Aligning Current VR/AR/MR Training with the Science of Learning

Jeffrey M. Beaubien, Ph.D.
Aptima, Inc.
Woburn, MA
jbeaubien@aptima.com

Evan Oster
Aptima, Inc.
Orlando, FL
eoster@aptima.com

Janet Spruill
Aptima, Inc.
Orlando, FL
jspruill@aptima.com

ABSTRACT

Recent years have witnessed an explosion of interest in the use of Virtual (VR), Augmented (AR), and Mixed Reality (MR) technologies in training. To date, no fewer than 11 empirical meta-analyses have been published on these topics. This high level of research interest is largely due to the affordances that these technologies bring to the learning environment: sensory immersion, interaction, real world annotation, spatial visualization, and contextual visualization, among others (Appelman, 2005; Santos et al., 2014). Unfortunately, there is a considerable gap between the published literature on so-called "extended reality" training – which has been largely spearheaded by technologists and content domain Subject Matter Experts – and the larger science of learning community. With few exceptions (Cook et al., 2013; Zendejas et al., 2013), there has been little attention paid to established principles of instruction – such as schedules of practice, blended instruction, and individualized learning – on learning outcomes. Moreover, when these principles are addressed, they are often considered singly and in the abstract, rather than within the context of more holistic instructional approaches from the fields of human factors, education, and sports psychology. In this practitioner-oriented paper, we identify five holistic instructional approaches – visual orientation, desirable difficulties, contrasting cases, peripheral detection, and stress exposure – that are well-suited for use during scenario-based, extended reality training. For each approach, we provide graphical examples, theoretical justifications, and practical guidance on how to implement them. Finally, we identify common pitfalls that can degrade the effectiveness of extended reality training, along with practical guidance for avoiding them.

ABOUT THE AUTHORS

Jeffrey M. Beaubien is a Distinguished Principal Scientist at Aptima, Inc. For the past 20 years, his work has focused on training and assessing leadership, teamwork, and decision-making skills. His research has been sponsored by the U.S. Navy, the U.S. Army, the U.S. Air Force, and the Telemedicine and Advanced Research Technologies Center, among others. Dr. Beaubien holds a Ph.D. in Industrial and Organizational Psychology from George Mason University, a M.A. in Industrial and Organizational Psychology from the University of New Haven, and a B.A. in Psychology from the University of Rhode Island.

Evan Oster is a Scientist at Aptima, Inc. with experience in courseware development, game-based training, and Augmented Reality-based instruction. His current work centers on leveraging cognitive science and instructional design principles to increase training effectiveness. His past work has involved designing game- and simulation-based training using immersive virtual environments for the Navy. Mr. Oster has also created several Augmented Reality games with a university games lab. He holds both a M.A. in Instructional Design and Technology and a B.S. in Human Development from Virginia Polytechnic Institute and State University, and a M.S. in Curriculum and Instruction from Radford University.

Janet Spruill is Vice President of Programs at Aptima, Inc. She has more than 25 years of experience leading simulation and training, performance support, knowledge management, and organizational development programs across the defense, intelligence, federal, and commercial sectors. She has deep expertise as a human performance technologist and learning solutions architect. Prior to joining Aptima, Ms. Spruill served as Principal Investigator on a Navy-funded effort to develop a new concept of operations for a Future AEGIS Schoolhouse, which was implemented as a Virtual Reality and Augmented Virtuality learning environment for Navy maintenance technicians on the AEGIS destroyer platform. The project included developing novel methods to elicit and codify experiential knowledge of maintenance Subject Matter Experts prior to their retirement. Over her career, she has supported or led more than 45 defense and federal learning and performance support programs.

Aligning Current VR/AR/MR Training with the Science of Learning

Jeffrey M. Beaubien, Ph.D.
Aptima, Inc.
Woburn, MA
jbeaubien@aptima.com

Evan Oster
Aptima, Inc.
Orlando, FL
eoster@aptima.com

Janet Spruill
Aptima, Inc.
Orlando, FL
jspruill@aptima.com

BACKGROUND

"*There are different types of learning that range from skill acquisition to identity formation, and it seems unlikely that a single pedagogy or psychological mechanism will prove optimal for all types of learning." (Schwartz, Chase, Oppezzo, & Chin, 2011)*

Recent years have witnessed an explosion of interest in the use of Virtual (VR), Augmented (AR), and Mixed Reality (MR) technologies for learning. VR systems, such as the Oculus Rift™, use head-mounted displays (HMDs) to provide the learner with a stereoscopic, 360-degree field of view of the simulated environment. Handheld sensors are used to interact with virtual objects using gesture-based interactions, and environmental sensors are used to track the user's location and physical movements. Because of the high processing power required, many VR systems are physically tethered to high-end personal computers or gaming consoles. By comparison, AR systems use camera-captured video of the real world, and then overlay virtual content on top it, for example using a HMD. The user then interacts with the virtual objects using gesture- or voice-based interactions. Finally, MR systems, such as Windows Mixed Reality™, merge virtual content into the real world, for example by integrating digitized objects into the real world that users can interact with, and which can occlude the real-world objects that are hidden behind them. This is the main distinction between AR and MR, in that MR provides the ability for the virtual and real world to interact in real-time. Like VR and AR systems, MR systems also use head-mounted displays. However, the field of view is typically constrained to around 100 degrees. While the popular press often speaks about these technologies as being conceptually distinct from one another, it would be more accurate to say that they form a continuum. In practice, they are also often integrated with other learning technologies, such as gaming.

