iLRN 2016 Santa Barbara

The Second Immersive Learning Research Network Conference

Colin Allison
Leonel Morgado
Johanna Pirker
Dennis Beck
Jonathon Richter
Christian Gütl (Eds.)

ISBN (e-book) 978-3-85125-472-3 **DOI 10.3217/978-3-85125-472-3**
2nd edition

Immersive Learning Research Network

Special Track 2: The Future of Education

An Online Immersive Reality Innovation-Lab

Vic Callaghan[a1], Marc Davies[2], Shumei Zhang[3]

[1]University of Essex, UK
[2]Creative Science Foundation, UK
[3]Shijiazhuang University, China

Abstract. This paper introduces the concept of online '*Innovations-Labs*' (i-Labs) as location-independent collaborative ideation spaces. We highlight the challenges and opportunities that disruptive innovations present to companies and society, and discuss how *Science Fiction Prototyping* and *Diegetic Innovation Templating* can provide a means to explore that space by acting as ideation process and a language for capturing and communicating innovations. A core hypothesis of this paper is that there are significant gains to be accrued from integrating *Virtual Reality*, *Science Fiction Prototyping*, *Diegetic Innovation Templating* and *Innovation Labs* to form an online immersive reality innovation-lab which both offers better affordances and access to people wishing to undertake innovation related activities. We present details of our initial implementation of an online innovation-lab (*Our HEX*) which takes the form of a virtual-reality space-station. We then conclude the paper by describing future directions of our work, principally, a venture which uses '*Our HEX*' space-station platform, plus a supporting textbook published by Tsinghua University Press, to teach '*English, Computing and Creativity*' to Chinese students. Finally the paper concludes with a summary and reflections on our work to-date.

Keywords: Virtual-Reality, Innovation-Labs, Ideation, Innovation, Science-Fiction Prototyping, Diegetic Innovation Templating, Creative-Science, EFL.

1 Introduction

It is generally agreed that innovation is an essential component for economic growth and productivity. A recent report by PriceWaterhouseCoopers, the largest professional services firm in the world, found that "*Five years ago, globalisation would have been the most powerful lever for growth and every business would have been talking about China. But now, the growth lever that has the greatest impact is innovation. Ninety three percent of executives tell us that organic growth through innovation will drive the greater proportion of their revenue growth*"[1]. Thus it's hardly surprising to find that governments around the world place a huge importance on supporting innovation activities although how they do that varies widely, depending on various political and financial factors. While innovation sometimes appears to be rooted in the individual (eg Steve Jobs) from a government perspective it is a product of a *National*

Innovation System (NIS) that includes all economic, political and other social institutions affecting innovation (eg education, financial structures, regulatory policies, labour markets; culture etc). For example, China operates a NIS derived from their 15-year national plan *(2006-2020)*, the '*National Outline*', which contains a section that focuses explicitly on creating nation-wide structures favourable to innovation [2]. In contrast the USA has not adopted a centralized approach rather, being a country that grew out of the notion of free enterprise and thinking, innovation was more easily established as it was part of the underlying 'DNA' of American culture. That is not to say that government policy does not play a role in fuelling American innovation, just a lesser one than in most other countries. It is difficult to measure a countries innovation capacity but one metric is the number of patents that are registered annually. Those statistics place the EU, USA, Japan and China in leading positions, aligning well with their economic performance. Because of the importance of innovation to companies and national economies, there is a huge incentive for companies to find tools that can aid the process of innovation. Once such tool is *Science Fiction Prototyping*, an ideation and communication tool that was first proposed by Brian David Johnson while he was working for Intel Labs in Portland. The basic principle of the method is that the stakholders of the innovation create futuristic fictions as a means of unleashing their imagination plus communicating and testing the ideas [3].Another tool is *Diegetic Innovation Templating* which uses existing fiction as an inspiration for new innovations (eg the flip-phone being inspired by the Star-Trek communicator) [4]. Innovation works better with a group of people where they can spark ideas off each other and the limited knowledge of an individual can be supplemented by others. One popular group-based approach is the *Innovation-Lab* (i-Lab) which offers a specially designed environment that is conducive to creative thinking [5]. For example, i-Labs provide participants with a relaxed comfortable setting where they can contribute ideas anonomously during ideation sessions. Generally, i-Labs require the participants to be physically present in the same location. However, the advent of vitual-reality has opened up the possibility of an i-Lab being located online in a virtual space which allows participants to be locacted anywhere in the world, and to ulilise tools that would not exist in the physical world. Thus, this is the aim of the work in this paper, to explore the potential arising from combining i-Labs, virtual-reality and science-fiction prototyping, diegetic innovation templating to create a novel online innovation facility which will be described in the following sections.

2 Related Work

2.1 Innovation Labs

An innovation-lab (i-Lab) has been described as an "*inspirational facility designed to transport users from their everyday environment into an extraordinary space encouraging creative thinking and problem solving*" [5]. The i-Lab concept was based on a model created by the UK Royal Mail's '*Futures and Innovation Group*' in 1997 for the purpose of helping their management teams brainstorm future possibilities. In doing this it became apparent that the interactions within the groups, together with the conversational and session management tools played a significant role in the effectiveness of the sessions, leading to the idea for providing specialist environments to support these activities.

In transferring the i-Lab concept from the original Royal Mail environment to the wider world there have been three notable projects. The first was the '*Learning the Habit of Innovation: Harnessing Technology for Strategic Planning*' (LHI) which was a collaboration between the UK Royal Mail and the universities of East Anglia, Cambridge, Essex, Bedfordshire plus Anglia Ruskin University. It was operated out of the University of East Anglia from 2001-2004 and funded by the Higher Education Funding Council for England [6]. The project sought to transfer the i-Lab model created by the UK Royal Mail into higher education and involved formalising a template that would form a minimum set of conditions to recreate an innovation environment. In brief they deduced that an i-Lab required three interlinking components namely the *environment*, the *technology* and the *facilitation mechanisms* to make it suitable for ideation and innovation activities. Furthermore, they determined that an iLab session comprised some mix of the following activities (most electronically supported):

- Icebreaker and reviver activities
- Discussion & getting other people's perspectives
- Brainstorming & voting
- Headlines, cut & paste collages and PowerPoint presentations
- Wall activities (collaborative writing, doodling etc)
- Scenario building
- Role play

They emphasised that creative thinking was not necessarily a rational, linear process and that revisiting and refining ideas could be a productive way to progress. At the core of the process was brainstorming, a technique for unleashing a flood of thoughts driven by members sparking ideas off each other, or carefully injected external stimulus. Having generated sufficient ideas a group would go on to categorise, rationalise and vote on the suggestions. Implementing the ideas is more challenging and occurs beyond the i-Lab session.

The two other notable ventures were EU Leonardo da Vinci collaborations between educational institutions from Poland, Greece, Romania and Turkey, coordinated by the University of Essex in the UK around two projects, namely '*The European i-Lab Competences Development Programme*' (2006–2008) and '*The Innovation Laboratories for the Quality Assurance of Vocational Education and Training'* (2012-2014) [7]. These projects led to the establishment of three innovation laboratories in Poland, Turkey and Romania and the production of a standard guide for i-Labs, namely the '*Innovation laboratory – Good Practice Guide*' [8] all of which aimed at the promotion of i-Lab use throughout Europe which, today, has resulted in over 100 globally-located i-Labs (from social to technical) created by organisations as diverse as the Standard Bank, Walmart, John Lewis, the UK National Health Service, Ryan Air and government (eg New York's 'Public Policy Lab' or the 'Social Innovation

Lab for Kent') [15]. In respect of this paper, one of the most significant i-Lab developments has been the introduction of web-based software which provides a much more efficient (and faster) ideation process together with providing an anonymity component [9]. Moreover, this computerisation has enabled i-Labs to move into Cyberspace, allowing participants to be freed from the need for physical co-location, a feature we build on in our online version of an i-Lab (*Our HEX*).

In our work, we use brainstorming as part of a product-innovation process called *Science Fiction Prototyping* that will be explained in the following section. In this we adopt a procedure procedure called an *Imagination Workshop* which was first proposed by Wu in 2013 and is similar to the brain-storming process used in an i-Lab except it uses science fiction and fantasy ideas to extrapolate forward current technologies, business and social practices by ten-plus years [10]. These concepts will be explained in the following section.

3 Creative Science

Creative Science refers to creative methods for supporting science, engineering, business and socio-political innovation through various imaginative activities. For the purposes of this paper those mostly concern *Science Fiction Prototyping* (SFP) and *Diegetic Innovation Templating* (DiT).

3.1 Science Fiction Prototyping

As was mentioned earlier, *Science Fiction Prototyping* was proposed by Brian David Johnson, Intel's then Futurist, as a response to a particularly difficult innovation challenge Intel faced in designing new generations of integrated circuits. Their challenge was that it takes between 7-10 years to take an integrated circuit from concept through to production and, during that period, there can be as many as 6 generations of potential applications for it. For example, new models of mobile phone can be released as frequently as every 18 months. Thus, chip designers needed to anticipate applications 7 years' ahead of specifying a chip (and possibly longer as the applications may live on for another 15 or more years) which, in a rapidly changing world, presents a formidable challenge! Of course an even bigger worry is the risk of disruptive technologies coming along. Thus, there was a compelling case for Intel to find a creative-thinking process that might come to their aid. Their solution was *Science-Fiction Prototyping*. Essentially, the method involves writing short fictional stories that imaginatively extrapolated current practices forward in time, leaping over incremental developments, exploring the world of disruptive product, business and social innovations. Because *Science-Fiction Prototyping* adopts a rich story-based structure it was able to create high-fidelity analogues of the real word, enabling it to act as a type of prototype to test the idea. Moreover, being a story it was accessible to anyone (aka the old adage '*everyone likes a story*') making it a perfect vehicle for conversations between all the stakeholders of the innovation, including society at large (the customers of innovations). The outcomes of *Science-Fiction Prototypes* are used to create new kinds of products, businesses or socio-political structures etc.

