

Augmented Reality Training of Military Tasks: Reactions from Subject Matter Experts

Roberto Champney¹(✉), Stephanie J. Lackey²(✉), Kay Stanney¹,
and Stephanie Quinn¹

¹ Design Interactive, Inc., Orlando, FL, USA
{roberto, kay, stephanie. quinn}@designinteractive.net

² Institute for Simulation and Training, Orlando, FL, USA
slackey@ist.ucf.edu

Abstract. The purpose of this research effort was to understand the training utility of augmented reality and simulation-based training capabilities in an outdoor field environment. Specifically, this research focused on evaluating the training efficacy of the Augmented Immersive Team Training (AITT) system, a portable augmented reality training solution that targets Forward Observer (FO) tasks associated with a Call for Fire (CFF) mission. The assessment focused on evaluating training utility, satisfaction, usability, simulator sickness, presence, immersion and appropriateness of the fidelity cues provided by the AITT system. Data were gathered via questionnaires. The results of this study provided insight for formative evolution of the AITT system design and may have implications to other similar technologies.

Keywords: Augmented reality · Training · Learning · Immersive virtual reality · Wearable technology · Mixed reality · Training systems

1 Introduction

Augmented Reality (AR), an emerging interactive technology that merges elements of the physical real-world environment with virtual computer generated imagery, promises to be a viable pedagogical venue for the training of a variety of domains. Kamarainen et al. (2013) demonstrated that the “authentic” participatory experience provided by AR training systems can increase the effectiveness of instruction, while improving trainee attention, engagement, and motivation. As the technology matures, AR is thus poised to quickly expand from a focus on information presentation and entertainment to applications for learning and exploring (Johnson, Smith, Willis, Levine, & Haywood, 2011). It is not uncommon, however, for new and innovative training technology such as AR to simply be incorporated into the latest educational applications without regard to how best the technology supports learning. Before AR receives widespread adoption into educational applications, it is thus necessary to understand what makes AR a promising technology for training. Bitter and Corral (2014) suggest the training effectiveness of AR will be highest in situations where the technology is aligned with solid educational theories. In this regard, Dunleavy and Dede (2014) have suggested that AR is particularly well suited to situated and

constructivist learning theories that involve authentic inquiry and active observation, which is supported by probes and scaffolding. This is important to understand because some past studies have failed to demonstrate the training value of AR Zhu et al., (2014). The current study examined the potential of AR to support learning of complex skills in an outdoor context.

2 Background

This study consisted of a formative evaluation of the Augmented Immersive Team Trainer (AITT) system. AITT is an AR training capability designed to support Forward Observers (FO) in conducting CFF tasks. AITT's innovation, as compared to other immersive virtual reality or PC based trainers, lies in its portable-outdoor capability. In contrast to most simulation based training systems that rely on an indoor lighting environment, the AITT allows training on-site – whether indoor or outside – via a wearable AR system. This allows the training to take place at the operational site or at relevant environments (e.g., live fire range).

Call for Fire. US Marines and other Services' ground Warfighters often employ artillery or mortars in support of their missions. To employ these weapons Warfighters need to coordinate with other teams to ensure the successful and safe use of these tools. A CFF is a message that contains specific information that is used to effectively conduct an attack on a particular target (U.S. Army, 1991). This message is created and communicated by an FO and contains the appropriate method of fire, which is determined by factors such as target type and location, potential friendly and civilian presence, and FO distance to target (Strensrud, Fragomeni, & Garrity, 2013). The Fire Direction Center (FDC) receives this message and initiates fire commands to a Firing Unit after the requested fire support is verified or modified. The FO conducts a CFF request in three transmissions in order to communicate detailed information that is necessary for the FDC to execute the mission. Each of the transmissions requires the FO to be highly observant and to utilize effective decision making skills in a high-risk environment. The first transmission consists of the FO identifying himself and declaring a warning order. A warning order contains the method of target location (e.g., shift from known point), type of mission (e.g., fire for effect), and size of requested fire (e.g., battalion). The second transmission consists of the target's location and the location accuracy, the latter of which is critical for the mission to be effective. The third transmission consists of a detailed target description, method of engagement, and method of fire and control. Once the FDC receives the information provided in the three transmissions, the fire commands are reviewed, approved and disseminated to the Firing Unit, which executes the specific methods at the target location. The FO must then observe and report height of round bursts, as well as communicate requests for fire adjustments on the target and the effect of the rounds on the target. These complex and coordinated tasks involve affective skills associated with listening, acknowledging, and responding to the information from the FDC, psychomotor skills associated with perceiving, calibrating, adapting to, and reacting to the current state of the environment,

and cognitive skills associated with evaluating, comparing, and predicting the fire adjustments needed.