Since 2006, no fewer than 11 empirical meta-analyses – which collectively summarize the results from nearly 800 journal articles and conference papers – have been published on extended reality training (Bacca, Baldiris, Fabregat, & Graf, 2014; Barron et al., 1998; D. B. Clark, Tanner-Smith, & Killingsworth, 2016; Consorti, Mancuso, Nocioni, & Piccolo, 2012; Cook, Erwin, & Triola, 2010; Cook et al., 2013; Haque & Srinivasan, 2006; Kirkman et al., 2014; Merchant, Goetz, Cifuentes, Keeney-Kennicutt, & Davis, 2014; Nagendran, Gurusamy, Aggarwal, Loizidou, & Davidson, 2013; Vogel et al., 2006; Zendejas, Brydges, Hamstra, & Cook, 2013). The high level of research interest is mirrored by the amount of money that is being invested to acquire and apply these technologies in training. For example, the Department of Defense is expected to invest nearly $11B in extended reality-based training over the next five years (Radu, 2018).

The interest in using these technologies for training is likely due to the affordances that they bring to the learning environment, such as providing learners with: 1) an immersive sensory experience that mixes realistic visual, auditory, and/or haptic cues; 2) the ability to interact with people and objects using their voice and/or naturalistic gestures, thereby reducing extraneous cognitive load; 3) the ability to superimpose digital content such as text, symbols, graphics, and animations onto environmental objects to enhance understanding, and; 4) immediate access to task- and/or contextually-relevant information such as digital checklists and schematics, thereby obviating the need to memorize infrequently-used factual information (Appelman, 2005; Santos et al., 2014). Unfortunately, there remains a considerable gap between the published literature – which has been largely spearheaded by technologists and content domain Subject Matter Experts – and the larger science of learning community. With few exceptions, there has been little attention to established theories and principles of instruction – such as schedules of practice, blended instruction, and personalized learning – on learning-related outcomes (Issenberg, Ringsted, Østergaard, & Dieckmann, 2011; Shepherd, 2017). Moreover, when these issues are addressed, they are often considered singly or in the abstract (Cook et al., 2013; Zendejas et al., 2013), rather than within the context of more holistic instructional approaches from the fields of human factors, education, and sports psychology.

Perhaps it is not surprising therefore that some end-users have become disillusioned with the technologies' inability to live up to their initial hype. For example, according to the 2017 Gartner Hype Cycle, both AR and VR have passed the initial "Peak of Inflated Expectations." AR is currently stuck in the so-called "Trough of Disillusionment," while VR is slowly clawing its way back toward meaningful use via the "Slope of Enlightenment" (Gartner Inc., 2017). As training professionals, it pains us to see the community repeat this cycle of initial hype followed by subsequent disillusionment every time that a new technology becomes mainstream. With this in mind, the purpose of this practitioner-oriented paper is to identify five holistic instructional approaches that are particularly well-suited for use in extended reality training. For each approach, we describe: 1) the types of skills that it is designed to support; 2) the underlying theoretical rationale behind its development; 3) the "core mechanic" by which it operates (Schwartz, Tsang, & Blair, 2016); 4) a brief summary of the evidence supporting its effectiveness, and; 5) a visual example of how it could be applied using extended reality technologies. Finally, we conclude by identifying common pitfalls that can degrade the effectiveness of extended reality-based training, and provide practical guidance for avoiding them.

HOLISTIC INSTRUCTIONAL APPROACHES

This section describes five instructional approaches that are particularly well-suited for use in extended reality training, despite the fact that they have not yet been implemented using such technologies. The five approaches were selected because of: 1) the extensive research literature supporting their effectiveness; 2) the wide range of skill types that they can be used to train, and; 3) their suitability to both Head-Mounted Displays (HMDs) and Cave Automatic Virtual Environments (CAVEs) visual rendering systems. To facilitate the reader's understanding, a common use case is used throughout: ship-based davit (crane) operators and maintenance technicians who are required to operate (e.g., to lift a rescue boat in and out of the water) and maintain (e.g., to perform both scheduled and emergency maintenance procedures) it, respectively. This use case was selected because it should be understandable to all readers, regardless of their military expertise.

