 16th Annual conference on Computer graphics and interactive techniques (SIGGRAPH '89), ACM Press, 23(3): 51-60, (1989).

[7] Arakawa, K., Krotkov, E. “Fractal Modeling of Natural Terrain: Analysis and Surface Reconstruction with Range Data”, Graphical Models and Image Procesing, 58(5): 413-436, (1996).

[8] Vemuri, B.C., Mandal, C., Lai, S.H. “A fast Gibbs sampler for synthesizing constrained fractals”, IEEE Transactions on Visualization and Computer Graphics, 3(4): 337-351, (1997).

[9] Danovaro, E., De Floriani, L., Magillo, P., Mesmoudi, M.M., Puppo, E. “Morphology-driven simplification and multiresolution modeling of terrains”, Proc. of the 11th ACM Int'l symposium on Advances in geographic information systems, ACM Press, 63-70, (2003).

[10] Belhadj, F. “Terrain modeling: a constrained fractal model”, Proc. of the 5th Int'l conference on Computer graphics, virtual reality, visualisation and interaction in Africa (AFRIGRAPH '07), ACM Press, 197-204, (2007).

[11] Zhou, H., Sun, J., Turk, G., Rehg, J.M. “Terrain Synthesis from Digital Elevation Models”, IEEE Transactions on Visualization and Computer Graphics, 13(4): 834-848, (2007).

[12] Van Pabst, J.L., H. Jense, H. “Dynamic Terrain Generation Based on Multifractal Techniques”, Proc. of the Int'l Workshop on High Performance Computing for Computer Graphics and Visualisation, 186-203, (1995).

[13] Hnaidi, H., Gurin, E., et al. “Feature based terrain generation using diffusion equation”, Computer Graphics Forum, 29(7): 2179-2186, (2010).

[14] Musgrave, F.K., Kolb, C.E., Mace, R.S. “The synthesis and rendering of eroded fractal terrains”, Proc. of the 16th Annual conference on Computer graphics and interactive techniques (SIGGRAPH '89), ACM Press, 41-50, (1989).

[15] Doran, J., Parberry, I. “Controlled Procedural Terrain Generation Using Software Agents”, IEEE Transactions on Computational Intelligence and AI in Games, 2(2): 111-119, (2010)

[16] Johnson, L., Yannakakis, G.N., Togelius, J. “Cellular automata for real-time generation of infinite cave levels”, Proc. of the 2010 Workshop on Procedural Content Generation in Games, ACM Press, Article No. 10, (2010).

[17] Hendrikx, M., Meijer, S., Van Der Velden, J., Iosup, A. “Procedural content generation for games: A survey”, ACM Transactions on Multimedia Computing, Communications, and Applications, ACM Press, 9(1), Article No. 1, (2013).

[18] Smelik, R.M., Tutenel, T., Bidarra, R., Benes, B. “A Survey on Procedural Modelling for Virtual Worlds”, Computer Graphics Forum, Willey, 33(6): 31-50, (2014).

[19] Oliva, A., Torralba, A., Schyns, P.G. “Hybrid images”, ACM Transactions on Graphics, 25(3): 527-532, (2006).

[20] Št'ava, O., Beneš, B., Brisbin, M., Křivánek, J. “Interactive terrain modeling using hydraulic erosion”, Proc. of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, 201-210, (2008).

[21] Anh, N.H., Sourin, A., Aswani, P. “Physically based hydraulic erosion simulation on graphics processing unit”, Proc. of the 5th Int'l conference on Computer graphics and interactive techniques in Australia and Southeast Asia (GRAPHITE '07), ACM Press, 257-264, (2007).

[22] Jákó B., Tóth, B. “Fast Hydraulic and Thermal Erosion on GPU”, Proc. of Eurographics 2011, Short Papers, 57-60, (2011).

[23] Lewis, C.H. “Using the 'Thinking Aloud' Method In Cognitive Interface Design”, Technical report, IBM RC-9265, (1982).

[24] Tomco, M. “Slide Canyon to Rock Canyon Hike”, Available at: http://www.thewildernesslife.com/slide-canyon-to-rock-canyon-hike/, Accessed at: 10/3/2015.

[25] Crowsnest Pass Photo: The Frank Slide Interpretive Centre and Turtle Mountain, Available at: http://www.tripadvisor.com/LocationPhotoDirectLink-g1472384-i39784428-Crowsnest_Pass_Alberta.html, Accessed at: 10/3/2015.

[26] Badlands, Globe Tourism, http://globe-tourism.com/badlands.html, Accessed at: 10/3/2015.

[27] Burgess, A. “Erosion by Water”, Available at: http://commons.wikimedia.org/wiki/File:Erosion_by_Water_-_geograph.org.uk_-_1178731.jpg, Accessed at: 10/3/2015.

[28] Takayasu, H., Inaoka, H. “New type of self-organized criticality in a model of erosion”, Physical Review Letters, 68(7), (1992).

[29] Somfai, E., Sander, L.M. “Scaling and river networks: A Landau theory for erosion”, Physical Review E, 56(1), R5(R), (1997).

[30] Génevaux, J.D., Galin, E., Guérin, E., Peytavie, A., Beneš, B. “Terrain generation using procedural models based on hydrology”, ACM Transactions on Graphics (TOG) - SIGGRAPH 2013 Conference Proceedings, ACM Press, 32(4), Article No. 143, (2013).

[31] Parberry, I. “Tobler's First Law of Geography, Self Similarity, and Perlin Noise: A Large Scale Analysis of Gradient Distribution in Southern Utah with Application to Procedural Terrain Generation”, LARC Technical Report LARC-2014-04, Laboratory for Recreational Computing, University of North Texas, June (2014).

[32] Jako, B. Toth, B. “Fast Hydraulic and Thermal Erosion on GPU”, Eurographics 2011, Short Papers, 57-60, (2011).

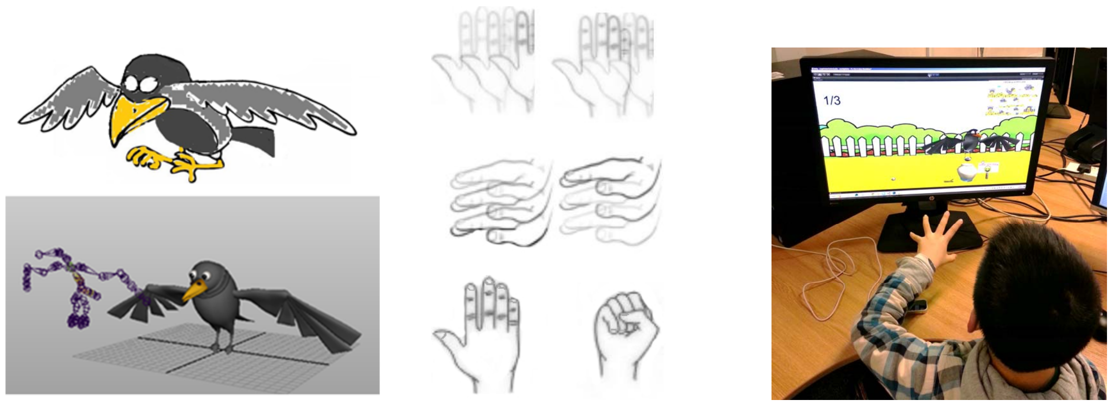
Back to Contents

Puppet Narrator: utilizing motion sensing technology in storytelling for young children

Hui Liang[1,2]
[1]National Centre for Computer Animation
Bournemouth University, Poole, UK
[2]Communication University of China,
Beijing, China
Email: liangh@bournemouth.ac.uk

Jian Chang*， Ismail K. Kazmi， Jian J. Zhang
National Centre for Computer Animation
Bournemouth University, Poole, UK
Email: jchang@bournemouth.ac.uk

Peifeng Jiao
Basic Medical School
Southern Medical University
Guangzhou, China



(a) (b) (c)

Fig. 1. (a) puppet crow sketch model and its 3D avatar; (b) hand gestures; (c) use hand gestures to control crow

***Abstract*— Using avatars in storytelling to assist narration has proved to be beneficial on promoting creativity, collaboration and intimacy among young children. Development of novel Human Computer Interaction (HCI) techniques provides us with new possibilities to explore the training aspects of storytelling by creating new ways of interaction. In this paper, we design and develop a novel digital puppetry storytelling system - Puppet Narrator for young children, utilizing depth motion sensing technology as the HCI method. More than merely allowing children to narrate orally, our system allows them to use hand gestures to play with a virtual puppet and manipulate it to interact with virtual items in virtual environment to assist narration. Under this novel pattern of interaction, children's narrative ability can be trained and the competencies of cognition and motor coordination can also be nourished.**

Keywords— Serious game; digital storytelling; virtual puppet; hand gesture recognizing; children's ability training

I. Introduction

In recognition for their considerable positive effects on pedagogy, educational games or serious games for training purposes have become immensely popular [1-2]. As a modern form of traditional storytelling, digital storytelling systems emerged over the last few years and have demonstrated powerful pedagogical functions, which enable children to express themselves and cooperate with others during narrating performance. Storytelling is essentially one of the original forms of teaching [3], which can be used as a method to teach ethics, values, and cultural norms and differences. Digital storytelling following the same well-known strategies similar to classical storytelling can help children to acquire several technological skills and work in groups and strengthen the bonds between each other. As another social benefit, digital storytelling can also help disabled children or students with learning difficulties to remove the barriers of communication with adults and peers and overcome the inability to focus on their feelings or thoughts by providing them with opportunities to play active roles [4]. At present, the major pedagogical benefit gained with digital storytelling is the ability to narrate [5-6].

However, storytelling is not just about narrative. In its basic form, storytelling is usually combined with gestures and expressions. Oral narrative can also be combined with other body movement, e.g. dancing to enhance the storytelling through remembrance and dramatic enactment of stories [7]. From this point of view, storytelling will not only benefit understanding narrative structures but also fertilize other abilities, such as cognitive competence and physical coordination during performance with the aid of different media.

The development of novel Human Computer Interaction (HCI) methods (e.g. depth motion sensing technology) provide us with new possibility for deriving storytelling's educational benefits by creating new ways of enabling interaction through hand gestures or other modalities. Since oral narrative is not the only thing that matters in storytelling, children's other capacities is also expected to increase during their storytelling

*corresponding author:jchang@bournemouth.ac.uk

performance, such as cognitive competence and physical coordination etc., especially when digital storytelling system could provide players more immersive interactions, such as using hand gestures to manipulate avatars to perform in virtual environments. Considering puppetry's positive benefits in education, we also study an interactive digital puppet model which works as a virtual actor to assist children's performance, as shown in Figure 1(c).

In this paper, we design and develop a novel digital puppetry storytelling system-Puppet Narrator for young children utilizing depth motion sensing technology which supports hand and finger motions as input but requiring no hand contact or touching. Considering a puppet's operation complexity for young children, we use hand tracking and gesture recognition technologies to simplify operations and provide intuitive interface, in which children can use hand gestures to manipulate virtual puppet to perform story. And different from precious research [8], in addition to narrative fertilization, we also devote to increase children's abilities on the aspects of cognitive development, and motor coordination ability during their storytelling performance with the help of depth motion sensing device.

In summary, this work has two main contributions:

- Introduce a novel narrative assistance with gesture control and computer animation by combining motion sensing technology to manipulating 3D puppet;
- Implement a prototype of the digital storytelling system to help young children to develop their related skills.

The remainder of this paper is organized as follows. Section 2 presents our system design. Section 3 describes system implementation including the architecture, input data processing, motion control, and output. Section 4 presents the examples and results. Discussion in Section 5 follows and Section 6 concludes the paper.

II. General Design

In this section we discuss the system design which involves the novel manifestation pattern as well as the pedagogical considerations we mentioned previously.

A. Target

Recently, there has been substantial amount of research undertaken on digital storytelling mainly investigating narrative abilities training, such as Toontastic, Kodu, Storytelling Alice, Wayang Authoring, et al [8]. However, as we mentioned before, storytelling is not just narrating. Beside oral narration, the development of other abilities is also vitally important for young children, e.g. space and object cognitive abilities. If powerful user friendly interaction methods, which are more intuitive and natural, are provided by digital storytelling systems, then the additional abilities (cognition and motor coordination) will also be considered.

There is a plethora of research in the psychology field which is used for development of important skills in children [9-12]. Research suggests that spatial-temporal reasoning and spatial visualization ability is an important indicator of achievement in science, technology, engineering and mathematics [13] and a pre-schooler's visual spatial attention ability predicts his future reading skills [14]. Researchers have also postulated a set of so-called "core domains" in cognitive development and suggested that children have innate sensitivity to specific kinds of patterns of information. Those commonly speculated core skills of cognition include: number [15], space [16], visual perception [17], essentialism [18], and language acquisition [19]. As an important aspect of children's psychosocial development, the significance of motor coordination competence has also been recognized in pedagogy long time ago. Children with poor motor coordination have been found to underachieve educationally and to experience difficulties with peer relationships [20].

Puppet Narrator is developed to target children between 5 and 8 years old. Our aim is to endow digital storytelling with a novel interaction method, which is more flexible and immersive. At the same time, our system is highly educational not only in terms of narrative competence but also cognitive ability and motor coordination. In our system, besides narrative ability training, we also pay attention to parts of these core skills developments: number, space and visual perception with the assistance of motion sensing technology in virtual environment. To enhance the children's motor coordination competence, by using depth motion sensor to track and recognize players' hand movement and gestures, our system enables children to use their hand motions to manipulate virtual puppet for interacting during narrating.

Our preliminary conception is: following the provided story plot, children will finish the whole story narration, and at the same time, children can simply use hand motion to control the movement of virtual puppet and interact with playthings in virtual scenario to assist narrating. Through this procedure, their narrative ability will be nourished. Using hand gestures for controlling the avatar, their motor coordination ability will be trained. In interaction with virtual items having different properties and roles in the story, their space and object recognition capability will also be developed.

Based on the above considerations, there are several aspects we should consider in our system design, as summarized below:

1. For the purpose of narrative ability training:
   a. Story topic should be carefully chosen, which should be familiar to the children and can be repeated and narrated easily.
   b. The story should have pedagogical meaning.
   c. The story plot should be provided to the children to follow before the story commences.
2. For the purpose of cognitive core skills' development:
   a. Numerical cognitive: players should finish a certain number of goals, denoted by a score within the game.
   b. Spatial cognitive: players should have a clear recognition of their hand's location in real world as well as the location of avatar and other playable objects in virtual environment. They will develop the ability to map the position of their hands from the real world to the position of the avatar in the virtual world.

Back to Contents

c. Visual perception: players should have the ability to distinguish different playable objects having different functions and uses.

3. For the purpose of motor coordination ability training:
   a. Players are required to perform hand gestures correctly and in sequence.
   b. Players are required to move their hands steadily and accurately.

The structure of our training target is illustrated in Figure 2.

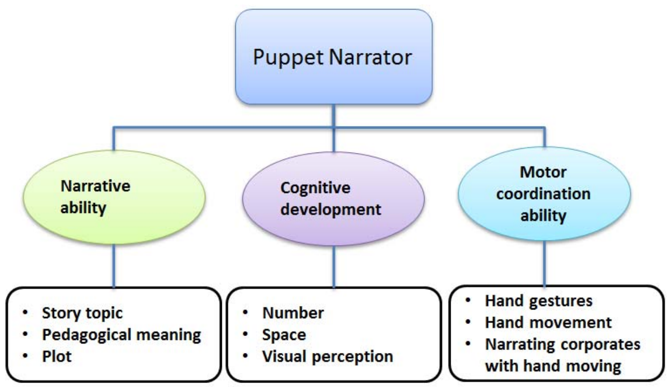


Fig. 2. Training target of Puppet Narrator

B. Story Topic

"The Crow and the Pitcher" is one most famous of Aesop's Fables. A thirsty crow found a pitcher with some water at the bottom out of the reach of its beak. The crow picks up pebble stones and drops them into the pitcher to raise the water level until it can drink the water. The fable is made by ancient Greek poet Bianor [21] and then collected by Avianus [22]. The fable emphasizes the virtue of thoughtfulness than brute strength and the value of the crow's persistence. Considering its popularity among young children as well as its positive pedagogical meaning, we choose "The Crow and the Pitcher" as the story topic.

During narration within our virtual environment young children could use their hands and a set of hand gestures to manipulate the puppet crow to pick up pebbles and drop them into the pitcher. The crow's actions, such as flying, grasping, and drinking will be presented through pre-recorded animations and controlled by hand gestures.

C. Pipeline

The component level interaction within the system is shown in Figure 3. First, a story plot is provided as storytelling hints to young players. Second, players use hand motion to manipulate the avatar through depth motion sensor device, which can automatically track hand motion and recognize hand gestures. Depth image data from the motion sensor are obtained and interpreted into motion control commands by the host computer. Finally, as the visual feedback, avatar's responding animation is provided to players and then, players adjust their hand gestures/movement to push the plot forward and narrate the story. Under this novel manifestation pattern, not only players' oral narrative ability but also their cognitive and motor coordination competence is expected to be developed.

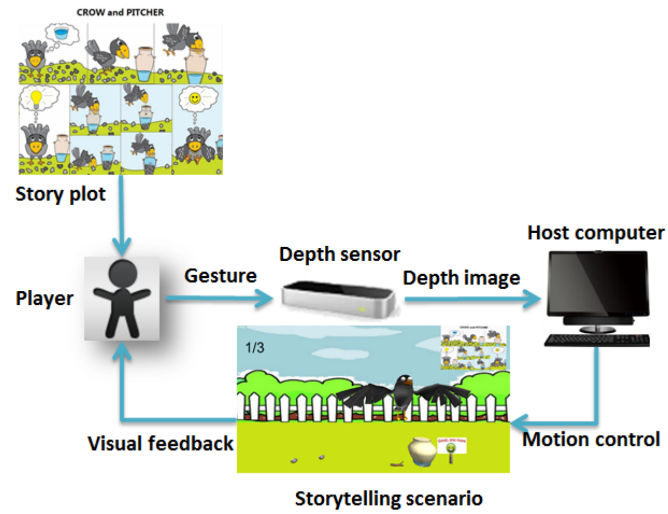


Fig. 3. System pipeline

III. Implementation

Puppet Narrator is mainly composed of three parts: Input, Motion Control, and Output, as illustrated in Figure 4. The input part processes the sensor data captured from motion sensor device through HCI and passes it to the next part. Motion Control interprets the data subsequently and determines avatar's location and posture. The output module updates avatar performance in virtual environment as the feedback.

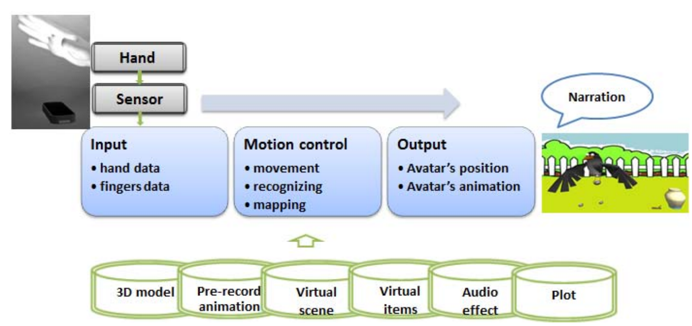


Fig. 4. System architecture

We utilized a Leap Motion controller in our system as the HCI sensor device to track hand gestures, which can provide a high fidelity finger tracking through an infrared depth sensor. We utilize the Leap Motion SDK provided by the Leap Motion Co. as the API to access the motion data of hands and fingers from the device. All the 3D models and animations were created in Maya 2014. We integrated and developed the entire system in Unity3D Pro V4.2.

A. Digital puppet crow design

To make our system more appealing for young children, considering puppetry's positive benefits in education, we use a digital puppet as an avatar to assist children's storytelling through animation technology.

1) Puppet geometry construction, which describes puppet shape, the rigging system, and its shading materials. The shape of the digital crow is modelled in Maya 2014 simulating the picture of the crow in the plot. Rigging the puppet defines its

behaviour with bones connecting joints. The sketched crow and the digital 3D crow model are shown in Figure 1(a).

2) User interaction, which defines the way players interact with the puppet, describing the performer expressions, and interaction interfaces.

3) Animation. Puppet crow uses pre-recorded animations to produce actions, which can be trigger by the players' hand gestures. Some screenshots are shown in Figure 5.

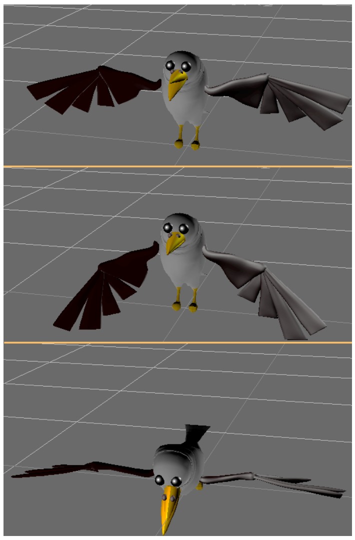


Fig. 5. Pre-recorded animations

B. Input data processing

The sensor data provided by Leap Motion controller contains a diversity of information about hands and pointables (such as fingers or finger-like tools defined by Leap Motion Co. [23]), in the virtual scene, which is updated by frame and can be represented as follows:

$$SensorData = \langle FR, H, P, T \rangle \quad (1)$$

where FR is frame rate; H represents the set of hands detected and P represents the set of pointables; and T is timestamp. Hand data H mainly contains the hand id, direction and different value about the palm position and status:

$$H = \langle id, dir, palmInf \rangle \quad (2)$$

And, pointables data P includes its id, direction and position information relative to the hand:

$$\mathrm{P} = \langle id, dir, handid, positionInf \rangle \quad (3)$$

C. Motion control

Recent research on neuroscience found that in human brain development there is a strong connection between perception, imagination and movement. According to previous studies [24-26], through the feedback of their own movement visual observation or the feedback of avatar's (3D virtual character's) motion [27], human can recognize and coordinate with their movements better.

The motion control module is the core of the system, which has two main functions: movement control and recognizing. Movement control function is responsible for mapping the play's relative hand position to the movement of the puppet crow. Recognizing function is in charge of identifying the pattern of hand movement and gestures implicating player's intention. Once a recognizable gesture is detected, recognizing function will trigger a pre-recorded animation or event (e.g. gripping a pebble or flapping wings).

1) Movement control mechanism

Since in a storytelling system children mainly focus on narrating, a complex puppet manipulation [28] as if we are stringing a puppet in a real show will be distracting or even hamper narrative. If young children pay much attention on puppet manipulation, they might forget the story line. From this consideration, we design a user-friendly puppet prototype used as a storytelling avatar with a simpler interaction manner and easier motion control mechanism.

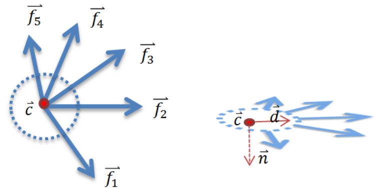
For young children to control the movement of the puppet/avatar, the most direct way is converting the translation of their hands position in Leap Motion coordinate system to the position of puppet in virtual environment. Considering the different scale between these two workspaces, a proper transformation matrix should be involved as a scaling into the coordinate translation. Once players' hands is not recognized by the Leap Motion device, puppet will keep its position in the last frame and then resume the movement immediately when hands can be detected again in the following frames. Within each frame, the position of the puppet is decided by the values of two coordinate vectors: the puppet's former coordinate in previous frame and the hand's relative translation generated by player's hand movement in current frame. And the scaling factor considered at the meanwhile, the puppet crow's coordinate in virtual workspace is computed as follows:

$$P_{crow} = M \cdot S \cdot R_{hand} + P'_{crow} \quad (4)$$

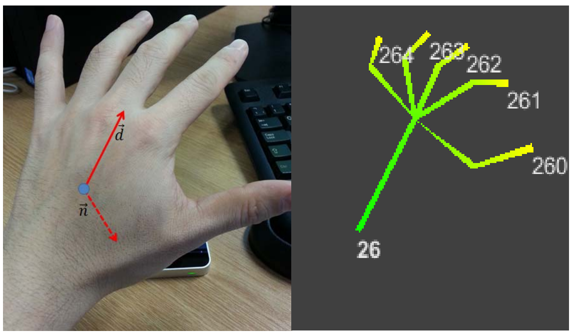
where P_{crow} is the puppet crow's position in current frame, R_{hand} is player's relative hand movement in real world in current frame comparing with the previous frame, which can be obtained in formula (2), M is the transformation matrix between the player's workspace and the puppet crow's coordinate system, S is the scaling matrix, and P'_{crow} presents puppet crow's position in the previous frame.

Hand model is shown in the right of Figure 6(a), where $\vec{f_i}$ ($i \in [1,5]$) presents the position of the finger tips (i is the number of recognized fingers), and $\vec{c}$ presents the centre of the palm. The plane of the hand is formed with the normal vector $\vec{n}$, and the directional vector $\vec{d}$. Vector $\vec{n}$ and $\vec{d}$ presents the normal and the directional vector of the hand plane. The player's hand and its virtual skeleton mapped are shown in Figure 6 (b).

For ease of controlling by young children, we define four intuitive motion controls: right, left, move downward and upward, which are mapped to different hand gestures, as shown in the upper two rows in Figure 1(b).



(a)



(b)

Fig. 6. (a) Hand skeleton model; (b) Player's hand and its virtual representation

2) Hand gesture recognizing

For young children, in game environment, using hand gestures as in input to control avatar to perform is more natural and intuitive than other HCI method, such as using keyboard and touching screen. Unlike a screen-based system, a hand gesture input is clearly visible to others and constitutes an expressive action in itself.

When designing the system, a most important consideration is choosing a most natural and intuitive gestural interaction manner to play with the avatar. Considering their simplicity and demonstrated effectiveness, we utilize detectable pointing gestures [23] into HCI to reduce the operation difficulty of young children. To reduce the degree of difficulty of young children's operation, some simple gestures are recognizable in our system, such as grip and stretch, as shown in the bottom row of Figure 1(b). Once a recognizable gesture is detected, recognizing function will call a pre-recorded animation of puppet crow linked with this gesture. And then, puppet crow will act as a real puppet and perform a pre-set action responding to player's hand gesture in the virtual environment. The scenario of player's using finger gestures to control the avatar is shown in Figure 1(c).

For the purpose of motor coordination training, players are required to perform hand gestures correctly, move hand smoothly and locate accurately, which contains two different aspects: hand movement and hand gestures. In our system, the data model of hand motion can be represented as follows:

$$Mot = \langle Mov, Ges \rangle \tag{5}$$

Hand movement includes move up (m_1), move down (m_2), move right (m_3) and move left(m_4).

$$Mov = \langle m_1, m_2, m_3, m_4 \rangle \tag{6}$$

And hand gestures include stretch (g_1) and grip (g_2).

$$Ges = \langle g_1, g_2 \rangle \tag{7}$$

Let J be the set of joints of the players' hand being tracked.

$$J = \langle j_1, j_2, j_3, \dots j_n \rangle \tag{8}$$

Each joint j_i has a number of motions associated with it which can be represented by M.

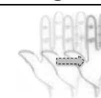
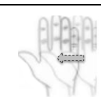
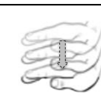
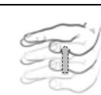
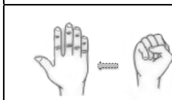
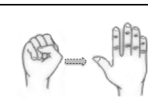
Then, each gesture g_i can be expressed as follows:

$$g_i = \langle j_i, m \rangle \tag{9}$$

where $j_i \in J$ is a joint, $m \in M$ is a motion associated to j_i.

With the definition of hand motion data model, we have designed a set of hand gestures which provide a natural and intuitive way for young children to interact with the playthings. Table I shows the mapping between hand gestures performed by players and the puppetry crow's action in virtual environment.

TABLE I. HAND GESTURES MAPPING

Hand gestures	Motion	Crow action
	Move right	Fly to the right
	Move left	Fly to the left
	Move down	Fly down
	Move up	Fly up
	Stretch	Hover
	Stretch to grip	Grasp pebble/ stick
	Grip to stretch	Drop pebbles/ stick

D. Output

Output module updates puppet crow's position calculated in motion control module and plays pre-recorded animations which have been linked with recognizable gestures in motion control. This module provides corresponding response to players as the feedback to adjustment of hand movement, which is vital for our cognitive development purpose as well as motor coordination ability training.

IV. RESULTS

The plot of "The Crow and the Pitcher" is presented first as the hint of the story, as illustrated in Figure 7. A scenario of Puppet Narrator during playing is shown in Figure 8.

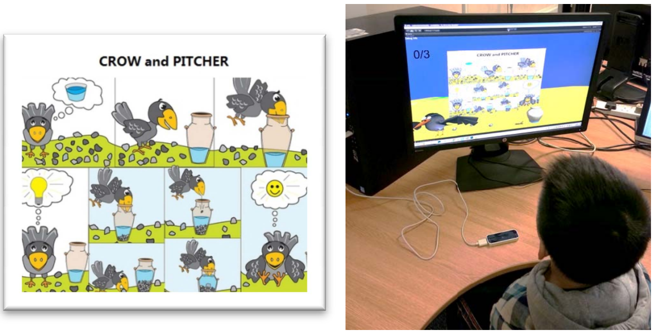


Fig. 7. Story plot*

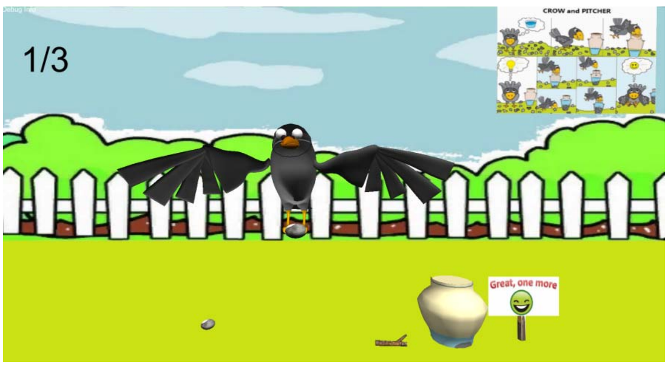
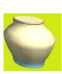


Fig. 8. A scenario of Puppet Narrator

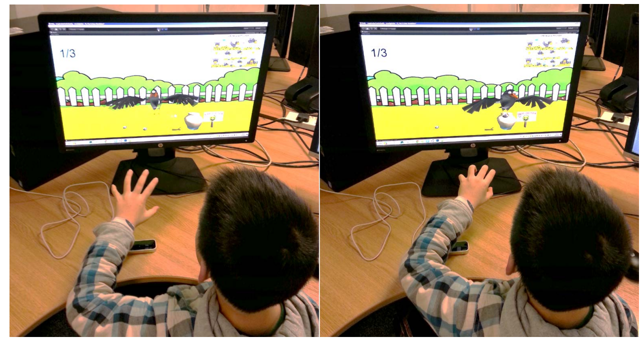
In the scenario, there are five kinds of virtual items, each of which has different function, as shown in Table II.

TABLE II. VIRTUAL ITEMS

Icons	Name	Function
	Pebble	Fill pitcher for water rising
	Stick	Confuse player which needs to be differentiated from pebbles
	Pitcher	Collect pebbles dropped by crow and feed crow water after three successful drops
 Sorry, unreachable / Great, one more / Congratulations	Signboard	Provide information to player
1/3	Counter	Calculate how many pebbles are in dropped in pitcher

In Puppet Narrator, young children can use their fingers to control the animation of the puppet. In Figure 9(a), we can see that the player stretches his hand and move it to control the movement of the puppet crow by mapping the palm position to the crow's position. In Figure 9(b), we can see that the crow has grasped a pebble successfully and is preparing to drop it into the pitcher.

* Fables of the world (https://www.pinterest.com/samargardezi/fables-of-the-world) accessed on 20 Mar, 2015.



(a) Stretch (b) Grip

Fig. 9. Finger gesture

The storytelling procedure finished by a 7 year-old boy is recorded in Figure 10.

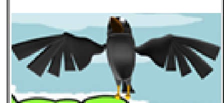
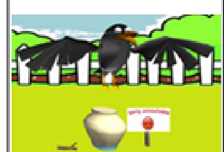
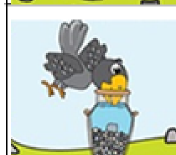
Plot	Virtual scene	Narrative record
		A crow is flying around on a hot summer day looking for water…
		He comes across a pitcher. Oops! He can't reach the water.
		He has an idea: collect pebbles and drop in the vase
		The crow drops pebbles in the pitcher. Water level rises.
		He takes lots pebbles and drops them into the Pitcher… He can drink water now!!

Fig. 10. Storytelling performed by a 7 year-old child

V. Discussion

Four 5-8 years old young children are invited to participate in our preliminary test. With the instruction of provided story plot, all of the children could finish the story narrative and perform with great interest. During playing, some minor frustrations (or difficulties) are observed, such as picking up the stick by mistake, having problem to move the avatar toward the pitcher, or using a second hand to support the main hand in operation etc. Getting around these difficulties also made the children feel rewarded and find the game interesting.

After twenty rounds of experiment (five for each child), we have noticed that:

All the children could finish the game with few faults after a couple rounds of repeating.

All the children could narrate the plot with more complex words through their own language and understanding in the final three rounds.

For evaluation purposes, we set up some metrics from different aspects of our training procedure, and the result of each round of trial is recorded. The trend plot of average values of metrics is presented in Figure 11 and the details are shown in table III.

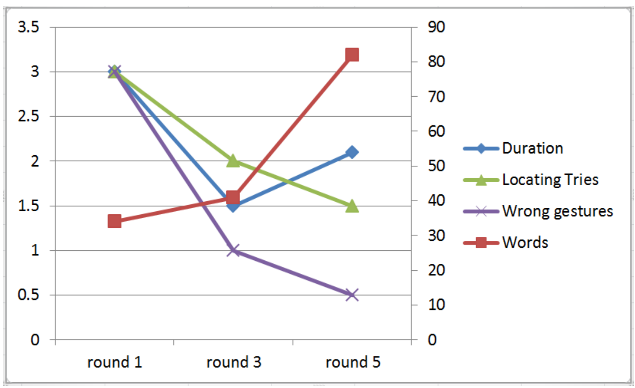


Fig. 11. Average values of metrics in experiment

TABLE III. EXPERIMENT RESULT

Metrics	Round 1	Round 3	Round 5
Duration of round (*M1*)	3 minutes	1.5 minutes	2.1 minutes
Narrative complexity (*M2*)	32 words	41 words	82 words
Number of tries to locate pebble (*M3*)	3 tries	2 tries	1.5 tries
Number of tries to pick up stick (*M4*)	1 try	0 tries	0 tries
Hand movement (*M5*)	disorder	smooth	smooth and rapid
Wrong hand gestures (*M6*)	3	1	0.5

Both Figure 11 and table III shows the quality of improvement of young participants' performance over several rounds of trials. Ability to successfully locate the pebble (*M3*) is improved by practicing picking up or dropping pebbles which can be indicated by the decline of locating tries (shown by green line in Figure 11). The average number of tries of picking up the stick by mistake (*M4*) drops from 1 to 0, which means that children can distinguish the stick from pebbles after 2 rounds of training. Hand movement (*M5*) and wrong hand gestures (*M6*) in the later stages of the experiment are improved over the earlier attempts.

We noticed from the experiment that young children were able to tell the story confidently and use more words and longer sentences in the last two rounds of trials, with some encouragement. One typical example is the sentence created by a 7 year-old boy at the start of the story: in the first trial, the boy used the sentence "A crow is looking for some water." In the third trial, he rephrased the sentence to "A thirsty crow is flying around looking for water." After adding more decorative words, the sentence finally transformed to "A thirsty crow is flying around on a hot summer day looking for water." It is obvious that the narrative ability was enhanced during storytelling, which is reflected by metric *M2* (red line in Figure 11). One interesting thing we found is change of duration of narrating the story (*M1*). Initially, participants used 3 minutes on average to complete the whole narration and then they could finish it within 1.5 minutes, which benefits from getting more familiar with the system. Finally, the duration of narration tends to last a litter longer. The reason is that once young children get used to the control mechanism and the interaction interface, they tend to pay more attention on the story line and the improvisation, which is vital for storytelling and more suitable for pedagogical purpose as a learning tool for children.

During the test, the aspects of children's cognitive competence and motor coordination ability were trained with obvious improvements.

Numerical cognition: all the three pebbles need to be dropped into the pitcher before the crow can drink water, which requires the players should have a basic numerical cognition to finish the mission.

Spatial cognition: pebbles could only be picked up within predefined area and dropped into the pitcher, which means the spatial cognition ability is required in locating the pebble.

Visual perception: there are several kinds of virtual items in the scenario and each of them has different property and usage, which means the players need to distinguish them from others and make correct choices, for example, only pebbles can make water to mount up and the stick will not work.

Motor coordination: only predefined hand gestures could be recognized by our system, which means players need to perform hand gestures correctly. For controlling the puppet crow, players need to use hand motions as the input to manipulate the crow in virtual environment and adjust hand gestures according to a visual feedback received from the crow's responding movement. This requires players to move their hands smoothly and steadily, which reflects a better motor coordination of limbs. During the performance, players' narration often incorporates with hands movement and changing gestures, which also requires perfect coordination.

We plan to exploit our system to conduct indicative experiments with more subjects to identify the pros and cons with quantitative data and analysis. We will also include other story templates in the game. This will introduce new gestures and pose additional operational complexity for young children.

VI. Conclusion

We have presented a novel digital storytelling system assisted with virtual puppetry, Puppet Narrator, providing young children with natural interaction/control and immersive experience when narrating story. The system is designed to support training of different cognitive skills and motor coordination through storytelling. It has been a novel attempt to include advanced motion sensing technology and computer animation as a medium for this development.

Besides its positive pedagogical significance as a digital storytelling system, our system also helps develop young children's skills and knowledge related to ICT.

The usability of the system is preliminary examined in our test and the results from the analysis are promising, which showed that young children can be benefited from their playing with Puppet Narrator. However, as only a limited number of subjects are tested, we will need to examine more cases and design a psychological experiment to affirm the conclusion. Further validation and analysis of the effectiveness of this approach is needed, but at the moment our observation and analysis can be supported with success of other parallel development in digital story telling [8].

The story telling in our test is a supervised learning process where an adult will provide guidance and helps, which is important for the young ones can accomplish their narrative task and receive proper trainings. Without the presence of supervision, it is possible that the virtual puppet may distract the story telling that the child would focus on the playing and controlling of the puppetry without practicing their narration. It will request a rewarding strategy in the future development of such system to automatically encourage and reward the players when they accomplish the narrative task properly.

ACKNOWLEDGMENT

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement n° [623883] - "AniM". The authors also acknowledge partial support from the FP7/2007-2013 ("Dr. Inventor", FP7-ICT-2013.8.1 611383) and the National Natural Science Foundation of China (NSFC, 31200708).

REFERENCES

[1] P. Backlund and M. Hendrix, "Educational games - are they worth the effort? A literature survey on the effectiveness of serious games," Games and Virtual Worlds for Serious Applications (VS-GAMES), 2013 5th International Conference on. IEEE, 2013.

[2] F.L. Greitzer, O.A. Kuchar and K. Huston, "Cognitive science implications for enhancing training effectiveness in a serious gaming context," ACM Journal on Educational Resources in Computing, 7(3), 2007.

[3] E. M. Pedersen, Storytelling and the art of teaching, In English Teaching Forum, 33(1), 1995, pp. 2-5.

[4] More, Cori, "Digital stories targeting social skills for children with disabilities multidimensional learning," Intervention in School and Clinic 43.3, 2008, pp. 168-177.

[5] N. Di Blas and B. Boretti, "Interactive storytelling in pre-school: a case-study," In Proceedings of IDC 2009, ACM, NY 2009, pp. 44-51.

[6] N. Di Blas, P. Paolini and A. Sabiescu, "Collective digital storytelling at school as a whole-class interaction," Proceedings of the 9th international Conference on interaction Design and Children. IDC '10. ACM, New York, NY, Indiana UP, 2010, pp. 11-19.

[7] D Eder, Life Lessons Through Storytelling: Children's Exploration of Ethics, Indiana University Press, 2010.

[8] P. Psomos and M. Kordaki, "Pedagogical analysis of educational digital storytelling environments of the last five years," Procedia-Social and Behavioural Sciences, 46, 2012, pp.1213-1218.

[9] R. D. Greer, "The teacher as a strategic scientist: A solution to our education crisis?" Behavior and Social Issues 1 (2), 1991, pp.25-41.

[10] R. D. Greer, "Designing teaching strategies: An applied behavior analysis system approach," 1st Edition, Academic Press, 2002.

[11] B. F. Skinner, "The Technology of Teaching," Appleton-Century-Crofts, 1968.

[12] E. A. Vargas, J. S. Vargas, "Programmed instruction: What it is and how to do it," Journal of Behavioral Education 1, 1991, pp.235-251.

[13] J Wai, D Lubinski and C.P. Benbow, "Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance," Journal of Educational Psychology 101(4), 2009, pp.817.

[14] S. Franceschini, S. Gori, M. Ruffino, K. Pedrolli, and A. Facoetti, "A causal link between visual spatial attention and reading acquisition," Curr Biol, 22(9), 2012, pp. 814-819.

[15] L.Feigenson, S. Dehaene and E. Spelke, "Core Systems of Number," Trends in Cognitive Sciences, 8(7), 2004, pp. 307-314.

[16] D Ness and S.J. Farenga, Knowledge Under Construction: The Importance of Play in Developing Children's Spatial and Geometric Thinking, Rowman & Littlefield Publishers, 2007.

[17] F Simion, L Regolin and H Bulf, "A predisposition for biological motion in the new born baby," Proceedings of the National Academy of Sciences 105(2), 2008, pp. 809-813.

[18] S.A. Gelman, The Essential Child: Origins of Essentialism in Everyday Thought, Oxford University Press, 2003.

[19] Best, John B., Cognitive Psychology, 5th ed. 1999, pp. 15-17.

[20] Skinner, Rosemary A., and Jan P. Piek, "Psychosocial implications of poor motor coordination in children and adolescents," Human movement science, 20(1), 2001, pp.73-94.

[21] The greek anthology. Vol. 1. W. Heinemann, 1916.

[22] Avianus 27, Retrieved on 2007-07-19.

[23] LEAPMOTION API, Available from http://www.leapmotion.com/.

[24] T. Beardsworth and T. Buckner, "The ability to recognize oneself from a video recording of one's movements without seeing one's body," Bulletin of the Psychonomic Society, 18(1), 1981, pp. 19-22.

[25] G. Knoblich and R. Flach, "Predicting the effects of actions: Interactions of perception and action," Psychological Science, 12(6), 2001, pp. 467-472.

[26] G. Knoblich and W. Prinz, "Recognition of selfgeneratedactions from kinematic displays of drawing," Journal of Experimental Psychology: human perception and performance, 27(2), 2001, pp. 456.

[27] A. Mazalek, M. Nitsche, S. Chandrasekharan, T. Welsh, P. Clifton, A. Quitmeyer, et al, "Recognizing self in puppet controlled virtual avatars," In Proceedings of the 3rd International Conference on Fun and Games ACM, 2010, pp. 66-73.

[28] H. Von Kleist and T.G. Neumiller, On the marionette theatre, The Drama Review: TDR, 1972, pp. 22-26.

Back to Contents

A literature review of gamification design frameworks

Alberto Mora*, Daniel Riera*, Carina González†, Joan Arnedo-Moreno*
*Estudis d'Informàtica, Multimedia i Telecomunicació
Universitat Oberta de Catalunya. Barcelona, Spain
{amoraca,drierat,jarnedo}@uoc.edu
†Ing. de Sist. y Aut. y Arq. y Tec. de Computadores
Universidad de La Laguna. La Laguna, Spain
{cjgonza}@ull.edu.es

***Abstract*—This paper presents a review of the literature on gamification design frameworks. Gamification, understood as the use of game design elements in other contexts for the purpose of engagement, has become a hot topic in the recent years. However, there's also a cautionary tale to be extracted from Gartner's reports on the topic: many gamification-based solutions fail because, mostly, they have been created on a whim, or mixing bits and pieces from game components, without a clear and formal design process. The application of a definite design framework aims to be a path to success. Therefore, before starting the gamification of a process, it is very important to know which frameworks or methods exist and their main characteristics. The present review synthesizes the process of gamification design for a successful engagement experience. This review categorizes existing approaches and provides an assessment of their main features, which may prove invaluable to developers of gamified solutions at different levels and scopes.**

***Index Terms*—Gamification, frameworks, game design elements, game design methods, game, review.**

I. Introduction

Games have been present in all human civilizations. Human beings have even been defined as *homo ludens*, a concept proposed by Huizinga in 1955 [1]. Based on this idea, he infers the concept of *game* as a free activity standing quite consciously outside of "ordinary" life, as being "not serious", but at the same time intensely absorbing to the player. The evolution of digital technologies, especially in its path from traditional games to video-games, has been essential in the growth of user enjoyment and engagement, as prove Brumels et al. (2008) [2]. In fact, with the widespread adoption of social media and mobile technology, the presence of games in our daily lives is more than an obvious fact in the 21st century society. Jesse Schell (2010) [3], a prestigious game designer, presents a hypothetical future where video games are part of our lives. It is a process with a point of no return.

Based on these precedents and considering the omnipresence of games, and therefore, the interiorization of game mechanics by society, gamification arises almost organically as a way to extract characteristics from games in order to incorporate them into other environments. A first approach comes from Nick Pelling in 2002 [4] defined as the application of game-like accelerated user interface design to make electronic transactions both enjoyable and fast. However, the term has much evolved since then, encompassing different aspects of game experience and design. Probably, the most widespread definition in the literature comes from Deterding et al. in 2011 [5], as "the use of game design elements in non-game contexts".

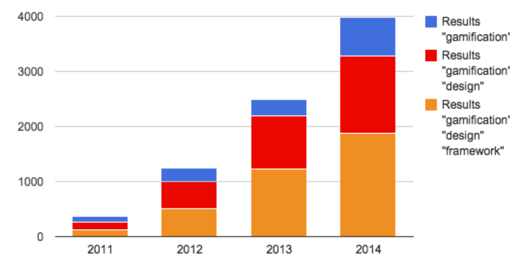


Fig. 1: Results of academical searches about gamification

Beyond the definitions and experiences, the application of gamification techniques in different contexts has increased in the last years, becoming a promising trend in many areas. Just a look at the emerging technologies hype cycle published by Gartner in 2013 [6] can help us to realize that the term had just reached the top of the wave. So much that M2 research 2011 predictions indicated that the gamification market would reach 2.8 billion dollars by 2016 [7]. However, Gartner also predicted that, by 2014, 80 percent of the gamified applications would fail to meet their business objectives, primarily due to poor design [8]. In Gartner's 2014 report [9], gamification was consequently moved towards the so called "Through of Disillusionment", indicating that it will take from 5 to 10 years to stabilize and reach what the "Plateau of Productivity", when the approach is finally considered mature.

The moral of the story from Gartner's point of view is obvious: a clear design strategy is the key to success in gamification. On that regard, as shown by the Google Scholar aggregate results per year and keywords in Figure 1, the community's interest in gamification design and frameworks is evident. This popularity encompasses all kinds of contexts: education and training, human resources, marketing, sales, health, etc.

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

The main goals of this paper are threefold: first, completing a state of the art on the gamification design process. Second, analysing the relationship between the gamification and game design processes. And third, identifying the existing gamification design frameworks and classifying them by their main features. Our contribution to this analysis is in the assessment of the shortcomings and the principles that are not being applied and may lead to failure.

This paper is structured as follows. Some background on the principles of game design and how they apply to most gamification approaches is presented in Section II. In Section III, a thorough review including a classification of gamification frameworks is developed. Finally, Section IV is devoted to sum up the conclusion of this work and the research questions and answers.

II. Game Design Theoretical Background

The purpose of gamification design elements is quite different from game design, the former being used to enhance the engagement in different contexts, whereas the latter is directed towards pure entertainment. Marczewski (2014) [10] makes an explicit distinction between game and gamification design and its features. First, the most common start for a game design is the basic idea of enjoyment, while gamification points towards a business objective. Secondly, the definition of metrics or game lines must happen in different stages of the design process.

Creating a gamified system will always be different to creating a game as a general process, although there is a thin connection (not well defined yet) between game and gamification design. Nevertheless, the basics of gamification heavily rely on the principles of game design theory. In this section, it has been provide a brief theoretical background on this topic, necessary to understand some of the common properties that can be found in most gamification design frameworks.

In the game context, what is known as "game design" could be proposed in a simple manner as "the action of making sense of things related to a game". This definition is not so far from Schell's description (2008) [11]: "the act of deciding what a game should be". In this regard, Salen and Zimmerman (2004) defined a set of game design fundamentals principles, which should be run using an iterative process [12]:

- Understanding design, systems, and interactivity, as well as player choice, action, and outcome.
- Including a study of rule-making and rule-breaking, complexity and emergence, game experience, game representation, and social game interaction.
- Adding the powerful connection between the rules of a game and the play that the rules engender, the pleasures games invoke, the meanings they construct, the ideologies they embody, and the stories they tell.

By improving upon these principles, Brathwaite and Schreiber (2009) [13] assert that, once the different elements of games have been identified, it is necessary to reflect about how to incorporate them. Since a chemical perspective, they define *game atoms* as "the smallest parts of a game that can be isolated and studied individually". Therefore, from an atomic point of view, the process of designing games as using a collection of atoms becomes clearer. This idea is used by Reeves and Red (2013) [14], that introduce ten *ingredients* to make a successful game design: self-representations, three-dimensional environments, narrative, feedback, reputations ranks and levels, marketplaces and economies, competition under rules, teams, communication and finally time pressure.

Once the game elements are already condensed into game design fundamentals, a standardized concept, practices and criteria are necessary for assembling them rationality under a framework's definition. Typically, a framework is a real or conceptual structure intended to serve as a support or guide for the building of something that expands the structure into something useful (taken from glossary). Nevertheless, it must be noted that some authors, such as Crawford (1984) [15], have concluded that game design is an activity too complex to be reducible to a formal procedure. In this regard, Julius and Salo (2013) [16] assert that it should be treated as an agile process that does not always follow a specific design framework (although they propose one of them).

The need for a formal and recognized proposal in game design contexts led to the development of the MDA (Mechanics, Dynamics and Aesthetics) framework, by Hunicke et al. in 2004 [17]: a formal approach to understanding games, which attempts to bridge the gap between game design and development, game criticism, and technical game research". According to this framework, games can be broken down into three elements: rules, system and fun. These elements are directly translated into the following design components, which must be defined when designing a game using this same order:

- *Mechanics*, describing the particular components of the game, at the level of data representation and algorithms.
- *Dynamics*, describing the run-time behaviour of the mechanics acting on player inputs and each others outputs over time.
- *Aesthetics*, describing the desirable emotional responses evoked in the player when interacting with the game system.

Thus, from a perspective of game experience, a model is only a fraction of the whole as proposed Cavillo-Gamez (2010) [18] in his "Core Elements of the Gaming Experience (CEGE)". He recites a set of necessary but not sufficient conditions to provide a positive experience while playing which must be considered in the design process: interface design pattern, design patter and dynamics, design principles and heuristics, models (i.e MDA and design methods). Thus, Zichermann and Cunningham (2011) [19] argue that game and user experience designers have been implementing these techniques for decades to create addictive games and engaging player experiences. Globally, Deterding et al. [20] describe the necessary game design actions for *gamefulness* in a set of levels: game interface design patterns, game design patterns and mechanics, game design principles and heuristics, game models and game design methods.

In summary, it has been showed the relevant game design features from an atomized breakdown like lens or ingredients until formal descriptions and models like MDA, usually proposed in the gamification design process.

III. Literature Review

Gamification as a concept brings together many disciplines and professionals including game designers, UX/UI designers, psychologists, sociologists, computer engineers and others. Our interest is focused on the gamification design process from a formal perspective and keeping what role each professional plays in mind. Based on the background described in Section II and taking into account that the gamification design process naturally assumes many game designing principles, we ask ourselves the following questions:

Q1: Which gamification design frameworks are available now in the literature and which are their main features?

Q2: Do the gamification frameworks inherit game design principles for their development?

Q3: What design considerations are not being applied, or only to a lesser extent, by the gamification designers?

A. Methodology

A literature review of works mainly indexed in Google Scholar, Scopus, and Web of Knowledge, was conducted. The keywords were *gamification, game, design, framework, models, methods* and *engagement*. The search of the literature on gamification design and frameworks was not established within a period of time although most of them date from the last four years. Case reports, review articles and studies found by keywords and the references taken from bibliography were short-listed, as well as frameworks definitions. A total of twenty-two candidate frameworks were initially reviewed, however, only eighteen of them met the study requirements. These requirements refer to a definition of a formal structure for a skeletal support used as the basis for something which is being constructed (partial or complete gamification design process for engagement purpose in generic or business contexts). Notice that in some cases for a single framework, multiple sources have been checked and even we have contacted their authors to clarify some aspects.

After studying the frameworks, they have been categorized according to a three-dimensional perspective as shown in Figure 2.

- Background: academic and non-academic.
- Scope: complete gamification processes and focused only on a specific part or step.
- Approach: applicable to a wide spectrum of environments (generic) or designed for an specific business context.

B. Results

This subsection summarizes the main properties of current gamification frameworks according to our literature review.

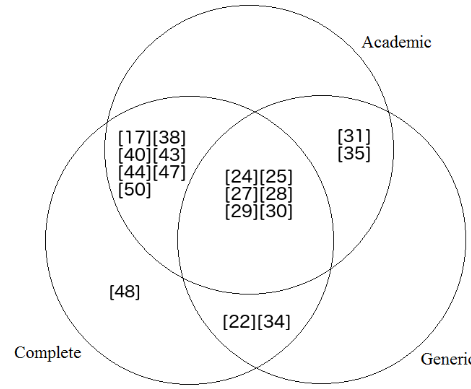


Fig. 2: Gamification framework's categorization

As follows, the frameworks have been split between two categories, being sorted by time, background and scope.

Generic frameworks

Di Tomasso (2011) defines a framework for Success [21] based on the Self-Determination Theory by Ryan and Deci (2000), known as SDT [22]. From a knowledge of individual player differences and social influences, he proposes the following steps: discover the reason to gamify (stakeholders and business objectives), identify players' profiles and motivational drivers, set up goals and objectives, describe skills, track and measure, define lenses of interest, desired outcomes (thanks to feedback and establishing the epic win state) and play-test, and *polish*.

However, the best-known design framework is presented in Six Steps to Gamification [23] by Werbach and Hunter (2012) and commonly known as 6D. This framework starts from a definition of business objectives and then proceeds to target the expected behaviours, describes the players, devises the activity loops without forgetting the fun, and finally, deploys the gamification system with the appropriate tools. Although not in a explicit way, takes a slight influence of Hunicke et. al's MDA game design framework. This can be seen in the *Pyramid of Gamification Elements*, which proposes the following relevant elements: mechanics, dynamics and components. It is the basis for several other gamification design frameworks.

Meanwhile, a simpler framework, called GAME [24], is proposed by Marczewski in 2012. It is based on two phases. Firstly, planning and designing, which includes the gathering, by means of a survey, of key information such as the users' types in the gamification context (Hedax user type, based on Bartle's (2005) [25]). Afterwards, the best solution for goals and engagement is designed, measuring user activities and outcomes. He applies an own motivation framework called RAMP (Relatedness, Autonomy, Mastery, Purpose). The design must be enriched over the time. Updates of this framework were published in different media afterwards and have been incorporated to this review after contacting the author.

Moreover, Marache-Francisco and Brangier (2013) define a

Gamification design process [26] based on Human-Computer Interaction (HCI) principles. They identify several dimensions outside the gamification components and practices which can be used to define a clear framework. Three dimensions are described: sensory-motor dimension, motivation emotion and commitment, and cognitive dimension of interaction. Based on these, the design process consists on two major iterative steps: the context analysis (*User-Centered Design*) and the iterative conception of the gamification experience. Moreover, a toolbox for gamification (named *Core Principles*) to help designers through the process is referenced.

On the other hand, De Paz (2013) proposes a set of steps or general guidelines to gamification [27] which can be applied to any type of project. The values of this framework seem highly influenced by Werbach and Hunter's Six Steps to Gamification. The proposal's guidelines are divided into three phases: the setting up of the business goals (preparation), the determination of the basic designing and the use of game elements. Implementation and maintenance consists on building the system and run it. This approach also recommends the use of metrics.

Another proposal comes from Robinson and Bellotti's taxonomy (2013) [28], who claim that different frameworks in the literature can be helpful for a gamification design, but they do not exactly meet their requirements. As the authors say: "*do not provide a concise, time-saving but reasonably comprehensive presentation of common gamification elements in terms of the various aspects of the user experience that they support*". They stablish six categories of gamification elements inspired from several valid sources on the literature. These categories are general frames, general rules and performance frames, social features, incentives, resources and constraints, and finally feedback and status information.

At this point, the approach in Francisco-Aparicio et al.'s framework (2013) [29] allows, on the one hand, to determine the type of game mechanics activities should incorporate to meet the psychological and social needs of human motivation (SDT). On the other hand, it aims to assess the effectiveness of the gamification process, based on the fun criteria: the properties characterising the playability and the degree of improvement in obtaining satisfactory results using a quality service mode. In this framework, games are divided into three parts (from their functional perception): game core, engine and interface. The essential activities proposed are: end-user analysis, main objectives and cross-cutting identification, implementation and analysis of the effectiveness.

Focused only on the ethical perspective, Versteeg (2013) defines a simplified framework for moral persuasive gamification design [30]. This combines a normative ethical framework (moral design) with the most relevant issues of the following methodologies. It is based on the moral design framework by Berdichevsky and Erik Neuenschwander (1999) [31] and its ethical golden rules that a designer should never exceed. Moreover, it incorporates a methodology for analysing the ethics of persuasive technologies like that proposed by Fogg (2002) [32]. The steps are: definition of moral principles and values, conceptual investigation, involvement the stakeholders, and evaluation and iteration.

Additionally, a Complete Gamification Framework called Octalysis (2013) [33] is proposed for Yu-kai Chou. For his point of view, the gamification is design that places the most emphasis on human motivation in the process. In essence, it puts on a Human-Focused Design (as opposed to function-focused design to get the job done quickly). The approach is based on an octagon shape with eight core drives represented by each side: epic meaning and calling, development and accomplishment, creativity and feedback, ownership and possession, social influence and relatedness, scarcity and impatience, unpredictability and curiosity and loss and avoidance.

To end this section about generic frameworks, we find Al Marshedi et al.'s propose (2015), "A Framework for Sustainable Gamification Impact" [34]. This approach aims to increase the sustainability of the desired impact of gamified applications. It is mainly based on three backgrounds: Csikszentmihalyi's Flow Dimension Theory (1990) [35], Pink's drive motivation elements (2011) [36] and SDT. Furthermore, it is focused on User-Centred Design (UCD). As the authors claim, it is a way to integrate purpose, mastery, relatedness and flow to competence and time; being as a guideline for designers that want to create relevant experiences that people will be engaged to in the long-term.

Business-specific frameworks

Purely for a business purpose, J. Kumar (2013) describes the "Player Centered Design Methodology" [37] as a practical guide for user experience designers, product managers and developers to incorporate the principles of gamification into their software. This approach is useful for the enterprise context and for specific applicability. The methodology is based on a *Player-Centered Design* (2004) [38], a related common point of view in other frameworks. The process focuses on good understanding of both the player and the mission. The following eight steps are described: understanding the player, understanding the mission, understanding human motivation, applying game mechanics, setting the game rules, defining engagement loops, managing-monitoring-measuring and considering legal and ethical issues.

Thus, the "Role-Motivation-Interaction Framework of Gears" (2013) [39] is a proposal based on the Constantine and Lockwood (1999) model and method of usage (software for use) [40] from UCD. Basic desires described by Reiss (2002) [41] are applied to the gamified system development process. This framework is based on a predefined architecture in order to make the process easier and provide a set of rules that cannot be broken. The recommended aspects to be considered for the design process are the description of the goals, objectives, business rules, behavioural norms, preconditions, actors and the course of these actions (gameful interactions).

In this regard, a "Gamification Framework" is proposed by Jacobs (2013) [42] for implementing *enterprise level gamification* within an organization. A good knowledge of the requirements determines the success of the gamification model, considering gamification as a *fluid subject*, constantly changing and evolving. This framework is based on a Goal-Model Design, distinguishing between short and long term goals. Several considerations are taken into account: understanding the goals and impact, defining the goals, considering user and social media, feed-backing and compilation of data for analysis, and finally, running the loop engagement.

Additionally, Julius and Salo (2013) propose a concrete framework for gamification [16] in the business context, focusing exclusively on a marketing environment. The authors consider an agile design process (which not always must use a design framework) created from a literature review and tested with an empirical study. Concretely, it was designed from the whole of Werbach's proposal, taking into account some of the special features in the marketing sector. This proposal inserts an additional third stage, called *market research*.

On the other hand, Li (2014) proposes the "Theoretical Model for Gamification in Workplace IS context" [43]. From an IT perspective, it is a theoretical framework for the process of gamification design and implementation in a workplace within an Information System (IS) environment. The model is based on the Technology Acceptance Model (1989), known as TAM [44], which deals how users come to accept and use the technology. According to Delone and McLean's Success model [45] takes an IS success measurement synthesis in order to provide a guidance to the future.

Moreover, A "Framework for Designing Gamification in the Enterprise" [46] is defined by N. Kumar (2013) as a prescriptive method for designing a gamification environment for the enterprise. As the author says, the process of gamification is very complex and involves multiples stages. Therefore, a framework is proposed to guide designers from concept to implementation and improvement. This approach is divided into three phases. The first phase includes the objectives definition, challenges and motivations understanding, and challenges management. The second phase, game design, includes the creation of the narrative, game mechanics, and the interface. The process concludes with the implementation of the gamification environment and its assessment in the third phase.

A business centred approach in our literature review can be found in the "Gamification Model Canvas" [47], a framework proposed by Sergio Jiménez (2013). It is an agile, flexible, and systematic tool to find and evaluate play based solutions in order to develop certain behaviours in non-game environments. It is based on the Business Model Canvas [48] design and the MDA game design framework. The result of this marriage is a new framework where a set of elements must be considered for the gamified design process: revenues, players, behaviours, aesthetics, dynamics, components, mechanics, platforms and costs.

Finally, related with Li's perspective, Herzig (2014) describes gamification development [49] as a Technology-Centred Design process. His approach is based on RUP (Rational Unified Process), an iterative software development process framework, adapted to the gamification domain. His proposal aims to visualize how gamification is introduced stepwise into arbitrary information systems, starting at the business modelling phase, and ending at the monitoring and improvement phases. This approach also considers roles definition (end-user, gamification, domain, business and, IT experts) as necessary. Herzig describes the following phases to be considered: business modelling, requirements, iterative design, provisioning, implementation, testing, deployment and, monitoring.

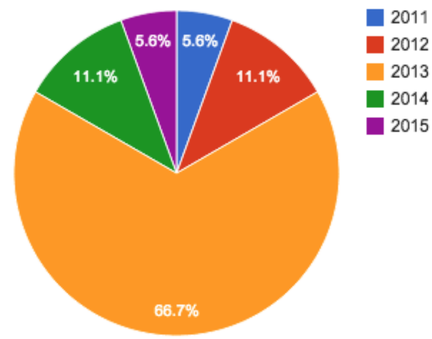


Fig. 3: Gamification design framework publish date

C. Assessment and discussion

Once the results of our literature review have been enumerated and described with a brief overview of each of the existing proposals (more detailed in Appendix A), it is possible to face our research questions.

Given the previous literature review, at this moment, we are able to answer the first question (Q1) proposed at the beginning of this review. Nowadays, there are a lot of original or based-on frameworks that try to formalise the design process from several point of view. We must also note that the publish dates of the literature on gamification frameworks are very recent, matching with the highest point of Gartner's Hype Cycle in 2013, as shown in Figure 3.

It is worth noting that most of the frameworks are based on a Human-Focused Design principles, taking into account the person as a main goal of the design. Psychological related aspects are very common items of great importance in most the frameworks proposed. Thus, SDT is a predominant approach for intrinsic motivation needs. This side is aligned with Zichermann's (2011) [19] theory who says that gamification is *75 percent of psychology and 25 percent of technology*. In addition, some of the frameworks are based on each other, as explicitly stated in their own definition or easily identifiable from careful reading. Werbach and Hunter's definition is the most referenced by other authors. At the same time, different approximations about game design principles are kept in mind for them, primarily the MDA game design framework described in Section II.

Thus, in order to provide a summary of the important properties in each proposal and better assess research questions

Q2 and Q3, it has been proposed a preliminary list of nineteen game design items taken from the literature, clustered and then organized into five categories:

1) Economic:

- Objectives: are the specific performance goals.
- Viability: a previous study, evaluation and analysis of the potential of applying gamification or refuse it.
- Risk: a probability or threat of damage, injury, liability, loss, or any other negative occurrence.
- ROI (Return On Investment): the benefit to the investor resulting from running a gamified experience.
- Stakeholders: a technique used to identify and keep in mind the people who have to interact with the design process.

2) Logic:

- Loop: the game mechanics combined with reinforcement and feedback in order to engage the player in the key system actions.
- End game / Epic win: a pre-established end of game or glorious victory in the system, usually stretching players to the limits of their abilities.
- On-boarding: the way of starting the new participants.
- Rules: the body of regulations prescribed by the designer.

3) Measurement:

- Metrics: the standards of measurement by which efficiency, performance, progress, process or quality.
- Analytic: the algorithms and data used to measure key performance indicators.

4) Psychology:

- Fun: the enjoyment or playfulness.
- Motivation: the behaviour which causes a person to want to repeat an action and vice-versa.
- Social: the interaction between players.
- Desired behaviours: the expected response of the players after the interaction.
- Ethics: a branch of philosophy that involves systematizing, defending and recommending concepts of right and wrong conducts.

5) Interaction:

- Narrative: the story and context created by designers.
- UI/UX: refers to everything designed into the gamified system which a player being may interact and the player's behaviours, attitudes, and emotions.
- Technology: the use or need of a software component for development.

All the items have been analysed and the ten most meaningful of them (in terms of results and heterogeneity) can be found in Appendix A. From this results, questions Q2 and Q3 can be answered. On the one hand, as previously seen, most of game design principles and components are being inherited for the gamification framework's description. Most of these items are present in lens of game design proposed by Schell (2008) [11], which is a world reference about game design and its components. So, mainly game design items are being used in the gamification process too. In the other hand, the way they are being applied is not the same as the game design environment. A set of new *new* steps or sequence is needed as Marczewski previously asserted.

By querying the table, Q3 answer can be inferred for the reader. Indeed, several aspects or factors are not being considered or extended by the authors.