Augmented Reality Training. Conventional classroom instruction of the affective, psychomotor, and cognitive skills associated with a CFF mission requires far-transfer: applying knowledge learned in the classroom situation to a far different outdoor context (Champney, Surpris, Carroll, & Cohn, 2014). The potential advantage of the AITT AR-solution for situated learning (Bossard and Kermarrec 2006; Lave & Wenger, 1990) of CFF mission is its ability to simulate real-world contexts that allow trainees to master authentic FO tasks in meaningful and realistic outdoor environments, which means that trainees must attain only near-transfer to achieve preparation for the operational (outdoor) environment. Beyond embedded learning within relevant environments, from a constructivist learning theory perspective, (Dunleavy & Dede, 2014) the AITT provides the following conditions that are likely to enhance learning: (1) allows for social negotiation between the FDC and FO, (2) allows for multiple perspectives (i.e., views of a limitless number of outdoor environments), (3) provides self-directed and active learning opportunities of FO tasks, and (4) supports and facilitates metacognitive strategies within the experience when coupled with an after action review that provides feedback. Within the armed forces domain, AR is also valuable for its ability to repeatedly provide opportunities to train otherwise hazardous or expensive tasks and thus have a high return on training investment.

3 Participants

A total of five (5) U.S. Marines participated in this study. All participants were male, had normal or corrected to normal vision, and ranged in age from 24 to 38 ($M = 31$, $SD = 6.8$). The average number of years of military service was 8.15 ($SD = 2.36$) and participants reported their current role as Forward Observer (FO), Expeditionary Warfare School (EWS) Student, or Joint Terminal Attack Controller (JTAC). The prior experience of CFF missions completed by the participants ranged from 30 to 400 ($M = 132$, $SD = 152$) missions. All participants reported they had trained others to perform CFF mission tasks.

4 Materials

The following tools were used to gather data during the study.

Training Utility Questionnaire (Task Execution): Participants answered 9 questions regarding the system's ability to support the execution and thus practice of CFF tasks. These questions were targeted at understanding the perceived utility of the AITT system for CFF training. Participants answered using a five-item Likert scale anchored by 1-Strongly Disagree and 5-Strongly Agree.

Satisfaction Questionnaire: The Net Promoter Score (NPS; Reichheld, 2003) method was used as a measure of satisfaction with the AITT system for use in CFF training.

The scale has defined cut-offs to indicate the likelihood of an individual to promote a product being assessed. Participants answered the following question: “*How likely is it that you would recommend the AITT for use as a CFF trainer?*”

System Usability Scale (SUS): The AITT’s global usability was evaluated using the SUS (Brooke, 1996), a 10-item Likert scale anchored by 1-Strongly Disagree and 5-Strongly Agree (provides a total score ranging from 0-100).

AITT Fidelity Questionnaire: This questionnaire sought participants’ impressions regarding the realism of the AITT system with regards to the multimodal (sensory) experience. Participants used a five-item Likert scale mentioned above (1-Strongly Disagree and 5-Strongly Agree) to assess the fidelity of the system. The responses for each individual question were averaged.

Simulator Sickness Questionnaire (SSQ): The SSQ (Kennedy et al., 1993) was used to assess any adverse symptoms associated with using the AITT system. The SSQ consists of a checklist of 26 symptoms, each of which is related in terms of degree of severity (none, slight, moderate, severe), with the highest possible total score (most severe) being 300. A weighted scoring procedure is used to obtain a global score reflecting the overall discomfort level known as the Total Severity (TS) score, along with three subscales representing separable but somewhat correlated dimensions of simulator sickness (i.e., Nausea [N], Oculomotor Disturbances [O], and Disorientation [D]).