Visual Orientation

The visual orientation method was developed to train visual perception skills, in general, and visual anticipation skills, in particular (Hagemann, Strauss, & Cañal-Bruland, 2006). It originated in the field of sports psychology based on research that compared expert-novice differences in performance. This body of research showed that experts possess well-honed, domain-specific visual perception skills. Such skills are acquired over time by observing regular patterns of environmental cues, many of which are quite subtle. In practice, perceptual learning is often implicit. The visual orientation technique was specifically designed to accelerate their development by deliberately drawing the learner's attention to those specific cues that experts use to perform the task.

Visual orientation training is based on the finding that a learner's visual attention can be cued toward specific, information-rich environmental cues (Magill, 1998) by systematically modifying their luminescence (Posner, 1980). Doing so highlights the critical cue without requiring the learner to expend additional cognitive resources. In practice, the visual orientation method frequently uses the "scaffolding" approach (Wood, Bruner, & Ross, 1976). On each practice trial, the critical visual cues are illuminated. Eventually, the scaffold is removed, and the learner is required to perform the task without any external support.

The core mechanics of the visual orientation technique are to: 1) identify the critical visual cues that experts use to perform the task, and; 2) highlight those cues by enhancing their luminescence vis-à-vis the background (Hagemann et al., 2006). The specific visual cues are typically identified by having experts perform the task and "talk aloud" while doing so. However, because experts often have difficulty verbalizing their cognitive processes, it is often necessary to have them perform the task with specific portions of the visual field occluded (Hagemann et al., 2006; Suss & Ward, 2013). The critical visual cues are then revealed when the experts' performance becomes degraded vis-à-vis baseline.

In a classic example of visual orientation training, Hageman and colleagues (2006) identified the critical visual cues that expert badminton players use to determine the direction of an incoming serve. They determined that during the wind up, experts focus their attention on their opponent's trunk. Immediately prior to contact between the racket and the shuttlecock, experts shift their attention to the opponent's arm and then to the racket, respectively. Based on this information, Hageman and colleagues developed a computer-based training program that presented learners with 200 randomly-ordered video clips, each of which depicted an incoming serve. During each clip, the critical environmental cues were highlighted by superimposing a semi-transparent red block over them. Immediately after each trial, the learner was asked to predict where the shuttlecock would land. Because each video clip lasted only a few seconds, the entire training event was completed in less than 45 minutes. As hypothesized, the training produced significant learning gains for novices, but not for experts (Hagemann et al., 2006).

The visual orientation method is fairly straightforward to implement in practice. After identifying the critical visual cues, the instructor systematically modifies their luminescence (see Figure 1). In this example, the instructor has highlighted the frapping lines, which are used to hoist the rescue boat in and out of the water, because they can become rusted or corroded. Practically speaking, it should be fairly easy to implement the visual orientation technique using VR methods, if for no other reason than the frapping lines' Cartesian (x, y, z) coordinates are already tracked by the simulation engine. By comparison, applying the visual orientation approach method using AR technologies could be more challenging, because external sensors would be required to track their location and movement in real time. Additionally, if the goal is to present the learner with numerous training trials in rapid succession (Hagemann et al., 2006), it may be practically simpler to batch script and present them using VR or MR methods. This is particularly true if the training scenarios will vary with respect to environmental conditions (e.g., ambient lighting, distracting noise, and simulated weather effects) or physical locations (e.g., observing the davit from many different distances and angles), which may be simpler to manipulate using VR and MR methods.

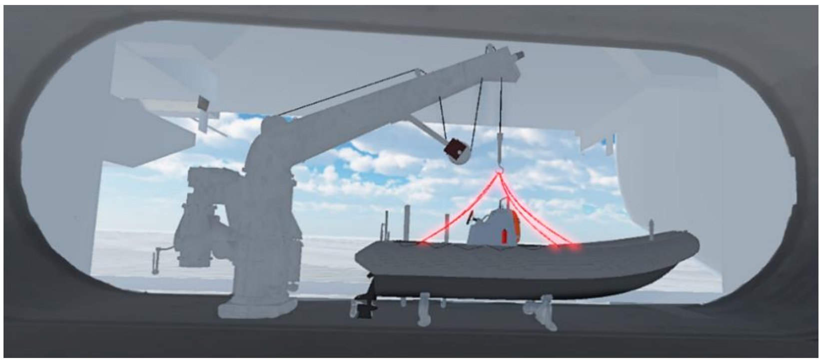

Figure 1. In the visual orientation method, a critical (but subtle) cue is visually highlighted to attract the learner's attention. In this example, the davit's frapping lines have been highlighted as a reminder to check for corrosion before hoisting the rescue boat into the water.

