

**Coventry
University**

Coventry University Repository for the Virtual Environment (CURVE)

Author name: Jarvis, S. and De Freitas, S.

Title: Evaluation of an immersive learning programme to support triage training

Article & version: Post-print

Original citation:

Jarvis, S. and De Freitas, S. (2009) 'Evaluation of an Immersive Learning Programme to Support Triage Training' Proc. of the 1st IEEE International Conference in Games and Virtual Worlds for Serious Applications. held: 23-24 March 2009. IEEE Computer Society: 117-122.

<http://dx.doi.org/10.1109/VS-GAMES.2009.31>

Statement required by publisher:

© 2009 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

This document is the author's final manuscript version of the journal article, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.

Available in the CURVE Research Collection: October 2011

<http://curve.coventry.ac.uk/open>

Evaluation of an Immersive Learning Programme to Support Triage Training

In-game Feedback and its effect on Learning Transfer

Steve Jarvis

SELEX Systems Integration Limited
Welwyn Garden City
steve.jarvis@selex-si-uk.com

Sara de Freitas

The Serious Games Institute
Coventry University
s.defreitas@coventry.ac.uk

Abstract—This paper describes the evaluation of a computer game to support triage training. Triage is a process for decision-making that prioritises mass casualties in terms of treatment. The main aim of the research was to test the hypothesis that participants using the game for practice would perform better in terms of the accuracy in applying the triage protocol than a group that practised triage with a table-top exercise. The method of giving in-game formative feedback to the learner was modified within the trial programme and that provided the opportunity to investigate whether changing the complexity and timing of feedback affected learning transfer through performance measured in a realistic assessment activity. The results showed that the participants who practised using the game were significantly more accurate for certain measures of performance in applying the triage protocol. The participants that received a modified in-game feedback that reduced complexity and the delay in giving feedback were also significantly more accurate for certain measures of performance in triaging the casualties. These findings will require further experimentation to determine which attributes of the in-game feedback have the greatest impact on the learning of the triage protocol for the given learner group.

Keywords- *evaluation; serious games; feedback; learning transfer*

I. INTRODUCTION

The potential of serious games to help increase the effectiveness of training and learning has been a subject of recent debate in the field of learning [1] [2]. While the literature-based evidence for research into the field of game-based learning is increasing, many of the studies being conducted have not produced conclusive results. This has led to a lack of high quality empirical evidence [1]. The research described in this paper, aims to address this lack of evidence and aims to help to establish when it may be appropriate to use game-based technology for training and learning.

Triage is a process for decision-making that prioritises mass casualties in terms of treatment. An initial triage protocol quickly sorts casualties into priorities based on airway, breathing and circulation parameters [3]. There is more than one triage protocol, but the one done at the scene is called sieve. One of the problems with current triage training is the high risk of skill fade that arises from insufficient opportunities to apply the learning in real situations after the training, as emergency incidents involving mass casualties are relatively uncommon and can be expensive to recreate in real training exercises [4].

Triage Trainer is a serious game¹ designed to improve the accuracy in applying the triage protocols defined by the Advanced Life Support Group (ALSG) and to provide increased opportunities for practice to address the issue of skill fade. The Advanced Life Support Group is a UK-based charity organisation with the mission to preserve life by providing training and education in life saving techniques. They deliver training in triage as part of a three-day training course covering the medical management of major incidents [3]. The prototype of Triage Trainer covers the triage sieve protocol.

The Triage Trainer prototype was developed by the Serious Games – Engaging Training Solutions consortia, which consists of TruSim, a division of Blitz Games Studios Limited, a major UK developer of entertainment games; SELEX Systems Integration Limited (formally VEGA), a Finmeccanica company, the Universities of Birmingham, Coventry, London and Sheffield. The project [5] was co-funded by the Technology Strategy Board's Collaborative Research and Development Programme, following an open competition. The Technology Strategy Board is a business-led executive non-departmental public body, established by the UK government. Its mission is to promote and support research into, and development and exploitation of, technology and innovation for the benefit of UK business, in order to increase economic growth and improve the quality of life. It is sponsored by the UK Department for Innovation, Universities and Skills (DIUS). Subject matter expertise was supplied by ALSG.