IV. Conclusion

In this work, we have carried out a review of the literature on gamification design and developed frameworks. It has been analysed a set of eighteen gamified design frameworks according to a nineteen related items. Although some of these items are not very common in gaming context, we have considered as special interest for the gamification design process. Thus, we present the conclusions grouped as follows:

Economic issues are important for a few authors. Usually, terms as risk, viability or ROI are low referred. For more than a half, the participation of the stakeholders in the design process is necessary, in contrast to the other half which do not consider it. However, the definition of business objectives is widespread.

From a logical view, while the importance of loop item is extended in more than a half frameworks, on the other hand, less than a half of them consider the on-boarding and endgame actions (entry and exit way) as relevant in their approaches.

Additionally, measuring is a relevant issue for gamification, from a static or dynamic re-designing of the gamified experience by changing the status and the necessary immediate feedback. Most of frameworks refer explicitly the user data and the importance of collecting these data. But the use of metrics is not widespread in all over the approaches as a tool for quantify data.

Moreover, from a psychological perspective, we have perceived a high significance of this topic in almost all of the frameworks. They agree this approach as an essential key that must be present in the design process. It is an evidence topic in the literature. Most of them are Human-Based, taking the person as the centre of their design.

Thus interaction fundamentals are referred for more than a half of frameworks analysed, emphasizing the importance of the user interface, user experience, and the need or recommendation of taking a software for its development. A couple of them are focused on a Technological-Based Design or Goal-Based in contrast to the main Human-Based design.

To conclude, we consider, for a further work, a development of a complete and generic framework from a new perspective (not currently found in the literature) and its application to different environments. Educational and training, business (commercial, marketing, human resources), govern, health, and life-day are the most common scenes. We assume that current approaches are on the right way, but do not take into account some necessary keys to get a more effective gamified process for success.

Appendix A
Framework's feature summary table

In this appendix, a table with the whole list of frameworks proposed in this study (rows) compared with the top ten most suitable items of interest for the review (columns) is published. It includes, the categories: economic (viability/risk/ROI, and stakeholders), logic (loop, endgame/epicwin, onboarding, and rules), analytic (metrics), psychology (ethics), and interaction (UI/UX and technology). The possible values of each table's cell are:

- E: explicit, the item has appeared in the framework's definition.
- I: implicit, the item has not appeared explicitly in the framework definition. Inferred by the authors or referred inside an academic work of the author.
- U: unavailable, the item has not appeared anyway.

Acknowledgments

This work was partly funded by Agència de Gestió d'Ajuts Universitaris i de Recerca (Generalitat de Catalunya) through the Industrial Doctorate programme 2014-DI-006 and the Spanish Government through the project TIN2013-45303-P "ICT-FLAG" (Enhancing ICT education through Formative assessment, Learning Analytics and Gamification).

References

[1] J. Huizinga, "Homo ludens: A study ofthe play element in culture," *Trans. RFC Hull.] Boston: Beacon*, 1955.
[2] K. Brumels and T. Blasius, "Comparison of efficacy between traditional and video game based balance programs," *Clinical*, 2008.
[3] J. Schell, "Dice 2010: "design outside the box" presentation videos - g4tv.com," 2010.
[4] N. Pelling, "The (short) prehistory of gamification," *Funding Startups (& other impossibilities). Haettu*, 2011.
[5] S. Deterding, R. Khaled, L. Nacke, and D. Dixon, "Gamification: toward a definition," in *Chi 2011*, 2011, pp. 12–15.
[6] Gartner, "Gartner's 2013 hype cycle for emerging technologies maps out evolving relationship between humans and machines http://www.gartner.com/newsroom/id/2575515 (accessed january 2015)," 2013.
[7] M2, "Gamification market to reach $2.8 billion in 2016 http://www.gamesindustry.biz/articles/2012-05-21-gamification-market-to-reach-usd2-8-billion-in-2016 (accessed march 2015)," 2012.
[8] Gartner, "Gartner says by 2014, 80 percent of current gamified applications will fail to meet business objectives primarily due to poor design http://www.gartner.com/newsroom/id/2251015 (accessed january 2015)," 2014.
[9] Gartner, "Gartner's 2014 hype cycle for emerging technologies maps the journey to digital business http://www.gartner.com/newsroom/id/2819918 (accessed january 2015)," 2014.
[10] A. Marczewsky, "Gamification design vs game design http://www.gamified.uk/2014/03/25/gamification-design-vs-game-design/ (accesed february 2015)," 2014.
[11] J. Schell, "The art of game design: A book of lenses," 2008.
[12] K. Salen and E. Zimmerman, "Rules of play: Game design fundamentals," 2004.
[13] B. Brathwaite and I. Schreiber, "Challenges for game designers," 2009.
[14] B. Reeves and J. Read, "Total engagement: How games and virtual worlds are changing the way people work and businesses compete," 2013.
[15] C. Crawford, "The art of computer game design," 1984.
[16] K. Julius and J. Salo, "Designing gamification," *Marketing*, 2013.
[17] R. Hunicke, M. LeBlanc, and R. Zubek, "Mda: A formal approach to game design and game research," *AAAI Workshop on Challenges in Game*, 2004.
[18] E. Calvillo-Gámez, P. Cairns, and A. Cox, "Assessing the core elements of the gaming experience," *Evaluating User Experience in Games*, 2010.
[19] G. Zichermann and C. Cunningham, *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. "O'Reilly Media, Inc.", 2011.
[20] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: Defining gamification\," pp. 9\–15\.
[21] D. DiTommaso, "Beyond gamification: Architecting engagement through game design thinking http://www.slideshare.net/ditommaso/beyond-gamification-architecting-engagement-through-game-design-thinking (accessed january 2015)," 2011.
[22] R. Ryan and E. Deci, "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being." *American psychologist*, 2000.
[23] K. Werbach and D. Hunter, *For the win: How game thinking can revolutionize your business*, 2012.
[24] A. Marczewski, *Gamification: a simple introduction*, 2012.
[25] R. Bartle, "Virtual worlds: Why people play," *Massively multiplayer game development*, 2005.
[26] C. Marache-Francisco and E. Brangier, "Process of gamification. from the consideration of gamification to its practical implementation," *CENTRIC 2013*, 2013.
[27] B. Merino de Paz, "Gamification: A tool to improve sustainability efforts," Ph.D. dissertation, 2013.
[28] D. Robinson and V. Bellotti, "A preliminary taxonomy of gamification elements for varying anticipated commitment," *Proceedings of the CHI'13*, 2013.
[29] A. Francisco-Aparicio, "Gamification: Analysis and application," *New Trends in Interaction, Virtual Reality and Modeling*, 2013.
[30] M. Versteeg, "Ethics & gamification design: a moral framework for taking responsibility," 2013.
[31] D. Berdichevsky and E. Neuenschwander, "Toward an ethics of persuasive technology," *Communications of the ACM*, 1999.
[32] B. Fogg, "Persuasive technology: using computers to change what we think and do," *Ubiquity*, 2002.
[33] Y.-k. Chou, "Octalysis: Complete gamification framework http://www.yukaichou.com/gamification-examples/octalysis-complete-gamification-framework/ (accessed january 2015)," 2013.
[34] A. AlMarshedi, G. Wills, V. Wanick, and A. Ranchhod, "SGI: A framework for increasing the sustainability of gamification impact," 2015.
[35] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*, 1990.
[36] D. Pink, "Drive: The surprising truth about what motivates us," 2011.
[37] J. Kumar, "Gamification at work: designing engaging business software," 2013.
[38] D. Charles and M. Black, "Dynamic player modeling: A framework for player-centered digital games," *International Conference on Computer Games*, 2004.
[39] D. Gears and K. Braun, "Gamification in business: Designing motivating solutions to problem situations," *CHI'13 Workshop Designing Gamification*, 2013.
[40] L. Constantine and L. Lockwood, "Software for use: a practical guide to the models and methods of usage-centered design," 1999.
[41] S. Reiss, *Who Am I?: 16 Basic Desires that Motivate Our Actions Define Our Persona*, 2002.
[42] H. Jacobs, "Gamification: A framework for the workplace," Ph.D. dissertation, 2013.
[43] C. Li, "Evaluation of a theoretical model for gamification in workplace is context," 2014.
[44] F. Davis, R. Bagozzi, and P. Warshaw, "User acceptance of computer technology: a comparison of two theoretical models," *Management science*, 1989.
[45] W. Delone and E. McLean, "The delone and mclean model of information systems success: a ten-year update," *Journal of management information*, 2003.
[46] N. Kumar, "A framework for designing gamification in the enterprise," *Infosys Labs Briefings*, 2013.
[47] S. Jiménez, "Gamification model canvas http://www.gameonlab.es/canvas/ (accessed january 2015)," 2013.
[48] A. Osterwalder and Y. Pigneur, "Business model canvas," *Self published*, 2010.
[49] P. Herzig, "Gamification as a service," Ph.D. dissertation, 2014.

Back to Contents

TABLE I: Framework's feature summary

CATEGORIES	Economic		Logic				Measurement	Psychology	Interaction	
FEATURES	Viability	Stakeholders	Loop	Endgame	On-boarding	Rules	Metrics	Ethics	UI/UX	Technology
A Framework for Success. Di Tommasso (2011) [21]	U	E	I	E	U	I	E	U	U	U
Six steps to Gamification. Werbach and Hunter (2012) [23]	I	I	E	E	E	E	E	I	E	E
Gamification Framework. Marczewsky (2012) [24]	E	E	E	U	E	U	E	U	U	E
Gamification Design Process. Marache-Francisco and Brangie (2013) [26]	U	U	E	U	U	E	U	E	E	U
Steps to Gamification. De Paz (2013) [27]	I	E	E	U	E	E	E	U	I	E
Robinson and Bellotti taxonomy (2013) [28]	U	U	U	U	U	E	U	U	U	U
Francisco-Aparicio et al. framework (2013) [29]	U	U	U	U	U	E	E	U	E	E
A moral framework for taking responsibility. Versteeg (2013) [30]	I	E	I	I	U	U	U	E	I	E
Octalysis: Complete Gamification Framework. Chou (2013) [33]	U	U	U	E	E	U	E	U	U	U
A Framework for Sustainable Gamification Impact. AlMarshedi (2015) [34]	U	U	E	E	E	I	I	U	I	U
Player Centered Design Methodology. J. Kumar (2013) [37]	U	U	E	U	E	E	E	E	I	E
Role-Motivation-Interaction Framework. Gears (2013) [39]	E	E	U	U	U	E	U	U	U	U
Gamification Framework model. Jacobs (2013) [42]	I	E	E	U	U	I	E	U	I	I
A framework for gamification suited for marketing. Julius and Salo (2013) [16]	I	I	E	E	E	E	E	I	E	E
Theoretical Model for Gamification in Workplace IS context. Li (2014) [43]	U	U	U	U	U	U	U	U	E	E
A Framework for Designing Gamification in the Enterprise. N. Kumar (2013) [46]	U	E	U	U	U	E	E	U	E	U
Gamification Model Canvas. Jiménez (2013) [47]	E	E	U	U	U	E	I	U	I	E
Gamification development process. Herzig (2014) [49]	I	E	I	I	U	E	E	I	E	E

Note:
E - Explicit: the item has appeared in the framework's definition.
I - Implicit: the item has not appeared explicitly in the framework definition. Inferred by the authors or referred inside an academic work of the author.
U - Unavailable: the item has not appeared anyway.

Back to Contents

Communication and knowledge sharing in an immersive learning game

David Panzoli*†§, Catherine Pons-Lelardeux*§ and Pierre Lagarrigue*‡§
*CUFR Jean-François Champollion, Albi, France
Email: firstname.lastname@univ-jfc.fr
†IRIT UMR 5505, Toulouse, France
‡Institut Clément Ader EA 814, Toulouse, France
§Serious Game Research Network

***Abstract*—Learning games are becoming a serious contender to real-life simulations for professional training, particularly in highly technical jobs where their cost-effectiveness is a sizeable asset. The most appreciated feature of a learning game is to provide in an automatic way to each learner an integrated feedback in real time during the game and, ideally, a personally meaningful debriefing at the end of each session. Immersive learning games use virtual reality and 3D environments to allow several learners at once to collaborate in the most natural way. Managing the communication on the other hand has proven so far a more difficult problem to overcome. In this article, we present a communication system designed to be used in immersive learning games. This innovative system is neither based on voice-chat nor branching dialogues but on the idea that pieces of information can be manipulated as tangible objects in a virtual environment. This system endeavours to offer the simplest and most intuitive way for several learners to acquire and share knowledge in an immersive virtual environment while complying with the requirements of a reliable assessment of their performance. A first experiment with nurse anaesthetist students gives evidence that this simple communication system is apt to support lifelike behaviours such as consultation, debate, conflict or irritation.**

***Keywords*—Immersive learning game, collaborative environment, data modelling.**

I. Context and related work

The 3D Virtual Operating Room project [1] (3DVOR) is a serious game dedicated to improving the communication inside the operating room between the surgeon, the nursing staff and the anaesthetist. Miscommunication accounts for a high percentage of the different failures arising during a surgery [2] and likely to provoke irreparable injuries to the patient, or death. 3DVOR is a collaborative and immersive experience, where the learners are expected to follow the procedures of a surgery (protocols, checklists, etc.) from the admission of the patient until their transfer into the recovery room. Doing so, the objective of the game is to highlight the importance of sharing knowledge and maintaining a good assessment of the current situation for the decision-making to be effective and efficient, even in emergency situations.

3DVOR is set in a realistic environment where several locations (OR, pre- and post-operating rooms) have been carefully modelled and furnished. All along a game session, the game spies on the users and records every bit of their activity in order to guide them through the scenario and deliver a debriefing at the end of the game. The interaction model proposed by the game to the users is very typical. Users are presented with a first-person perspective of the environment which reflects realistically the actual locations of the surgeon, the nurse or the anaesthetist. The player is not enabled to move but the environment can be scrolled left or right, simulating the head movements. Objects and other characters can be interacted with by means of predefined actions and interactions (open/close a drawer, power on/off an appliance, read a document, ask something to someone, and so on). Upon being clicked, an object displays a specific contextual menu listing the interactions as textual labels. Clicking on a label triggers the corresponding interaction, which is expected to have an impact on the environment and possibly entail further interactions. Unlike the interactions, the communications between users, however, is nothing typical and has necessitated the design of an innovative system.

Traditionally in games, communication systems must compromise between traceability and naturalness.

The most natural communication system that comes to mind is natural explicit communication outside of the game, meaning talking to each other. Voice-chatting through headset and microphone is a very common way to let people communicate. Besides, this interface makes it easy for the game to record the communications. However, understanding natural language is far from trivial for a computer, let alone understanding the context and the meaning of each utterance. Natural language understanding (NLU) is still considered prone to recurring failures, and therefore traceability is compromised. As a result, most games focusing on communication skills and team-working knowledgeably use a voice-chat system and give up on the possibility to automate – even partially – the debriefing. This is the case for Clinispace [3] (Innovation in Learning Inc.) and 3DiTeams [4] (Duke Medical Center and Virtual Heroes), two learning games for healthcare training inside which the human supervisor must be part of the game in order to listen to the conversation and use them for debriefing the players once the session is over. In spite of the difficulty, one successful usage of NLU in a game must be noted, in the game Façade [5], where the player can talk naturally to the non-playing characters (NPCs) and would get an appropriate response most of the time. This suggests that such a system could as well be used for debriefing a game session, however unreliably. Besides, related domains of application like embodied conversational agents, which are virtual agents able to demonstrate verbal and non-verbal communication [6], and conversational intelligent tutoring systems [7] have reported significant advances in natural language processing techniques, and the benefits of using them are increasingly advocated [8].

Chat systems are easier to manage since the voice recognition stage is unnecessary. However, understanding the content remains as much a problem. Moreover, chat is less natural, less efficient, since at least voice-chat keeps the hand of the player free for actually playing the game. Chat systems are nonetheless very common in games. Historically, Lucasfilm's Habitat [9] was the first game to allow multiple human players to communicate in a shared virtual environment via text-chatting. In second life, a chat console is at hand for the players to communicate with each other or with chat-bots. Chat-bots are virtual characters controlled by a script and whose answers are based on the syntactic analysis (i.e. parsing keywords) of the learner's utterances. For instance, in the Indiana University Medical School Virtual Clinic [10], one can converse with a virtual patient in order to investigate their condition and formulate a diagnosis.

The pinnacle of traceability in games consists in using dialogue trees. In a dialogue tree, every utterance, question or answer is scripted in a tree-like structure. The system is very common in single-player adventure games to design the dialogues between the player and a non-playing character. Each line of dialogue from the NPC calls for several responses from the player, each of which continues the dialogue the same way a tree is being explored by an algorithm. Obviously, the drawback of this technique is the work required to think ahead and write every line of dialogue. This is even more complex when both the interlocutors must be proposed several choices. Therefore, in a multi-player context, not only is the task Herculean but it seems near-impossible to provide for every discussion that the players are likely to engage in, even in a controlled context where the topics of discussion are controlled. Despite the limitations of this technique, traceability is optimal since the objects manipulated have been designed in advance and are therefore known and easily recorded. Predefined dialogues are therefore frequently in use in learning games, provided adequate authoring tools are resorted to in order to ease the writing [11].

The fact is unquestionable that in a collaborative and team-working-focused learning game, communication between the learners is a feature that cannot be put aside. The requirements of such a communication system are defined as follows: Firstly, the ways of communication must be intuitive enough for the learners to engage in conversations naturally; Secondly, every information shared must also be easily captured and understood by the game so as to deliver the most relevant feedback to the learners individually or to the team as a whole; Thirdly, in a collaborative learning game, NPCs are likely to be resorted to to replace missing players or to play uninteresting roles, educationally-wise [12]. Those NPCs must be considered as fully equal partners (FEPs) [13]. Therefore, not only must they be in capability to understand the communications of the learners as much as their actions, but they must as well be able to participate in those communications.

In this paper, we describe a communication system which attempts to mimic the way information is shared and spread in a group, although restricted to a very specific context. The system has been designed with the goal of being the simplest and the most usable model able to comply with the above-mentioned requirements. As a consequence, the reader must keep in mind that the system has been deliberately designed to present some limitations with respect to how communication is usually understood in a general context. Particularly, the communication system presented in this paper does not intend to simulate natural communication, either verbal nor non-verbal. In concrete terms, we define communication in this research as a set of means and skills to acquire, share and use knowledge related to the training activity and the game objectives. The next section describes the data model of the virtual interactive environment. The knowledge used for the players to communicate is grounded on this data model.

II. The interactive environment

The game (illustrated in Fig. 2) runs on a web browser and is merely controlled using a mouse. Although each player sees and interacts with a subjective view of the scene, depending on their avatar's location inside the environment, the data model of the environment is centralised and hosted on a server. Recent technologies (Node.js and socket.io) allows for the application to run in real-time and interactions undertaken by each player are sent to the server which performs adequate changes in the model and broadcast them onto each user's client applications within milliseconds. This allows several players to carry out collaborative sequences of actions in real time and very naturally.

The model of the virtual environment has been described in details in [14]. Synthetically, the environment is represented as a set of objects, each of which being itself represented by a set of attributes. Attributes refer to variable features of the objects which can take several values (on/off, open/closed, wrapped/unwrapped, full/empty, dead/alive and so on). For practical reasons, the value of an attribute is a boolean. For instance, `ECG.on=true` means the ECG (object) is powered on. Attributes may refer to visual or functional features so that objects inside the environment can be displayed differently depending on the values of (some of) their attributes. Attributes can be changed by the users by means of interactions. Interactions are presented to the player as labels on a contextual menu displayed when the object is clicked. Interactions can be allowed or not depending on preconditions that must be evaluated in real time when the contextual menu is requested by the player on an object.

The model has proved to work reliably in single- and multi-player mode, allowing for multiple player's collaboration while successfully managing to maintain an up-to-date and consistent "state" of the environment for each player. In the next sections, we demonstrate that the model also caters for the content of the communications between the players.

III. Communication management

The specifications for the communication system were established keeping in mind the critical requirement for the game to analyse the communications and their content in an automated way. The communication system was also thought in accordance with the skills of the audience targeted by the

TABLE I. A PIECE OF INFORMATION CAN BE PRESENTED DIFFERENTLY FOLLOWING THE CONTEXT.

information : Patient.anxious	
context	**label**
positive	The patient is anxious.
negative	The patient is calm.
inspect	Evaluate the anxiety of the patient.
request	Is the patient anxious?

TABLE II. COLOURS ARE ASSOCIATED TO INFORMATION BUBBLES IN ORDER TO HELP THE PLAYER DURING THE RETRIEVAL PROCESS

color	**meaning**
blue	information concerns the patient
green	information concerns a conversation involving the patient
violet	information is about an equipment
yellow	information refers to a vote
orange	information refers to a document or a field within a document

game. A preliminary unpublished study showed that surgeons, doctors, nurses and medical students were mostly unfamiliar with computer games and uncomfortable with exotic set-ups like mouse-following point-of-view or keyboard and mouse combination. A modelling choice was made to keep the communications as simple to use as the interactions, that is to develop a system where only a mouse was required. The system is based on two principles: giving information a “virtually tangible” representation for the players to see, grasp and manipulate it like objects in the environment, and providing a new set of interactions for manipulating information and knowledge. This way, we understand communication, in the context of a learning game, as an uninterrupted flow of matter-specific information circulating among the players. Uncontextual communication like small talking is obviously out of the scope of this definition.

A. Information representation

Pieces of information allowed in the game for learners to communicate are facts about the environment. Facts, straightforwardly issued from the objects, are pairs of attribute/value, meaning that every attribute from every object is likely to be used as information. For instance, `ECG.on=true` and `patient.asleep=false` both represent information (the ECG is powered on; the patient is awake). For the sake of intelligibility, a piece of information is associated to a label before being displayed to the player. Depending on the context, one piece information can be translated into four different labels. There are 4 contexts: when the value is true (positive information) or false (opposite information), when the value is unknown (must-be-inspected information) and when the piece of information is meant as a question (request information). For instance, Table I lists the different meanings associated to the attribute `Patient.anxious` depending on these contexts.

Inside the virtual environment, every piece of information is represented as a floating bubble where the label is displayed (illustration in Fig.1) along with the source(s) or sender(s) of the information which are depicted by thumbnails representing the corresponding characters. The background colour of the bubble also gives a hint regarding what or whom is concerned by the information. Table II lists the colours used in the game.

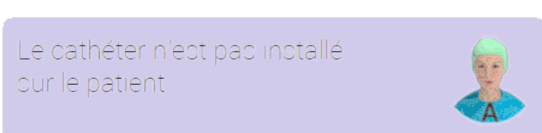


Fig. 1. An information bubble representing ‘the catheter is not installed on the patient” which was sent to the player by the anaesthetist nurse.

The bubbles are listed on a specific panel on the right and fittingly named the memory panel since it holds every piece of knowledge known by the learner (in the context of the game obviously). Incoming information bubbles pile down in the memory list so that the most recent are on top, in direct sight of the player, whereas the least recent are quickly pushed down the hidden zone of the scrollbar. This mechanism has two interesting properties with regards to the game educational objectives. Firstly, only the most recent information is at hand of the player for quickly sharing with team-mates or using in the environment. The knowledge of the procedures is therefore a valuable asset for anticipating the importance of each information and being one step ahead of the team, which in a collaborative training context is a rewarded behaviour. Secondly, the large amount of bubbles continuously piling down the list makes it time consuming to scroll and seek a piece of information within less recent knowledge. Filters have been designed in order to facilitate the retrieval of a specific piece of information on the basis of contextual cues given by the player. Indeed, knowledge in memory can be filtered by means of ticking check-boxes (filtering on the roles involved) and/or pointing at the objects in the environment (filtering on the object concerned). The ability to identify the context of the information needed (by whom was it sent, whom or what object was concerned, etc.) and to use the filters adequately is another rewarded skill helping improve the performance of the learner.

Playing the game, the team of learners will be given the opportunity to explore several scenarios, all of which are dealing with hazardous and adverse situations that can be avoided or recovered by an efficient communication. Besides a good knowledge of the protocols, the patient safety in the operating room is tied to a handful of good practices: to collect information from the environment (objects, team-mates) in order to maintain an up-to-date knowledge of the current situation; when made aware of a new information, to be able to assess how important it is to each other team-mate (with regard to their occupation in the OR) and share the information accordingly; and finally, to refer to one’s knowledge of the situation when a collaborative decision needs to be made. For every one of these tasks, an interaction was designed and implemented.

B. Gathering information

Information can be collected from the objects (furniture, medical equipment, documents) or from the other players (see next section). To collect a piece of information from an object, the player has to click on it in order to display the contextual menu. Inside the menu, a list of attributes is displayed along with the interactions available for this object. In the contextual menu, the values are always hidden to the player as only the “inspect” labels of the attributes are displayed (see Table I). In order to learn about its value (i.e. get the entire meaningful information), the player must click on the label and collect

Fig. 2. This image is a screenshot from the game. The main panel is the 3D environment from the point of view of the player, including the objects, furniture and other characters. Every interactive object is highlighted when hovered and a tag with its name is displayed around the mouse cursor. The menu bar at the top contains links to open the documents (which are not spread throughout the environment for an easier access). The second to last icon on the bar can be used to trigger a vote about one among several predefined topics. The "memory" panel at the right of the screen contains every piece of information gathered by the learner. When the screenshot was captured, a piece of information was being transmitted from David the anaesthetist (the player) to Jules the surgeon. The tag says: "the anaesthetist is equipped with gloves".

the information. That way, the game keeps a record of every information acknowledged by the player during the game session. This mechanism is essential since letting the players see and learn new information without the system knowing about it would hinder the accuracy of the debriefing.

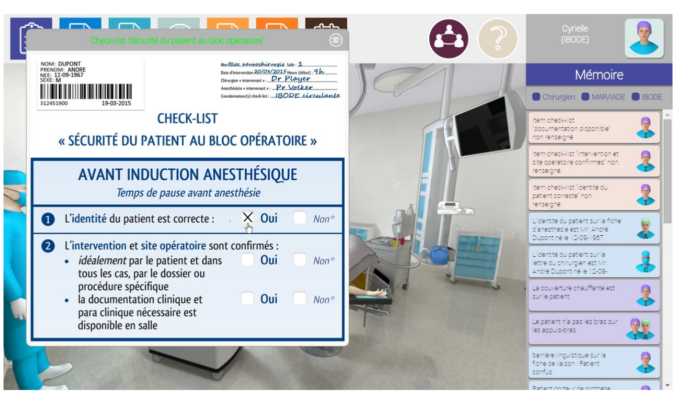


Fig. 3. Clicking a document icon on the game screen's top bar displays a realistic depiction of the document (for instance in this screenshot, the security check-list used in France). Documents are objects that can be interacted with (changing values, ticking boxes, etc.) and from which information can be collected.

C. Communication between players

Sending information to a team-mate is simple as dragging and dropping the corresponding bubble to their character. In Fig. 2, a piece of information is being sent by the player to another character. When a player is being talked to, a pop-up appears in the middle of their game screen. Merely clicking on the pop-up acknowledges the communication and the information bubble is placed on the memory panel. A player can also broadcast information to everyone at the same time by dropping the bubble in the environment, or on the speaker icon (top-left of the game screen).

Upon being received, an already existing information in memory is pulled to the top of the panel. The object/attribute couple is what makes two pieces of information come under scrutiny every time a new information is received. The value of the attribute and the source are two varying properties of a piece of information. Depending on them, various interpretations are likely to be made by the learner, as Table III shows. When the exact same piece of information is repeated, it is simply pulled up to the top without any other form of processing. When the entering piece of information updates the previous one, the bubble is updated, pulled to the top and flashes for a few seconds. When an existing piece of information is confirmed by a new one, the corresponding bubble inside the learner's panel is added a thumbnail depicting the sender or the player's avatar, depending on

TABLE III. A PIECE OF INFORMATION IS INTERPRETED DIFFERENTLY DEPENDING ON THE CONTEXT.

	same value	different value
same source	information is being repeated	information is being updated
different source	information is being confirmed by a third party	conflicting information, some of which is necessarily inaccurate

whether the piece of information was sent by a team-mate or collected by the player themselves. Finally, when an entering piece of information causes a conflict, both the new and the old bubbles are pulled to the top and flash for a few seconds. It is the player's responsibility to investigate and solve the problem.

Sending information is an intentional action undertaken by a player when they feel some knowledge they have acquired is of any importance to another player and therefore should be shared. This proactive behaviour denotes a good knowledge of the situation and/or a good experience, although in practice a significant part of the communication is likely not to be anticipated but delivered on request. To that end, the communication system offers the ability for a player to ask some information to another player. The interaction process is similar to collecting information from an object, only that the value of the information is not available directly. In practice, say player A means to ask player B some information. The list of available questions is presented to A by the contextual menu associated to B. The questions are almost straight translations of all available pieces of information in the memory of B, only put in the interrogative form using the request label (as described in Table I). At this stage, the actual value of the piece of information is hidden to A, since only the objects and the attribute are necessary. Information unknown to B is absent from the list and therefore unavailable for A to ask. The pending request is notified to B by a window popping-up overlaying their game screen, just like any other information sent. The request window however contains two additional buttons for player B to send a quick acknowledgement of receipt translating their intent. "It's not my role" intends to tell player A that their question is very likely to remain unanswered whereas "I'm on it" supposedly means the information is to be sent shortly. In whatever case, whether player B will indulge or not is out of the responsibility of the player alone.

D. Collaborative decision making

Votes can be cast at any time during the game by any player. A vote is a question on a selected topic about which all the players are requested to give their opinion. The vote is limited in time (in the experiment of section IV, the time limit was set to 90 seconds). During the vote, the game is paused and disabled. The players are free to select an answer, and a thumbnail of their character is placed in the corresponding box for the other players to see each other's decisions in real time. Each player is also allowed to drag and drop information bubbles in provided spaces. They stand for arguments or evidence to support their vote or convince team-mates.

When the time is out, the player who triggered the vote is responsible for selecting the final decision. Whether or not the final decision reflects the opinion of the majority is the responsibility of this player. Similarly, the other players have the right not to acknowledge this decision and take a counteraction. Making decisions and acting in the stead of the players is out of the scope of the vote mechanism. However, recording the arguments and the outcome in order to use these data for debriefing is clearly an added benefit of the learning game.

Fig. 4. The players are voting. The surgeon has a different opinion than the anaesthetist and tries to use arguments to plead their cause. The nurse has not voted yet.

IV. Experiment

Each scenario in 3DVOR has been designed to be played in standalone mode (without trainer's intervention), in supervised mode (with teacher's intervention) or in blended mode (with asynchronous trainer's intervention).

A. Context and practice

The experiment took place during a course at the anaesthetist nurse school of Toulouse in March 2015. The learning game was used by a teacher to evaluate their student's knowledge of procedures, as part of the curriculum. The experiment had no impact on their grades. The experiment consisted of 3 game sessions, each with a different group of students. Each group played the same scenario – a hazardous situation where both the risks of operating the wrong patient and of missing the surgical site are high – which is specifically dealing with risk-management and teamwork. In the scenario, each role has access to a limited number of documents of the patient records and therefore the players are encouraged to communicate to share this fragmented knowledge. The objective consists in looking after the patient from their arrival in pre-operating room until the end of the anaesthesia procedure. In order to assess the performance of the students, the scenario embeds a set of metrics to measure how well the standard procedures are applied. Communication with the patient is an important element of this scenario as well. Positive communication, like informing the patient or telling them jokes, must be used to counter effect the many anxiety-provoking actions of the procedure and balance the patient's anxiety within a comfort zone.

Considering that all the learners were inexperienced anaesthetist nurse students, the teacher asked them to pair-up so that each team would be composed of 3 teams of 2 students (see Fig. 5). Each team would then have to play a role in the game: the surgeon, the anaesthetist and the nurse. The rules

Fig. 5. Three (times two) players and the trainer take part in each session of the experimentation. While the learners are playing, the trainer supervises the game in real time and uses the supervisor's tools to take control of the session when necessary.

Fig. 6. The supervisor's tools are grouped in a panel where the game can be observed in real time or taken control of.

of the experiment were clearly stated at the beginning of each session. Oral communication was allowed within a team but forbidden outside, as only the game communication system must be used.

B. Supervisor's tools

During each game session, the trainer uses a supervisor's panel (illustrated in Fig. 6) which contains a set of tools to follow in real time the progress of the scenario. Visualisation tools include:

- A detailed view of the actions being carried out and the pieces of information being exchanged by the learners, and
- A more synoptic view of the state of progress of the scenario against the objectives.

Control tools enable the trainer to actually take control of the game:

- The game can be paused/resumed for the trainer and the learners to review the situation if it gets confused or if the team is unable to advance the scenario (lack of experience or lack of trust in someone's decision).
- The game can also be reset to start over.
- To encourage the team to work collaboratively, the trainer can cast a vote on a topic of his/her choice and doing so force the learners to make a decision.

Finally, the supervisor's tools also include a debriefing panel where the objectives of the scenario are synthesised at the end of a session. The completion or failure of each objective is mentioned so that the learners are led to understand their mistakes on their own. It can also be used to support a more elaborate (i.e. at the scale of the team) analysis of the scenario in retrospect, conducted by the supervisor. The debriefing tools, the data model representing the objectives and the technical know-how used to link them to the scenario and the communication are deliberately not detailed in this article.

C. Playtests and results

During the three sessions of the experiment, every interaction within the game was computer-recorded for analysis. Every session (see Fig. 5) and post-session individual interviews were recorded on camera for assisting and corroborating the analysis of quantitative data.

In a general way, the data collected express a strong involvement of all the learners towards the game, which is confirmed by the recordings showing enthusiastic and lively behaviours. Table IV presents the interaction count per session sorted by category (due to a technical issue, the records for sessions 3 are missing). Based on these figures, several observations and hypotheses can be formulated.

- No interaction has been left unused, which indicates the different interactions seem to have been understood by the learners.
- The quantity of information collected from objects is significantly higher than other related interactions like transmissions or requests. This behaviour denotes a systematic information scavenging of the environment by the learners and points out that on several occasions, the team may have temporarily lost track of the scenario. This problem is independent from the communication system and can be explained by the fact the learners in this experiment were not experienced surgeons, anaesthetists and nurses but students.

Table V counts how many times each document has been accessed by each learner. Some documents were made inaccessible to specific roles in the game to reflect the fact that for instance the anaesthesia record can only be read and understood by the anaesthetist. Besides, the game in general and this scenario in particular are centred on sharing information and therefore letting all the practitioners in the OR have access to every information would be nonsensical. In the table, an inaccessibility is mentioned as "non-applicable" (n/a). Unlike the information inside the environment (see paragraphs above), information from the documents were accessed parsimoniously, as the low figures in the table indicate. This indicates that the learners were well aware of the interest and the utility of this information and therefore the documents were only accessed on purpose. Again, this is consistent

TABLE IV. COMMUNICATION-RELATED ACTIONS UNDERTAKEN BY THE LEARNERS SORTED BY CATEGORY.

	Dialogue		Action	Collection	Listening	Transmission	Vote				
	requests	answers to requests	actions on object	info. collected from object	info. listened	info. transmitted	votes	answers to votes	votes validated	arguments	arguments withdrawn
session 1	23	17	38	234	11	15	10	31	13	47	3
session 2	18	6	29	286	5	15	5	21	6	31	0
mean values	20.5	11.5	33.5	260	8	15	7.5	26	9.5	39	1.5

TABLE V. READINGS OF THE PATIENT FILE AND OTHER DOCUMENTS RECORDED DURING THE SECOND SESSION.

	Checklist	MRI	Surgical planning	Anaesthesia record	Liaison form
Nurse	18	n/a	5	n/a	3
Anaesthetist	1	n/a	4	2	0
Surgeon	5	2	0	n/a	0

with observations made in the operating room and with the expectations of the scenario.

The “talk to everyone” feature, represented by a speaker at the top left of the screen, was very scarcely used and perhaps most of the learners could not figure how to use it properly and safely preferred the one-to-one communication scheme. However, this cannot be interpreted as a failure to collaborate as the vote feature was on the other hand often used. On average, the team of learners took 7,5 collaborative decisions (votes) per session. It was observed that during a vote in the game, the learners tended to argue much more than in real life, and they clearly failed to identify the most relevant information likely to rest their case unquestionably. As a result, deadlocks were reached on some occasions and the intervention of the trainer was necessary.

In a more general context, the communication system revealed incapable to solve ingrained conflicts. When facing adversity, some learners were clearly and firmly disagreeing with the rest of the team, and would refuse to communicate by systematically answering “It’s not my role” to every question. Video records show irritated gestures from the learners on these occasions. Interviews conducted after the sessions have revealed the learners wish they could have used some chat system ultimately.

V. CONCLUSION

Natural collaborative action in an immersive environment is certainly an important feature for learning games to close the gap to the level of realism offered by role playing, real life simulators and training set-ups of the like. Communicating and sharing knowledge is even more important for the learners to actually be trained efficiently. Yet, this aspect of a game has received very little attention so far. In this paper, we have presented a communication system offering a way for players to exchange knowledge in real time, which is to our belief the very purpose of communication in the workplace. Indeed, our definition of communication, as manipulating pieces of information, is clearly specific to the objectives of a learning game. Yet, although it is not as expressive a communication system as a chat or a voice-chat, it enables for the game to understand the exchanges between the players and use that knowledge for debriefing the players, or at least facilitating the task of the trainer. Besides, how each piece of information is interpreted remains the responsibility of the players sending and receiving it.

A first experiment with the system has revealed the strengths and gaps of our communication system. The game has received a positive welcome from the audience and the data recorded from the sessions have confirmed the successful appropriation by the players of the various interaction abilities designed for them to communicate. On the negative side, we learnt experimentally that the communication system did not grant enough expressiveness to the learners to help them solve every conflict. Conflictual situations, we consider, are likely to thrive in a collaborative environment if the communication is not optimal. Future work will therefore address this issue by conducting further experiments dedicated to better understanding collaborative decision making.

ACKNOWLEDGEMENT

The steering committee of 3DVOR is composed of Pr. Pierre Lagarrigue, M.D. Ph.D. Vincent Lubrano, M.D. Ph.D. Vincent Minville and Catherine Pons-Lelardeux. The scenarios used in the game were designed by Thomas Rodsphon and Vincent Lubrano from the Toulouse University Hospital. The experiment described in section IV has been conducted under the supervision of Christiane Paban, Hoang-Minh Truong and Amélie Lafitte. The authors are also grateful to the designers and programmers of the game: Cyrielle Guimbal, Jules de Guglielmi and Romain Régis.

REFERENCES

[1] P. Lagarrigue, V. Lubrano, V. Minville, and C. Pons-Lelardeux. (2012) The 3dvor project. [Online]. Available: http://3dvor.univ-jfc.fr/

[2] L. Kohn, J. Corrigan, and M. Donaldson, *To err is human: building a safer health system*. Washington: National Academy Press, 2000.

[3] D. Parvati and W. L. Heinrichs, “Clinispace: A multiperson 3d online immersive training environment accessible through a browser,” *Medicine Meets Virtual Reality 18: NextMed*, vol. 163, p. 173, 2011.

[4] J. M. Taekman, N. Segall, E. Hobbs, and M. Wright, “3diteams–healthcare team training in a virtual environment,” *Anesthesiology*, vol. 107, no. A2145, p. A2145, 2007.

[5] M. Mateas and A. Stern, “Natural language understanding in faade: Surfacetext processing,” *Proceedings of the Conference on Technologies for Interactive Digital Storytelling and Entertainment (TIDSE)*, 2004.

[6] J. Cassell, *Embodied conversational agents*. MIT press, 2000.

[7] V. Rus, S. DMello, X. Hu, and A. Graesser, “Recent advances in conversational intelligent tutoring systems,” *AI Magazine*, vol. 34, no. 3, pp. 42–54, 2013.

[8] B. Hennigan, “Making the case for nlp in dialogue systems for serious games,” *8th International Conference on Natural Language Processing (JapTAL), 1st Workshop on Games and NLP*, 2012.

[9] C. Morningstar and R. F. Farmer, “The lessons of Lucasfilm’s Habitat,” in *The First International Conference on Cyberspace*, M. Benedikt, Ed., May 1990. [Online]. Available: http://www.fudco.com/chip/lessons.html

[10] M. T. Johnson and M. Clary, “A second life virtual clinic for medical student training,” 2008, second Life Education Community Conference (SLEDcc’08).

[11] D. Mori, R. Berta, A. De Gloria, V. Fiore, and L. Magnani, “An easy to author dialogue management system for serious games,” *J. Comput. Cult. Herit.*, vol. 6, no. 2, pp. 10:1–10:15, May 2013. [Online]. Available: http://doi.acm.org/10.1145/2460376.2460381

[12] M. Sanselone, S. Sanchez, C. Sanza, D. Panzoli, and Y. Duthen, “Control of non-playing characters in a medical learning game with Monte Carlo Tree Search (regular paper),” in *IEEE Conference on Computational Intelligence and Games, Dortmund, Germany, 26/08/2014-29/08/2014*. http://www.computer.org: IEEE Computer Society, aot 2014, pp. 208–215.

[13] D. Thomas and L. Vlacic, “Collaborative decision making amongst human and artificial beings,” in *Intelligent Decision Making: An AI-Based Approach*. Springer, 2008, pp. 97–133.

[14] D. Panzoli, M. Sanselone, S. Sanchez, C. Sanza, C. Lelardeux, P. Lagarrigue, and Y. Duthen, “Introducing a design methodology for multi-character collaboration in immersive learning games (regular paper),” in *Proceedings of the Sixth International Conference on Virtual Worlds and Games for Serious Applications (VS-Games14)*. University of Malta: IEEExplore digital library, 09–12/09/2014 2014, p. (electronic medium).

Back to Contents

A Feasibility Study for Gamification in Transport Maintenance

Requirements to implement gamification in heterogeneous organizations

Vinutha Magal Shreenath, Maksims Kornevs, Jayanth Raghothama, Sebastiaan Meijer
School of Technology and Health
KTH Royal Institute of Technology
Stockholm, Sweden
vinutha, kornevs, jayanthr, smeijer @ kth.se

***Abstract*—Gamification has been successfully applied in many domains, but mostly for simple, isolated and operational tasks. The hope for gamification as a method to radically change and improve behavior, to provide incentives for sustained engagement has proven to be more difficult to get right. Applying gamification in large networked organizations with heterogeneous tasks remains a challenge. Applying gamification in such enterprise environments posits different requirements, and a match between these requirements and the institution needs to be investigated before venturing into the design and implementation of gamification. The current paper contributes a study where the authors investigate the feasibility of implementing gamification in Trafikverket, the Swedish transport administration. Through an investigation of the institutional arrangements around data collection, procurement processes and links to institutional structures, the study finds areas within Trafikverket where gamification could be successfully applied, and suggests gaps and methods to apply gamification in other areas.**

Keywords—gamification, feasibility, requirements, data mining, procurement

I. Introduction

Over the last couple of years, gamification has earned a fair amount of positive attention as a means of supporting engagement and incentivizing behaviour change, towards increased user activity, social interactions or improving the quality and productivity of tasks. Gamification is supposed to create these changes by providing positive, inherently motivating “playful” experiences through the use of interactive technologies embedded into daily tasks and services [1]–[3].

Gamification was largely considered to be a tool for supporting engagement, and applied mostly in Business to Consumer (B2C) applications [4], [5]. People also considered applying gamification in enterprise contexts, within organizations to improve motivation and quality of tasks performed. Despite the large number of gamification projects and applications, evidence shows gamification to be only a partial success, with difficulties in sustaining the initial benefits accrued through gamification over a long time [6]–[8].

Irrespective of the context, “gamified” applications try to include game mechanics in real life situations, in an attempt to evoke the psychological and behavioural outcomes of games in the real life context. Such playful designs typically reward the players’ activities with points and badges, stratifying them into levels based on accumulation of rewards and creating competitions among players [9].

Many such applications are for routine, simple and isolated tasks, even within enterprise applications, such as learning, or reviews of documents or interacting on social networks and so on [10]. Applying gamification in large networked organizations with heterogeneous tasks posits different challenges and requirements, such as reconciling the heterogeneity of tasks, the influence of institutional structures, and institutional arrangements on the provisions of data and so on. It is crucial to investigate these different requirements before venturing into the design and implementation of gamification in an enterprise context, especially to tasks central to the organizations performance [11]–[13].

In the following sections, the authors present a study on the feasibility of implementing gamification within Trafikverket, the Swedish transport administration. The study is undertaken in the context of implementing gamification to incentivize energy efficient behaviour, as part of a project called ELSA. ELSA is an enterprise information system, to manage Trafikverket’s energy. We do this by first listing the requirements for implementing gamification within ELSA, outlining the method for investigating the feasibility, and the results of the investigation. The results point to various areas and methods by which Trafikverket can implement gamification and challenges therein.

II. ELSA

ELSA stands for Energy Management System for Installations (in Swedish: Energiledningssystem för anläggningar). The overall goal is to create an energy management system that is compliant with the ISO 50001:2011 standard to ultimately reduce the amount of energy used in large-scale installations within the rail and road sector. Gamification is conceived as an approach to becoming more energy efficient by influencing the behaviour of those in the power to actually change a part of the system. This may be operators who could drive more efficient, maintenance mechanics who need to choose between different repair

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

options, and planners who influence the operations in a particular region. ELSA in this case would act as a mediator to introduce game mechanics such as points or relevant feedback to players [5].

Energy management in Trafikverket is an approach that encompasses many departments, institutional levels and tasks, and even extends to the entire infrastructure sector due to the dominant agency position the organization has. It is influenced by operations (for example maintenance tasks such as ploughing snow), planning (the order of roads in which snow is ploughed), procurement (the type of equipment to buy for maintenance) and so on. Personnel at different levels in the institution conduct these tasks. Further, many operations at Trafikverket are (sub) contracted out to different companies, which might have different equipment, standards and operating mechanisms within the framework of the contract between the company and Trafikverket.

Gamification in a sector-wide way in the context of a networked organization (with a multi-actor, multi-level nature) should encompass all these departments, companies, institutional levels and people. This implies a gamification approach where everybody involved in energy management, directly or indirectly should be players in the gamified world[1]. Since the game has to incentivize energy efficient behaviour, their tasks during the course of work have to be quantified and measured for relevant metrics such as energy efficiency or carbon emissions. The metrics can then become the basis of the gamification system, and points, rewards and competitions can be setup based on these metrics.

Construction of these metrics requires the availability of data, through the ELSA system. This data collection should be automated to a high degree, and the metrics should be relevant to the players work, providing valuable feedback. Data collection is done through the ELSA system, but the construction and operation of the gamification system should not interfere with working models and organization processes, so as to not have negative impacts on the organization.

In short, the game is a layer augmenting existing processes and structures within Trafikverket that incentivizes energy efficient behaviour.

A crucial factor in the design of the game is the heterogeneity in the organization. Players are at different levels in the organization and conduct different types of tasks (such as operations, planning, and procurement). Trafikverket also manages through various (sub) contractors different types of maintenance tasks, such as snow ploughing, road lighting, railway switch maintenance and so on. These tasks are all different, in seasonality, planning, material required, savings on energy and influence on the overall energy efficiency.

Encompassing this heterogeneous organization with its multitude of tasks, people and companies within a single game is a challenge. Achieving this will mean that this heterogeneity is reconciled, or at least managed in some fashion. Whether ELSA can be managed as a single game, or multiple games that mix and match levels and tasks is a question to be investigated. For example, is it possible to build a gamification system to compare a maintenance mechanic with a planner, or a snow plough driver in one region of Sweden with another, or a snow plough driver with a train driver?[2]

It is important to understand the availability of data, the impact of gamification on organizational processes, and the limitations of what can be achieved within the current institutional arrangements.

III. Methodology

The feasibility of gamification within ELSA was investigated over the course of four months, focusing on the following issues:

- Availability of data: What is the data, and the level of detail that is available, and that can be used to construct the different metrics for players at different levels in the organization?
- Game Models: Based on the availability of data, what are the possibilities and types of combinations of games for different tasks and levels?
- Procurement: What is the potential impact of gamification on organization processes, such as procurement?

The feasibility was investigated through a series of workshops with relevant stakeholders. The first workshop was with the heads of different units in Trafikverket, all of whom dealt with energy and with people from the IT units who were responsible for the collection and management of data in Trafikverket. The second workshop was conducted with representatives from different contractor companies, which perform a significant amount of maintenance tasks for Trafikverket, to investigate the availability of data within these companies and to obtain insights on how these companies perceive gamification. The third and final workshop was conducted with representatives from the transport administrations of all Nordic countries to get insights on how these processes differ across countries. In between the workshops, the researchers followed leads into the organization and topic on issues that were raised by the participants.

The feasibility was investigated through three distinct maintenance scenarios: snow ploughing, road lighting and railway switch maintenance. Initially, the third scenario was that of ballast cleaning along railway tasks, but it was discovered already in the first workshop that there was very little data about this task since it occurred very rarely and infrequently. The scenario of railway switch maintenance was proposed instead.

The first management scenario is that of ploughing snow, where roads are cleared of snow whenever snowfall happens. The behaviour of the driver driving the plough influences the

[1] To simplify, we refer henceforth to the gamification concept as a game, and people who will be users of this gamification concept as players.

[2] In the following sections, we refer to each such comparison as a different competition.

energy efficiency of the task itself, but also saves energy for all vehicles driving on the road.

The second scenario is that of road lighting, where regular maintenance and procurement of more efficient lamps and control systems saves the electricity consumed.

The third scenario is that of maintaining switches on the railroad, where regular maintenance and immediate repairs reduces delays and saves fuel. The heating of switches in winter is handled by control systems, which turn on or off heating based on weather. However, the implementation of such control systems is a strategic decision, and the players at this level can also be included in the game.

The snow plough scenario is seasonal, whereas the other two are not. The behaviour of the maintenance personnel has the most influence on energy efficiency in the snow plough scenario, a more moderate influence in the switch maintenance and very little influence in the road lighting scenario. The road lighting scenario is mostly about planning, whereas the other two are not.

IV. Results

Three aspects in the feasibility of the ELSA concept were investigated: the availability of data to measure players' influences, the feasibility of constructing relevant metrics from available data for all players and construct different kinds of levels or games within ELSA, and the processes within Trafikverket that will be impacted. The following sections describe the results of the investigation.

A. Data

Data on different aspects of maintenance operations is already being gathered by Trafikverket in different systems. For each of three scenarios we document the data already available, and the data that is missing.

The data requirements and availability is analysed based on the institutional structure. For players at every level in the institutional structure as conceptualized in Figure 1, we look for data that is needed to quantify their influence on energy efficiency. Potential data requirements are identified through a conceptual systems dynamics diagram for that scenario. For example, Figure 2 is a broad, conceptual systems dynamics diagram representing the various influences that affects the amount of snow that can be cleared, emissions and monetary savings that can be achieved by a snow plough driver.

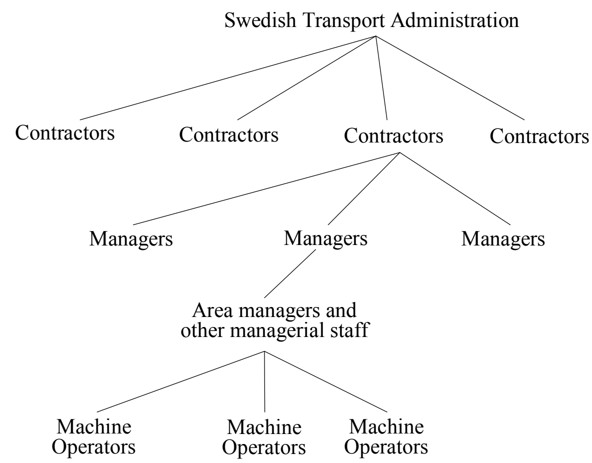


Fig. 1. General Institutional Structure

Further, for each scenario and set of data requirements, we investigated the availability, detailed parameters, the stakeholder currently responsible for maintaining the data, the source agency or system holding the data and the extent of accessibility to Trafikverket and ELSA.

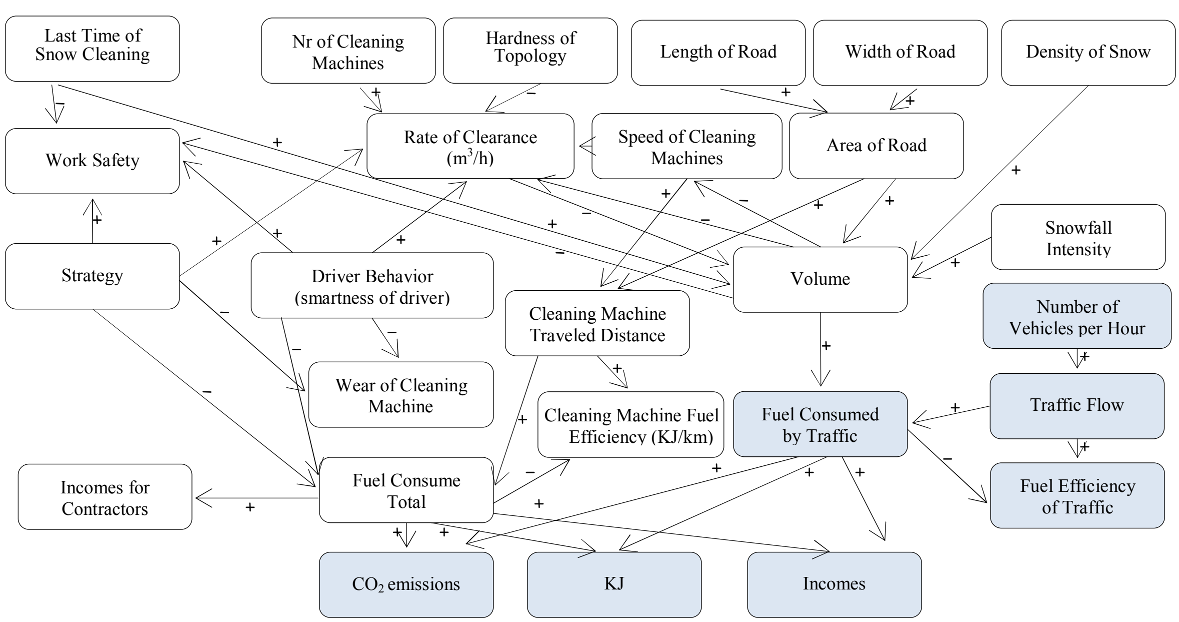


Fig. 2. Systems Dynamics Model for Snow Plough

Back to Contents

TABLE I. DATA REQUIREMENTS FOR SNOW PLOUGH

Data	Parameters	Stakeholder responsible	Source or Agency	Available to Trafikverket	Purpose
Road network: GIS Topology Classification	Link information, Annotated with driver survey data on hardness of topology	Trafikverket	NVDB	Yes	Link information, co-related with surveys of drivers to annotate GIS database to give topology. Classification to give indication of traffic flow.
Weather data	Snowfall (cms/hr), density, start time, end time, per geography area	Trafikverket	VVIS	Yes	To normalize weather conditions,
Machines or fleets	Speed, Fuel consumption, gear changes, idle time etc., co-related with GIS and time	Contractors	Unknown	No	Ping from machines with relevant data , which help estimate driver behaviour
Traffic flow	Vehicles/hour, per link in GIS database	Trafikverket	Tindra (DB storing statistics)	Yes	To estimate energy saved
Operational data	Machines (machine identities) which perform a task (winter/summer maintenance) at a specific time, for long how, the route taken to within a particular area complete this task	Trafikverket	Mipmap	Yes	To estimate driver behaviour and to calculate points.
Personnel data	Data on operators of machines → managers → contractors → division with Trafikverket responsible	Contractors and Trafikverket	Within Trafikverket: Contractors and some IT systems	Partially	To estimate driver behaviour, to attribute points to managers and contractors correctly.

1) Snow plough scenario

Snow is cleared from the road network based on the priority of the roads, rate of snowfall and other criteria, listed in contracts between Trafikverket and contractors. Apart from snow being cleared as agreed to in the contract, workers familiar with the area also take proactive de-icing measures to prevent accidents. As depicted in Figure 2, apart from external influences, there are some parameters which can be directly controlled by the driver to plough the snow with less utilization of CO2, and these can be used to encourage energy efficient behaviour. Table 1 presents necessary parameters, data sources and the relevant stakeholders for this scenario.

2) Road lighting

Trafikverket defines policies for energy efficiency for lighting roads across the road network. These policies could be procuring energy efficient lamps, control systems to monitor and change energy usage etc. Energy efficiency of road lighting can be measured across the road network in terms of counties, regions, and contractors. The road lighting scenario is less dynamic and changes in this scenario less frequent than that of the snow plough. The institutional structure is similar to that of the snow plough scenario, but the decisions that affect the efficiency are taken at higher levels. The energy efficiency is more dependent on the equipment and policies than on the personnel who repair the equipment, although quality of repair can be estimated as well from this data.

3) Railway Switch Maintenance

The personnel responsible for maintenance of switches conduct routine manual checks on the switches and file a report if they find a fault. The maintenance crew fills out a form to get required equipment. The crew then takes out vehicle(s) to the site and fixes the fault.

From the first workshop, we discovered that the fuel efficiency in this scenario is directly related to the quality of maintenance conducted, since higher quality of maintenance implies less maintenance. Gamifying improvement of quality of maintenance will have proportional impact on energy efficiency. The quality of maintenance and hence the fuel efficiency can be measured by in terms of:

- Faults per switch OR faults per rail passage OR faults per tonnage
- Delay per fault
- Number of times the same fault repeats

This process goes on around the year; the frequency of data collection is at the same as the frequency of the event (of maintenance).

4) Summary

Operational data needed for the ELSA game is available across various systems. Data on the stakeholders or owners responsible for some of these systems is not accessible, but the data itself is collected in some form. Operational data is collected on an automated basis in most cases, recorded from sensors or other devices, and entered by the maintenance personnel in other cases, but is collected regularly for each mission or task etc.

In all three scenarios analysed, some data is required to build the game for one level, and other data to build correction factors so other competitions can be set up, such as competitions between two drivers, or a driver and a manager etc. These correction factors are required to make the competition fair by normalizing factors not entirely in the control of the players, such as weather conditions and the state of the machine etc. Although operational data is available, data on the activities of people at other levels of the institutional structure is not readily available. For higher levels, data is available only in the form of contracts, between different units in Trafikverket and contractors. To make the game feasible across levels, new forms of data collection and metrics of performance need to be established within the upper levels of the organization (managers and planners).

A lot of the data collection today is already done in real time, but some of the data is collected on a batch basis, usually after the maintenance task. This frequency of collection has implications for the game design and game play. If data is not available real time, real time game play will not be possible. Players can only play the game on a similar schedule as that of the data collection, i.e. game play can move forward only when new data is available. This should be considered when creating infrastructure for new forms of data collection.

The IT systems within Trafikverket seem to be of at least two types, as discovered during the feasibility study:

Internal systems/tools built for Trafikverket – examples are Mipmap and Bessy. These systems seem to be built for the unique requirements of Trafikverket. These systems will be easier to examine and data from these systems can be collected with more authority.

- Systems used by Trafikverket – example for this is Maximo (an assets management database). Maximo is an enterprise product by IBM. The input, the data standards and the workflows are fixed. Maximo especially seems to be able to work with internal systems as well. In general, with such licensed software, there might be limitations of querying with this product's database to collect the data needed for the game.
- The differentiation in the type of systems raises the issues of availability and collectability as they are no longer the same. Operational data for many scenarios might be available, but if the data lies in a licensed software product's database (such as Maximo), it might be hard to get to as there will be licensing and cost issues when querying the data in a way that is not expected in the software product, or in creating a new interface on the product to collect data for the game.

The software interfaces to pull data from different sources to the game would need to be designed, while considering privacy and anonymity of players, and connectivity to other interfaces which calculate points.

B. Game Models

Metrics for every player in the game, on their influence on energy efficiency during or after performance of their task is considered as a score. Every player in the game can be scored on three basic attributes: money, energy and emissions, the emissions reduced or created, energy saved and monetary benefits by saving energy. The scoring is for each individual mission or day or unit of time, which can be different for every scenario. For example, while ploughing snow, the score can be calculated per day or per hour or between activation and deactivation of the plough, or between driving out the machine and returning it to a garage. The scores are based on the quantity of money and energy saved, and the quantity of emissions. While some savings can be done during the maintenance task itself, most of the savings accrue by people using the infrastructure being maintained. Hence, the money, energy and emissions saved or not saved by all the users of the maintained infrastructure also factor into the score calculation.

Given that data is available only for operational tasks, the scores for all players in the game have to be calculated based on operational data. Thus, scores for managers (within a contractor organization), planners (within the transport agencies), and other higher levels in the institutional structure have to be inferred from this operational data, or a combination of operational data and other means. For the operations personnel, and their managers, the scores are calculated based on the operational data available. The personnel at other levels, scores need to be calculated based on operational data, and expert panels that can infer how plans and strategies influenced operations. This also implies that if more relevant and precise metrics for people at higher levels in the institutional structure needs to be constructed, some standardized way of collecting and analysing contracts and plans has to be developed.

Once scored, every player can be compared against his or her own performance. For competitions with other players at the same institutional level, some correction factors need to be applied, so the competition can be fair. For example, two snow plough drivers can compete with each other if the geography, weather and machines are the same. These correction factors might also need to be applied for historical comparison of the same player, since the maintenance tasks will not be constant but depend on weather and the machine and so on. Therefore, those factors will need to be corrected for if two drivers in different countries can compete against each other. Similarly, other correction factors might need to be applied for different levels in the organizations. For competitions across different levels, the scores have to be calculated differently.

Each maintenance scenario is different, in terms of the actual operations, the way it is planned, its impact on energy consumption, frequency of occurrence and seasonality. The data collected for each task is different, and the availability is different as well. This implies that the scoring mechanism is different for each scenario, and has to be built separately. As more scenarios are added into the game, the mechanisms for that scenario will have to be added. Comparison across these scenarios will be difficult without correcting for even more factors and collecting more data.

C. Procurement

Organisational functions within a game space needs to be integrated using one tool across all levels and different partners. Considering that i) access to all systems for data collection across the organization requires a legal support system and ii) there is a need to influence external partners, a legal framework can serve as a tool for supporting organisational function. Such a legal framework that also can affect partners can be implemented though a procurement system.

Procurement is the set of measures taken by a contracting authority with the aim of awarding a contract or concluding a framework agreement regarding products, services or public works [14].

Swedish law, similar to most developed countries, specifies that a procurement contract is a legal document which binds both parties (in this case Trafikverket and a contractor) to fulfil their obligations. These obligations include completing the assignment by satisfying technical and functional specifications or paying for said service. It could require descriptions of equipment used in terms of characteristics and fuel consumption of machinery or requests to provide records for every time that a service task is performed. However these obligations also describe other processes and aspects related to the procured item. One such aspect is related to data and its availability.

Contractors can only make sure that they meet all requirements by measuring their performance. This implies that they collect data on their actions, and the same fact was discovered in the second workshop. This could include data about:

- environment and conditions (weather, traffic, geography, road quality),
- available resources (vehicles, workers, salt or sand mix, knowledge),
- work performance (usage of petrol, utilization of time, travelled distance), and
- work result (quality measures)

Contractors currently use the data they produce to determine if their procurement requirements are fulfilled. Some data is reported to Trafikverket according to the contract, but a fair amount of data that they collect is often not reported. In some cases they report only aggregated data. For example, typical road maintenance contracts state: “The contractor shall report sand and salt consumption, for the previous month”. Information about location, time and reasons for using sand and salt does not go to Trafikverket.

The same applies to other data. Thus, the situation right now is that contractors make a lot of measurements to assess whether they are satisfying the requirements in the contract, but Trafikverket does not have access to all this data because it is not specified in the contract.

One solution could be to make procurement processes and contracts more stringent by requiring more data. This implies that the data requirements are identified and defined before the procurement contract is developed. It might be hard to make such changes for all types of contracts in all regions, and could also make some contractors less inclined to participate in tenders reducing the effectiveness of the entire procurement process.

The ideal solution would be to get contractors interested in data sharing by including them in the game, providing game results that could help to increase profits and other benefits. Including contractors in the game implies that some elements in the procurement contracts need to be changed, to make room for additional tasks and data collection for the game, but also to ensure that these changes don’t have adverse effects on the actual performance of the maintenance tasks.

V. Conclusions

Trafikverket has significant influence on the energy efficiency of Sweden as a whole, since they singularly administer all transport maintenance and operations. For Sweden to meet climate goals, Trafikverket, and its associated contractor companies, is an important actor.

The investigation on data availability reveals that a lot of the data is already being collected by Trafikverket. Some of these data can be used readily for gamification, such as operational data for snow ploughs. The forms of other data sources, such as contracts or maintenance logs for switches will need to be changed to fit into the gamification concept.

The data resides in multiple systems, as described earlier. We have described the different data requirements and sources for the gamification of just one scenario: snow ploughing. Further, these are the data requirements and sources for only one level within the institutional structure: the level of the machine operators. While this data is fairly well defined and mostly available, it will still be a challenge to inter-operate among the different sources to create metrics and games for the snow plough drivers.

Even with this inter-operation, it is still one sub-game for snow plough drivers. Without correction factors, it will mean that sub-games will have to be created for every scenario, per region per country. For example, if there are ten snow plough clearing contractors in Stockholm, there will be ten games running for each of these companies, and the same for other counties and for other tasks. Most gamification concepts are flat spaces; players all engage in the same tasks and earn the same rewards. Tasks in large organizations need to be either mapped onto the same conceptual flat spaces, through boundary objects such as interface systems (such as weather and topology data) and through mechanisms such as correction factors for gamification to work across boundaries.

The same concept applies for gamification across levels. However, at higher levels in the institutional structure, the data collection mechanisms are not mature enough yet to implement gamification mechanisms that deliver relevant feedback. Contracts and other forms of data need to be automated before this can be achieved. Further, higher levels in the institutional structure work on different time scales, so engagement over a long period of time is also a concern.

The data collection and implementing gamification needs to be supported by a strong and robust procurement mechanism. Processes such as procurement will be influenced by gamification, since the behaviour changes of people will not be predictable. These processes will also influence the gamification, since they will be the facilitators of the implementation.

As evidenced by the investigation, it is recommended to start implementing gamification with certain kinds of operational tasks in the ELSA gamification, reflecting the current trends in gamification projects. However, in such large organizations implementing even such ostensibly simple projects can become quite complicated. It is also recommended to start small, targeted proof-of-concept implementations, in this complex real-world setting. The Swedish Transport Authority is currently discussing the possibility of doing a pre-commercial procurement of proof-of-concepts for the snow plough and railway examples mentioned. By documenting the development path of such trial projects, the current study can be validated.

Given the enormous resources within and outside Trafikverket that can be leveraged for gamification, and the enthusiasm shown by stakeholders, implementing gamification in a participatory way with stakeholders can be very promising, and can institute a process that will eventually shape Trafikverket's performance and practices on energy.