3.2 **Science Fiction Prototypes Style**

The most common size for a *Science Fiction Prototype* is 6-12 pages (referred to as a *mini-SFP*) which is of a similar size to a conference paper [10]. However, 6-12 pages can take many days to write so for innovation sessions, that need to take place in less than a day, an even shorter form of *Science-Fiction Prototype* was developed; the *Micro-SFP* (or µSFP) [11] which will be described in the following section.

3.3 Science Fiction Prototyping Workshops.

Typically, science fiction prototyping based innovation sessions take the form of an *Imagination Workshop* [14]. It involves gathering together a group of participants, specifying a goal (eg a new business or product etc), providing a context (eg business, home etc), setting a timeline (eg usually 10+ years into the future) and offering support for brainstorming about possible futures. A World Café approach is adapted to stimulate brainstorming and discussion with participants being placed in small groups (eg 5-7 members). Most other aspects are similar to an i-Lab.

3.4 **µSFP- A Shorthand Innovation Language**

There is no agreed specification for micro-fiction but, given the close relationship of *Science Fiction Prototyping* to technology perhaps it is not surprising to discover a popular size for a *µSFP* is one that fits mobile phone text (160 characters) or Twitter messages (140 characters) which, in English language, equates roughly to 25-30, words. Since *µSFPs* are short, they have the advantage of being quick to write, enabling users to capture and create many ideas in a short time period, in a similar timescale to brainstorming. Thus, *µSFPs* are seen as being complementary to brainstorming, providing a means to wrap a brainstormed idea in a more story-like framework which provides added meaning. From another perspective *µSFPs* are an interim step between a raw idea and a full *Science Fiction Prototype*. By way of an illustration of the principle of *µSFPs*, consider the following example:

> *Zoe, you've been my life-long friend on* <u>SentiBook</u>*; today the news feed reports most social network friends don't exits, are you real?* (22 words, 133 characters)

This *µSFP* extrapolates forward in time the current trend of companies adopting ever-more more automated customer call handling systems but explores the consequences of such technology reaching out more widely, for example into email and social messaging systems. It raises the question about whether we will know, or even care, if the parties we are communicating with are real or artificial. In this particular example the *µSFP* observes that our lives are becoming increasing virtualised through, for example, friendships on social networks with people we may never have met physically. As AI advances, machines will be better able to mimic real people, raising all kinds of new opportunities and conundrums.

Following the creation of a *µSFP* the next step would be to expand it into a *mini-SFP* (a 6-12 page version with a rationale and comments), followed by the usual product development cycle involving pre-production prototypes etc.

3.5 Diegetic Innovation Templating

Diegetic Innovation Templating (DiT) is a process of extracting creative ideas (eg innovations) from fictions created for the purpose of entertainment, rather than for technology, social or business innovation. Thus they are typically science fiction or fantasy movies or TV series such as, for example, Star-Trek that taps into the creative abilities of great authors and filmmakers as source of creative ideas. The term '*diegetic*' is borrowed from film studies and refers to things which are embedded into a fiction, playing an integral role in the story, such as the use of a gadget by one of the characters, and seen through their eyes. The artistic nature of such productions makes them particularly useful for non-technical applications or for situations where writing bespoke fictions is not a good option. For example it has been used by one of China's leading fashion design houses (Sunfed) where it levers the advantage from popular fiction being embedded into socio-cultural contexts (ie the firms marketplace) aiding branding and marketing efforts [12].

3.6 Out of the Box and into 'Our HEX'

By way of a summary of this section, we introduced *Science Fiction Prototyping* and *Diegetic Innovation Templating* as tools to support the early ideation phase of the innovation process by providing a means to engage people's imagination in thinking '*out of the box*' about future possibilities. *Science Fiction Prototyping* also allows the ideas to be tested within a plausible narrative and provides a way of opening dialogues, independently of specialist domain knowledge, with all the key stakeholders. In the next section we will describe '*Our Hex*' a virtual spacestation which provides an online facility to host i-Lab activities based around the *Creative Science* concepts we have presented above.

4 **The Virtual Spacestation** (on online Innovation-Laboratory)

4.1 A Spacestation Based i-Lab

Since *Science-Fiction Prototyping* concerns thinking about high-tech futures, the idea to base the online i-Lab on a simulation of a spacestation was born. The first version was funded by the *Creative Science Foundation* as a way to explore the concept of '*free will*' raised in Brian Johnson's original *21st Century Robot* science fiction prototype [13]. Our current online innovation lab is a modification of that early virtual-reality spacestation and consists of a large central arrival area (Social Deck) leading to an, essentially, unlimited number individual rooms, each outfitted to resemble an i-Lab.

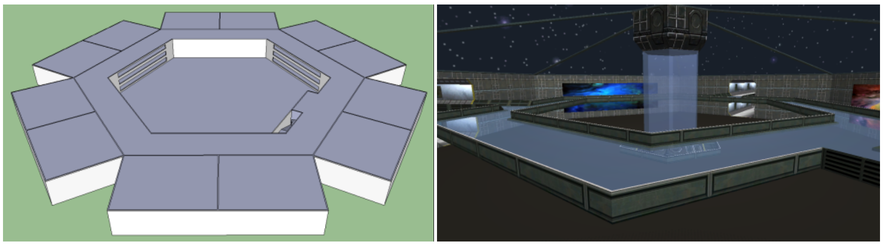

Fig.3. ‘*Our ‘HEX*’ Spacestation (Layout & Prototype Interior).

The spacestation structure was inspired by the *Hexagon Restaurant* (affectionately referred to as “*Our HEX*”) at Essex University (now defunct) which is shown with 6 pairs of i-Labs (Fig 1) but, in practice, since i-Labs are simply software instances, there is no fixed number as they can be created on-the-fly, as required. In keeping with the list of functionalities listed earlier, each simulated i-Lab includes a communal electronic white-board, a set of anonymised editing stations (so ideas and comments can be written to the white-board without identifying the writer) and facilitator tools for managing and archiving the sessions.

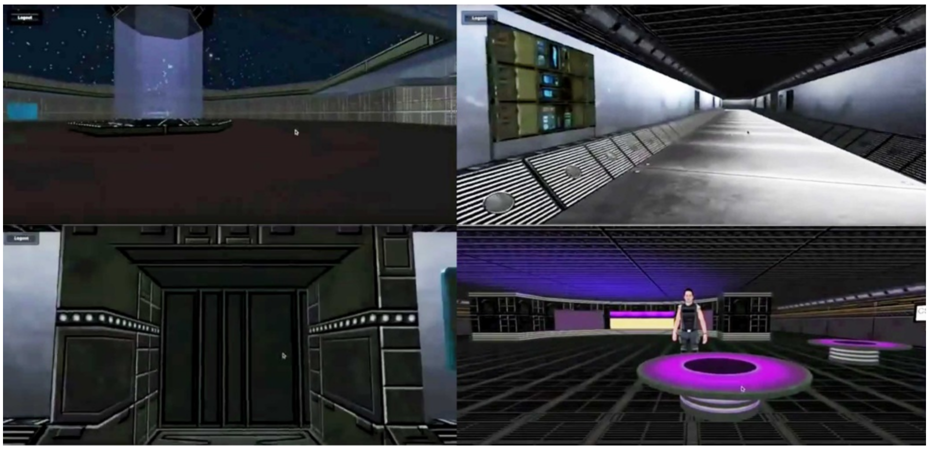

Fig. 2 – The Unity 3D Prototype iLab space station (clockwise from the top left there is the Social-Deck, one of the radial connecting corridors, an i-Lab entrance and a view of an i-Lab)

With reference to figure 2, each user who accesses the virtual world (ie logs in) first appears in the central arrivals area (the Social Deck). From that location they are free to walk around the environment; interacting with any displays they encounter (eg display boards showing outputs from earlier science fiction prototyping, diegetic innovation templating sessions, or interactive display boards where they can participate in competitions to evaluate innovation outputs, or just read notices of other events). The central area has corridors leading to each of the different i-Labs. In each i-Lab, users are able to participate in Imagination Workshop sessions (described earlier). Teachers and facilitators are able to observe, assist and rate student work.

The prototype of '*Our HEX*' was implemented using Unity-3D, an online gaming engine. Being an MMO cloud based virtual world, users are able to log into the environment via a link from the website of the Creative Science Foundation (CSf). The spacestation's i-Lab server resources are provided by a cloud based system. The execution-engine currently supports a Java runtime environment structured in a modularised client / server arrangement to facilitate future expansion. While a working prototype of the spacestation has been built (a video walkthrough is available at http://www.youtube.com/watch?v=-i6ki5YHGZc) there are a number of aspects that require completion before the system can be publically deployed, most notably creating a full gamut of i-Lab facilitation tools plus completing a formal evaluation with students. In addition the platform's user-guide needs to be integrated with the Tsinghua University Press textbook. Thus, '*Our HEX*' is a 'work-in-progress' task with functionality being added continually in response to user needs. To provide an insight to our immediate work-plans, the following section describes our next steps.

5 Deployment Plans

Currently '*Our HEX*' is being operated with a closed group of students at Shijiazhuang University, China, who follow a *Computer English* course [17] based on a carefully crafted Tsinghua University Press textbook [18].

By way of some background, in China it's mandatory for universities to teach "*Public English*" to all their students as this is seen as a necessary skill for them to thrive in a global business environment. For computer science students this requirement is translated into the provision of a specialized English module called '*Computer English*' that is usually delivered to students in their 3rd or 4th year [19]. By combining English Language with Computer Engineering, the course is made relevant to the student's studies [20] [21].

Beyond learning English, another vital skill for a workforce with aspirations to compete in global markets is an ability to innovate, which *Science-Fiction Prototyping* supports. Thus the proposition to integrate learning English Language, Computer Science and Innovation via an engaging new course was born, leading to a pilot trial being conducted by Zhang at Shijiazhuang University during the period 2014-2016 [16]. Following the success of this trial (student motivation and performance were demonstrated to sharply increase, with one student even publishing his SFP in an international workshop [22]) the team worked with Tsinghua University Press to produce a textbook that has been made available across China [18]. In support of this venture, we are planning to use the '*Our HEX*' spacestation platform as a means to widen access to innovation-lab facilities across China and the rest of the world. As part of this vision, in the longer-term, we plan to address other languages such as Spanish.