Presence Questionnaire: Presence, the sense of “being” in a different place than where a user is physically located (Witmer & Singer, 1998), was reported on four subscales. The Involvement/Control subscale reflects a user’s psychological state resulting from focusing attention on stimuli and is comprised of control, realism, and sensory factors. The Natural Interaction subscale addresses control and realism factors affecting the match between virtual and real objects and environments. The Resolution subscale solely focuses on sensory factors. The Interface Quality subscale accounts for distraction and control factors. Participants answered all subscales using a seven-item Likert scale anchored by 1-Strongly Disagree and 7-Strongly Agree.

Immersion Questionnaire: A relevant subset of questions from Jennett et al. (2008) was used in this study, which evaluates a user’s feelings of cognitive absorption and flow. Participants answered using a five-item Likert scale anchored by 1-Strongly Disagree and 5-Strongly Agree.

5 Method

Kirkpatrick’s (1994) first of four levels of evaluation (i.e., 1. trainee reactions, versus 2. learning, 3. transfer and 4. impact), served as the basis for a formative evaluation of the AITT system. Specifically, the present evaluation sought to evaluate reactions from Subject Matter Experts (SMEs) regarding the constructs of training utility, satisfaction, usability, fidelity, simulator sickness, presence, and immersion via questionnaires, as well as evaluate task execution via observation (including notes taken by the

Experimenter while he/she observed participants interacting with the AITT system). The evaluation took place in a large, open field in Quantico, VA. The field was lined with trees and a building and parking lot were at the North end of the field.

Upon arrival to the field, the participant was asked to read and sign an informed consent form, which described the tasks involved in the study and notified him that his participation was completely voluntary. The participant was then asked to fill out the Simulator Sickness Questionnaire (SSQ) Kennedy et al., (1993). Next, the participant was fitted with a portable AITT setup and was allowed to familiarize himself with the system.

The participant was then asked to execute three (3) CFF missions in the following order: (1) Grid method, (2) Polar method, and (3) Shift from Known Point method. The Grid method consisted of a scenario in which the CFF was to be conducted on one squad of enemy dismounts. The Polar method consisted of a scenario in which the CFF was to be conducted on an enemy vehicle in the parking lot next to the field. The Shift from Known Point method used the location of the enemy vehicle from the Polar method in order to execute the CFF on an additional enemy vehicle. All scenarios had targets that were Danger Close, and thus participants were told to ignore Danger Close procedures for purposes of the testing. All three CFF missions allowed the participant to use a simulated VECTOR 21 (for coordinates and range finding, as well as real-world tools including a map, protractor, notepad and pencil. In addition, all CFF missions were self-paced and self-directed. Figure. 1(a) and (b) show participants interacting with the AITT system. The time immersed in the AITT system ranged from 29 to 46 min ($M = 39.4$ min., $SD = 7.0$).

Upon completion of the three CFF missions, the participant immediately filled out the training utility, satisfaction, usability, fidelity, simulator sickness, presence, and immersion questionnaires, The experimenter was present while the participant filled out the questionnaires in order to answer any questions and to take notes of comments the participant provided. The duration of the entire experiment was approximately 1.5 h per participant.



(a) Kneeling



(b) Prone Position

Fig. 1. Participant interacting with AITT system

6 Results and Discussion

The sections below detail the findings from the study.

Training Capability. With regard to training utility, questionnaire responses ranged from 3.6 to 4.7 ($M = 4.19$, $SD = 0.68$), with these scores indicating that most participants agreed (4) that the AITT has the ability to support execution and practice of the tasks necessary to correctly and accurately conduct a CFF, with the exception of one aspect. The average response to the statement, “*The system allowed me to properly utilize all the tools necessary for a CFF mission (e.g., map, compass, Binovs/VECTOR 21, radio)*” resulted in an average score of 3.6 ($SD = 1.52$), which was a neutral score. It should be noted that the only simulated tool utilized by the participants was the VECTOR 21; all other tools (i.e., map, protractor, notepad and pencil) were real-world tools. Further review of comments and notes highlighted the possible cause of this rating as being the interference caused by the Head Mounted Display (HMD) with these real-world tools. While the participant transitioned from observing the AR environment and utilizing the real-world CFF tools, participants needed to ‘flip’ the HMD up to utilize these tools or find other workarounds (e.g., raise head while looking down under the HMD), which made it hard to make use of those other tools. This indicates a need to better merge the elements of the physical real-world environment that are to be coordinated with virtual computer generated imagery.