Acknowledgment

The authors would like to acknowledge the support and encouragement of Martin Strid, Trafikverket, through the duration of the study. The authors would also like to acknowledge the participants of the various workshops, who provided valuable insights into the problem. This feasibility study was conducted while the authors were affiliated with the Department of Transport Science, KTH Royal Institute of Technology.

References

[1] S. Deterding et al., “From Game Design Elements to Gamefulness: Defining ‘Gamification,’” in Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, New York, NY, USA, 2011, pp. 9–15.

[2] K. Huotari and J. Hamari, “Defining Gamification: A Service Marketing Perspective,” in Proceeding of the 16th International Academic MindTrek Conference, New York, NY, USA, 2012, pp. 17–22.

[3] J. Hamari and V. Lehdonvirta, “Game Design as Marketing: How Game Mechanics Create Demand for Virtual Goods,” Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 1443907, 2010.

[4] G. Jipa and I. Marin, “Enterprise gamification in business to consumer (B2C) engagement model,” in Proceedings of the 8th International Management Conference “Management Challenges for Sustainable Development,” 2014, pp. 489–496.

[5] P. Herzig et al., “A Generic Platform for Enterprise Gamification,” in 2012 Joint Working IEEE/IFIP Conference on Software Architecture (WICSA) and European Conference on Software Architecture (ECSA), 2012, pp. 219–223.

[6] R. Farzan et al., “Results from Deploying a Participation Incentive Mechanism Within the Enterprise,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, 2008, pp. 563–572.

[7] J. Thom et al., “Removing Gamification from an Enterprise SNS,” in Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, New York, NY, USA, 2012, pp. 1067–1070.

[8] P. Herzig et al., “Implementing Gamification: Requirements and Gamification Platforms,” in Gamification in Education and Business, T. Reiners and L. C. Wood, Eds. Springer International Publishing, 2015, pp. 431–450.

[9] S. Deterding et al., “Designing gamification: creating gameful and playful experiences,” CHI’13 Extended Abstracts on Human Factors in Computing Systems, 2013, pp. 3263–3266.

[10] J. Hamari et al., “Does Gamification Work? – A Literature Review of Empirical Studies on Gamification,” in 2014 47th Hawaii International Conference on System Sciences (HICSS), 2014, pp. 3025–3034.

[11] E. T. Chen, “The Gamification as a Resourceful Tool to Improve Work Performance,” in Gamification in Education and Business, T. Reiners and L. C. Wood, Eds. Springer International Publishing, 2015, pp. 473–488.

[12] M. Rauch, “Best Practices for Using Enterprise Gamification to Engage Employees and Customers,” in Human-Computer Interaction. Applications and Services, M. Kurosu, Ed. Springer Berlin Heidelberg, 2013, pp. 276–283.

[13] F. Oprescu et al., “I PLAY AT WORK—Ten Principles for Transforming Work Processes Through Gamification,” Front Psychol, vol. 5, Jan. 2014.

[14] Finansdepartementet UR, “Swedish Public Procurement Act.” Regeringskansliets rättsdatabaser, 2007.

Mapping between Pedagogical Design Strategies and Serious Game Narratives

Frederik Van Broeckhoven*, Joachim Vlieghe†, Olga De Troyer‡
Department of Computer Science - WISE*†‡
Vrije Universiteit Brussel
Pleinlaan 2
B - 1050 Brussels, Belgium
Email: Frederik.Van.Broeckhoven@vub.ac.be*, Joachim.Vlieghe@vub.ac.be†, Olga.DeTroyer@vub.ac.be‡

***Abstract*—Successful serious games include a compelling narrative context and empirically validated pedagogical intervention methods. In order to create such games, design teams must consist of a multidisciplinary group of technical and pedagogical experts. In this paper, the authors show how the domain specific modeling language ATTAC-L facilitates communication between designers with different expertise, thus enabling and stimulating multidisciplinary collaboration. As a serious game design tool, ATTAC-L creates a link between the processes of pedagogical design and narrative modeling through its elaborate annotation system. As such, this modeling language enables designers to concentrate on aspects related to their field of expertise without losing oversight of the serious game as a whole. To support these tentative claims, the author present illustrations of how ATTAC-L is used in combination with a specific pedagogical design strategy (i.e. the Intervention Mapping Protocol) for the development of a serious game against cyber-bullying.**

I. Introduction

Serious games that provide a narrative context are generally more effective at conveying knowledge and achieving behavior change. This claim is supported the empirical findings of several studies on effective game design (see see [1] for an overview). Therefore, serious game designers choose to present pedagogical content and interventions through an interactive narrative. To ensure effective learning in such a serious game, the developers are not only faced with the challenge of creating a compelling narrative, but also with the additional challenge of incorporating suitable learning strategies into their narratives. Thusfar, designers are faced with a lack of general guidelines or standards for making such combinations.

To come to well-grounded and effective serious games, different parties should be involved in the development of serious games: game developers as well as pedagogical experts and subject-matter experts. Subject-matter experts bring in their knowledge about the subject of the serious games (e.g., cyber-bullying), pedagogical experts share their knowledge on how to ensure the learning that is aimed for (e.g., knowledge acquisition, attitude change, and/or behavior change), while game developers contribute their experience on how to develop a challenging game. However, pedagogical experts are usually no game or software engineers, and game developers are usually not trained in pedagogical design. This often results in a communication gap that can seriously hinder the proper development of these games [2]. In addition, there is little common knowledge on how to incorporate pedagogical principles and techniques into game narratives. Although some research has attempted to map pedagogical principles onto game mechanics (see e.g., [3]), clear guidelines for designing games in an interdisciplinary setting are lacking.

In previous publications [4], [5], we have argued for the use of a Domain Specific Modeling Language (DSML) to support people with limited technical backgrounds in participating in the specification of serious games narratives. This kind of a modeling language is often visual in nature and its syntax contains the vocabulary of the domain under consideration [6]. We proposed such a DSML - called ATTAC-L - for specifying the storyline, as well as the pedagogical aspects of educational video games. It combines flow chart principles and a natural language based syntax that facilitates participation by non-technical experts in the narrative modeling of a serious game. The output generated by the DSML results in a formal specification that can be processed automatically to generate code.

In previous publications, we focused on the visual representation of the DSML and its natural language-based syntax [4]. In [5], we proposed a mechanism, i.e. pedagogical annotations, to model pedagogical aspects in the narrative of a serious game. These annotations are specified on top of the storyline and allow the modeler to associate pedagogical issues, such as objectives and actions, with particular parts of the story. In these publications, however, we did not yet consider methods or guidelines for identifying these pedagogical aspects or incorporating existing learning strategies. In this paper, we tackle this issue by showing how a well-documented pedagogical design strategy, i.e. the Intervention Mapping Protocol (IMP) [7], provides such guidelines and how its outcomes can be linked explicitly with the narrative of the serious game using the annotations of ATTAC-L.

The paper is structured as follows: Section II discusses related work, section III explains IMP and section IV briefly describes ATTAC-L. Section V discusses how the modeling concepts of ATTAC-L and the annotations in particular can be used to link the narrative to the outcome of an IMP process. Section VI presents conclusions and future work.

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

II. Related Work

IMP has already been successfully applied for the development of serious games and digital intervention platforms. 'PR:EPARe' [8] (Positive Relationships: Eliminating Coercion and Pressure in Adolescent Relationships) is a game developed as part of an intervention program for teaching relationships and sex education for young adolescents in the UK. 'The Gay Cruise' [9] and 'QueerMasters' [10] are both intervention programs addressing HIV-prevention in the Dutch homosexual community. Although the latter two are not explicitly presented as serious games, they are similar in the sense that these cases use virtual narratives in which the characters introduce the participants to the issues at hand.

The main advantages of IMP are the involvement of a wide range of collaborators (technical as well as non-technical) in the development process and the sound foundation the protocol provides for validating the effect of the intervention. The project presented above where successful at reconciling a pedagogical design strategy - in this case IMP - with the process of game development without the use of a dedicate design tool. In other cases, this proved to be a very challenging undertaking due to a large communication gap between the experts involved [11]. While there is a growing need for specialized game authoring tools that take into account pedagogical design principles, current research on this topic is very sparse. Furthermore, to our knowledge, tools specifically based on the effective design strategy of IMP are non-existent at this moment.

Different authoring tools can be used for designing scenario-based serious games, such as interactive digital storytelling tools, e.g., StoryTech[12], Scenejo[13], e-Adventure[14]. Also, several DSMLs were developed for the same purpose, e.g., WEEV[15] and GLiSMo[16]. Most of these systems concentrate on one aspect, mostly the game story. Here, we will not discuss them, we will review work that also aimed for linking pedagogical design principles to the narrative.

EDoS (Environment for the Design of Serious Games) [17] is an interactive authoring environment for serious games. Its purpose is similar to that of ATTAC-L: to help an interdisciplinary team in designing a serious game through a number of standardized steps, from formalizing the pedagogical objectives, to elaborating a scenario and modeling user interactions. The outcome is also "a structured scenario that will be automatically executed by an engine" [17, p. 1]. EDoS focuses on the reuse of available components of different granularity and the creation of serious games for teaching engineering skills. The design process builds on 3 models. The first one is a model of the targeted pedagogical objectives, e.g. professional competences for an engineer. The second model relates pedagogical objectives to pedagogical activities in order to form a pedagogical serious game scenario. These scenarios are created using an adapted version of the IMS-LD (Instructional Management Systems - Learning Design [18] language, and only describe the pedagogical content of the serious game. The third model helps to include the entertaining elements, i.e. the task model that describes the screens with which the users will interact. In contrast with the approach of ATTAC-L, the EDoS approach relies on a specific learning design, i.e. IMS-LD, thus providing limited flexibility. Furthermore, publications related to EDoS do not provide guidelines or methods for identifying pedagogical objectives or constructing serious game scenarios.

TABLE I: Example of matrix for performance objective 'Always comfort the victim', determinants and change objectives

Performance Objective: ***PO*** - Always comfort the victim	
Knowledge	*K1* - Recognize that by comforting the victim, you are making the victim feel better *K2* - Describe ways to comfort a victim that are in line with your personality
Self-efficacy	*Se1* - Express confidence in being able to comfort or provide advice to the victim
Outcome expectancies	*Oe1* - Expect that by comforting the victim, he/she will feel better
Perceived social norms	*Sn1* - Recognize that your friends expect you to comfort or provide advice to the victim

III. Intervention Mapping Protocol

The Intervention Mapping Protocol (IMP) has been developed to aid in the systematic planning and design of behavioral change programs. The protocol stimulates an ecological approach to the design of behavioral change programs focused on health issues [19]. It recognizes the importance and bi-directional influence of individual and environmental factors of behavior (e.g., peers, family relations, school policy). The aim of the IMP is to increase the efficacy of the design process as well as the intervention program itself. It does this by means of a set of six clearly defined steps which include iterative cycles of reviewing evidence of problem-related determinants, selecting and implementing theory based strategies, and consulting stakeholders [7]. IMP encourages its users to document the design process and to create detailed descriptions of the foundations and different steps of the intervention. As such, the protocol also meets the recent and popular demands for more thorough reporting.

In what follows, we first provide a detailed description of IMP and then briefly explain how IMP can be used for the development of serious games.

A. IMP Outline

IMP consists of the following steps: needs assessment, preparing matrices of change objectives, selection of theory-informed intervention methods and practical strategies, development of the intervention program, planning for adoption, implementation and sustainability, and development of an evaluation design.

Step 1: Needs Assessment

The first step in IMP is to define the program goal or health problem(s) that the intervention will tackle. This includes identifying the population at risk and developing an understanding of their environmental context. The program goal is refined into program objectives. Each of these objectives is defined in terms of a desired outcome and has a

Back to Contents

priority. The priorities are set based on the objective's level of relevance, desirability, changeability, as well as the required means and efforts to achieve it [7].

Step 2: Preparing Matrices of Change Objectives

This second step compromises an investigation of the behaviors that can help to reduce the problem and attain the program objectives. In the literature related to the IMP, the term '*performance objectives*' is used to refer to the set of desired behaviors. For each of these performance objectives the program designers must assess which factors influence the performance of the desired behavior. Based on this assessment, *behavioral determinants* are identified. What needs to be changed in relation to these determinants in order to achieve the performance objectives is then formulated in terms of *change objectives*. These change objectives create the basis for the development of the actual intervention steps. For each performance objective, the determinants and change objectives are formulated in a matrix of change objectives. Table I shows and example fragment of such a matrix for a program to decrease cyber-bullying. The performance objective considered is '*always comfort the victim*'. Behavioral determinants are 'knowledge', 'self-efficacy', 'outcome expectation', and 'social norms'. For each of these determinants, the change objectives are given. So for example, for the determinant '*knowledge*' the following change objectives are targeted: '*K1: Recognize that by comforting the victim, you are making the victim feel better*' and '*K2: Describe ways to comfort a victim that is in line with your personality*'.

Step 3: Selection of Theory-informed Intervention Methods and Practical Strategies

During the third step, different methods are selected from a body of available literature and assessed in light of the change objectives. This means that for each of the change objectives, the program developers try to find a method that has been tested and said to impact the type of behavioral change that they intend to achieve [7]. These are then matched with other methods to form practical strategies.

Step 4: Development of the Intervention Program

In this fourth step, all information of previous steps is combined to develop an Intervention Plan. Information from step 4 is also used to revise decisions made in step 3.

Step 5: Planning for Adoption, Implementation and Sustainability

Although listed as step 5, implementation planning runs throughout the whole development process. To ensure that the finished product would be feasible to use in practice, a group of stakeholders (e.g., teachers, schools, school counselors, youth advisory centers) is set up at the onset of the project to provide feedback.

Step 6: Development of an Evaluation Design

Similarly, step 6 runs throughout the whole development process. This step covers the evaluation of the intervention program, which is performed by conducting formative research and assessing effectiveness together with end users.

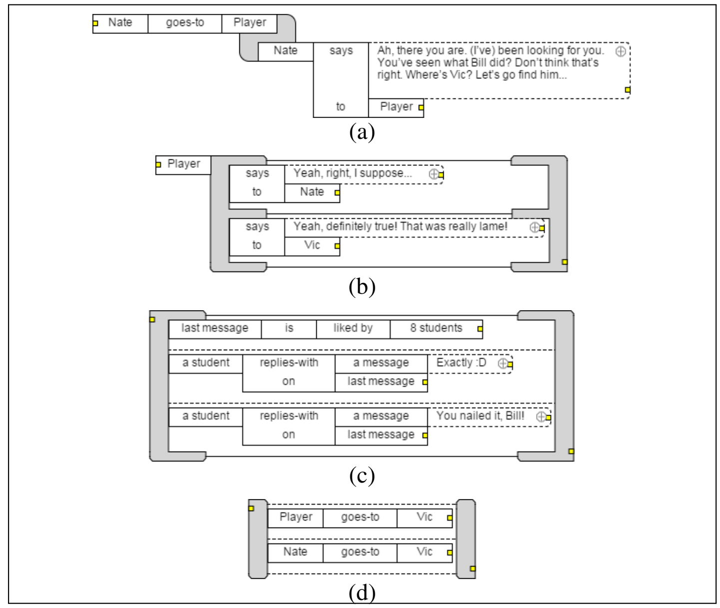


Fig. 1: Several ATTAC-L control structures (fragments taken from figure 5).

B. IMP for the Development of Serious Games

IMP allows for the systematic planning and design of behavioral change programs. A serious game could be part of such a program as a way to accomplish the performance objectives, though it may still be useful to complement it with other components, such as media campaigns, lectures, lessons, etc. . . The serious game should then be designed (as part of step 4) to target the associated change objectives using identified intervention methods and practical strategies (in step 3). The implementation of the serious game is part of step 5 , while its evaluation is part of step 6. Note that step 6 actually runs throughout the whole development process, which is in line with good software engineering practices. IMP is also in line with a user-centered software development approach [20] given the incorporation of stakeholders.

IV. ATTAC-L

ATTAC-L is a Domain Specific Modeling Language for specifying the storyline of educational video games. It combines a syntax based on natural language with flow chart modeling principles to allow both technical and non-technical people to model (i.e., describe in a formal way) the narrative of an educational video game. The output, a formal specification of a game narrative, can be processed automatically to generate code.

A first basic modeling concept in ATTAC-L is a *game move*, i.e. a single action in the narrative [21]. It represents one individual step in the game narrative, either performed by the player or 'automatically' by a non-playable character (NPC). To form a story, game moves should be linked to each other to denote their relative order in the narrative. For

Back to Contents

this, ATTAC-L adopts principles from flow-chart modeling (e.g., UML[1]). The modeler can express sequence (i.e., game moves following each other), choice (i.e., branching, defining alternative story flow paths), and concurrency (i.e., story flow paths that are performed in parallel). In addition, ATTAC-L provides an extra control mechanism to increase its expressiveness, i.e. ‘order independence’. This mechanism allows designers to determine that particular story flow paths must all be performed regardless in what order.

A second important modeling concept is a *brick* (adopted from the StoryBricks[2] framework). Bricks are the basic building blocks used in ATTAC-L to compose the game moves as well as the overall flow. Two classes of bricks are distinguished: *regular bricks* and *control-bricks*.

Regular bricks are used to construct game moves. They correspond to the smallest meaningful unit that exists in the context of a story. This can be an act to be performed, a tangible object that can perform or undergo the act, a state, or a value. Game moves are constructed by interconnecting bricks according to rules that are based on a controlled natural language [22]. The result is a construct that reads as a simple sentence and denotes a game play activity. A regular brick is graphically represented by a rectangle containing a word or word-group that gives it a meaning. In figure 1, the white rectangles containing words are regular bricks. As can be seen, the bricks are combined to form simple sentences, i.e., the game moves. For example, the game moves in figure 1a are “`Nate | goes-to | player`” and “`Nate | says | ‘Ah, there you are ...’ | to | player`”. We refer to [22] for details about constructing game moves.

Control bricks are used to express temporal relationships between game moves. As stated earlier, ATTAC-L uses a flow-based structure for this because empirical evidence shows that this is more suitable for non-technical users [4]. Because the target users are not familiar with the typical flow-chart notation used in modeling languages such a UML, we have decided to use bricks as well to express these control structures. In figure 1, the grey bricks that interconnect game moves are control bricks. We distinguish bricks to model sequence, choice, order independence and concurrence. A sequence-brick is visualized by a chair-like brick that interconnects game moves. In Figure 1a it is used to specify the order between the two game moves, so that Nate first goes to player after which he says ‘Ah, there you are. I’ve been looking…’ to player. A choice-brick encapsulates alternative storyline paths. Figure 1b shows a choice for the player between saying ‘Yeah, right I suppose’ or ‘Yeah, definitely true! That was really lame!’. Similarly, an order independence-brick encapsulates storyline paths that must all be performed, but in any order. Figure 1c expresses that 8 students like the message and two reply with ‘Exactly :D’ or ‘You nailed it, Bill’, but the order in which they do it is irrelevant. Finally, a concurrence-brick encapsulates storyline paths that are performed in parallel. Figure 1d specifies that both player and Nate go to Vic at the same time.

[1] Unified Modeling Language

[2] http://www.storybricks.org (project discontinued since 03-05-2015)

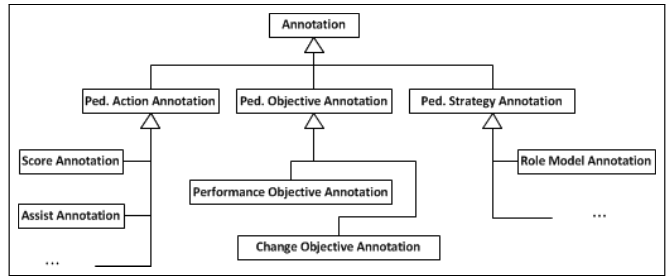


Fig. 2: Annotation class-hierarchy

A third modeling concept is called a scenario. This concepts was introduced to deal with the complexity of large models. It allows designers to divide a large narrative into smaller logical units. The so called scenario-bricks refer to different scenarios by name. They function as placeholders for scenarios and can be arranged like game moves in a story line. Figure 5 shows a storyline decomposed into scenarios. The overall scenario ‘Introduction to Comforting’, contains two scenarios ‘Vic’s bullying situation’ and ’Nate demonstrates comforting’. The corresponding scenario-bricks (shown at the top of the example) are linked to each other using a sequence-brick. This means that first the scenario ‘Vic’s bullying situation’ is performed, next the scenario ‘Nate demonstrates comforting’.

The fourth modeling concept in ATTAC-L enables the modeler to specify additional information (e.g. pedagogical relevant information) to parts of the storyline model, i.e. annotations [5]. Annotations are represented graphically by means of small and square-like bricks, called annotation-bricks. Each brick contains an icon that denotes its meaning. They can be attached to game moves and scenarios. Annotations allow the modeler to specify and add relevant non-storyline related aspects on top of the storyline model, such as pedagogical aspects (e.g. pedagogical interventions) . This prevents that the specification of different aspects (here learning and gaming) are entangled. It allows for a clear separation between the narrative content and the educational aspects while the modeler can still relate the latter to the story flow.

Currently, ATTAC-L distinguishes between pedagogical action annotations, pedagogical objective annotations, and pedagogical strategy annotations. Pedagogical action annotations are used to specify particular pedagogical oriented actions that should be performed in the story, such as providing additional information, assistance or feedback (i.e. pedagogical interventions). Pedagogical objective annotations are used to explicitly relate behavioral change objectives to scenarios. An example of such an objective could be “to practice the multiplication tables of 1 to 10”, but also “to realize the impact of cyber-bullying”. Pedagogical strategy annotations are used to associate pedagogical strategies to the story or scenarios. Examples of such strategies are ‘drill & practice’, ‘trial and error’, or ‘learning by doing’.

V. Linking Narrative and IMP Outcome

As explained in section III, a serious game could be used to achieve the performances objectives identified by through IMP. When a serious game is part of a program developed

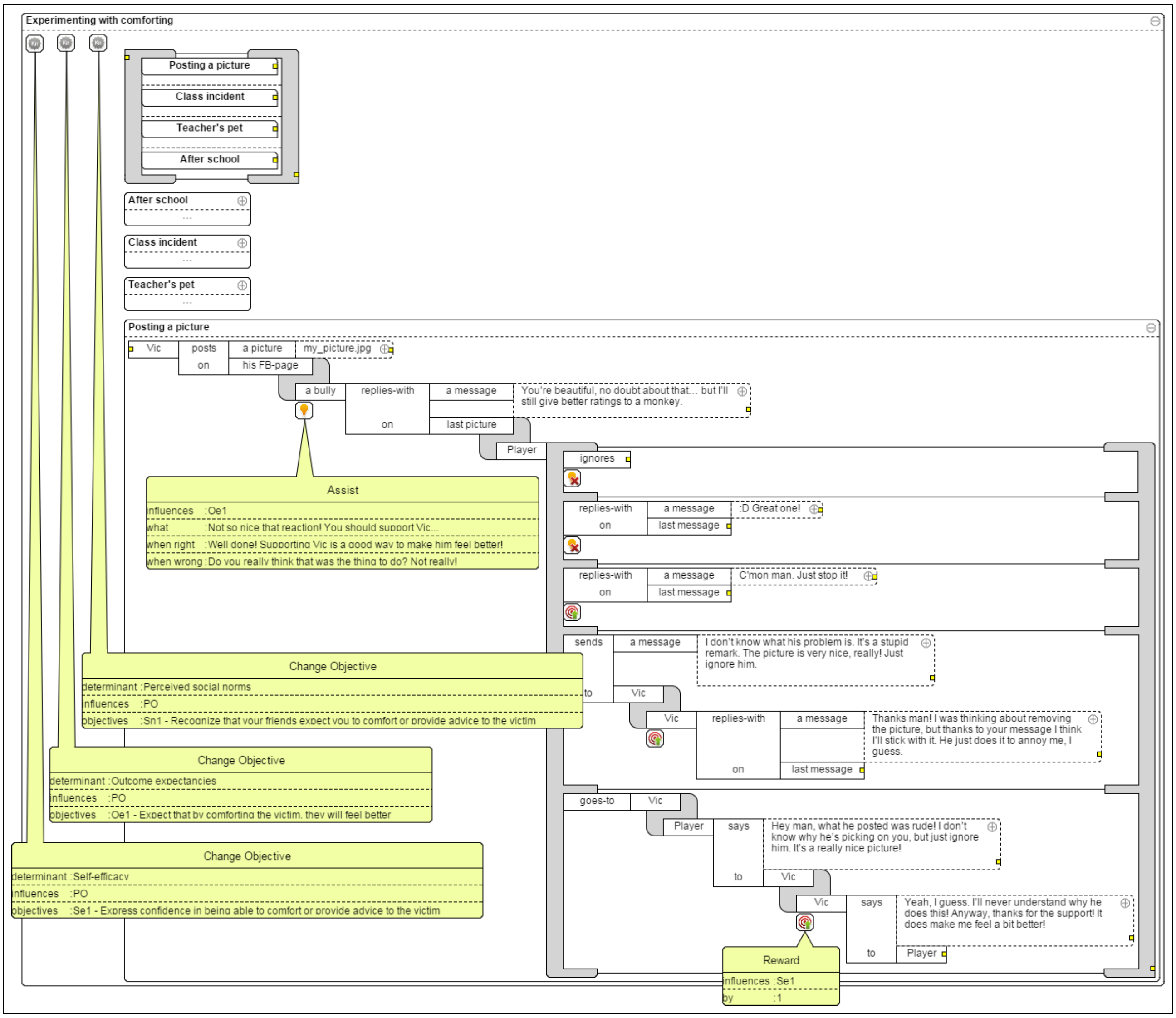


Fig. 4: A scenario with change objective annotations for determinants 'self-efficacy', 'outcome expectations' and 'perceived social norms' and ped. action annotations for 'scoring' and 'assistance'.

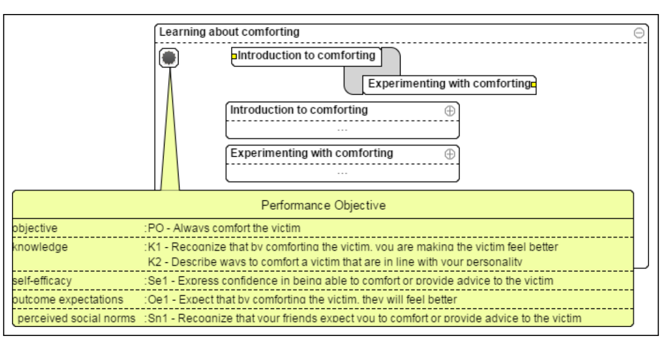


Fig. 3: A scenario with a performance objective 'always comforting the victim' and two sub-scenarios (both detailed in figures 5 and 4)

with IMP (in step 4), its components and their functions should be closely related on the outcome of the IMP process (step 1 to step 3). Stated differently, the serious game should target one or more of the change objectives associated with the performances objectives (outcome of step 2), while using intervention methods and practical strategies identified in step 3 of the IMP process.

The modeling concepts of ATTAC-L can be used to link the narrative to the outcome of the IMP process. This creates two important advantages: first, allowing the designers to verify whether the serious game and the IMP outcomes are in accordance; secondly, automatically providing documentation of the design process as required by IMP and other pedagogical design strategies.

To illustrate our approach, we use the storyline models given in figure 5 and 4. The context of our example is the Friendly ATTAC project [23], which aims to develop a behavioral change program for cyber-bullying based on IMP. One of the program goals is to teach youngsters to understand the concept of cyber-bullying and to react in an adequate way when confronted with cyber-bullying incidents. The scenario that we use in this paper, focuses on the performance objective described as "comforting a victim after witnessing a cyber-bullying incident". Using IMP, the behavioral determinants and change objectives associated with this performance objective were identified (i.e., the matrix of changes objectives, see table I – we refer to the work of our colleagues [24] for more details about the IMP program developed), as well as a set of intervention methods.

We use the annotation concept of ATTAC-L to specify how the outcome of the IMP process can be incorporated into the narrative model of the serious game, i.e., the performance objectives, change objectives, and intervention methods. This way, modelers (and stakeholders) can clearly identify which parts of the storyline attribute to which performance objectives and behavioral determinants. The annotations also help to determine the connections that exist within the serious game between the developed intervention methods and the narrative.

We start by explaining how the performance objectives and their further refinement into change objectives can be specified for the narrative model. We then do the same for the techniques and strategies of the intervention methods.

A. Linking Performance Objectives and Change Objectives

Pedagogical objective annotations can be used to specify which part of a storyline model is associated with a particular performance objective. For this purpose, we have created the performance objective annotation, which is a subtype of a pedagogical objective annotation that is associated with a scenario (also see figure 2, which provides an illustration of the different annotations and their subtypes dedicated to IMP using an UML class diagram). A performance objective annotation has a set of parameters that define the performance objective and the determinants and change objectives related to this performance objective. In figure 3, an example of such a performance objective annotation is presented. The performance objective is 'always comfort the victim' (see table I). The behavioral determinants that should be influenced by the scenario are 'knowledge', 'self-efficacy', 'outcome-expectations', and 'perceived social norms'.

ATTAC-L also allows the modeler to reflect the refinement of a performance objective into change objectives. A scenario annotated with a performance objective can be divided into different sub-scenarios, each dealing with one or more change objectives and their corresponding behavioral determinants. This means that in each of these scenarios', specific behavioral determinants will be tackled. The scenarios will include intervention methods that affect determinants with the aim of achieving the change objective. All sub-scenarios are marked to indicate with which specific change objectives they are dealing. They are annotated with another type of pedagogical objective annotation called a change objective annotation (see figure 2). The sub-scenarios of the scenario 'learning about comforting' used in figure 3 are given in figures 5 and 4. The first one is annotated with change objectives related to the determinant 'knowledge'. It states that the scenario contains intervention methods specifically for increasing the knowledge of the player about comforting a victim. The second one has an analogous purpose for the determinants 'self-efficacy', 'outcome expectations' and 'perceived social norms'. The specification of how and where the intervention methods are used is explained in the next section.

B. Linking Intervention Methods

Intervention methods can be embedded in a scenario in two ways: (1) as game mechanics that are unrelated but complementary to the storyline, guiding the player in a certain way. For example by presenting a pop-up to the player with extra hints or by giving a positive score after making a correct choice. Or (2) as events expressed directly in the narrative. For example a character demonstrating the correct behavior or giving positive feedback during a conversation.

In order to express an intervention method through game mechanics (method type (1)) designers can use the existing pedagogical action annotations. Recall that pedagogical action annotations are used to specify required actions that have a particular pedagogical intention, such as providing additional information, assistance, or feedback. Different subtypes of pedagogical action annotations have already been defined, like the score annotation that changes a player's score and the assist annotation that gives hints to the player.

To create the link with the change objective, the pedagogical action annotation has been extended to include an argument that denotes the targeted change objective. Figure 4 shows the use of the score annotation as a method for achieving the change objective 'Se1' (see Table I). Similarly, the same figure uses an assist annotation for linking the change objective 'Oe1' (see table I) to the method of giving hints to the player about the correct choice.

Intervention methods can also be embedded by expressing them as a part of the narrative, e.g., involving characters or showing specific situations that help the player to obtain the intended behavior (method type (2)). To model this we use the existing pedagogical strategy annotation. In this case, the scenario containing the intervention should be annotated as a whole. Since many different intervention methods are possible – each with different characteristics – this annotation type is an abstract one (i.e. we cannot define all of its properties; cf. abstract class in UML). Concrete subtypes, whose properties have all been specified, should be defined for different methods. For instance, the role-model annotation has been defined for the 'modeling' or 'observational learning' principle used in the Social Cognitive Theory (SCT) [25]. This annotation is illustrated in figure 5. In the example, a friend of player 'Nate' shows the correct behavior of comforting the victim. The role-model annotation is attached to the scenario to indicate that the scenario as a whole implements this role-modeling method. The annotation itself contains the parameters to indicate the change objectives it influences ('K1' and 'K2', see table I) and which in-game character is acting as the role-model.

VI. Conclusions and Future Work

In this paper we established a link between the process of pedagogical design based on IMP and the narrative

Back to Contents

modeling in ATTAC-L. By combining pedagogical design and narrative modeling, designers can maximize the efficacy of serious games. Furthermore, the integration of elements from the pedagogical design process into ATTAC-L through an elaborate annotation system facilitates the communication between technical and non-technical people. As such, ATTAC-L represents a serious game design tool that enables and stimulates multidisciplinary collaboration.

We showed that relevant outcomes of a pedagogical design process based on the IMP (i.e., performance objectives, change objectives, and intervention methods) can be linked to the story model using corresponding pedagogical annotations. ATTAC-L and its elaborate annotation system allow designers to specify different aspects like learning goals and story flow separately. The clear separation of narrative and educational aspects helps to increase the designers' overview of the different elements of the serious game. As a consequence, designers with different expertise can concentrate their attention on particular aspects of the game, while maintaining a clear view of its relations to all others aspects of the game.

Currently, ATTAC-L is used in the Friendly-ATTAC project [23] which applies the IMP to tackle the issue of cyber-bullying. Using ATTAC-L, the project team is developing a serious game for youngsters based on intervention methods identified with IMP. The output of ATTAC-L is used to generate parts of the code to help speed up the development process. Thus far, the pedagogical annotations have not been included in the generated code. We believe that it should be the ambition of future research endeavors to develop a standard for transforming the pedagogical annotations into functional coding that can be integrated into the main flow of the serious game code. Another path to consider is the linking between ATTAC-L and pedagogical design strategies other than IMP. We believe that even though the approach presented in this paper focused specifically on IMP, the principles for linking pedagogical design strategies and narrative modeling are generally applicable and therefore provide lots of flexibility.

Acknowledgment

This research is funded by IWT (Institute for Science and Technology)[3] under the Friendly ATTAC project.

References

[1] M. Dondlinger, "Educational Video Game Design: A Review of the Literature," *Journal of Applied Educational Technology*, vol. 1, no. 4, pp. 21–31, 2007.

[2] O. De Troyer and E. Janssens, "Supporting the requirement analysis phase for the development of serious games for children," *International Journal of Child-Computer Interaction*, vol. 2, no. 2, pp. 76–84, May 2014.

[3] T. Lim, M. B. Carvalho, F. Bellotti, S. Arnab, S. de Freitas, S. Louchart, N. Suttie, R. Berta, and A. De Gloria, "The lm-gm framework for serious games analysis."

[4] F. Van Broeckhoven and O. De Troyer, "ATTAC-L: A Modeling Language for Educational Virtual Scenarios in the Context of Preventing Cyber Bullying." in *2nd International Conference on Serious Games and Applications for Health*. IEEE, May 2013, pp. 1–8.

[3]http://www.iwt.be

[5] ——, "Specifying the Pedagogical Aspects of Narrative-Based Digital Learning Games Using Annotations," in *Proceedings of the 9th International Conference on the Foundations of Digital Games*. Society for the Advancement of the Science of Digital Games, 2014.

[6] S. Kelly and J. Tolvanen, *Domain-specific modeling: enabling full code generation*. John Wiley & Sons, Mar. 2008.

[7] L. K. Bartholomew, G. S. Parcel, G. Kok, N. H. Gottlieb, and M. E. Fernández, *Planning Health Promotion Programs: An Intervention Mapping Approach*. John Wiley & Sons, 2011.

[8] S. Arnab, K. Brown, S. Clarke, I. Dunwell, T. Lim, N. Suttie, S. Louchart, M. Hendrix, and S. de Freitas, "The development approach of a pedagogically-driven serious game to support Relationship and Sex Education (RSE) within a classroom setting," *Computers & Education*, vol. 69, pp. 15–30, Nov. 2013.

[9] G. Kok, P. Harterink, P. Vriens, O. de Zwart, and H. Hospers, "The gay cruise: Developing a theory- and evidence-based Internet HIV-prevention intervention," *Sexuality Research & Social Policy*, vol. 3, no. 2, pp. 52–67, 2006.

[10] J. Mikolajczak, G. Kok, and H. J. Hospers, "Queermasters: Developing a theory- and evidence-based internet hiv-prevention intervention to promote hiv-testing among men who have sex with men (msm)," *Applied Psychology*, vol. 57, no. 4, pp. 681–697, 2008.

[11] O. De Troyer, E. Janssens, and H. VandeBosch, "How to kick off the development process of a serious game?" in *CHI workshop: Lets talk about Failures: Why was the Game for Children not a Success?*, 2013.

[12] S. Gbel, L. Salvatore, R. Konrad, and F. Mehm, "Storytec: A digital storytelling platform for the authoring and experiencing of interactive and non-linear stories," in *Interactive Storytelling*, ser. Lecture Notes in Computer Science, U. Spierling and N. Szilas, Eds. Springer Berlin Heidelberg, 2008, vol. 5334, pp. 325–328.

[13] S. Weiss, W. Mller, U. Spierling, and F. Steimle, "Scenejo an interactive storytelling platform," in *Virtual Storytelling. Using Virtual Reality Technologies for Storytelling*, ser. Lecture Notes in Computer Science, G. Subsol, Ed. Springer Berlin Heidelberg, 2005, vol. 3805, pp. 77–80.

[14] J. Torrente, A. del Blanco, E. Marchiori, P. Moreno-Ger, and B. Fernandez-Manjon, "¡e-adventure¿: Introducing educational games in the learning process," in *Education Engineering (EDUCON), 2010 IEEE*.

[15] E. J. Marchiori, A. del Blanco, J. Torrente, I. Martinez-Ortiz, and B. Fernández-Manjón, *Journal of Visual Languages & Computing*, no. 6, pp. 443–452, dec.

[16] E. M. Hirdes, N. Thillainathan, and J. M. Leimeister, "Towards modeling educational objectives in serious games," *Pedagogically-driven Serious Games (PDSG 2012)*, 2012.

[17] C. Tran, S. George, and I. Marfisi-Schottman, "Edos: An authoring environment for seriousgames design based on three models," in *4th Europeen Conference on Games Based Learning ECGBL2010*, october 2010, pp. 393–402.

[18] R. Koper and B. Olivier, "Representing the Learning Design of Units of Learning." *Educational Technology & Society*, vol. 7, no. 3, pp. 97–111, Nov. 2003.

[19] J. F. Sallis, N. Owen, and E. B. Fisher, "Ecological models of health behavior," *Health behavior and health education: Theory, research, and practice*, vol. 4, pp. 465–486, 2008.

[20] B. Gransson, J. Gulliksen, and I. Boivie, "The usability design process integrating user-centered systems design in the software development process," *Software Process: Improvement and Practice*, vol. 8, no. 2, pp. 111–131, 2003.

[21] C. A. Lindley, "Story and narrative structures in computer games," 2005.

[22] F. Van Broeckhoven, O. De Troyer, and J. Vlieghe, "Using Controlled Natural Language for Specifying the Narratives of Digital Serious Games," Vrije Universiteit Brussel, Tech. Rep., 2015.

[23] Friendly ATTAC, "http://www.friendlyattac.be/en/."

[24] A. Desmet, K. Van Cleemput, S. Bastiaensens, K. Poels, H. Vandebosch, M. Verloigne, G. Vanwolleghem, L. Mertens, and I. De Bourdeaudhuij, "Mobilize bystanders against cyberbullying among adolescents the friendly attac project." Submitted.

[25] A. Bandura, "Social cognitive theory of self-regulation," *Organizational behavior and human decision processes*, vol. 50, no. 2, pp. 248–287, 1991.

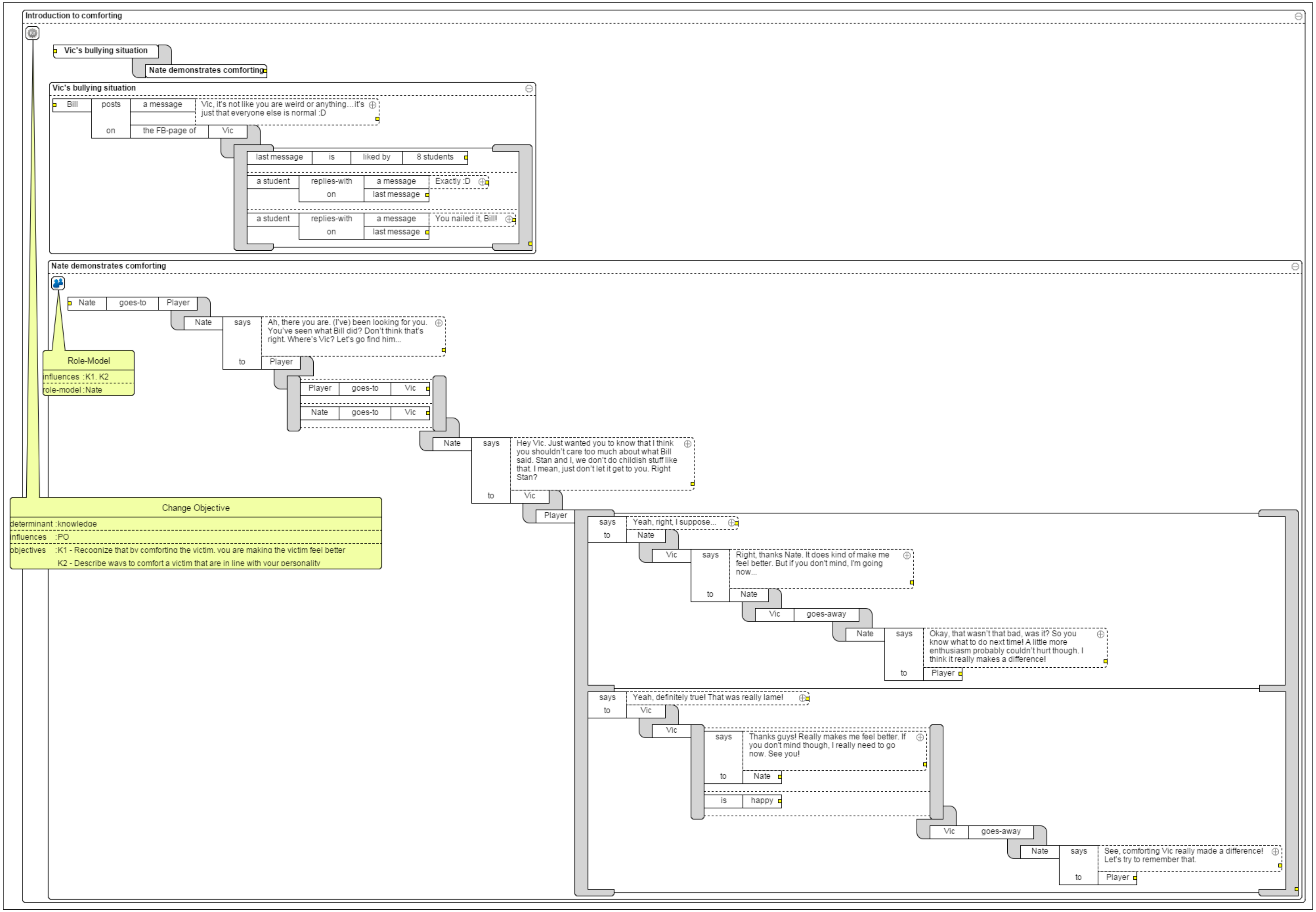


Fig. 5: A scenario with a change objective annotations for determinant 'knowledge' and a ped. strategy annotation for 'role-modeling'.

A systematic mapping study on gamified software quality

Juan Vargas-Enríquez, Lilia García-Mundo
Institute of Technologies and Information Systems
University of Castilla-La Mancha
Ciudad Real, Spain
{JuanAntonio.Vargas, LiliaCarmen.Garcia}@alu.uclm.es

Marcela Genero, Mario Piattini
Institute of Technologies and Information Systems
University of Castilla-La Mancha
Ciudad Real, Spain
{Marcela.Genero, Mario.Piattini}@uclm.es

***Abstract*— Gamified software is currently very popular, and it is expected that it will be widely adopted over the coming years. The social impact of gamified software will probably be very high, and we therefore believe that the assessment and improvement of gamified software quality may be necessary. The aim of this paper is to present a systematic mapping study (SMS) carried out to discover the current state of the research on software gamification quality, in order to identify gaps that merit rigorous future investigation. This SMS allowed us to select 35 papers found in five digital libraries up to April 2014. This paper summarizes the main issues of the planning and the conducting of the SMS. The main results of the data synthesis are detailed and future work is also outlined.**

Keywords—gamification; gamified software quality; ISO/IEC 25010; systematic mapping study

I. INTRODUCTION

The term "Gamification" originated in the digital media industry in 2008, but was not widely adopted until the second half of 2010. Gamification is defined as the use of game design elements in non-game contexts [1]. Game design elements are constituent parts that are found in most games, readily associated with games, and found to play a significant role in gameplay [1]. Examples of game design elements are: points, levels, leaderboards, badges, challenges/quests and engagement loops, among others [2]. The non-game context refers to the use of elements of games for purposes other than their normal expected use as part of an entertainment game [1]. The non-game context can be as varied as crowdsourcing, social networks, loyalty programs, marketing, industry, education, health initiatives, etc. This definition explicitly excludes serious games, another emerging technology that uses game elements, as serious games are complete games, whereas gamification is a way of designing applications and services [3]. However, the boundary between "game" and "gamification" can often be blurry [1].

Gamification is a valuable approach as regards making non-game products, services or applications more enjoyable, more motivating and/or engaging to use [1]. Gamified Software has undergone a very important growth; Gamified software is currently very popular and this means by which to engage users is becoming a growing trend. It is therefore expected that this key emerging technology will be widely adopted over the coming years. M2 Research estimates that the market spend on gamification solutions, applying game mechanics and behavioral analytics in non-traditional applications will reach 2.8 billion USD by 2016 [4]. According to Gartner's 2013 Hype Cycle Special Report [5], gamification is at the top of the peak of inflated expectations, with 5-10 years of mainstream adoption, which is when "Early publicity produces a number of success stories—often accompanied by scores of failures. Some companies take action; many do not". The growing interest in gamification is also reflected in the academic context [6]. The aforementioned aspects reveal that the social impact of gamified software on users and many areas of society will probably be very high, and we therefore believe that the assessment and improvement of gamified software quality may be necessary. Measures with which to measure gamified software quality characteristics are therefore required. These measures will permit the evaluation of gamified software quality and help determine, for instance, whether one gamification solution for a particular piece of software is better than another.

The goal of the research presented in this paper is to discover the current state of the research on gamified software quality and to identify gaps that merit rigorous future investigation. In order to achieve this goal we carried out a systematic mapping study (SMS) following the guidelines proposed by [7] [8] [9].

The remainder of the paper is organized as follows: Section II presents the related work. Section III provides a description of how the SMS was planned, while Section IV explains how the SMS was conducted. Section V sets out the data synthesis and results. An analysis of the threats to the validity of the SMS is presented in Section VI, and finally, our conclusions and future work are presented in Section VII.

II. RELATED WORK

As related work, we will introduce some existing literature reviews that are to some extent related to that presented herein.

Xu [10] conducted a literature review on web application gamification and analytics. In this review the author surveyed 4 gamified applications and 3 computer games. He concluded that much can be learned from the principles and practices of gaming, although the focus must be kept on the limits and potential traps embedded in gamification.

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

Hamari, Koivisto, and Sarsa [6] conducted a literature review focused on the empirical evidence of the motivational affordances (gamification design elements) implemented and the psychological (enjoyment, engagement, motivation, etc.) and behavioral (participation, effectiveness of learning, etc.) outcomes measured. They found 24 papers published between 2008 and 2013. They concluded that even though gamification seems to have positive effects, these are greatly dependent on both the context in which gamification is being implemented and the users who employ it.

Pedreira, Garcia, Brisaboa, and Piattini [11] conducted a SMS on gamification in software engineering. As a result of the SMS they found 16 primary studies, published between 2011 and 2013. They came to three major conclusions; 1) More research efforts analyzing the impact of gamification on software engineering are needed, 2) Most studies share a lack of methodological support that would make their proposals replicable in other settings, and 3) The integration of gamification into an organization's existing tools is also an important challenge that needs to be addressed in this field.

In a previous work we conducted an SMS focused on serious game quality [12]. In that work we found 112 papers published up April 2013. We came to three conclusions: a) it is necessary to address the quality of serious games from early stages of development, b) more empirical evidence is needed as regards proposals that address the quality of serious games, and c) the experiments need to be replicated by researchers other than those who proposed the serious game.

The literature review presented in this paper is different from those previously described in several respects:

- Its goal: our goal is to collect all the existing literature on gamified software quality, and not only the empirical studies.
- Procedure: the SMS presented in this paper has been carried out using the guidelines proposed for SMSs in [7] [8] [9].
- Period of time: the time period is longer and more recent; more studies have therefore been included: we searched for all papers published up to April 2014 without setting a start date.

III. PLANNING THE SMS

The protocol developed to conduct this SMS includes objectives, research questions, search strategy, selection strategy (inclusion/exclusion criteria), study selection procedure, data extraction strategy, and data synthesis.

The main research question that will lead to the objective of this SMS being achieved is:

What is the state-of-the-art of the research on gamified software quality?

Since the main research question is quite general, it was subsequently divided into 6 research questions (see Table I), in order to provide a more specific understanding of the topic being investigated.

TABLE I. RESEARCH QUESTIONS

Research questions	Main motivation
RQ1. What gamified software quality characteristics have been addressed?	To identify the quality characteristics and to map them onto the quality characteristics proposed in ISO/IEC 25010 [14].
RQ2. What research methods have been used when investigating gamified software quality?	To determine whether or not the research has been validated and to discover what research method was used to validate it.
RQ3. What has the outcome of the research been?	To discover the outputs that are produced when investigating software gamification quality.
RQ4. On which software artifacts from gamified software has the research on quality been focused?	To discover whether gamified software quality has been researched throughout the whole software development lifecycle or whether it has focused solely on certain parts.
RQ5. What gamification elements have been used in gamified software?	To discover what gamification elements have been considered in the initiatives related to gamified software quality.
RQ6. What have the application areas of gamification been?	To determine the application area on which gamified software quality initiatives have been focused.

The search string used to collect the papers from sources was constructed using the steps described in Brereton et al [13]. The major search terms would, at first sight, appear to be “quality” and "gamification", but as quality is a multidimensional term, we decided to use the search term “evaluation” and its related terms, because we found that these terms were frequently used in other papers dealing with software quality. The other major search term was “gamification”, but we also used the terms “gameification”, “ludification” and “funware”, since these terms were also used before the term “gamification” was widely adopted. The major search terms were therefore “evaluation” and “gamification” and we built the search string with the alternative terms within the search terms, or with synonyms, as shown in Table II. We searched in 5 of the most important digital resources: Scopus, Science@Direct, IEEE Digital Library, ACM Digital Library, and Springer. The search string was applied to the title, abstract and keywords.

The papers which were included were those dealing with gamified software quality that had been written in English. We decided to include journals, conferences and workshop papers published up to April 2014 without fixing a starting year in order to make this SMS as complete as possible. We excluded

TABLE II. MAJOR SEARCH TERMS AND THEIR ALTERNATIVE TERMS

Major Terms	**Alternative terms**
Evaluation	(evaluat* OR assess* OR measur* OR test*)
Gamification	(gamif* OR gameif* OR ludif* OR “funware”)

papers related to serious games or game-based learning, or in which the paper´s contributions was not related to the evaluation/assessment/measuring or testing of quality characteristics.

The mechanics of the entire selection procedure were: the first author reviewed the paper and another author then verified it. Any discrepancies were resolved by means of a consensus being reached between the four authors, taking into account the full text of the paper.

In order to ensure that the same criteria were used for the data extraction and classification of the selected papers, a strategy based on providing a classification scheme based on the research questions was defined. This classification scheme consisted of 6 dimensions (one for each research question) with several dimensions each one. For research question 1, the quality characteristics categories were defined on the basis of the ISO/IEC 25010 standard [14] due to it is the current standard of software product quality. For research question 2, the classification of research methods suggested in [15] was used. The categories for the remaining dimensions were initially defined prior to the data extraction and needed to be refined. A detailed description of the classification scheme can be accessed at http://alarcos.esi.uclm.es/SMS-GamificationQuality/.

IV. Conducting the SMS

Four researchers took part in the whole process. The planning of the SMS began in December 2013. All papers related to gamified software quality published up to April 2014 were retrieved in April 2014. The paper selection process is shown in Fig. 1; 254 papers were found, and of these 39 were discarded because they were duplicate papers (the same paper in a different source). The title and abstract of each paper were subsequently reviewed and the number of papers selected was reduced to 56. Inclusion and exclusion criteria were applied to the full text, and 30 more papers were discarded; 26 papers were eventually selected as primary studies. Some months after our paper selection process finished, the literature review carried out by Hamari, Koivisto, and Sarsa [6] was published. Upon considering the similarity between the topic addressed by [6] and our SMS, we decided to check the papers selected in [6] to discover whether or not they coincided with ours. We found that only 5 of the papers selected in [6] coincided with our selection, and believe that this lack of coincidence may be owing to differences in the search string. However, Hamari, Koivisto, and Sarsa [6] did not explicitly describe the search string. In order to be fair and rigorous, we checked all 19 of the papers that did not coincide and applied the inclusion/exclusion criteria (see Table 4) to them. We found 9 more papers that met the inclusion/exclusion criteria of this SMS, and they were therefore included. Most of the papers from Hamari, Koivisto, and Sarsa [6] that were not included were focused on serious games and not strictly on gamification. A detailed explanation of the reasons that led us to include or exclude the papers in [6] in this SMS, and the list of the 35 primary studies selected in this SMS can be accessed at http://alarcos.esi.uclm.es/SMS-GamificationQuality/.

Although the protocol was initially defined during the planning of the SMS, it was necessary that all the authors review and refine it during the execution of the SMS. The identification and selection of primary studies was performed by the first two authors. In order to reduce the risk of a publication being incorrectly included in or excluded from the SMS, each paper was reviewed by at least two authors. In those cases in which these two authors had conflicting views, it was necessary for a third and a fourth author to review the publication and make a final decision.

Some search engines in databases or digital libraries have limitations when using complex Boolean search strings. When a database or digital library did not allow the use of complex Boolean expressions, the search string was split or modified to accommodate the limitations of the search engine. The aim was to obtain the same results that had been achieved using the original search string.

V. Data Synthesis and Results

In this section, the answers to each of the research questions shown in Table I are presented. Some other additional results are detailed at the end of the section.

A. *RQ1. What gamified software quality characteristics have been addressed?*

The process used to match the characteristics shown in the ISO/IEC 25010 standard [14] with the characteristics investigated in each paper is described as follows. The full text of the paper was read in order to search for the quality characteristics that were addressed by the researchers, and the standard was then viewed in order to find the characteristic or characteristics that, in the author's opinion, best matched the characteristics found in the paper. In the review of the full text of the selected papers it was found that the majority of them did not explicitly mention that an assessment of the quality of gamification was being made. The results revealed that most researchers are interested in assessing various aspects of the use of the gamified software, and particularly its effectiveness as regards motivating and engaging users in desired behaviors.

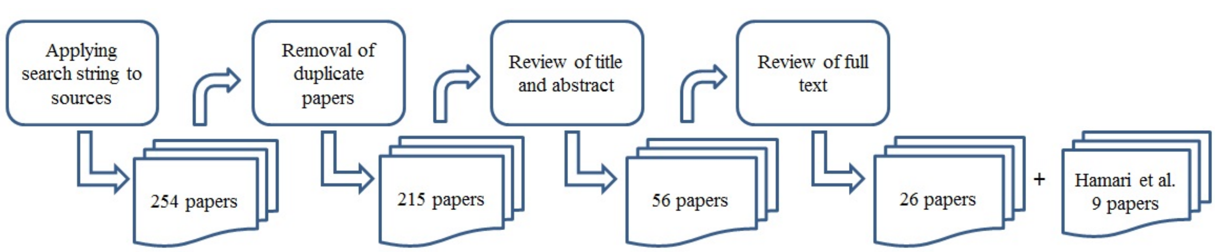


Fig. 1. Paper selection process

Another of the most frequently evaluated aspects was user satisfaction with the gamified software. Most of the papers (32, 91.43%) dealt with the quality characteristics of the quality in use model. The characteristic most frequently addressed was effectiveness (30 papers, 85.71%). In these papers, effectiveness was a measure of how the gamified application achieved its goals to motivate and engage users to achieve the desired behavioral change. These results were then compared with those concerning non-gamified software to observe whether there was any kind of improvement. Another important finding is that the main factor behind the use of gamification is user engagement (12 papers, 34.29%), followed by engagement/motivation (9 papers, 25.71%) and motivation (9 papers, 25.71%). The remaining 5 papers did not mentioned why gamification was used. In papers focused on engagement, the researchers attempted to encourage users to continue doing certain tasks or activities, while in those focused on motivation the researchers wished to increase the users' participation in certain tasks or activities. Following effectiveness is satisfaction (20 papers, 57.14%). The most frequently investigated sub-characteristics as regards satisfaction were pleasure (18 papers, 51.43%) and usefulness (8 papers, 22.86%). These papers assess the effect of gamification on several aspects of user satisfaction such as enjoyment, fun, or the perception of usefulness of the gamified software. A summary of the quantitative results of the characteristics of the quality in use model is shown in Table III.

Only 11 papers (31.43%) dealt with the quality characteristics of the product quality model. The characteristics of the product quality model that were most frequently researched were usability (8 papers, 22.86%) followed by functional suitability (3 papers, 8.57%). The most frequently investigated sub-characteristic of usability was operability (7 papers, 20%), followed to a far lesser extent by the sub-characteristics of user interface aesthetics and learnability (1 paper each). The papers related to usability, focused on the usability of the entire application, not just those aspects related to gamification. A summary of the quantitative results of the characteristics of the product quality model is shown in Table IV.

A detailed description of the process used to match the characteristics in the ISO/IEC 25010 standard [14] with the

TABLE III. DISTRIBUTION OF PAPERS ACCORDING TO CHARACTERISTICS OF THE QUALITY IN USE MODE

Quality Characteristic	Number of papers	Primary studies references
Effectiveness	30	P1, P2, P4, P5, P6, P7, P8, P9, P10, P11, P12, P13, P15, P16, P17, P18, P19, P20, P21, P23, P24, P25, P26, P27, P28, P29, P30, P31, P33, P34
Satisfaction-Pleasure	18	P1, P2, P4, P5, P6, P7, P8, P12, P13, P15, P17, P20, P26, P28, P29, P31, P32, P35
Satisfaction-Usefulness	8	P5, P7, P17, P23, P25, P28, P31, P32
Total	**32**	

TABLE IV. DISTRIBUTION OF PAPERS ACCORDING TO CHARACTERISTICS OF THE PRODUCT QUALITY MODEL

Quality Characteristic	Number of papers	Primary studies references
Functional suitability-Functional appropriateness	3	P3, P22, P29
Usability-User Interface Aesthetics	1	P17
Usability-Operability	7	P5, P7, P12, P13, P14, P31, P35
Usability-Learnability	1	P14
Total	**11**	

characteristics investigated in the papers can be accessed at http://alarcos.esi.uclm.es/SMS-GamificationQuality/.

B. *RQ2. What research methods have been used when investigating gamified software quality?*

The classification scheme of research methods proposed by [15] was used as recommended in [9]. This classification scheme makes it possible to classify empirical research into either validation or evaluation, and non-empirical research in the categories of proposal papers, philosophical papers, opinion papers and personal experience papers. The definitions of these categories can be found in the classification that can be accessed at http://alarcos.esi.uclm.es/SMS-GamificationQuality/.

The results of RQ2 showed that most of the papers are empirical studies (32 papers, 91.43%), and of these, 28 (80%) fall into the category of validation and 4 (11.43%) into the category of evaluation (see Fig. 2). In validation papers the research methods used were quasi-experiments (18 papers, 51.43%) or experiments (10 papers, 28.57%). Most of the quasi-experimental papers used a between subject design with only a posttest (P9, P13, P23, P24, P28, P29), while the remainder used a between subject design with pre and posttests. In all those papers in which an experiment was

TABLE V. CLASSIFICATION OF THE QUANTITATIVE MEASURES OF ENGAGEMENT

Categories	Primary studies references	Measure example
Participation	P2, P5, P7, P11, P19, P23, P24, P29, P34	- number of interactions with the learning platform
Performance	P2, P4, P5, P6, P7, P8, P9, P10, P12, P13, P15, P16, P17, P18, P19, P20, P21, P23, P24, P25, P26, P27, P28, P29, P30, P31	- the percentage of branch coverage of test cases
Communication /socialization	P2, P11, P16, P21, P31, P34	- contributions to forums and other participative media
Frequency of achievements	P6, P16, P17, P18	- Badges awarded

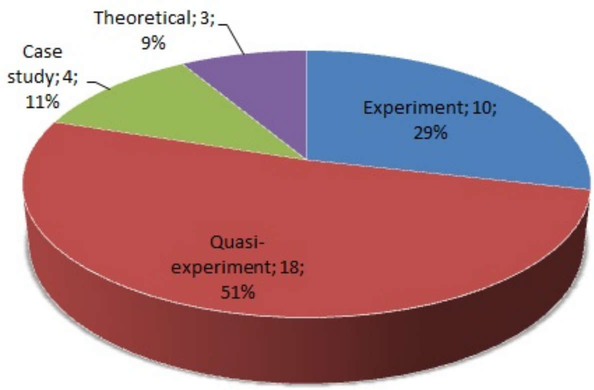


Fig. 2. Distribution of papers by research method

conducted, the researchers used a between subject design with a posttest. In this type of design, the participants in the treatment group used the gamified software and the participants in the control group used the non-gamified software, after which either they were required to fill out a posttest, or usage data were extracted and analyzed. Experimental and quasi-experimental designs were used to assess the effectiveness of the gamified approach, mainly using various quantitative measures applied to the usage data. The objective of these measures was to measure several aspects of user engagement. Although these measures vary according to the type of application, they could be classified into the categories proposed in Table V. The measures of participation aim to assess the user's involvement with the gamified application or the effect of the gamified application on users in terms of the number of actions or transactions performed, while measures of performance are intended to assess how well certain tasks or desired behavior are fulfilled. Communication/socialization measures are intended to assess the exchange or sharing of comments, opinions, or information among users inside or outside the gamified software, while the frequency of achievement measures aim to directly measure engagement through the quantification of points, badges and generally all kinds of achievements.

Satisfaction, usability and functional suitability were in most cases assessed using surveys based on Likert scales.

The evaluation papers carried out case studies (4 papers, 11.43%). A summary of the empirical studies found in this SMS can be accessed at http://alarcos.esi.uclm.es/SMS-GamificationQuality/. With regard to the non-empirical papers, all of them (3 papers, 8.57%) are philosophical papers that propose frameworks for the design and evaluation of gamified applications (P1, P3) or for the design and evaluation of gamification elements (P22).

One positive aspect found is that nearly all of the primary studies (32 papers, 91.43%) are empirical; in all of them, however, the empirical studies were conducted by the same researchers who had proposed the gamification, and none of the studies have been replicated.

C. *RQ3. What has the outcome of the research been?*

The results showed that there are only two types of research outcomes in the papers reviewed. The most common outcome is "knowledge" (30 papers, all papers excluding P1, P3, P13, P22 and P33, or 85.71%), followed to a far lesser extent by "framework" (5 papers, 14.29%). Papers whose outcome is "knowledge" do not present something whose result is "tangible" but rather use empirical studies to confirm whether gamification works, and the outcome of the research in these papers is the "knowledge" about the evidence acquired. With regard to the "framework", 3 papers are presented for the design of gamified applications (P1, P3, P13), one for the design of gamification elements (P22), and one (P33) is a conceptual framework that can be used to explore purchase intentions.

D. *RQ4. On which software artifacts from gamified software has the research on quality been focused?*

The results showed that 32 (all papers excluding P3 and P22, or 91.43%) papers dealt with gamified software quality after the product had been developed, or when a final version was ready. Only 2 (7.69%) papers dealt with the gamified software quality at the design stage. One paper presents a framework based on Bartle's [1] model of player types and on Maslow's [17] hierarchy of needs, and is intended to evaluate the engagement potential of a game design by mapping player motivations onto the various game mechanics (P3), while the other is a framework consisting of a definition of "game achievement" and a three-part model for the evaluation and design of game achievements (P22). As additional results we found that 57.14% of the implementations of the gamified approach are carried out along with the development of the application whilst in 42.86% of the cases this is done by means of a plug-in. The results show that the evaluation of quality in current gamification development practices is often put off until the later stages of the life cycle. These results are not surprising, since researchers have focused on determining the effectiveness of gamified applications and this usually occurs (as is expected in a quality-in-use evaluation) once the software product has been completed or is at least a fully functional prototype. Another possible reason why software gamification quality has focused on the final product is that nearly 43% of gamification implementations are carried out using pre-coded plug-ins. Nonetheless, we consider that it is desirable to consider quality from the early stages of the development of the gamified applications in order to obtain higher quality gamified applications.

E. *RQ5 What gamification elements have been used in gamified software?*

We found 15 different types of gamification elements in the primary studies reviewed. The gamification elements most frequently used are points and badges (23 papers each, 65.71%), closely followed by leaderboards (18, 51.43%) and levels (14, 27.45%), and to a much lesser extent rewards (6 papers, 17.14%) and unlocking (2 papers, 5.71%). There are other gamification elements, but these were mentioned in only one paper each and were classified in the 'other' category (see Fig, 3). Several papers specifically addressed the use of badges, 4 investigated their effects and 1 their design; 3 papers focused

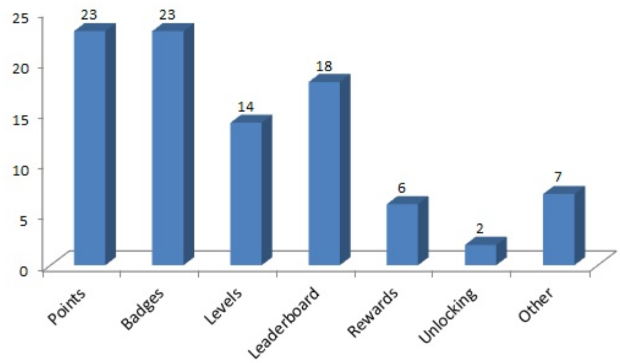


Fig. 3. Distribution of papers by use of gamification elements

on investigating the modification of user behavior (P16, P17, P18), and one paper investigated the effect of badges on user engagement (P6) while another focused on the design of badges (P22).

The results also showed that the combination of points and badges is that most frequently used (16 papers; 45.71%); another frequently used combination is that of points, badges and levels (8 papers; 22.86%). With regard to the use of the gamification elements, several papers claim to deal with gamified software. However, they have only used one or two gamification elements, which are mainly badges or points. We believe that these applications could be improved by adding other gamification elements that could encourage desirable behaviors, thereby making the application more effective [1].

Although it would be interesting to know the level of gamification of the implementations reviewed in this study, unfortunately this is outside of the scope of this paper, owing to the fact that the majority of the papers barely mention how these applications were designed and developed. The researchers have mainly focused on the use of gamified software rather than on its design or development, as is shown by the results of research question 4.