Desirable Difficulties

The "desirable difficulties in learning" instructional method was originally developed to train the recall of factual knowledge, but has since been applied to include a range of skill types, including complex cognitive and motor skills. It originated in the fields of educational and cognitive psychology, based on research which demonstrated that short-term performance during initial skill acquisition is a poor proxy for long-term retention and transfer. For example, certain instructional strategies that improve performance during initial skill acquisition – such as massing the practice sessions over a brief period of time, and providing learners with performance feedback after every trial – actually inhibit long-term retention and transfer (Bjork & Bjork, 2011; Schmidt & Bjork, 1992).

The desirable difficulties method is based on research which has shown that retrieval practice, also known as the "testing effect," produces greater retention and transfer than does the amount of initial skills practice (Roediger & Butler, 2011). Over time, researchers have identified several instructional strategies that capitalize on the benefits of retrieval practice. Collectively, these strategies are referred to as "difficult" because they make short-term skill acquisition more difficult and/or error-prone, thereby giving the mistaken impression that learning is not occurring. The truth, however, is revealed when long-term measures of retention and transfer are collected. Desirable difficulty-based instructional strategies include: distributing the practice sessions over an extended period of time; withholding performance feedback after every practice trial; interleaving the instructional content across two different content areas; systematically varying the conditions of practice; and the frequent use of tests that require learners to generate a response (Bjork & Bjork, 2011; Schmidt & Bjork, 1992).

The core mechanics of the desirable difficulties method are to force the learner to: 1) repeatedly retrieve information that has faded from memory, and; 2) deeply process the learned material, for example by self-diagnosing the causes of success or failure (Schmidt & Bjork, 1992). Desirable difficulties can be incorporated into extended reality-based training in various ways. For example, the instructor could occlude naturally-occurring visual or audio feedback cues during each training trial, and instead provide summary feedback only after every n^{th} trial. Doing so would force the learner to actively diagnose the causes of his/her own success or failure, rather than relying on instructor-provided feedback. Alternatively, the instructor could eschew the all-too-common "crawl, walk, run" instructional sequence, and instead present the learning trials in a quasi-random mix of low, medium, and high difficulty levels (Bjork & Bjork, 2011). Doing so would force the learner to modulate their behaviors in response to changing environmental cues and demands. Finally, the instructor could interleave the practice of two distinct tasks or skills to capitalize on the naturally-occurring process of forgetting and subsequent relearning. In every case, the goal is to make initial skill acquisition more effortful in ways that enhance retention and transfer.

In a classic example of the desirable difficulties method, researchers had a group of children perform a "bean bag toss" game. Visual feedback was withheld from the learners; however, they were provided with knowledge of the result (hit or miss) for each practice trial. One half of the learners practiced throwing from a criterion distance to the target. The other half practiced from a mix of different distances, but never from the actual criterion distance. At the end of training, all learners were assessed on their performance throwing from the criterion distance. As hypothesized, learners who practiced at a range of distances from the target outperformed those who had practiced exclusively from the criterion distance (Kerr & Booth, 1978).

The desirable difficulties method can be implemented in a variety of ways, all of which are fairly straightforward. For example, in order to apply the "interleaved instructional content" approach, the instructor would need to develop training scenarios for two different tasks, such as operating the davit and troubleshooting a system failure. The instructor would then alternate the practice of these two tasks (see Figure 2). Other "desirable difficulty" interventions could be integrated with this approach, such as by distributing the learning trials over time. In the left hand example, the learner practices operating the davit. In the right hand example, the learner practices troubleshooting a system failure with the aid of a digitized system schematic. This process is then repeated until each task is learned to criterion. As noted previously, while the desirable difficulties approach will likely slow the process of short-term skill acquisition, it will significantly enhance long-term retention and transfer, thereby minimizing the need for frequent refresher training. Therefore, the instructor must carefully weigh short-term constraints, such as the time available for initial skill acquisition, with the long-term benefits such as resistance to forgetting and transfer.

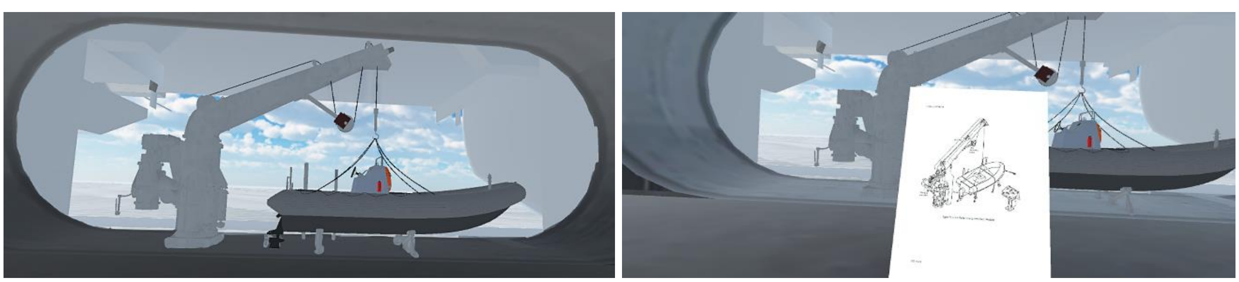

Figure 2. In this example of the desirable difficulties method, the content of two topics has been interleaved. The image on the left depicts a scenario where the learner operates the davit. The image on the right depicts a scenario where the leaner must diagnose a system fault with the aid of a digitized system schematic.