Thus, "Our HEX" functions as an online school to teach '*English as a Foreign Language*' (EFL) based around Creative Science, which brings the additional bonus of training students in creative thinking and innovation. In terms of the potential for this venture, the market for teaching English is estimated to be worth some $5 billion or more. In China alone there are an estimated 250 million English learners, increasing by 20 million per year, with a requirement for 1 million English teachers, which has led to the emergence of a plethora of enterprises seeking to satisfy these needs. Examples include *Ivy League English*, founded in 2009 by graduates of the Massachusetts Institute of Technology, which provides an app that connects students with USA-based business coaches for real-time roleplay activities (www.ile-china.com/), the 2013 Kickstarter funded start-up, *Influent*, that created a video game designed to introduce foreign vocabulary to learners by them exploring an interactive 3D environment filled with hundreds of selectable objects (www.playinfluent.com) through to full blown MOOCs learning platforms such as the Shanghai based *Hujiang* which has grown to over 90 million registered users since starting in 2001 (www.hujiang.com/). Hence, this venture joins a fairly crowded marketplace but differentiates itself by offering a novel combination of science, creative-thinking and language learning, especially tailored for university based Computer Engineering students through a supporting Tsinghua University Press textbook.

From the earlier sections it can be understood that creative science exercises English language by requiring students to read and write short stories plus undertake group work via brainstorming and presentations (and, as a by-product, getting other useful skills such as creative thinking and product innovation). Because, this involves group-work there is a space issue since, ideally, each group would have their own dedicated space (room). Clearly, in most situations that is impractical. For example, in the case of Shijiazhuang University's '*Computer English*' course, their 160 students would require some 23 rooms (assuming maximum group sizes of 7 students). Thus, '*Our HEX*' overcomes these space limitations as well as broadening participation to students, independently of their geographical location. In addition, given the virtual nature of the space, it is simple to outfit it with simulations i-Lab tools (ie an electronic white-board, anonymised editing stations and computerised facilitator tools) making it a virtual innovation-lab that can be replicated with little cost.

While our current focus is on creating an online "*English as a Foreign Language*" school we have been considering other longer-term possibilities for '*Our HEX*'. In terms of language training it would be possible to enrich the activities by including online role-play [23] [24]. Beyond language training, clearly one major application is as an online Innovation-Lab which would aim to satisfy the growing commercial demand for innovation services and we are working with a Taiwanese start-up, LivingPattern Technology Inc to explore these possibilities [25]. Other possibilities include collaborating with the Creative Science Foundation to host an online version of their vacation '*Entrepreneurship Schools*' (http://www.creative-science.org) or working with FortiTo Ltd to create online '*Maker Schools*' (www.fortito.mx)..

5.1 Deployment Platforms

A key issue is the cost of accessing this service. As a consequence we developed the system to work with a range of technologies to better fit the user's resources. These range from commonplace technologies such as mobile phones, pads, laptops and desktops, to more sophisticated devices such as virtual and augmented reality glasses (see figure 3).

Being a virtual-reality environment, '*Our HEX*' has the potential to simultaneously offer a number of different user experiences, depending upon how an individual chooses to interface and interact with the world. For example, whether the world is viewed from a first or third-person perspective can significantly alter the relative experiences of individual users, especially when working with others in team-based exercises. Furthermore, technologies such as VR headsets, (e.g. the *Oculus Rift*, or *HTC Vive*) could be used to generate a more immersive experience in the minds of users, allowing them to move around '*Our HEX*', with the impression of actually being transported inside the artificial world. Mixed reality interfaces, such as the Metavision's Meta-2 or Microsoft's *HoloLens* system, could also potentially be used to superimpose fragments of the spacestation onto the real world, effectively turning a physical room or other location into an extension of the '*Our HEX'* environment. Such an arrangement could facilitate interaction between groups of people where several are sharing the same physical space but wish to interact with other remote users present elsewhere in '*Our HEX'*.

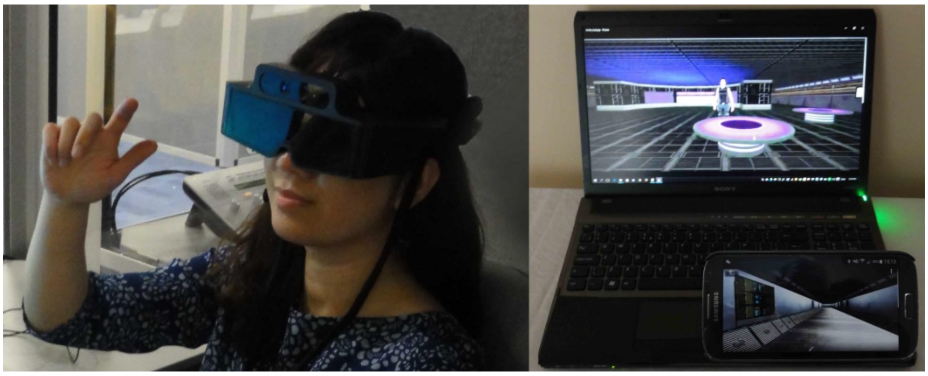

Fig. 3 – Some platforms for "*Our HEX*" (picture courtesy of Dan Chen)

As mentioned earlier, '*Our HEX* was implemented using *Unity 3D*, a professional tool used for the creation of computer games. The decision was made to use a game engine as an implementation platform in order to take advantage of some of the available graphics, physics, networking and other technologies developed by advancements in the computer games industry. Another reason was to give users some familiarity via a common interface, with many of the controls being identical to those used in PC games, (e.g. WSAD movement controls). By making the user as comfortable and immersed as possible in the '*Our HEX'* environment, their user experience should be enhanced and hopefully create a more productive innovation or education session. Other computer games technologies that may be beneficial to a learning/innovation environment are also being explored for potential integration with the '*Our HEX'* system. For example, live streaming services, such as *Twitch*, could be invaluable for a teaching experience, as users could both visually see a live representation of their teacher and provide feedback or ask questions via the text chat feature. From a business perspective, live streaming services could have potential benefits such as revenue generation from advertising and subscriptions or tips from users. Recordings of past broadcasts can also be played back on-demand by users.

6 Summary

This paper has described how we developed an online creative space which integrated *virtual reality*, *science fiction prototyping*, *diegetic innovation templating* and *innovation-lab* concepts to create a novel shared ideation space. We argued that the synergy derived from this linkage introduced significant new opportunities for those seeking to undertake innovation activities. For instance virtual reality both provides a more engaging and functional space, together with widening participation. We also argued that the inclusion of creative science tools provides a particularly good approach for exploring disruptive innovations as it levers people's imagination through the use of futuristic science fiction to offer more radical perspectives on the future. We also explained that a story based narrative provides an effective way to facilitate communication between professionals and lay-members of society, who frequently lack a shared vocabulary to converse (articulated by the mantra "*everyone likes a story*"). Finally we described how, in support of the book we have published with Tsinghua University Press in China, we are exploring the application of the '*Our HEX*' spacestation platform as an aid to students learning a combination of English language and innovation. Clearly this work is at an early stage and we will look forward to reporting on further progress in later conferences.

Acknowledgements

We are pleased to acknowledge financial support from the Creative Science Foundation (www.creative-science.org) that ennabled the development of the first version of the virtual spacestation.

References

1. R. Shelton, D. Percival, "*Breakthrough innovation and growth*", PricewaterhouseCoopers report, September 2013
2. C. Cai, R. Suttmeier, D. Simon "*China's 15-year science and technology plan*", Physics Today, 12th December 2006, 38–43.
3. V. Callaghan "*Creative Science: Injecting Innovation into the IT Industry*", Oxford Journals, ITNOW, Volume 57, Issue 2, pp. 52-55.

4. P. Zheng V. Callaghan "*Diegetic Innovation Templating*", Creative Science 2016, London, 12-13th September 2016
5. I. Susnea, E. Pecheanu, C. Tudorie, "*Initiatives towards an education for creativity*", The 6th International Conference Edu World 2014 "Education Facing Contemporary World, Issues", 7th - 9th November 2014., Elsevier Procedia - Social and Behavioural Sciences 180 (2015) 1520 – 1526
6. G. Gill, S. Oldfield, "*Scribbling on the Walls: Introducing Creativity at UEA, the story so far*", Building Cultures of Creativity in Higher Education Conference, University of Wales, Cardiff, January 8-10 2007
7. I. Susnea, E. Pecheanu, C. Tudorie, A. Cocu, "*The education for creativity–the only student's tool for coping with the uncertainties of the future*", Multidisciplinary Academic Conference on Education, Teaching and E-learning (MAC ETEL 2014), Prague , 16-17th October 2014
8. i-Lab2 Partnership "*Innovation Laboratory - Good Practice Guide*" Institute for Sustainable Technologies, Radom, Poland, October 2014, ISBN: 978-83-7789-322-7
9. L. Powell "*The where factor: New Environments for Innovative Thinking*", White paper available from https://www.facilitate.com/
10. B.D Johnson "*Science Fiction for Prototyping: Designing the Future with Science Fiction*", Morgan & Claypool Publishers, 2011, ISBN-10: 1608456552
11. V. Callaghan "*Micro-Futures*", Creative-Science 2014, Shanghai Jiaotong University, China, 30th June 2014
12. P. Zheng, V. Callaghan 'T*he Role of Science Fiction and Fantasy in Facilitating Innovation: Empirical Evidence from China*", Futures 2016
13. S. Egerton, M. Davies, B. Johnson, V. Callaghan, "*Jimmy: Searching for Free-Will*", Creative Science 2011, Nottingham, 26th July 2011
14. H.Y. Wu, "*Imagination workshops: an empirical exploration of SFP for technology-based business innovation*". Futures, 50(3), 44–55, 2013.
15. Rockefeller Foundation "*2014 Innovation Labs Survey*", Bridgespan Group / Rockefeller Foundation, September 2014,Bridgespan.org/Publications-and-Tools/Innovation-Labs-Insight-Center.aspx
16. S. Zhang and V. Callaghan "*Improving English as a Foreign Language Education in China with Science-Fiction & Immersive Reality*", 2nd Int'l Symposium on Future Intelligent Educational Environments and Learning (SOFIEE'16), London, 12-13th September 2016
17. S. Zhang, V. Callaghan, "*Using Science-Fiction Prototyping as a Means to Motivate Learning of STEM Topics and Foreign Languages*", Intelligent Environments 2014, pp.353-356, 2014
18. S. Zhang, V. Callaghan, H. Wang, B. Hedong "*Computer English Teaching Using a Creative Science Approach*, Tsinghua University Press, 2016, ISBN 9787302422730
19. X.F. LU, D.C.WU "*On Computer English Teaching*", Computer Knowledge and Technology, *36*, 085, 2009
20. M Xian, L. F. Zhou. "*Discussion and Research of Continuity Mode of Computer Professional English*". Computer Knowledge and Technology, 17, 058, 2010
21. J.I.A. Wen-Fang, "*Discussion on Teaching Models of Computer English*", Journal of Computer Knowledge and Technology, *5*, 063, 2009
22. P. Liu, S. Zhang, "*Some Thoughts About Future of Intelligent Buildings and Smart Transportation*", Workshop Proceedings of the 10th International Conference on Intelligent Environments, pp 247-255, Shanghai, China, 2-4 July 2014.
23. C.D Chen, F. Myles, V. Callaghan, "*The 21st Century Interpreter: Exploring the use of smart-glasses for technology-augmented interpreting*", Immersive Learning 2016 (iLRN'16), Santa Barbra, USA, June 27th – July 1st, 2016
24. M. Gardner, B. Horan "*Using virtual worlds for online role-play*" iED'11 (the 2010 Immersive