F. RQ6. What have the application areas of gamification been?

The results showed that the main application area of gamification is "Education" (13 papers, 37,14%), and most of these papers deal with gamified courses (11 papers), that is, the use of gamification elements such as points, badges and leaderboards, amongst others, to motivate and engage students in activities such as attending lectures, workshops or labs, assignment submissions, etc. The other two papers address topics such as orientation for new students (P12) and the development of virtual human patients (P19). The most common subjects of gamified courses are: Programming languages (P17, P18, P29), information and communication technologies (P5, P7), healthcare (P6, P19), software engineering (P24), information systems and computer engineering (P2), mathematics for computer scientists (P15), and computer organization and cloud computing (P23).

The other application areas are, but to a far lesser extent, "Work" with 4 papers (P10, P13, P16, P27, 11.43%), followed by "Environment" with 3 papers (P14, P25, P26, 8.57%),

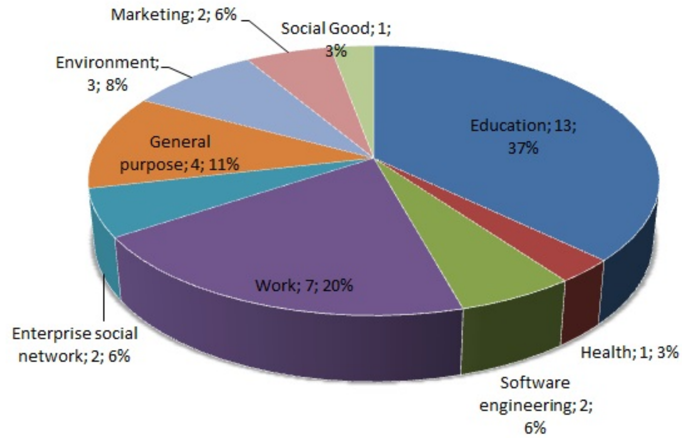


Fig. 4. Application areas

Enterprise Social Network with 2 papers (P11, P34, 5.71%), and Software engineering with 2 papers too (P9, P30, 5.71%). In this fourth category, gamification elements are used throughout the software development life cycle in order to engage and motivate developers when doing their work. The gamified course "software engineering" in the "Education" category should not be confused with the "Software engineering" application area; they are two different things. In the former, gamification elements are used in the teaching process, while in the latter the gamification elements are used during the software development cycle. There are also 4 papers (P1, P3, P22, P33, 11.43%) that do not have a particular application area, signifying that the gamification proposals in these papers are general purpose. It was also found that only 7 papers (P4, P8, P11, P16, P30, P31, P34, 20%) deal with gamification in real world settings, such as social networks, business, enterprises, banking, etc. These results show the importance of gamification and its wide acceptance in the academic world, but they simultaneously highlight the lack of research works on gamified software quality in industry. Figure 4 shows the distribution of papers by application area.

G. Aditional results.

The results reveal that "Education" is the application area in which it is reported that the gamification has had the most

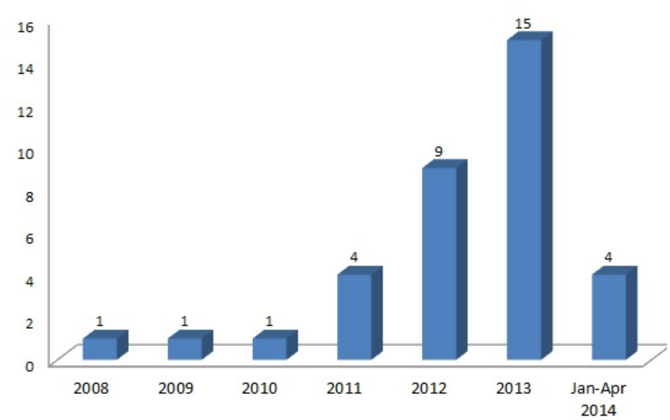


Fig. 5. Distribution of papers by publication year

positive effects, followed by the application area "work". In most of the papers the researchers report the positive effects of gamification in all application areas except the "Enterprise Social Network" category, in which the results were negative owing to the removal of a point based incentive system, and in the "Environment" category in which the results were reported as being inconclusive.

The results also reveal that since 2011, the number of publications related to gamified software has been growing steadily, and doubling every year, as shown in Fig. 5. The number of publications in 2014 was lower, because this study only considered publications up to April 2014. The results therefore show that since 2011, gamified software has undergone a very significant growth, and has in recent years become a "hot topic", thus making software gamification quality an area of opportunity for future research.

VI. Threats to Validity

The main threats to validity that should be considered for this SMS are publication selection bias, inaccuracy in data extraction and misclassification [18].

Covering everything that has been written about any topic is impossible. In spite of this, we have done everything possible to gather all the relevant information related to gamified software quality. We therefore also checked the papers included in a related literature review [6] that was published after the execution of the current SMS, resulting in the inclusion of 9 additional papers. With regard to publication selection bias, we searched in 5 of the most important digital resources, including journals, conferences and workshops that are important forums for the disclosure of gamification. We believe that the scope of influence of the journals, conferences and workshops included in this SMS is sufficient to ensure the completeness of the field of software gamification quality. Grey literature such as technical reports or PhD theses were not included since most grey literature originates from or will eventually become peer-reviewed papers. It is possible that some relevant papers have been omitted but this is, to the best of our knowledge, unlikely.

In order to tackle the issue of inaccuracy in data extraction, we defined the research questions in advance, organized the selection of articles as a multistage process, involved four researchers in this process, and documented the reasons for inclusion/exclusion as suggested in [19]. As mentioned previously, the selection of papers to be included as primary studies in this SMS was a rigorous process in which four researchers participated. The data extraction and classification of prose was difficult owing to the lack of a standard terminology for gamification design elements and with which to define quality characteristics in gamification. This may have resulted in some inaccuracies in the data extraction which may in turn have led to a misclassification. However, we believe that the extraction and selection processes were rigorous, as explained in Section IV. We are also of the opinion that the participation of two expert researchers (the last two authors) when performing the classification reduced the risk of misclassification.

VII. Conclusions and Future Work

The results of the SMS presented herein reveal that research interest in the evaluation of gamified applications has been growing steadily and rapidly since 2011. However, gamification is still an emerging technology and researchers have to date been principally interested in demonstrating its effectiveness, thereby confirming the findings of Hamari, Koivisto, and Sarsa [6]. Although engagement and motivation were assessed in several papers, on most occasions this assessment was not made directly, but by means of an evaluation of the effectiveness of the gamified application as regards achieving its objectives, such as increasing participation or another type of behavioral change.

Although satisfaction and usability have been addressed in several papers, but to a much lesser extent than effectiveness, other quality characteristics have been neglected, especially those related to product quality, such as performance efficiency, compatibility, reliability, etc. Moreover, in most of the papers quality is barely mentioned as a topic of interest for researchers. We believe that this lack of interest in software gamification quality is owing to both the fact that gamification is a new field, and also that existing software quality models are not suitable for the accurate assessment of the particular characteristics of gamified software. We believe it is time for researchers in this area to reach a consensus on which characteristics best characterize the quality of software gamification, and which are the most suitable measures as regards evaluating it. All of the aforementioned aspects highlight the need to adapt existing quality models and thus have a specific quality model with which to assess the quality of gamified software more accurately.

One positive aspect that was found as regards the research method is that nearly all the primary studies (32 papers, 91.43 %) presented experiments. This to some extent reveals that researchers concerned with software gamification initiatives recognize the need to support their proposals with empirical evidence and do not only rely on general wisdom. In most of these papers researchers have reported positive effects of gamification in all areas of application, but in Education is where the best results have been obtained. However, there is a lack of corroboration of the findings obtained by other researchers, signifying that the external replication of the empirical studies is needed in order to obtain more solid and general findings.

It was also clear that very little research has been carried out into the quality of gamified software for businesses or enterprises, and most efforts have focused on the area of education. This shows that there is a gap between industrial practice and academic research that should be addressed by researchers, bearing in mind that gamified software for enterprises represents the biggest segment of the market for new growth, with 25% of the market [4].

The findings obtained have allowed us to identify some possible opportunities for future research which are, among others:

- A quality model for gamified software: Existing software quality models need to be adapted and extended (if necessary) in order to obtain a consensus on which quality characteristics are relevant as regards evaluation and improvement in the context of gamified software, and how to measure them.

- Need for replication: It is necessary to replicate the empirical studies because in all cases the experimentation was carried out by the same researchers who had proposed the gamified software quality initiative. The material used in the experimentation could therefore be made available in order to encourage other researchers to carry out replications.
- Research in other contexts: Research on gamified software quality should also be carried out in other contexts apart from education, such as businesses and enterprises.
- Address from the early stages: Quality assurance methods that incorporate quality issues from the early stages of the development of the gamified software should be provided.
- Focus on the "good" design of gamified software: Research into what the most effective gamification design elements (or combination of them) are would also be interesting. This task is particularly complex, since the design of gamified software is a multidisciplinary activity that requires skills related not only to software development but also to psychology, among others.

Acknowledgments

This work has been funded by the following projects: GEODAS-BC (Ministerio de Economía y Competitividad and Fondo Europeo de Desarrollo Regional FEDER, TIN2012-37493-C03-01) and IMPACTUM (Consejería de Educación, Ciencia y Cultura de la Junta de Comunidades de Castilla La Mancha, y Fondo Europeo de Desarrollo Regional FEDER, PEII11-0330-4414). We would also like to thank the Instituto Tecnológico de Ciudad Victoria and PRODEP for granting the two first authors the scholarship that made it possible to complete the research work presented in this paper.

References

[1] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: defining "gamification"," Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, pp. 9-15, September 2011.

[2] G. Zichermann, and C. Cunningham, Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps. O'Reilly, Sebastopol, 2011.

[3] M. Oja, and J. Riekki, Ubiquitous Framework for Creating and Evaluating Persuasive Applications and Games," M. Rautiainen, T. Korhonen, E. Mutafungwa, E. Ovaska, A. Katasonov, A. Evesti, H. Ailisto, A. Quigley, J. Häkkilä, N. Milic-Frayling and J. Riekki (Eds.), Grid and Pervasive Computing Workshops, vol. 7096, pp. 133-140, 2012.

[4] M2 Research, "Gamification in 2012: Market Update, Consumer and Enterprise Market Trends,", 2012. Retrieved September 5th, 2014, from http://gamingbusinessreview.com/wp-content/uploads/2012/05/Gamification-in-2012-M2R3.pdf

[5] Gartner, "Gartner's 2013 Hype Cycle for Emerging Technologies,", 2013. Retrieved September 5th, 2014, from http://www.gartner.com/newsroom/id/2575515

[6] J. Hamari, J. Koivisto and H. Sarsa, "Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification," Proceedings of the 47th Hawaii International Conference on System Sciences, pp. 3025-3034, January 2014.

[7] B. Kitchenham, D. Budgen and P. Brereton, "Using mapping studies as the basis for further research – A participant-observer case study," Information and Software Technology, vol. 53 (6), pp. 638-651, 2011. doi: http://dx.doi.org/10.1016/j.infsof.2010.12.011

[8] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," Technical Report EBSE-2007-01, Software Engineering Group of Keele University Durham UK, 2007.

[9] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic Mapping Studies in Software Engineering," Conference on Evaluation and Assessment in Software Engineering , pp. 26-27, June 2008.

[10] Y. Xu, "Literature Review on Web Application Gamification and Analytics," Technical Report 11-05, Department of Information and Computer Sciences University of Hawaii, Honolulu, HI 2011.

[11] O. Pedreira, F. Garcia, N. Brisaboa, and M. Piattini, "Gamification in software engineering - A systematic mapping,", Information and Software Technology, vol.57, pp. 157-168.

[12] J. Vargas, L. García-Mundo, M. Genero, and M. Piattini, "A Systematic Mapping Study on Serious Game Quality," 18th International Conference on Evaluation and Assessment in Software Engineering (EASE´14), pp. 15, May 2014.

[13] P. Brereton, B. Kitchenham, D. Budgen, M. Turner, and M. Khalil, "Lessons from applying the systematic literature review process within the software engineering domain,", Journal of Systems and Software, vol. 80, pp. 571-583, 2007.

[14] ISO/IEC, ISO/IEC IS 25010: Systems and Software Engineering - Systems and Software Quality Requirements and Evaluation (SQuaRE) - System and Software Quality Models, ISO (International Organization for Standardization), 2011.

[15] R. Wieringa, N. Maiden, N. Mead, and C. Rolland, "Requirements engineering paper classification and evaluation criteria: a proposal and a discussion," Requirements Engineering, vol. 11, pp. 102-107, 2006. doi: 10.1007/s00766-005-0021-6

[16] R. A. Bartle, Designing Virtual Worlds. New Riders, 2004.

[17] A. H. Maslow, Motivation and Personality (Third Ed. ed.), Addison-Wesley, New York, NY, 1987.

[18] D.I. Sjøberg, J. E. Hannay, O. Hansen, V.B. Kampenes, A. Karahasanovic,N.K. Liborg, A.C. and Rekdal, "A survey of controlled experiments in software engineering," IEEE Transactions on Software Engineering , vol. 31 (9), pp. 733-753, 2005. doi: 10.1109/tse.2005.97

[19] J. Liu, J. Dehlinger, and R. Lutz, "Safety analysis of software product lines using state-based modeling," Journal of Systems and Software, vol. 80 (11), pp. 1879-1892, 2007. doi: 10.1016/j.jss.2007.01.047

Back to Contents

The Effect of Prior Gaming Experience in Motor Imagery Training for Brain-Computer Interfaces: A Pilot Study

Athanasios Vourvopoulos
Madeira-ITI
Universidade da Madeira (UMa)
Funchal, Portugal
athanasios.vourvopoulos@m-iti.org

Fotis Liarokapis
Human-Computer Interaction Laboratory
Masaryk University
Brno, Czech Republic
liarokap@fi.muni.cz

Mon-Chu Chen
Madeira-ITI
Universidade da Madeira (UMa)
Funchal, Portugal
monchu@andrew.cmu.edu

***Abstract*— Brain–Computer Interfaces (BCIs) are communication systems which translate brain activity into control commands in order to be used by computer systems. In recent years, BCIs had been used as an input method for video games and virtual environments mainly as research prototypes. However, BCI training requires long and repetitive trials resulting in user fatigue and low performance. Past research in BCI was mostly oriented around the signal processing layers neglecting the human aspect in the loop. In this paper, we are focusing at the effect that prior gaming experience has at the brain pattern modulation as an attempt to systematically identify all these elements that contribute to high BCI control. Based on current literature, we argue that experienced gamers could have better performance in BCI training due to enhanced sensorimotor learning derived from gaming. To achieve this a pilot study with 12 participants was conducted, undergoing 3 BCI training sessions, resulting in 36 EEG datasets. Results show that a strong gaming profile not only could possibly enhance the performance in BCI training through Motor-Imagery but it can also increase EEG rhythm activity.**

Keywords—Brain-Computer Interfaces; Serious Games; Virtual Reality; Motor Imagery

I. INTRODUCTION

Brain–Computer Interfaces (BCIs) are communication systems which translate brain activity into control commands [1]. In a BCI system, brain modulation patterns can be extracted and analyzed in order to determine the mental state of the user. These states can be translated with the help of signal processing algorithms and machine learning into a control signal that could act as an input for computers or external devices (e.g. robots). BCI technology is a rapidly growing field of research and has been shown to be very promising for controlling agents within virtual environments [2]–[4].

Currently, three main techniques are used in BCI systems for user interaction and control including: (a) Steady State Visual Evoked Potentials (SSVEP), (b) P300 BCI and (c) Motor-Imagery (MI) or ERS/ERD BCI. SSVEP is caused by visual stimulation by flashing lights to the user and occur at the primary visual cortex of the brain [5]. On the other hand, P300 BCI is generated by measuring the brain evoked response after stimulus onset, positive and negative deflections in the EEG signal after 300ms (hence the name) [6]. Finally, ERS/ERD stands for event related synchronization/desynchronization of the mu (μ) rhythm. Mu is located at the motor and somatosensory cortex of the brain where patterns of electrical activity control voluntary movement [7]. Motor imagery (MI) BCI training is based on visuo-motor imagination and has been widely used as a BCI paradigm in research [8]. MI is relying on the same brain systems that would be used for actual performance of the task by activating the same brain areas as actual action execution or action observation. Results from previous studies have proven mental practice of action to be useful in MI-BCI with beneficial effects in motor control of patients with paralysis [9].

BCI games or neurogames have recently become increasingly more advanced by incorporating immersive virtual environments [10], multiple user objectives, and hybrid control systems integrating both conventional input devices and multiple BCI techniques [11]. Several BCI surveys have analyzed and reviewed BCI games in terms of different approaches, including human–computer interaction (HCI) [12], and immersive virtual reality (VR) [13]. Unfortunately, the fundamental issue of BCI illiteracy/skill where, regardless of the duration of the training session, users are unable to have a stable control, is present in MI-BCI.

The aim of this paper is to examine the effect that prior gaming experience has at the brain pattern modulation during MI training in order to identify the elements that contribute to high BCI control. The hypothesis is based on that experienced gamers could have better performance in MI-BCI training due to enhanced sensorimotor learning derived from gaming. A pilot study with 12 participants, undergoing 3 MI-BCI training sessions was performed. Initial results indicate that a strong gaming profile could possibly enhance the performance in classification accuracy during a MI-BCI training, and additionally increase EEG rhythm modulation in the Alpha, Beta, Theta and Gamma bands. In addition, a relationship with the demographic data provides useful pointers for the trainee profile and its effect to the training outcome.

The rest of the paper is structured as follows. Section II presents relevant background information with several case studies. Section III describes the research hypothesis whereas

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

section IV the methodology applied. Section V illustrates the results of the experiment. Finally, section VI presents the discussion of the results and section VII the conclusions and future work.

II. Background

The state-of-the-art of BCIs in the context of games and entertainment as well as interactive systems has been recently reviewed [14], [15]. BCIs in games and virtual environments (VEs) can be categorized into: passive, reactive and active [16]. Passive BCIs are not controlling the games, but just monitors the user's mental states automatically. Reactive BCIs embed information with respect to the user's intention in the response signal to stimulation. Active BCIs are typically used direct control of games and VEs and this is where the focus of this research is. Motor imagery is the paradigm used most often in BCI research [17]. In one of the earliest examples, users controlled the movement of First-Person Shooter game either to the left by producing low mu or to the right by producing high mu [3]. Results indicated that subjects learn to control levels of mu very quickly, but especially when this learning involves producing similar mu levels (whether high or low) over each hemisphere.

The Berlin Brain–Computer Interface (BBCI) which presented setups where the user provided with intuitive control strategies in plausible gaming applications that use biofeedback [18]. The BBCI used motor imagery to play Pacman and Pong and similarly familiar games such as Tetris. Moreover, a one-channel EEG-based controlled 'Run and Jump' game to allow subjects to get pulse width modulated control was implemented [19]. Training procedure that allows subjects to produce one brain pattern, elicited with motor imagery, of two different durations. Results with 5 users showed that they were able to jump with a high accuracy (about 94.5%) over the hills having low number of false positives (39%) between the hills. Another example is pinball BCI game that controlled the left and right paddles through brainwaves [20]. The system was designed to examine two classes of motor imagery (left and right hand movement). More lately, the use of cheap commercial BCIs was evaluated when users played the same serious game. Results indicated that both BCI technologies offer the potential of being used as alternative game interfaces prior to some familiarization with the device and in some cases some sort of calibration [21].

Furthermore, there have been several cases where online game or VEs can be controlled with BCIs. For example Google Earth was successfully controlled using BCIs based on a platform called Brainloop interface [22]. In the same study, a computer-game-like application was also proposed where subjects have to navigate through the environment and collect coins by autonomously selecting navigation commands. Three subjects participated in these feedback experiments and each learned to navigate through the VE and collect coins. Two out of the three succeeded in collecting all three coins. A novel interface allowed users with severe disabilities to freely explore Second Life's virtual world and control their avatar within it [23]. To activate a command, the user focuses their attention on the corresponding icon on a screen, such as 'Lights On', while the EEG cap records their P300. The icons are flashed randomly, one at a time, and it is possible to tell which icon they are looking at by correlating a spike in the P300 with the timing of when that icon flashes.

In terms of immersive VEs, an interesting study presented two experiments exploring the different constraints posed by current BCI systems when used in VR [24]. In the first experiment, participants made free choices during the experience and compared their BCI performance with participants using BCI without free choice. In the second experiment the system allowed users to control a virtual body with motor imagery, with accuracy ranging from 72 to 96%. A more recent study, examined a multimodal BCI experiment situated in a highly immersive CAVE [11]. Users controlled the main character of a virtual reality game (triggered a jump action via the BCI) but also additional steering with a game controller as a secondary task was tested. Results showed that a transfer of skills is possible in spite of the changes in visual complexity and task demand and that the use of a secondary motor task did not deteriorate the BCI performance during the game.

Unfortunately, the fundamental issue of BCI illiteracy is always present. Regardless of the duration of the training session, users are unable to master the control capabilities of a BCI system. Unfortunately, this limitation is being more common with motor imagery [25]. Motor imagery can also be affected by insufficient attention due to user distraction or frustration [26]. As frustration is present within many video games (triggered either from false mechanics within the game design or player's limited skill), the need for maintaining the entertainment and gameplay challenges within a brain-controlled game is still a challenge. A recent approach is by making use of tutorial levels with training elements adopted in motor imagery BCI games in order to “disguise” or encapsulate user training inside the game so as to enhance the user's experience [27].

Despite the increased attention that BCI‘s had with the launch of low-cost EEG devices for gaming in the last few years, BCI technology is hardly used outside laboratory environments [28]. This is mainly due to the fact that current BCI systems lack reliability and good performance in comparison with other types of gaming interfaces [29]. As a result, long training sessions can contribute in user fatigue and reduced performance. In addition, prolonged training is problematic in generating EEG oscillatory rhythms modulated during MI, such as mu (μ) and beta (β) rhythms [30]. Therefore, it is essential to identify new strategies for a successful MI-BCI training and control.

III. Research Hypothesis

To date, it has been shown that people regularly exposed to video-games have improved over time their visual and spatial attention, memory, mental rotation abilities [31][32] and enhanced sensorimotor learning, enabling better performance in tasks with consistent and predictable structure [33]. Extensive video-game practice has also been shown to improve the efficiency of movement control brain networks and visuo-motor skills of the users [34]. Since these type of skills are used in current mental tasks used to control a Brain-Computer Interface (BCI) [e.g., mental rotation of geometric figures, motor imagery, remembering familiar faces [35], this suggests BCI users might improve their mastery of BCI by performing training tasks that do not involve the BCI system, such as by playing various videogames and improving in an indirect way their visuo-motor capabilities. So far, relationship between videogame practice, player profile and BCI performance has not been yet shown for BCI based on mental tasks, and there is currently no available literature to support this hypothesis [28]. So far, it has been observed for BCI based on Steady-State Visual Evoked Potentials (SSVEP) [36] but not in motor-imagery (visuo-motor imagination). Motor imagery BCI's requires long training trials per session and it is user-specific. Although mental practice of action is used widely in BCIs, long and repetitive training sessions can result in user fatigue and declining performance over time. This suggests that having BCI users practicing (non-BCI-based) videogames might be a promising indirect training method to improve their BCI control skills and minimize the overall training time.

In this study we attempt to identify if the enhanced sensorimotor capability of experienced gamers is reflected in MI-BCI training. For this, we designed an experimental setup including a standard BCI training paradigm and two different user groups based on their previous gaming experience.

IV. Methodology

A. Experimental Setup

The experimental setup was composed by a desktop computer (OS: Windows 8.1, CPU: Intel® Core™ i5-4440 at 3.3 GHz, RAM: 8GB DDR3 1600MHZ, Graphics: Nvidia GT 630 1GB GDDR3), running the BCI training task. In addition the Vuzix iWear VR920 (Vuzix, NY, USA) head mounted display (HMD) was used by the participants (see Figure 1(a)). The HMD includes 640x480 twin LCD displays, 32-degree field of view (FOV), 3/4" eye relief and 5/16" eye box.
The BCI set up comprised, by 8 active electrodes equipped with a low-noise bio-signals amplifier and a 16-bit A/D converter (256 Hz). The spatial distribution of the electrodes followed the 10-20 system configuration [37] with the following electrodes over the sensory-motor areas: Frontal-Central (FC3, FC4), Central (C3, C4, C5, C6), and Central-Parietal (CP3, CP4) as it is illustrated in Figure 2. The g.MOBIlab biosignal amplifier (gtec, Graz, Austria) was connected via bluetooth to the desktop computer for the EEG signal acquisition and processing through OpenVibe platform [38]. For all user data, a Common Spatial Patterns (CSP) filter was used for feature extraction, and a Linear Discriminant Analysis (LDA) for the classification of MI from EEG data.

B. Participants

The study consisted of a total of 12 participants (mean age of 28 ± 2 years old, 8 male, 4 female). Participants were a voluntary sample, recruited based on their motivation to participate to the study. In addition, no previous known neurological disorder, nor previous experience in BCIs was reported. All participants were University students and academic staff. Furthermore, all participants signed an informed consent form before participating in the user-study. The assessment took place in a controlled environment at a neurorehabilitation lab.

C. Motor-Imagery BCI Training

The visual stimulation was based on the Graz-BCI paradigm [39] with a standard bars-and-arrows feedback (see Figure 1(b)). When an arrow appears on the screen, the user has to perform a mental imagery of the corresponding hand (based on arrow direction) performing a task (grasping, throwing, waving, etc.). The visualization needs to be consistent during the whole duration of the training session in order to train the linear classifier to distinguish left from right hand imagery. Each participant went through 3 complete training sessions (1 per day) within one week. On each session

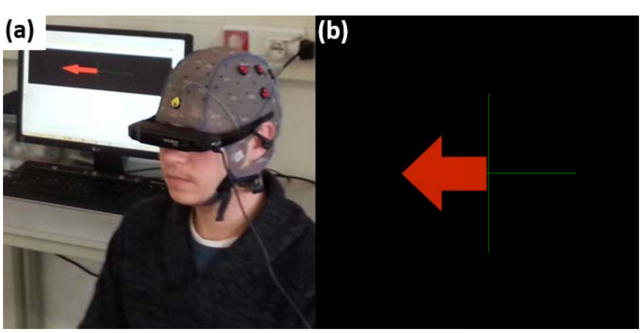


Figure 1. (a) Experimenters station responsible for EEG signal acquisition, processing and feedback control. (b) Participants view through the HMD during Motor-Imagery training

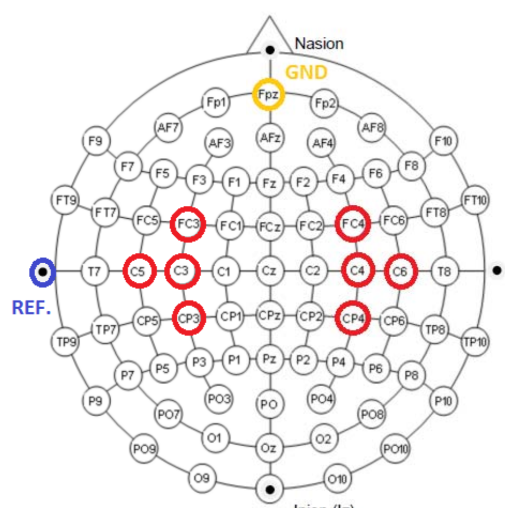


Figure 2. Electrode location on the 10-20 system. Left side is covered by the FC3, C5, C3 and CP3 electrodes and the right side by FC4, C4, C6 and CP4. The ground electrode is placed in Fpz location and the reference at the left ear love.

the participant had to perform a 30 seconds baseline measurement, followed by cue based motor-imagery training, 20 times per class (left or right). Each feedback duration (unidirectional arrow) lasted for 4 seconds followed by 1.5 second pause. After the completion of the training session, a 5 minute rest was followed by an online session with a trained classifier for acquiring the accuracy score. Finally, for all 3 sessions, from 12 participants, 36 EEG datasets were gathered for further analysis.

D. Questionnaires

Prior the BCI training session, participants were tested for their handedness using the 'Edinburg handedness inventory' [40]. The handedness inventory classify people in left handed (-100% to -40%), ambidextrous (-40% to 40%) and right handed (40% to 100%). The higher score corresponds to higher level of handedness either left or right. Following the handedness classification, the Vividness of Movement Imagery Questionnaire-2 (VMIQ2) [41] was used in order to assess the feeling of the participant to perform an imagined movement (Kinaesthetic Imagery). The VMIQ [42] has been previously used to determine differences in neural activation patterns between vivid and non-vivid imagery. The Kinaesthetic Imagery (KI) questions involve both upper and lower limb movements ranging from 1 ('no kinesthetic sensation'/'no image') to 5 ('as clear as executing an action'/'image as clear as seeing').

In addition, for creating a profile based on previous gaming experience of the participants, we first used a modified version of the Game Addiction (GA) questionnaire [43] comprised of 30 questions with different weights based on the answer. The Likelihood of Video Game Addiction index was giver after calculating the final score and it ranged between 0-80 (for low: 0 – 20 points, moderate: 21 – 40 points, high: 41 – 60 points, and very high: 61-80 points).

Finally, as currently there are no validated questionnaires that measure the impact of the overall gaming experience of users, we used a 15 factor classification questionnaire for calculating the Gamer Dedication (GD) [44]. The measurements were in a liker scale between 1 to 5, in which participants were asked whether they "strongly disagree," "disagree," "neither disagree nor agree," "agree," or "strongly agree" with a series of statements. The classification score was between 0-100 and included the following scales: >30 for non-gamers, 30-45 for casual, 46-55 for transitional/moderate, 56-70 for hardcore, and >70 for ultra-hardcore.

E. Grouping Gamers

To group users based on their previous gaming experience, the GD questionnaire was used as a statistical method for classifying the participants as explained in the previous section. Through this method the GD score was calculated based on the following formula:

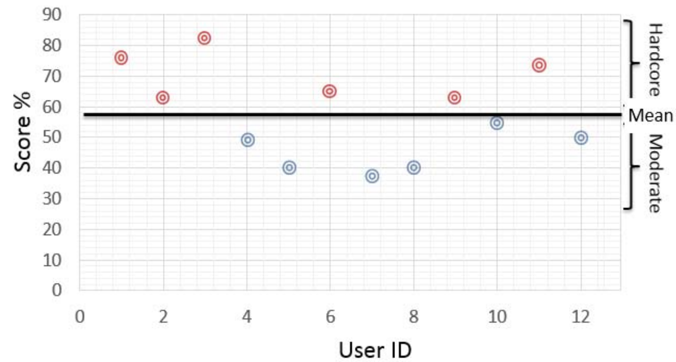


Figure 3. Gamer Dedication clusters. Upper level, above the mean line (M=57.65) is the hardcore group score (range: 62.6-82.2%) separated by the moderate group (range: 37.2-54.4%)

$$GD = \frac{\sum_{j=1}^{n} w_j \, s_j}{\sum_{j=1}^{n} 5w_j} \tag{1}$$

Where s = self-ranked score; and w = weight.

Following the recruitment, two groups had been formed based on their score: (1) Hardcore group (with score between: 62.6-82.2%) and (2) Moderate group (with score between: 37.2-54.4%). The hardcore group was created by merging the "Hardcore" (56-70) and "Ultra-Hardcore" (>70) groups from the GD scale, and the moderate group by merging the "Casual" (30-45) with the "Moderate"(46-55) groups. This allowed us to have a distinct separation of two equally numbered groups due to the low user numbers (see Figure 3).

F. Extracting EEG Rhythms

The raw EEG signals were processed by extracting the Power Spectral Density (PSD), averaging all electrodes in Matlab (MathWorks Inc., Massachusetts, US). The power was extracted every 500 ms using Welch's method [45] with windows of 128 samples for the following frequency bands: Alpha (8 Hz -12 Hz), Beta (12 Hz- 30 Hz), Theta (4 Hz- 7 Hz), and Gamma (25 Hz- 100 Hz). In addition, to remove signal artifacts related with eye movement, EMG, ECG and power line noise, a component rejection process was carried out using Independent Component Analysis (ICA) with the help of the EEGLAB toolbox [46]. Finally, the Mean Power (dB) had been extracted for all bands and grouped among all three sessions, for both gamer groups.

V. RESULTS

In this pilot, we tried to gain insights on how previous gaming experience and user profile can reflect to BCI performance during training with the use of MI. During all sessions a set of data has been acquired that can provide information about the participant in three parts: through subjective answers (questionnaires), training outcome (LDA

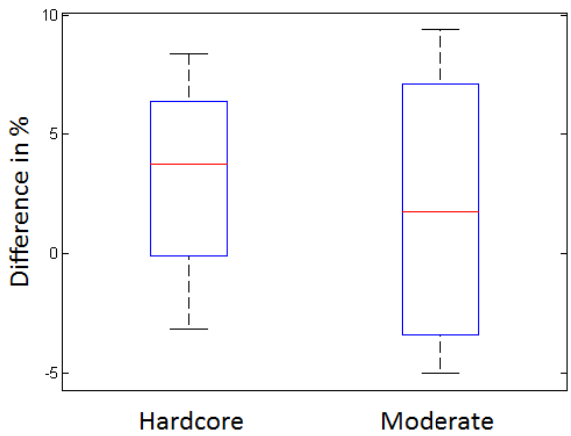


Figure 4. Difference in performance between first and last session of MI-BCI training between hardcore and moderate participants.

classification result), and electrophysiological data (EEG rhythms).

A. Can experienced gamers increase their performance faster?

From the LDA classification output between the first and the last training session, on the "Hardcore" group, we observe an improvement of performance of 3.16% (*STD = 3.9*). On the "Moderate" group, the increase is almost half at 1.53% (*STD = 5.15*) with higher standard deviation. On average, we observe that "Hardcore" users improve their performance 1.63% faster compared to the "Moderate" (see Figure 4).

B. Can different gamer groups modulate different EEG patterns?

From the EEG bands output between the first and the last training session, on the "Hardcore" group, Alpha band increased 7.76% compared to the first session, Beta decreased -1.02%, Theta increased 3.68%, and finally Gamma decreased -4.3%. On the "Moderate" group, Alpha decreased -4.68%, Beta decreased -3.66%, Theta decreased -3.41%, and Gamma decreased -1.13%. Overall, on the "Hardcore" group Alpha

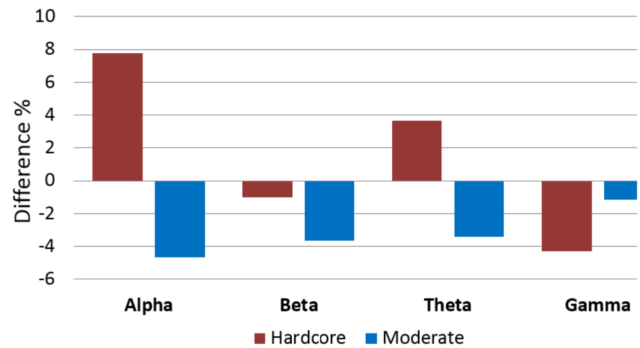


Figure 5. Increased or decreased EEG rhythm activity between first and last training session for both groups.

and Theta bands increased with Beta and Gamma to be decreased. On the "Moderate" group, all bands decreased. In relation between the two groups, we observe that on Alpha band, "Hardcore" group is 12.4% higher than "Moderate", 2.64% for Beta, 7.09% for Theta, but reduced -3.17% on Gamma. On average, we can observe that users belonging to the "Hardcore" group had increased EEG bands modulation of 4.75% compared to the "Moderate" group (see Figure 5).

C. Relationship between demographics and EEG pattern modulation.

From all the EEG data in all participants, gender-specific correlations have been found with EEG bands. In particular, significant correlations are found in Beta (r=0.7848, p<0.05), Theta (r=0.6111, p<0.05), and Gamma rhythms (r=0.6302, p<0.05) with women to have higher modulation in these bands than men (see Figure 6). This difference is also visible through t-test statistics, with a significant effect for beta t(10) = -4.004, p < 0.05, theta t(10) = -2.441, p < 0.05, and Gamma t(10) = -2.566, p < 0.05. These results suggest that gender could have an effect on the EEG bands modulation on specific frequencies. Finally, a weak correlation had been found between age groups (18-24, 25-34, 35-44, 45-54, 55-64, 65-74, and 75 years or older) and Alpha band (-0.5642, p = 0.05). This suggests that older people could have higher Alpha modulation during training (see Figure 7).

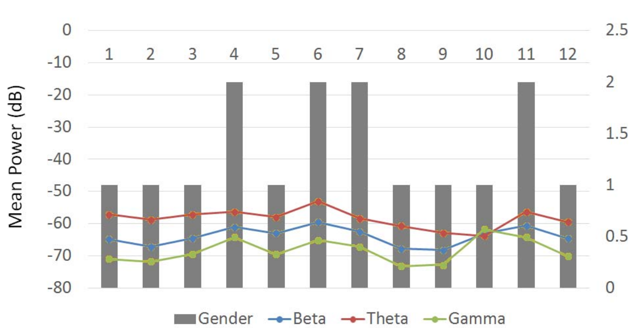


Figure 6. Relationship between Gender (1 for male, 2 for female) and EEG bands. Significant correlations between Beta, Theta, Gamma bands.

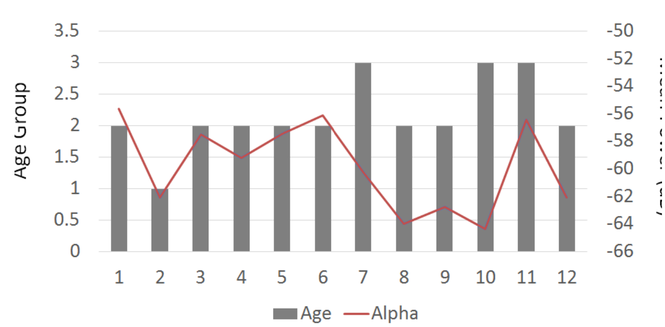


Figure 7. Relationship between Age Groups and Alpha Band Modulation. Elder groups show a weak correlation between the modulation of the Alpha.

Back to Contents

D. Relationship between subjective reports and brain activity.

To assess the relationship between the user profile - the subjective experience as reported through the questionnaires including demographics - and the elicited brain activity patterns, we decided to use a stepwise multilinear regression modeling approach. With this regression model we can determine statistical relationships –through the predictive variables- between the generated EEG rhythms during the MI-BCI conditions (Alpha, Beta, Theta, and Gamma bands), the answers of the questionnaires (GD, GA, handedness and Kinaesthetic Imagery) and the demographic data. Table I summarizes the findings of the model, including only the significant relationships.

From GD answers, we can see that participants that play games over many long sessions (Coeff. = 0.1871, $p<0.05$) have increased Theta activity and also participants who discuss games with friends and/or bulletin boards show an increase on the Alpha band (Coeff. = 0.3683, $p<0.05$). From GA, participants that have unsuccessfully tried to reduce the amount of time that play video games (Coeff. = 0.2746, $p<0.05$), eat meals while playing video games (Coeff. = 0.1779, $p<0.05$) and get headaches, red eyes, sore fingers, or wrist pains from playing video games have increased Beta (Coeff. = 0.0912, $p<0.05$). In addition, participants that stay up late to play video games and are tired the next morning, have an increase in Alpha band (Coeff. = 0.0715, $p<0.05$). Moreover, we can observe that there is a relationship between the ability of kinaesthetic imagery on kicking a stone with a higher activation of the Gamma band (Coeff. = 0.1547, $p<0.05$) and game genre preference with Alpha (Coeff. = -0.3987, $p<0.05$). Finally, high level of right-handedness (Coeff. = 0.1484, $p<0.05$) and gender (Coeff. = 10.9204, $p<0.05$) contribute with higher activation on Beta band.

Table I. Multi-linear stepwise regression model. The table shows the coefficients of the significant contribution in the regression model.

EEG Rhythms / Subjective Self-Reports	Alpha	Beta	Theta	Gamma
Gender	-	**0.1484**	-	-
Game Genre	**-0.3987**	-	-	-
GD: I play games over many long sessions	-	-	**0.187**	-
GD: I discuss games with friends/bulletin boards	**0.3683**	-	-	-
GA: I have unsuccessfully tried to reduce the amount of time I play video games	-	**0.2746**	-	-
GA: I stay up late to play video games and as a result, I am tired the next morning	**0.0715**	-	-	-
GA: I am an active member in activities or clubs	-	-	-	**-0.2118**
GA: I eat meals while playing video games	-	**0.1779**	-	-
GA: I get headaches, red eyes, sore fingers, or wrist pains from playing video games	-	**0.0912**	-	-
KI: Kicking a stone	-	-	-	**0.1547**
Handedness	-	**10.920**	-	-

VI. Discussion

Although the data sample cannot provide concrete statistical significance, current results provide interesting pointers related to gamer profile and brain activity during a MI-BCI training. Initially, on the Alpha band, we observe that participants which discuss games and stay up late to play video games have an increased modulation. Gaming experience leading to increased Alpha activity can be the reason that older participants had higher Alpha. In addition, gamers with favorite genre the Action games, had increased activity. Past research showed positive correlations between good game performance (less time to complete a task) and, on average, the production of Alpha waves [47]. On the Beta band, high addiction users produce higher the activity. In particular, gamers that unsuccessfully tried to reduce the amount of time they play video games, eat meals while playing, and get headaches, red eyes, sore fingers, or wrist pains, have increased Beta activity. In addition, participants with higher level of right-handedness and females had increased modulation. So far we know that, voluntary movement results in a desynchronization in the upper alpha and lower beta band oscillations, localized over sensorimotor areas, for example, imagery of right and left hand movements results in desynchronization of mu and beta rhythms over the contralateral hand area [48], therefore high levels of addiction could result into higher band oscillations over the sensorimotor area. Furthermore, current results show that participants which play games over many long sessions have increased Theta modulation during training. Cortical theta is observed frequently during meditative or relaxing states, but also it has been shown that the level of theta brainwave activity in the prefrontal cortex predicts whether people will be able to overcome ingrained biases - choosing an action that is counter to habit - when is required to achieve a goal [49]. This trait could be reflected during BCI training in order to produce better results in a repetitive bar-and-arrow feedback.

Moreover, participants with high kinaesthetic imagery related with “kicking a stone” visualization, have increased Gamma activity but participants of the “Hardcore” group had decreased activity compared with the “Moderate” group. So far, invasive studies with electrocorticography (ECoG) have shown that motor imagery elicits high Gamma activity in the motor cortex (cortical activity during motor execution, motor imagery, and imagery-based online feedback) although the high frequency Gamma band has a low signal-to-noise ratio (SNR) in EEG recordings due to the attenuation of electrical potentials when they diffuse through the intervening tissues (skull, and scalp) to the surface recording electrode [50]. Overall through the LDA classification results, we can observe a tendency of “Hardcore” gamers to increase their performance faster during training.

So far, with the current results, we can distinguish a trend between the two gamer groups as classified by the GD formula. These results suggest that a strong gaming profile could possibly enhance the ability to use a BCI system in terms of increased visuo-motor imagination and faster learning outcome overtime within a neurofeedback loop. Results revealed differences between all EEG bands, but also classification percentages with increased performance. This user profiling becomes very relevant for the design of future neurogames where electrophysiological data is used as the main input.

VII. Conclusions and Future Work

This paper has presented a pilot study aiming in identifying the reasons that contribute to high BCI control for experienced gamers through motor-imagery training. A user study with 12 participants indicated that players with a strong gaming background have the potential of using more effectively an MI-BCI through increased EEG rhythm modulation and increased score progression. Although due to sample size the statistical error margin is wide, we gain confidence that current results provide a first step into user-centered neurogame design for platforms using eletrophysiological data as an input. This unexplored possibility for BCI training is not aiming only for games in the entertainment domain but also explore the potential of utilizing these techniques in the health domain for people enrolled in rehabilitation programs with the use of virtual tools and serious games.

In terms of future work, the user-base needs to be extended by including participants with low or no gaming experience. This will help us to generalize the findings in a greater extend and provide also a solid statistically significant outcome. In addition, a 3D immersive environment will be used, including an HMD with higher resolution or a stereoscopic projection system. In addition, the use of an EEG system with additional electrodes is necessary. This will allow us not only to gain higher spatial resolution over the somatosensory area but also to improve the quality of the feature selection before training the classifier. These steps are consider necessary in order to gain more insights on how different areas are activated on "hardcore" gamers and how this information can be used to optimize the BCI-Game experience.

Acknowledgments

This work is supported by the European Commission through the RehabNet project - Neuroscience Based Interactive Systems for Motor Rehabilitation - EC (303891 RehabNet FP7-PEOPLE-2011-CIG), and by the Fundação para a Ciência e Tecnologia (Portuguese Foundation for Science and Technology) through SFRH/BD/97117/2013 and Projeto Estratégico - LA 9 - 2014-2015. Authors would also like to thank HCI Lab members at Masaryk University for their support and inspiration.

References

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, Jun. 2002.

[2] A. Vourvopoulos and F. Liarokapis, "Evaluation of commercial brain–computer interfaces in real and virtual world environment: A pilot study," *Comput. Electr. Eng.*

[3] J. A. Pineda, D. S. Silverman, A. Vankov, and J. Hestenes, "Learning to control brain rhythms: making a brain-computer interface possible," *Neural Syst. Rehabil. Eng. IEEE Trans. On*, vol. 11, no. 2, pp. 181–184, 2003.

[4] J. D. Bayliss, "Use of the evoked potential P3 component for control in a virtual apartment," *Neural Syst. Rehabil. Eng. IEEE Trans. On*, vol. 11, no. 2, pp. 113–116, 2003.

[5] Y. Wang, X. Gao, B. Hong, C. Jia, and S. Gao, "Brain-Computer Interfaces Based on Visual Evoked Potentials," *IEEE Eng. Med. Biol. Mag.*, vol. 27, no. 5, pp. 64–71, Sep. 2008.

[6] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger, "How many people are able to control a P300-based brain–computer interface (BCI)?," *Neurosci. Lett.*, vol. 462, no. 1, pp. 94–98, Sep. 2009.

[7] C. Neuper, M. Wörtz, and G. Pfurtscheller, "ERD/ERS patterns reflecting sensorimotor activation and deactivation," *Prog. Brain Res.*, vol. 159, pp. 211–222, 2006.

[8] J. Li and L. Zhang, "Active training paradigm for motor imagery BCI," *Exp. Brain Res.*, vol. 219, no. 2, pp. 245–254, Jun. 2012.

[9] N. Birbaumer and L. G. Cohen, "Brain–computer interfaces: communication and restoration of movement in paralysis," *J. Physiol.*, vol. 579, no. 3, pp. 621–636, Mar. 2007.

[10] B. van de Laar, H. Gurkok, D. Plass-Oude Bos, M. Poel, and A. Nijholt, "Experiencing BCI Control in a Popular Computer Game," *IEEE Trans. Comput. Intell. AI Games*, vol. 5, no. 2, pp. 176–184, Jun. 2013.

[11] R. Leeb, M. Lancelle, V. Kaiser, D. W. Fellner, and G. Pfurtscheller, "Thinking Penguin: Multimodal Brain #x2013;Computer Interface Control of a VR Game," *IEEE Trans. Comput. Intell. AI Games*, vol. 5, no. 2, pp. 117–128, Jun. 2013.

[12] D. P.-O. Bos, B. Reuderink, B. van de Laar, H. Gürkök, C. Mühl, M. Poel, A. Nijholt, and D. Heylen, "Brain-Computer Interfacing and Games," in *Brain-Computer Interfaces*, D. S. Tan and A. Nijholt, Eds. Springer London, 2010, pp. 149–178.

[13] A. Lecuyer, F. Lotte, R. B. Reilly, R. Leeb, M. Hirose, and M. Slater, "Brain-Computer Interfaces, Virtual Reality, and Videogames," *Computer*, vol. 41, no. 10, pp. 66–72, 2008.

[14] A. Nijholt, B. Reuderink, and D. O. Bos, "Turning Shortcomings into Challenges: Brain-Computer Interfaces for Games," in *Intelligent Technologies for Interactive Entertainment*, A. Nijholt, D. Reidsma, and H. Hondorp, Eds. Springer Berlin Heidelberg, 2009, pp. 153–168.

[15] A. L. S. Ferreira, L. C. de Miranda, E. E. C. de Miranda, and S. G. Sakamoto, "A Survey of Interactive Systems based on Brain-Computer Interfaces," *SBC J. Interact. Syst.*, vol. 4, no. 1, pp. 3–13, Aug. 2013.

[16] M. Ahn, M. Lee, J. Choi, and S. C. Jun, "A review of brain-computer interface games and an opinion survey from researchers, developers and users," *Sensors*, vol. 14, no. 8, pp. 14601–14633, 2014.

[17] H.-J. Hwang, S. Kim, S. Choi, and C.-H. Im, "EEG-Based Brain-Computer Interfaces: A Thorough Literature Survey," *Int. J. Hum.-Comput. Interact.*, vol. 29, no. 12, pp. 814–826, Dec. 2013.

[18] R. Krepki, B. Blankertz, G. Curio, and K.-R. Müller, "The Berlin Brain-Computer Interface (BBCI) – towards a new communication channel for online control in gaming applications," *Multimed. Tools Appl.*, vol. 33, no. 1, pp. 73–90, Apr. 2007.

[19] G. Müller-Putz, R. Scherer, and G. Pfurtscheller, "Game-like training to learn single switch operated neuroprosthetic control," in *BRAINPLAY 07 Brain-Computer Interfaces and Games Workshop at ACE (Advances in Computer Entertainment) 2007*, 2007, p. 41.

[20] M. Tangermann, M. Krauledat, K. Grzeska, M. Sagebaum, B. Blankertz, C. Vidaurre, and K.-R. Müller, "Playing pinball with non-invasive BCI," in *Advances in Neural Information Processing Systems*, 2009, vol. 21, pp. 1641–1648.

[21] F. Liarokapis, A. Vourvopoulos, A. Ene, and P. Petridis, "Assessing Brain-Computer Interfaces for Controlling Serious Games," presented at the 5th International Conference on Games and Virtual Worlds for Serious Applications (VS-Games'13), Bournemouth, UK, 2013.

[22] R. Scherer, A. Schloegl, F. Lee, H. Bischof, J. Jansa, and G. Pfurtscheller, "The self-paced graz brain-computer interface: methods and applications," *Comput. Intell. Neurosci.*, p. 79826, 2007.

[23] "Control your home with thought alone - tech - 05 July 2011 - New Scientist." [Online]. Available: http://www.newscientist.com/article/mg21128195.200-control-your-home-with-thought-alone.html#.VRWHY_msV8F. [Accessed: 27-Mar-2015].
[24] D. Friedman, R. Leeb, G. Pfurtscheller, and M. Slater, "Human–Computer Interface Issues in Controlling Virtual Reality With Brain–Computer Interface," *Human–Computer Interact.*, vol. 25, no. 1, pp. 67–94, Feb. 2010.
[25] A. Vuckovic, "Motor imagery questionnaire as a method to detect BCI illiteracy," in *2010 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL)*, 2010, pp. 1–5.
[26] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, and G. Pfurtscheller, "How many people are able to operate an EEG-based brain-computer interface (BCI)?," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 145–147, Jun. 2003.
[27] D. Marshall, S. Wilson, D. Coyle, and M. Callaghan, "Games, Gameplay, and BCI: The State of the Art," *Comput. Intell. AI Games IEEE Trans. On*, vol. PP, no. 99, pp. 1–1, 2013.
[28] F. Lotte, F. Larrue, and C. Mühl, "Flaws in current human training protocols for spontaneous Brain-Computer Interfaces: lessons learned from instructional design," *Front. Hum. Neurosci.*, vol. 7, Sep. 2013.
[29] F. Lotte, "On the need for alternative feedback training approaches for BCI," presented at the Berlin Brain-Computer Interface Workshop, 2012.
[30] E. Niedermeyer and F. H. L. da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins, 2005.
[31] C. S. Green and D. Bavelier, "Action video game modifies visual selective attention," *Nature*, vol. 423, no. 6939, pp. 534–537, May 2003.
[32] J. Feng, I. Spence, and J. Pratt, "Playing an action video game reduces gender differences in spatial cognition," *Psychol. Sci.*, vol. 18, no. 10, pp. 850–855, Oct. 2007.
[33] D. G. Gozli, D. Bavelier, and J. Pratt, "The effect of action video game playing on sensorimotor learning: Evidence from a movement tracking task," *Hum. Mov. Sci.*, vol. 38C, pp. 152–162, Oct. 2014.
[34] J. A. Granek, D. J. Gorbet, and L. E. Sergio, "Extensive video-game experience alters cortical networks for complex visuomotor transformations," *Cortex J. Devoted Study Nerv. Syst. Behav.*, vol. 46, no. 9, pp. 1165–1177, Oct. 2010.
[35] E. V. C. Friedrich, R. Scherer, and C. Neuper, "Long-term evaluation of a 4-class imagery-based brain-computer interface," *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.*, vol. 124, no. 5, pp. 916–927, May 2013.
[36] B. Z. Allison, D. J. McFarland, G. Schalk, S. D. Zheng, M. M. Jackson, and J. R. Wolpaw, "Towards an independent brain-computer interface using steady state visual evoked potentials," *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.*, vol. 119, no. 2, pp. 399–408, Feb. 2008.
[37] "Report of the committee on methods of clinical examination in electroencephalography: 1957," *Electroencephalogr. Clin. Neurophysiol.*, vol. 10, no. 2, pp. 370–375, May 1958.
[38] Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, O. Bertrand, and A. Lécuyer, "OpenViBE: an open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments," *Presence Teleoperators Virtual Environ.*, vol. 19, no. 1, pp. 35–53, 2010.
[39] G. Pfurtscheller, C. Neuper, G. R. Müller, B. Obermaier, G. Krausz, A. Schlögl, R. Scherer, B. Graimann, C. Keinrath, D. Skliris, M. Wörtz, G. Supp, and C. Schrank, "Graz-BCI: state of the art and clinical applications," *IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc.*, vol. 11, no. 2, pp. 177–180, Jun. 2003.
[40] R. C. Oldfield, "The assessment and analysis of handedness: The Edinburgh inventory," *Neuropsychologia*, vol. 9, no. 1, pp. 97–113, Mar. 1971.
[41] R. Roberts, N. Callow, L. Hardy, D. Markland, and J. Bringer, "Movement imagery ability: development and assessment of a revised version of the vividness of movement imagery questionnaire," *J. Sport Exerc. Psychol.*, vol. 30, no. 2, pp. 200–221, Apr. 2008.
[42] D. F. Marks and A. R. Isaac, "Topographical distribution of EEG activity accompanying visual and motor imagery in vivid and non-vivid imagers," *Br. J. Psychol. Lond. Engl. 1953*, vol. 86 (Pt 2), pp. 271–282, May 1995.
[43] C. Brent, "Video Game Addiction Test for Parents - Is Your Child Addicted?" [Online]. Available: http://www.techaddiction.ca/video-game-addiction-test.html. [Accessed: 02-Mar-2015].
[44] E. Adams and B. Ip, "From Casual to Core: A Statistical Mechanism for Studying Gamer Dedication." [Online]. Available: http://www.gamasutra.com/view/feature/131397/from_casual_to_core_a_statistical_.php. [Accessed: 05-Jan-2015].
[45] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *Audio Electroacoustics IEEE Trans. On*, vol. 15, no. 2, pp. 70–73, 1967.
[46] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, vol. 134, pp. 9–21, 2004.
[47] L. Pugnetti, M. Meehan, and L. Mendozzi, "Psychophysiological Correlates of Virtual Reality: A Review," *Presence*, vol. 10, no. 4, pp. 384–400, Aug. 2001.
[48] G. Pfurtscheller and C. Neuper, "Dynamics of Sensorimotor Oscillations in a Motor Task," in *Brain-Computer Interfaces*, B. Graimann, G. Pfurtscheller, and B. Allison, Eds. Springer Berlin Heidelberg, 2009, pp. 47–64.
[49] "Theta brainwaves reflect ability to beat built-in bias," *ScienceDaily*. [Online]. Available: http://www.sciencedaily.com/releases/2013/05/130507195648.htm. [Accessed: 30-Mar-2015].
[50] G. Buzsáki, C. A. Anastassiou, and C. Koch, "The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes," *Nat. Rev. Neurosci.*, vol. 13, no. 6, pp. 407–420, Jun. 2012.

Back to Contents

Accessible Game Culture using Inclusive Game Design

Participating in a visual culture that you cannot see

UIf Wilhelmsson, Henrik Engström, Jenny Brusk, Per-Anders Östblad
School of Informatics
University of Skövde
Skövde, Sweden
ulf.wilhelmsson@his.se

***Abstract*— In this paper, we present the result of an experiment, in which we compare the gaming experience between sighted players and visually impaired players playing the same game. Specifically we discuss whether they experience the same story construed from the plot elements that are either manifested by audio and graphics in the case of sighted players or primarily by audio in the case of visually impaired players. To this end, we have developed a graphical point-and-click adventure game for iOS and Android devices. The game has been designed to provide players with audio feedback that enables visually impaired players to interact with and experience the game, but in a manner that does not interfere with the overall appearance and functionality of the game, i.e. a design that is fully inclusive to both groups of players and that is as invisible for sighted players as possible without hindering visually impaired players to share the same gaming experience when it comes to story content. The study shows that the perception of the story was almost identical between the two groups. Generally it took visually impaired players a little longer to play the game but they also seem to listen more carefully to the dialogue and hence also build a slightly deeper understanding of the characters.**

Keywords—***inclusive game design, narrative, story, plot elements, visually impaired players, sighted players***

I. Introduction

Playing video games has become more common during the last 10 years, and the number of children having access to different platforms for playing video games, including computers, game consoles, handheld devices, and especially smartphones has substantially increased during this period [1, 2, 3]. According to [1], almost all of Swedish children ages 9-12 play video games (98% of the boys and 97% of the girls). While boys continue to play in their teenage, girls tend to favour social media instead. The amount of time spent on playing video games also differs substantially between these groups. While less than 10% of the girls play 3 hours or more per day, almost every other teenage boy play at least 3 hours per day. The prior study [2] from 2012/2013 reported that while girls preferred to play solo, boys in this age group tend to play games in a social context with friends, siblings, and to a lesser degree, parents or other grown-ups and the percentage of players playing video games in solitude has decreased.

The vast majority of mainstream commercially available video games rely heavily on their graphical components as the main communication with the player and is used for defining and diversifying game characters, creating the game environment, and separating interactable objects from non-interactable objects. Most often videogames are designed to use visual cues such as written language and symbols as an integral part of both the game environment and the game interface. Even the term video game in itself is part of a visually dominated culture.

Very few of the games on the market are accessible to visually impaired players and those that are, are usually designed primarily for the visually impaired, rendering them less attractive for players that are not [4, 5, 6]. One of the reasons is possibly that the economical incitements for producing commercial video games primarily for visually impaired players are small. This is not a mass market why it seems that a majority of video game companies do not address this subject at all or to a very small extent.

Organizations, such as the *AbleGamers Foundation*, have had a positive effect on some of the larger game developers to make more inclusive game designs concerning different kinds of impairments. Only quite recently, well-functioning support systems have been built into the operating system on computers running a graphical user interface or into mobile units such as smartphones or tablets to make these more accessible. Apple’s iOS VoiceOver system is an example hereof.

This lack of inclusive games creates inaccessibility to a progressively important social context – game culture – for the visually impaired. There is a risk that an already marginalized and to some extent socially excluded group will be even more marginalized if they are not given the same possibilities and opportunities to participate in this particular social context and share the same kind of experiences on as equal terms as possible. We find this problematic, especially concerning visually impaired children since this constitutes an obstacle for them to play games, in the broad sense of the concept of games, and hence to gain access to the mediated culture communicated through the act of playing games of different kinds.

978-1-4799-8102-1/15/$31.00 ©2015 Crown

In an attempt to bridge this gap, we have developed a graphical point-and-click adventure game called *Frequency Missing* for iOS and Android devices. The game has been designed to provide players with audio feedback that enables visually impaired players to interact with and experience the game, but in a manner that does not interfere with the overall appearance and functionality of the game, i.e. a design that is fully inclusive to both groups of players and that is as invisible for sighted players as possible without hindering visually impaired players to share the same gaming experience when it comes to story content.

II. Background

Play and games possess a social dimension in that even solitary play and games often turn into competitive play between potential players [7]. This can manifest itself by players comparing individual skills such as high scores, chat about their experiences and progress in games in online forums etc. In other words, play and game, even though competitive or played solitary, are part of an important sociocultural context. Games are vehicles of social and cultural content and there is a constant interplay between games on the one hand and society/culture on the other [7]. Furthermore, "Play is not merely an individual pastime. It may not even be that as frequently as is supposed." [7].

Video games are in respect to this not all that different from other types of games and they have rapidly become a large part of everyday life for children [1] as well as adults [8].

As Embrick, Wright, and Lukacs notes [9], "There is a tendency to dismiss online virtual programs such as *World of Warcraft* and *Everquest* as 'games' that people play either for fun, or to escape the realities of everyday life. What many of us fail to recognize is that these communities, although virtual, are extensions of the society in which they reside.". Video games replicate social structures outside of games and to be accepted as a gamer on equal terms within a community of gamers requires a player that is able to enter the social context of game playing. This might be problematic for visually impaired gamers.

There has been a trend to integrate visually impaired and sighted pupils in the same educational situations during the last three decades [10]. In Sweden, prior to 1986, visually impaired children without other disabilities mostly attended the special boarding school for the blind, but since almost thirty years they are part of the ordinary school system and hence have classmates that are not visually impaired. The main reason for this was to promote social inclusion and reduce marginalization [11]. Special education for different groups of pupils does not always lead to good results as far as integration goes. As has been pointed out by Söderqvist Dunkers [12] this attempt to include visually impaired children was not unproblematic. There are thresholds that need to be overcome in the social relations in the school environment's peer culture. According to Svensson referred to by [12] only one third of the visually impaired pupils had good social relations with other pupils in their initial years in ordinary schools.

The same is valid for games and play. Special games and play designed to address specific disabilities at the cost of overall game design is not inclusive but separating even more. Play has in its essence a potential of integration. It could serve as a tool for inclusion. Corsaro notes [13] that in order to successfully enter play in a peer culture, a player needs to first position him/herself in a socially accepted distance to the players. The next step is to find out what the theme of the play is, what the content and scope of the play is and then make a useful contribution to the play activity that the peers playing can accept. For a blind player, who needs to rely on auditory information and touch, finding out what the theme is can be difficult in many play situations. Making a useful contribution is likewise problematic for the same reason.

Games targeted at visually impaired players are often not inclusive with respect to sighted players, or are designed to use only audio to communicate with the player which most sighted players are unfamiliar with [4, 5, 6, 14].

To completely rely on audio in game design is still uncommon and could be considered as a subculture in itself. Two rather well known and well received games that are designed to use only audio to play the game are *Papa Sangre* [15] and *Papa Sangre II* [16]. The first of these was awarded Most Innovative Game at Mobile Gaming Awards in 2011, and Papa Sangre II had a rating of 92% on the iOS platform in 2013, making it the best-reviewed game on iOS that year. But awards and positive ratings do not necessarily indicate that these two games have or will set a new standard for how games are primarily designed. They are novelties and very interesting games, but they have not yet had too much impact on the video game industry at large.

III. Point and Click Adventure/puzzle Games as a Point of Departure

Adams defines adventure games as: "*[...]an interactive story about a protagonist character who is played by the player. Storytelling and exploration are essential elements of the game. Puzzle solving and conceptual challenges make up the majority of the gameplay. Combat, economic management, and action challenges are reduced or non-existent."* (his italics) [17]. We chose this specific type of game since inherent in these games' basic structure is a reliance upon text and spoken words to communicate the key plot elements of the game. The first adventure games, such as *Colossal Cave Adventure* (a.k.a ADVENT) [18] and *Zork* [19], prove this point as they relied solely on written language as a means to play the game. Games like these are story driven and rather linear i.e. they start from the very beginning of the chain of events and the player follows that path to the end of the game (which might vary due to the choices the player makes during game play). Later games that combine graphical manifestations of the game environment with written language and spoken word, sound effects and music, have commonalities with the text only games: riddles, puzzles, problem solving and exploring the game environment. Such games, like *The Secret of Monkey Island* [20] and *Day of the Tentacle* [21], also established a structure how written language is displayed on the screen, how the player can control what to say in a dialogue with game characters, how they can store objects they pick up along the path in an inventory etc. A game such as *Day of the Tentacle* is due to its basic design and organization of interactable objects on the screen, a good point of departure to construe a game that can be played without any graphical content as long as the structure of the interface is consistent throughout the game. As long as the player can correctly target the interactable objects on the screen the game can be played.

Quite a few games in the genre of point and click adventure/puzzle games use a voice over from a first person point of being to communicate with the player when clicking objects on the screen. This approach mimics the narrative style of other media such as novels and film where this is quite common and well established. Used in a game, this approach provides yet another level of communication; it makes possible a prosodic emphasis on certain qualities of the object when playing the game with sound. In a written narrative, the emphasis could be signalled by utilizing different typographical styles (bold, italics etc.) or providing the reader this kind of information by writing how a specific utterance is said. In point and click adventure games it is part of the genre conventions to let the playable character say things like "I can't use these things together!" [23] to help the player notice specific qualities of interactable objects.

We deliberately chose a story driven design since that allowed us to study how well the content of the game was understood by the players that tested the game. Narratives within the same genre are quite uniform due to genre conventions developed over time. A mystery story usually contains the same kind of plot elements making up the story.

IV. Designing the Game

Frequency Missing is a mystery point and click adventure/puzzle game designed for iOS and Android tablets or smartphones. These are the dominating platforms among visually impaired as well as sighted users. It was therefore something that from the very outset connected visually impaired and sighted people, a common base level of using the same kind of devices. It also made possible, a design that is touch based, i.e. the way to interact with objects on the screen is to touch them directly why there is no need to use a keyboard or mouse to play the game.

The story is the recountable chain of causal events constructed from the experienced plot elements, which are the actual interactable events presented to the player over duration of time. As such, the story is the mental construction that players make from all the plot elements, and is furthermore subject to speculation on what might happen in the next series of events triggered by the player´s actions within the game environment. That is to say, that while playing, the story is under construction by the player when choosing between the plot elements present and due to this, recounting the story after a play session, is affected by how many of the plot elements that have been part of the interaction with the game. This is not unique to games but is equally valid in how an audience interpret a story from narratives at large.

The story of the game is centered entirely on a female reporter named *Patricia*. She is the focalizer [22] of the events why it is from her presence in the game environment alone that the plot elements are possible to engage in. Patricia is the player's agent, the game ego, but she is not a visible character since the game is enacted from a mix of first person and God's perspective in its graphical construction.

There are no other playable characters in the game but quite a few non-playable characters that Patricia can interact with, mainly through dialogue organized in a branching tree structure. Patricia can also trigger reactions from the other characters by using interactable objects in the game environment.