Contrasting Cases

The contrasting cases instructional method was developed to train the ability to discriminate subtle patterns of perceptual cues, and to leverage these patterns during decision making (Schwartz & Bransford, 1998). It originated in the field of educational psychology, and was developed based on research which revealed that novices are often unprepared for classroom-based instruction because they do not have sufficiently well-developed mental models of the to-be-learned factual material. This body of research demonstrated that providing learners with a problem-based discrimination task prior to the classroom-based instruction resulted in better retention and transfer than either the classroom instruction alone, or the classroom instruction supplemented with read-ahead materials (Schwartz &

Bransford, 1998; Schwartz et al., 2016). The contrasting cases method works by requiring learners to actively process a set of stimuli, to invent an explanation that explains the patterns or structural relationships among the stimuli, and to iteratively revise this explanation by applying it to new cases and observing the results.

The method is based on research which demonstrates that with increased domain expertise, humans can make increasingly fine discriminations among a set of subtle stimuli – such as expert sommeliers who can differentiate wines in terms of their sweetness, acidity, and body – particularly when those stimuli are presented side-by-side or in close temporal proximity (Gibson & Gibson, 1955). It is also derived from problem-based learning (PBL) theory, which suggest that active learning methods – such as sorting, comparison and contrast, formulating hypotheses, and testing the veracity of hypotheses using new data – promote the "deep learning" of complex concepts (Barron et al., 1998). In practice, learners are often given multiple stimuli to compare and contrast, and these stimuli differ on a limited number of key characteristics. When attempting to train very complex patterns, it may be necessary to develop multiple sets of contrasting cases, with each set isolating only one or two key characteristics (Schwartz et al., 2016).

The core mechanics of the contrasting cases method are to: 1) identify a set of stimuli that have differing surface properties, but which share the same underlying structural property (or vice versa); 2) provide learners the opportunity to actively process the stimuli, to invent an explanation that summarizes the structural property, and to reflect upon the adequacy of this explanation, and; 3) provide corrective feedback, as appropriate (Barron et al., 1998; Fowlkes, Norman, Schatz, & Stagl, 2009). The contrasting cases method is well-suited to use in extended reality before, during, or after a simulation-based training event given the easy of environmental manipulation. For example, a comparison-and-contrast exercise could be conducted prior to the simulation with the goal of priming the learner to look for critical environmental cues in the upcoming scenario. Alternatively, the learner could perform two training scenarios – which have very similar surface characteristics – immediately back-to-back. Afterward, the instructor would lead an integrated debriefing that compares and contrasts the two scenarios, with the goal of highlighting how subtle differences in environmental cue patterns across the two scenarios would suggest the need for fundamentally different behavioral responses (Fowlkes et al., 2009).

In a classic example of the contrasting cases instructional method, Schwartz and colleagues (Schwartz et al., 2011) taught two groups of students the concept of mathematical ratios such as speed and density. The control group was provided with a formula for each specific ratio, a set of worked examples, and a physics lecture that described the various ratios in detail. This approach closely mirrors the typical instructional approach that is used in the U.S. public school system. The learners then practiced this method for five days. The contrasting cases group was only provided with guidance on how to invent their own contrasting cases. They then practiced inventing and applying contrasting cases to a set of unworked examples. They did this each day for five consecutive days. On the fifth day, the contrasting cases group finally received the same physics lecture that was provided earlier to the control group. While both groups performed equally well on tests of "surface recall," the contrasting cases group performed significantly better on measures of understanding the "deep structure," and were better at generalizing what they had learned to a novel problem (Schwartz et al., 2011).

The contrasting cases approach is fairly simple to implement. After identifying the critical stimuli and the relationships among them, the instructor then develops two training scenarios which are completed back-to-back (see Figure 3). In this example, the instructor has identified two critical situational variables: atmospheric wind speed and load weight, which are displayed in the lower left hand side of the visual field. The instructor's goal is to demonstrate how wind speed and load weight interact to cause the cargo to shift as it is being hoisted in and out of the water. In the first example (low wind speed and medium load weight), the risk of sudden or catastrophic shift is minimal, especially if the operator is following established safety practices. In the second example (high wind speed and low load weight), the risk of sudden or catastrophic shift is much higher, and the need to follow established safety practices is even more critical. Afterward, there is an integrated debriefing to discuss how the critical cues and their relationships should be used to aid the decision-making process.