The scenario presented to the learner in the Triage Trainer is the aftermath of a bomb explosion in a city centre resulting in a number of casualties located around the scene. The player is the first medically qualified person at the scene and is required to follow the ALSG triage sieve protocol to tag each casualty with the appropriate priority. The player clicks on an icon beside each casualty to enter an Examination Mode to then apply the checks required to implement the triage sieve protocol. Fig. 1 shows Triage Trainer in Examination Mode with the player performing the Respiratory Rate check. For this check the learner is presented with an animated image to simulate breathing rate and an analogue clock to time the rate.

When the learner has applied the triage sieve protocol to all the casualties in the given scene, an After Action Review (AAR) is displayed that provides formative

¹ One simple definition for a serious game is a computer game developed for a serious or non-leisure purpose such as training

feedback to the learner on their performance in applying the triage sieve protocol to each casualty.

The primary research objective was to compare the performance of participants practicing the triage sieve protocol using Triage Trainer with participants practicing with the current triage sieve practical exercise to determine any difference in the learning transfer between these two groups. The current triage sieve practice involves a table-top exercise to sort casualties in terms of priority for treatment based on the given airway, breathing and circulation parameters written on cards.

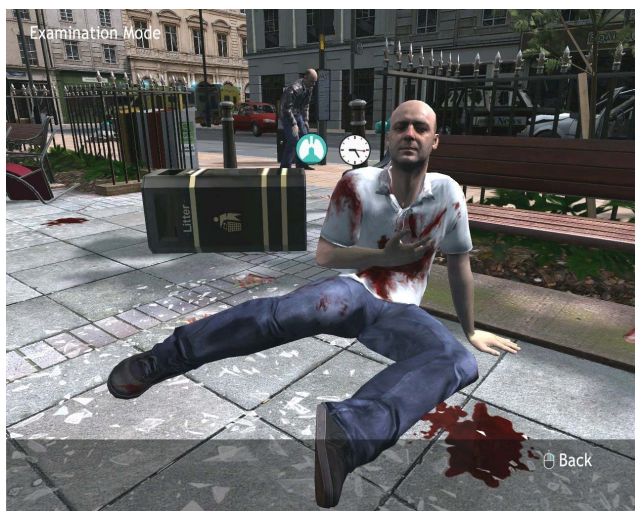


Figure 1. Triage Trainer in Examination Mode (Respiratory Rate Check)

The comparison of different training media has been subject to criticism, because of the many different variables that impact learning (e.g. characteristics of the learner and the method of providing formative feedback) and the difficulty in conducting experiments that enable the comparison of only one of these variables [6]. Studies have shown that it is not the training media (delivery method) that impacts learning, but rather the training method (instructional strategy) [6] [1], although some research describes a more complex relationship between training media and training method and their impact on learning [7].

The decision was taken to follow a deductive approach for the experiment to compare Triage Trainer with the table-top triage exercise. This approach involved testing a priori hypothesis of difference between the two training media based on analyses of their different attributes [8]. Our hypothesis was that Triage Trainer would produce improved performance in the application of the triage sieve protocol compared to the table-top exercise, as measured in terms of tagging accuracy and step accuracy due to an increased level of engagement and more effective formative feedback. Tagging relates to the placing of a particular tag on the casualty that indicates the initial priority for treatment as defined by the triage sieve protocol. Step accuracy refers to whether the correct steps, according to ALSG's protocols, were taken without omissions or unnecessary steps that would waste vital time.

The method of giving in-game formative feedback to the learner through the AAR was modified within the trial

programme and that provided the opportunity to investigate whether changing the complexity and timing of feedback affected learning transfer through performance measured in a realistic assessment activity. The AAR that was used for the first two trial events² required the learner to navigate several screens. A summary AAR screen appeared first that presented information on the percentage of correctly and incorrectly tagged casualties for each of the four priorities. The learner could then select one of the four priorities to obtain more information on the nature of any errors for particular casualties, such as that shown in Fig. 2. The supposition was that this would allow the learner to focus where their performance was weak, and so aid learning. Fig. 2 shows a time-line to communicate the steps taken for triaging each casualty and the time taken for each step. There were 10 casualties in each of the three scenes. The AAR was changed following the analysis of feedback from the participants. The modified AAR is shown in Fig. 3 and was used for the third and fourth trial events. The left side of the modified AAR provides the learner with information on whether they tagged the casualties correctly and whether they made any step errors. The right side is where the learner obtains more information about the nature of any step errors. The learner was presented with five scenes with 3, 4, 5, 8 and 10 casualties in each scene. This would mean that a participant would receive their first in-game feedback after triaging just 3 casualties, compared to the 10 casualties in the initial AAR.