The game begins with Patricia being trapped in a secluded dark environment and the player can explore this environment by touching the screen. After having solved the first puzzle the next plot element is Patricia's first day at her new work at a radio station, which is preceded by a voice over and a text setting us back three days in time. Patricia learns that her friend *Richard*, who helped her get the job at the radio station, has mysteriously disappeared. The objective of the game is to find out what happened to Richard. The setting of the remainder of the first chapter is the radio station, consisting of a lobby from which three offices and a newsroom are accessed. At the station, the player gets introduced to her new colleagues: *Karl*, a technician and Patricia's office mate; *the manager* (he is not given a name). *Monica*, the receptionist; and, *Stephanie*, the news anchor. Through dialogue with these characters, the player can find out more about the colleagues and the circumstances around Richard's disappearance. In addition to dialogue, the story is narrated by means of Patricia's inner voice, a common feature of point-and-click adventures. The story is inspired by film noir and is set in the 1960's, which is manifested in the visual style (see Figure 1) and the dark jazz music played in the game. The main motivation for using this particular setting was that it opened up the possibility to use older iconic sounds like bell telephones and typewriters instead of their contemporary counterparts mobile phones and computers.

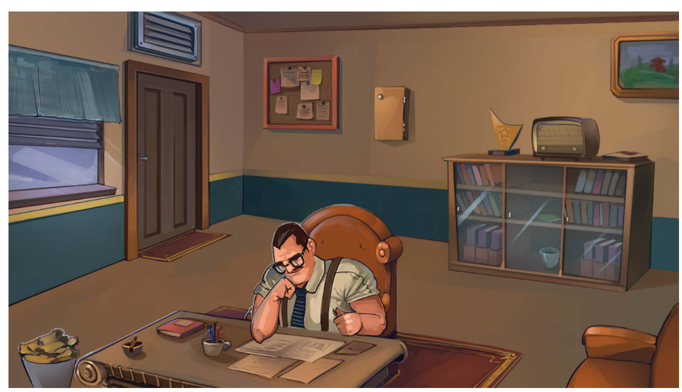


Fig. 1. A screen capture from *Frequency Missing*, illustrating its visual style.

A major challenge was to achieve clarity in the sound design and audio mix. Since the game must be equally playable independently of whether it is played with or without graphics, all information and events that are conveyed by graphics must also be conveyed by sound. All rooms, for example, have unique ambiance tracks that reflect the setting of the scene, e.g. in the lobby there is low murmur of voices etc.

Similarly, all interactable objects have their own unique 3D sound, meaning the closer you are with your finger to the object, the stronger its sound volume becomes and you can tell if the object is to the left or right of the finger. In addition to 3D sounds, when an interactable object is located, there is a clear and audible interface sound. Every action and event that is not diegetic (directly connected to the game environment) is highlighted by an interface sound. All interface sounds are separated and made audible in the mix by being spread very wide in stereo width.

All dialogue is voice acted by experienced actors, following a hypothesis that using text-to-speech could reduce the level of immersion the player could experience. To make all speech clear and audible, since audio-only players can't access the subtitles, all background sounds (music, ambiance) are lowered in volume and filtered in their high frequencies whenever a character speaks.

V. User Test

The game has been evaluated through formal user tests where subjects have been asked to play a section of the game followed by an interview. We have used a qualitative approach for evaluating the subjects' progress in the game and their perception of the story with the aim to find out whether the subjects shared the same experience of the narrative in the game. The experiment was conducted according the following procedure:

1. The subject is given a brief introduction to the purpose of the study and information on what data that are collected and that their participation is voluntary and anonymous.
2. The subject is equipped with headphones. After an adjustment of the volume according to the subject's preference, the game is started.
3. The subject plays a short tutorial and then the first chapter of the game. If a subject does not complete this stage within 25 minutes, the test leader ends the session.
4. The subject answers open questions regarding the experience including aspects related to the story and interaction.
5. The subject answers questions regarding background parameters such as age and experience of gaming.

The gaming device used during the tests was an iPod touch 5G, with a 4" touch display. The headphones were a pair of AKG 240 Studio. All responses from subjects were transcribed directly by the experiment leader. No audio or video recordings were made. The total time to complete the whole procedure was 50-60 minutes. As a compensation for participation all subjects were given a €15 iTunes gift card.

During the play session, the progress of the subject was monitored and if a subject spent three consecutive minutes without making any progress, the experiment leader provided some hints. The nature of the hints given depended on the type of obstacles a subject experienced. It could for example be related to the game logic (e.g. to search for items), the feedback system (e.g. to listen for interface sounds) or the subject actions (e.g. to make slower swipes on the touch screen).

The progress subjects made in the game has been evaluated in terms of how far they reached in the *Critical Path* (CP) shown in Figure 2. The CP consists of 31 steps, see Figure 2, of which the first four are tutorial steps before the actual game begins. These 31 steps are the absolute minimum number of interactions that has to be completed to finish the chapter.

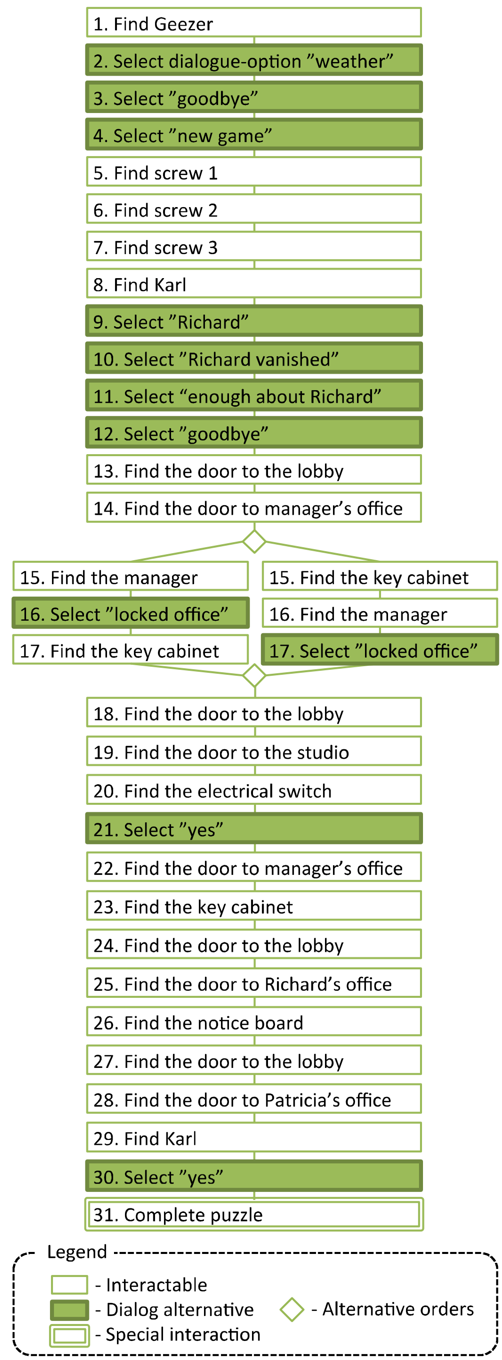


Fig. 2. The objectives of the 31 steps in critical path for the test session

The open questions in stage 4 of the test session have been coded in different categories of which perception of setting, characters and story are the main focus for this paper.

In total, 19 subjects participated in the test, divided into two main groups: sighted subjects (S-subjects) consisting of five men and five women and visually impaired subjects (VI-subjects) consisting of five men and four women. These groups are not homogeneous. There are for example variations in age (from 15 to 39 years of age) and previous experience of using smartphones and games. From a methodological perspective this is a strength rather than a weakness since our focus was to qualitatively investigate whether diversified players could share the same gaming experience. Of the VI-subjects the degree of impairment varied: some were blind from birth, some could distinguish light from dark and some could read the text on the screen.

For the scope of this paper, eight open questions from the test are of particular interest:

1. Your spontaneous reactions (to the game)?
2. What is the game about?
3. What is the setting of the game?
4. What are the characteristics of the protagonist of the game?
5. Which characters are there in the game and what roles do they have?
6. What do you think will happen next in the game?
7. What do you think has happened to the colleague who has disappeared?
8. Who do you think is the villain?

These questions all aim to explore to what extent the content of the game has come through to the subjects and if they after playing the first chapter or parts hereof, were able to make assumptions on how the story will unfold and if the subjects enjoyed playing the game. Questions two and three are mainly concerned with the framing story and the location in which it takes place. Questions four and five address the characters in the game and how their characteristics and respective roles in the game have been perceived by the players. Questions six to eight address to what extent the game's plot elements have been communicated in order for the player to construe a story from them. For example, have the plot elements fired the players' imagination so that they are able to draw speculative conclusions from the various plot elements they have interacted with?

VI. Test Results

A. Game Progress

In the group of sighted players all but one reached step 31 in the CP. Among the VI-subjects the result is more diversified, ranging from CP step 13 to 31. Only four subjects in this group reached step 31 of the CP during the test session of 25 minutes.

For S-subjects who finished the chapter, the average time used was 18 minutes. This was slightly longer for the VI-subjects who used 22 minutes on average to finish the chapter. It is notable though, that some VI-subjects completed the game in a shorter time than some S-subjects. The remaining subjects, i.e. those that did not reach step 31 in the CP, used all the allotted time. None choose to voluntarily end the game session.

Two S-subjects were provided with hints during their game session, according to the three minute rule presented above. The hints related to problems to locate items in the introductory step (CP 5-6) and problems with the dialog system. Four VI-subjects were provided with help and for three of those, it related to game logic. The fourth subject experienced problems with the introductory stage and dialog system.

B. Gaming Experience

Question 1: Your spontaneous reactions (to the game)?

Both groups were mainly positive towards the game. Most importantly there were no explicit spontaneous comments from the S-subjects that the inclusive design interfered with the gameplay in a disturbing way. One of the VI-subjects did not in general like touch screen interfaces since she found those hard to navigate but was despite this positive towards the game.

Among S-subjects, four of them made comments on the graphics such as "I liked how it looked. It was nice looking", "It was a bit difficult in the beginning to understand that you should draw [your finger] in order to get the circle. I'm an impatient person so I read faster than they talk.", "If one didn't look at the screen it was hard to find the doors, hard to find objects with background sounds. It felt like more sounds than I was looking for. The ploink when missing the menu was muddled.", and "Didn't notice it was a scrollbar."

All VI-subjects had positive comments about the content of the game and the gaming experience. Two of them wanted clearer instructions concerning how the game interface works. One subject said "First I didn't understand what one should do. Where to go. It took a while to understand that one should talk to Karl. The logic of getting into Karl's office was hard to understand." On the other hand, another subject said "Interesting. It was… one immediately understood what one was supposed to do. It wasn't hard." Both these subjects reached CP 31 in less than 22 minutes.

One subject (CP 31) said "I thought it was a good game. It was clear but just enough. Otherwise I think that other games [for visually impaired players] can be humiliatingly clear."

These spontaneous comments indicate that our design is not overtly or patronizingly supportive in a disturbing manner for either S-subjects or VI-subjects. The players seems to have found the game treating them as players and not impaired players in the case of VI-subjects on the one hand and not sacrificing playability on the other in the case of S-subjects. There is of course a risk that the subjects have been polite and do not really put forth their true opinion of the game. However, the answers on most questions are relatively rich and consistent which indicates that the game design works as intended.

Question 2: What is the game about?

All subjects in the test identified the main objective of the game i.e. to find out what has happened to the protagonist's missing friend and colleague Richard, without any problem. This is one of the first things that are communicated through the plot elements, why also subjects that only reached CP step 13 could identify this objective. The answers are more or less elaborate and the VI-subjects are more verbal and elaborate in their answers. Three subjects in each group said that Richard has been missing for three days.

Question 3: What is the setting of the game?

All subjects except two VI-subjects mention a radio station as one of the settings. One says "media station" and one "in an office space." The latter later mentions the radio station setting in another answer so it is not the case that the subject didn't identify the setting but rather misunderstood the question. The location is explicitly introduced at step 8 of the CP why all subjects in the test have been exposed to it. The cabinet in which Patricia is locked in, is part of a later location, a club, but there are no plot elements revealing this in the chapter played.

Question 4: What are the characteristics of the protagonist of the game?

There are not any notable differences between the two groups of subjects when they describe the protagonist's core traits. For example, 7 of the S-subjects and 5 of the VI-subjects describe the protagonist as curious. She is also described as talkative and stubborn by subjects in both groups. Most subjects in both groups actually say very few things about her. The protagonist is designed to have a core personality that will bias the player to interact with the game environment in a specific way but that also allows for the players to identify with her as their agent in the game.

Question 5: Which characters are there in the game and what roles do they have?

In total there are seven characters that can be identified by name and/or role in the chapter played. There are also a number of noninteractable prop characters in the lobby of the radio station that have an audiovisual presence in the environment but that only serve as a means to make the environment more vivid. Out of these characters, four are audiovisual characters. Karl (sound technician and tech guy, colleague), Monica (the receptionist), Stephanie (news anchor) and the manger. In the chapter, Richard – the missing colleague – is mentioned by name as is *duke Silverglans* who is said to own a social club. Of course the protagonist Patricia is also part of the first chapter but as noted she is not visually represented but only present by audio and through the player's manipulation of the game environment. The number of identified characters ranged between 4 and 6 for S-subjects and 3 and 6 for VI-subjects. There is in other words only a small difference between the groups, which can be explained by progress in the CP.

The VI-subjects as a group provide slightly more comments concerning the characters' personalities than the S-subjects. The basic roles in the game are more protruding in the group of S-subjects. What really comes through in the data is that the efforts put in the character design of the manager worked (Figure 1 shows his visual appearance). All subjects identify him as one of the characters and he is the character that has the most comments regarding his personality. He is deliberately designed as a stereotype of a manager (not too friendly towards his employees, uninterested, concerned about the image of his enterprise etc.). He is also designed to attract the attention of the players as a main suspect behind Richard's disappearance. The VI-subjects generally also provide more comments on his personality than the S-subjects. The roles of the characters are primarily described by all subjects in relation to their subjective construction and interpretation of the story. As an example hereof one of the S-subject (CP 31) answered "Patricia who was looking for his friend and colleague. Karl in the first room, also colleague. The three felt like they were working in a team. She who was in the lobby, receptionist. Her role I do not know. The manager may know more than he wants to talk about. [...] she was down there who broadcast radio, she was mentor to Patricia." Compared to one of the VI-subjects (CP 31) who answered "On the one hand Patricia who was the main character. Her role was to find Richard and bring the story forward. She had a co-worker in the same office called Karl. His role was to be Patricia's backup and supported her. A person named Stephanie who was Patricia's mentor. You were not allowed to meet her so much. Was told that she was cocky and competent. The manager was targeted. nasty. One person I do not remember her name, she was kind and informative [Monica]. Richard but we don't meet him and one does not get to know much about him. Respected employee who ensured that Patricia got the job." Both these subjects cover the same main topics and the interrelation between the characters and they have a similar experience of playing the game.

Question 6: What do you think will happen next in the game?

All subjects in the test could speculate on the forthcoming events in the game from their respective position in the CP. The answers vary from speculating on the next chapters to what could happen within the first chapter depending on this. The answers also vary from what they thought would happen very close in time to speculations on how the game as such might end. As in the previous question the answers vary in accordance to how the subjects have construed a story. One of the S-subjects (CP 31) notes that the game environment is built like a big library and hence "Maybe you will find hidden passages behind book shelves. Did not get the feeling that there was something supernatural. More a point click where you have to be a detective without becoming too rebellious towards the manager." None of the VI-subjects have comments about what will happen next in ways similar to this manner. Instead they are rather more to the point in finding clues to solve the mystery. This is probably to some extent an effect of how far they have reached in the CP but even the VI-subjects that reached step 31 provide less speculative answers than S-subjects.

Question 7: What do you think has happened to the colleague who has disappeared?

The most common answer to this question is that Richard has been abducted/kidnapped. There were eight S-subjects and five VI-subjects who suggest this. Other suggestions are that he is voluntary absent (one S-subject), locked in at the radio station by the manager and Stephanie (one S-subject). One VI-subject suggests that he has been murdered. Six of the S-subjects and two of the VI-subjects provide answers that, rather than only addressing the question, suggest *why* Richard is missing. The most common suggestion is that he has been researching something that might be a threat to someone which is a quite common theme in this genre of narratives.

Question 8: Who do you think is the villain?

The game begins (step 5 to 7 in the CP) with a puzzle sequence in which the protagonist is locked into a cabinet and needs to find a way to get out. This is indicative that someone has actually locked her in. Six S-subjects and two VI-subjects already had begun to speculate who might be a villain, as question 7 indicates. Three S-subjects and one VI-subject think it is someone that is not yet introduced in the game. The manager is the most common suspect but the speculations spans over all characters that are part of the chapter, and even beyond that since there are also speculations that the villain has not yet been present in the game. As noted above (question 5), the manager is the character that is deliberately designed to attract attention from the players

VII. Conclusions

A. Shared Experience Through a Common Ground

From the test we can conclude that the two groups, sighted players and visually impaired players, are able to share the experience of playing the game since they construe the same basic story communicated by the plot elements. Actions and interactable objects in a game do not need to be visual. By using dialogue and narrative voice over as well as clear and well thought out sound design, actions can be defined by audio.

By choosing a game genre which has its primary origin in a long tradition of storytelling, a mystery story, we also chose a common ground between sighted and visually impaired players. The test results indicate that both groups of players have been able to construe similar basic stories, the same setting and have a common idea about the personality of the playable character Patricia from the plot elements in the game. They also speculate and put forth similar prognosis concerning what will happen in the next parts of the game. There are differences in how detailed the descriptions are but the core plot elements are recounted by all participants in the test: Richard has disappeared and Patricia wants to find out what has happened to him.

B. Inclusive Game Design

We have also succeeded in designing a game that is inclusive for both groups in that it does not incorporate any obvious distractions to make the game playable and enjoyable for visually impaired players since the strategy we have adopted does not interfere with the gameplay for the sighted player, but is an integral part of the game. The setting of the game at a radio station which is a predominantly audio culture rather than visual culture, as backdrop for the events, makes it possible to integrate audio puzzles as crucial plot elements. From the data gathered in the tests, we can also conclude that the time spent playing the game might pose a problem when sighted and visually impaired players play together. Most sighted players finished the game in shorter time than the visually impaired players in the test. This might pose a problem when a sighted and a visually impaired player plays this kind of game together or at least concurrently, since the sighted player might be impatient towards the slower progress of a visually impaired player. This might be yet an area where an even more inclusive design might be necessary, a design in which different competences complement each other rather than put one competence in favour of another, and in which the time aspect is less of a limitation. For instance, question 2 (see section VI.VI.B, above) indicates that visually impaired players listen more carefully to what is said in the dialogue sequences of the game. A game designed for both groups could for example be more focused on cooperation and benefit from the differences between the two groups of players making them more interdependent.

C. Accesible Game Culture

As noted, the use of digital games is widespread and increasing among youngsters [1]. With our project, we have begun to make the continuously growing game culture and its social dimensions more accessible for visually impaired players. We have found ways of designing video games in a manner that makes them accessible not only for either visually impaired or those that are not but for both these groups. Accessibility is however not the only important issue, but more importantly we have found ways to make a video game joyful to play and emotionally rewarding for both these groups. The design does not frighten sighted players off, and it is not patronizing towards visually impaired players. Both groups can contribute to the act of playing since they share the same basic experience.

D. Future Work

We will need to perform more tests to explore whether the game will make it possible for visually impaired players to join a group of sighted players by following the steps that Corsaro [13] suggests for entering a play situation. That has been beyond the scope of the current study. However, as a first step of designing a game that is enjoyable and comprehensive for sighted and visually impaired players alike, our design has crossed an important threshold; it creates a common ground between these two groups by using the force that lies within narrative as basis for the design. The project has the potential to make digital games even more widespread since we have found ways to design such games in an inclusive manner. As part of the next step of our project on inclusive game design we will design a game where a sighted player and a visually impaired player can play together co-operatively via network using one iOS device each.

ACKNOWLEDGMENT

This project has been co-funded by The Swedish Post and Telecom Authority and the University of Skövde.

REFERENCES

[1] Swedish Media Council, Kids & Media 2015, (Statens Medieråd, Ungar & Medier 2015, Stockholm, 2015)

[2] Swedish Media Council, Kids & Media 2012/13, Stockholm, 2013 (Statens Medieråd, Ungar & Medier 2012/13, Stockholm, 2013)

[3] Dataspelsbransc hen, Game Developer Index 2014, 2014

[4] M. T. Atkinson, S. Gucukoglu, C. H. C. Machin, and A. E. Lawrence. “Making the mainstream accessible: Redefining the game”, in Proceedings of the 2006 ACM SIGGRAPH Symposium on Videogames, Sandbox ’06, New York, NY, USA, 2006.

[5] I. Ekman. “Sound-based gaming for sighted audiences–experiences from a mobile multiplayer location aware game”, in Proceedings of the 2nd Audio Mostly Conference, pp. 148–153, 2007.

[6] D. Grammenos, A. Savidis, and C. Stephanidis. □”Designing universally accessible games”, Comput. □Entertain., vol 7(1):8, pp. 1–29, Feb. 2009.

[7] R. Caillois, Man, Play and Games. Urbana and Chicago, Illinois University Press 1961/2001

[8] Entertainment Software Association, Essential facts about the computer and video game industry, 2015

[9] D. G. Embrick, T.J. Wright, and A. Lukacs , Eds., Social Exclusion, Power and Video Game Play: New Research in Digital Media and Technology. New York, NY: Lexington Books (Division of Roman & Littlefield). P. xv, introduction, 2012

[10] K-G. Stukat. Three Decades of Integrated Education. in Education & Children with Special Needs – From Segregation to Inclusion, Eds., S. Hegarty, M. Alur, pages 173-180, Tejeshwar Sing for sage Publications India Pvt Ltd 2002 (third edition 2005)

[11] Synskadades riksförbund. Skolor och undervisning http://www.srf.nu/om-oss/synskadades-museum/Synskadades-historia/Skolor-och-undervisning/

[12] T. Söderqvist Dunkers, Talande möten mellan ungdomar: Villkor för delaktighet i kamratgemenskap - ungdomar med synnedsättning berättar, Master thesis, Dept. for pedagogy and didactics, Stockholm: University of Stockholm, 2011 (English title: Talking encounters: What do young people with severe visual impairment say about terms of participation in the peer culture?)

[13] W. Corsaro, The Sociology of Childhood. Thousand Oaks, London, New Delhi: Pine, Forge Press. 1997

[14] M. A. Oren. Speed sonic across the span: building a platform audio game. In *CHI’07 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2007.

[15] Somethin’ Else, Mobile game, *Papa Sangre*, 2010.

[16] Somethin’ Else, Mobile game, *Papa Sangre II*, 2012.

[17] E. Adams, Fundamentals of game design, 2nd ed., Berkley, CA, New Riders, 2010.

[18] CRL, Computer game*, Colossal Cave Adventure*, 1976

[19] Infocom, Computer game, *Zork I: The great underground empire, 1980*

[20] Lucas Arts, Computer game, *The Secret of Monkey Island*, 1990.

[21] Lucas Arts, Computer game, *Day of The Tentacle, 1993*

[22] G. Genette, Narrative Discourse, Ithaka, N.Y, Cornell University Press, 1983

[23] Lucas Arts, Computer game, *Sam & Max Hit the Road,* 1993

Sort Attack: Visualization and Gamification of Sorting Algorithm Learning

Alfa R. Yohannis, Yulius D. Prabowo
Faculty of Computer Science and Communication Science
Institut Teknologi dan Bisnis KALBIS
Jakarta, Indonesia
alfa.yohannis@kalbis.ac.id

***Abstract*—Algorithms are commonly perceived as a difficult subject, which is quite an irony as they have a fundamental role in computer science. Failure to master this subject will inhibit students' capabilities as they advance to higher levels. Algorithm visualization, as an effort to overcome the problem, that has been growing towards gameful visualization recently it is presumed to be able to engage the learners longer and more intensely. However, integrating algorithm visualization, game elements, and instructional design is not a trivial task as it requires a careful design. Hence, a conceptual model of how algorithm learning instructions, algorithm visualization, and gamification improve learning outcomes was developed. While instructional design concerns with developing the best strategy for learning, algorithm visualization functions as a cognitive support provider and gamification works by delivering engagement. Both cognitive support and engagement moderate the learning instructions that concern with enhancing learning outcomes. Principles and elements originating from the three domains have to be taken into consideration during the designing process to produce an artifact that can deliver the essential functions of each domain. A working artifact was then constructed, as the instantiation of the model, to validate whether the idea of integrating algorithm visualization and gamification into algorithm learning instructions is capable enough to improve learning outcomes. Based on our respondents' learning outcomes, it was found the artifact can significantly improve the procedural knowledge of learners that are indicated by their increased capability in solving sorting algorithm problems.**

Keywords—algorithm visualization, gamification, algorithm learning, instructional design, sorting algorithms

I. INTRODUCTION

Algorithms are effective, finite, and deterministic problem-solving methods or can be described as procedures for solving problems. Algorithms are essential for computer science as they are fundamental objects of study in the discipline [1]. Less proficiency in algorithms will likely cause computer science students to face difficulties in completing their degree, as most of computer science subjects have a sense of intrinsic algorithm problems. Their perception of algorithms will affect their current attitude towards computer science as well as their future career paths.

For new computer science students, algorithms are commonly perceived as a subject that is difficult to learn. Algorithms, by their nature, are abstract and often complex, thus requiring a high cognitive load to understand them. This problem is similar to a problem in mathematics where much of the concepts can only be accessed through semiotic representations [2]. We call this problem as the abstract barrier and algorithm visualization is one of the several approaches that have been conducted so far in order to pass the barrier. Algorithm visualization works by concreting the abstract, so learners can directly interact with algorithms and perceive the outcomes.

On the other hand, in the last two decades, the use of games for learning has been growing and attracting many attentions of researchers from various domains. Serious games [3] and gamification [4] lately are two popular key terms coined that represent this research trend. By using games or game elements, learners are believed to become more motivated and engaged with their subjects and therefore they might improve their learning outcomes. This paper will use the gamification term more often as the umbrella term that comprises both the serious games and the gamification itself, since gamification can be perceived as a process which then can transform a gameless entity into a serious game, a complete gameful entity but with a balance in its inherent characteristics [5], and gamification is a term that is broader than serious games [6].

There is a prospect for algorithm visualization to be unified with gamification, providing not just a cognitive support for learners to understand and explore algorithms, but also a motivating environment to improve learners' engagement with the algorithms. Both algorithm visualization and gamification are prospective means to support instructional design in improving knowledge acquisition. Unifying them facilitates cognitive support and engagement to be both delivered by a specifically designed artefact; in this case, a game is believed to be able to improve the learning outcome of learners. Therefore, to answer such expectation, we derive two questions that are addressed in this research. First, how do we design an artifact which unifies gamification and visualization, and instructional design which delivers to be not only engaging but also promoting cognitive support and knowledge acquisition for learners to understand algorithms? Second, to what extent does the instance of artifact improve learners' learning outcomes overall?

This paper is organized as follows, first, this paper starts with an introduction that presents the background and motivation of this research. Next, the research methods that are

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

employed are going to be presented. Afterward, the related works and theories are investigated to identify the main pattern of algorithm visualization development and to derive a conceptual model of how algorithm visualization, gamification, and learning instructions interact to improve learning outcomes. After that, the design consideration of the artifact of this research is presented. Subsequently, the significance of the artifact—to what extend it improves learning outcome—is examined. Finally, this paper ends with conclusions as well as the next challenges faced by future works.

II. Research Methods

There were three phases in conducting this research. In the first phase, a literature study was performed to derive a model based on existing related works and theories. The underlying important constructs and their relationships were required to be recognized first and understood well before they could be applied as a conceptual model to design the artifact. Therefore, in section IV, how our conceptual was derived is explained. The conceptual model itself originates from three main domains, instructional design, algorithm visualization, and games, and consists of six constructs, algorithm visualization, cognitive support, learning instructions, gamification, engagement, and learning outcomes.

The second phase was designing the artifact. The identified constructs had to be supported by the artifact, so later how they affected learning outcomes were possible to be examined. Therefore, the artifact had to be designed and developed carefully by incorporating elements that could support the presence of all the identified constructs. Several principles and elements originating from the three domains were considered and integrated into the design of the resulted artifact. Our design consideration is explained in section V.

In the third phase, the artifact was tested on learners to measure to what extent the artifact, as the representation of cognitive support and engagement, influences learners' outcome. We took a sample consisting of 59 informatics students from our university, from the second-year and the third-year students. The sample was then separated into two groups, the book-first group, 31 students, and the game-first group, 28 students. Both groups were given 30 minutes to study sorting algorithms materials. However, book-first group studied the materials only using the textbook, the traditional way while game-first book studied the materials using the artifact. After 30 minutes, both groups were tested—the first test—by giving them a set of problems, several series of unordered numbers, and then they had to sort them using three sorting algorithms, namely insertion sort, selection sort, and bubble sort in 15 minutes. After the first test, both groups were asked to change their learning media. The book-first group then studied using the game and game-first group studied using the textbook. They were given 30 minutes to study again the same materials. After that, both groups had to take the second test for 15 minutes. Using this kind of test enabled us to understand the significance of the artifact in improving the learning outcome. Wilcoxon and Mann-Whitney methods were used to test the significance. The results of this phase are discussed in section VI.

III. Related Works

Several approaches have been conducted to bring algorithms in a more concrete form that can be easily understood by learners. Visualization is one approach that is commonly taken and programming, an activity that translates the algorithms into working programs, has been taught using following approaches: traditional lectures and labs, robots, problem-based learning, cognitive apprenticeships, software visualization, microworlds, as well as programming environments and support tools [7] [8]. Visualization translates algorithm concepts into organized visual forms to depict the state of the process steps while the algorithms run and portray the dynamics of algorithm's processes over time. This approach reinforces learners to grasp the concepts in a more concrete way [9].

Along with the growth of technology, algorithm visualization is improved by its presentation and sophistication. In the beginning, algorithms are visualized using static visualization. This is what we usually found in almost every algorithm textbook where algorithm concepts, to the depth of their process steps, are explained by using static illustrations along with their codes, flowcharts, and explanation texts. Animated visualization then appears to deliver a more dynamic presentation to explain algorithm concepts, as performed by [10] for instance. Surprisingly opposite from the expectation, [9] and [11] found that the benefits of algorithm visualization is not significant and fails to be a standard in computer science learning. The dynamic visualization then moves forward into a more interactive visualization where learners can engagingly navigate through algorithm process visualization, reproduce the processes, move from one state to another state, inspect all the involved variables, tune the variable values, and observe the outcomes [12] [13]. This approach provides a more engaging, holistic, and deeper view for learners to understand algorithms. However, interactive visualization has weaknesses in motivating learners to engage continuously, frequently, and actively to learn algorithms, though it is an important aspect to have a greater impact on learning [12] [14]. This condition then elevates game as one approach, a.k.a. game-based learning, to learn or to teach algorithms. This approach provides students a new way of learning algorithms that is not only interactive but also engaging enough to keep them learn continuously.

A comprehensive study of applying game-based learning for learning algorithms has been conducted by [15]. They developed Algorithm Visualization using Serious Games (AVuSG) Framework that illustrates how computer games can be used in algorithm learning and visualization, including embedding learning theories and models to their framework. However, their work did not provide a deep comprehension on how game elements or mechanics should be integrated into algorithm learning. The gamification of algorithm learning was not discussed in depth.

Many applications intended to teach and learn algorithms are available in the market, e.g. Algorithms App in iTunes and Sorts in Play Store. However, most of them run as a learning tool and not as a serious game. There are some serious games with dedicated 3D environments that teach programming to their players, such as C-Sheep [16] and Prog&Play [17].

Although they can also be used for learning algorithms, their main concern and design are for interactive programming and not intended to teach specific algorithm concepts. Exception for Sortko, it incorporates game elements into its design to teach basic sorting algorithms. Unfortunately, its game elements are limited to only points as a motivational module [18]. Other game elements have not been incorporated extensively in algorithm learning applications, which deliver prospect to study the influence of other game elements in algorithm learning. Nevertheless, incorporating game elements into algorithm learning is not trivial; it requires a careful design. The development of algorithm visualization is summarized in TABLE I. The list does not mean to be extensive but intends to depict the trend of algorithm visualization, its types, and concerns.

The design concern of the gamification is in line with the question prompted by [19] that the 'how' of learning games design is essential to be addressed. Likewise, Perotta et al. proposed a research challenge that a more analytic approach needs to be developed on how game elements contribute to learning [20], including a better understanding on how games fit with desired outcomes and how games integrate into user's learning experience [21]. This condition also aligns to what Deterding et al. also prompted that much work needs to be done on gamification design, particularly on how to integrate existing design patterns and dynamics of games into non-game contexts [22]. Based on these statements, the design aspect is perceived as a crucial part of this research.

TABLE I. SUMMARY OF ALGORITHM VISUALIZATION DEVELOPMENT

Year	Type	Concern	Work
< 1980s	Static visualization	Presenting abstract concept	Traditional textbook
1980s-1990s	Animated/ dynamic visualization	Portraying dynamic part of abstract concept	[10] [9] [11]
1990s-2010s	Interactive Visualization	Exploration, engagement	[23] [13] [12]
> 2010s	Gameful Visualization	Motivation, engagement	[15] [18]

This research tries to integrate elements which come from three different domains—instructional design, algorithm visualization, and games—into an artifact that will be used for an algorithm learning purpose. The difference of elements' domains demands the artifact to be carefully designed. Subsequently, the underlying essential constructs that each instructional design, algorithm visualization, and gamification tries to deliver must be identified and understood first before they can act as a theoretical model to design the artifact.

IV. COGNITIVE SUPPORT, ENGAGEMENT, LEARNING INSTRUCTIONS, AND LEARNING OUTCOMES

The essential value of algorithm visualization is to deliver cognitive support. It means that the hidden or abstract parts of a concept are made concrete or visible, so learners are facilitated to access the concept. The visualization shows the process of algorithms to assist learners in comprehending how the algorithms work and their comparison. It delivers a visible illustration of the notion and procedures which are intrinsic in algorithms and depicts the dynamics of time-view progression [9]. In other words, learners can observe the relationship of causality and the structure of an algorithm and finally deduce what the algorithm is currently executing [10]. Visualization functions as an external memory and, therefore, lessens the cognitive complexity of programming activities [11]. The visualization brings the inner mechanism of algorithms more observable and facilitates its users to investigate the computation of the algorithms [23].

However, the capability of the visualization to portray abstract concepts will not automatically bring benefits if learners do not actively engage with the visualization [12]. To improve the engagement, learners' motivation has to be escalated. Hereinafter, the term of engagement will also comprise motivation since learners cannot engage frequently and completely if they do not have enough motivation to do the intended activities. To improve learners' engagement, one available approach is to use a game or game elements or gamification. Achieving engagement is one of several main use cases of gamification [4]. Through a literature study, games were identified as having a positive impact on knowledge acquisition, motivation, and engagement [20]. Some use cases of using games or game elements into the design of artifacts are as follow. A point system that usually exists in games was included in a mobile application of sorting algorithms as its motivation module [18]. Considering the theory of motivation, games were used to motivate learners to learn algorithms by providing encouraging feedback and delivering challenges through establishing well-defined goals with proper difficulties [15].

Both employment of algorithm visualization and gamification in learning have the same end goal, improving the learning outcome. While algorithm visualization attempts to deliver a cognitive support to improve learning outcomes, gamification seeks to leverage learners' engagement. Nevertheless, both approaches should not be applied solely separated from instructional design, since up to now, it has been a fundamental process and a mature discipline that deals with learning quality improvement before the early use of interactive visualization and gamification in learning. The aim of instructional design is directing learners to obtain certain capabilities, including the acquisition of knowledge or skills [24] [25]. To facilitate the acquisition, instructional design produces optimum learning instructions, comprising of instructional activities, materials, information resources, and evaluation, as an embodiment of instruction and learning principles, after a reflective and systematic process [24]. The 'optimum' word itself can be interpreted as effective, efficient, and appealing [24] [25]. A good instructional content will enhance learning outcomes [26].

In a learning context, it is better not to use visualization solely without designing the instruction first since visualization is an instruction aid, not the instruction itself [9]. The same rule also applies to gamification. A game or its elements are only a moderator which moderates learners' engagement towards learning [26]. Whereas instructional design is dealing with the overall strategy of how to optimize a learning process, algorithm visualization and gamification provide support to learners while they are interacting with learning content—

algorithm visualization delivers cognitive support and gamification leverages learners' engagement. The learning instructions drive learning processes to produce learning outcomes. Both cognitive support and engagement act as moderators, reinforcing learning instructions to leverage learning processes in order to improve learning outcomes. The relationships are summarized in Fig. 1.

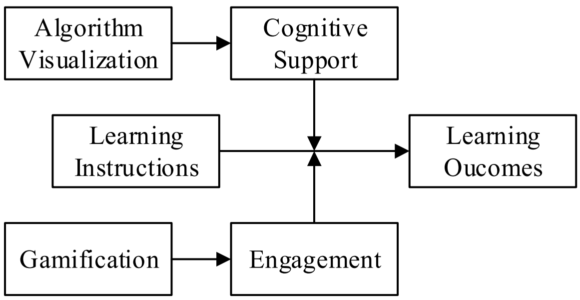


Fig. 1. A conceptual model of how algorithm visualization, gamification, and learning instructions interact to improve learning outcomes

V. Design Consideration

Optimum learning instructions, cognitive support, and engagement are three constructs that influence learning outcomes in this research's context.

To produce an artifact that improves learning outcomes, the artifact must embody the three constructs. To accomplish the embodiment, principles that originate from algorithm instructional design, algorithm visualization, and game domains must be taken into account during designing. The three domains are considered comparable to the three components in the Levee Patroller design in that algorithm instructional design, algorithm visualization, and game domains correspond to the meaning, reality, and play components of Levee Patroller design respectively [27]. Subsequently, the principles drive the design process of the artifact to deliver gamified algorithm learning, in this case, in the form of a game. Elements from the three domains also have to be incorporated into the artifact design to deliver the three constructs. This way of abstraction is depicted as a framework in Fig. 2. The conceptual framework is then employed to guide the design of our game to deliver gamified algorithm learning.

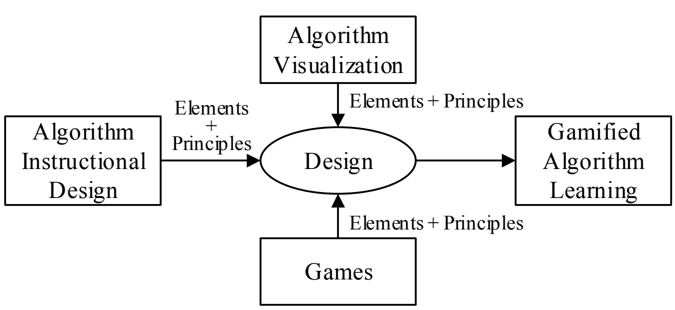


Fig. 2. A conceptual framework of how to design an artifact which delivers gamified algorithm learning

Algorithm Instructional Design. Instructional design concerns with learning strategies and arrangement, what topics are delivered first and what topics should be given later, to optimize knowledge or skills of acquisition of learners. From our investigation of top three algorithm textbooks [1] [28] [29], simple and basic topics are delivered first to support learners in developing early access and fundamental understanding of algorithms before moving to more advanced topics. Basic sorting algorithms—insertion sort, selection sort, and bubble sort—were selected as the topics delivered by the artifact since they are three of the most basic, simple, and frequent topics taught in algorithm textbooks and courses.

Considering the first 3 categories—remember, understand, and apply—of the revised Bloom's Taxonomy [30], the artifact needs to have a feature that support learners to remember, understand, and apply relevant knowledge regarding sorting algorithms. Consequently, the artifact is designed to have an introduction feature that includes a brief description, a relevant everyday example, brief operational explanation, and code example of each sorting algorithm. This introduction feature provides the factual knowledge, basic and overview knowledge to acquaint sorting algorithms [30]. Moreover, the game also presents a tutorial feature of how each sorting algorithm operates by providing a simulation—an animated visualization—accompanied with its textual explanation. The tutorial feature delivers procedural knowledge, the course of mechanisms performed by the sorting algorithms [9] [30], and addresses the understand category of the revised Bloom's Taxonomy. Both introduction and tutorial are presented as the first and second level consecutively on each algorithm.

After completing the introduction and tutorial, learners can move to the next levels where they can apply or practice their knowledge by sorting a set of unordered numbers represented by an array of rectangles through dragging and dropping using their fingers. This feature facilitates the development of the apply category of the revised Bloom's Taxonomy [30] as well as the construction of knowledge (constructivism) [31] through the interaction between learners' existing understanding and their real experiences (experiential learning) [32]. During the interaction, learners perform learning using the visual, spatial mechanism via the visual presentation of the unordered numbers in an array of rectangles, the kinesthetic mechanism through ordering the numbers using their fingers by dragging and dropping, the mathematical, logical mechanism through comparison to determine the order of the numbers, and by textual mechanism via reading the explanation and instruction of the algorithm. Different mechanisms of learning are expected to enhance learning [33].

The mechanisms are also supported with feedback, an important aspect of learning, to let learners know whether their decision in applying their knowledge to run the algorithms is correct or not. Feedback is one of Gagne's nine events of instructions that brings reinforcement in the internal learning process [34]. The feedback indirectly supports learners to check, verify, and validate whether their current understanding of the respective algorithm is correct or not and thus allowing them to adjust their current understanding about the algorithms. The feedback facilitates the fifth category of the revised Bloom's Taxonomy, the evaluate category [30].

The principle of flow also needs to be considered [35]. Levels given have to be designed carefully, so they do not provide learners with over difficult or too easy materials and challenges. As the capability of learners grows, the materials and challenges have to be adjusted. The other option is learners are provided with a freedom to choose materials and challenges that match their capabilities. The difficulty of the artifact is designed to gradually increase as the level goes up. The increasing difficulty is the result of a combination between the quantity of the unordered numbers, the number of digits of a number, and the order of the numbers.

Algorithm Visualization. The abstract concept of sorting algorithm has to be visualized to make it more concrete. A set of unordered numbers that will be sorted is commonly delivered in the form of an array. The array and the value of each array element have to be visualized in order for the abstract concept of the array and numbers to be perceived visually by learners, reinforcing their mental image of entities. To realize this, the array is depicted as an array of holding rectangles and the numbers as an array of rectangles with values. This visualization enables learners to observe the states of the array and the numbers, including the temporary storage that is visualized also as a holding rectangle to store a number for swap purpose, during the execution of the sorting algorithms.

The visualization has to provide learners with interactivity in order for them to interact—simulate, animate, and explore—with the abstract concept of sorting algorithms. Learners can touch and drag, and then the drop rectangles allowing them to simulate the sorting algorithms on their own step by step. Additionally, animation accompanied with textual information, is also used in the tutorial feature to show learners how the respective algorithm works. The interactivity given provides learners a freedom to explore what is the next most appropriate step to operate the intended sorting algorithms.

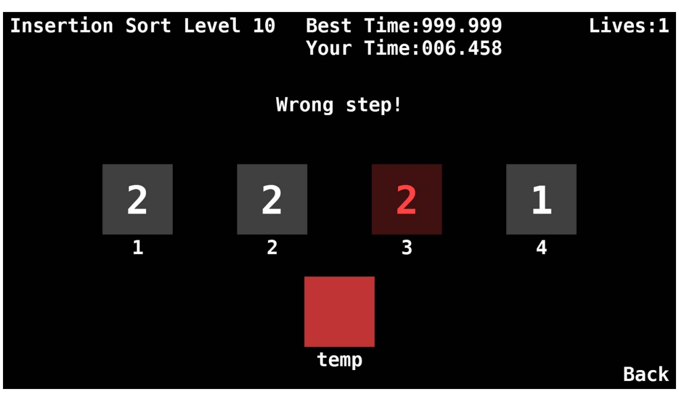


Fig. 3. Sort Attack, the visualization and gamification of sorting algorithm learning

The visualization also has to provide feedback to learners in order for learners to know the outcomes or their actions. The artifact provides feedback by displaying blinking green color on the correct rectangle when dropped to the empty holding rectangle, accompanied with a 'positive' sound effect. Oppositely, when learners drop the incorrect rectangle to the empty holding rectangle, the blinking color displayed will be red, together with a 'negative' sound effect. Using this mechanism, learners are notified if their steps are correct or not. One screenshot of the artifact is displayed in Fig. 3.

Games. To deliver a gameful experience, game principles and elements need to be incorporated into the artifact. Nine characteristics—player, environment, rule, challenge, interaction, goal, emotional experience, quantifiable outcome, and negotiable consequences—of games mentioned in [5] are used as the basis to embed and explain the gamefulness that presents in the artifact.

- *Players*. Learners are the players themselves since they are the actors that interact with the valued rectangles and sort them to complete a level.
- *Environment*. The artifact provides a dedicated environment and its constituent components where and to which learners can perform interaction. Learners interact with the valued rectangles by sorting them and some indicator elements exist in the artifact providing learners with feedback about the current conditions.
- *Rules*. The mechanisms of the sorting algorithms are the rules which learners must understand, follow, and perform to complete the game. Lives, where learners are given three excuses to make mistakes, can also be perceived as rules.
- *Challenges*. The rules mentioned previously can also be perceived as challenges since the rules determine the next correct step to execute the intended algorithm while for learners, if they have not understood the algorithm, they will not know exactly which rectangle should be dragged and dropped next to the empty holding rectangle.

  Challenges also are presented in the form of levels. When learners complete a level, they are allowed to move to a more advanced level with a higher difficulty. The difficulty of the levels is defined by the number of unordered numbers, the number of digits of the unordered numbers, and the order of the unordered numbers. Increasing the number of unordered numbers to be sorted decreases the chance of a valued rectangle to be a correct rectangle that should be moved next to the empty holding rectangle. The number of digits or the unordered numbers increase the mental load when perceiving the numbers and when deciding which number is smaller or bigger than other numbers. The high randomness of the numbers' order makes the sorting activities become more complex and consume more time.

  Time record and lives can also provide challenges. Time record challenges learners to complete the sorting as fast as they can and break the record of previous best time record. Lives component limits learners' chances of making errors during executing the sorting algorithms.
- *Interaction*. To deliver an engaging interaction in the artifact, game mechanics need to be presented. Game mechanics are the main processes—consisting of algorithms, data, and representations—that move a game forward leading player's actions and, therefore, might produce player engagement [36] [37] [38]. The

artifact's game mechanics is learners are required to sort a set of unordered numbers, represented as an array of rectangles, to become ordered. However, they have to sort them using the rules of sorting algorithms they have selected. If they choose selection sort, they need to follow the steps of selection sort. If they can sort the unordered numbers without breaking the rules, one round of the game is completed and they can move to the next level.

The game also needs to provide feedback which enables learners to keep track of the progression when the game is moving forward. Learners are also notified of the results of every action they made to the dedicated environment. Blinking red or green color, accompanied by 'negative' or 'positive' sound effect, when valued rectangle drops into an empty holding rectangle, is an indicator whether the dropped valued rectangle is the correct one or not.

- *Goals*. The goal of playing using the artifact is completing the sorting of a set of unordered numbers using a certain sorting algorithm without making mistakes.
- *Emotional Experiences*. Challenges delivered should raise learners' tension during sorting the unordered numbers. Failure in completing the sorting should wake the curiosity of learners to try again and success should bring a feeling of fun, satisfaction, and curiosity of challenging the next level.
- *Quantifiable Outcomes*. In the artifact, quantifiable outcomes such as rewards, i.e. badges and points, are not explicitly given. Nevertheless, the success of completing a level and move to more advanced levels, beating the previous time record, completing a sorting without a mistake, and master the algorithms are the implicit quantifiable outcomes that can be achieved by learners.
- *Negotiable Consequences*. Learners that fail at one level can retry many times until they can complete the level. The unlimited retrial provides learners negotiable consequences to explore and then understand the sorting algorithms. Additionally, lives can also be perceived as a chance to explore the algorithms in a limited manner. Using trial and error method, learners can test their understanding about the next correct step but limited to a certain number of mistakes.

Learning from the design process of the artifact, it is discovered that there are areas where the three domains—instructional design, algorithm visualization, and games—intersect. The intersections deliver opportunity for the three domains to be more easily integrated into the design of the artifact. For instance, principle of simple, basic topic first and flow concept in the instructional design can be integrated with the level element of games. Another example is the action aspect—actions are needed to move forward—on the procedural knowledge of sorting algorithms, the interactive part of algorithm visualization and the game mechanics of games provide an area where the three domains can be integrated. For example, learners play the artifact by applying the procedural knowledge of the respective sorting algorithms in an interactive manner, i.e. by dragging and dropping.

VI. Significance of The Artifact

This research has measured the significance of the improvement of learners' learning outcomes after using the artifact through executing the research method previously described in Section II. The learning outcome itself is the mastery of the procedural knowledge of the respective sorting algorithms and exhibited by their capability in solving given problems. It was found on average that learners who had used the artifact experienced significant improvement in their learning outcome. The results are exhibited in TABLE II. , TABLE III. , and Fig. 4. Fig. 4 summarizes the average scores in TABLE II. in the form of a bar chart.

TABLE II. The Average Scores between Book-First Group and Game-First Group after The First Test and The Second Test

Group	**Book-First**			**Game-First**			**Diff.**	
Type	***1st Test***	***2nd Test***	***Diff.***	***1st Test***	***2nd Test***	***Diff.***	***1st Test***	***2nd Test***
Bubble	8.71	10.00	1.29	8.93	10.00	1.07	0.22	0.00
Insertion	7.31	10.00	2.69	9.64	10.00	0.36	2.33	0.00
Selection	7.42	10.00	2.58	10.00	10.00	0.00	2.58	0.00
Overall	7.81	10.00	2.19	9.52	10.00	0.48	1.71	0.00

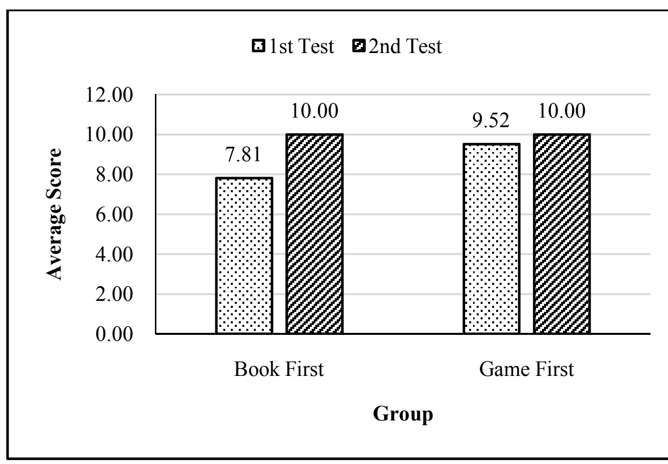


Fig. 4. The average scores between book-first group and game-first group after the first test and the second test

Average scores for both book-first group and game-first group generally experienced improvement based on the first test and the second test. Nevertheless, book-first group's average score's improvement was higher than the improvement of the game-first group's average score, the improvement from the first test result to the second test result. Both book-first and game-first groups' overall average scores increased by 2.19 and 0.48 points with bubble sort's, insertion sort's, and selection sort's average scores for book-first group went up by 1.29, 2.69, and 2.58 points respectively, while for the game-first group, the bubble sort's and insertion sort's average scores inclined by 1.07 and 0.36 points. Interestingly on selection sort, game-first group after using the artifact could immediately solve the given problems perfectly, as indicated by the 10.00 points and the attainment was consistently sustained when they passed the second test flawlessly with the same score after learning again from a textbook.

The game-first group produced better learning outcomes than the book-first group in the first test. Averagely, the game-first group scored 9.52 points, 1.71 points higher than 7.81 points scored by the book-first group. In the first test, book-first group scored 8.93, 9.64, and 10.00 points in bubble sort, insertion sort, and selection sort correspondingly, whereas book-first group scored 8.71, 7.31, and 7.42 points for bubble sort, insertion sort, and selection sort respectively. The results indicate the artifact could more immediately support learners in achieving higher learning outcomes than by learning only using a textbook.

In the second test, the game-first group and the book-first group produced the same learning outcome, as indicated by the equal test scores—both groups in all types of sorting algorithms examined reached the perfect score, 10.00 points. The result demonstrates that a combination between the artifact and textbook ultimately supports learners to master sorting algorithms.

To support all the findings, Mann-Whitney test and Wilcoxon test were employed to test whether the scores between book-first group and game-first group for each test and between the first test and the second test for each group are significantly different. The results can be found in TABLE III. From the table, it was found that the overall average scores of the first test and the second test of the book-first group were significantly different, including on every type of the sorting algorithms, since the p value of their every Wilcoxon test was below 0.05. In contrast, there was no significant difference between the average of the first test scores and the average of the second test scores of the game-first group, as none of the p values of the overall average and the average of each type of the sorting algorithms was below 0.05.

TABLE III. THE SIGNIFICANCE OF SCORE DIFFERENCES BETWEEN THE 1ST TEST RESULT AND THE 2ND TEST RESULTS FOR EACH GROUP AND BETWEEN BOOK-FIRST GROUP AND GAME-FIRST GROUP FOR EACH TEST

Type	1st Test vs 2nd Test Book-First Group [a]	1st Test vs 2nd Test Game-First Group [a]	Book-First vs Game-First 1st Test Result [b]	Book-First vs Game-First 2nd Test Result [b]
Bubble	0.024*	0.083	0.423	1.000
Insertion	0.004*	0.317	0.006*	1.000
Selection	0.005*	1.000	0.004*	1.000
Overall	0.001*	0.102	0.004*	1.000

* Significantly different, $p < 0.05$, [a] using Wilcoxon test, [b] using Mann-Whitney test.

In the first test, the average scores of the book-first group and the game-first group were significantly different ($p < 0.05$ using Mann-Whitney test), except the for the bubble sort's average scores that had p higher than 0.05. Exception for bubble sort, the scores between the book-first group and the game-first group were not significantly different. The reason that can be given is bubble sort is easy to remember and to code [39]. Therefore, by only using a textbook, the book-first group could grasp the concept of bubble sort effortlessly, reaching scores that were likely indifferent from the scores that the game-first group had. The second test shows us that the average scores of the book-first group and the game-first group were not significantly different. They all had the equal value of p that was 1.000. The result exhibits that a combination of the artifact and the textbook can produce the maximum learning outcome.

VII. CONCLUSIONS

Algorithm visualization is one of the common approaches that has been taken so far for concreting the abstract concept of algorithms. It has been growing from a static visualization, dynamic visualization, interactive visualization, and recently to a gameful visualization. In conclusion, we argue that in order to improve learning outcome, algorithm learning instruction can be supported by algorithm visualization and gamification. Algorithm visualization works by providing cognitive support while gamification functions through developing engagement. Both are moderators for algorithm learning instructions that focus more on preparing the optimal strategy to learn algorithms. Principles and elements that come from the instructional design, algorithm visualization, and games have been taken into account and integrated into design to yield an artifact that is capable to address the main concern of the three domains. Finding the similarities or the intersections of the three domains can simplify the integration of the principles and elements of the three domains into the design of the intended artifact. From this research's significance measurement, learners who have used the artifact in learning sorting algorithms experience significant improvement on their learning outcome, specifically their procedural knowledge, indicated by their competence in solving sorting algorithm problems. Learners who have used the artifact as their first media to learn sorting algorithms significantly perform better than learners who use the textbook. Combining both of them has brought out the most to ultimately support the learners in becoming proficient at sorting algorithms.

Nevertheless, the result of the evaluation presented here does not indicate whether the artifact, as the representation of gameful visualization, can outperform dynamic or interactive visualization and therefore it remains an open question that will be addressed in the next future works. Additionally, improving the aesthetics of the artifact and investigating its effect to engagement and learning is an interesting activity to work on since existing works found that aesthetic influence learner's engagement [16]. Other future works that are challenging to perform are designing gamification and visualization of algorithm learning that support the development of conceptual and metacognitive knowledge of algorithms. The knowledge addressed in this research is limited to factual and procedural knowledge. Conceptual and metacognitive knowledge has not been addressed yet. Conceptual knowledge refers to an understanding of the interrelationships between the components of an algorithm within a broader structure that allow them to work together or simply a comprehension of the algorithm's structures and implications [9] [30]. Metacognitive knowledge is one's comprehension of his/her own knowledge and cognition and thus can be used to develop a strategy to acquire new knowledge [30]. Addressing both domains of knowledge is difficult since the development is using extensive inference and implication and not based on learning factual data or prescriptive procedures. From the gamification design viewpoint, selecting appropriate game elements and designing suitable game mechanics are not trivial tasks and we presume it requires a lot of creativity more than just knowledge about the game domain.

ACKNOWLEDGEMENT

We thank the classes of '13 and '14 of Informatics Program at KALBIS Institute and to Stephanie Mew, Wulan, and Christ. This research was funded by KALBIS Institute and Ministry of Education and Culture of the Republic of Indonesia. The artifact can be found at http://goo.gl/2M5TIV.

REFERENCES

[1] R. Sedgewick and K. Wayne, Algorithms, Pearson Education, Inc., 2011.

[2] R. Duval, "A cognitive analysis of problems of comprehension in a learning of mathematics," *Educational Studies in Mahematics,* vol. 61, pp. 103-131, 2006.

[3] D. Michael and S. Chen, Serious games: Games that educate, train, and inform, Thomson Course Technology PTR, 2006.

[4] S. Deterding, D. Dixon, R. Khaled and L. Nacke, "From game design elements to gamefulness: Defining gamification," *MindTrek,* 2011.

[5] A. R. Yohannis, Y. D. Prabowo and A. Waworuntu, "Defining Gamification: From lexical meaning and process viewpoint towards a gameful reality," in *International Conference on Information Technology System and Innovation*, Bali, 2014.

[6] K. M. Kapp, The gamification of learning and instruction: Game-based methods and trategies for training and education, San Francisco: John Wiley & Sons, 2012.

[7] E. Coestelloe, Teaching programming the state of the art, The Center for Research in IT in Education. Dept. of Computer Science Education. Dublin: Trinity College, 2004.

[8] A. Pears, S. Seidman, M. Lauri, M. Linda, A. Elizabeth, B. Jens, D. Marie and P. James, "A survey of literature on the teaching of introductory programming," in *ACM SIGCSE Bulletin*, vol. 39, ACM, 2007, pp. 204-223.

[9] M. D. Byrne, R. Catrambone and J. T. Statsko, "Do algorithm animations aid learning?," *Tech. Rep. GIT-GVU-96-18,* 1996.

[10] R. Baecker, "Sorting out sorting: A case study of software visualization for teaching computer science," *Software Visualization: Programming as a Multimedia Experience,* vol. 1, pp. 369-381, 1998.

[11] C. Hundhausen, S. Douglas and J. Stasko, "A meta-study of algorithm visualization effectiveness," *Visual Languages and Computing,* vol. 13, no. 3, pp. 259-290, 2002.

[12] T. L. Naps, G. Rößling, W. Dann, V. Almstrum, R. Fleischer, C. Hundhausen and A. Korhonen, "Exploring the role of visualization and engagement in computer science education," *ACM SIGCSE Bulletin,* vol. 35, no. 2, pp. 131-152, 2002.

[13] T. Lauer, "When does Algorithm Visualization Improve Algorithm Learning?–Reviewing and Refining an Evaluation Framework.," in *Proceedings of Informatics Education Europe III*, 2008.

[14] T. L. Naps, "Jhave: Supporting algorithm visualization," *IEEE Computer Graphics and Applications,* vol. 25, no. 5, pp. 49-55, 2005.

[15] S. Shabanah and J. X. Chen, "Simplifying algorithm learning using serious games," in *Proceedings of the 14th Western Canadian Conference on Computing Education*, 2009.

[16] E. F. Anderson and L. McLoughlin, "Do Robots Dream of Virtual Sheep: Rediscovering the" Karel the Robot" Paradigm for the" Plug&Play Generation".," in *Fourth Game Design and Technology Workshop and Conference GDTW2006*, 2006.

[17] M. Muratet, P. Torguet, F. Viallet and J. P. Jessel, "Experimental feedback on Prog&Play, a serious game for programming practice," *Computer Graphics Forum,* vol. 30, no. 1, pp. 61-73, 2011.

[18] I. Boticki, A. Barisic, S. Martin and N. Drljevic, "Teaching and learning computer science sorting algorithms with mobile devices: A case study," *Computer Applications in Engineering Education,* vol. 21, no. S1, pp. E41-E50, 2013.

[19] M. Prensky, "Computer games and learning: Digital game-based learning," *Handbook of computer game studies,* vol. 18, pp. 97-122, 2005.

[20] C. Perotta, G. Featherstone, H. Aston and E. Houghton, Game-based Learning: Latest Evidence and Future Directions (NFER Research Programme: Innovation in Education), Slough: NFER, 2012.

[21] T. M. Conolly, E. A. Boyle, E. MacArthur, T. Hainey and J. M. Boyle, "A systematic literature review of empirical evidence on computer games and serious games," *Computers and Education,* vol. 59, pp. 661-686, 2012.

[22] S. Deterding, S. L. Björk, L. E. Nacke, D. Dixon and E. Lawley, "Designing gamification: creating gameful and playful experiences," in *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, 2013.

[23] W. L. Hibbard, B. E. Paul, D. A. Santek, C. R. Dyer, A. L. Battaiola and M. F. Voidrot-Martinez, "Interactive visualization of earth and space science computations," *Computer,* vol. 27, no. 7, pp. 65-72, 1994.

[24] P. L. Smith and T. J. Ragan, Instructional Desgin, New York: Wiley, 1999.

[25] D. M. Merrill, L. Drake, M. J. Lacy, J. Pratt and ID2 Research Group, "Reclaiming instructional design," *Educational Technology,* vol. 36, no. 5, pp. 5-7, 1996.

[26] R. N. Landers, "Developing a Theory of Gamified Learning: Linking Serious Games ad Gamification of Learning," *Simulation and Gaming,* vol. 1046878114563660, pp. 1-17, 2015.

[27] C. Harteveld, R. Guimarães, I. S. Mayer and R. Bidarra, "Balancing Play, Meaning and Reality: Meaning and Reality: of LEVEE PATROLLER," *Simulation & Gaming,* vol. 41, no. 3, pp. 316-340, 2010.

[28] T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein, Introduction to Algorithms, 3rd ed., MIT Press, 2009.

[29] S. S. Skiena, The algorithm design manual, Springer, 2008.

[30] D. R. Krathwohl, "A revision of Bloom's taxonomy: An overview," *Theory into practice,* vol. 41, no. 4, pp. 212-218, 2002.

[31] D. J. Cunningham and T. M. Duffy, "Constructivism: Implications for the design and delivery of instruction," in *Handbook of research for educational communications and technology*, 1996, pp. 170-198.

[32] A. Y. Kolb and D. A. Kolb, "Learning styles and learning spaces: Enhancing experiential learning in higher education," *Academy of management learning & education,* vol. 4, no. 2, pp. 193-212, 2005.

[33] F. Coffield, D. Moseley, E. Hall and K. Ecclestone, "Learning styles and pedagogy in post-16 learning: A systematic and critical review," Learning and Skills Research Centre, Trowbridge, Wiltshire, 2004.

[34] M. P. Driscoll, Psychology of learning for instruction, 2nd ed., Allyn and Bacon, 2000.

[35] M. Csikszentmihalyi, Finding flow: The psychology of engagement with everyday life, Basic Books, 1997.

[36] K. Werbach and D. Hunter, For the win: How game thinking can revolutionize your business, Wharton Digital Press, 2012.

[37] G. Zichermann and C. Cunningham, Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps, O'Reilly Media, Inc., 2011.

[38] R. Hunicke, M. LeBlanc and R. Zubek, "MDA: A formal approach to game design and game research," in *Game Developers Conference, San Jose 2001-2004*, 2004.

[39] W. A. Martin, "Sorting," *ACM Computing Surveys (CSUR),* vol. 3, no. 4, pp. 147-174, 1971.

Your answer will make an impression

Using Quiz Game Mechanics for the Collection of Visitor Data at an Exhibition

María Guadalupe Alvarez Díaz
Ulf Wilhelmsson
Mikael Lebram
Marcus Toftedahl
School of Informatics
University of Skövde
Sweden

***Abstract* — this paper presents the initial results from a project that aimed to collect visitor data at a traveling exhibition starting at the Regional Museum in Kristianstad, Sweden during 2014-2016. The project was intended also to contribute to the creation of an atmosphere "About time", which was the subject of the exhibit. We built a system that was integrated as an interactable part of the exhibition by using elements of quiz game mechanics in combination with elements of data based tracking applications and elements of visual art installations. The data provides statistics which are used to visualize the current status of the visitors' attitude toward specific questions about time, imprinting the visitors themselves an integral part of the exhibition. Visitors build a visual Game Ego when answering questions and at the same time provided statistical data that can be monitored and extracted from the system. The results show that we succeeded to some degree but more can be done towards incorporating game design elements to engage the user, such as feedback and challenge.**

Keywords— quiz games; survey techniques; art installations; digital applications in museum environments; tracking applications

I. Introduction

In January 2012 we were contacted by a representative from the regional museum in Kristianstad, Anna Hadders, and asked if our researchers and students would be interested to engage in a project concerning time. The project consisted of a long term study relative to the manner people experience time, through an interactive approach with teenagers: workshops, seminars, and the production of a traveling exhibition. Our participation in the exhibit consisted of the development of games and other artifacts that would put a light on humans' perception of and relation to time.

II. Inspiration in Time-clock Machines

An idea we were asked to address was the use of a time clock as the emblematic time shift between buying goods and buying people's time to produce goods, by designing a system that allowed visitors to check in at the exhibition with a time clock, and to check out when leaving the exhibition. We found this to be quite interesting but also quite possible to do something more of: a system integrating some game mechanics powered with survey technics and visualizations, referring to abstract film animations in modern art. [1] [2]

III. The Quiz-Tracking Console

The main question for us to address initially was how we would encourage visitors to interact with the time clock console and provide these data for us. Inspired by Higinbotham's Tennis for Two game, [3] which made use of people's playfulness, we decided to build the interaction based on a quiz. Having been working on slightly similar projects, we wanted to give visitors the possibility to modify or personalize the image represented in the screen while answering questions.

The questions shaped in a combination of riddle and a questionnaire would tailor the exchange. The formulation of the questions would trigger some thoughts and reactions as much as answers would provide input about and along the experience.

We chose to create an application able to read bar-codes or QR-codes from a business card; a time-credit card that would relate to the novel Momo [4], which was an inspiration for the exhibition in the first place, there people are lured to put their time into a time savings account; a punch-checking point where the visitor is instigated to an action that physically and mentally would relate it to the fact that her personal time is being counted.