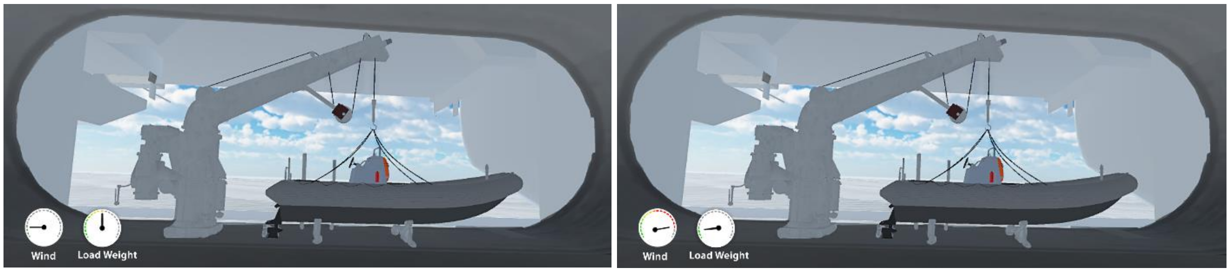

Figure 3. In the contrasting cases method, two training scenarios are presented back-to-back. Afterward, an integrative debriefing helps the learner understand how patterns of critical cues – in this case, atmospheric wind and cargo weight – should influence their decision-making. The example on the left depicts nominal conditions of low wind and average load weight. Under such conditions, the cargo is unlikely to suddenly shift. The example on the right depicts dangerous conditions of high wind and low load weight. Under such conditions, there is a high risk of a sudden or catastrophic shift.

Peripheral Detection

The dual-task method, of which peripheral detection (PD) is just one example, is used to train a continuous motor skill (which is called the primary task) to the point of automaticity. Dual-task methods originated in the fields of cognitive and experimental psychology based on extensive research which has demonstrated that cognition changes qualitatively with increased expertise. During the earliest stages of skill acquisition, decision-making occurs entirely at the conscious level. It is relatively slow, requires large amounts of working memory, and operates primarily by applying textbook- or instructor-provided rules of thumb to each new situation. With increased expertise, decision-making becomes more automatic, requires less working memory, and operates by associatively comparing the current situation to one's accumulated corpus of prior experiences from long-term memory (Dreyfus & Dreyfus, 1980; Evans & Stanovich, 2013). Like all dual-task methods, the peripheral detection approach works by having learners perform two tasks – the primary task and a secondary task – simultaneously (Jahn, Oehme, Krems, & Gelau, 2005). When both tasks can be performed to criterion without interference, the learner has achieved automaticity on the primary task.

The peripheral detection method is based on research which demonstrates that high levels of cognitive workload tend to produce "tunneled" attention, whereby the individual focuses intently on a very limited subset of environmental cues – typically what is directly in front of him or her – and pays substantially less attention to other stimuli in the periphery. This is particularly true when the operator is required to perform two tasks that use the same attentional channel, such as two visual tasks or two auditory tasks (Wickens, 2008).

The core mechanics of the Peripheral Detection method are to: 1) present the learner with a secondary decision task that appears in the learner's peripheral vision; 2) measure the learner's performance on both tasks, and; 3) practice performing both tasks until they can be performed to a pre-established criterion. For example, while the learner is driving a simulated vehicle, a panel of LED lights are displayed in the lower left hand corner of the windshield. Every time that the yellow LED is illuminated, the driver is required to press a micro-switch that is mounted to the steering wheel. Standard measures of performance – such as distance from the vehicle in front – are collected for the primary task, along with measures of reaction time and accuracy for the secondary task (Jahn et al., 2005).

In a classic example of using peripheral detection to assess workload, Jahn and colleagues (2005) tasked professional drivers with performing a PD task while driving a simulated car using different navigational aids. Parts of the driving route were classified as "high information processing demand," because there is a great deal of traffic and the driver was required to yield the right of way to pedestrians and other drivers. Other parts of the driving route were classified of "low information processing demand," because there were fewer pedestrians or other drivers with whom to interact. As hypothesized, the PD method was sensitive to the traffic route workload manipulation.

Implementing the peripheral detection technique is fairly straightforward. The instructor could simply insert a panel of simulated LED lights in the periphery of the visual field (see Figure 4). The lights should illuminate randomly for a very brief period of time (approximately one second), with each illumination being followed by a brief pause during which all of the lights are dimmed. As the learner performs the primary task – such as raising or lowering the davit to the water – he or she is required to make a rapid, a non-verbal response – such as a button press on the handheld input device – in response to a specific colored light. Given that the primary task will require several minutes to perform, there will be numerous measures of response time and accuracy on the secondary task during each training trial. In practice, it can be a challenge to find the appropriate level of illumination for the secondary task. In our experience, doing so requires several "dry runs" with a sample of novice and expert performers. The level of illumination should be such that experts – who can perform the primary task with minimal cognitive load – can perform the secondary task with greater than chance accuracy on their first few practice trials. At the same time, novices – who require a great deal of cognitive resources to perform the primary task – should perform worse than chance on their first few practice trials, but should be able to demonstrate improvements with repeated practice.