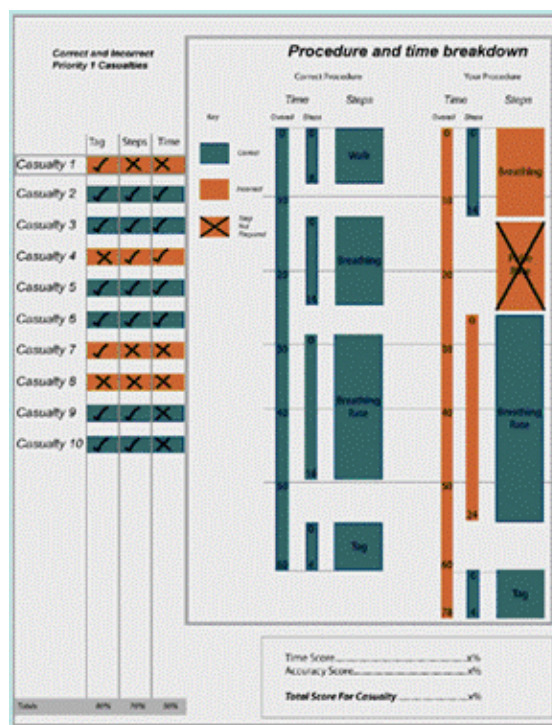


Figure 2. Part of the Initial AAR

Recent comprehensive literature reviews of feedback research have produced inconclusive evidence as to the feedback variables that have the greatest impact on

² The trial programme comprised four trial events. Each trial event was delivered within one of ALSG's Major Incident Medical Management and Support training courses

learning [9] [10]. In terms of feedback timing, there is some evidence that delayed feedback may produce improved learning transfer, especially in relation to concept-formation tasks, whereas immediate feedback may be more efficient for procedural learning [11]. Various studies on the timing of feedback have concluded that the effectiveness of feedback depends not on the main effect of timing but on the nature of the task and the capability of the learner [12]. What is needed is further exploration on the most effective ways to match feedback (e.g. timing, complexity) to learning tasks and individual learner characteristics. The research that will now be described is aimed at informing this area with game-based technology as the training media.

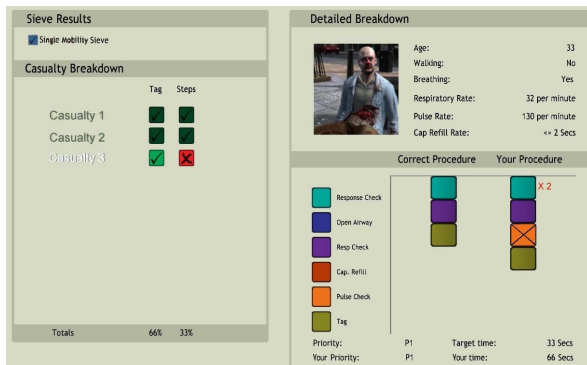


Figure 3. Modified AAR

II. METHOD

The most widely accepted model for evaluating training is the Kirkpatrick model [13]. The aim of the study was to evaluate Triage Trainer on the reaction and learning levels of the Kirkpatrick model (levels 1 and 2). This would involve gathering information on the participants' reaction to using Triage Trainer or the table-top exercise and the effect on subsequent performance, as measured by an assessment activity. It was impracticable to evaluate at Kirkpatrick's levels 3 or 4, which look at job behaviour and benefits to the organisation, which would have been very difficult to measure as they would have required the participants to attend real emergency incidents.

The evaluation of Triage Trainer was undertaken during Major Incident Medical Management and Support (MIMMS) training courses at Manchester, Plymouth, Croydon, and Newcastle. These are the four trial events mentioned earlier. Two pilot trial events were conducted prior to these four, to test the assessment method and obtain more user feedback on the game. For each course all participants were randomly assigned into one of two groups. One group practised triage sieve with Triage Trainer and the other group performed the table-top triage sieve practical exercise. The latter group will be referred to as the control group. The trials used an independent samples design [14]. Typically, each ALSG course has 24 delegates, implying 12 participants in each experimental group. The delegates included paramedics, doctors and other healthcare workers requiring knowledge of the medical management of major incidents. Both groups completed a questionnaire at the start of the course that asked about their attitude to computers; their use of

computers; their gaming experience and also their learning preference. Delegates received a manual before the course and a lecture on day one of the course to introduce the triage process.