Satisfaction. With regard to satisfaction, the AITT’s NPS was 40 % (percentage of Promoters [60 %] minus percentage of Detractors [20 %]), reflecting relatively high support for the AITT’s training capability. This NPS is higher than the average NPS of 20 % for popular products, including Microsoft Word, Google Calendar, Dropbox and Adobe Photoshop (Sauro, 2014). The results of the NPS indicate that although the majority of the participants were promoters of the AITT system, there was one detractor and one passive participant. A review of these individuals’ comments and reported perceptions may indicate what elements of the AITT may be driving their dissatisfaction or hesitation to promote the system. Specifically, the detractor was concerned with the physical aspects of the AITT system, stating that the prototype appeared fragile and cumbersome (“Marines would break it”). At the same time, this participant also had positive feelings regarding the system and its use for training CFF. While it is understood that the version of the AITT system evaluated is a prototype, participants had a difficult time understanding what would be the quality and ruggedness of the final operational version. As such, their comments regarding the durability of the system were based on what they observed and experienced. The takeaway from this feedback is that the system must be very ruggedized before fielding. The passive participant also had positive feelings, stating that the system seems like a great tool for teaching basic CFF. He gave several recommendations for improvement, such as making the displays more adjustable and blocking ambient light. These comments point to the need to more adjustable and immersive AR headset.

Usability. With regard to usability, the AITT’s SUS score was a 54, which would correspond with an adjective measure of “OK” usability (Bangor, Kortum, & Miller, 2009); scores above 68 correlate with systems having high usability. Based on the

responses, it was clear that while participants were motivated to use the system, they were hesitant about their ability to use the system unaided (e.g.,: “I think I would need the support of a technical person to be able to use this system” and “I need to learn a lot of things before I could get going with this system” were the highest rated negative statements in the SUS). It must be noted that the version of the AITT evaluated was a prototype and requires additional refinement to ensure a high level of usability.

Simulator Fidelity. Overall, participants agreed that the AITT scenarios were believable and the visual detail of simulated objects was presented realistically, with the exception of Ordnance Damage. Participants tended to be neutral with regard to the degree of realism of the sounds and behaviors of simulated objects. While participants agreed that the Binos looked, felt and behaved as expected in the real world, the VECTOR 21 functionality was rated neutral. Additionally, participants rated the visual simulation update-rate when turning their head as adequate (see Table 1).

Table 1. Degree of realism of AITT and specific objects

Question	Average Response (SD)					
	Ground Vehicles	People	Ordnance (e.g., smoke, explosion, etc.)	Ordnance Damage	Binos	Vector 21
1. The visual detail of the simulated objects was presented realistically in the system	4 (0)	4.2 (0.45)	4.2 (0.84)	3.4 (1.52)	–	–
2. The sounds of the simulated objects were presented realistically in the system	3 (1.79)	–	3.2 (0.45)	–	–	–
3. The behavior of the simulated objects is realistic (e.g., how they moved, reacted, etc.)	2.4 (0.55)	3.5 (0.58)	3.8 (0.84)	3.4 (0.89)	–	–
4. The simulated tools looked, felt and behaved as expected in the real world	–	–	–	–	4 (0)	3.4 (0.89)