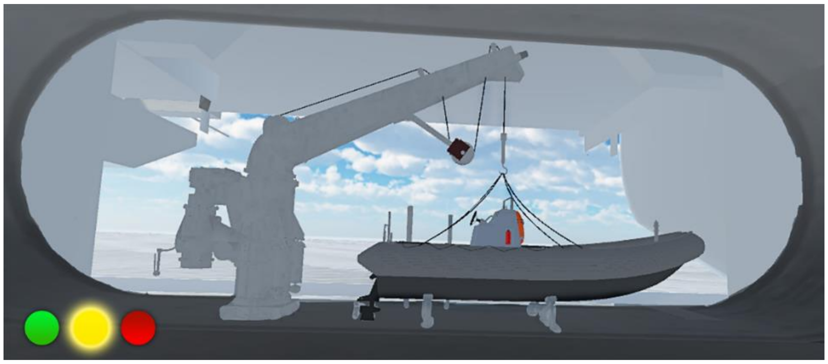

Figure 4. In the dual-task paradigm, the learner must perform the primary task, such as operating the davit, along with a secondary task, such as pressing a button every time the light turns yellow. Performance on the primary task reaches automaticity when the learner can perform both tasks without interference.

Stress Exposure

The Stress Exposure method is designed to help learners maintain high levels of task performance even under highly-stressful, real-world conditions. Stress Exposure Training (SET) originated in the field of human factors psychology based on extensive research which has shown that it is possible to "inoculate" individuals from the negative effects of stress (Driskell & Johnston, 1998). There are three primary phases in SET. The first phase provides learners with factual information about: the types of stressors that they will encounter; the physical, cognitive, and emotional effects of these stressors, and; how their performance may suffer as a result. The second phase provides cognitive and behavioral skills training to minimize the negative effects of stress, such as: attentional-focus training (which helps keep the learner focused on task-relevant stimuli), cognitive control training (which helps the learner regulate distracting emotions such as fear and anxiety), and time-sharing skills training (which helps the learner perform multiple concurrent tasks). The third and final phase provides learners with opportunities to practice these new skills under conditions of increasingly greater and/or varied stressors (Driskell & Johnston, 1998).

The stress exposure method is based on research which demonstrates that stressors produce predictable physiological reactions such as rapid heartbeat, shallow breathing, and muscle tension; cognitive reactions such as increased reaction times, restricted attention span, and a constrained behavioral repertoire; and emotional reactions such as fear or anxiety. While these are normal effects of stress on the human body, they can significantly degrade task performance. For example, the stressors may distract the operator by focusing their attention on task-irrelevant stimuli. Similarly, stressors may cause the operator to experience intrusive thoughts such as fear or self-doubt (Driskell & Johnston, 1998; Driskell, Salas, & Johnston, 2001).

The core mechanics of the SET method are to: 1) identify the critical stressors that learners will likely experience in the actual work environment, such as distracting noises, distracting visuals, or time pressure; 2) develop a series of training scenarios in which the stressor – alone or in combination with other stressors – becomes increasingly more challenging, and; 3) have the learners practice the training scenarios in a sequence of increasingly greater stress. This approach allows the learners to become gradually exposed to more, and a greater variety of, environmental stressors thereby preventing them from becoming immediately overwhelmed. It also enhances their sense of self-confidence by successfully rising to a variety of different challenges.

In a classic example of SET, Driskell and colleagues tasked a group of maintenance personnel with performing two computerized tasks (spatial orientation and memory search, respectively) under different stressors (ambient sound and time pressure, respectively). After assessing the participant's baseline proficiency, the participants completed the three-phased Stress Exposure Training described above. The participants were then required to practice under different task/stressor combinations, and their performance was measured on every trial. Results suggest that the SET intervention generalized to both different tasks and stressor types (Driskell, Johnston, & Salas, 2001).

The third and final phase of SET is the most amenable to training using extended reality methods. A common work-related stressor across jobs is time pressure. To simulate the effects of time pressure – such as when a group of Sailors has fallen overboard and the davit operator must quickly launch the rescue boat to prevent them from drowning – a stopwatch can be presented prominently in the visual field. To help make the time stressor more salient, it can be paired with a corresponding audio countdown, thereby making it nearly impossible for the learner to ignore (see Figure 5). In practice, it may take some "trial and error" to identify the precise stimuli characteristics – the size and location of the timer, the volume of the audio countdown – given the unique task and environmental demands. In this respect, VR technologies may provide a significant advantage, because the environmental conditions (e.g., ambient light and noise) are known with a high level of precision. As a result, precise modifications to the stressor can be made. By contrast, AR and MR technologies may not be as accurate in sensing the environmental conditions, and modifying the stressors accordingly. In all cases, it will be necessary to recruit a sample of experts and novices to help determine the precise stimuli characteristics, as well as the sequence of increasingly more challenging training scenarios. Generally speaking, the stressors should exert a greater negative effect on the novices' performance than on the experts' performance. However, both groups should exhibit some performance declines vis-à-vis "nominal" conditions without the stressor present.