During the morning of day two a triage practical exercise required delegates to apply the ALSG triage sieve protocol to casualties with different physiological parameters written on cards, to sort into the correct priority for treatment. The control group practised triage using these cards. This exercise lasts for one hour. For our trial, the 12 participants in the game group first received a tutorial to learn how to use Triage Trainer and then were given the remaining time to practise triage by applying the ALSG triage sieve protocol to each casualty in each scene. The total number of casualties in all scenes was 30. For the first two trial events at Manchester and Plymouth, a more complex AAR was used as previously mentioned. Three scenes had been created with 10 casualties in each scene. The participant had to triage all 10 casualties before receiving feedback on their performance. For the last two trials at Croydon and Newcastle, a simplified After Action Review was used, and the player was presented with five scenes with 3, 4, 5, 8 and 10 casualties in each world (retaining the same total of 30 casualties as the earlier AAR). This would mean that a participant would receive their first in-game feedback after triaging just 3 casualties, compared to the 10 casualties in the earlier trials.

In the afternoon of day two, all participants were assessed for their mastery of the triage sieve protocol by tagging 8 casualties played by actors. This assessment was recorded on video for later analysis.

III. RESULTS

Table I shows the results for tagging and step accuracy for the 8 casualties, combining the assessment data from all four trial events. There were 44 participants in the control group and 47 in the game group.

TABLE I. PERCENTAGE OF PARTICIPANTS CORRECTLY TAGGING AND FOLLOWING CORRECT STEP PROTOCOL WHEN TRIAGING THE EIGHT CASUALTIES

No. of casualties correctly triaged	Tagging Accuracy		Step Accuracy	
	Table Top	Game	Table Top	Game
0/8	0.0	0.0	4.5	4.3
1/8	0.0	0.0	4.5	6.4
2/8	2.3	0.0	11.4	0.0
3/8	0.0	0.0	6.8	4.3
4/8	0.0	4.3	18.2	14.9
5/8	6.8	4.3	18.2	19.1
6/8	11.4	0.0	13.6	12.8
7/8	25.0	19.1	15.9	10.6
8/8	54.5	72.3	6.8	27.7

A Chi squared analysis using the Likelihood ratio (SPSS, 15.0) indicated that the game group performed significantly better than the control group for tagging accuracy [Chi = 13.136, df = 5, p<0.05].

The number of participants that achieved an 8/8 for step accuracy can be compared with those who attained a step accuracy of less than eight correctly triaged casualties out of eight. Using the Likelihood ratio, Chi squared analysis shows that the game group has significantly more participants scoring the maximum than the control group [Chi = 7.29, df = 1, p<0.01].

Table II shows the affect of the modified AAR on the percentage of participants in the game group correctly triaging the 8 casualties with the correct tag and following the correct protocol steps. The total number of participants in the game group for the first two trials with the initial AAR was 24, with a total of 23 participants in the game group for the last two trials with the modified AAR. The modified AAR had no affect on tagging accuracy (p>0.05). However, for step accuracy, Chi squared analysis using the Likelihood ratio (SPSS, 15.0) indicated that participants performed significantly better with the modified AAR than the initial AAR [Chi = 16.44, df = 7, p<0.05], with more participants triaging 8/8, 7/8 and 6/8 correctly after playing the game with the modified AAR.

TABLE II. AFFECT OF MODIFIED AAR ON THE NUMBER OF PARTICIPANTS (%) CORRECTLY TRIAGING THE 8 CASUALTIES WITH THE CORRECT TAG AND FOLLOWING THE CORRECT PROTOCOL STEPS

No. of casualties correctly triaged	Tagging Accuracy		Step Accuracy	
	Initial AAR	Modified AAR	Initial AAR	Modified AAR
0/8	0.0	0.0	8.3	0.0
1/8	0.0	0.0	0.0	13.0
2/8	0.0	0.0	0.0	0.0
3/8	0.0	0.0	8.3	4.3
4/8	4.2	4.3	25.0	14.9
5/8	4.2	4.3	20.8	17.4
6/8	0.0	0.0	8.3	17.4
7/8	20.8	17.4	4.2	17.4
8/8	70.8	73.9	25.0	30.4

A second paper is being submitted for publication that focuses on the characteristics of the participants in each group and presents the results of an analysis of the completed participant questionnaire and correlations with assessment performance. This analysis indicated that there were no significant correlations between assessment performance of tag and step accuracy with any of the self-ratings of computer skill.