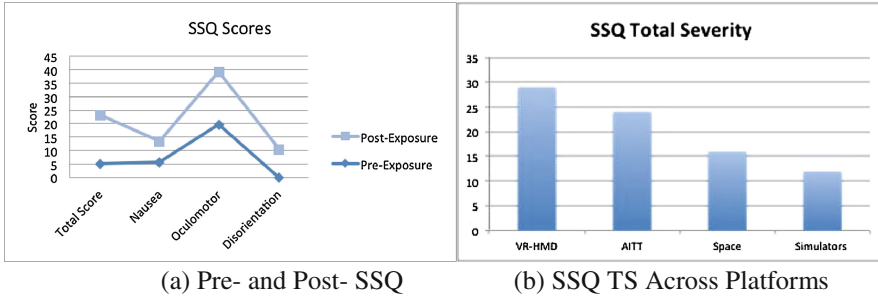


Fig. 2. Simulator sickness results

Simulator Sickness. Figure 2a below shows the average pre- and post- exposure SSQ scores. Exposure to the AITT system resulted in higher total severity (TS), as well as nausea (N), oculomotor (O), and disorientation (D) scores for 3 of the 5 participants. It is interesting to note that the SSQ profile for AITT is $O > N > D$, which is the same as the typical profile for simulators and different from that of virtual reality systems, which is $D > N > O$ profile (Stanney et al., 1998). The significance of this finding is unclear and requires further study. Figure 2b shows that the sickness associated with the AITT AR-technology is a bit lower than that experienced with VR systems but higher than space and simulator sickness (Stanney et al., 1998).

Presence. The subscale means for Involvement/Control and Resolution indicate positive participant perceptions on average compared to the maximum score possible (see Table 2). However, the means of the Natural Interaction and Interface Quality subscales require further study due to the bimodal nature of the results shown in Figs. 3 and 4. Recent efforts investigating the impact of wearable interfaces upon Presence generally report higher levels of Presence than was found in this AITT study (Taylor & Barnett, 2013). The awkward merging of the virtual and real-world discussed in the Training Capability section above, the AR headset adjustability and ambient light issues noted in the Satisfaction section above, and the sickness symptoms reported in the Simulator Sickness section above may be collectively driving down presence. These results point to the need to enhance the naturalness of interaction (the bridging of real and virtual) and

Table 2. Presence subscale ratings

Results				
Presence subscale ratings				
Descriptive statistics	Involvement/control	Natural interaction	Resolution	Interface quality
Mean (<i>M</i>)	54.2	12.2	9.6	12.8
Standard Deviation (<i>SD</i>)	9.83	4.48	1.67	4.21
Maximum Score Possible	84	21	14	21

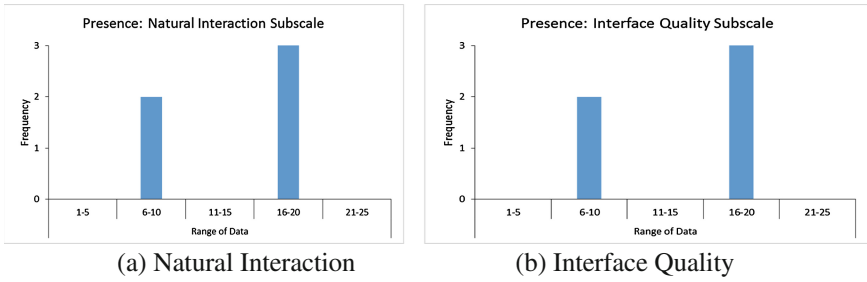


Fig. 3. Immersion subscale frequency charts

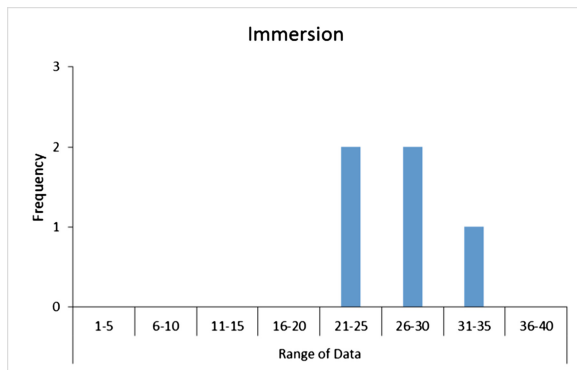


Fig. 4. Immersion frequency chart

interface quality (the AR headset adjustability and ambient lighting conditions) of future AR training solutions.

Italics indicate average rating of Neutral or below.

Immersion. Figure 4 shows the number of participants whose average immersion scores fell within the range categories provided. The full data range is included on the horizontal axis to aid the reader’s understanding of responses relative to the range of possible outcomes. An average immersion score of 27.2 ($SD = 2.86$) out of a maximum of 40 was reported on the modified Immersion questionnaire, which equates to a 68 % immersion rating on the full survey. Previous results (Jennett, et al., 2008) based upon the full survey showed a comparable percentage (58-75 %), which indicated a relatively strong level of cognitive absorption and flow in the AITT system.