We initiated with the alternative to develop a short, simple track-report database to store visitor's data with the possibility to represent a graphical path of each visitor in a ground plan, displayable on a regular flat TV-screen. The graphical display could become an interesting visualization when multiple visitor paths were added to the image, contributing to a spontaneous collective expression [5]. The interface would be an artifact with physical elements, e.g., the card that would be personal, something to keep in the pocket while roaming around; a number of stations would be placed at several positions in the exhibition area. Those would be boxes containing a card slot, a screen and an arcade machine style button. Visitors would check-in their cards in the box. The card-id, the station-id and the time would be sent via Wi-Fi to a server. The server would generate graphical statistics about the visitors. The images of the statistics could be displayed on a large screen and streamlined to a web page.

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

A. Developing a prototype

The basic requirement was to construct a robust and easy replaceable artifact; identically constructed and interchangeable. The number of consoles didn't need to be limited, it could work equally well with any number. The size and style of fonts should be readable to any kind of public; a bigger size of the console screen would facilitate the communication with the user. These requirements motivated the use of a computer with a monitor screen in combination with a bar-code reader instead of a smart phone or other type of mobile terminal. Having a display, optical reader, and Wi-Fi in a single unit was one of the main advantages of using a smart phone in the initial design. However, the requirement of an external display led to a new design, based on a single board computer (Raspberry Pi) running Linux. The cost of the hardware would then decrease significantly per unit, where the screen represents more than half of the cost. The Raspberry Pi facilitates HDMI output for the display and generic IO pins for the external buttons. Instead of using a camera and QR code, a standard bar code reader was modified and connected to the computer via USB. The bar code reader emulates a keyboard, which means that it sends a series of keystrokes corresponding to the code that is read. A website was made available to enter the card-code after the visit to see the answers chosen and its personal graphics. The database containing answers to questions from each individual generates statistical information about participation, counting information from the questions and track user movements.

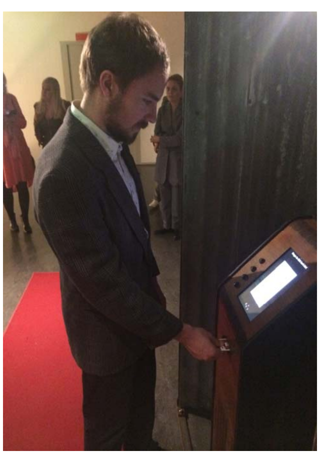


Figure 1. Check-in at the exhibit's entrance

B. Designing an interface

The alternatives explored in designing the exchange between the application and the card-user (the visitor) were a combination of texts in the form of instructions, dialogue, questionnaire or riddles and a graphic representation displayed on the console screen. By creating a specific query sequence in each individual record via the card-id, we supported the freedom of moving around the hall, checking-in at the various consoles and answering as many questions as they choose at any moment at any console.

C. The big screen - A collective creation

A visual component of the experience is a TV display with large dimensions, with a graphic visualization of Game Egos [6] participating in the quiz, forming polychromatic compositions based on answer statistics. The choices visitors make affect the form, the size, the color, the shape and the velocity of the movement of the objects on the big screen. More interaction with the user can be developed on this basis in future phases of this project following the tradition of creative experiences with digital forms. [7]

Libraries were built in the application to contribute with the normative criteria, containing catalogs for icons, questions and features, such as colors and scales for sizes, movement and distance. The icons were chosen to manifest the Game Ego of each individual; the uses of small icons such as emoticons are common in chats to express everyday actions and feelings in simple standard symbols. The normative approach contributes to create rules in the application as how the image would get constructed.

An editor was also incorporated to input the questions and to configure the graphics. This interface allows the user to pre-define the relation between the questions (input) the answers (selections) and its manifestation (graphics).

IV. Testing and Production

There was an opportunity to test the prototype in the environment of the Scandinavian Game Developers Conference organized in June 2014 by our university.

We observed that users had a smooth interface interaction by using the cards, checking-in and moving ahead across the questions; except that users frequently tended to answer by touching the screen instead of pressing the buttons. In most cases participants used one console only.

The big screen was intriguing at first, until viewers figured that patterns and shapes repeated its movements, hence it derived the idea of expanding the configuration of images.

Statistical results of this test showed that 773 answers were received from over a hundred participants, which means an average of 6, 5 answers per user. It also presents that 26% of all users answering the first question did not continue to the second question.

A. Production

By entering the production phase with the exhibition's art designers Eric Langert and Per Petersson the physical console adopted a shape inspired on the steam punk theme of the exhibition. It also required adjustments to fulfill visitor's impairments. The result was a wooden old-fashion check-in terminal in the form of an old radio with black buttons aligned vertically; also a smaller rectangular screen replaced the big computer display used in the prototype of the console.

B. Features and Visualizations

The overall system works similarly to its prototype in terms of data acquisition, database as well as data visualization. Differences with the prototype strive on that the Game Ego built shows up momentously at the first check-in at the console screen, providing the user with a direct reference to how choices can influence the Game Ego. At the bottom of the screen the user see a tiny reference to the living Game Egos distributed in the other consoles at the exhibit hall. The user could check-in only once at the console situated in the entrance hall, allowing the user to continue to other stations; it also helped to avoid long queues when larger groups were visiting.

The application is web browser based; every time the visitor checks-in generates a transaction modifying its register in the database in real time. The big screen visualization is also web browser based. It displays the visitors' Game Egos. It is possible to customize the image through the user interface, making easy to adjust the icons catalog, the chromatic parameters and the texts providing unlimited usability options.

Several modifications were necessary to adjust the big screen image to the exhibition environment, such as the background and the shifting between multiple visualization models. Two big screens and four consoles were displayed along the exhibition hall. TV-screens were also framed and re-contextualized to fit the museum installation.

V. Initial Results From the Exhibition in Kristianstad

It was observed during the exhibit that users, maybe captured by curiosity about the console and the card, were willing to initiate in a quiz format with the first question. Statistics show that 12% of the users didn't continue to the second question.

The data collected is sorted and reported in tables, making the application a very rich source of information on mood and roaming of visitors in the exhibit. The statistics shows that Sunday and Wednesday were days when visitors showed more willingness to answer than other days of the week, being Monday the day with a drastic drop. It also showed an important drop after the fifth question, we believe that this single question, in which the visitor gets feedback, was interpreted as the end of the play.

A. The Aesthetic Component

The big screen produced the effect of a collective expression built out of the visitor answers [7]. In the following image each Game Ego makes satellite movements leaving a track created by brushstrokes. The Game Ego movements

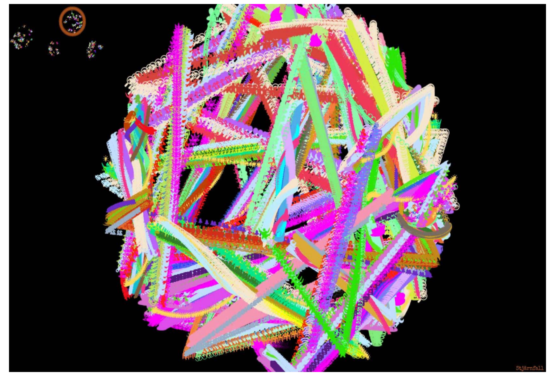


Figure 2. Example of a screen composition

evolved into switchable scenes forming compositions about the mood at various locations in the exhibit hall.

"Falling star" is the title of the composition above with a small mapping of other consoles in the hall in the upper left side.

The image below is a different composition named "Upline". When this image is in movement it can be seen each

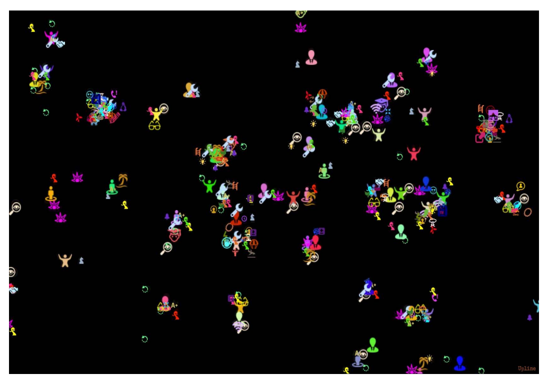


Figure 3. Visualizing Game Ego's manifestations

Game Ego built of multiple icons; each icon correspond to one answer. The assembling of icons rotates around the Game Ego, and moreover, they can combine movements with other Game Egos as part of the variations established in the parameters for each composition.

VI. Conclusions

With this interactable application participants subscribe themselves in the creative collective artwork. Alternative roles of audiences and creators speak of the need for appropriate ways to incorporate motivations, interests and perspectives of the public in the museum proposals to accommodate their

Back to Contents

participation. While there is an open disposition to new experiences the public is not necessary conscious of the possibility to live it collectively.

It is possible to know the flow of visitors to the exhibit hall in detail. The system reports data collected from each visitor registered. However, to contribute to a valid and useful data, is important that the access entry to the exhibit hall is designed for a fluent individual check-in at the first console - and to subsequent consoles - allowing a personal use of the cads.

We interpreted the results of a period of one week in the museum as evidence of the mood of the public visiting the exhibition.

The application and its consoles incorporate elements from quiz mechanics, reflections on the subject of the exhibition, elements of art installations and survey questions. It is a process of creation of expressions, adding to the living state of the environment. The application based its features on the collection and sorting of data of living events producing chromatic combinations of specific moments experienced in collectivity during the exhibit.

VII. Future Work

Approaching to public in the museum with these consoles is a starting point to achieve dynamic interventions on the content of an exhibition from any participant. It needs adjustments in the incorporation of its parts to reach a balance in the experience: as a game, as an artistic installation and as a survey tool.

It is obvious that our installation has some flaws and that the number of answers drop off steeply after just a couple of questions. It needs to be more of a thought through quiz game design in its essence and not only on its surface. As a quiz game, it does not really work as well as we hoped maybe for the lack of feedback given to visitors. We also think that more playful questions will result in better response rates.

The awareness of users to intervene the visual compositions may rise by increasing the interactivity with the screen console, facilitating feedback to visitors. Existing reporting capabilities allow statistical information and the Wi-Fi connection supports real-time counting. Those are features to consider in future design.

A following step in collective compositions is the incorporation of sound and music to the installation, especially as the console cabinet for this exhibition already has the shape and the space to incorporate speakers.

It would be interesting a deeper exploration on the aesthetic dimension of this project. The consoles, the screen and the application can be seen as components of an art installation, expressing and sharing experiences between participants.

The application seems suitable for development in areas of marketing survey; database, application capabilities and interface can be customized for diverse events involving many participants. The external shape of the end terminals can be adapted to the particularities of the event, taking the form of game consoles, information terminals or some other external appearance.

References

[1] V. Eggeling, ”Symphonie Diagonale,” 1923. [Online]. Available: https://vimeo.com/42401347. [Visited 18 June 2015].

[2] J. Whitney, ”Catalog,” 1961. [Online]. Available: https://www.youtube.com/watch?v=TbV7loKp69s. [Visited 18 June 2015].

[3] W. Higinbotham, Tennis for Two, Analog computer/oscilloscope game, USA: Brookhaven National Laboratory, 1958.

[4] M. Ende, Momo, Stuttgart: K. Thienemanns Verlag, 1973.

[5] D. Cox, ”Metaphoric Mappings: The Art of Visualization,” in Aesthetic Computing, Leonardo red., Cambridge, Massachusetts; London, The MIT Press, 2006.

[6] U. Wilhelmsson, "Game Ego Presence in Video and Computer Games," in *Extended Experiences*, Rovaniemi, Lapland University Press, 2008, pp. 58-72..

[7] C. Paul, ”Databases, Data visualization and mapping,” in Digital Art, London, Thames & Hudson, Ltd, 2003, pp. 174-189.

Back to Contents

Enhancing Immersion with Contextualized Scenarios

Role-playing in prehospital care training

Per Backlund, Henrik Engström,
Mikael Johannesson, Mikael Lebram
University of Skövde
Skövde, Sweden

Magnus Andersson Hagiwara,
Hanna Maurin Söderholm
University of Borås,
Borås, Sweden

***Abstract*—This paper discusses the method used in a field experiment with 12 paramedic teams (n=24) exploring how they perceive a novel training approach. Much simulation-based training in prehospital care is decontextualized, meaning that medical care is trained without taking other characteristics of prehospital care into account. In this paper we suggest how a richer setting (contextualization), which includes more of the complicating aspects of prehospital care, can be introduced and evaluated in prehospital training.**

Keywords—simulation; simulation-based training, scenario-based training; role-playing; prehospital care

I. Introduction

Serious games, in particular role playing games, have a long tradition of being employed in many areas, with military training as one prominent example. Simulation and simulator-based training are also used in other areas, e.g. for training of airline and fighter pilots, and in the field of medical training [1]. Due to the complexity of the prehospital care process, current training approaches are not sufficient for effective training. At present, different aspects are typically trained in isolation, e.g. medical skills using patient simulators [2] or team training focusing on coordination and cooperation without performance of realistic medical tasks.

In this short paper, we report on work in progress which aims to develop concepts for prehospital training taking the entire process and its complexity into account. One important aim of this project is to study the impact that different levels of fidelity and contextualization/ decontextualization have on the level of immersion.

II. Role playing in rich contexts

A. Simulation, Fidelity and Immersion

Live role-playing is commonly used for many types of training. According to Lankoski and Järvälä [3] live role-playing is about shared fiction created by the use of the participants' imagination. Training with live role-playing scenarios is an example of a training activity in which some aspects of reality are abstracted away for the benefit of practicality and safety [4]. In this sense, military maneuvers and field exercises are live role-playing games with explicit learning goals.

Wyatt, Archer and Fallows [5] and Alexander et al. [6] refer to three types of fidelity: environmental fidelity, equipment fidelity and psychological fidelity. Environmental (or physical) fidelity is defined as the degree to which the physical simulation looks, sounds, and feels like the operational environment in terms of visual displays, controls, and audio. Equipment (or functional fidelity) is defined as the degree to which the simulation reacts like the operational equipment to the tasks executed by the trainee. Psychological fidelity is the degree to which the simulation replicates the psychological factors (e.g., stress and fear) experienced in a real-world environment, thus engaging the trainee in the same manner as the actual equipment would in a real world situation.

Jennett, Cox, Cairns et al. [7] claim that all successful (computer) games have one thing in common: an ability to draw people in, to engage them in the experience so that they lose track of time and things around them. Jennett et al. [7] refer to this as immersion which they claim has the following features: (1) lack of awareness of time; (2) loss of awareness of the real world; (3) involvement and a sense of being in the task environment.

B. Vocational Training

The challenges of vocational training and in-job training are many. For example, Billet [8] points out that there is a concern about the lack of transfer of knowledge from educational institutions to workplaces. Transfer is referred to as the extent to which knowledge from one situation can be used in another situation less (far transfer) or more (near transfer) similar to the original situation. However, the concept of transfer has been criticized [9] for not being a sufficient model of explanation. One problem is decontextualization [10], which is referred to as "a process of detachment from conditions that constrain the generality of meanings and actions." [10, p. 136]. van Oers [10] proposes the notion of recontextualization as a better model of explanation as it takes the activity in its context into account. Recontextualization can be interpreted as alternative realizations of a well-known activity in a new context.

The use of simulation training and patient simulators has a long tradition in medicine and medical training [1]. Even though simulation training is a way of role-playing which takes activity into consideration, i.e. *learning by doing* under safe conditions, patient simulators in prehospital training are most often used in a decontextualized way. That is, the patient

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

simulator is used in a clinical setting or in a regular training space where none of the complicating contextual factors pertaining to prehospital care are present. Examples may be disturbances on the scene of the accident such as relatives and upset bystanders, or physical threats to the personal security, as well as noises and physically uncomfortable working environment. Hence, one issue to address in prehospital training is to create a more intense learning experience by using role-play in a richer context, henceforth referred to as a *contextualized scenario*.

III. A CONTEXTUALIZED SCENARIO FOR PREHOSPITAL TRAINING

A. Setting

The project described here is a multi-disciplinary and multi-organizational collaboration between researchers from information technology, the emergency medicine sector, and a regional ambulance service organization [11]. The project is aimed at developing technology support for live role-play training in the prehospital context, i.e. all activities taking place from an alarm call until a patient is delivered at the hospital emergency unit. This is a complex process [12] comprising communication and teamwork skills as well as medical skills and decision making. Emergency medicine is characterized by an uncontrolled stream of patients, high level of acuity, time pressure, lack of information and a wide range of clinical conditions [13]. What distinguish prehospital care from hospital emergency care is that prehospital care is provided far from medical support, in changing settings, and that an ambulance mission is conducted of several phases [14]. The typical ambulance mission involves receiving the call, arriving at the address, performing an on-scene assessment, performing an initial patient assessment, move the patient to the ambulance, performing further assessment and treatment during transport, arriving to the hospital and handing over the patient to the emergency room. The parts of the mission where most critical decisions are made are in the on-scene assessment and initial patient assessment, but every phase influences the decision making in the other phases. For example, already in the phase where the ambulance personnel receive the dispatch call they start to prepare for the mission by imagining the scenario, based on past experiences [15]. Current practice is that different aspects are trained in isolation, e.g. medical skills using patient simulators. This is feasible for clinical contexts but as the prehospital work context is richer, there is a need to incorporate more of it in the training. The case we present aims to capture more of the *complexity* of the prehospital process in order to provide more adequate training. Naturally, this makes training harder to plan and carry out and hence the need for integrated technology support.

B. Technology Support for Live Role Playing

Medical training with patient simulators is a complex matter in itself. We use the SimMan 3G patient simulator (www.laerdal.com). It is composed of a full size manikin which simulates a variety of vital parameters, such as pulse and saturation as well as eye-movement and breathing to provide a realistic perception of a patient. Data from interventions are captured via RFID tags and logged so that the complete record of all interactions with the simulator can be analyzed afterwards. The simulator is operated via WiFi and the operator will typically play the role of the patient by communicating via the speaker system.

As we need to provide a richer context for the training activity, we have utilized a mixed-reality approach where some parts of the environment were recreated by physical props, e.g. using a real ambulance as interface to a simulated driving scenario. We also created different home environments (as examples of dangerous versus safe environments) which were projected to the walls of a conference room and complemented by some physical props. Examples can be found in the right column of Figure 1.The contextualized scenario is designed to better represent all three types of fidelity as presented above (IIA).

Fig. 1. Views from the traditional (decontextualized) and contextualized enactment of a scenario.

C. Immersion in Live Role Playing Scenarios

The aim of the method is to capture and classify indicators of immersion in a role-play training scenario. This is motivated by the theories of Alexander et al. [6] and Morris et al. [16] that immersion may impact retention and transfer from simulated environments. In order to be able to determine the effect of various factors on the level of immersion of the participants we set up a field experiment with two different versions of the same scenario. One was referred to as the traditional set-up and the other one as the contextualized set-up. The scenarios differ both with respect to their scope and how they represent certain details. The latter may serve several purposes: e.g. to enhance immersion of the team and providing information that may be useful in terms of diagnosis or serve as indicators for necessary safety precautions. The general goal of the contextualized scenario is that all information that the team needs to enact the scenario should be delivered within the scenario itself, e.g. they will get information from a simulated SOS Alarm unit rather than orally from the instructor; from the drivers compartment they are exposed to a large screen visualization of the actual turnout; they will get information about the general situation from props, sounds and visualizations of an apartment interior; they will interact with systems and tools in a manner as realistic as possible.

With respect to scope, both scenarios cover medical treatment, care taking and communication but the contextualized scenario also includes actual activities during transportation. However, the major difference in richness is that the traditional scenario is enacted on the scene of accident only whereas the contextualized scenario includes transportation to the scene of accident, activities on the scene of accident, transportation from the scene of accident and treatment in the ambulance caring space including actual transfer between the different environments. Furthermore, the medical equipment was used in a more realistic way in the contextualized scenario. For example, when delivering a medicine to the patient simulator the medication was tagged with the correct dose and the interaction was registered by the simulator. In the traditional scenario, the paramedic would show the medication and say that they deliver a dose.

TABLE I. MAPPING BETWEEN MISSION PHASES AND HOW THEY ARE REPRESENTED IN THE TRADITIONAL AND CONTEXTUALIZED SCENARIOS

Mission Phase	Traditional (decontextualized)	Contextualized
Alarm call	Delivered from the instructor during introduction.	Using a realistic two-way communication system.
Turn-out	Oral information from the instructor.	Full visualization of an actual turn out from the station to the address (4 minutes). Communication with SOS Alarm possible.
Arrival at scene of accident	The instructor informs the team that they are at the scene and that they can start working. Equipment is already in place.	Paramedics have to get an access code from SOS Alarm before being able to enter the building. Crew physically relocate themselves and equipment to the patients' apartment.
Time on scene	The team interacts with the simulator in an office environment that represents the home of the patient. Medicine is delivered by informing the instructor of the action.	The team enters an apartment with props and with interiors projected on the walls - which may indicate e.g. the life-style of the patient or give clues about the situation. A dog is barking behind a semi-closed door, indicating that they have to shut the door as a precaution before interacting with the patient simulator. Medicine is delivered in a realistic way using RFID tags.
Load to ambulance	The patient simulator is loaded to the stretcher. The team stays in place but report that they load the patient into the ambulance.	The patient simulator is loaded to the stretcher, brought out and then loaded into the ambulance.
Transport to the emergency unit	Not a part of the scenario but the team can discuss the case and inform the instructor of what they would have done during transport.	The simulated trip to the emergency unit takes 7 minutes during which additional treatment is carried out in the ambulance caring space.
Hand-over to the emergency unit	The team reports information from the scene to the attending emergency physician in a realistic way.	The team reports to the attending emergency physician in a realistic way.

IV. A METHOD FOR EVALUATING ROLE-PALY SCENARIOS

The method is developed in cooperation between researchers from University of Borås (prehospital care and ICT), University of Skövde (ICT). It was tested in cooperation with one ambulance unit in Region Västra Götaland, Saab Venture and Laerdal Medical during 4 days in November 2014.

The method has a within subject, posttest design comparing immersion between two conditions, one traditional and one contextualized scenario. These were organized in blocks in order to vary the following factors: 1) the type of medical scenario ("elderly man with respiratory distress" or "drug addict with respiratory distress") in each of the scenarios (contextualized/traditional), and 2) the order in which participants did the scenarios.

We recruited 24 participants (specialist ambulance nurses) from four different healthcare organizations in the surrounding region. The participants were recruited as teams from 8 different ambulance units and thus knew each other beforehand. This was with the exception from two cases where stand-ins had to replace an original team member. As all participants (including stand-ins) were trained specialists and used to switching team members this is not considered to impact the outcome. Participants were subjected to the two different scenarios, one in a contextualized setting and one in a traditional setting. Before each scenario, participants were introduced to the scenario by the scenario leader and given time to familiarize themselves with the equipment provided. During each scenario participants was working through an ambulance mission as described in section III.C. Meanwhile, the performance of each participant was graded by an emergency medicine expert according to a standardized protocol, and there were also one researcher observer present

Back to Contents

focusing on non-medical interaction aspects. In addition to the observers, the whole scenario session, including debriefing, was recorded by a number of video recorders and one handheld audio recorder. The traditional scenario was captured by one fixed and one handheld camera. The contextualized scenario was recorded by 6 fixed and one handheld camera.

The contextualized scenario took on average 34 minutes, ranging from 28 to 39 minutes. The traditional scenario took on average 15 minutes, ranging from 10 to 20 minutes. Immediately after each scenario participants completed a post questionnaire on their immersion during the scenario. This was followed by a debriefing with the attending emergency physician. In the traditional scenario, this was a short feedback session where the team got feedback on their performance. In the contextualized scenario, the debriefing session was supported by video recordings from the scenario. Next, participants completed a self-efficacy questionnaire. The scenario/block was concluded by a semi structured team interview with open questions related to how participants perceived the scenario and their own performance. Participants were treated for lunch and refreshments during breaks. In all, each team spent 5 hours at the ambulance facility where the study was carried out.

In all, the study had a rich collection of data ranging from researcher observations, individual questionnaires, team interviews, expert grading to analysis of video recordings. This allows for a large amount of different types of analyses. As immersion is a complicated concept to measure much attention has been paid to construct subjective/introspective (i.e. questionnaires) as well as external (structured observation) instruments. Other relevant variables are self-efficacy as well as professional and gaming experience.

V. CONCLUSION

This work in progress reports on our work towards constructing an instrument to evaluate the impact a richer (training) context has on immersion. The instrument captures data concerning both vocational skills and immersion, from a subjective as well as a more objective observer perspective. Early results show that the contextualized (richer) scenario has potential to create deeper immersion. Immersions is however a complex phenomenon to evaluate and we see the need for a more objective way to explore, measure and analyze immersion, especially in contextualized live role-playing training scenarios. Future work includes the construction and application of an instrument to analyze immersion from a more objective point of view, including tool support. This instrument should be validated and thus also applicable in other domains where immersion is expected to matter. Thus such an instrument can be an important knowledge contribution to the serious games community.

ACKNOWLEDGMENT

The authors would like to thank the ambulance training unit at Skaraborg's regional hospital and the ambulance crews who were kind enough to participate.

REFERENCES

[1] D.M. Gaba, "The future vision of simulation in health care," Quality and Safety in Health Care, vol. 13(suppl 1), 2004, pp. i2-i10.

[2] R.E. Hall, J.R. Plant, C.J. Bands, A.R. Wall, J. Kang and C.A. Hall, "Human Patient Simulation Is Effective for Teaching Paramedic Students Endotracheal Intubation," Academic Emergency Medicine, vol. 12(9), 2005, 850-5.

[3] P. Lankoski and S. Järvälä, "An Embodied Cognition Approach for Understanding Role-playing," International Journal of Role-Playing, Issue 3, 2012, pp. 18-32.

[4] C. Nikendei, A. Zeuch, P. Dieckmann, C. Roth, S. Schäfer, M. Völkl, D. Schellberg, W. Herzog, J. Jünger, "Role-playing for more realistic technical skills training," Medical Teacher, vol. 27, no. 2, 2005, Pages 122-126.

[5] A. Wyatt, F. Archer and B. Fallows, "Use of simaulators in teaching and learning: Paramedics' evaluation of a Patient Simulator," Journal of Emergency Primary Health Care, vol. 5 (2), 2007, article 990168.

[6] A.L. Alexander, T. Brunéy, J. Sidman, and S.A.Weil, "From Gaming to Training: A Review of Studies on Fidelity, Immersion, Presence, and Buy-in and Their Effects on Transfer in PC-Based Simulations and Games," Report from DARWARS Training Impact Group, 2005, Accessed on 2015-03-16 from: http://www.aptima.com/publications/2005_Alexander_Brunye_Sidman_Weil.pdf

[7] C. Jennett, A. Cox, P. Cairns, S. Dohparee, A. Epps, T. Tijs and A. Walton, "Measuring and defining the experience of immersion in games," International Journal of Human-Computer Studies, vol. 66, 2008, pp. 641-661.

[8] S. Billett, "The transfer problem: distinguishing between levels of social practice," Australian and New Zealand Journal of Vocational Education Research, vol 6, no 1, 1998, pp.1-26.

[9] M. Packer, "The Problem of Transfer, and the Sociocultural Critique of Schooloing," The Journal of the Learning Sciences, vol 10(4), 2001, pp. 493-514.

[10] B. van Oers, "The Fallacy of Decontextualization," Mind, Culture, and Activity, vol 5(2), 1998, pp. 135-142.

[11] B. Alenljung and H. Maurin Söderholm, "Designing Simulation-Based Training for Prehospital Emergency Care: Participation from a Participant Perspective," 17th International Conference on Human-Computer Interaction, HCI International, in press.

[12] E. Söderström, J. van Laere, P. Backlund and H. Maurin Söderholm, "Combining Work Process Models to Identify Training Needs in the Prehospital Care Process, Perspectives in Business Informatics Research: 13th International Conference, BIR 2014 Proceedings, September 2014, pp. 375-389.

[13] S.G. Campbell, P. Croskerry and W.F. Bond, "Profiles in patient safety: A "perfect storm" in the emergency department," Academic Emergency Medicine: Official Journal Of The Society For Academic Emergency Medicine, vol. 14(8), 2007, pp. 743-9.

[14] M. Hagiwara, B.O. Suserud, A. Jonsson and M. Henricson, "Exclusion of context knowledge in the development of prehospital guidelines: results produced by realistic evaluation," Scandinavian Journal Of Trauma, Resuscitation And Emergency Medicine, vol. 21(1), 2013, 46.

[15] B.W. Sundström and K. Dahlberg, "Being Prepared for the Unprepared: A Phenomenology Field Study of Swedish Prehospital Care," Journal of Emergency Nursing, vol. 38(6), 2012, pp. 571-577.

[16] C.S. Morris, P.A. Hancock and E.C. Shirkey, "Motivational effects of adding context relevant stress in PC-based game training," Military Psychology, vol. 16, no. 2, 2004, pp. 135–147.

Proposal for Non-contact Analysis of Multimodal Inputs to Measure Stress Level in Serious Games

Fernando Bevilacqua
Federal University of Fronteira Sul
Chapecó, Brazil
Email: fernando.bevilacqua@uffs.edu.br

Per Backlund and Henrik Engström
University of Skövde
Skövde, Sweden
Email: {per.backlund,henrik.engstrom}@his.se

***Abstract*—The process of monitoring user emotions in serious games or human-computer interaction is usually obtrusive. The work-flow is typically based on sensors that are physically attached to the user. Sometimes those sensors completely disturb the user experience, such as finger sensors that prevent the use of keyboard/mouse. This short paper presents techniques used to remotely measure different signals produced by a person, e.g. heart rate, through the use of a camera and computer vision techniques. The analysis of a combination of such signals (multimodal input) can be used in a variety of applications such as emotion assessment and measurement of cognitive stress. We present a research proposal for measurement of player's stress level based on a non-contact analysis of multimodal user inputs. Our main contribution is a survey of commonly used methods to remotely measure user input signals related to stress assessment.**

Keywords*—*Serious Games, Emotion Assessment, Remote Sensing, Computer Vision, Multimodal Input, Affective Computing.

I. Introduction

Serious games are commonly used as a tool whose objective is not solely to entertain the player, but to teach her/him a specific topic. The use of a game mechanic can improve the learning process, since the concept of games is usually appealing and familiar to the users. That same concept, however, can add an overhead to the teaching experience. A stressful game scene, for instance, does not necessarily align with the objectives of a serious game and it can distract the user from learning the desired topic. If a game is able to measure the current emotional state of the player, however, it can dynamically adapt its difficulty level to maximize engagement [1]. Emotion monitoring can be an important factor in the learning process of serious games, because engaged users are more likely to keep focused on the task [2]. The process of monitoring user emotions in human-computer interaction, known as affective computing [3], is a challenging endeavour however. A typical work-flow for measuring or monitoring emotions is usually obtrusive, because it is based on sensors that are physically attached to the user. Sometimes those sensors completely disturb the user experience, such as finger sensors that prevent the use of keyboard/mouse and are highly affected by the movements of the hand.

Research in affective computing aims to replace physical sensors with non-obtrusive approaches. By using computer vision and a video stream captured by a webcam, for instance, one can measure vital signals of a subject without the use physical sensors. The information can be used to infer the subject's emotional state, for instance. A variety of techniques are emerging to remotely obtain information from a person, such as heart rate, respiratory frequency, blood pressure and oxygen level. Those multimodal inputs can be used to make serious games aware of the players emotional state, enhancing the learning process.

This short paper presents a research proposal for measurement of player's stress level based on a non-contact analysis of multimodal user inputs, e.g. respiratory frequency. Our main contribution is a survey of commonly used methods to remotely measure user input signals related to stress assessment. We used a combination of the search terms ”emotion”, ”stress”, ”assessment”, ”computer vision” and ”games” to collect the approaches, however we included only the ones that could be used to remotely infer any subject's state. In the remainder of this work we will: enumerate a list of remote multimodal sensing techniques and their use to measure the user emotional state; explain our proposal for non-contact stress monitoring; present our conclusion and future work.

II. Remote Multimodal Sensing

Multimodal sensing consists of reading different signal channels from a subject. A person presents a wide set of those channels, such as blood pressure, pulse oximetry, facial expressions, pupillary variation, among others. The analysis of such data can be used to infer the subject's emotional/stress state, for instance. Usually biometric sensors attached to the subject are used to read the channels, however remote approaches are proving to be feasible. The following subsections present some of those channels and the sensing techniques used to extract them. All of them are non-intrusive, working by analyzing a video of the subject with no physical contact.

A. Heart rate and correlates

The heart rate (HR) and heart rate variability (HRV), related to beat-to-beat differences, are commonly used for physiological sensing. They are connected with the nervous system and are affected by cognitive tasks [4], [5]. Usually HR and HRV are physically measured by electrocardiogram (ECG), which requires the user to attach small electrodes to the chest. The first approach regarding the remote measurement of HR and HRV was based on photo-plethysmography (PPG) [6]. The pressure of the cardiac activity causes the blood vessels to change its volume (because of the pulse), which makes the light absorption change accordingly. PPG is a time-varying signal resulted from such differences in the light absorption in

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

live human tissue. The use of PPG, however, requires specialized hardware which is usually placed close to the subject. A work that tries to overcome those limitations was proposed by Sereevoravitgul&Kondo [7], a technique to detect HR using a video sequence captured from an USB camera attached to a personal computer. Using common illumination (fluorescent bulb and sun light through the window), the authors record a 20 seconds video of each subject, which remain still during the whole process. After a manually selected region of interest (ROI) of the video, such as the forehead, is analyzed over time. Within the ROI area, the mean of intensities of each color channel (red, green, blue) and hue is calculated. The variation of the mean over time produces a signal, which is analyzed and worked with a discrete Fourier transform (DFT) to extract a spectrum that reveals the HR. According to the authors, the forehead and the cheek area present the best ROI, as well as the use of the hue and the green components for the analysis.

Similarly McDuff et al. [8] show an approach to measure multiple physiological parameters, such as HR and blood volume pressure (BVP), using a basic webcam. Computer vision is used to automatically detect and extract a ROI containing part of the face, whose color channels are also averaged over time. The three signals (red, green, blue) are decomposed using independent component analysis (ICA) and the one containing the highest peak is selected. Using a combination of interpolation and a custom algorithm, peaks are detected and used to identify the BVP. The HR and HRV (and its correlates) are obtained from calculations based on the discovered BVP. The authors also use the identified information to estimate the respiratory rate (RR) from the HRV power spectrum. According to the authors, the root-mean-squared error of the HR and the RR measurement was 1.24 and 1.28 bpm, which is significantly in agreement with a physical sensor used as the ground truth. Datcu et al. [9] present a technique based on the same approach differing only in the ROI used. Using a 30 seconds video from motionless subjects, the face is automatically detected by Viola&Jones [10] and active appearance model (AAM) [11] is used to segment the face into regions of interest. Each video frame is processed using the average of the color channels intensities within the face area segmented by the AAM and the 10 ROIs created from it. According to the authors, using an analysis window of 15 seconds, the best ROIs for HR detection were achieved using the left inner cheek (error of 1.47 bmp) and the forehead (error of 1.49 bpm) areas. The tests suggest that the ROIs originated from the use of AAM instead of the use of whole face ROI improves the accuracy rate and reduce computational time. The ROI corresponding to the whole face does not necessarily provide the best result for HR measurement.

Balakrishnan et al. [12] present a technique to detect pulse in videos by using head motions instead of the intensity of colors. According to the authors, the head performs small movements as a result of the blood flow being pumped through the arteries by the heart. By using a camera placed in front of the subject in an environment with varying light, a video is recorded and analyzed. A region containing the head is selected and feature points are tracked for all frames of the video. By monitoring the vertical component of each feature point over time, it is possible to create a signal containing the head movement frequency, which is processed by a peak detection algorithms that identifies the HR. The method was tested with 18 subjects, varying in gender (7 female, 11 male) and color. The results regarding HR measurement present a mean error of 1.5%, which is similar to the true pulse measured with a physical device. For 16 of the 18 subjects there was no statistically difference between the HR measured with the technique and the one measured physically by ECG. As pointed by the author, an advantage of this technique is that it works even when no skin is exposed, since all calculations are based on motion tracking, not color averaging. As a consequence, the author was able to correctly measure the HR of subjects based on videos of the back of their heads or with subjects wearing a mask. The technique is improved by Irani et al. [13] by using a moving average filter applied to the trajectory of the feature points being tracked to remove the noise produced by other sources of motion, e.g. respiratory activity. Additionally a Discrete Cosine Transform (DCT) is used instead of DFT to estimate the HR signal, which allows the technique to more accurately measure HR, such as when the subject is moving the head or is speaking.

B. *Pulse oximetry*

Additionally to the HR signal, the analysis of blood gases concentration is also an important physiological measurement. It can be used to check the oxygenation status of a person, an important information in the context of a hospital, for instance. The level of oxygen in the blood stream is usually measured by a pulse oximeter, a low-cost device that must be physically attached to the subject. Freitas [14] presents an experiment to check the viability of remote and camera-based pulse oximetry. As explained by the author, the feasibility of remote oximetry relies on two main components: the previously explained PPG signal and multi-wavelength analysis. The latter is related to the use of channels with different wave information to calculate the oximetry, such as the ones recorded by a camera. Of-the-shelf webcams, for instance, usually record videos in three channels (RGB), which contain information from different wavelength. The author performed an experiment similar to the ones already described, based on the average of the intensities of colors in the frames of the video. The only difference is that no ROI was used, each frame was divided in 20x20 pixel regions instead. The results show that some pixel regions produced a better PPG signal than others, for different color channels. According to the author discussion, pulse oximetry is possible under common illumination conditions (ambient light, for instance) and a non-contact and remote measurement device is feasible.

Kong et al. [15] confirm the results of Freitas [14] by presenting a remote method to detect oxygen saturation, under ambient light. In the experiment, the authors use two low light monochrome cameras, attached to a personal computer and mounted in front of the subject. The heart rate is discovered using the previously mentioned technique of averaging the color intensities over time. The oxygen saturation, however, is calculated based on an equation that uses the average information from two different wavelengths (660nm and 520nm), each one obtained by a different camera. The calculation uses the same principle of a physical oximeter, which analyses the light absorption in two different wavelengths in order to determine the oxygen saturation. According to the authors, the results of their non-contact technique is in agreement with the measurement made by finger sensor attached to the subjects, however it can't reach the level of clinical application.

III. Usage of Multimodal Sensing

The different channels obtained with multimodal sensing can be processed and combined to obtain complex information from the person being analyzed, for instance her/his emotional state. Channels based on physiological signals, e.g. heart rate, are considered reliable sources since they are hard to fake, differently from facial expressions [16], for instance. Chanel et al. [17] demonstrate the use of multimodal input analysis to measure emotions and game involvement, using physical sensors to read the input channels. Participants played Tetris in different difficulty conditions while being monitored by a variety of sensors, which included a respiration belt and an EEG system. Each participant, under monitoring, had to rest for a few minutes, play the game then answer a questionnaire about their emotional state and level of involvement. The data from the EEG and the other peripheral signals were computed into two groups of features, which were used to train classifiers to recognize the states: boredom (low pleasure and low pressure, associated with easy difficulty), engagement (higher pleasure and motivation, associated with medium difficulty) and stress (anxiety and low pleasure, associated with hard difficulty). The model was tested for each participant, whose classifier was trained using data from the other participants (excluding the subject). The classification accuracy of the model varies between 48% and 55% for different input signals, classifiers and feature selection.

McDuff et al. [18] present a work of remote measurement of cognitive stress via heart rate variability. Using a digital camera placed at a distance of 3m of the subject, two videos are recorded (both with subjects sitting motionlessly): one with subject resting and other with subject silently performing a mental arithmetic task. Facial landmarks are automatically detected and a ROI containing part of the subject's face is selected for analysis. Using a spacial average of the pixel intensities in each frame, a PPG signal is calculated through independent component analysis (ICA) and BVP is extracted. Based on the discovered BVP, several other physiological parameters are calculated, such as HR, RR, HRV low frequency (LF) and HRV high frequency (HF). The remote measurement of all physiological signals is in agreement with a physical device used as ground truth. Using those parameters, the authors constructed a classifier, modeled with Naive Bayes and support vector machine (SVM), for predicting whether the subject is under cognitive stress. The input features used for the models were mean heart rate, mean respiratory rate, normalized HRV LF power, normalized HRV HF power and HRV LF/HF power ratio for each session. Classifiers were trained for each subject using data from all other participants, excluding the subject being tested. Results show that the prediction accuracy was 85% using SVM, demonstrating that the input signals are sensitive enough to measure the cognitive stress state. The HRV components and the RR were the strongest predictors, while the HR was not significantly different between the two detectable states.

Hjortskov et al. [19] propose a model that integrates unobtrusive multimodal sensing to identify user emotion. The main goal of such approach is aimed at monitoring patients mental health states. The model was created by weighting different features created from input signals, all obtained unobtrusively from a video of the subject's face and the textual content he/she created during the session (e.g. a reply to a tweet). The features are physiological (HR and pupil radius), physical (head movement rate, facial expression, eye blinking rate), based on human-computer interactions (mouse and keyboard usage rate) and based on sentiment analysis of social content (images and textual tweets). All features are used to train a multi-class classifier (based on logistic regression), which learns about the three possible emotion states: negative, neutral and positive. The approach is tested for each subject by training the classifier using data from all other participants, except the one being tested. The user is then presented to a feed of real social texts and images, previously labeled by sentiment analysis. The subject has to report about his/her feelings every two minutes while reading the material. The emotion inference is performed in real-time and based on a probability rule: the state suggested by the classifier with the highest intensity is assumed to be the current subject's emotional state. The results show an accuracy rate of 89% for negative, 56% for neutral and 78% for positive state identification.

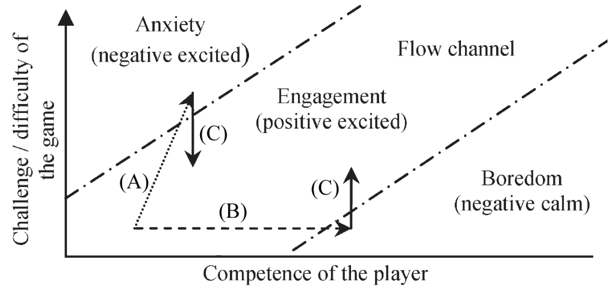


Fig. 1. Automatic adaption of game difficulty based on player's emotional status [17]. (A) change in game's difficulty; (B) change in player's competence; (C) automatic adaptation of the game difficulty.

IV. Proposal for Non-contact Stress Evaluation

As previously explained multimodal sensing uses different input channels to measure or monitor a particular state of a person, e.g. emotional state. We propose a model that is able to infer the current stress level of a user playing a serious game, so it can dynamically adapt its difficulty level in order to maximize engagement. Since an engaged user tends to keep focused on the game's current challenge, it is more likely that the subject will learn about the topic being covered [20]. The subject's input channels will be extracted from a video feed captured using of-the-shelf hardware, such as a webcam, and will be fed into the model. The result will be a stress score S, which is a number in the interval $[-1, 1]$ representing where the subject's stress level is regarding the context of Figure 1. A value of $S = 0$ informs that the subject is engaged (middle area in the figure), while values between $[-1, 0[$ and $]0, 1]$ indicate anxiety (upper-half in the figure) and boredom (bottom half in the figure), respectively. The game can use S to make adjustments to the difficulty level in place.

We plan on using a combination of five signals as input channels for the model: heart rate variability (HRV), respiratory rate (RR), blood volume pulse (BVP), pupillary variation (PV) and facial expression (FE). Other channels will be investigated, such as body/head motion and eye gaze tracking. Our approach will be based on the identification and analysis of changes in the input channels. A subject will be asked to relax and remain still for a few seconds before the analysis, while the input signals are used to create a profile that describes the subject's neutral state. The stress score will be calculated based on a comparison of the subjects current signals against its neutral profile. We envision the calculation of stress score S as:

$$S = \sum_{i=1}^{n} \frac{\overline{\Delta_i}}{T - t_i} w_i \tag{1}$$

where Δ_i is the average of changes of input channel i within the analysis window, w_i is the weight for that channel, T is the current time and t_i is the timestamp of the last time Δ_i changed significantly. The value of S can be interpreted as a weighted sum of the average of the changes of each input channel, divided by the the amount of time elapsed since the last significant change in that channel. Repeatable changes will be labeled as a pattern or a state, however the model might not be able to identify what the added label actually means. As a consequence the game being monitored will be probably required to provide a few hints about its context so the model can correctly tag the changes that are happening. For instance the game can inform that the subject just scored or was hit by an enemy. Our approach, however, will probably be affected by a few potential obstacles. One of them is that most of the techniques previously mentioned require the users to be still during the signals measurement, which might not happen in a game session. We speculate that common movement patterns (e.g. head tilting, speaking, etc.) will cause noise in the input reading, possibly leading to imprecise results. Another problem is the accuracy to measure the subject's neutral state. The emotional state and physical setting prior to in-game measurement is critical because many factors outside of the serious game itself could be in effect when the game begins.

V. Conclusion and Future Work

This short paper presented techniques used to remotely measure different signals produced by a person, e.g. heart rate, through the use of a camera and computer vision analysis. The approaches are mainly based on the use of motion tracking or changes in the intensity of the colors in the pixels to extract a PPG signal. It is then processed and used to calculate several physiological signals, including the oxygen level in the blood flow. The analysis of a combination of those signals (multimodal input) can be used in a variety of applications such as emotion assessment and measurement of cognitive stress. Based on the feasibility of such techniques and their usage, we proposed a research proposal to measure the stress level of a subject while playing (serious) games. The measurement relies on the variations of each user signal over time, aiming to provide the game with information regarding the subject's engagement, anxiety and boredom state. As our future work, we plan to implement and test our approach by creating several experiments to explore suitable technologies and approaches that support our ideas. The results of such experiments will be used to understand the relation between psycho-physiological measures and their connection with the subject's emotions.

References

[1] C. Liu, P. Agrawal, N. Sarkar, and S. Chen, "Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback," *International Journal of Human-Computer Interaction*, vol. 25, no. 6, pp. 506–529, 2009.

[2] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, "Emotion recognition in human-computer interaction," *Signal Processing Magazine, IEEE*, vol. 18, no. 1, pp. 32–80, 2001.

[3] R. W. Picard, *Affective computing*. MIT press, 2000.

[4] N. Hjortskov, D. Rissn, A. Blangsted, N. Fallentin, U. Lundberg, and K. Sgaard, "The effect of mental stress on heart rate variability and blood pressure during computer work," *European Journal of Applied Physiology*, vol. 92, pp. 84–89, 2004.

[5] Z. B. Moses, L. J. Luecken, and J. C. Eason, "Measuring Task-related Changes in Heart Rate Variability," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2007, pp. 644–647.

[6] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological measurement*, vol. 28, no. 3, p. R1, 2007.

[7] T. Sereevoravitgul and T. Kondo, "A comparative study for heart rate measurement in video sequences," in *5th International Conference on Information and Communication Technology for Embedded Systems (ICICTES2014)*, Jan 2014.

[8] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 1, pp. 7–11, 2011.

[9] *Noncontact Automatic Heart Rate Analysis in Visible Spectrum by Specific Face Regions*, vol. 767, Ruse, 2013.

[10] P. Viola and M. Jones, "Robust real-time object detection," in *International Journal of Computer Vision*, 2001.

[11] G. J. Edwards, C. J. Taylor, and T. F. Cootes, "Interpreting face images using active appearance models," in *Proceedings of the 3rd. International Conference on Face & Gesture Recognition*, ser. FG '98. Washington, DC, USA: IEEE Computer Society, 1998, pp. 300–. [Online]. Available: http://dl.acm.org/citation.cfm?id=520809.796067

[12] G. Balakrishnan, F. Durand, and J. Guttag, "Detecting pulse from head motions in video," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, June 2013, pp. 3430–3437.

[13] R. Irani, K. Nasrollahi, and T. Moeslund, *Improved Pulse Detection from Head Motions Using DCT*. Institute for Systems and Technologies of Information, Control and Communication, 2014.

[14] U. S. Freitas, "Remote camera-based pulse oximetry," in *eTELEMED 2014, The Sixth International Conference on eHealth, Telemedicine, and Social Medicine*, Mar 2014, pp. 59–63.

[15] L. Kong, Y. Zhao, L. Dong, Y. Jian, X. Jin, B. Li, Y. Feng, M. Liu, X. Liu, and H. Wu, "Non-contact detection of oxygen saturation based on visible light imaging device using ambient light," in *Optics Express 21*, 2013.

[16] A. Landowska, "Emotion monitoring verification of physiological characteristics measurement procedures," *Metrology and Measurement Systems*, vol. 21, no. 4, pp. 719–732, Dec 2014.

[17] G. Chanel, C. Rebetez, M. Btrancourt, and T. Pun, "Emotion Assessment From Physiological Signals for Adaptation of Game Difficulty," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 41, pp. 1052–1063, 2011.

[18] D. McDuff, S. Gontarek, and R. Picard, "Remote measurement of cognitive stress via heart rate variability," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, Aug 2014, pp. 2957–2960.

[19] D. Zhou, J. Luo, V. M. Silenzio, Y. Zhou, J. Hu, G. Currier, and H. Kautz, "Tackling mental health by integrating unobtrusive multimodal sensing," in *Proceedings Of The 29th AAAI Conference On Artificial Intelligence And The 27th Innovative Applications Of Artificial Intelligence Conferenc*, April 2015.

[20] M. Csikszentmihalyi, *The psychology of optimal experience*. Harper & Row, 1990.

Back to Contents

Continuous Collision Detection Using Tetrahedral Structures

Jan Čejka
Faculty of Informatics
Masaryk University
Brno, Czech Republic
xcejka2@fi.muni.cz

***Abstract*—This work presents a new method for computing continuous collision detection between two rigid bodies. It introduces a new acceleration structure that divides the object space into tetrahedra, and describes how this structure can utilize temporal coherence in computing collision detection. It explains how to detect collisions between basic primitives using this structure, especially vertex-face, edge-edge, and face-vertex collisions.**

***Index Terms*—Collision detection, CCD, tetrahedron, rigid bodies**

I. Introduction

Virtual words need not only to display the data (using visual, acoustic, or other output), but often also to simulate it. These simulations need to know if and how objects interact with each other.

Many simulations compute the movement of bodies, and as an intrinsic part of their computation they need to determine whether objects collide or not. If so, the collision detectors must provide the time and place of the collision, and other information depending on the simulator.

Discrete collision detection (DCD) checks for collisions between all objects at discrete times. Most existing implementations test whether objects collide in the final position, and based on this information they deduce all the data about the collision. DCD is fast and sufficient in many applications, including computer games and real-time simulations. However, it has several drawbacks. The most important is that the DCD cannot determine the exact time and place of the collision.

This is not true for the continuous collision detection (CCD). It simulates the movement of the object through the whole continuous time interval, and searches for the first collision during this movement. Of course, it usually spends more time for computation than DCD.

Despite that, CCD starts to appear in games at places where a more precise collision detection is necessary. For instance, Catto [1] shows its usage in Blizzard game Diablo III. CCD is also supported in some physics libraries, for example in Bullet [2] or Havok [3].

Because of the discussed advantages, we focus our work on CCD. We are interested in simulations that perform small movements, but demand precise results. These simulations span a large area of problems, from assembling components of cars and other machines to precise manipulation of small objects in medicine. Making these simulations fast allows us to add haptic force feedback and create very good training aids.

Some algorithms detecting collisions for haptics simulations already use CCD. Redon et al. [4] present a method of CCD that uses bounding volume hierarchy (BVH) and interval arithmetics. The original method operates in real-time, although not yet fast enough for haptics. In [5] It is improved and adapted it for haptic interactions.

Small movements are handled faster by algorithms that utilize temporal coherence. Lin [6] proposes a simple method to accelerate collision detection between two convex objects. It is improved in [7] which divides non-convex objects into a set of convex objects and thus allows to use the previous algorithm.

In this paper we present an algorithm that uses the temporal coherence and is suitable for any objects, including separate primitives. Its basic concepts are described in Section II. Sections III and IV explain the algorithms for solving basic collision detection queries, and the final section V discusses unsolved issues and future work.

II. Basic concepts

Our algorithm consists of two phases: a precomputation phase and a collision phase. In the precomputation phase, we divide the whole space of and around a static obstacle into convex polyhedra, or *cells*, that do not overlap and fill the space. The faces of these polyhedra are a superset of faces of the obstacle, as illustrated in Figure 1. We also do the same for a moving probe.

The collision phase may be further split into four parts. In the first part, we detect whether moving vertices of the probe collide with faces of the obstacle. To do that, we check every vertex individually whether it collides with faces of a cell of the obstacle space in which it is located. A collision between the objects occurs only when the vertex collides with a face that is one of the faces of the original obstacle and not with the faces of the tetrahedral structure.

In the second part of the collision phase, we detect collisions between moving faces of the probe and vertices of the obstacle. We transform this problem into the previous problem by switching the role of a probe and an obstacle.

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

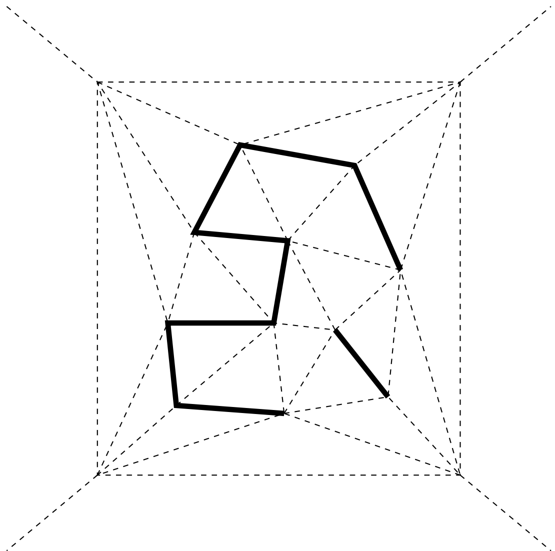


Figure 1: Tetrahedralizing the space of an object. We start with faces of the object (thick lines), and divide the space of its enlarged bounding box into tetrahedra by adding additional triangles (dashed lines). We also divide the space outside the box into infinite tetrahedra.

In the third part, we detect collisions between probe edges (called *segments*) and obstacle edges. A collision between the probe and the obstacle may again happen only when the segment collides with one of the edges of cells the segment is inside, and if the edge it collides with is an edge of the original obstacle.

In the final fourth part of the collision phase, we perform the part of the movement without any collisions, and update the information about in which cells the primitives are. We use this information in the next frame of the simulation.

All collisions are computed in the space of convex polyhedra. Thanks to this, we may simplify many algorithms to work with infinite lines and planes, instead of line segments and faces.

We also simplify the movement of objects as a linear movement between starting and ending position. The movement of a probe is treated as a relative movement from an obstacle, allowing us to make obstacles immovable.

A. Dividing a space into polyhedra

We use Tetgen library [8] to divide the space of objects into tetrahedra. We use tetrahedra because they are trivially convex, and all polyhedra we get have the same number of faces.

Tetgen library divides only the space inside an object, so we place the object into its enlarged axis aligned bounding box and tetrahedralize the space of this box, see Figure 1. This effectively creates a single structure containing all disjoint parts of the object.

We also need to divide the space around this bounding box. We use a technique similar to the one described in [6]. We create infinite tetrahedra starting at the faces of the box and expand them from the center of the bounding box.

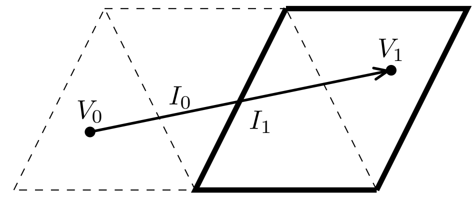


Figure 2: $V{\times}F$ collision detection: A vertex starts at position V_0 and moves to position V_1. At intersection I_0, it changes its active cell. At intersection I_1, the vertex collides with the obstacle, and stays in its active cell.

III. Detecting collisions between vertices and faces

We start with detecting collisions between a moving vertex and faces of the obstacle. We denote this situation as $V{\times}F$. To compute the collisions, we start in the cell in which the vertex is positioned (*active cell*), and compute, whether it intersects one of the faces of the cell during its movement (see Figure 2). Since we work with convex polyhedra, we can treat all faces as planes.

We need to handle some special situations. First, we consider only faces for whose the vertex position at the end of the movement is outside its plane. However, we do not do the same for the starting position of the vertex. Removing such planes would fail in situations when the vertex is at the face at the end of the previous animation frame, and is treated as outside at the beginning of the current frame, due to numerical errors.

The next issue we must solve is the problem when the movement is parallel to the faces of the cell. We detect these faces and exclude them from computation.

We compute the time t of intersection for all faces that were not removed in the previous steps as

$$t = \frac{V_0 p}{V_0 p - V_1 p}$$

where V_0 is the position of the vertex at the beginning of the movement, V_1 is the position of the vertex at the end of the movement, and Vp is the signed distance of a point V from the plane p of a given face.

We choose one face from all intersected faces which corresponds to the smallest time of intersection. If it is a face of the original obstacle, we report a collision. Otherwise, we take a cell behind that face as a new active cell, and compute the collision again using its faces. In case the vertex intersects no face, it does not leave the cell and no collision occurs.

A. Commit vector

As already mentioned, we use the last active cell in current animation frame as the initial active cell in the next frame. Unfortunately, we cannot save this information directly, because the probe's movement may be interrupted by a collision of another vertex or edge.

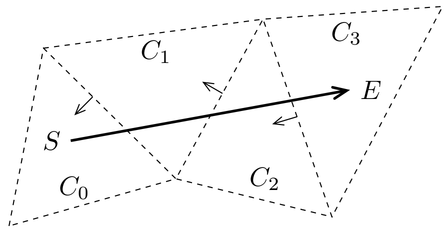


Figure 3: Validity of active cell list, planes between C_0 and C_1, C_1 and C_2, and C_2 and C_3 must head toward the starting point S of the line segment SE.

To solve this issue, we introduce a concept of a commit vector. The commit vector is an array of pairs of the time of an intersection and a change of active cell that happens at that intersection. During the collision computation, we store all changes of active cells in this commit vector and redo them in the final part of the computation, when collisions of all primitives are already known.

IV. Detecting collisions between edges

Detecting collisions between edges of the probe and the obstacle (denoted as $E{\times}E$) is similar to that between vertices and faces. We detect collisions in cells in which the edge currently is.

We associate a list of active cells with each line segment. This list of cells must satisfy the condition that no cell must be presented twice, the neighbouring cells in the list must share at least a face (sharing a vertex or an edge is not sufficient), and the face must be facing toward the starting point of the segment (the dot product of the direction from the starting point to the ending point of the segment and the normal of the plane it intersects in the preceding cell in the list of active cells must be non-positive), see Figure 3. These cells are sorted in a way that the first cell contains the starting point of the segment, and the last cell contains the ending point.

A. Detecting collisions between a line segment and edges of a cell

First, we detect collisions between the line segment and edges of the face between two neighbour cells in the list of active cells. We do this for all cells in the list.

We compute the intersection of the face and the edge in homogeneous coordinates, using the following formula:

$$Q_p = S^p E - E^p S$$

where Q_p is the point of intersection in homogeneous coordinates, S and E are the starting and ending points of the segment and p is the plane we intersect. S^p is a dot product of two four-dimensional vectors, representing homogeneous coordinates of a point S and coefficients of a plane p.

Signed distance $d(Q_p, a)$ of the intersecting point Q_p from another plane a is

$$d(Q_p, a) = \frac{Q_p^a}{Q_p.w}$$

where $Q_p.w$ denotes the fourth homogeneous coordinate of the point Q_p, and Q_p^a is a dot product of Q_p and a. The sign of $Q_p.w$ is positive if the plane heads towards the starting point, zero if the face is parallel to the line segment, and negative if it heads away from the starting point. Thanks to the ordering of cells in the list of active cells, $Q.w$ is always non-negative.

A collision between the segment and the edge on the intersection of planes p and a occurs when $d(Q_p, a) = 0$. Since $Q_p.w$ is always non-negative and S and E changes linearly, we solve a quadratic equation to get all situations in which $Q_p^a = 0$. Although an intersection happens at any moment when $Q_p^a = 0$, we use only the moments in which the derivation of Q_p^a is negative. This negative derivation indicates the situations in which the segment starts intersecting the plane a and leaves the plane p (the plane between the neighbour cells).

If $Q_p.w$ is zero (the segment is parallel to the face), we switch the planes p and a. We compute the point of intersection Q_a, its distance to p, the moments when this distance is zero, and use positive derivation. Since $Q_p^a = -Q_a^p$, we can use the same result described in the previous paragraph.

We again need to deal with special cases. A problem appears when the edge collides with two edges at one of the cell's vertices (we refer to this situation as $E{\times}V$ collision). In such situation, we need to choose only one of these edges. To decide it, we compute the intersection of the segment and both faces, and choose which point recedes from the other plane using derivation of its distance from the plane (the derivation is positive). If both derivations are zero, we may choose the face arbitrarily.

We may also detect a collision with several faces at the same time. This happens when $Q_p.w$ is zero, since the algorithm reports a collision with all three faces at the same time. We solve this issue by computing the point of intersection with all these planes, and choosing a face with intersection which is inside the cell and whose fourth homogeneous coordinate is positive.

B. Updating the list of active cells

The second part of the algorithm updates the list of active cells after a collision happens in any cell. We differentiate between two situations, a collision in the middle of the edge (we call this *pure* $E{\times}E$ situation), and a collision at end points of the edge ($E{\times}V$ situation).

First, we focus at *pure* $E{\times}E$ situations. If the collided edge is one of the edges of the original obstacle, we report a collision and stop the computation. Otherwise, we need to update the list of active cells after the segment traverses the edge. We find the first and the last cell in the list of active cells that contains this edge (we call these cells *entering* and *exiting* cells, respectively), see Figure 4. Then we remove all cells between these two cells and add new cells located at the opposite side of the edge.

Before adding new cells, we must check whether their intersected faces facing towards the starting point of the line segment. If they do not, we cannot allow this transition.

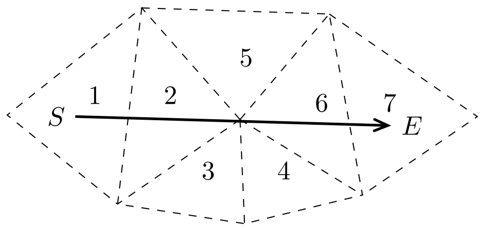


Figure 4: Updating *pure* $E{\times}E$ situation, the line segment is moving upward. Cells before update $(1, 2, 3, 4, 6, 7)$, entering cell (2), exiting cell (6), cells to be added (5), cells to be removed $(3, 4)$, cells after update $(1, 2, 5, 6, 7)$

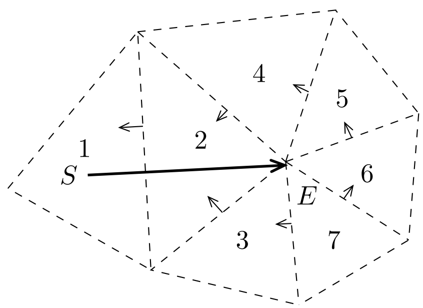


Figure 5: Invalid $E{\times}E$ update, the line segment is moving upward. Cells before update $(1, 2, 3, 7)$, entering cell (2), exiting cell (7), cells to be added $(4, 5, 6)$, cells to be removed (3), cells after update $(1, 2, 4, 5, 6, 7)$. Orientation of the face between 6 and 7 is invalid. To make this $E{\times}E$ transition valid, we must first move the end point from cell 7 to cell 6.

This happens in situations when the edge is traversed by an end point of the segment, as illustrated in Figure 5, and is processed later.

The $E{\times}V$ situation is more complicated. We again check whether the vertex is not a vertex of the obstacle (and report a collision if so), and find the first and the last cell that contains the vertex (again called as *entering* and *exiting* cell). Unlike *pure* $E{\times}E$ situation, each plane between the entering and exiting cells is intersected in a vertex, so there are two edges in each plane, from which we must choose which to traverse. We must do this decision consistently for all these edges.

We start with the plane that first reported the collision, mark its edges as *Traverse* or *NotTraverse*, based on which edge our algorithm decided to traverse, and propagate this information to all other planes. If an edge is marked as *Traverse* in the previous or the next cell, it must be marked as *Traverse* in the current cell. The same holds for edges marked as *NotTraverse*. If an edge is marked as *Traverse*, the other edge must be marked as *NotTraverse*, since we cannot traverse both edges at the same time. Finally, when we have an edge marked as *NotTraverse*, we use derivations to decide whether the segment leaves the face. If it does, we mark the other edge as *Traverse*. If it does not leave it, we mark it as *NotTraverse*.

When we have all edges marked, we start traversing them. We choose any edge marked as *Traverse*, and traverse it in the same way we traverse an edge in *pure* $E{\times}E$ situation, including the test for plane orientation validity. If the transition is not valid (faces would not face toward the stating point), we postpone this transition and choose another. With each transition we remove and add new cells, and update flags of affected cells. We mark the traversed edge as *NotTraverse* to avoid traversing them back, and other edges as described before.

The whole $E{\times}V$ transition is allowed only when all edges marked as *Traverse* are traversed. If some of these edges are still present and nothing can be traversed, we disallow the whole ExV transition. This again happens when traversing with end point of the line segment.

C. *Processing end points of the segment*

A change in the list of active cells of a segment also happens when its end points change their active cells. When this change occurs, we update the list using the data from the commit vector of the vertex.

We must again perform a test whether the added face heads towards or away from the starting point. Such situation may happen when the end point traverses the face at edge of at vertex. In such situation, we analyze this edge or vertex, perform $E{\times}E$ or $E{\times}V$ transition (which was disallowed or would be disallowed in future) and perform the transition again, as illustrated in Figure 5.

V. Conclusion and future work

We presented a new method for continuous collision detection between two rigid bodies. Our solution uses a space tetrahedrization to utilize the temporal coherence to accelerate the computation. We proposed algorithms that detect collisions between moving vertices and edges of one object in tetrahedra of another object.

Our method is currently very dependent on proper solution of special cases, which occur during simulations. Though we solved many of them, some problems still appear. They are mostly connected with solving parallel lines and end points of segments. These problems are the subject of our future work.