IV. DISCUSSION

The analysis of the trials has shown that the game group performed significantly better on tagging accuracy than the control group. There was also significantly better performance by the game group in achieving an 8/8 step accuracy compared to less than 8/8. How should we interpret these findings? Clearly the number of participants in the sample is small, and other variables could be influencing the results, such as a different quality of instruction at the different course venues. We would suggest that one possible explanation is that the context for

learning is closer to the real world for the triage practice with the game, as compared to the current training using the casualty data written on cards. This would likely benefit the game playing participants in the assessment.

Although, clearly further research would be needed to test the hypothesis that increased immersion contributed to the improved performance for this computer game supporting triage training, one explanation for the improved performance of the game over the table-top method may relate to levels of immersion. In a number of studies, game-based learning has exhibited improvements over traditional learning in areas such as motivation and engagement. The main reasons for this have been outlined as greater learner control, improved engagement and higher levels of motivation [15]. In general game-based learning can adopt many of the characteristics of real life learning, which is the best method of training due to the minimal requirement of learning transfer. Edgar Dale's Cone of Learning (Fig. 4) indicates how real life learning is the most effective form of learning, followed by simulated experiences and a dramatic presentation [16]. This is due to the multimodal experience whereby different senses are utilised.

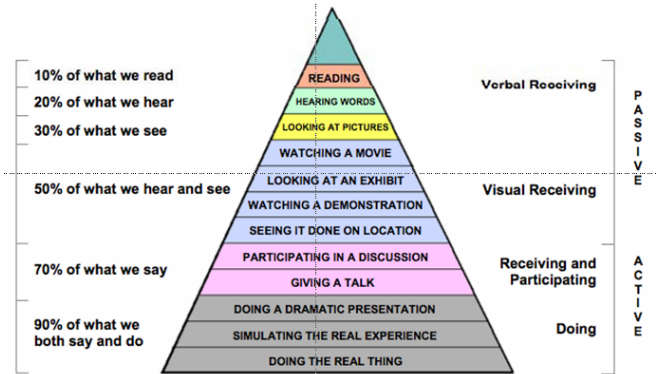


Figure 4. Cone of Learning

Learning on the job can also be more effective for retention of students on courses. Chi et al. indicate this in their work (Fig. 5) [17]. To a great extent this is because of how we learn and how we recall what we learn, and this in turn is based upon the crucial aspect of how we interface with the environment within which we interact. Fig. 6 demonstrates that relationship [18]. Sensory interactions with the environment are processed into our working memory which is then stored as long term memory. The link between physiological and cognitive functioning is a primary dimension of what we are trying to understand in the use of immersive learning experiences such as that evaluated in the Triage Trainer. If we can incur the same sets of sensory inputs as we obtain in the real environment then we are more likely to retain the training aims and to be able to accurately reproduce it in the event of a real life event. Learning fade is also less likely.

Therefore in future research we would like to identify separate measures for testing learner control, motivation and engagement and test this hypothesis by conducting studies two months after training, four months and six months to test the difference between traditional methods of learning against serious game learning over longer durations. We hypothesise that the learning recall will be

better for the game-based learners over the traditional learners, due to the more multimodal learning experience.

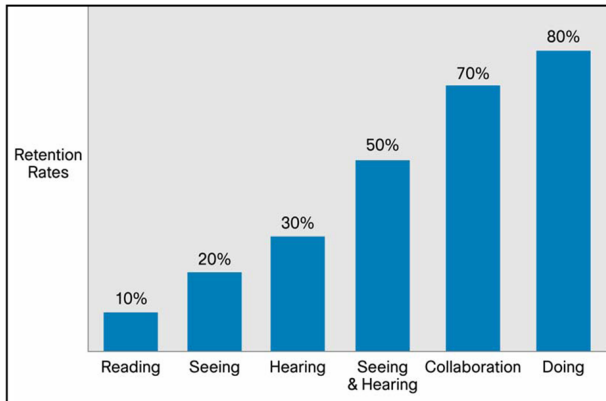


Figure 5. Learning Methods and Retention Rates

In other studies we would like to test a range of different evaluative methods to determine whether other performance indicators can skew accurate evaluation methodologies.