Education Summit), Universidad Carlos III de Madrid, Spain, Summit from 28th to 29th November 2011

25. H.Y. Wu, V Callaghan, "From Imagination to Innovation: A Process Point of View", Creative science 2016, London, 12-13th September 2016

Virtual Observation Lenses for Assessing Online Collaborative Learning Environment

Samah Felemban [a,b], Michael Gardner [a] , Victor Callaghan [a]

[a] University of Essex, Colchester, UK
[b] Umm Al-Qura University, Makkah, Saudi Arabia

ssyfel@essex.ac.uk

Abstract. The purpose of this paper is to introduce a new approach for assessing learning outcomes from collaborative work in 3D virtual environments. It represents a novel computational framework that improves recording and observing collaborative activities between students to evaluate learning outcomes. The framework includes a virtual observation model that maps observing learners in classrooms with observing and assessing the students in 3D spaces. This can be accomplished by applying a mechanism that combines natural agents and software agents to support collecting learning evidences from virtual activities and simulate the educators' observation(s). Such a novel framework will solve issues that could develop from evaluating students' performance, interaction, skill and knowledge in collaborative virtual learning environments.

Keywords: E-learning; 3D Virtual Worlds; Assessment; Virtual Observation; Collaborative Learning; Learning Evidence; Software Agents; Natural Agents.

1 Introduction

The power of networks and computers has invented technologies that support learning and connect geographically dispersed learners to enhance learning experiences. Several educational technologies have been widely applied that connect scholars and educators to provide different types of activities and to access learning sessions remotely without requiring physical attendance. By using online environments, organisations could easily educate learners and support collaborative learning without offering physical place or hiring educators.

A great technology that enables virtual collaborative learning is the immersive environment, the 3D virtual worlds (3D VWs). The 3D spaces are increasing in popularity because of many features that distinguish them from other online systems. They connect students in real-time and enhance interactivity, exploration, and engagement between them. Moreover, they facilitate investigation of ideas, situations and places that cannot be reached physically; delivering learning processes; providing

realism of interaction, discussions and activities of even the most complicated topics in simpler conditions with less cost.
Collaborative learning can help students to achieve learning through working with their peers, who support them to enhance their information and skills, resulting in constructing new knowledge and experiences. Learners usually obtain new knowledge while participating in learning sessions, so evaluating learners in a group should not be applied just after the last learning session, but it should also be applied during the learning process. Wells [1] also stated that educators should evaluate the whole learning process when performing collaborative learning activities rather than look at the final artefact as evidence of learning.

However, numerous issues can arise when assessing learning outcomes for a group of students in the 3D environments. Firstly, observing users' behaviour dynamically and collecting evidence of learning are complex tasks in VWs. Secondly, various skills, including communication and negotiation skills, can be gained from collaborative activities, but it is difficult to automatically detect evidence of them in these spaces. Thirdly, labelling and recognizing the evidence of many users in real-time is difficult because several students are contributing at the same time, which makes tracking the evidence much more complex. Therefore, finding an event detection method that can dynamically recognise users' behaviour, collect learning evidence data, and analyse events to measure the learning outcomes, is necessary. Gardner and Elliott [2] indicated that 'learning within technology creates a pedagogical shift that requires teachers to think about measuring outcomes in non-traditional ways'.

The purpose of this paper is to introduce a new approach for assessing learning outcomes from collaborative work in 3D virtual environments. It represents a novel computational framework that improves recording and observing collaborative activities between students to evaluate learning outcomes. The framework includes a virtual observation model that maps observing learners in classrooms with observing and assessing the students in 3D spaces. This can be accomplished by applying a mechanism that combines natural agents and software agents to support collecting learning evidences from virtual activities and simulate the educators' observation(s). Such a novel framework will solve issues that could develop from evaluating students' performance, interaction, skill and knowledge in collaborative virtual learning environments.

2 Related Work

2.1 Identifying Learning Evidence in Virtual Environments

Identifying learning evidence is simple in the multiple choice online test format, but it becomes more problematic in 3D VWs or educational games, because of the large number of observational variables and the complex relationship between these variables and students' performance [3]. Although technological improvements assist in recording data, even for difficult situations, understanding and analysing the composite data that results involves more complex processes.

Certain approaches have been used to assess modelling learners' skills and knowledge in simulation learning spaces. The approaches can be categorised into two groups: 1) knowledge engineering/ cognitive task analysis approach and 2) machine learning/data mining approach. The knowledge engineering approach formulates logical rules to assess and group particular students' behaviours. The rules are also applied to differentiate the level of students' skills such as the study by [4]. In the machine learning/data mining approach, learners' behaviours are recognised by analysing data and extracting learners' performance from the log files that are auto-generated while students are participating. For example, learning evidence has been collected through analysing users' log data by applying cluster analysis algorithms to determine the key feature of students' performance in educational game environments [5].

However, the log files save all the players' responses to the given educational problems which creates enormous amounts of data that provide a serious obstacle for researchers when collecting learning evidence from immersive environments [6]. This makes it very difficult to capture individual students' learning, knowledge, and skills and challenging to identify the actions and performance that represent learning. Moreover, collecting data in simulation or virtual environments without consideration of how the data will be assessed or scored is an ineffective method for creating assessments. Hence, designing the learning environment from the beginning to enable assessment and collecting learning evidence is more preferable [7].

Additional issue with identifying learning evidence is that technologies cannot capture all of the acquired skills. Several skills can be gained from collaborative activities, but it is complicated to automatically detect evidence of them [8]. For example, the quality of the interaction skills between students including teamwork, collaboration, negotiation, and communication are hard to measure with regular assessments. The study [9] proposed techniques that permits assessing learning outcomes (skills, knowledge, and competencies) by using elements such as smart objects and avatars in 3D spaces. However, these techniques lack in measuring the quality of learning in collaborative environments.

Analysing various users' behaviour/data, identifying the meaningful actions, and combining those actions into learning evidence to determine the learning outcomes are very complex processes in such environments. Consequently, discovering techniques that could dynamically recognise learning evidence and analyse events to measure the quality and quantity of learning outcomes is advantageous. Developing such mechanisms will help to identify and gather proof of learning during collaborative activities in immersive worlds and correlate the evidence with learning objectives, to assess the overall outcomes of the learning processes.

According to Thompson and Markauskaite [10], 'educators need to move beyond traditional forms of assessment and search for evidence of learning in the learner interactions with each other and the virtual environment, and artefacts created.' Hence, we have considered another assessment method such as classroom observation which greatly assists educators to evaluate students by collecting evidence about their learning. We have mapped the physical observation to the 3D spaces to provide more insights of what evidence could be collected from students' performance. Section (2.2) gives more explanation of the observation method in learning.

2.2 Observation

2.1.3. 'Teacher observation occurs continually as a natural part of the learning and teaching process and can be used to gather a broad range of information about the students' demonstrations of learning outcomes' [11]. Observation takes place in several settings and with a variety of methods. It can help teachers gather information about the individuals' and groups' behaviours and skills. To distinguish the observation levels in classrooms, Gray [12] introduced conceptual frameworks that follow educational standards to define the basic frames for observing. Because observing classrooms is very complex, he suggests that each teacher should select a specific frame or 'lens' to gain more insight into a specific classroom characteristic. Such 'lenses' are summarised in Table 1.

Table 3. The Observable Signs Pertaining to the Eight-Question Areas [12]

1. The learning climate	• Degree to which students can express their feelings and opinions • Frequency with which student responses are used and extended • Amount of interaction and sharing among learners
2. Classroom management	• Use of preestablished classroom rules • Use of instructional routines • System of incentives and consequences
3. Lesson clarity	• Frequency of examples, illustrations, and demonstrations • Percentage of students who can follow directions given • Use of review and summary
4. Instructional variety	• Use of attention-gaining devices • Changes in voice inflection, body movement, and eye contact • Use of a mix of learning modalities (visual, oral)
5. Teacher's task orientation and content presentation	• Orderliness of transitions • Teacher's preorganization of administrative tasks • Cycles of review, testing, and feedback
6. Students' engagement in the learning process	• Use of exercises and activities to elicit student responses • Monitoring and checking during seatwork • Use of remedial or programmed materials for lower-performing
7. Student success	• Number of correct or partially correct answers • Number of right answers acknowledged or reinforced • Number of delayed corrections vs. immediate corrections
8. Students' higher thought processes and performance outcomes	• Use of teaming, pairing, or other cooperative activities that encourage student problem solving • Display of student products and projects • Opportunities for independent practice and application

Adopting these 'lenses' when observing students can determine what could be evaluated and monitored when assessing students. They can help to observe students learning and to recognise the type of evidence should be collected when measuring the learning outcomes. Furthermore, creating a virtual observation hierarchy model to determine the granularity levels of observing learning activity in collaborative virtual environments can assist designers and developers to identify the learning evidence that can be captured and help to apply it in the virtual environment. Suskie stated that ‘the more evidence you collect and consider, the greater confidence you will have in your conclusions about students learning’[13].