References

[1] E. Catto, “Continuous Collision,” presented at the Game Developer Conference (GCD 2013), 2013.
[2] Bullet Physics Library. [Online]. Available: http://bulletphysics.org
[3] Havok Physics. [Online]. Available: http://www.havok.com/products/physics
[4] S. Redon, A. Kheddar, and S. Coquillart, “Fast Continuous Collision Detection between Rigid Bodies,” *Computer Graphics Forum*, 2002.
[5] M. Ortega, S. Redon, and S. Coquillart, “A Six Degree-of-Freedom God-Object Method for Haptic Display of Rigid Bodies,” in *Virtual Reality Conference, 2006*, March 2006, pp. 191–198.
[6] M. C. Lin, “Efficient Collision Detection for Animation and Robotics,” Ph.D. dissertation, University of California, 1993.
[7] M. Ponamgi, D. Manocha, and M. Lin, “Incremental Algorithms for Collision Detection Between Polygonal Models,” *Visualization and Computer Graphics, IEEE Transactions on*, vol. 3, no. 1, pp. 51–64, Jan 1997.
[8] H. Si, “TetGen, a Delaunay-Based Quality Tetrahedral Mesh Generator,” *ACM Trans. Math. Softw.*, vol. 41, no. 2, pp. 11:1–11:36, Feb. 2015. [Online]. Available: http://doi.acm.org/10.1145/2629697

Back to Contents

Drink & Drive: A Serious but Fun Game on Alcohol-Induced Impairments in Road Traffic

Frank Gaibler, Simon Faber, Sarah Edenhofer, Sebastian von Mammen
Organic Computing, Dept. of Computer Science
University of Augsburg, Germany
{frank.bernhard.gaibler,simon.alexander.faber}@student.uni-augsburg.de
{sarah.edenhofer,sebastian.von.mammen}@informatik.uni-augsburg.de

***Abstract*—In this paper, we present *Drink & Drive*, a serious game about the effects of alcohol in individual transport. Although there were a few projects in place, we were eager to offer a more attractive gaming experience compared to the preceding projects, especially as the effects of drunk driving should be communicated to a young audience. *Drink & Drive* implements a competitive third-person racing game relying on well-known game mechanics. The goal of *Drink & Drive* is the motivation of a safe arrival by systematically avoiding any alcoholic beverages if driving. Accordingly, collectibles on the track can influence the driver's state of perception, for the better and for the worse. The delicate balance between serious contents and gaming fun is established by placing great emphasis on a clear, playful design and the introduction of effective gamification elements that support an immersive, intrinsically motivating experience. In this paper, we motivate and present the design of *Drink & Drive*, we embed it into the context of related works and we provide preliminary results based on its presentation to forty students and faculty.**

I. Introduction

Due to technological advances, driving cars is becoming increasingly safe—serious accidents are mainly caused by human error. Hereby, the reduction of the drivers' perceptional and cognitive abilities through alcohol and other consumable substances such as THC [1] plays a considerable role. Although widely known, costly campaigns are run by governments to further raise awareness about this fact. In Germany, especially younger drivers (between 14 and 27 years) seem to underestimate the influence of alcohol on their driving abilities, see for instance [2]. A lack of internalisation and ineffective ways of teaching about the problem are two hypotheses that explain the gap between knowledge and self-assessment. As always, teaching facts without providing real experiences renders it hard to actually appreciate the stakes of a poor decision. Serious games can appropriately address this challenge [3]. Virtually, the impairment of a driver that results from a certain blood alcohol concentration (BAC) can be simulated considering perceptive distortion and prolonged reaction times. Hence, it is possible to provide a comprehensive, multi-sensory, potentially emotionally supported learning experience and not only bare facts about a hypothetical situation.

In this paper, we present our approach of tackling this challenge. Our implementation, *Drink & Drive*, is a serious game that wraps the issue of drunk driving in an appealing, accessible disguise. In particular, at a first glance, *Drink & Drive* resembles the popular video game title "Super Mario Kart" [4]. Despite of *Drink & Drive*'s cartoonish looks and its simplistic user interface, it is built on three goals that are motivated by its serious background. (1) The impairment of alcohol in a driving situation needs to be communicated. (2) The conveyed information needs to be based on scientific facts. (3) In order to render the teachings effective, *Drink & Drive* needs to be accessible and fun.

The main contribution of this paper is the concept of *Drink & Drive* enriching serious contents with actual fun gaming elements. Before we dive into the design details in Section III, we present related works in Section II. Here, we also outline differences and commonalities to our approach. Afterwards, we discuss our results and present some ideas around future work in Section IV.

II. Related Work

Elaborate research work has been published on the impairments to driving abilities induced by alcohol. The effects to the human body caused by a certain BAC are well examined and documented, for a quick introduction, see for instance [5].

Motivated by the adverse effects of alcohol in individual traffic, several attempts have been made to use simulation games to establish awareness about the associated risks.

The German non-governmental organisation "Alliance against Alcohol and Drugs in Road Traffic" (B.D.A.S.) provides simulators that offer a highly realistic experience of drunk driving [6]. Their first generation of simulators are turnkey driver cab's with space for a single person, steering wheel, manual shifting lever and pedals. A wide-screen and speakers provide for the feedback. A similar setup with extensive means of configuration and analysis has been developed by researchers from Seoul in collaboration with the local traffic authority [7]. B.A.D.S.' latest generation of drunk-driving simulators makes use of compact Smart or Renault Twizzy cars and augments them with an external projection. Following a parametric model, the users are shown how their perception and reaction are affected by different levels of BAC. Similarly to the mentioned simulators, *Drink & Drive* creates an experience of drunk driving based on a parametric distortion model of the user's perception and reactivity dependent on his (simulated) BAC. Yet, *Drink & Drive* does not rely on a comprehensive set of hardware gadgets. Instead, we try to provide an immersive experience engaging the users by

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

means of gamification. In this way, we also ensure that *Drink & Drive* is affordable and accessible to the broad public.

Another approach was presented by students and faculty members of the University of Calgary [8]. In their gamified simulation ”The Booze Cruise” (TBC) they challenge the users in travelling the way back home from a party by car, while being under the influence of alcohol. Like our approach, TBC does not simply confront the user solely with an alcohol-augmented driving experience, but rather it tasks him with a specific goal (in this case: getting home safely). Also, TBC integrates game elements such as surreal obstacles to render the simulation more exciting, and thus more attractive to the user. We see two shortcomings in the design choices of TBC: The user steers the car from a third-person perspective. Given the large number of first-person racing games, we are convinced that a first-person perspective, i.e. sitting at the simulated steering wheel, does not jeopardise joyful gaming, at the same time it provides for a more realistic driving experience. Moreover, TBC uses a preset BAC which does not change while driving. Due to this rather limiting experience, we decided to introduce collectible items on the driving tracks in *Drink & Drive* which raise or lower the BAC also during the ride. By doing so, the user's initial loss of control may result in a vicious circle of alcohol misuse that boosts the driving challenge. The analogy to excessive alcohol consumption is a welcomed side-effect of our design.

Another virtual drunk driving simulator, the “DUI Simulator”, was developed as part of the Global Game Jam in 2014 [9]. This simulator offers a third-person perspective and challenges the user to drive a car under the influence of alcohol while avoiding obstacles. The game ends when the car leaves the track. The simulated BAC as well as the car's speed are increasing with the achieved distance. This game offers an adequate simulation of alcohol-induced impairment and also supports a virtual reality headset. Its greatest shortcoming is the limited gameplay options due to the linear challenge. The game also suffers from the lack of a clear concept of competition.

The multiplicity of driving simulations, especially also frequently deployed systems such as [7] and [6], which teach the effects of alcohol-induced impairments emphasise their utility. Yet, we built on the preceding works in order to reach a broad audience and in order to effectively instil the dangers of drunk driving into a younger audience.

III. Methodology

This section shows how *Drink & Drive* stimulates the user's interaction and it covers the underlying models, including game mechanics and alcohol-induced impairments.

A. Interaction Elements and Interface Concept

Since *Drink & Drive* is a driving simulation, the player takes control of a car and interacts through it with the world around him. The final result is represented by Figure 1. As *Drink & Drive* needs to appeal to a broad group of users and especially to the younger audience, the user interface has to be simple and provide precise feedback. For our first implementation, we decided on simple visuals and an automatic car, which can accelerate forward and backward, turn left and right, and break – all by pressing standard computer keys, i.e. the arrow keys and the space bar, respectively.

Fig. 1: *Drink & Drive*'s first-person default view is reduced to a simple steering wheel dashboard and a few icons that represent the time left to complete the track (the heart icon in the upper-left corner), the achieved score (the diamond icon next to the heart icon), and the alcohol blood concentration (to the right-hand side).

In order to reflect the intoxication level of the driver, we applied different post-processing filters to the rendering viewport and we increased steering and acceleration effects to correspond to perceptional and cognitive lags. With an increased intoxication, these effects get stronger and more difficult to ignore. The underlying scientific model is detailed as part of the next section. At a rather high BAC, the driver faints and the simulation ends.

Gamification elements have been introduced to engage the user. We deployed a high score system to get the user motivated to interact with the environment and to collect diamonds that appear on the track (Figure 2). The scoring system takes advantage of the basic human urge for sportive competition and will encourage him to repeatedly play the simulation and to improve his performance over time. Information about the game mechanisms, including the scoring theme, are conveyed to the user in a simple textually-guided tutorial, see Figure 3. To counterbalance the ease of achieving high scores, and to allow for a non-linear intoxication experience, we decided to design an interactive mode of alcohol consumption, too. Therefore, alcohol is represented as beer cans that can be picked up like diamonds (see also Figure 2). Of course, their uptake does not contribute to the user's score but increases his BAC and thus they hinder the driver from following an optimal route and reaching high scores. Due to the cans' placement at road constrictions or round a bend, the player either needs to drive very slowly to avoid them, which conflicts with the requirement to finish the track fast, or handle the driving mechanics rather well and follow a challenging trajectory. In the end, especially novices to the game quickly experience the simulated impact of alcohol, whereas advanced players hone their skills to avoid them.

Without further incentives, the user could simply drive as slowly as possible and, thereby lever the difficulty of the game.

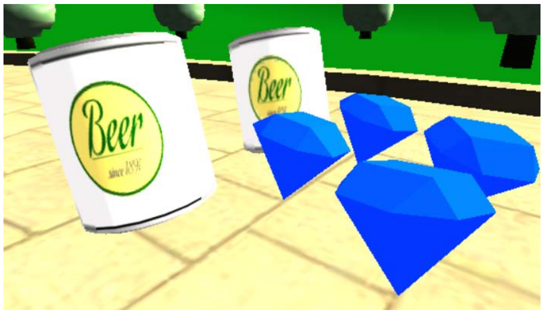


Fig. 2: Alcoholic beverages and diamonds can be picked up from the road - the first increases the driver's blood alcohol concentration, the latter his score.

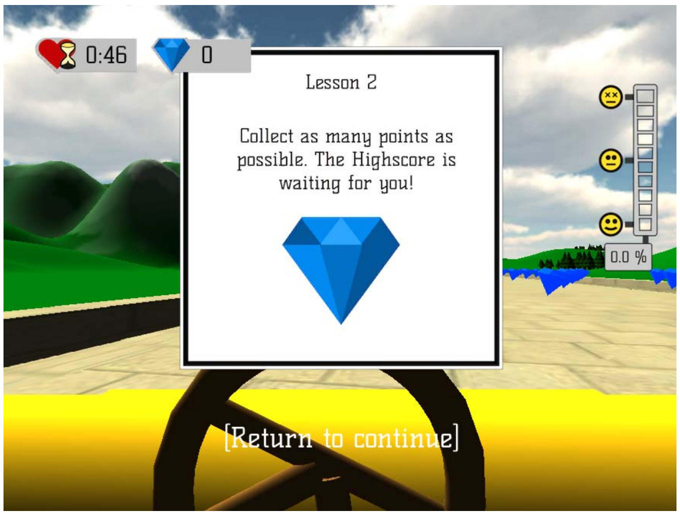


Fig. 3: A tutorial quickly steps through the game mechanics.

Fig. 4: We developed a trendy style to attract our young target audience (14 to 27 yrs).

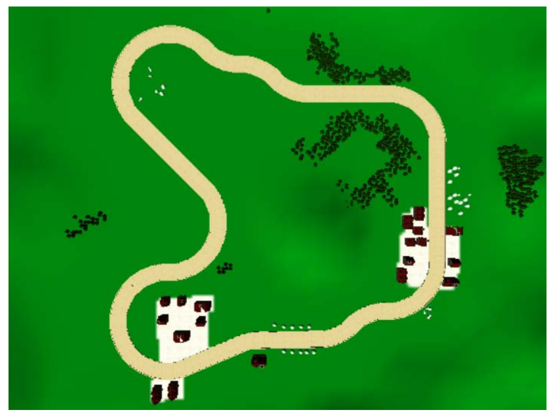


Fig. 5: The default track seen from above. The windings impose user interaction without pushing the challenge too far.

Therefore, some form of restriction had to be implemented. We decided to give the player a certain amount of time to complete one lap on the track. If he does, the timer is reset, otherwise the simulation ends. Similar to trendy endless runners and physics puzzlers such as Angry Birds, an infinite play is only prohibited by practical reason. As a result, there is always an incentive in returning to the simulation and in trying for a new high score.

Next to non-linear interaction with collectible items and to constant time pressure, we also introduce a linearly growing challenge factor (with an upper boundary). At later stages of the game, computer-generated opponents spawn on the track. They drive in the opposite direction, are not contenders and do not collect any of the track items. But they simply impede driving on an optimal route by blocking the player and pushing him around. Their anti-social interference may result in the driver's involuntary alcohol consumption. They spawn in greater numbers as the game progresses, however, there is an upper limit to maintain a playable level of difficulty.

To render *Drink & Drive* attractive for our target audience, we developed a flashy, eye popping style that we deployed consistently throughout the game, see for instance Figure 4. Simulating three different daytimes creates atmosphere and allows the user to experience the effects of high BAC under different conditions. The immersion is rounded off using an easy listening tune for the background music and brief, cartoonish sounds to underpin various user interactions such as picking up items.

B. Simulation Model and Scenario

The car represents one of the most important aspects of *Drink & Drive*'s game mechanics. Again, we needed to balance realism and simplicity. The car is able to interact with everything on the track, but it cannot somersault and it does not easily lose traction. The user is therefore always able to continue driving and any potential frustration due to a lack of experience driving real cars is avoided. Along the same lines, a realistic damage model would have been counterproductive due to the expectedly great degree of frustration by the user. We designed the track in such a way that it would support the learning process about intoxication. This determined its relatively short distance and its slightly curvy flow (Figure 5).

With the track being static, other aspects of the scenario had to be flexible and versatile, as to not lead to a repetitive

BAC in ‰	Effects
0.3	- Slight decrease in vision performance, of concentration and judgement - Impaired reflexes
0.5	- Speed is miscalculated - Vision decreases by about 15% - Reduced hearing
0.8	- Vision decreases by about 25% - Tunnel vision - Reaction time prolonged by 30 to 50% - Balance disorders
1.0-2.0	- Further deterioration of vision performance - Disorientation and problems with spatial vision - Significantly disrupted responsiveness - Loss of the ability to accept criticism
2.0-3.0	- Strong balance and concentration problems - Memory gaps and confusion - Vomiting and muscle relaxation
at 3.0	- Unconsciousness and memory loss - Weak breathing and loss of reflexes
at 4.0	- Paralysis, coma and death

TABLE I: The adverse effects of alcohol on the human body.

and ultimately boring experience. To keep an element of surprise, diamonds and beer cans are randomly distributed along the track. In order to ensure that the driver is not forced into any obstacles, i.e. to ensure that skilled driving provides for a smooth escape from any beer cans on the track, we defined spawning areas for the two types of icons (which could overlap) and allowed the spawned icons to shift and rotate within predefined ranges. This technique is similar to calculating potential solutions to inverse kinematics based on time and space constraints, see for instance [10]. Cars driving the opposite direction also add to a varied gaming experience. They implement a simple behaviour, following a given path as closely as possible.

The realistic demonstration of the effects of alcohol is an important part of the model. It is based on data from the German federal centre for health education [11]. Table I summarises the physiological impairment. We focused on visual effects. With an increase in BAC, our visual perception blurs, gets shaky, loses colour and the edges get darker. Impacts on our sense of hearing is another important facet. Sounds appear muffled and less clear. Consequently, the feedback from and the knowledge about our environment is heavily diminished. To infer concrete numbers, we assumed a 20 year old, 1,75m tall male driver, weighing 75kg. We further assumed a beer can to contain half a litre of (Bavarian) beer with 5% alcohol strength. This leads to one beer equaling about 0.3‰ of BAC. The body shuts down at around 3‰. As a result, collecting ten cans of beer results in loosing the game.

There are no impediments at the beginning of the game. The player always starts the ride sober. Intoxication first results in a slight deterioration of clear-sightedness. Next, serious impairments start taking effect. At 0.6‰, the vision gets darker at the edges and the hearing starts to get dull. Camera and sound filters are used to simulate this. For instance, a fisheye effect distorts the image and a vignette is laid over the picture to darken the edges (see Figure 6). An increase in alcohol results in prolonged reaction times, diminished ability to assess speed, and heightened readiness to take risks. We translated these changes into a faster simulation. Given the relatively long period of alcohol degradation, the player cannot experience this process in the game.

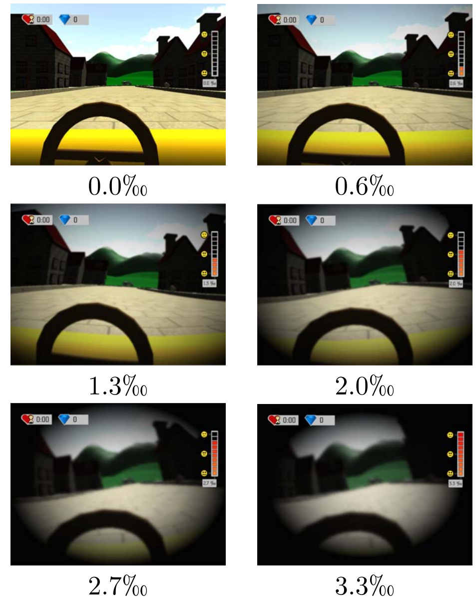


Fig. 6: Succession of increasingly low performing vision with increasing blood alcohol concentration.

IV. Results & Future Work

Different from its precursors, *Drink & Drive* offers a fun and informative gaming experience about alcohol misuse in road traffic. It implements a scientifically backed model of perceptional and reaction impairments similar to numerous serious, non-game simulators but it also introduces gamification elements to engage young players. The impairment model is based on scientifically determined facts, but the corresponding visualisation can only be an approximation. In general, *Drink & Drive* is not set out to achieve great realism but rather to integrate the serious contents that are valuable and the game elements that trigger intrinsic motivation to engage the players. In particular, we tried to realise some of Koster's work, speaking to the users' competence, to establish relationships with the game contents but also social links providing a competitive environment, and to ease the user into playing, maintaining his autonomy as much as possible [12].

In a competition on interactive simulations, we presented *Drink & Drive* to about 40 people, most of them students. They voted the game to be the best out of 15 projects, including interactive simulations about ants foraging, bee colony defence, and medical surgery. Criteria in the competition comprised the complexity of the scientific model, usability, and visual appeal. As a next step, we will present *Drink & Drive* to a younger subset of the target audience (at ages 14 to 18) and inquire about their findings. We hope this study to allow for various finishing touches before releasing *Drink & Drive* to the public.

There are several directions that we would like to pursue as future work. For instance, we would like to explore the benefits (and challenges) of using virtual reality headsets such as the Oculus DK2 [13]. Along the same lines, we have also been deliberating about a more natural motion-based input control [14] – a mobile version could utilise built-in accelerometers for steering. A mobile version would also allow for a much wider audience to experience the benefits of driving sober.

References

[1] A. Ronen, P. Gershon, H. Drobiner, A. Rabinovich, R. Bar-Hamburger, R. Mechoulam, Y. Cassuto, and D. Shinar, "Effects of thc on driving performance, physiological state and subjective feelings relative to alcohol," *Accident Analysis & Prevention*, vol. 40, no. 3, pp. 926–934, 2008.

[2] Statistisches Bundesamt, "Verkehrsunfälle: Unfälle unter dem einfluss von alkohol oder anderen berauschenden mitteln im straßenverkehr." Annual Report, Wiesbaden, October 2014.

[3] L. Zamboni, L. Gamberini, A. Spagnolli, S. Cipolletta, G. De Giuli, and I. Tion, "Serious games in social intervention: Designing technologies to promote safe and healthy behaviors," in *Proceedings of the 9th ACM SIGCHI Italian Chapter International Conference on Computer-Human Interaction: Facing Complexity*, CHItaly, (New York, NY, USA), pp. 139–142, ACM, 2011.

[4] R. P. McMahan, A. J. D. Alon, S. Lazem, R. J. Beaton, D. Machaj, M. Schaefer, M. G. Silva, A. Leal, R. Hagan, and D. A. Bowman, "Evaluating natural interaction techniques in video games," in *3D User Interfaces (3DUI), 2010 IEEE Symposium on*, pp. 11–14, IEEE, 2010.

[5] The Drinkaware Trust, "drinkaware.co.uk–for the facts." Published online https://www.drinkaware.co.uk/, June 2015.

[6] Bund gegen Alkohol und Drogen im Straßenverkehr, "Alkoholfahrsimulatoren." Published online http://www.bads.de/agieren/alkoholfahrsimulator/, April 2015.

[7] I.-K. Hong, J.-B. Ryu, J.-H. Cho, K.-H. Lee, and W.-S. Lee, "Development of a driving simulator for virtual experience and training of drunk driving," in *3rd International Conference on Road Safety and Simulation*, 2011.

[8] J. R. Parker, N. Sorenson, N. Esmaeili, R. Sicre, P. Gil, V. Kochlar, L. Shyba, and J. Heerema, "The booze cruise: Impaired driving in virtual spaces.," *IEEE Computer Graphics and Applications*, vol. 29, no. 2, pp. 6–10, 2009.

[9] T. Suominen, M. Luimula, A. Peltola, A. Autio, and P. Ketola, "Dui simulator." Published online http://globalgamejam.org/2014/games/dui-simulator, 2014.

[10] M. Shaker, N. Shaker, M. Abou-Zleikha, and J. Togelius, "A projection-based approach for real-time assessment and playability check for physics-based games," in *Applications of Evolutionary Computation*, pp. 430–442, Springer, 2015.

[11] Bundeszentrale für gesundheitliche Aufklärung, "Kenn dein Limit." Published online http://www.kenn-dein-limit.info/, October 2014.

[12] R. Koster, *Theory of fun for game design*. O'Reilly Media, Inc., 2013.

[13] Oculus VR, Inc., "Oculus rift - virtual reality headset for immersive 3d gaming." Published online https://www.oculus.com/rift/, January 2015.

[14] IGN Entertainment, Inc., "Wii remote." Published online http://www.ign.com/wikis/wii-u/Wii_Remote, December 2012.

Junk-Food Destroyer: Helping adolescents with Down syndrome to understand healthy eating through serious game

Paul D Hatzigiannakoglou
Department of Educational and Social Policy
School of Social Sciences, Humanities and Arts
University of Macedonia
Thessaloniki, Greece
Email: pxatzi@uom.gr

***Abstract*—One research study carried out in adolescents with Down syndrome (DS) indicated that they consume fewer calories, vitamins and trace elements than they ought to according to the Recommended Daily Intake. A different study pointed out that people with Down syndrome start using computers from early childhood, displaying a liking for educational software, as well as for Video Games. In an attempt to make the most of this motivating occupation, and, at the same time, implement the positive effects resulting from playing Serious Games (SG) that promote a healthy diet to non-disabled children, I designed and developed a Serious Game of First Person Shooting (FPS) type (played either with a Wii-Remote or a mouse), which intends to help behavior change of adolescents with Down syndrome, so that they would adopt a more balanced nutrition. To avoid further low-calorie intake, the game presents only the benefits resulting from healthy food and beverage consumption compared to unhealthy ones, without mentioning quantity or caloric value.**

Keywords*— serious games, intellectual disability, health behavior, diet, nutrition

I. Introduction

A research [1] that was carried out in a sample of 49 persons with Down syndrome (DS) in late adolescence or older ages, who lived at home with their parents or guardian, showed that 32.1% of males' and 66.7% of females' weight exceeds by 15% the ideal body weight; further, 25.0% of males and 42.9% of females are obese. From the same study it becomes evident that people with DS, compared to the general population, consume fewer calories and less cholesterol and saturated fat. Another research study [2] demonstrated that, besides consuming fewer calories, people with DS also take 20% lower amounts of vitamins and trace elements than the Recommended Daily Intake (RDI). Increased body weight in people with DS does not seem to be directly related to their nutrition [1, 13], but rather to the nature of the syndrome itself, a compromised fat metabolism and the way of life in general. Other authors [3, 4] suggest that children and adolescents with DS should adopt a low-fat balanced nutrition. Such a recommendation led me to create a Serious Game (SG) which would emphasize the benefits of balanced nutrition compared to junk food. The two strong motivating factors for designing my game were the scientifically recorded beneficial effect of similar games on people with Intellectual Disability (ID) [5, 31, 33], along with the fact that numerous SG were thus far developed for such individuals [6]. While designing and developing my SG, I kept in mind that although the game is educational, it has to be entertaining as well, in order to yield an efficient training [5]. That realization led me to choose a widely spread type of game (First Person Shooting, FPS), particularly if adolescents were to be involved. To enhance the fun, I gave users the option to play the game with a Wii-Remote. In case they do not own a Wii-Remote, they can play it by using a mouse, since it is common knowledge that people with DS, even at very young ages, are absolutely capable of using a mouse to play Video Games or SG on the computer [11, 12].

II. Literature Review

In reviewing previous data for the scope of the present paper, I examined the results of three review articles in the field of SG and health in general.

According to Wattanasoontorn [14], 108 SG have been created regarding health and health care from 2004 to 2012. This research proposed the following subcategories:

- Health Monitoring
- Detection focuses on analysis or tracing of irregular symptoms of the patient(s).
- Treatment or therapy is used to remedy a health problem
- Rehabilitation is a restoration of health and life skills after illness, such as neuropsychological rehabilitation
- Education in health/self-directed care

Because the SG I developed is classified in the subcategory Education in health/self-directed care, I filtered the conclusions

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

of my review and I only present those SG that belong to this subcategory.

Wattanasoontorn's research [14] established that nine SG (Table I) that concern nutrition and related diseases were developed. More specifically, three of them dealt with obesity, five involved diabetes and one discussed nutrition.

TABLE I

Author	Disease
Fuchsclocher [16]	Diabetes
Nauta & Spil[17]	Diabetes
The Diablotines[18]	Diabetes
Diehl [19]	Diabetes
Archimage[20]	Diabetes
Scarle et al. [21]	Obesity
Fatworld.org[22]	Obesity
HopeLab[23]	Obesity
Janomedia[24]	Nutrition

Besides, Ricciardi [15] found six SG (Table II) that belong to this subcategory, four related to diabetes and two pertaining to healthy eating.

TABLE II

Author	Disease
Guillaume [25]	Diabetes, dietary terms, insulin injection, and physics activity
Guillaume [26]	Diabetes, pump treatment
Guillaume et al. [27]	Diabetes, dietetic knowledge
Diehl [28]	Diabetes, insulin therapy
Inglés-Camats [29]	Healthy eating habits
Thompson [30]	Increase fruit and vegetable consumption

Moreover, research conducted by Papastergiou [33] in 2008, presented, among others, five SG (Table III). Four of them discuss nutrition and the fifth associates nutrition with obesity.

TABLE III

Author	Disease
Turnin [34]	Nutrition
Cullen, Watson, Baranowski, Baranowski, and Zakeri [35]	increasing Fruit, Juice and Vegetable intake.
Baranowski [36]	increasing Fruit, Juice and Vegetable intake.
Munguba, Valdes, and da Silva [37]	Nutrition, obesity
Silk [38]	Nutrition

Because none of the above listed software had people with ID as a target group, I focused on SG that fall under the category Education in health/self-directed care and have as end users people with ID. After thorough research, I found articles that refer to SG under the topics "Healthy eating", "Nutrition" and "Diabetes". I searched the international online bibliographic databases Institute of Electrical and Electronics Engineers (IEEE), PubMed, and Science Direct. The search key words were: "serious game", "intellectual disability", "healthy eating", "nutrition", and "diabetes". In that search, I only found one relevant SG [10]. Apparently, this finding supports my decision to create my own SG.

III. GAME DESIGN

Because end users of SG are adolescents with DS, at the stage of design and development I had to take into account a series of factors that would make accessibility and use of the game easier for this particular group of users. My first and foremost concern was the fact that children with DS present a unique differentiation in comparison with children of the general population or children with other developmental disabilities [11]. Their distinct behavior lies in the fact that each child with DS may show a varying degree of cognitive, motor, and perceptual disabilities [11].

Still, despite the encouraging fact that the above mentioned disabilities are usually mild, the issue that the child's computer usage is subject to the combined impact of limitations in multiple channels, making it more difficult to accommodate than those cases where only one single channel is affected, set an obstacle to my project [11].

Eventually, I proceeded to the implementation of the project, as the very same research that was discussing the disabilities of children with DS, and another similar study showed that people with DS become capable computer users from early childhood, use the mouse effectively, can run educational software and play games on the computer [11, 12]. I selected First Person Shooting game type for two main reasons. Firstly, FPS is one of the most popular types of games among adolescents. Secondly, Point-and-click type (which includes FPS) is considered to be an accessible type of game for people with cognitive disabilities [32].

Another point that strengthened my decision was the idea of substituting mouse click for a switch device and, thus, expanding the accessibility of the game by motor-impaired persons [32].

My next concern was the selection of the means of communication between software and user, that is, whether I should use text, voice, sound, images or symbols. I finally decided to rule out text and to use vocal instructions and images, because children with DS face intense problems with text reading and, as research results indicate, they often get disappointed and abandon the use of software just because they do not comprehend the instructions [11]. Parents of children with DS noticed that there are four factors [11] that cause frustration and disappointment to their children and, in the end, lead them to give up. These factors are the following:

Navigation. Navigation throughout computer software or on the internet is really challenging for these children [11]. They find it rather difficult to use menus and to select an option. In my software there is no need for the user to navigate. From the startup, the user is being transferred automatically from one level to the next. When one reaches the last stage, one simply has to shoot a hamburger in order to start over.

Trouble shooting. Children with DS demonstrate zero tolerance for any kind of trouble that might arise while using

Back to Contents

software [11]. They hardly understand the problem, let alone solve it. The technical problems that might appear are often complicated and in many cases they do not occur because of the software itself, but due to other factors, such as hardware, operation system or even other software. While developing my SG, I repeatedly tested it by running it on different computers with different hardware configurations and different Windows versions without noticing any problem. However, to minimize the occurrence of trouble these trials had to be done on a larger scale at this stage.

Lack of patience. Impatience is a common trait of children with DS [11]. They cannot deal with software that delays at start-up and they fail to play games or educational software which involve questions with no obvious answers. User's instructions on the software have to be simple and brief. Another hindrance factor in using computers is the difficulty they face with the keyboard. They prefer using the mouse, most likely because they can do it easily and more directly. In my effort to avoid testing the users' tolerance level I managed to reduce both start-up time and transition time between stages as much as possible. The game can be played either with a Wii-Remote or a mouse without using a keyboard at all. Instructions on the use of the software are verbal, quite simple, and short.

Design flaws. Parents of children with DS have observed that the main consequence of design flaws is that they cause confusion when they appear suddenly [11]. Some typical examples of design flaws are when a function key produces a different result when used in different software or when the information provided to the user is not relevant to the content of the software, or when there is not a clear distinction between user's instructions and software feedback.

For all those reasons I completely rejected the idea of using the keyboard and preferred to use the mouse or the Wii-Remote. I tried to ensure that verbal instructions in each stage are strictly related to its content. Additionally, I created such a feedback that would make evident the direct reference to the user's current action.

I proceeded to the second phase of design by adopting those attributes of SG that would boost behavior change in the users [7]. The features I embedded are interactivity, feedback, characters, story, and sound effects [8, 9].

Interactivity. By default video games are interactive and allow the user to readily acquire applicable knowledge which can be used to achieve the best possible result, e.g. greater score [8, 9]. The vocal/sound performance feedback seemed to me as the most appropriate because, as shown by research, this kind of feedback is getting easily perceived by users [9]. In my game the user, after listening to vocal instructions on every action during the game, one receives immediate vocal feedback that is rewarding (Good job!) or informs someone about one's mistake (error sound).

Characters. In my game I adopted the hero model in order to stimulate the user to imitate him and help to change the dietary behavior, while keeping the game challenging and interesting. The game presents an adolescent in two different situations. In one scene he looks leisurely, obese and he likes to eat unhealthy food; in the other, he is fit and he prefers to consume healthy food.

Story. The theme of the game was selected with the expectation that end users would find it fascinating, and by developing emotional ties with the hero, they would consequently try to help him consume healthy food.

Sound effects. The sound effects convey meaning and direction without words. In my software, apart from images and verbal instructions, I use sounds through which users become aware of their mistakes. Also, I have included sounds of blast or gun shooting for greater amusement.

IV. THE GAME

My project is a Cartoon-Style Game. To minimize the cognitive load I created simple-layout stages that consist of few elements and contain the least animation elements possible. Additionally, to avoid the case of causing an epileptic seizure to the user, I have excluded the usage of flashing elements [5].

Trying to promote healthy nutrition almost for all meals throughout the day and to stress its beneficial results on children's health, I created four mini games (Breakfast, Snack, Lunch, Beverages). I decided to include Beverages in order to emphasize the importance of proper hydration for the body.

Every mini-game begins with the presentation of a group of healthy food/beverages in contrast to a group of unhealthy food/beverages. Following this presentation, the user is asked to participate actively.

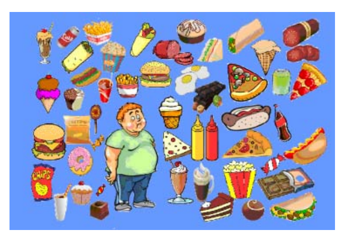


Fig. 1

The story. The game starts with a two-stage Introduction that presents the story of the game and introduces the hero. The hero is an obese adolescent surrounded by junk food (see Fig.1) and all sort of food that could prove to be harmful to his health.

Next, the same adolescent is shown with normal weight, surrounded by food (see Fig.2) that is considered to be healthy. During these two stages a voice is heard describing how food can affect the adolescent's health.

Fig. 2

Getting acquainted with the environment. In the following scene the user takes action. At random points of the screen, there appear items of unhealthy food, which one has to point at and destroy. The aim of this stage is to familiarize the user with the environment and the sounds of the game and, on the other hand, to train him/her at hitting the targets with the Wii-Remote or the mouse. The stage comes to an end when the user succeeds in destroying 60% of the targets and not because one has run out of time, a situation that can cause anxiety and frustration.

In the mini games that follow I adopted the Theory-Action pattern. At the beginning, healthy food items from each group (Breakfast, Snack, Lunch, Beverages) are projected as images (see Fig.3) and described vocally.

Fig. 3

Then the user takes action. On the screen appear images of healthy and unhealthy food in pairs, while at the bottom part of the screen the adolescent appears in both versions (fit - obese) (see Fig.4).

The user is asked via vocal instruction to identify and shoot the unhealthy food. For each group of food or beverages, there appear a total of 13 pairs. As soon as the user makes a choice, he/she receives immediate vocal/sound feedback which informs one whether the action was right or wrong. To provide more intriguing feedback to the user I added a visual cue (see Fig. 4). Every time the user destroys an item of unhealthy food, the fit boy gets taller and his score increases. Accordingly, the same thing happens with the obese boy whenever the user destroys an item of healthy food. At the end of each mini-game, the score appears. If the player successfully destroyed at least 7 unhealthy food items, one is awarded a prize; otherwise one hears a vocal message that informs him of the mistake and at the same time encourages one to continue.

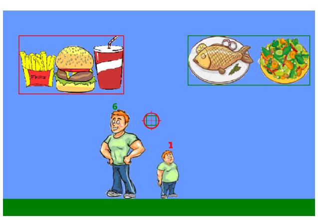


Fig. 4

The order of appearance of the mini-games is Breakfast, Snack, Lunch, Beverages. Although the above scenes are accompanied by sounds of explosion, gun-shooting and explosion animation, they are considered as static, compared to modern action games, because they lack motion and the user has as much time as needed to reach the final stage. To prevent possible disappointment and boredom, I created two additional scenes, one between the mini-games and one at the end of them. These stages have a lot of action, many sound, and animation elements. Unlike the four mini games, the user has to finish them within a given time limit.

Acknowledgment

I would like to express warmest thanks to my colleague Catherine Notaridou for the amount of time and work she contributed in editing this article.

References

[1] Fujiura, G. T., Fitzsimons, N., Marks, B., & Chicoine, B. (1997). Predictors of BMI among adults with Down syndrome: the social context of health promotion. *Research in Developmental Disabilities*, *18*(4), 261-274.

[2] Luke, A., Sutton, M., Schoeller, D. A., & Roizen, N. J. (1996). Nutrient intake and obesity in prepubescent children with Down syndrome. *Journal of the American Dietetic Association*, *96*(12), 1262-1267.

[3] Pueschel, S. M. (1990). Clinical aspects of Down syndrome from infancy to adulthood. *American Journal of Medical Genetics*, *37*(S7), 52-56.

[4] Bull, M. J. (2011). Health supervision for children with Down syndrome. *Pediatrics*, *128*(2), 393-406.

[5] Lanyi, C. S., & Brown, D. J. (2010). Design of serious games for students with intellectual disability. *IHCI*, *10*, 44-54.

[6] Saridaki, M., Gouscos, D., & Meimaris, M. (2009). Digital Games-Based Learning for Students with Intellectual Disability. Game-Based Learning Advancements for Multi-Sensory Human Computer Interfaces: Techniques and Effective Practices. Information Science Reference Publications.

[7] Thompson, D. (2014). What serious video games can offer child obesity prevention. *JMIR Serious Games*, *2*(2), e8.

[8] Thompson, D., Baranowski, T., Buday, R., Baranowski, J., Thompson, V., Jago, R., & Griffith, M. J. (2008). Serious video games for health: how behavioral science guided the design of a game on diabetes and obesity. *Simulation & Gaming, August 2010; vol. 41, 4: pp. 587-606., first published on December 24, 2008*

[9] Baranowski, T., Buday, R., Thompson, D., Lyons, E. J., Lu, A. S., & Baranowski, J. (2013). Developing games for health behavior change: Getting started. *GAMES FOR HEALTH: Research, Development, and Clinical Applications*, *2*(4), 183-190.

[10] Rodriguez Isasi, A.; Lopez Basterretxea, A.; Mendez Zorrilla, A.; Garcia Zapirain, B., "Helping children with Intellectual Disability to understand healthy eating habits with an IPad based serious game," *Computer Games: AI, Animation, Mobile, Interactive Multimedia, Educational & Serious Games (CGAMES), 2013 18th International Conference on* , vol., no., pp.169,173, July 30 2013-Aug. 1 2013

[11] Feng, J., Lazar, J., Kumin, L., & Ozok, A. (2010). Computer usage by children with down syndrome: Challenges and future research. *ACM Transactions on Accessible Computing (TACCESS)*, *2*(3), 13.

[12] Feng, J., Lazar, J., Kumin, L., & Ozok, A. (2008, October). Computer usage by young individuals with down syndrome: an exploratory study. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 35-42). ACM.

[13] A guidebook for Teens and Adults with Down Syndrome. “http://services.medicine.uab.edu/publicdocuments/genetics/ADSC/living_healthy_booklet.pdf”

[14] Wattanasoontorn, V., Boada, I., García, R., & Sbert, M. (2013). Serious games for health. *Entertainment Computing*, *4*(4), 231-247.

[15] Ricciardi, F., & De Paolis, L. T. (2014). A Comprehensive Review of Serious Games in Health Professions. *International Journal of Computer Games Technology*, *2014*.

[16] Fuchslocher, A., Niesenhaus, J., & Krämer, N. (2011). Serious games for health: An empirical study of the game “Balance” for teenagers with diabetes mellitus. *Entertainment Computing*, *2*(2), 97-101.

[17] Nauta, H., & Spil, T. A. (2011, November). Change your lifestyle or your game is over: The design of a serious game for diabetes. In *Serious Games and Applications for Health (SeGAH), 2011 IEEE 1st International Conference on* (pp. 1-7). IEEE.

[18] The Diablotines, L’ Affaire BIRMAN 2011. ‘‘http://www.gluciweb.com/’’accessed 20 November 2012.

[19] Diehl, L. A., Lehmann, E., Souza, R. M., Alves, J. B., Esteves, R. Z., & Gordan, P. A. (2011, November). A serious game prototype for education of medical doctors and students on insulin management for treatment of diabetes mellitus. In *Serious Games and Applications for Health (SeGAH), 2011 IEEE 1st International Conference on* (pp. 1-4). IEEE.

[20] Archimage, Inc., Escape from Diab. <http://www.escapefromdiab.com/>, 2006 (accessed 04.01.2013).

[21] Scarle, S., Dunwell, I., Bashford-Rogers, T., Selmanovic, E., Debattista, K., Chalmers, A., ... & Robertson, W. (2011, May). Complete motion control of a serious game against obesity in children. In *Games and Virtual Worlds for Serious Applications (VS-GAMES), 2011 Third International Conference on* (pp. 178-179). IEEE.

[22] Fatworld.org (2007) Fatworld. ‘‘http://fatworld.org/’’, accessed 26 March 2012.

[23] HopeLab (2010) Zamzee: A game that gets families moving. ‘‘https://www.zamzee.com’’, accessed 20 November 2012.

[24] Janomedia (2006) Terveellinen ateria. ‘‘http://www03.edu.fi/oppimateriaalit/healthy_meal’’, accessed 26 March 2012.

[25] A. Guillaume, “Affaire birman”, http://www.zippyware.com/

[26] A. Guillaume, “Meli-melo glucique”, http://www.zippyware.com.

[27] A. Guillaume et al., “Timeout”, http://www.zippyware.com/.

[28] Diehl, L. A., Souza, R. M., Alves, J. B., Gordan, P. A., Esteves, R. Z., Jorge, M. L. S. G., & Coelho, I. C. M. (2013). InsuOnline, a serious game to teach insulin therapy to primary care physicians: design of the game and a randomized controlled trial for educational validation. *JMIR Research Protocols*, *2*(1).

[29] Inglés-Camats, G., Presno-Rivas, M. M., Antonijoan, M., Garcia-Panella, O., & Forrest, T. (2011). Yummy tricks: a serious game for learning healthy eating habits. *Studies in Health Technology and Informatics*, *172*, 185-190.

[30] Thompson, D., Bhatt, R., Lazarus, M., Cullen, K., Baranowski, J., & Baranowski, T. (2012). A serious video game to increase fruit and vegetable consumption among elementary aged youth (Squire’s Quest! II): Rationale, design, and methods. *JMIR Research Protocols*, *1*(2).

[31] Martins, T., Carvalho, V., Soares, F., & Moreira, M. F. (2011, November). Serious game as a tool to intellectual disabilities therapy: Total challenge. In *Serious Games and Applications for Health (SeGAH), 2011 IEEE 1st International Conference on* (pp. 1-7). IEEE.

[32] Torrente, J., Del Blanco, Á., Moreno-Ger, P., & Fernández-Manjón, B. (2012, January). Designing serious games for adult students with cognitive disabilities. In *Neural Information Processing* (pp. 603-610). Springer Berlin Heidelberg.

[33] Papastergiou, M. (2009). Exploring the potential of computer and video games for health and physical education: A literature review. *Computers & Education*, *53*(3), 603-622.

[34] Turnin, M. C., Tauber, M. T., Couvaras, O., Jouret, B., Bolzonella, C., Bourgeois, O., ... & Hanaire-Broutin, H. (2008). Evaluation of *microcomputer nutritional teaching games in 1,876 children at school.* Diabetes & metabolism, 27(4 Pt 1), 459-464.

[35] Cullen, K. W., Watson, K., Baranowski, T., Baranowski, J. H., & Zakeri, I. (2005). Squire's Quest: intervention changes occurred at lunch and snack meals. *Appetite*, *45*(2), 148-151.

[36] Baranowski, T., Baranowski, J., Cullen, K. W., Marsh, T., Islam, N., Zakeri, I., ... & demoor, C. (2003). Squire’s Quest!: Dietary outcome evaluation of a multimedia game. *American Journal of Preventive Medicine*, *24*(1), 52-61.

[37] Munguba, M. C., Valdes, M. T. M., & da Silva, C. A. B. (2008). The application of an occupational therapy nutrition education programme for children who are obese. *Occupational Therapy International*, *15*(1), 56-70.

[38] Silk, K. J., Sherry, J., Winn, B., Keesecker, N., Horodynski, M. A., & Sayir, A. (2008). Increasing nutrition literacy: Testing the effectiveness of print, web site, and game modalities. *Journal of Nutrition Education and Behavior*, *40*(1), 3-10.

Back to Contents

Distinction of Intrinsic and Extrinsic Stress in an Exercise Game by Combining Multiple Physiological Indices

Yoshimasa Ohmoto
Kyoto University
Yoshida-Honmachi, Sakyo-ku
Kyoto 606-8501 Japan
Email: ohmoto@i.kyoto-u.ac.jp

Seiji Takeda
Kyoto University
Yoshida-Honmachi, Sakyo-ku
Kyoto 606-8501 Japan
Email: stakeda@ii.ist.i.kyoto-u.ac.jp

Toyoaki Nishida
Kyoto University
Yoshida-Honmachi, Sakyo-ku
Kyoto 606-8501 Japan
Email: nishida@i.kyoto-u.ac.jp

***Abstract*—In a serious game, agent characters that interact with players often appear on the game scene. These characters are usually not regarded as expert teachers or familiar friends. To develop a reliable expert teacher agent, students' inner states can be estimated for providing efficient training individually. As a first step, the present study investigated clues to distinguish whether the cause of changes in a human's activities are intrinsic factors or extrinsic stimuli based on physiological indices and game context. We conducted an experiment using an exercise game in which participants either provided spontaneous strategic behavior or followed a communication partner's advice as a result of continuous interaction. We measured physiological indices and compared the responses. As a result, significant differences were found in the responses of the physiological indices. We can suggest that it is possible to distinguish intrinsic and extrinsic stress based on the physiological responses and game context.**

***Keywords*—Human-agent interaction, physiological indices, stress estimation, exercise game.**

I. Introduction

The use of serious games using a virtual world is spreading widely in education, medical services, welfare and fitness. By using virtual games, we can conduct rehearsals or practice in specialized situations, extra-ordinary environments, and circumstances where no mistakes can be allowed. In addition, an individual's motivation to practice hard can be improved, as although serious, the game includes an entertainment aspect. In a serious game, agent characters that interact with players often appear on the game scene. These characters are regarded as multi-modal interfaces that provide useful information; and are not regarded as expert teachers or familiar friends. An expert teacher agent needs to be able to bring about desirable effects that the user cannot achieve in the training alone. For example, in an actual fitness club, trainers often estimate students' motivation from their facial expressions, nonverbal behavior, exercise heart rate and sharpness of body motion. This fitness support, including inner state monitoring, increases the efficiency of training and also improves the reliability of the trainers. It is important to provide adaptive game content through the estimation of players' inner states, including mental stress and physical loads. The estimation of stress is also useful to evaluate the effect of a serious game.

However, it is difficult to estimate inner state (including mental stress and physical load) by subjective reports and observations. In recent years, progress has been made in developing microelectromechanical systems, and wearable devices that can measure a human's physiological indices (e.g., wristwatches that measure heart rate and shirts that obtain electrocardiogram measurements) appear in daily life. In the present study, we aimed to estimate human inner state by synthetic use of physiological indices and analysis of game context in continuous tasks.

The long-term goal of our research is to develop a reliable expert teacher agent that can estimate students' inner states and provide efficient training individually adapted to each student depending on the daily conditions in a long-term task with continuous interaction. As a first step in achieving this goal, the present study investigated a method to distinguish whether the cause of changes in a human's activities are intrinsic factors related to spontaneous mental activities, or extrinsic stimuli generated by circumstances in continuous interaction, based on physiological indices and game context. These results could be used to evaluate the adequacy of an interaction agent's behavior and the difficulty of the task set for the user.

II. Related works

It is difficult to estimate human internal states only by behavior observations. We used physiological indices for estimating human inner states during interaction. There are various studies on estimating the states by measuring physiological indices (e.g. [1]). There are also several studies that use measured physiological indices for effective human-agent interaction.

Lin et al. [2] reported physiological data correlate with task performance data in a video game: with a decrease of the task performance level, the normalized SCR increases. Tanaka and Kawahara [3] reported two participants' heart-rate fluctuations were corresponding according to progress the game. These indicated that the physiological indices were useful to estimate participants' stress in serious games.

Nikolic-Popovic and Goubran [4] measured electrocardiogram, respiration, and skin conductance in a set of five isometric exercises. It was shown that there is a correlation between motion and the physiological indices. Masuko and Hoshino [5] used heart rate to properly set the degree of difficulty of an exercise game. The results showed that maximum effective exercise that elevated heart rate above the effective heart rate was effective in improving the sense of accomplishment. This means that the control of serious game scenario based on the physiological indices is useful for effective learning.

As mentioned above, physiological indices are helpful to estimate inner states of people even in exercise. However, the previous works detected whether the (inner) conditions were changed or not, but the sources of the stress were not estimated. In this study, we used skin conductance responses (SCR) and electrocardiograms (LF/HF values) to detect mental stress and tried to estimate the sources of the stress in a virtual exercise game with continuous interaction.

III. Estimation of mental stress in continuous interaction

In previous work, we have investigated continuous interaction in decision-making tasks (e.g. [6]). In the investigation, the distinction of the stress source is important for evaluation of the interaction and learning effects. It is useful to distinguish types of stress: intrinsic stress arising from spontaneous mental activity and extrinsic stress generated by outside stimuli such as other people and the environment. When intrinsic stress is identified, we can determine whether the interaction behavior is spontaneous or not and whether the recent action was caused by the target interaction or not. When extrinsic stress is identified, we can determine the influence and adequacy of the interaction partners' behavior. The present study aimed to identify clues to distinguish intrinsic and extrinsic stress in a continuous interaction using physiological indices.

A. Physiological indices used to estimate stress

In continuous interaction, physiological responses are often found in unexpected situations. A reason for this is that stress responses are sometimes made in advance, or the source of stress is ignored because people in continuous interaction often plan their actions and pay attention to the circumstances. Therefore, the physiological responses made at a time other than when extrinsic stimuli are applied cannot be simply regard as intrinsic stress. We investigated the method of distinguishing intrinsic and extrinsic stress in continuous interaction using SCR and LF/HF values, and estimated the person's inner state based on the interaction context.

As the underlying mechanisms of SCR and LH/HF are different, we expected to find different responses from different sources of stress. Emotional perspiration can also be elicited by emotional stimuli, intellectual strain or painful cutaneous stimulation. Mechanisms underlying SCR are related to anticipation, expectation and concentration of attention [7]. Therefore, we expected that SCR could be observed when people anticipated dealing with an unexpected or thrilling situation. This means that SCR would relatively reflect intrinsic stress. The LF/HF value was calculated from instantaneous heart rate. Heart rate variability is controlled by the sympathetic and parasympathetic nervous systems. Lacey and Lacey [8] suggested it was caused by sensory intake and sensory rejection. Therefore, we expected LF/HF to show reactive responses based on external stimuli. This means the LF/HF value would relatively reflect extrinsic stimuli.

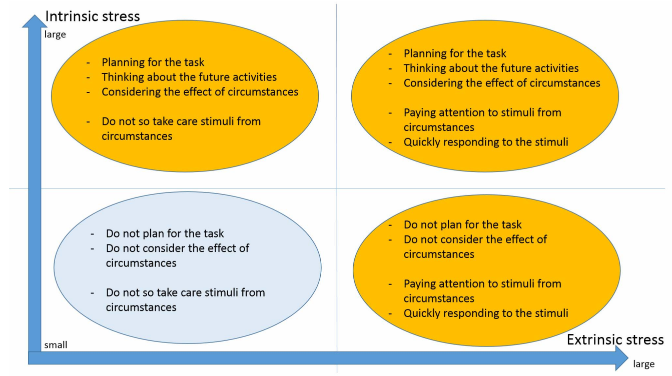


Fig. 1: The estimation model of inner state of a person.

B. Inner state estimation model

We conducted a preliminary experiment to confirm the response characteristics of SCR and LF/HF. We determined that SCR responses could be found when participants considered a calculation task and a puzzle task in exercise. In addition, we confirmed that LF/HF responses could be found in an exercise task and that the influence of exercise on LF/HF responses was limited. However, when we use the responses for human-agent interaction, we have to consider the human inner states.

We expected that the effect of stresses would be different depending on the inner state of a person when they felt stress. To estimate the inner state, we proposed a brief inner state estimation model using an exercise task (Figure 1). If the source of extrinsic stress is not changed, the different responses can be attributed to the difference in intrinsic stress. By analyzing these differences, we expected to find clues to distinguish intrinsic and extrinsic stress based on physiological indices and interaction behavior in continuous interaction.

IV. Experiment

In this experiment, we investigated whether the physiological responses to a task after provision of advice differed depending on the participant's inner state at the time the advice was given. When a difference was found, we investigated whether the source of stress (intrinsic or extrinsic) could be distinguished based on physiological responses. We compared the physiological responses to two types of agents.

A. Task

Each participant performed a virtual exercise game, playing with two rival characters. An advice agent that supported game play accompanied the participant. The game was an action video game in which three players (a participant and two automated game characters) compete for points which can be won in two different ways; carrying a teddy bear character to point zones or hitting a virtual ball at other players.

Players could spontaneously decide a strategy to maximize their points. The players had to consider their own strategy depending on the strategies of the other players and their own

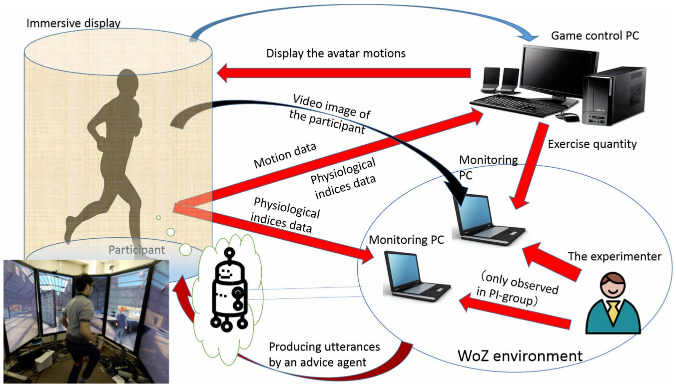


Fig. 2: The experimental setting.

state. In the game, a main source of intrinsic stress was strategy planning, and sources of extrinsic stress were game actions and reactions, such as throwing or being hit by a ball, chasing a teddy bear and encountering other players.

B. Experimental setting

The experimental setting is shown in Figure 2. We used Immersive Collaborative Interaction Environment (ICIE) [9] to easily look around in the virtual space similar to the real world. The participant's virtual avatar could be intuitively controlled by their body motions using motion sensors placed on their dominant arm, both feet and waist. To determine the participant's inner state, exercise quantity was estimated from the stepping motion and SCR and LF/HF were recorded by a device for measuring the physiological indices. This information was sent to the game system in real time.

First of an experiment, the participant was instructed on the experimental procedures. Electrodes for measuring SCR and LF/HF values were attached to the participant. After 2-minute relaxation period, the participant performed a practice session of the game. The participant was introduced the game rules and how to control the avatar by an advice agent in the practice session. This session worked as an icebreaker. After that, they intermittently performed three game sessions. Each game session lasted 8 minutes, with 2-minute intervals between each game session. In the interval, the advice agent gave opinions about the previous session. The interactions were useful to increase perceived reliability of the agent's advice.

C. Advice agents

Two types of advice agents were implemented: one was an exercise quantity base agent (EQ-agent) that provided advice related to game strategy when exercise quantity reduced to a certain level, and the other was a physiological indices based agent (PI-agent) that provided advice related to game strategy when there were no SCR and LF/HF responses. Advice was provided in the first half of the second and third sessions, as we expected that the effect of the advice was reduced after the interval. One type of agent accompanied each participant.

As participants were refreshed after each interval and received help from an advice agent in the first half of the second and third sessions, we determined that the reduced

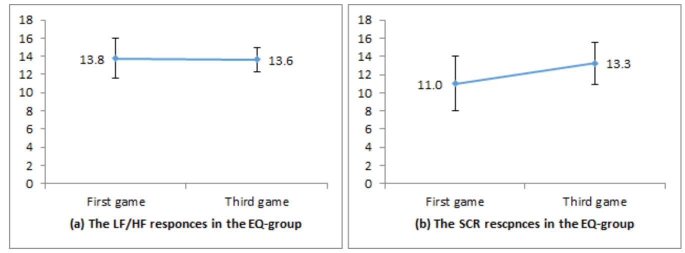


Fig. 3: The physiological responses in the EQ-group.

exercise quantity was related to the participant's inner state becoming lightly absorbed in the game. Proposing a new game strategy in this situation therefore would encourage absorption. Participants who received this kind of advice in the situation provided higher responses in both intrinsic and extrinsic stress.

When there were no SCR and LF/HF responses, the participant experienced no stress. No stress after one or two game sessions indicated that the participant had settled on a game strategy after trial and error. The proposal of a new game strategy in this situation encouraged continuous game play, but only if the participant followed the proposed action. This suggested that participants who received new game strategy advice provided higher responses in extrinsic stress but lower responses in intrinsic stress.

Twenty students (13 males and 7 females) participated in the experiment. All participants were undergraduate students (average of 21.9 years old). Ten participants (7 males and 3 females) played the game with the EQ-agent (EQ-group) and the remainder played the game with the PI-agent (PI-group).

D. Results

The purpose of this analysis was to investigate any differences in physiological responses between the EQ-group and the PI-group, and determine if it was possible to distinguish intrinsic and extrinsic stress based on physiological responses and game situation. To explore this, the numbers of SCR and LF/HF responses were counted. A peak of the SCR wave above 1.0 from the baseline level was regarded as a SCR response. A peak of the LF/HF wave above 5.0 was regarded as a LF/HF response. We hypothesized that SCR relatively reflected the effect of intrinsic stress and LF/HF relatively reflected the effect of extrinsic stress. Therefore, we expected that in the EQ-group, the number of SCR and LF/HF responses would increase after the advice; and in PI-group, the number of SCR responses would decrease and the number of LF/HF responses would increase after the advice.

Figure 3 and 4 show the numbers of the SCR and LF/HF responses for the EQ- and PI-groups for the first and third game sessions. SCR could not be measured for two participants in the EQ-group and one participant in the PI-group. Data for these participants were eliminated from the analysis.

Paired t-tests were conducted on EQ- and PI-group data to compare the number of responses in the first sessions with those in the third sessions. We obtained the following results: First, for the EQ-group, the number of SCR responses in the first session was marginally significantly less than that in the

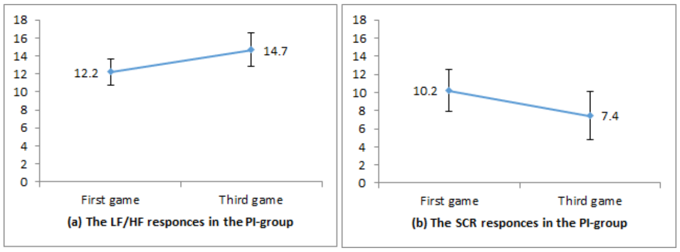


Fig. 4: The physiological responses in the PI-group.

third session (p = 0.065). Second, for the PI-group, the number of LF/HF responses in the first session was significantly less than that in the third session (p = 0.0075), and the number of SCR responses in the first session was significantly more than that in the third session (p = 0.018). In addition, we conducted t-tests to compare the number of responses of the EQ-group with those of the PI-group in the first and third sessions. The results showed that the PI-group had significantly more SCR responses than the EQ-group in the third session (p = 0.00042).

From these results, we can find the differences in the physiological responses. The SCR responses changed inversely with each other, with the difference in the experimental conditions being the inner state of the participant when the agent provided advice. This indicates that the effect of the advice differed depending on the inner state of the participant; and suggests that advice from the agent cannot create adequate learning effects without appropriate evaluation of a person's inner state.

V. Discussion

Notably, there were different changes in SCR and LF/HF responses in each group. In the PI-group, SCR and LF/HF responses changed inversely, whereas in the EQ-group, SCR responses increased but no tendency was seen in LF/HF responses. This suggests that it is possible to distinguish intrinsic and extrinsic stress based on physiological responses and game context. In addition, these results were as we expected, with the exception of LF/HF responses in the EQ-group. We therefore propose that the inner state estimation model described in the previous section is appropriate to some extent.

We also asked participants to rank the game session depending on subjective absorption. As a result, we could find that that the ranking of subjective absorption increased in the second and third sessions in the PI-group, but there was no tendency in the EQ-group. From the results of these and the physiological analysis, subjective reports of absorption mainly reflected on LF/HF responses. We inferred that the participants regarded the state of paying attention to extrinsic stimuli as being absorbed. A previous study reported that SCR did not change when participants became absorbed in a task. This suggests that when participants are absorbed in a game, they automatically respond to extrinsic stimuli and do not consider the possibilities of various patterns of interaction. When we take into account these results, absorption may not always indicate a good state for learning and training situations.

VI. Conclusions

In the present study, we investigated clues to distinguish if various types of stress sources in continuous interaction are intrinsic factors or extrinsic stimuli, based on physiological indices and game context. We conducted an experiment to measure SCR and LF/HF values in an exercise game with continuous interaction. As the mechanisms underlying SCR and LF/HF are different, we expected to find different responses from different sources of stress. We implemented two types of agents to provide advice on game strategies when a participant's inner state was different. Significant differences were found in the responses of the physiological indices and subjective absorption. In addition, we found that it is possible to distinguish intrinsic and extrinsic stress based on the physiological responses and game context.

In future work, we intend to quantify intrinsic and extrinsic stress, and further develop the method to maintain user motivation for the exercise through continuous agent interaction, and implement a method to estimate inner states in general situations. Following this, we aim to develop a reliable expert teacher agent that can estimate students' inner states and provide efficient training adapted to each student in a long-term task with continuous interaction.

Acknowledgment

This research is supported by the Center of Innovation Program from Japan Science and Technology Agency, JST, AFOSR/AOARD Grant No. FA2386-14-1-0005 and Grant-in-Aid for Scientific Research (A) (KAKENHI No. 24240023) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] R. L. Mandryk and K. M. Inkpen, "Physiological indicators for the evaluation of co-located collaborative play," in Proceedings of the 2004 ACM conference on Computer supported cooperative work. ACM, 2004, pp. 102–111.

[2] T. Lin, M. Omata, W. Hu, and A. Imamiya, "Do physiological data relate to traditional usability indexes?" in Proceedings of the 17th Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future. Computer-Human Interaction Special Interest Group (CHISIG) of Australia, 2005, pp. 1–10.

[3] H. Tanaka and T. Kawahara, "Conformity in heart rate variability under online game without conversations," in ICCAS-SICE, 2009. IEEE, 2009, pp. 2121–2124.

[4] J. Nikolic-Popovic and R. Goubran, "Measuring heart rate, breathing rate and skin conductance during exercise," in Medical Measurements and Applications Proceedings (MeMeA), 2011 IEEE International Workshop on. IEEE, 2011, pp. 507–511.

[5] S. Masuko and J. Hoshino, "A fitness game reflecting heart rate," in Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology. ACM, 2006, p. 53.

[6] Y. Ohmoto, M. Kataoka, and T. Nishida, "The effect of convergent interaction using subjective opinions in the decision-making process," in Proc. the 36th Annual Conference of the Cognitive Science Society, 2014, pp. 2711–2716.

[7] K. Hugdahl, Psychophysiology: The mind-body perspective. Harvard University Press, 1995.

[8] B. C. Lacey and J. I. Lacey, "Two-way communication between the heart and the brain: Significance of time within the cardiac cycle." American Psychologist, vol. 33, no. 2, 1978, p. 99.

[9] Y. Ohmoto, M. Kataoka, and T. Nishida, "Extended methods to dynamically estimate emphasizing points for group decision-making and their evaluation," Procedia-Social and Behavioral Sciences, vol. 97, 2013, pp. 147–155.

Cloud Computing - An Educational Game on Weather Phenomena

Julia Rauscher, Katharina Eckardt, Sarah Edenhofer and Sebastian von Mammen

***Abstract*—We experience weather every day, yet many people do not know how different weather phenomena come about. In this work, we present "Cloud Computing", an educational game that instills a meaningful subset of the according relationships and mechanisms. In the game, weather phenomena are animated based on environmental parameters entered by the user. The user is engaged by "weather quests", for example the task to create a thunderstorm. The user improves his highscore and levels up solving such quests. The challenge of setting the parameters rises with an increasing level, but exuberant difficulty is diligently avoided. Intermittent tests and evaluations involving potential users determined the final design of the game. The available scientific data significantly shaped its basic structure as well as its aesthetics. Cloud Computing was thoroughly evaluated by more than thirty persons in order to ensure its effectiveness—both in terms of engagement and learning targets. The survey is included towards the end of this paper.**

Keywords*— *interactive simulation, educational games, serious games, weather phenomena, clouds

I. INTRODUCTION

Weather prediction is of great importance and defines a considerably large domain of research. It is discussed, for example, at the American Meterological Society Conferences [1] and the World Weather Open Science Conference [2]. Only little changes of climate conditions can cause great weather alterations. To interpret these indications properly and to know which weather phenomena will occur, a person has to understand the scientific background of weather formation. Cloud Computing, the educational game presented in this paper, aims at imparting the necessary knowledge. The game's initial idea was influenced by the Pixar short movie "Partly Cloudy" [6]. While watching the movie it appeared that the common belief that clouds shape the weather is widespread. It is a common misconception that clouds are the *cause* for certain weather phenomena [3], whereas their formation is only a visually noticeable side-effect of specific weather phenomena [4]. Cloud Computing highlights this insight and conveys the variables and parameters that can be considered the actual underlying factors. The game is dedicated to and was developed for adults. For scientific research, the level of detail is not appropriate. For children, the presentation of ideas and learning content might be too complex and not appealing.

The game shows islands, which represent different climate zones that correspond to the real world, and the different levels. Each island has an increasing number of quests. This makes each level harder to successfully complete it than the previous level. In the game, the user is encouraged to manipulate certain sliders, e.g. changing the ambient temperature [5], to make animated weather phenomena occur. If the user struggles to create a certain weather phenomenon, hints are available to help him find the right adjustments of the sliders. This paper describes the development of the game as well as the scientific requirements and the background. The research target of this work can be summarised as conveying the influence of different conditions on emergence of certain weather phenomena in the context of an educational game.

The remainder of this paper is structured as follows. In Section II a brief overview of the related work is provided. In Section III, the methodology including the concept idea as well as our agile, user-centred design process are presented. At the end of the paper, the results are discussed in Section IV, the stated claims are supported by means of a user survey, and an outlook to possible future work is given.

II. RELATED WORK

Weather forecasting is a heavily studied research topic. As an example, the Thor Project, founded in 2002, addresses the challenging task of forecasting thunderstorms especially to improve aviation safety [13].