While clearly further research is needed to determine which factors related to in-game feedback have the greatest influence on triage performance. It is suggested by the study that giving earlier and more frequent feedback should accelerate learning (particularly if time for practice is limited). Displaying the most relevant information for performance and having a layout that enables the learner to identify errors quickly should also support more efficient learning. This research seems to suggest that improved learning transfer may be possible with a simple procedural learning task (such as the triage sieve protocol), when immediate feedback is given at the start of practice, representing the situation of a novice who may lack confidence and need the reassurance given by immediate feedback. An increasing delay in feedback may be preferential for later practice when the learner is close to mastering the procedure. This hypothesis will require further experimentation to determine more precisely which attributes of the in-game feedback have the greatest impact on the learning of the triage protocol for the given learner group, and whether these findings can be extended to a wider set of learning scenarios.

The user feedback from the first pilot trial identified issues with the game that needed addressing, such as the complexity of the AAR. We had involved representatives of the user community in the design of the game, but the feedback from the pilot trial exposed the need for even greater user involvement earlier in the development process.

As discussed earlier, one of the challenges of conducting media comparison studies is being able to control all variables, beyond those included in the experiment, which could impact the performance between the two groups. As an example, we allowed an additional 15 minutes to learn how to use the Triage Trainer game, but this took up to 25 minutes to complete, without permitting additional practice time. Also, the game generated more discussion regarding undertaking the checks that was not the case with the training media and method for the control group. The consequence was that

the game group actually had less time to practise than the control group.

V. CONCLUSIONS

The evaluation of the Triage Trainer serious game has produced evidence for significantly greater learning transfer for this game over the current training media and method (table top exercise). The authors suggest that this may be the result of an increased level of engagement, but that the design of the in-game formative feedback also has an important impact on the learning.

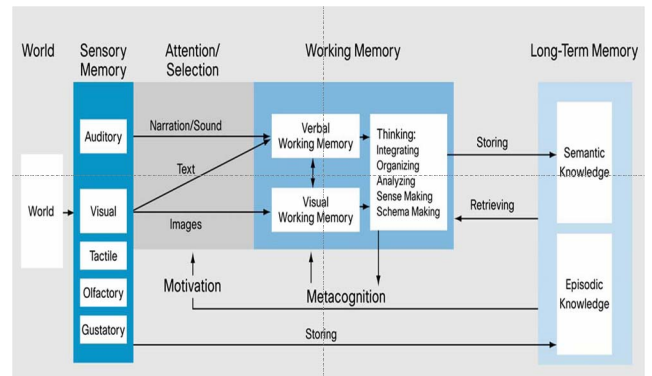


Figure 6. Thinking: Physiological and Cognitive Functions

The evaluation of the Triage Trainer showed the importance of providing effective in-game feedback. The results are consistent with the evidence from earlier studies that suggest that immediate feedback may be more efficient for procedural learning [11].

The study has presented some empirical evidence for the efficacy of game-based learning over traditional methods of learning. The study opens up new avenues for future studies, and presents an evaluation method against which other methods may be benchmarked. The study has demonstrated powerful evidence that game-based learning is effective over traditional methods where the approach allows for high levels of fidelity and an immersive training experience.

This study together with other evidence emerging from leading edge research being undertaken internationally, is leading to the conclusion that there is a significant difference between text-based learning methods and immersive 3D experiences in terms of efficacy of learning, particularly in the areas of learner control, motivation and engagement. The work is the first of many studies to indicate this conclusion (new studies from the PULSE!! Project [19] for example are being evaluated). Larger studies are needed to support these findings; however the power of multimodal learning is being demonstrated here.

We believe that game-based learning has real capabilities for supporting training, particularly in areas of training where learning fade could be critical and involve the loss of lives, such as in emergency response training. In the future, the research objectives will involve the development of metrics and purpose developed tools for evaluating efficacy against learning outcomes. In addition, there is clearly the potential of the virtual learning environment for triage training to provide an individualised experience by adapting what is presented to the learner based on identified individual differences in

behaviour and performance, through the use of AI and other techniques/technologies. This may provide an important area for future research, because of the ability to assess in real-time the learner behaviour and performance in the game and modify what is presented to the learner in the game based on an analysis of this information.