Observations and Notes. Based on the observations and notes that were taken during the study, the value of AR training for CFF was evaluated. As aforementioned, past research has suggested that AR provides a rich contextual learning environment to aid in acquiring complex task skills, such as decision making and asset allocation, while providing learners with a more personalized (self-direct /self-paced) and engaging learning experience. These findings continued to hold true in the present study.

In general, one of the things the SMEs liked best about the AITT system was its inherent training value. A typical related comment was – “I really like that you can train marines to CFF and force them to make adjustments in a force-on-force now live fire scenario.” Note that this comment refers to all three dimensions of skills being trained in the AITT system: affective skills associated with communicating and coordinating the CFF, psychomotor skills associated with making adjustments that react to the current force-on-force conditions, and cognitive skills associated with decision making within a live fire scenario. The SME’s comments showed concern, on the other hand, that this training value could be limited by the cumbersome nature of the AR technology – both with regard to form and fit /comfort issues and the clunky merging of the virtual with the real-world tools. A typical comment in this regard was a SME that indicated what he liked least about the AITT system was “the comfort and gear accountability.” Based on the observations in this study it is clear that while AR technology holds promise for enhancing training of complex skills such as those associated with a CFF, there are technology hurdles that must first be overcome. Further, this study continued to support the premise that to derive the most benefit out of AR training, there is a need to use learning theories (e.g., situated learning theories, constructivist theories, etc.) to guide the design of AR solutions, as some past efforts have failed to show tremendous promise (Zhu et al., 2014), which may be why AR has yet to have a compelling value proposition (Nguyen & Blau, 2014).

7 Recommendations

Based on the results obtained from this study, the following recommendations for improving AR training systems are proposed.

1. *Tool Utility:* Participants consistently rated the utility of the AITT as high for its ability to support execution and training of CFF tasks. Nevertheless, the AR system’s equipment affected the participant’s ability to interact with the real world CFF tools necessary for task execution (e.g., map, compass, etc.). During the evaluation, it was noticed that participants were often frustrated with the need to “flip” the headset up in order to use the CFF tools. *As such, there is a need to identify ways to seamlessly transition between virtual and real-world views in AR training applications to reduce user frustration.*
2. *Satisfaction:* Participants were frustrated with the lack of adjustability of the headset and the ambient light in the outdoor environment. *It is recommended that AR displays be designed with adjustability in mind, i.e., provide an appropriate eye relief and lock in angle to make the display easier to see; provide a shield to block ambient light above the display to help with viewing the display on bright days.*
3. *Usability:* Participants indicated feeling intimidated by the highly technical setup required for use of the AITT system and feared it would be cumbersome in its current form to cost effectively support AR training, as technical support would be needed to keep the system up and running. *It is recommended that an AR training solution be coupled with a training management system that can walk instructors and trainees through how to easily set-up and use the system.*

4. *Fidelity*: Participants provided high ratings for the visual realism of ground vehicles, people, and the Binos look, feel, and function. Yet participants provided neutral responses to the visual realism of ordnance damage. Battle damage assessment (BDA) is crucial in determining CFF mission effectiveness. BDA involves evaluating the type, quantity, and location of damage. Thus, BDA tasks involve psychomotor and cognitive skills, such as using sensory cues from the environment to comprehend the battle scene, then estimating the battle damage and interpreting what the impact is to mission effectiveness. *While high fidelity for all cues in an AR training system is not necessary, it is important to identify the set of tasks similar to ordnance damage, which rely on sensorial discrimination and other such skills that require high fidelity to support task performance.*
5. *Simulator Sickness*: The amount of time an individual is immersed in an interactive training system has been found to be directly related to symptoms of simulator sickness (Nelson, Roe, Bolia, & Morley, 2000). After an average of 39.4 min of being immersed in the AITT system, symptoms of simulator sickness were higher than baseline levels, space sickness, and simulator sickness (see Figs. 2a and 2b), particularly with regard to oculomotor disruption (e.g., eyestrain, inability to focus). Until the visual discomfort associated with AR displays can be resolved, *it is recommended that the amount of time in an AR headset be limited (< 30 min) and regular break schedules be imposed in order to reduce the adverse effects related to protracted AR immersion.*

Acknowledgements. This material is based upon work supported in part by the Office of Naval Research (ONR) under contract N00014-12-C-0216. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views or the endorsement of ONR.