3 Proposed Observation Technique in 3D VWs

We propose the Virtual Observation Portal (ObservePortal), which is a 3D virtual environment that can track users' behaviour and capture real-time evidence from collaborative activities. The environment employs real classroom observation lenses and applies each lens to the virtual world. The observation level can be stated in the learning design by the teacher to identify which lens should be activated to evaluate the learners. It determines the levels of granularity for observing learning activity in virtual environments to capture the learning evidence, beginning with general observation to in-depth observation (more details in section 5.4).

To capture the learning events, the platform utilises some techniques from agent systems to track users’ actions and predict the learners' acquired skills and knowledge. It has two different types of agents: software agents and natural agents. The software agents track learners and collect different users’ clicks and actions, while the natural agents perform peer evaluations of each other to evaluate the quality of performance. These agents are employed to record both implicit and explicit data that will be analysed to determine the learning evidence and students’ performance. All agents will work together in real-time to collect the learners’ evidence (more details in section 5.3).

3.1 The Learning Environment

The virtual world (ObservePortal) is the environment in which the students will perform the activities. To implement the research prototype, the InterReality Portal will be used, a project developed by a member of the Immersive Learning Lab, Anasol Pena-Rios, at the University of Essex (Figure 1) [14]. It is built upon the Unity [1] platform, a flexible development platform for assembling 2D and 3D collaborative games and environments. The environment was developed using the C# programming language. We chose to apply the prototype within this environment because it supports collaborative programming activities and assists in setting up learning tasks that help students understand the concepts and functionality of embedded systems in smart homes.

1 https://unity3d.com/unity

Fig. 4. Graphical User Interface (GUI) – InterReality Portal [14]

4 Conceptual Framework

Based upon the literature, observing and measuring online collaborative learning outcomes, both dynamically and on the fly, within 3D virtual worlds is scarce. As a result, we have proposed a Mixed Intelligent Virtual Observation (MIVO) conceptual framework that mixes learning models and computational models for observing and evaluating collaborative learning in 3D VWs. The framework consists of five models: user, learning, observation lenses, mixed agents and presentation (Figure 2). Each model will be discussed in the following section.

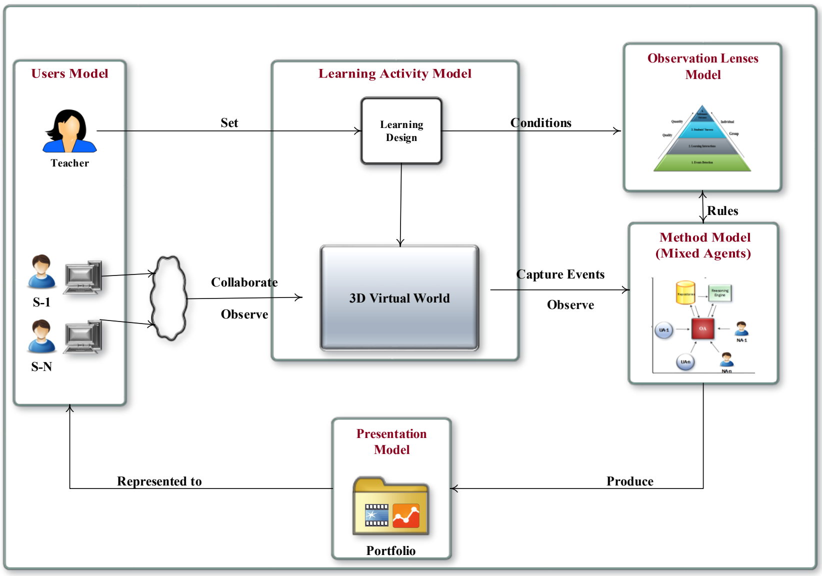

Fig. 5. Mixed **I**ntelligent **V**irtual **O**bservation (MIVO) Conceptual Framework for Collaborative Learning Environment

4.1 Users Model

This model identifies who the users are and their roles within the learning activity. Users will be either learners or teachers, and the specific user interface will be displayed based upon the user's identity and role. For example, instructors have a customisable interface that allows them to design learning activities. Moreover, a teacher can view learners' portfolios to evaluate their performances and review their work. From the learners' viewpoint, the user interface will enable them to interact with the environment and with other students' avatars. All participants will then work together on the simulation learning activities in the 3D environment. They can participate in the activities, evaluate others, obtain learning feedback from the system and view their portfolios.

4.2 Learning Activity Model

This model consists of two parts: the learning design and the environment that contains the collaborative learning practices. The learning design is defined as the learning scenarios that can be shared in the system and that can be planned and adjusted by the teachers. Moreover, the teachers can specify the observation criteria for evaluating the learning outcomes. Also, this model includes the virtual environment that students will participate in.

4.3 Mixed Agents Model (MixAgent)

This model identifies the method of gathering different types of evidence to illustrate individuals' and groups' learning outcomes. We expand the concept of software agents to include natural agents (users). The software agents will be needed to automatically track users' behaviour and collect data from real-time events as users interact with each other and with objects in the virtual world. Two types of software agents are used: user agents and ontology agent. In addition, the natural agents will be combined with them to enhance the capture of evidence. All agents, software and natural agents, will collaborate and work towards one central goal together, to produce evidence that evaluates the quality and quantity of students learning and performance (see Figure 3). In the following section, the agents' capabilities including their particular assessment abilities will be discussed.

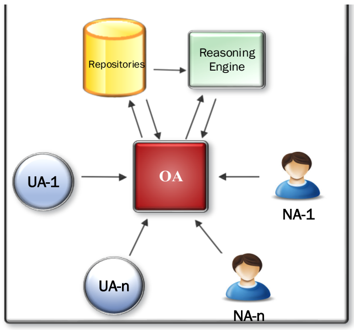

Fig. 6. Mixed Agents Model (MixAgent)

- **User Agents (UA).** These agents will be created once a student is authorised in the environment. There will be an agent for each learner. This agent can trace the user's actions in real time, translate any behaviour into data and send them to the ontology agent. They will monitor users' log data, behaviour and history.
- **Natural Agents (NA).** Peer evaluation could assist in capturing implicit learning evidence that is hard to capture with technology [8], and it would be useful to secure it from people directly to distinguish students' performance. To this end, learners will be considered natural agents. These agents can produce learning evidence by regularly assessing the quality of each other's communication, negotiation, teamwork, and active learning skills. While students are working together, there will be sliding scales scored from 1 to 5 will allow natural agents to act and rate other learners regularly. When the natural agents produce evidence and trigger the system, messages will be sent to the ontology agent. The ontology agent will receive the data and store them in the ontology repository. Employing natural agents will permit capturing the quality of learning outcomes that are too complicated to be identified by technology.
- **Ontology Agent (OA).** This agent is based on a semantic web and ontology approach that models different elements in the VW. Ontologies typically consist of object classes, the relationship between these objects and the properties that the objects have [15]. With ontologies, we can set up all the relationships between objects so that devices can understand the meaning of concepts. They can offer a standardised vocabulary to describe a knowledge domain by developing connected semantics and sets of vocabularies that can be reasoned. Thus, we have proposed this agent which has the ability to receive data from other agents and send them to the repositories. It will act as a communication agent and a bridge between all agents in the learning environment, so the collected data from other agents can be analysed based on logical rules that could assist in retrieving learning evidence. This agent will infer the relationship between the collected data and what it means in term of learning evidence through using a reasoning engine. Moreover, the logical rules will permit reasoning the repositories and parsing more meaning from the data gathered by each agent.

4.4 Observation Lenses Model (OLens Model)

This model determines how we can analyse the data that is captured by the agents. To observe the students in the classrooms, educators should consider numerous criteria, aspects and frames to gain more insight into the students' learning and improve their education. However, not all learning outcomes and skills mentioned can be easily observed and identified in virtual environments. Depending on the observation framework [12], we adopt particular 'lenses' to our model and applied them to the 3D VW to evaluate what could be monitored in these environments. The virtual observation model defines the levels of granularity for observing students and recording evidence of collaborative learning, commencing with high-level to low-level observation (see Figure 4). The observation layers are: events detection, learning interactions, students' success and performance outcomes.

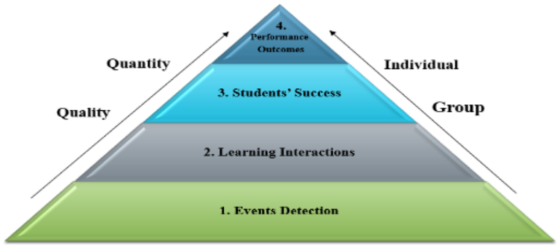

Fig. 4. Observation Lenses Model (OLens Model)

Describing the model lenses and their pedagogical meaning, beginning with the lower level of the hierarchy is Events Detection lens. This simulates an instructor when he/she watches a collaborative activity from high altitude, but without looking deeply into what is happening. In the VW, the automated observer monitors the activity by recognising that a sequence of events is occurring and capturing these events without judging. The second level is Learning Interactions lens, which considers a deeper view of the social and environmental interactions. In our case, the social interactions are between peers, and the environmental interactions are between students and the VW. Evaluating the quality and quantity of collaborations and interactions infers whether the learners have valuable interactions and if they are active learners in their groups. It determines the amount of sharing and interaction among students. The third level is the Students' Success lens. It represents teachers when they are observing the students' success by counting the number of correct answers, the number of right answers reinforced or acknowledged, and the number of delayed corrections. The fourth level is Performance Outcomes, which simulates the observer tracking the students in-depth to identify the skills and knowledge that they have acquired from the learning activities.

These frames help to measure the individual's and the group's performance, and the quality and quantity of each learning outcome. The following sections provide examples of how one can map some of the pedagogical lenses to collect evidence or to create logical rules that can be applied to the VWs.