ACKNOWLEDGMENT

The authors thank the following people and organisations for their significant contributions to this research: Richard Smithies of TruSim and all the TruSim team; Bryan Tregunna of SELEX Systems Integration Ltd; Dr James Knight of the University of Birmingham; the Advanced Life Support Group and in particular Dr Simon Carley and Professor Kevin Mackway-Jones; Daniela Romano of the University of Sheffield and Les Moseley formally of the Coventry University Centre for Disaster Management.

REFERENCES

- [1] R. T. Hays, "The effectiveness of instructional games: A literature review and discussion", NAVAIR Technical Report 2005-004, Naval Air Warfare Center Training Systems Division, 12350 Research Parkway, Orlando Florida 32826-3275
- [2] S. de Freitas, "Using games and simulations for supporting learning", *Learning, Media and Technology*, December 2006, Vol. 31, Issue, pp. 343-358
- [3] T. J. Hodgetts and K. Mackway-Jones, Eds. *Major Incident Medical Management and Support*, Advanced Life Support Group, 2nd ed., BMJ Books, 2002
- [4] Y-F. Chen, G. Rebolledo-Mendez, F. Liarokapis, S. de Freitas and E. Parker, "The use of virtual world platforms for supporting an emergency response training exercise", 2008, Proceedings of the CGAMES conference, Wolverhampton, UK
- [5] S. de Freitas and S. Jarvis, *Serious Games – Engaging Training Solutions: A research and development project for supporting training needs*. *British Journal of Educational Technology*, 2007, Vol. 38, No. 3, pp. 523-525
- [6] R. E. Clark, "Reconsidering research on Learning from media", *Review of Educational Research*, Winter, 1983, Vol. 53, No. 4, pp.445-459
- [7] R. Kozma, "Learning with media", *Review of Educational Research*, Summer 1991, Vol. 61, No. 2, pp.179-211
- [8] S. M. Ross and G. R. Morrison, "Experimental research methods", in *Handbook of Research on Educational Communications and Technology*, 2nd ed., D. H. Jonassen, Eds. Lawrence Erlbaum Associates, 2004, pp.1021-1043
- [9] V. J. Shute, "Focus on formative feedback", *Review of Educational Research*, March 2008, Vol. 78, No. 1, pp. 153–189
- [10] E. H. Mory, "Feedback research revisited", in *Handbook of Research on Educational Communications and Technology*, 2nd ed., D. H. Jonassen, Eds. Lawrence Erlbaum Associates, 2004, p.745-783
- [11] M. L. Schroth, "The effects of delay of feedback on a delayed concept formation transfer task", *Contemporary Educational Psychology*, 1992, Vol. 17, No. 1, pp. 78–82
- [12] S. A. Mathan and K. R. Koedinger, "An empirical assessment of comprehension fostering features in an intelligent tutoring system", in *Intelligent Tutoring Systems, 6th International Conference, ITS, S. A. Cerri, G. Gouarderes and F. Paraguacu (Eds.)*, 2002, New York: Springer-Verlag, Vol. 2363, pp. 330–343
- [13] D. L. Kirkpatrick, *Evaluating Training Programs: The Four Levels*, San Francisco: Berrett-Koehler, 1994
- [14] D. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, 3rd ed., 2004, CRC Press
- [15] R. Garris, R. Ahler and J. E. Driskell, "Games, motivation, and learning: a research and practical model", *Simulation and Gaming*, 2002, Vo. 33, No. 4, pp. 441-467
- [16] D. Edgar, *Audiovisual Methods in Teaching*, 3rd ed., 1969, Holt, Rinehart and Winston
- [17] M. T. H. Chi, M. Bassok, M. W. Lewis, P. Reimann and R. Glaser, "Self-explanations: how students study and use examples in learning to solve problems", *Cognitive Science*, April-June 1989, Vol. 13, No. 2, pp. 145-182
- [18] R. Moreno and R. Mayer, "Interactive multimodal learning environments", *Educational Psychology Review*, September 2007, Vo. 19, No. 3, pp. 309-326, Springer Netherlands
- [19] Texas A&M University Corpus Christi. Pulse!! – The Virtual Clinical Learning Lab [online]. 2008 [accessed 26.11.2008]. <http://www.sp.tamucc.edu/pulse/home.asp>