References

- Bangor, A., Kortum, P., Miller, J.: Determining what individual sus scores mean: adding an adjective rating scale. *J. Usability* **4**(3), 114–123 (2009)
- Bitter, G., Corral, A.: The pedagogical potential of augmented reality apps. *Int. J. Eng. Sci. Invention* **3**(10), 13–17 (2014)
- Bossard, C., Kermarrec, G.: Conditions that facilitate transfer of learning in virtual environments. *J. Inf. Commun. Technol., ICTTA 2006* **1**, 604–609 (2006)
- Brooke, J.: SUS: a quick and dirty usability scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, A.L. (eds.) *Usability Evaluation in Industry*. Taylor and Francis, London (1996)
- Champney, R.K., Surpris, G., Carroll, M., Cohn, J.: Conducting training transfer studies in virtual environments. In: Hale, K.S., Stanney, K. (eds.) *Handbook of Virtual Environments: Design, Implementation, and Applications*, 2nd edn, pp. 781–795. CRC Press, Boca Raton, FL (2014)
- Dunleavy, M., Dede, C.: Augmented reality teaching and learning. In: Spector, J.M., Merrill, M.D., Elen, J., Bishop, M.J. (eds.) *The Handbook of Research for Educational Communications and Technology*, 4th edn, pp. 735–745. Springer, New York (2014)

- Hale, K.S., Stanney, K.: Effects of low stereo acuity on performance, presence and sickness within a virtual environment. *Appl. Ergonomics* **37**, 329–339 (2004)
- Jennett, C., Cox, A.L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., Walton, A.: Measuring and defining the experience of immersion in games. *Int. J. Hum Comput Stud.* **66**(9), 641–661 (2008)
- Johnson, L., Smith, R., Willis, H., Levine, A., Haywood, K.: The 2011 Horizon Report. The New Media Consortium, Austin, TX (2011)
- Kamarainen, A.M., et al.: EcoMOBILE: Integrating augmented reality and probeware with environmental education field trips. *Comput. Educ.* **68**, 545–556 (2013)
- Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G.: Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *Int. J. Aviat. Psychol.* **3**(3), 203–220 (1993)
- Kirkpatrick, D.L.: Evaluating Training Programs: The Four Levels. Berrett-Koehler, San Francisco, CA (1994)
- Lave, J., Wenger, E.: *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge, UK (1990)
- Nelson, W.T., Roe, M.M., Bolia, R.S., Morley, R.M.: Assessing Simulator Sickness in a See-Through HMD: Effects of Time Delay, Time on Task, and Task Complexity (ADA: 430344). Air Force Research Laboratory, Wright-Patterson Air Force Base, OH (2000)
- Nguyen, T.H., Blau, B.: Market guide to augmented reality. Gartner report No: G00268639. Gartner, Inc (2014)
- Reichheld, F.F.: The one number you need to grow. *Harvard Bus. Rev.* **81**(12), 46–55 (2003)
- Sauro, J.: Software usability and net promoter benchmarks for (2014). 14 October 2014. www.measuringu.com/blog/software-benchmarks14.php. Accessed 8 November 2014
- Stanney, K.M., et al.: Aftereffects and sense of presence in virtual environments: formulation of a research and development agenda. report sponsored by the life sciences division at NASA headquarters. *Int. J. Hum.-Comput. Interact.* **10**(2), 135–187 (1998)
- Stensrud, B.S., Fragomeni, G., Garrity, P.: Autonomy requirements for virtual JFO training. In: *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (IITSEC)*. NTSA, Arlington, VA (2013)
- Taylor, G.S., Barnett, J.S.: Evaluation of wearable simulation interface for military training. *Hum. Factors* **55**(3), 672–690 (2013)
- Army, U.S.: FM 6-30: Tactics, Techniques, and Procedures for Observed Fire. Department of the Army, Washington, DC (1991)
- Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: A presence questionnaire. *Presence* **7**(3), 225–240 (1998)
- Zhu, E., Hadadgar, A., Masiello, I., Zary, N.: Augmented reality in healthcare education: an integrative review. *PeerJ* 2:e469; doi:10.7717/peerj.469 (2014)