- **Events Detection Lens.** This level focuses on observing the activity from a high level and collecting different events that demonstrate interactions between students and their surroundings. Examples of the events that can be observed and collected from students and group activities include the following:

  Avatar Actions:
  Avatar Log: <AvatarID, AvatarName, LogInTime, LogOutTime, Date, GroupNo>
  Chat Log: <AvatarID, DialogueTime, DialogueText>
  Touched Object: <AvatarID, ObjectID, ObjectName, TouchedType, Time>
  Rating: <AvatarID, RatedAvatar, RateScore, Time>

  Group Actions:
  Group Log: <GroupID, GroupMembers, StartTime, EndTime, Date>
  Group Dialogue: < GroupID, GroupDialogueText >
  GroupRating: <GroupID, GRateScore >

- **Learning Interactions Lens.** In this level, we are extending the teachers' judgements of group interactions in a physical setting to understand the interactions between the group and individuals in the virtual environment. It is possible to infer the quantity and the quality of the learners' interactions by creating rules based upon the teachers' viewpoints. Table 2 gives examples of the rules that can be created in this lens.

Table 4. Examples of the observation rules

	Quantity	**Quality**
Individual	The number of a learner's contributions in using the virtual objects during a period compared with other learners.	The rating scores for a student from other members in a period. 5 = Excellent; 4 = Good; 3 = Average; 2 = Fair; 1 = Poor
Group	The number of the group's contributions in the activities compared with other groups.	The average rating scores for all members in one group.

4.5 Presentation Model

The final model in the framework illustrates how evidence of the learning outcomes will be presented to teachers and learners. From the evidence gathered by agents and applied observation rules, the evaluation model will demonstrate how the performance of individuals and the groups was rated. The observation methods will allow analysing the learning outcomes from the activities and will correlate them to the learners’ portfolios. These portfolios can demonstrate students’ performances through any type of method, for example, it can include a feedback dashboard displaying when performance was either high or low, to allow teachers to evaluate the group as a whole and as individuals. Another example is that the performance could be reviewed by video snaps that map between time stamps of evidence and video recording to enhance the learning affordances of the immersive environment through visualising and reviewing the learning outcomes.

5 Conclusion and Future Work

In this paper we have introduced and described the Mixed Intelligent Virtual Observation (MIVO) conceptual framework for the collaborative learning environment. It consists of several models: user, learning activity, mixed agents, Observation Lenses (OLens), and presentation. The **MixAgent** and the OLens models play important roles to observe and recognise events that are occurring during the learning activity to evaluate the students learning.

This is a work-in-progress paper and there is much research still needed to be completed. Currently, we are commencing with the technical experimental phases to investigate the appropriate mechanism, based upon the complexity of observing and assessing learning in 3D VWs. The aforementioned collaborative environment, InterReality Portal, is used which allows students, worldwide, to participate in learning activities. In the future, the mixed-agents approach, namely, the combination of the natural agents (users) and software agents will be implemented to provide better results for collecting evidence and evaluating students. Hence, this phase will demonstrate how software agents can be combined with natural agents to improve the collection of learning evidence.

The next phase of the experimental phase will explore how to observe students' activities in the virtual world by applying methods from physical educational settings. The mixed agents approach helps observe and recognise events that are occurring during the learning activity and record them without evaluating the students. To analyse and translate these events, we will examine the frames of the OLens Model to create virtual observing rules that can infer learning outcomes in such environments.

The final experimental phase amalgamates all previous phases and explores the observation system implementation within the design of the collaborative learning activities, constructing learners' portfolios based on the evidence-gathering mechanisms, and analysing this data based upon the observation layers in the model in real-time.

Beside the experimental phases, the evaluation of our work is an essential component which is considered for the future progress. The research framework and models will be evaluated through user-based and expert-based evaluations. We are looking forward to report the results for the experimental and evaluation phases in future events and conferences.

Acknowledgment.

We are pleased to acknowledge Dr Anasol Pena-Rios for providing us with the great 3D virtual environment (InterReality Portal) and her technical support with this environment.

References

1. Wells, G., Action, talk, and text: Learning and teaching through inquiry. Vol. 16. 2001: Teachers College Press.
2. Gardner, M. and J. Elliott, The Immersive Education Laboratory: understanding affordances, structuring experiences, and creating constructivist, collaborative processes, in mixed-reality smart environments. EAI Endorsed Transactions on Future Intelligent Educational Environments, 2014. 14(1): p. e6.
3. Frezzo, D.C., et al. Psychometric and Evidentiary Approaches to Simulation Assessment in Packet Tracer Software. in Networking and Services, 2009. ICNS '09. Fifth International Conference on. 2009.
4. Schunn, C.D. and J.R. Anderson, The generality/specificity of expertise in scientific reasoning. Cognitive Science, 1999. 23(3): p. 337-370.
5. Kerr, D. and G.K. Chung, Identifying key features of student performance in educational video games and simulations through cluster analysis. JEDM-Journal of Educational Data Mining, 2012. 4(1): p. 144-182.
6. Mislevy, R.J., R.G. Almond, and J.F. Lukas, A brief introduction to evidence-centered design. ETS Research Report Series, 2003. 2003(1): p. i-29.
7. Tesfazgi, S.H., Survey on behavioral observation methods in virtual environments. research assignment, Delft Univ. of Tech, 2003.
8. Csapó, B., et al., Technological Issues for Computer-Based Assessment, in Assessment and Teaching of 21st Century Skills, P. Griffin, B. McGaw, and E. Care, Editors. 2012, Springer Netherlands. p. 143-230.
9. Ibáñez, M.B., R.M. Crespo, and C.D. Kloos, Assessment of knowledge and competencies in 3D virtual worlds: A proposal, in Key Competencies in the Knowledge Society. 2010, Springer. p. 165-176.
10. Thompson, K. and L. Markauskaite, Identifying Group Processes and Affect in Learners: A Holistic Approach to. Cases on the Assessment of Scenario and Game-Based Virtual Worlds in Higher Education, 2014: p. 175.
11. Maxwell, G.S., Teacher Observation in Student Assessment. 2001: Queensland School Curriculum Council.

12. Borich, G.D., Observation skills for effective teaching. 2016: Routledge.

13. Suskie, L., Assessing student learning: A common sense guide. 2009, San Francisco: Jossey-Bass.

14. Pena-Rios, A., Exploring Mixed Reality in Distributed Collaborative Learning Environments, in School of Computer Science and Electronic Engineering. 2016, University of Essex. p. 305.

15. W3C, Semantic Web, in Ontologies. 2015, W3C.

The 21st Century Interpreter: Exploring the use of smart-glasses for technology-augmented interpreting

Chantel Dan Chen[1], Florence Myles[1] and Victor Callaghan[1]

[1]University, of Essex
cchenn, fmyles, vic @essex.ac.uk

Abstract. In this 'work in progress' paper we set out the case for how smart-glasses can be used to augment and improve live Simultaneous Interpreting (SI). We do this through reviewing the relevant literature and identifying the current challenges faced by professional interpreters, such as cognitive load, memory constraints and session dynamics. Finally, we describe our experimental framework and the prototype smart-glasses based system we are building which will act as a testbed for research into the use of augmented-reality smart-glasses as an aid to interpreting. The main contributions of this paper are the review of the state of the art in interpreting technology plus the smart-glass experimental framework which act as an aid to Simultaneous Interpreting (SI). Later papers will report of other phases of our work.

Keywords: Simultaneous Interpreting, Translation, Languages, Augmented Reality, Smart Glasses, Meta, glossary-building, term extraction, multi-media learning, multitasking

1 Introduction

Interpreting is to orally translate the spoken words in language 'A' into language 'B'. Modern interpreting gained its professional status as early as the establishment of League of Nations, the forerunner to the United Nation [1], where interpreters were required to render oral languages between French and English, the two working languages of the organization.

Interpreters work in two different modes: consecutive and simultaneous. A consecutive interpreter listens to the source spoken language and renders it into the target language when the speaker stops for interpreters to deliver the messages to the listeners. A simultaneous interpreter renders the spoken language into the target language to the listeners in real-time while the speaker is delivering a speech. In this paper, we will only discuss simultaneous interpreting, as the smart-glasses will be applied to simultaneous interpreting only. Nowadays, simultaneous interpreters work in many different settings. International organizations, such as the United Nations and the European Commission, employ their own in-house interpreters, managed by a specific department (United Nationals DGACM n.d.), which oversees management of interpreting services for their on-going programme of international conferences and meetings.

Interpreting services are considered an ancillary service of the Meeting Incentives Conferences Exhibitions (MICE) industry [3]. Along with the development of MICE industry around the world [4], in order to engage multi-national participants in conferences and meetings, there is a growing need of professional interpreters. As such, there are already a large number of freelance interpreters, especially in the mega cities, providing interpreting services to international conferences, seminars and multi-language meetings.

The growing trend and demand are reflected by the university education system. In China alone, more than 100 universities have master level interpreters' education programmes. In the UK, the U.S and the European countries more and more universities provide master level interpreters' education. In order to provide a near-native working environment, universities invest large amount of funding in building interpreters' lab with a conference setting with a large conference table and delegate positions. The conference participants listen to the interpretation at the delegate positions through headsets.

2 Simultaneous interpreters' technical working environment

2.1 Inside the simultaneous interpreter's booth

The physical working environments of simultaneous interpreters are fixed and mobile booths. Simultaneous interpreters usually work in pairs in a booth (Fig.1). Each booth is set up with two user consoles (Fig. 2), which are each provided with a microphone and a headset. Interpreters listen to the source language through the headset and deliver the interpretation via the microphone at the same time. The interpreters take turns to interpret at every 20 – 30 minutes. The listeners outside booth listen to the interpretation from the wireless receivers or at the delegate positions. All the audio feeds are connected to a mixing console which is controlled by an audio-visual technician on site.

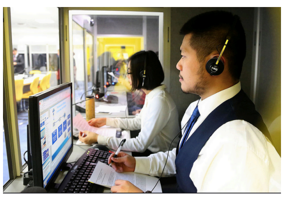

Fig. 1. Interpreters working in pair in a booth **Fig. 2.** Interpreter's console
(the Interpreting Lab in the University of Essex)

In order to maintain the quality of an interpreters' working environment, ISO - standards [5] have been established for both mobile booths and fixed booths. The European Commission [6] has also published a technical specification for booths in conference rooms. The standards and specifications require a booth technician onsite to guarantee the two-way communication in and outside the booth. Three core metrics aim to reduce unnecessary cognitive load on the interpreters' thereby improving their performance:

- The input sound quality (to provide clearer speech)
- The quietness of the booth (so interpreters can concentrate), and
- A good view of the conference/meeting proceedings.