A great number of games already revolve around weather and climatic topics. Yet, only few are based on scientific facts, and, as far as the authors know, there is not a single one that is based on animating weather phenomena. Addressing the animation aspect, however, some titles are similar to the presented approach. In [7], for instance, the weather can be changed and influenced by varying the ambient temperature at the ground and three levels above. According to the settings, the result of the simulation of the precipitation is altered. The simulation differs from Cloud Computing as it does not offer any game mechanics, there are no tasks presented that need solving. Similar to [16] or [17], all weather effects were rendered as particle systems from the Asset Store in Unity3D [15]. A recent discussion of cloud-generation and –rendering, that provides a good survey of this field, is [18]. Another related title is "The Weather Game" [8]. Here, the task is to answer weather-related questions. "The Weather Game" includes no animation and no interaction with weather conditions or phenomena are offered. Computer-based board games like "Wild Weather Adventure" [9] by NASA educate children about weather phenomena and climatic processes, for example by asking questions about components of clouds or prevailing winds in certain regions. The game "Weather Maze" [10] is rather similar to Cloud Computing with respect to solving quests. However, here the user cannot proactively play a role in the emergence of weather phenomena. Furthermore, "Weather Maze" is limited

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

Back to Contents

to weather icons instead of animations. In general, there are many games dealing with the topic weather, especially for children. Creating it as an interactive edutainment game with real weather to be animated is a novel approach. The scientific foundation and the distinct target group of our approach differ substantially from other titles in the field. Due to these distinct features, it can be classified as an edutainment game [14].

III. Methodology

In this section, the methodology of developing the game Cloud Computing with the game engine Unity3D [15] is decribed.

A. Overview of the Game Mechanics

This section gives a brief introduction of the game. We describe what the user sees first and what he is supposed to do at which stage. First in the game, the start menu is loaded. This menu contains the two buttons “Start Game“ and “Introduction“. In the introduction the description of the game and an explanation of the basic functionalities are provided.

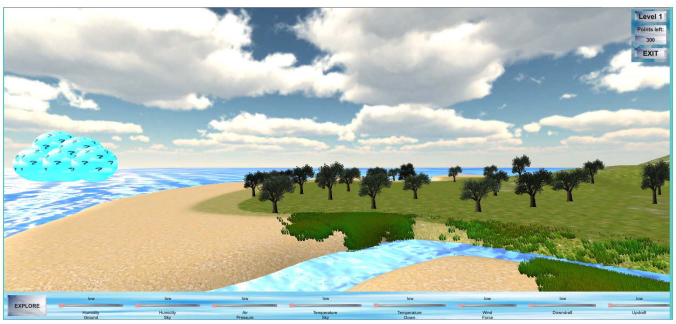


Fig. 1. The main screen of Cloud Computing.

After starting the game, the first level is loaded and the GUI for user interaction is displayed, as can be seen in Fig. 1. A 3D island is shown to the user. It is a closed environment with a few quests represented by 3D cloud icons. The user can decide whether he wants to solve these quests or explore by himself setting sliders and clicking the “Explore” button. The sliders on the bottom represent the changeable weather conditions and are always visible to allow for self-guided learning via the “Explore” button, as displayed in Fig. 4. The sliders each have three levels (low, medium and high), regulating the humidity of the air, air pressure, temperature, the wind force, as well as the up- and downdraftThe other permanent GUI element is in the top right corner of Fig. 1 and shown in detail in Fig. 2b). By clicking on the “Level” button, the next level is loaded, if all quests on an island have been solved successfully and the reward for the quest has been added to the user’s score. In order to quit the game, an “Exit” button is included to this GUI. The player clicks on one of the cloud icons to start a quest. The quest details are displayed in a pop-up window. Compared to [19], Cloud Computing does not allow the user to adjust settings of one phenomenon in different manifestations, but of different phenomena altogether. To further exploit this learning opportunity, a quest starts by asking for the proper names of a visualized phenomenon, or vice versa. Afterwards, the user can work through the presented phenomena one after the other.

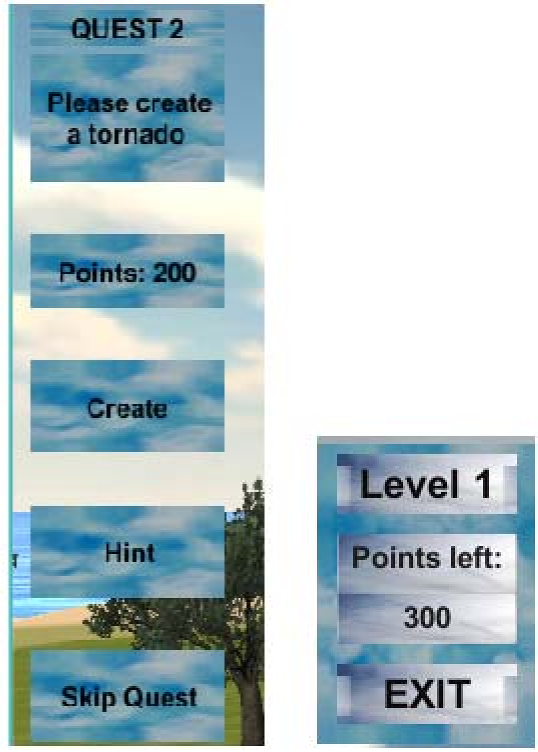


Fig. 2. a) Interaction bar during solving quests (left)
b) GUI element is in the top right corner (right)

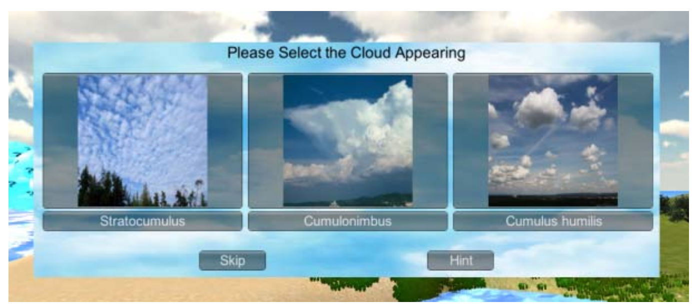


Fig. 3. Task after clicking on a quest cloud: Select right cloud formation according to the weather phenomena

At that point, a hint may be given to continue. After this task was solved successfully, the sliders have to be set to the right positions in order to start the animation of the weather phenomenon. When solving the quest, a GUI bar appears to the left (see Fig. 2a): It contains interactive and non-interactive elements. The first three labels from the top inform the user about the quest to be solved, including the points that can be achieved. The three buttons below have the following functions: By clicking “Create”, the weather phenomenon is animated if the sliders were adjusted correctly. . If only one slider has the wrong settings, the desired weather phenomenon is not animated. An example is depicted in Fig. 5, where a tornado is created, Fig. 4 shows the correct settings for the tornado. Other weather phenomena are thunderstorms, winter storms, rain, hail, fog, tornados, and snow. By clicking the button ”Hint”, further explanations and instructions are shown. “Skip” abandons the currently active quest.

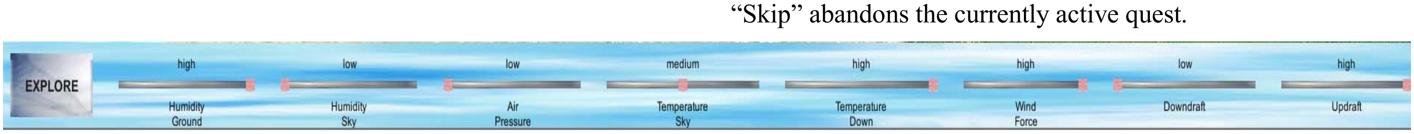


Fig. 4. The slider bar with settings of a tornado

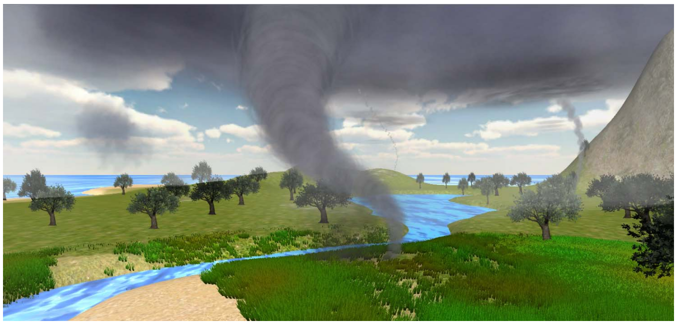


Fig. 5. The triggered weather phenomenon – a Tornado – is animated

After the weather animation, level and score are updated. The actual simulation model of Cloud Computing is based on a comprehensive database (derived from [12]), comprised of several rules that link the adjustable weather parameters to (a) certain changes in the parameter settings themselves, (b) changes in the weather, and (c) emerging weather phenomena as a whole.

B. *Concept User Experience*

While the player levels over time, his knowledge grows. He receives feedback after testing the slider adjustments if he is either solving a specific task or experimenting on his own agenda. The feedback is served as an animation of the effected weather. The game consists of different levels, which aims at an increase of complexity over time and provides guidance for the user. The camera movement is restricted in order to focus the user's attention on the important parts of the game. There is no movement of the camera while the user is adjusting the sliders or clicking buttons. Zooming in and out to see details of the level scenario and the animations is allowed for the user's comfort and it is possible to rotate the camera. The "Explore" button emphasizes the autonomy of the user, wheras the combination of questions about cloud types, the interaction with weather conditions, and animated weather phenomena ensure the immersion into the targeted educational domain.

IV. Results and Discussion

The edutainment game Cloud Computing instills the underlying mechanisms of the formation of various weather phenomena. In the following paragraphs, the complexity and elaboration of the finished game are elucidated.

A. *Science, Gamification and Complexity*

The simulation provides an insight into the development of various weather phenomena. As described in Section III, the user learns what weather conditions have to be given to create, e.g., thunderstorms, rain, or hail. The weather phenomena are presented in an abstract form and only selected influences can be changed. This, in combination with the underlying scientific basis ensures a focused learning experience. The necessary criteria for each weather phenomenon are based on scientific research to ensure the correct visualisation and a knowledge increase. Interaction is required in Cloud Computing through the already mentioned quests and levels. Experiencing dangerous weather phenomena in a safe, virtual environment can be fun. Animation, immediate feedback and hints provide for a smooth gameplay. The motivation of the user is supported by a decent difficulty of the tasks. They are neither too easy, nor too challenging.

In the game, weather is simulated in an abstracted form, by focusing only on some aspects. But they are still recognizable for the user. The abstraction was implemented by developing weather phenomena in a shorter period of time, since their formation would need more time in nature. Complex processes like seasons are not taken into account.

B. *Aesthetics*

The novelty aspect of the game is the new approach of a holistic simulation of interactively changeable weather. Thereby, we wanted to cover more aspects of the topic than other games. The abstracted presentation of varying islands with trees, mountains, rivers, lakes, sandy beaches, and much more contributes to the aesthetics of the game. The coloured and uniformly presented quest clouds suit the uniform style of the game. We tried to create a clear and suitable presentation of details at the specific levels to accomplish an aesthetic view and facile understanding. The quest presentation as well as the self-explaining three levels for each slider were created to increase the sleek and simple appearance. The introduction shall prepare the user for the game efficiently, so no further explanation is needed during the game. Furthermore, solved quests disappear to contribute to a clean and structured overview. To provide a comprehensive overview throughout the game, a restricted camera navigation and first person view is used. With the appearance and consistency of the label, buttons and windows we tried to contribute to the clarity of the game. The informativeness of the game is high (Section IV.C), as only needed information is presented and hints can be used if necessary. The different buttons are self-explaining as well as the score level on the right hand side as shown in Fig. 6.

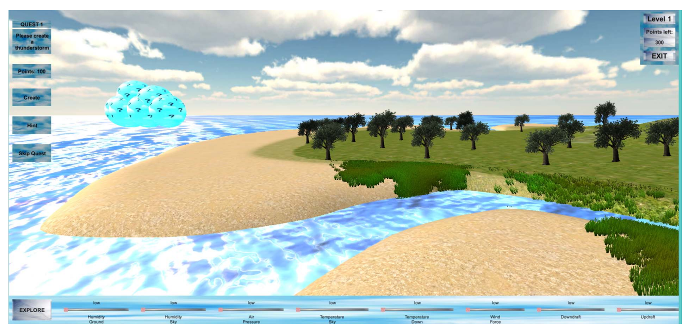


Fig. 6. Scene with Sliders and Quest Cloud before choosing a quest to solve

C. *User Survey*

To get feedback about the game, a user survey was conducted. 34 persons provided feedback, the age of the experimentees ranged from 16 to 31. First, the knowledge of the experimentees about serious games, weather in general and weather phenomena in particular were identified by asking about prior knowledge on the mentioned topics. Therefore, the survey started with the questions "Do you know the term `serious games'?", "Have you ever played a serious game?", and "Do you have detailed knowledge about the formation of weather?" 50% of the experimentees knew the term "serious games", 44.88% already played a serious game. Only 18.18% had detailed knowledge about the formation of weather phenomena. Afterwards, the experimentees were asked to experience the game and answer questions about the game play. In detail, we asked about the navigation throughout the game, intuitiveness, representation, fun, learning factor, and willingness to recommend. The answer possibilities were each "poor", "fair", or "good". All criteria were mainly evaluated with "fair" to "good", details can be seen in Table 1. Noteworthy is the high fun value (46.88% rated "good") and that only 6.25% criti-

cized a "poor" learning factor. As a last, the experimentees were asked to overall rate the game, answer possibilities were "excellent", "good", "fair", "poor", and "very poor". A majority rated the game as "good" or "excellent" (48.39%, respectively 22.58%), no one said it was "very poor", the remainder rated the game "fair" or "poor" (19.35% respectively 9.68%).
With these results, further improvements, e.g., on the navigation throughout the game, can be made.

	Poor	*Fair*	*Good*
Navigation	15.63%	**46.88%**	37.5%
Intuitiveness	9.38%	**50%**	40.63%
Representation	6.25%	**53.13%**	40.63%
Fun	12.5%	40.63%	**46.88%**
Learning	6.25%	**59.38%**	34.38%
Willingness to recommend	15.63%	**43.75%**	40.63%

Table 1. Results of the user survey.

V. Conclusion and Future Work

The aim of the project was to create an educational game. The presented game idea was compared to related games, the concept idea and design decisions were explained in detail. In addition, the complexity of the topic, the simulation and the user experience were introduced. All in all, the learning game is ideal to get a basics knowledge about and understanding weather. It is by no means usable, however, for weather research as an accordingly complex representation and necessary details are not integrated. For future work, it would be possible to expand the abstracted model or create real world scenarios with more complex phenomena like ocean currents or the coriolois force. Therefore, additional sliders or "bonus quests" could be added. In addition, it would be possible to integrate seasons, climate change and different climate zones, as they are not represented in this model. A great extension would be the combination with other games. For example, a game with the goal to grow plants could use the sliders to change the weather and therey the growth in specific niches. In a real-time strategy game, the power to influence the weather could become a strong ally in epic battles – all, of course, inspired by actual historic events.

References

[1] American Meteorological Society, Founded in 1919, annual Conferences

[2] WWOSC, "World Weather Open Science Conference," Canada, 2014.

[3] Löhr, Verena. "Der Wandel im frühneuzeitlichen Umgang mit Klimaveränderungen unter dem Einfluss der einsetzenden Aufklärung.", 2012.

[4] H. Häckel, Wolken, Eugen Ulmer, 2010.

[5] K. Brandt, Das Wetter - Beobachten, verstehen, voraussagen, Anaconda, 2012.

[6] P. Sohn, Director, *Partly Cloudy*. [Film]. USA: Pixar Animation Studios, 2009.

[7] SWR/WDR, "Interactive Wolkensimmulation," [Online]. Available: http://www.planet-schule.de/sf/php/mmewin.php?id=56. [Accessed November 2014].

[8] Wheeling Jesuit University/Center for Educational Technologies, "The Weather Game," 1999-2015. [Online]. Available: http://www.cotf.edu/ete/modules/k4/online/Wonline1.html. [Accessed November 2014].

[9] K. Mansfield, "Wild Weather Adventure :: NASA's the Space Place," NASA, JPL, California Institute of Technology, International Technology and Engineering Education Association, [Online]. Available: http://spaceplace.nasa.gov/wild-weather-adventure/en/#/review/wild-weather-adventure/game.html. [Accessed November 2014].

[10] Cambridge English Online, "Weather Maze," [Online]. Available: http://learnenglishkids.britishcouncil.org/en/fun-games/weather-maze. [Accessed November 2014].

[11] R. Hamblyn, Welche Wolke ist das?, Franck Kosmos Verlag, 2009.

[12] M. Schulze, Wetterkunde: Wind, Wetter, Wolken und Temperatur, Schulze Media, 2014.

[13] S. B. Smith, D. Pace, S. J. Goodman, D. W. Burgess, D. Smarsh, R. D. Roberts and M. M. Wolfson, "The THOR Project - Reducing the impact of thunderstorms on aviation and the general public through a multi-agency effort," *Conference on weather analysis and forecasting,* no. Vol. 18, pp. 314-316, 2001.

[14] Egenfeldt-Nielsen, Simon. Beyond edutainment: Exploring the educational potential of computer games. , 2005.

[15] Unity Technologies, Unity3D [Online] Available: http://unity3d.com/ [Accessed June 2015]

[16] T. Nishita, H. Iwasaki, Y. Dobashi and E. Nakamae, "A Modeling and Rendering Method for Snow by Using Metaballs", In: *Computer Graphics Forum* Vol. 16, No. 3, 1997

[17] M. J. Harris and A. Lastra, "Real-Time Cloud Rendering", In: *Computer Graphics Forum* Vol. 20, No. 3, 2001

[18] L. McLoughlin, "Realistic Natural Atmospheric Phenomena and Weather Effects for Interactive Virtual Environments", 2012

[19] L. McLoughlin and E.F. Anderson, "Reverse Skew-T – A Cloudmaking Tool for CG", In*: Proceedings of the Eurographics 2013 – Posters,* 2013

A Social-centred Gamification Approach to Improve Household Water Use Efficiency

Zhenchen Wang
Department of Computer Science
Brunel University
London, UK
zhenchen.wang@brunel.ac.uk

Andrea Capiluppi
Department of Computer Science
Brunel University
London, UK
andrea.capiluppi@brunel.ac.uk

***Abstract*— The research community is showing a growing interest in gamification and there are works showing the usefulness of gamification in different problem domains. Recently, a special interest has been given to the gamification design on systems addressing natural resource consumption issues such as to encourage efficient household water consumption. Despite the potential benefits, the gamification design method for such system is not conclusive. In this paper, we proposed a social-centred gamification approach to improve household water use efficiency. The approach firstly identified the water use related social network activities based upon existing popular social network activities. The approach then gamified each identified activity in terms of traditional instruments for improving water use efficiency and gamification rewards. The approach also used a set of indicators to explicitly detect and monitor both online social network activities and offline water use activities. With this approach the gamification effectiveness can be better traced and evaluated.**

Keywords*— gamification, social network gamification, gamification design, efficient water use

I. Introduction

The persuasive technologies are emerging as useful tools to support behavioural change in daily life [13]. In recent years, a special interest is given to the technologies designed to support natural resource consumption behavioural change such as efficient household water use (e.g. [16] [17]). Thanks to the advancement of digital sensors, researchers and companies are developing smart metering systems that can collect real life water consumption data and present to users in different visualised formats. E.g. in [24] and [25], individual household water use data can be collected from smart meter and water use information can be further presented to users based upon the collected data. Such kind of system can be named ‘personal informatics system’ according to Li et al.’s work [14], which is defined as collecting personally relevant information for the purpose of gaining self-knowledge. The personal informatics systems allow users to explore and reflect the collected information and the reflection can lead the individual to reconsider and possibly change attitudes or behaviours [15]. A more recent development trend of persuasive technology is to present the systems to users in a game like manner, which is called the gamification. Gamification is generally defined as the use of game design elements in non-game contexts [18]. With a growing popularity of gamification systems in the markets and research fields, gamification is argued to be a next generation method for marketing and customer engagement [20] and has been proposed as a design pattern for persuasive systems [22].

Despite the envisioned benefits brought by gamification, there are still challenges with respect to the gamification design and evaluation. The gamified systems are argued to be neither a pure functional software nor a full-fledged game [23] and hence there are no well-established methods to design such systems. In Hamari et al.’s work [21], 24 research works were reviewed in terms of the effectiveness of gamification. Even though the majority of the reviewed studies reported positive results from gamification, the authors concluded there were still limitations that might affect the qualification of the previous works’ results. E.g., one of them is that controls between implemented elements of gamification were often lacking, so multiple elements were investigated as a whole. As a result, it remains unclear how gamification can be used and evaluated to support natural resource consumption behavioural change.

In the light of the criticisms and challenges on the gamification design and evaluation, we proposed a social-centred gamification design approach to develop the gamification specially for improving household water consumption efficiency. Our proposal is researched and developed in the context of the FP7 EU Project ISS-EWATUS [31], where a social network (SN) platform is provided to increase water consumption awareness and improve household water use efficiency. One of the fundamental differences between a SN based persuasive system such as in ISS-EWATUS and a sensor based individual persuasive system is the former entails additional set of user-user activities that are envisioned can enable offline water use change. Hence, it firstly requires SN activities to be defined and mapped in the context of problem domain, i.e. water use efficiency, so that the activities can be better justified to engender the target behavioural change. In order to implement the gamification and to evaluate the gamification effectiveness in such SN platform, there should also be mechanisms to link users’ online activities and offline water use, i.e. the activity is not only merely between online users and the SN, but also the between the external, physical systems and the users. This aspect is of course specially complicated, in particular when the external system interfaces are unknown or incompatible with the SN system interfaces.

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

In this paper, a social-centred gamification approach is presented to address these gamification design and evaluation challenges, i.e.

- to define social activities used in a SN to enable user behavioural change
- to gamify the SN activities to incentivise SN users to enact expected online activities and offline water use change.

The approach is currently being used in the ISS-EWATUS project for promoting efficient household water consumption behaviour. There are also important issues such as the impact of privacy and the cultural considerations on designing gamification for a SN platform which are outside the scope of this work.

The rest of the paper is organised as follows. In section II, Related works are firstly reviewed. The proposed approach is presented in Section III. Section IV concludes and outlooks the work.

II. Related Work

There are lots of research works on using ICT tools to promote efficient water use. Here, we reviewed the most relevant works starting with social networks and behavioural change and then we reviewed the gamification motivational affordance on social networks.

A. Social Networks and Behavioural Change

Behaviour can be influenced through **social norms**: people are more likely to participate in an activity if others around them are participating in the same activity [5] [4]. In [26], it showed that water conservation could be influenced by the presence of people modelling the promoted water saving action of turning off the shower while soaping. It was found that before the study 6% of shower users carried out the action while 93% were aware of the sign suggesting the action. Following the presence of one model 49% carried out the action. The presence of two models resulted in 67% participation. Studies including [7] [12] have also indicated that SNs can be used to motivate or to give certain pressure to people to behave in certain ways.

Behaviours can also be influenced via **designed intervention**. In Rolls' work [8], a concept called CBSM (Community Based Social Marketing) referred to those intervention programs that attempt to apply a structured approach and the insights of social psychology when influencing community behaviour. CBSM involves fours steps including: 1) to identify barriers which prevent the action, e.g. installing a low-flow showerhead, from being taken, 2) to design a programme to overcome the selected barrier to achieve the behavioural change. The important aspect of the design stage is to target interventions very specifically towards the identified barriers, drawing on social-psychological devices such as commitments, prompts, or signals, to promote the desired behaviours. 3) To pilot and test the programme and 4) evaluate its effectiveness before it is applied on a wider scale. Based on CBSM, Jackson [27] further found that habitual or routine change is much more difficult to influence compared to changing a one-off behaviour.

B. Gamification Motivational Affordance on Social Networks

Gamification is commonly known as the use of game design elements in non-game contexts [18]. Gamification is also defined as a process of enhancing services with affordances in order to invoke gameful experiences and further behavioural outcomes [19]. Reward, as a motivational affordance [28], is one of the key composing elements in a gamification system. In [9] it was found that behavioural change was more likely to occur if **physical (i.e. real) rewards** were offered, e.g. it was illustrated that more actions were taken on use of free products compared to personal initiatives. **Virtual reward** systems (such as scores, stars, reputations and badges) are a common practice in many gamified SN systems. Reasons of why such a reward system is successful can be explained by using the behavioural model proposed in [3]. However, the model in [3] also argued that an activity that is easy to do, such as play an online game, will require only a low level of motivation for a person to do it. When an activity is more challenging to do, it will require a much higher motivation. The virtual reward systems can be also viewed as a motivational affordance to trigger social comparison which refers to the process of evaluating one's own abilities and opinions by comparing them to the abilities and opinions of others [30]. The underlying assumption is that users would become more physically active in order to outperform others. Another type of reward is the **private personal rewards** which are dependent upon individuals and they do not have explicit forms. These rewards can be i) the gratification received from entertainment, ii) information and knowledge retrieval, and iii) social activity [11] [2].

III. Social-Centred Gamification Approach

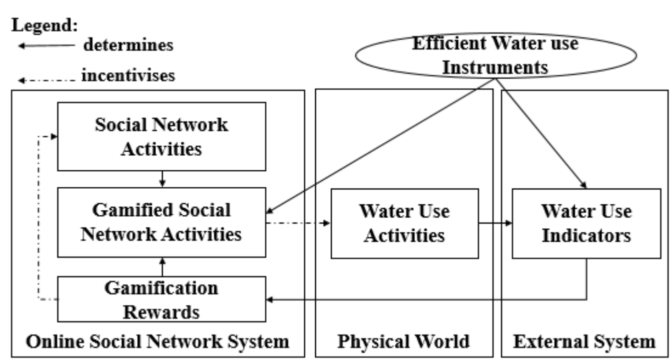


Fig. 1 Overview of Social-centred gamification approach

The proposed approach (See Fig. 1) models five components distributed across online SN system, physical world and external system. It also models the efficient water use instruments as an external input to contextualise the SN activities in efficient water use domain. The approach aims to transform the existing SN activities to 'SN version' efficient water use instruments supported by gamification. It also aims to support gamification evaluation, i.e. better trace and monitor the gamification via links between individual gamification process and individual SN activity (SNA). In the next few sections, we start by defining efficient water use instruments and then each of other components will be elaborated respectively.

A. Efficient Water Use Instruments

In water use domain, the WCM (water conservation management) instruments [6] is a well-known water use management theory. The WCM provides a good theatrical base for designing efficient water use instruments. The **efficient water use instruments** therefore can be categorised into five Es: 1) engineering (i.e. physical WCM equipment); 2) economics (i.e. water price related information such as rebates); 3) enforcement (i.e. penalty measurements on water waste); 4) encouragement (i.e. endorsement on water conservation behaviour); 5) education (i.e. raising efficient water use awareness).

B. SN Activities

In the ISS-EAWATUS, we entail a selected list of common activities extracted from a comprehensive list of **SN activities** (SNAs) identified in [1]. Each of these activities are tailored to a set of water user related SN activities which were supported by potential system users in Greece Poland and UK [31]. (See **Table 1**).

Table 1 SN Activities

#	Common SNAs	Water Use Related SNAs
1	Sharing information	Sharing water use pattern to friends; Sharing water bills with friends
2	Getting recommendations	Water use tips search
3	Organising social events	Water use topics discussion e.g. surface water drainage rebate
4	Playing games	Completing a task driven water use activity, e.g. start a shower between 8pm and 8:30pm and finish it less than 10 minutes
5	Keeping in touch with friends	Inviting friends to take part in water use related SN activities.

C. Gamified SN Activities and Gamification Rewards

Gamified SN activities (GSNAs) are a set of social activities that can be used to incentivise online and offline activities. A GSNA is determined by both a SNA and gamification rewards, hence defining a GSNA can be viewed as a two-step process.

The first step is to classify the defined SN activities in terms of the efficient water use instruments. As an example, Table 2 shows how SNAs in Table 1 are defined in ISS-EWATUS as GSNAs. The second step is to associate a GSNA with the **gamification rewards**. The decisions on which types of reward to be used, as reviewed in section III.B, depend upon the instrument nature of a SNA, e.g. the physical rewards could be more suitable for engineering instrument related SN activities whereas private personal rewards may apply to all instruments. For the evaluation purpose, the assignment of rewards should also take account of the relationship between rewards so that rewards can be traced down to specific SNA.

In ISS-EWATUS [31], each GSNA is assigned a distinguishable rewarding mechanism, i.e. different rewarding scores/badges for different achievements for an activity. As part of the evaluation process, we will evaluate the effectiveness of rewards or reward conversions by using both qualitative approach such as survey and interview and quantitative approach in which gamification rewards data and real water consumption data will be retrieved from both the online SN system and external systems (See section D).

Table 2 ISS-EWATUS GSNAs (See **Table 1** for social activity ids)

Related Efficient Water Use Instruments	GSNAs (Table 1)
Engineering	(4)
Economics	(1) (3)
Enforcement& Encouragement	(2)
Education	(5)

D. Water Use Activities and Indicators

To enable water use behaviour change, **water use activities** should be detected and monitored. However, the monitoring task is beyond the boundary of an online SN system and it is usually difficult to learn the exact water use activity without consulting the users.

Here, we use indictors to explicitly assess whether an expected offline water use task has been done. The **indicators** refer to the quantitative measurements associated with specific GSNA. In ISS-EWATUS, the calculation of a GSNA's gamification rewards are associated with the indicators. E.g. if the GSNA is to share water meter readings with friends, then the indicators can be the frequency of uploading a user's water readings and the number of people shared with. The indicators can also be retrieved from external systems, and this is especially true for the offline water use activities. E.g., in the water use related activity (4) demonstrated in Table 1, the water use patterns in terms of time and total water use from a smart water meter can be the indicators in this case.

IV. Conclusion and Further Work

We presented a social-centred gamification approach helping the design, implementation and evaluation of a SN system for promoting household water use efficiency. We demonstrated how the proposed approach is used in the ISS-EWATUS project via modelling the gamified social network activities, gamification rewards and indictors used for monitoring both online SN activities and offline water use activities.

The next research task is to evaluate the proposed approach. We aim to find out how effective the designed GSNAs can help change households water use pattern and we will also assess the effectiveness of each individual GSNA. This will be done via two methods, one is via continuous collecting and analysing the quantitative data from the systems including both SNs and associated external systems; and the other is to analyse the qualitative data periodically retrieved from questionnaires and focus groups within and outside the project consortium

ACKNOWLEDGMENT

This work has been undertaken within the framework of the ISS-EWATUS, Integrated Support System for Efficient Water Usage and Resources Management, FP7 project (grant no. 619228), funded by the European Community.

REFERENCES

[1] A. Binsted, R. Hutchins,The role of social networking sites in changing travel behaviours. TRL Reports, p.38, 2012.

[2] Á. Dunne, M. Lawlor, & J. Rowley, Young people's use of online social networking sites - A uses and gratifications perspective. Journal of Research in Interactive Marketing, 4(1), pp. 46–58, 2010.

[3] B. Fogg. A behavior model for persuasive design. In Proceedings of the 4th International Conference on Persuasive Technology (Persuasive '09). ACM, New York, NY, USA, Article 40, 7 pages, 2009.

[4] D. Foster, C. Linehan, B. Kirman, S. Lawson, and G. James. Motivating physical activity at work: using persuasive social media for competitive step counting. In Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments, pages 111–116. ACM, 2010.

[5] D. McKenzie-Mohr, Fostering Sustainable Behavior: An Introduction to Community-Based Social Marketing. New York: New Society Publishers, 2011.

[6] H. Savenjie, and Z. Van, Water as an economic good and demand management, paradigms with pitfalls. International water resources association, Water international, 27(1), pp. 98-104, 2002.

[7] J. Cotterell, Social networks and social influences in adolescence. London: Routledge. 1996.

[8] J. M. Rolls, A review of Strategies Promoting Energy Related Behavioural change. International Solar Energy Society Solar World Congress. Adelaide, South Australia 25th November-2nd December, 2001.

[9] J.Grizzell, Behavioural change Theories and Models: relating to health promotion and education efforts. American College Health Association, 2003.

[10] R. Wirfs-Brock, “Designing scenarios: making the case for a use case framework” SmallTalk Report, November-December, 1993.

[11] T. Jung, H. Youn, & S. Mcclung, Motivations and self-presentation strategies on Korean-based “Cyworld” weblog format personal homepages. CyberPsychology & Behavior, 10(1), pp. 24–31, 2007.

[12] W. Barry, Are personal communities local? A Dumptarian reconsideration, Social Networks, 18(4), pp. 347-354, 1996.

[13] Fogg B Persuasive technology: using computers to change what we think and do. Morgan Kaufmann, San Francisco, 2002.

[14] I. Li, Dey A, Forlizzi J A stage-based model of personal informatics systems. In: Proceedings of the SIGCHI conference on human factors in computing systems (CHI 2010). ACM, New York, pp. 557–566, 2010.

[15] B. Ploderer, W. Reitberger, G. Oinas-KukkonenH, Social activity and reflection for behavioural change. Pers Ubiquit Comput, pp.1667-1676, 2014.

[16] H. Lunzer,, "Intelligent Metering," Industrial Informatics, 2007 5th IEEE International Conference on , vol.2, no., pp.1215,1219, 23-27 June 2007.

[17] T. Shoji, W. Hirohashi, Y. Fujimoto, Y. Hayashi, "Home energy management based on Bayesian network considering resident convenience," Probabilistic Methods Applied to Power Systems (PMAPS), 2014 International Conference on , vol., no., pp.1,6, 7-10 July 2014.

[18] S. Deterding, D. Dixon, R. Khaled, L. Nacke, From game design elements to gamefulness: defining ‘‘gamification’’. In: Proceedings of the 15th international academic MindTrek conference: envisioning future media environments (MindTrek 2011). ACM, New York, pp 9–15, 2011.

[19] J. Hamari, “Transforming Homo Economicus into Homo Ludens: A Field Experiment on Gamification in a Utilitarian Peer-To-Peer Trading Service”, Electronic Commerce Research and Applications, 12(4), pp. 236- 245, 2013.

[20] G. Zichermann, and C. Cunningham, Gamification by design: Implementing game mechanics in web and mobile apps, Sebastopol, CA, O'Reilly Media, 2011.

[21] J. Hamari, J. Koivisto, H. Sarsa, "Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification," System Sciences (HICSS), 2014 47th Hawaii International Conference on, pp. 3025 - 3034, 2014.

[22] O. Michael, A. Tuomas, and O. Harri, Persuasive software design patterns for social influence. Personal Ubiquitous Comput. 18(7), pp.1689-1704, 2014.

[23] S. Deterding, S. Bjo¨rk, L. Nacke, D. Dixon, E. Lawley, Designing gamification: creating gameful and playful experiences. In CHI’13 extended abstracts on human factors in computing systems (CHI EA 13). ACM, New York, pp 3263–3266, 2013.

[24] O. Masia, L. Erasmus, "Smart metering implementation for enabling Water Conservation and water demand management: An investigation in Gauteng, South Africa," AFRICON, 2013 , pp.1,5, 9-12 Sept. 2013.

[25] R. Cardell-Oliver, "Discovering water use activities for smart metering," Intelligent Sensors, Sensor Networks and Information Processing, 2013 IEEE Eighth International Conference on, pp.171 - 176, 2013.

[26] M. Mohr, D. Fostering Sustainable Behavior: An Introduction to Community-Based Social Marketing. New York: New Society Publishers, 2011.

[27] T. Jackson, Motivating Sustainable Consumption: A review of evidence on consumer behaviour and behaviour change. Report to the Sustainable Development Research Network. Centre of Environmental Strategy, University of Surrey. Guilford, Surrey, 2005.

[28] J. Webster, and R.T. Watson, “Analyzing the past to prepare for the future: writing a literature review”, MIS Quarterly, 26(2), pp. xiii-xxiii. 2002.

[29] E. Deci, R. Koestner, R. Ryan, Extrinsic rewards and intrinsic motivation in education: reconsidered once again. Rev Educ Res 71: pp.1–27, 2001.

[30] A. Kruglanski, O. Mayseless, Classic and current social comparison research: expanding the perspective. Psychol Bull 108: pp.195–208. 1990.

[31] Deliverable 5.1 Design documentation of the architecture of the system, available at http://issewatus.eu.

Explicit Fun, Implicit Learning in Multiplayer Online Battle Arenas

Methodological proposal for studying the development of cognitive skills using commercial video games

Carlos Mauricio Castaño Díaz *(1st Author)*
Heinrich Hussmann *(3rd Author)*
Jan-Willem Strijbos *(4th Author)*
Ludwig-Maximilians Universität München
Munich, Germany.
mauricio.diaz@psy.lmu.de
heinrich.hussmann@ifi.lmu.de
jan-willem.strijbos@psy.lmu.de

Birgit Dorner *(2nd Author)*
Katholische Stiftungsfachhochschule München
Munich, Germany.
birgit.dorner@ksfh.de

***Abstract*—the present study aims to evaluate how commercial video games can foster hypothetico-deductive reasoning in the everyday life. For this, we studied Multiplayer Online Battle Arenas while developing a methodology based on multi-design and multi-level data triangulation between game videos, eye tracking, game talks, and cued retrospective protocols. The studies take place in a simulated-naturalistic environment.**

Keywords—video games; cognition; reasoning; cognitive skill development; mixed methods

I. Introduction

The present work is an interdisciplinary study linking the fields of Psychology, Game Studies, and Learning Sciences. It aims to understand the process for which hypothetico-deductive reasoning [1] can be fostered by commercial games. The study focuses on one specific genre of video games: Multi-player On-line Battle Arenas (MOBAs). This genre was chosen because MOBAs are today, along with Real Time Strategy games (RTSs), Fighting Games, and First Person Shooters (FPSs), considered not only to be games but a type of sport (eSport) [2]. So, playing these games can be both a leisure-time activity as well as a professional one.

II. Rationale

From an evolutionary perspective, playing is considered best activity for different species to develop behaviours and abilities necessary for adapting to both their physical and social environment [3]. Information technologies have become the centre of the contemporary world for humans, changing the way we engage with our surroundings, socialise, and play [4][5]. In a double-folded fashion, not only the current technologies have changed the way people play, but the later has also evolved allowing people to adapt better to the new information society.

During the last two decades research focusing on the influence of video games in our everyday life has increased. Anthropological [6], cultural [7], social [8], learning [9], and psychological studies [10] aim to understand different facets of the video game phenomenon, like the creation of gamer communities, cultural changes, literacy, learning, and development of cognitive skills.

Although the number of studies investigating the phenomenon of video games has increased in the last years, there are still not many models and research strategies that can be adapted by new researchers in the field allowing to conduct replicable studies [11]. Furthermore, there still exist methodological problems with the study of video games and skill acquisition, like the problem of over generalising a research to all video games [12], the problem of transfer [13], amongst others. Also, the methods to study the influence of video games sometimes differ to how they are presented in the everyday life and socio-environmental context of players, thus questioning the ecological validity of the studies.

III. Proposal and Methodology

To tackle these problems, the present research aims to develop a methodology to study hypothetico-deductive reasoning in games (MOBAs) using mixed methods[14]. The research makes use of two different designs, (1) exploratory design [15] which allows to identify patterns of action and thought linked to the gameplay. (2) A microgenetic design [16][17] which allows to identify emergence and development of skills over time. The exploratory study is being carried with ranked (expert) players of the games League of Legends [18] and Defence of the Ancients 2 [19],

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

while the microgenetic study is being carried with people who has never played any MOBA (novices).

For the study game recordings, eye tracking data [20], in-game communication, and cued retrospective reports [21] are used as a means of analysing the different reasoning processes of the players in a simulated-natural environment. The simulated-natural environment is an environment created to be in between a natural environment and a laboratory environment. Therefore, players can feel like playing a normal game at home while in a laboratory setting (i.e. wearing eye trackers). This combination of a natural environment with a laboratory setting allows for a natural development of the game without the synthetics of the laboratory setting and the invasiveness of the field studies [22].

The research aims to use the cued retrospective reports as a means to triangulate the in-game behaviour, eye tracking data and the game speech in order to understand how players reason while playing the game, as well as understanding which cognitive processes are being used in order to ace the game.

IV. Reach

This research aims to understand better how people develop hypothetic-deductive reasoning in an everyday context (video gaming). Thus contributing to a better understanding on how video games impact people cognitively in their everyday life and contexts. It is expected that the findings of this research will shed some light on how learning can be fostered in informal learning settings and how video gaming can contribute to this. The methodology developed within the study aims to create a more accurate and ecologically valid way to study video games, avoiding the artificiality of the laboratory and the intrusion of natural observations.

V. Partial results

In the current phase of the project, it has been possible to identify some trends for both the expert and the novice groups.

- Expert players make hypotheses, both in conscious and unconscious ways. Their hypotheses are related to opponent behaviour, opponent intentions, opponent location, and the probability of the occurrence of a specific event. In contrast, novice players do not make high-level hypotheses; their hypotheses are mostly based on patterns gathered from the opponent-computer.

- Expert players try to acquire more information to test their hypotheses. Nevertheless, they tend to assume their beliefs to be true and act based on this. Novice players are too focused on particular events to search for new information.

- For the experts, map awareness and manipulation of information are key elements for hypothesis and strategy generation. On the other hand, novice players create strategies by imitating the computer behaviour. They follow this until they find a winning strategy, and once they have found it, it becomes stable and inflexible.

VI. Conclusions and further research

Although the study is still in progress, the current methodology has shown great potential to understand the use and development of cognitive skills. It has also depicted a good way to understand how players play MOBAs. Further on, it is expected to expand the research so that cognitive models and schemas can be generated.

Acknowledgments

This research is being carried with the support of the Elite Network of Bavaria (ENB), The Munich Center of the Learning Sciences (MCLS), and the Catholic School of Applied Sciences (KSFH).

References

[1] K. Popper, *The logic of scientific discovery*, 2002nd ed. New York: Routledge Classics, 1959.

[2] B. Popper, "Field of streams: how Twitch made video games a spectator sport," *The Verge*, www.theverge.com, 30-Sep-2013.

[3] G. Burghardt, *The genesis of animal play*. Cambridge, United States of America: MIT Press, 2005.

[4] C. Steinkuehler, "Cognition and literacy in massively multiplayer online games," in *Handbook of research on new literacies*, New York, USA: Lawrence Erlbaum Associate, 2008, pp. 611–634.

[5] J. Huizinga, *Homo Ludens. A study of the play element in culture*, 2014th ed. Connecticut, USA: Martino Publishing, 1950.

[6] J. Folkerts, "Playing Games as an Art Experience: How Videogames Produce Meaning Through Narrative and Play," *Emerg. Pract. Cyberculture Soc. Netw.*, pp. 99–117, 2010.

[7] J. P. Gee, *What video games have to teach us about learning and literacy*. New York, USA: Library of Congress, 2007.

[8] E. Abrudan and G. Prundaru, "Social Representations and Ideologies in Digital Games," *J. Media Res.*, vol. 5, pp. 31–41, 2009.

[9] L. L. Galarneau, "Authentic Learning Experiences Through Play: Games, Simulations and the Construction of Knowledge," Rochester, NY, 2005.

[10] F. Blumberg, E. Altschuler, D. Almonte, and M. Mileaf, "The Impact of Recreational Video Game Play on Children's and Adolescents' Cognition," in *Digital Games: A Context for Cognitive Development. New Directions for Child and Adolescent Development*, Blumberg and Fisch, Eds. 2013, pp. 41–50.

[11] E. Aarseth, "Playing Research:Methodological Approaches to Game Analysis," *Fine Art Forum*, vol. 17, 2003.

[12] S. Engl and L. E. Nacke, "Contextual influences on mobile player experience – A game user experience model," *Entertain. Comput.*, vol. 4, pp. 83–91, 2013.

[13] S. Barnett and S. Ceci, "When and where do we apply what we learn? A taxonomy for far transfer.," *Psychol. Bull.*, vol. 128, no. 4, pp. 612–637, 2002.

[14] C. Steinkuehler, E. King, C. Martin, Y. Oh, S. Chu, C. Williams, A. Ochsner, S. Harris, E. Owen, and G. Anton, “Mixed methods to study games and learning,” in *Proceedings of the 7th international conference on Games+ Learning+ Society Conference*, 2011, pp. 222–228.

[15] R. Stebbins, *Exploratory research in the social sciences*. United States of America: Sage Publications, Inc., 2001.

[16] R. Siegler and K. Crowley, “The microgenetic method: a direct means for studying cognitive development,” *Am. Psychol.*, vol. 46, no. 6, pp. 606–620, 1991.

[17] D. Kuhn, “Microgenetic study of change: What has it told us?,” *Psychol. Sci.*, vol. 6, no. 3, pp. 133–139, 1995.

[18] Riot Games, *League of Legends*. Riot Games, 2009.

[19] Valve Corporation, *Dota 2*. Valve Corporation, 2013.

[20] A. Poole and L. Ball, “Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects,” in *Encyclopedia of human computer interaction*, Pensilvania: Idea Group, 2005, pp. 211–219.

[21] T. van Gog, F. Paas, J. van Merriënboer, and P. Witte, “Uncovering the problem-solving process: cued retrospective reporting versus concurrent and retrospective reporting,” *J. Exp. Psychol. Appl.*, vol. 11, no. 4, pp. 2037–244, 2005.

[22] A. Bryman, *Social Research Methods*, 4th ed. USA: Oxford University Press, 2012.

Back to Contents

Program with Ixquic

How to Learn Object-Oriented Programming with a game.

Carmen Ramos[1]
[1] Department of Visualization and Virtual Reality, Ixtli, UNAM.
National Autonomous University of Mexico, UNAM
Circuito Exterior, Nº 3000, Ciudad Universitaria – Coyoacán
cramos@unam.mx

Tania Patiño[1]
[1] Department of Visualization and Virtual Reality, Ixtli, UNAM.
National Autonomous University of Mexico, UNAM
Circuito Exterior, Nº 3000, Ciudad Universitaria – Coyoacán
tania@ciencias.unam.mx, taniapc@live.com.mx

Abstract— **In this paper we want to describe advances reached in the game design and implementation related to *Serious Game: Program with Ixquic*. We present a work in progress about this videogame.**

Program with Ixquic*, has two main purposes: developing programming skills for every user that interacts with this videogame, where virtual scenes created in 2D dimensions presents *Object-Oriented Programming* concepts through the development of examples and exercises in *Java Programming Language.

Keywords—Educational Games, Serious Games, Game Design, Learning Object-Oriented Programming, Games, Virtual Reality.

I. INTRODUCTION

The name of this videogame *Program with Ixquic* comes from Quiché Mayan Folktale of the Sun and the Moon, a prehispanic tale where is related how in the beginning of the time was the creation of the Earth, Sun and Moon also *Ixquic* is the name of a Mayan princess.

The game aims to improve a user's programming ability when interacting with it. The objective is that while the user plays at the same time improves cognitive skills and learns programming techniques understanding OOP concepts [1],[3]. With this game we wanted to address several questions:

1. Is it possible enhance and improve user skills in certain subjects and programming concepts through a Serious Game?

2. What is the relationship between Serious Games and learning?

3. How can players be motivated to be apprentices, and how to find an optimal balance between entertainment and learning?

Serious Games:

"Serious Games are a mental contest, played with a computer in accordance with specific rules, that uses entertainment to further government or corporate training, education, health, public policy, and strategic communication objectives. [1]"

Virtual Reality:

"Virtual": the virtual word comes from medieval latin virtualis, which in turn derives from virtus: strength, power. The virtual is not opposed to reality, could be part of the reality. [4]"

II. PROGRAM WITH IXQUIC

However, we know that there are millions of games that just entertain but there are other directions like create educational experiences with games. We are taking this direction focus on developing an educational game that permits to improve and create specific programming skills in the field of Object Oriented Programming using Java Programming Language.

In *Program with Ixquic,* it is possible to resolve simple puzzles, questions and create blocks of simple control structures with touch screen interaction in a Galaxy Note tablet with Android system or in iPad tablet with iOS system.

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

Figure1. Intro Program with Ixquic

This game consist of two levels, where every user can have an interactive experience with each level, where learns basic concepts about O.O.P, control structures and have fun with *Ixquic* [5].

Figure2. Control structures Program with Ixquic

The game is at this moment a work in progress. The Figure3 is the main interface about *Program with Ixquic*.

Figure3. Program with Ixquic

III. SERIOUS GAMES TOPICS

In this section we describe the various aspects that are related to serious games:

Educational Games: with the multimedia explosion in 1990, those games were created with educational proposes but soon decreased because the poor quality of the games or because of a growing interest in the Internet (Michael & Chen, 2006), [1].

Healthcare Games: those games are applications about health and healthcare. Which ones influence direct or indirectly in the physiology and psychology of individuals.

Military Games: are used for training, these games have simple rules and allow the users to be better planners for battles in an individual form or in group training [1].

Government Games: training and simulation that concern a number of different tasks and situations, like dealing with terrorist attacks, biohazards, city planning, traffic control, firefighting, ethics training, defensive driving, etc.

Corporate Games: corporate training is a great industry that sometimes requires the creation of Serious Games which permits computer assisted training that allows corporations to cut cost for training staff, special equipment, places, etc. This kind of games started on 1990's with multimedia PC, CD-ROM's and Internet (Michael & Chen, 2006), [1].

Our objective with this game *Program with Ixquic*, is to create an educational game that permits to learn OOP concepts and basic control structures in Java Programming Language.

ACKNOWLEDGMENTS

We thank all the team and friends at the *National Autonomous University of Mexico, UNAM, specially the Department of Visualization and Virtual Reality, Ixtli, UNAM.* Where the development of this *Serious Game: Program with Ixquic,* is in progress. All the *Game Design* was made at my studio called *Madagascar Computer Games and Systems S.A. de C.V.*

REFERENCES

[1] Mitchell, D. & Chen, S. (2006). Serious Games: Games that educate, train, and inform. Boston, M.A. Thomson Course Technology.

[2] Wolz U., Barnes T., Parberry I., Wick M., Digital Gaming as a Vehicle for Learning. Dept. of Computer Science & Engineering. University of North Texas Denton,TX. USA., University of North Carolina-Charlotte, Charlotte, NC USA. ian@unt.edu, tbarnes@uncc.edu.

[3] J.C. Adams. Chance-It: An object-oriented capstone project for CS-1. In Proceeding of the 29 th SIGCSE.Technical Symposium on Computer Science Education, pages 10-14, ACM Press, 1998.

[4] CBurdea, G., & Coiffet, P. (2003). Virtual reality technology. *Presence: Teleoperators and virtual environments*, *12*(6), 663-664.

[5] Cox, B. J. (1985). Object oriented programming.

Tašlihan virtual reconstruction

Interactive digital story or a serious game

Selma Rizvić
Faculty of Electrical Engineering Sarajevo
University of Sarajevo
Bosnia and Herzegovina

Irfan Prazina
Faculty of Electrical Engineering Sarajevo
University of Sarajevo
Bosnia and Herzegovina

***Abstract*—During the Ottoman period, Tašlihan was the largest accommodation complex in Sarajevo, Bosnia and Herzegovina. Today, only one wall remains as a memento of its existence. In this paper, we compare user appreciation of an interactive digital story about this building and of a serious game about Tašlihan to see which application offers more knowledge and immersion while bringing this monument to life in the collective memory of the people.**

***Keywords*—serious games for cultural heritage, virtual reconstruction, virtual reality, interactive digital storytelling, user evaluation studies**

I. Introduction

Time is not a friend of cultural monuments. Many of them disappear under new construction projects and some exist only as remains at archaeological sites. People forget them or cannot imagine what they used to look like in the past. Digital technologies are an efficient tool to bring these buildings back to life in collective memory. They enable us to virtually explore 3D reconstructions of cultural heritage sites and learn about events and people from their history.

The scientific community is still researching the most suitable form for virtual heritage applications. Serious games for cultural heritage combine realistic virtual reconstructions of cultural heritage sites with stories and tasks to be performed by users in order to succeed in playing.

In this paper, we compare two virtual cultural heritage applications we created to bring the Tašlihan building back to life in collective memory: an interactive digital story and a serious game.

II. Related work

The thorough research on serious games for cultural heritage is presented in [1]. In [2], the authors present an illustrative example of a serious game for cultural heritage and in particular for museum environments focusing on the younger generation. Guidelines for effective design of serious games are presented in [3]. The authors offer an overview of the factors that make serious games effective in terms of maximizing the learning impact. A serious game application for cultural heritage entitled Admotum [4] was developed within the EU FP7 Network of Excellence “Virtual Museum Transnational Network”. This application presents Rome, Amsterdam, Alexandria and Sarajevo during the Roman period through a quest for objects placed invirtually reconstructed archaeological sites. It represents an efficient combination of storytelling and virtual reality and is very educational for users. Evaluation of serious games projects for cultural heritage is performed through user studies. In [5], the evaluation was related to the learning aspect of Icura, a serious game about Japanese culture and etiquette. The authors concluded that educational objectives should be embedded as game tasks and not simply transmitted by textual information. The initial findings of the Admotum evaluation have shown the great appreciation of users towards this concept, a high level of learning through the game and a strong feeling of immersion.

III. Tašlihan cultural heritage site

Tašlihan was the largest inn in Sarajevo during the Ottoman period. It was built between 1540 and 1543 as an endowment of Gazi Husref Bey, governor of the Bosnian province within the Ottoman Empire. Presently, only one wall of Tašlihan remains standing (1b) inside the garden of Hotel Europe. The only visualization of Tašlihan’s original appearance can be seen as part of the physical model of Sarajevo Old Town in the 14th century, which is exhibited at the Museum of Sarajevo (1a).

A. *Interactive digital story*

We created the virtual reconstruction of Tašlihan (1c) in Maya and exported it to Unity, where textures and illumination were adjusted and optimized for online use. The geometry of the model is based on the scientific work of archaeologists and historians who excavated the remains of the building.

In order to breathe life into the 3D geometry, we created an interactive digital story that combines movies with the interactive virtual environment of Tašlihan [7]. The method of hyper-structured interactive digital storytelling that we developed for this project is presented in [8]. Here, we will briefly summarize the main characteristics of that method and its implementation.

Hyper-structured interactive digital storytelling is based on structuring the story scenario so that it consists of the main story offering a short overview of the topic and sub-stories presenting particular important aspects, previously mentioned in the main story, in more detail. The main story is hyperlinked with sub-stories and users can choose the parts they want to watch or come back to watch the parts they missed. Apart of sub-stories, interactive virtual environments can also be linked in this hyper-structure.

In the Tašlihan application, the main story is divided into seven thematic clusters, each shortly presenting one aspect of the Tašlihan building (endowment, the builder, the decay etc). After watching each of these clusters, users can click on the link to the corresponding sub-story and/or the interactive virtual

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

Fig. 1: Tašlihan: a) physical model, b) remains, c) interactive virtual model

model of the site. The main objectives of this method are: optimization of the time spent by the viewer in the application; adjusting the form of the narrative to concepts familiar to a contemporary Internet-era audience; and combining different media into a unique digital storyline.

B. Serious game

Much like the interactive digital story, the Tašlihan serious game application [7] is implemented in JavaScript using Unity. The difference between the two applications is that in the interactive digital story, the virtual model of Tašlihan is linked to a sub-story and can be viewed if the user selects that sub-story. In the case of the serious game application, the user can see the interactive model only if he/she answers correctly the majority of questions from a quiz offered after watching the interactive stories. The questions are related to the information presented in the stories.

IV. Comparative user evaluation

Apart from evaluating the success of both Tašlihan applications, the purpose of the preliminary user study we performed was to determine whether users prefered the interactive digital story or the serious game and why.
Our preliminary user study was performed on 8 users, out of which 7 were Bosnian and one was Italian. All of them had better than average computer literacy and no visual or hearing impairments. The users were interviewed after watching the interactive digital story and/or playing the game. They were asked to give detailed comments on the learning aspect of the applications, level of immersion they felt during use and their appreciation of the hyper-structured storytelling concept compared to linear storytelling. At the end of the interview, they were asked which application they preferred and why.
All users highly appreciated the learning aspect of both applications. They stated they learned new facts about the history of the building which they had no opportunity toattention to the information presented in the stories. All users felt immersed in the past while watching the stories. They said that the feeling of immersion was enhanced by the fact that the storyline contains a character/narrator, Murad Bey, a friend of Gazi Husref Bey, the sponsor of the building. From these results we can conclude that the serious game adds educational value to a cultural heritage presentation.

V. Conclusion

In this paper we described two forms of virtual heritage presentation of a site which today exists only as remains. Both the Tašlihan interactive digital story and the serious game have kindled the interest of users to learn about the past and brought the building to life in their collective memory. The serious game application was found to be more educational than the interactive digital story, because the possibility to see the 3D reconstruction of the building after correctly answering quiz questions motivated users to pay more attention to the information presented in the stories.

References

[1] E. F. Anderson and L. McLoughlin and F. Liarokapis and C. Peters and P. Petridis and S. de Freitas *Developing serious games for cultural heritage: a state-of-the-art review*, Virtual Reality. December 2010, Volume 14, Issue 4, pp 255-275

[2] Anastasios Doulamis and Fotis Liarokapis and Panagiotis Petridis and Georgios Miaoulis, *Serious Games for Cultural Applications*, Intelligent Computer Graphics 2011, Studies in Computational Intelligence Volume. Volume 374, 2012, pp 97-115

[3] C. E. Catalano and A. M. Luccini and M. Mortara, *Guidelines for effective design of serious games*, International Journal of Serious Games. Vol. 1, No. 1, 2014

[4] S. Pescarin and D. Pletinckx and W.M. Hupperetz and M. Farouk and S. Rizvic et al, *Keys to Rome: Roman culture*, Virtual museums, exhibition catalog Rome, ISBN 978-88-902028-2-7 Edizioni CNR ITABC

[5] M. Mortara and C. E. Catalano and G. Fiucci, M. Derntl, *Evaluating the Effectiveness of Serious Games for Cultural Awareness: the Icura User Study*, In Proceedings of Second International Conference GALA 2013, Paris, France

[6] Sarajevo Graphics Group 2015, *Tašlihan interactive digital story*, http://h.etf.unsa.ba/han/sd/Build.html

[7] Sarajevo Graphics Group 2015, *Tašlihan serious game application*, http://h.etf.unsa.ba/han/game/

[8] I. Prazina and S. Rizvic, *Unity Hyperlinked Interactive Digital Storytelling*, In Proceedings of Central European Seminar on Computer Graphics CESCG 2015, Smolenice, Slovakia

[9] C.B.Seaman, *Qualitative Methods in Empirical Studies of Software Engineering*, IEEE Transactions on Software Engineering Vol. 25, No. 4, 1999.

The Development of TASTER, a Cognitive Training Game Using Human-Centered Design, Tailored for Children with Global and Specific Cognitive Impairments

Nigel Robb[1], Annalu Waller[2], Kate A. Woodcock[1]

[1]School of Psychology
Queen's University Belfast
Belfast, UK

[2]School of Computing
University of Dundee
Dundee, UK

Abstract—Children with Prader-Willi syndrome often exhibit challenging behavior in response to changes to routine. This phenomenon has been linked to a deficit in task switching ability which has been observed in children with the syndrome. TASTER is a cognitive training game which is being designed with input from a group of children with Prader-Willi syndrome, which aims to train task switching ability and thus reduce associated challenging behavior.

Keywords—human-centered design; serious games; Prader-Willi syndrome; cognitive training; task switching

I. Introduction

Prader-Willi syndrome (PWS) is a rare genetic disorder usually associated with mild to moderate intellectual disability (ID) and resistance to change [1]. In particular, temper outbursts due to changes to plans or routines have been linked to a deficit in task switching ability in individuals with PWS [2], [3], [4]. Video games have produced gains in specific cognitive processes in several typical and atypical populations [5]. Additionally, research in cognitive neuroscience has demonstrated a link between playing action video games (AVGs) and an increase in task switching ability [6], [7].

II. Specific Application

We are developing TASTER, a video game designed to train task switching in children with PWS. By increasing task switching ability, we aim to decrease occurrences of temper outbursts associated with changes to plans or routines. As such, our game aims to have a positive effect on clinically-significant behavior by targeting a specific cognitive impairment that has been linked to the behavior.

III. Design and Development

It is essential that the game produced is tailored to the preferences and capabilities of children with PWS, who – due to cognitive impairments and behavioral traits associated with the syndrome – may not be capable of playing, or motivated to play, typical entertainment games. It is also important to fully utilize the insights afforded by the aforementioned research on the cognitive benefits of AVGs. Therefore, our iterative development process incorporates both human-centered design in the form of observational studies, and a domain analysis of AVGs. Observational studies took place before any design decisions had been made (using existing entertainment games), and will continue throughout development (using prototypes). The children can therefore be considered design informants [8].

A. Observational studies

1) Participants and procedure: Seven children with PWS (age 7 – 15; six female, one male) played a selection of existing games, in order to observe how they interacted with existing technology [8]. Children played the games during sessions lasting at least 30 minutes, with multiple sessions taking place over two weeks. To overcome difficulties faced by children with ID when expressing their preferences [9], most questions were presented as a forced choice between two alternatives, with images used where possible.

2) Results: Six children preferred playing games on a handheld tablet. Children preferred games in which they were tasked with collecting items (e.g. coins) (Table I), and preferred controlling games using simple touch gestures (Table II). Scores were calculated from responses to forced-choice questions comparing features; scores show the number of times a feature was preferred as a proportion of the total number of times the feature was considered.

TABLE I. Overall gameplay preferences

	Gameplay feature			
	Collecting items	***Controlling characters***	***Shape-based puzzles***	***Building and creating***
Score	0.64	0.57	0.48	0.36

TABLE II. Overall preferences for control system

	Control system				
	Touching screen	***Keyboard***	***On-screen buttons***	***Mouse***	***Moving the tablet***
Score	0.79	0.71	0.48	0.29	0.14

The authors would like to thank the Foundation for Prader-Willi Research, for sponsoring the research.

B. Domain analysis

While AVGs show benefits in task switching ability, typical AVGs are not suitable for children with ID due to the complexity of the gameplay, the use of themes unsuitable for young children (e.g. violence) and the complexity of the control systems. The domain analysis aimed to identify features of AVGs that could be incorporated into a game suitable for children with ID, for example:

- Monitoring on-screen information during gameplay.
- Adapting goals due to unforeseen circumstances.
- The presence of distracting stimuli which must be filtered and ignored.

C. Conceptual design

A game was designed in which players must control a character to collect items (the preferred gameplay identified during observational studies). The items are creatures which can be identified in terms of their color (red or blue) or shape (cuboid or pyramid) (Fig. 1). Players must collect only one kind of creature at a time, with the current target specified as either a shape or a color; by changing how the target is specified, players are forced to switch how they conceptualize the creatures (the task switch). The game incorporates the three features of AVGs mentioned above: information panels and messages are displayed on screen; large rocks, which must be avoided, appear at random intervals in the player's path; and distracting stimuli occur in the form of sound effects, particle effects, and a screen shake effect. The player's character is controlled by touching the screen at the location which they wish to move to, in line with the preference for simple touch controls expressed by children in observational studies.

D. Prototype

A prototype game was developed for handheld tablets, the preferred distribution platform of the children who took part in the observational studies.

E. Playtesting

1) Participants and procedure: To date, five children with PWS (age 7 – 15; three female, two male) have tested the game. Children played the game for at least 30 minutes, and answered questions designed to gauge the usability of the prototype, their motivation to play it and their level of comprehension of the gameplay.

2) Results: Initial results show that players achieve a high level of satisfaction from the game, are motivated to play it, and find the interface and gameplay easy to understand.

IV. Conclusion

Children with ID can play the role of design informants [8] in the design of cognitive training games. Furthermore, the design of serious games in general can benefit from analysis of entertainment games, to identify features that might contribute to the effectiveness of the serious game.

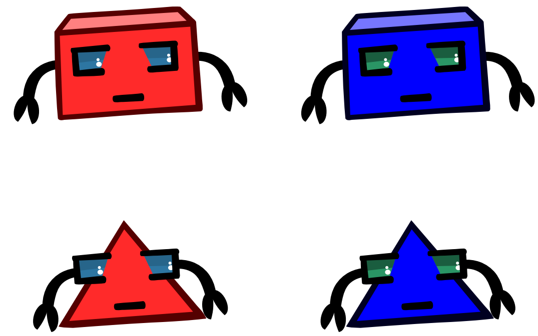


Fig. 1. The creatures which the player must collect. Each creature can be identified in terms of its color or its shape.

V. Future Work

A. Development

The prototype will be refined in an iterative process, incorporating feedback from playtesting sessions, using further observational studies of children with PWS as they interact with the prototype game.

B. Validation

The final game will be evaluated for its effect on task switching ability in a placebo controlled, cross-over design over ten weeks, with participants' task switching ability assessed – using a custom battery of task switching assessments tailored for children with ID – at baseline, after five weeks, and after ten weeks.

References

[1] J. Whittington, 'Population prevalence and estimated birth incidence and mortality rate for people with Prader-Willi syndrome in one UK Health Region', *J. of Medical Genetics*, vol. 38, no. 11, pp. 792-798, 2001.

[2] K. A. Woodcock, C. Oliver and G. Humphreys, 'Task-switching deficits and repetitive behaviour in genetic neurodevelopmental disorders: Data from children with Prader–Willi syndrome chromosome 15 q11–q13 deletion and boys with Fragile X syndrome', *Cognitive Neuropsychology*, vol. 26, no. 2, pp. 172-194, 2009.

[3] K. A. Woodcock, G. Humphreys, C. Oliver and P. Hansen, 'Neural correlates of task switching in paternal 15q11–q13 deletion Prader–Willi syndrome', *Brain Research*, vol. 1363, pp. 128-142, 2010.

[4] K. A. Woodcock, C. Oliver and G. Humphreys, 'The relationship between specific cognitive impairment and behaviour in Prader-Willi syndrome', *J. of Intellectual Disability Research*, vol. 55, no. 2, pp. 152-171, 2011.

[5] T. Klingberg, H. Forssberg, and H. Westerberg, 'Training of working memory in children with ADHD', *J. of Clinical and Experimental Neuropsychology*, vol. 24, no. 6, pp. 781-791, 2002.

[6] C. S. Green, M. Sugarman, K. Medford, E. Klobusicky and D. Bavelier, 'The effect of action video game experience on task-switching', *Comput. in Human Behavior*, vol. 28, no. 3, pp. 984-994, 2012.

[7] C. S. Green and D. Bavelier, 'Learning, Attentional Control, and Action Video Games', *Current Biology*, vol. 22, no. 6, pp. R197-R206, 2012.

[8] A. Druin, 'The role of children in the design of new technology', *Behaviour & Inform. Technology*, vol. 21, no. 1, pp. 1-25, 2002.

[9] H. Prosser and J. Bromley, 'Interviewing people with intellectual disabilities', in *Clinical Psychology and People with Intellectual Disabilities*, 2nd ed., E. Emerson, K. Dickson, R. Gone, C. Hatton, J. Bromley and A. Caine, Eds. New York: John Wiley & Sons, 2012, pp. 107-120.

Author Index