Interpreters also bring their own technological devices such as a laptop, tablet computer and/or smart phone to booth. Such personal devices are used to (1) display session materials (i.e. agenda, presentation files) plus a self-prepared glossary and (2) facilitate searches on the Internet.

2.2 Alternative conference interpreting equipment

In recent years, alternative equipment has been used in conference venues, mainly to reduce the cost of equipment. For example, the Tourguide system with one-way communication channel is sometimes used for small scale conferences/meetings. With this system, booths, interpreters' consoles and the mixing console are not required. Audiences listen to the interpretation through wireless receivers. To have good audio reception, interpreters need to sit near the loud-speakers or near the human speakers. Though it saves the cost of equipment hiring, such a working environment can greatly affect the interpreters' performance due to uncontrollable audio input.

A recent innovation was the introduction of a mobile phone application which, together with Bluetooth, is used to transmit interpretation services to individual listeners, replacing the wired equipment [7]. Audio input and output for both interpreters and audiences are controlled by the application. The application claims to ease the job of *conference equipment manager*, not that of the interpreters, however.

2.3 Multimedia learning context at conferences/meetings

Conferences and meetings often have a theme or correlated themes. Invited speakers talk around the theme with the aid of presentation files, often in one of the two formats PowerPoints or pdf. The introduction of the theme, the speakers and the speakers' topics are presented on the conference/meeting agenda. The purposes of conferences and meetings are to disseminate information and exchange ideas. The process of dissemination and interaction is actually a learning process for the participants. Therefore, interpreters work not just across different subject knowledge, topics and cultures but also in different learning contexts. Recent years have seen large advances in the provision of technological support for conferences and meetings.

Compared with 20 years ago, conference speakers no longer use transparent plastic slides but instead use computer based presentation files, large rich multimedia displays (i.e. screen panels), fancy lighting, and more reliable and clearer sound systems help to enhance the multimedia learning experience of the conference/meeting participants.

Along with the development of software and applications, it becomes much easier and faster to design and create graphical information. Presenters add audio and video clips, complex diagrams, and figures to their presentations for better demonstration and explanation and to compress complex ideas within their presentations. The multimedia display of information and the more complex content in a presentation constitute a "*multimedia cognitive load*" for interpreters [8]. The implication is that while comprehending the presenter's messages in real-time as well as delivering it in the another language, interpreters will have to make use of much or all of the limited capacity of their working memory to comprehend, process and express the message in another language. There will be very little capacity left for interpreters to follow up the presenter-designed learning process for audiences.

To facilitate comprehension of a particular presentation, interpreters study the text and diagrams on slides to form understanding of the speaker's presentation and main ideas prior to the conference/meeting. In order to accurately render the speech and maintain a good flow of delivery, good views of the presentation file and the conference proceedings are essential for interpreters in the booth at the conference/meeting.

3 The role of the glossary for simultaneous interpreters

While preparing for an interpreting task, an interpreter usually compiles a bilingual glossary, which is formatted as two parallel columns, with one column presenting language-A and the other the equivalent word or phrase in language-B. The glossary usually contains unfamiliar words, technical terms and proper names extracted from *the speakers' presentation files, conference/meeting agenda and relevant readings* during the preparation phase. Professional interpreters, including the interpreters from the Association Internationale des Interprètes de Conférence (International Association of Conference Interpreters AIIC), consider glossaries to be of paramount importance.

AIIC is a global association of conference interpreters with over 3,000 professional members from across the world. The organization was established more than 60 years ago. Their web magazine regularly publishes articles about hot issues in the interpreting world, glossaries being one of the popular topics. The association has given guidance on glossary building in their *Practical Guide for Professional Conference Interpreters* [9]. This guide suggests the process of glossary building is a learning process which helps the interpreter to understand and remember terminologies and concepts.

A recent article in AIIC [10] presented the results of "*A survey of glossary practice of conference interpreters*". The results confirmed the importance of the learning process during glossary building, describing the process as one to "*learn about issues and concepts*". In the survey, professionals agreed that most of the glossary comes from presentations, the agenda and information linked to the agenda [10]. Moreover, the survey indicated that instantaneously retrieving the glossary from (1) the interpreter's memory or (2) a glossary list, are the only ways to use the prepared terms in the process of real-time rendition and delivery. This survey, not only emphasized the significance of the glossary list, the presentations, the agenda and interpreter's memory, but also illustrated a dynamic relationship and links between them.

3.1 Technologies for extracting terms and build up glossary

The ways to search for accurate translations of terminologies and proper names have changed from using traditional dictionaries to online dictionaries, and/or massive cloud services and databases [11, 12]. Xu and Sharoff [13] reviewed methods using comparable corpora to extract terminologies from conference documents and web content. They claim when the accuracy of the generated term lists is high, the use of automatic term lists could improve the preparation efficiency of interpreters.

More applications are also available to interpreters. Costa et-al [14] reviewed the available software for interpreter's terminology management to be used prior to an interpreting task. They also described "unit conversion" applications for mobile phones which are helpful when converting between currencies and measuring units.

3.2 Are technologies assisting interpreters in the right way?

This is a serious question raised by researchers and practicing interpreters [12, 15]. Technologies can be helpful, but with conditions and constraints. Various issues raised include how much time interpreters might spend on finding the resources and trainings required to learn and adapt to the new technologies, the familiarity required to use the new technologies, and the cognitive capacity available when working for using these technologies. For example, when an interpreter works in the booth, with a laptop to read the slides, a tablet showing terminologies, and a mobile phone at hand ready for looking up new terms, the interpreter will have to shift attention and increase processing capacity when using different media to search for information.

4 Challenges to Interpreters

4.1 Cognitive challenges

Cognitive challenges are also widely acknowledged and discussed theoretically by researchers and practicing interpreters. The last two decades has seen considerable discussions concerning the cognitive challenges faced by interpreters, firstly from a linguistic perspective [16–19], and secondly from a psychological perspective [20–22].

This research has shown that modern presenting methods and rich-media contexts bring additional cognitive challenges, the extent of which are dependent on the content in the presentation files and on the nature of the technological environments.

Brook Macnamara [23] from Princeton University reviewed all the cognitive aptitudes required of an interpreter, and identified the cognitive functions required for interpreting. She used five complex diagrams to illustrate the required skills, abilities, intelligence, and memory from "operational, perspicacity, processing, and second language learning" perspectives (see Macnamara's paper for details), which in turn evidently reflects the cognitive challenges often experienced in interpreting.

4.2 Multitasking, attentional control and memory

Simultaneity of cognitive tasks (listening, processing and speaking) is known as multi-tasking, which is a foundational skill of Simultaneous Interpreting (SI). Attentional control allows interpreters to appropriately allocate attentional resources: (1) to attend to the useful stimuli to "logically reason, analyse and store information in memory", (2) to activate a functional working memory for processing information and form renditions in the target language [23]. With the additions of presentation files, the use of glossary list and other conference/meeting materials, the interpreters also need to allocate attentions to visual aids so as to assist comprehension and rendition. Technological advances in the personal devices are intended to support the interpreters with better management and easy alignment of additional visual information. However, the diversified applications and formats of the conference materials require the interpreter to allocate cognitive capacity and shift attentional resources for managing and processing different visual materials. For example, in a case when an interpreter needs to find a term in the glossary (prepared from the presentation materials), the interpreter's attention shifts to finding the term in the long list of glossary.

As suggested by Macnamara [23], in the process of simultaneous interpreting, attention is allocated to different tasks simultaneously. Familiarity of tasks reduce cognitive load. The extreme development of familiarity is automation (as cited in [23]). In the previous case of 'term searching' in the glossary, an automated search for terms in the glossary illustrates one form of automation. Later in this paper we will present a system (hypothesis) which explores both opportunities for reducing cognitive load through use of automation and a better designed Human Computer Interaction (HCI).

4.3 Challenges caused by the location of booth

We will illustrate the challenges facing interpreters by studying one of the settings of our training facilities in the University of Essex. LTB6 (Lecture Theatre for teaching) in the University of Essex was built with fixed booths. This facility is used to host mock conferences to train interpreters. The venue comprises a large lecture room with a capacity for 300 people. The booths are fixed on one side of the upper floor (see Fig. 5).

When the interpreters go into the booths to setup the workstation, they turn on a laptop which displays a glossary list together with the speakers' presentation. In this particular context, the interpreters need to constantly check the main auditorium screen to follow the presenter's speech. As the screen concerned is about 30 meters' to one side of the booth (Fig. 3 and Fig. 5), the interpreters have difficulty reading text on the screen. To have a view of the conference proceeding, the interpreters need switch their gaze from the main auditorium to their personal laptop from time to time. Another difficulty is that the interpreter is not always able to realize immediately when the presenter changes slides, especially when the display on the projector is unclear (Fig. 4). In cases where speaker's jump slides, there is a risk of negative psychological effects on interpreters who feel they have lost track of the presentation.

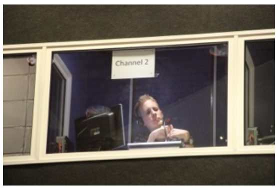

Fig 3. Interpreter looks at the projecter from booth

Fig 4. Projector's view from booth

Fig 5. Booth position in LTB6

The pre-prepared glossary list can have thirty (or more) pairs of specialized terms in two languages. When the presenter mentions a term which was included in the prepared list but which the interpreter cannot remember the exact translation of, she/he needs to refer to the glossary list. Finding the term from the glossary list means re-focusing their attention away from the speaker and the list (adding to their cognitive load), until the term is located. In a case when multiple unremembered terms appear within one sentence, the interpreter needs to find all of them from the glossary list, occupying a great amount of the interpreter's cognitive capability and risking delays in interpreting.

Thus, from this setting we argue that cognitive loading (or overloading!) of an interpreter is a major factor in determining how well an interpreter performs. In particular, for any technology to be adopted by interpreters it needs to lower, rather than increase their cognitive load. The two most important aspects of cognitive loading for interpreters is 1) their working memory, and 2) their speed of reasoning. The first of these can be supported by creating computer supported glossaries of terms, with fast search methods to access them (essential extending working memory) and the second of these can be improved by good human-computer interaction design making information and control simple and intuitive (essentially simplifying any reasoning activities). By way of a theoretical basis, for the first we are building on the concept of working memory, for the second we build on the notion of elementary mental discriminations, or the Stroud number. Exploring how technology, and in particular smart-glasses, could positively augment an interpreter's capability is the aim of our research. Our approach to this is described in the following section

5 Interpreting in booth with augmented reality glasses

Fig 6. Chantel (interpreter) in a booth

Fig 7. Chantel (interpreter) wearing meta-1 glasses

As was explained in the previous section, we have set out to explore how smart glasses may be used to reduce the cognitive load on interpreters, in order to improve their performance. Thus, a project was initiated in the University of Essex to undertake research on potential solutions to the challenges described in the previous section for 21st century interpreters using augmented reality smart-glasses. At this stage we are hypothesising that smart-glasses can overcome the problems we have described, so our mission is to characterize the challenge (one of the purposes of this paper), create some theoretical models for the pedagogy and computer architecture (another aim of this paper) and then finally test the hypothesis by experimenting with a real system (an aim of a future paper). Our hypothesis is not simply a binary question (does it hold or not) but rather an exploration of the variables at work especially regarding HCI parameters such as size, position, colour and mode of control of the interpreting session data. Thus our experimental architecture seeks to accommodate as much customisation as possible, allowing the interpreters to change as much of the appearance and operation of the system as is practical. Explaining this in another way, we are arguing that by placing a pre-prepared glossary, together with other session information in the interpreter's field of view (Fig. 8) using augmented reality glasses (with appropriately designed Human Computer Interaction), interpreters will be able to reduce their cognitive effort and concentrate more on rendering information and messages from different sources.

At this stage we are prototyping the system, starting with an electronic mock-up of the user interface which is shown in the diagram below:

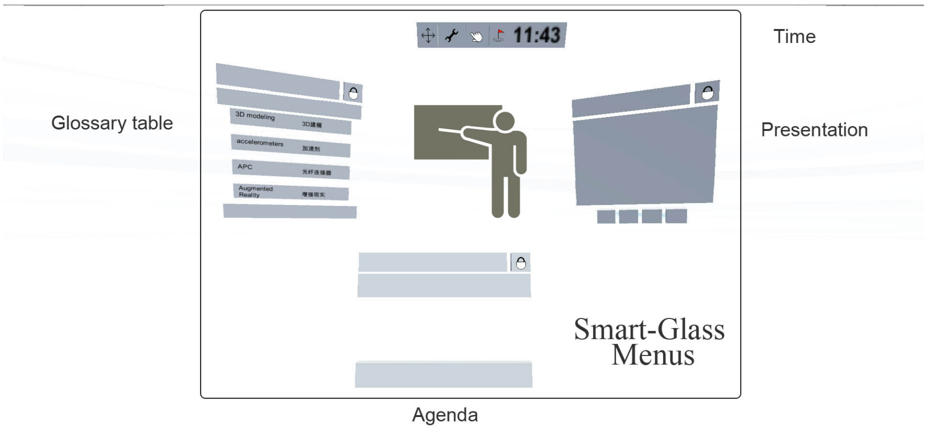

Fig 8. AR-Language Interpreting smart-glasses screen

We envisage the smart-glasses will be worn by the interpreters during live sessions allowing them to simultaneously view the real event and virtual screens containing supplementary materials positioned to one side of their field of view. The virtual screens are relatively large (a metre or so at a distance of a few meters) and contain information such as the glossary of terms, the agenda, the presenters' slides, the time and an auxiliary window that could, for example, be used by the supporting (second) interpreter who could provide additional and unplanned information. We also envisage that the second interpreter would wear a set of smart-glasses which they could use to manipulate information at key moments; to assist the main interpreter (eg undertake an online search for unknown vocabulary arising from a Q&A with the audience). This is very much an experimental system, and so one of its purposes is to allow the interpreter to customize the environment as much as possible so new research data can be gathered from how the system is personalized or used in live interpreting sessions. Thus there are many hidden functionalities concerned with personalizing the environment.

This framework forms a model for interpreting that we call SmARTI (Smart Augmented Reality Technology for Interpreters). The Meta glasses we are using were designed for individuals to wear, but have proved to be little heavy for prolonged use. Thus, one of the ideal specs for of smart-glasses for interpreters would be lightness; other features being no wires (not tethered), fashionable appearance, excellent sound, long battery life (at least a half day) etc. The current state-of-the-art in wearable AR glasses has some way to go before they would meet an ideal specification for interpreters since they are tethered, a little on the heavy side for prolonged use, and the geeky appearance might not be appealing to all interpreters! To popularize the use of this technology, interpreters will require further hardware improvements which this work will also aim to throw light on.

6 Summary & Reflections

This paper introduced the booth environment for simultaneous interpreters. It argued that insufficient assistance is given to the interpreters in booths to reduce the cognitive load caused by the increasing use of technology and the ever-increasing complexity of contexts at conferences and meetings. In particular, we identified that extending working memory and easing reasoning tasks were key areas where technology might be used to improve an interpreter's performance. We also proposed that wearable smart-glasses might provide a useful simultaneous interpreting environment and, have described some preliminary studies we are undertaking using Meta-1 augmented-reality glasses. This is a work-in-progress project and at this stage we have framed the problem space through a literature review, identified the research issues to be explored, proposed a solution (with hypothesis), created an operational model (SmARTI – Smart Augmented Reality Technology for Interpreters) and built a simple prototype all of which we have reported on in this paper. Our longer-term aim is that we hope to be able to create what is, in effect, a virtual (and wearable) interpreting booth that is designed in such a way as to reduce the cognitive load on interpreters, thereby improving their mobility and performance. Our aim is to refine this design through ongoing work, further exploring the issues and reporting on those at later conferences.

Acknowledgements

The authors wish to thank Thomas Hayward and Ahmed Alzahrani for providing the hardware and software assistance in designing the interpreter's interface. We also acknowledge the University of Essex's UROP fund which enabled us to implement this prototype. Finally, we wish to thank the Department of Language and Linguistics University of Essex for supporting the project and interpreter lab facilities described in this paper.

References

[1] Whispering and the Origin of Simultaneous Interpreting (cont.). http://unprofessionaltranslation.blogspot.co.uk/2012/12/whispering-and-invention-of.html. Accessed 12 Apr 2016
[2] United Nations DGACM. http://www.un.org/depts/DGACM/interpretation.shtml. Accessed 12 Apr 2016
[3] (2014) CBI Product Factsheet : MICE from the EU.
[4] Vijay S (2015) What's trending in the MICE Industry? http://www.thesmartcube.com/insights/strategy/item/what-s-trending-in-the-mice-industry-2. Accessed 14 Apr 2016
[5] ISO (1998) ISO 2603:1998(en), Booths for simultaneous interpretation — General characteristics and equipment. https://www.iso.org/obp/ui/#iso:std:iso:2603:ed-3:v1:en. Accessed 18 Apr 2016
[6] European commission (2016) Main Technical Specifications for Conference Rooms.
[7] Interactio (2016) New Technologies and Conference Interpretation.

http://webcache.googleusercontent.com/search?q=cache:8rdUrhH5dPwJ:www.port.ac.uk/media/contacts-and-departments/slas/events/tr08-farwick.pdf+&cd=11&hl=en&ct=clnk&gl=uk. Accessed 9 Apr 2016
[8] Sorden SD (2012) The cognitive theory of multimedia learning. In: Brown G, Lara-Alecio R (eds) Handb. Educ. Theor. Information Age, pp 1–31
[9] AIIC (2004) Practical Guide for Professional Conference Interpreters. http://aiic.net/page/628#22. Accessed 18 Apr 2016
[10] Jiang H (2015) A survey of glossary practice of conference interpreters. http://aiic.net/page/7151/a-survey-of-glossary-practice-of-conference-interpreters/lang/1. Accessed 18 Apr 2016
[11] Kelly N (2009) Moving toward machine interpretation. Transl. localization
[12] Farwick JS (2010) New Technologies and Conference Interpretation Judith S . Farwick. In: Lee-Jahnke, Hannelore Prunc E (ed) Am Schnittpkt. Von Philol. Und Transl. Festschrift Zu Ehren Von Martin Forstner. Peter Lang, pp 70–79
[13] Xu R, Sharoff S (2014) Evaluating Term Extraction Methods for Interpreters. In: 4th Int. Work. Comput. Terminol. pp 86–93
[14] Costa, Hernani; Pastor Gloria Corpas; Munes ID (2014) Technology-Assisted Interpreting. Multilingual 25:27–32.
[15] Pym A (2011) What technology does to translating. Int J Transl Interpret 3:1–9.
[16] Roy CB (1993) A sociolinguistic analysis of the interpreter's role in simultaneous talk in interpreted interaction. Multilingua 12:341–364. doi: 10.1515/mult.1993.12.4.341
[17] Darò V, Fabbro F (1994) Verbal memory during simultaneous interpretation: Effects of phonological interference. Appl Linguist 15:365–381. doi: 10.1093/applin/15.4.365
[18] Gile D (2008) 2008 Local Analysis Forum. In: Forum 6. pp 2.59–77.
[19] Hanaoka O (2002) JAIS Proper Names in the Media : Problems for Translators / Interpreters and L2 Learners. Interpret Stud 2:28–42.
[20] Gerver D (1975) A Psychological Approach to Simultaneous Interpretation. Meta J des traducteurs 20:119–128. doi: 10.7202/002885ar
[21] Moser-Mercer B, Künzli A, Korac M (1998) Prolonged turns in interpreting: Effects on quality, physiological and psychological stress (Pilot study). Interpreting 3:47–64. doi: 10.1075/intp.3.1.03mos
[22] Kurz I (2003) Physiological Stress During Simultaneous Interpreting: a comparison of experts and novices. Interpret Newsl 12:51–67
[23] Macnamara BN (2012) Interpreter Cognitive Aptitudes. J Interpret 19:9–31.