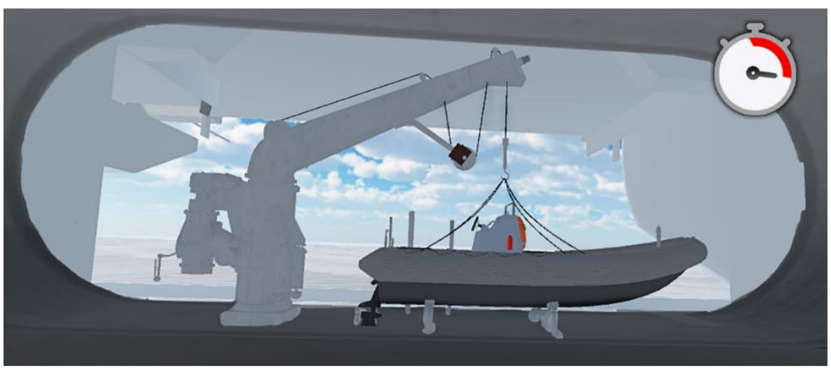

Figure 5. In the stress exposure approach, the instructor inserts a stimulus, such as a countdown timer, to induce the feeling of stress. To be maximally effective, the stressor should utilize multiple attentional channels simultaneously so that it cannot be ignored by the learner.

COMMON PITFALLS TO AVOID

This section describes common pitfalls that can degrade the effectiveness of extended reality-based training, along with guidelines for preventing them. The previous use case – a crew of davit operators and maintenance technicians who are required to operate (e.g., to lift a rescue boat in and out of the water) and maintain (e.g., to perform both scheduled and emergency maintenance procedures) it – is again used to illustrate key concepts.

Distraction

Distraction occurs when extraneous information creates a conflict between focusing on the learning task and focusing on the other information (Harp & Mayer, 1998). Defaulting to a "more is better" approach with regard to information presentation is one of the most common pitfalls that instructional designers make (Ritz, 2015). For example, while following a procedural checklist about how to lower a rescue boat into the water, the davit operator may become distracted by a link to a compelling story about a Sailor who was seriously injured when the davit was not handled safely. In such a case, rather than retaining information about the procedures required to safely deploy the rescue boat, the learner may end up retaining the information about the incident.

Therefore, it is important for instructional designers to determine what information should be "active" at any given time. For example, if the goal is to familiarize the learner with the components of the davit system and its basic operation, supplementary information – such as access to quick reference cards and technical manuals, on-screen "tags" that identify each piece of equipment by name, links to animations that describe how the various system components work, and the like – may be appropriate. However, once the learner has mastered the basics of davit operation, and is now required to practice these skills under increasingly realistic, varied, and difficult scenarios, most of those additional cues should be removed. Therefore, the instructional designer should develop a clear and scalable strategy for presenting contextually-relevant information to the learner. Being clear about the specific purposes of training – such as learning basic procedures vs. advanced skills practice – should help the instructional designer to determine which features need to be active at any given time.

Disruption

Disruption occurs when the learning process is impeded. It may result from a major break in task continuity, or an interruption in the student's sense of "flow." Disruption typically occurs when an interruption occurs in the transition from one main learning idea or task to the next (Harp & Mayer, 1998). For example, in order for a davit operator to construct a coherent mental model of the sequence of tasks needed to successfully deploy the rescue boat, links between the various steps in the causal chain must be constructed in working memory. When disruptive details are presented between those steps in the causal chain, the learner may fail to recognize the critical task linkages. As a result, he or she may interpret each step as a separate, independent event, rather than as part of an integrated chain.

Every time that the learner switches from one task to another, there is a chance for the disruption of learning to occur. In extreme cases, this can result in an error of *prospective memory*, which is defined as forgetting to perform an intended future action. For example, task switching – such as responding to a poorly-timed radio call – may cause the davit operator to lose his or her place in a safety checklist and, by extension, inadvertently skip one or more of the checklist items. When this happens, the learner thinks that he has actually performed the checklist item, when in fact he merely intended to perform it (Loukopoulis, Dismukes, & Barshi, 2009). Errors of prospective memory – in this case, which could be caused by a poorly-designed training scenario – have been known to produce catastrophic outcomes in the real world.

Instructional designers should pay close attention to avoid inadvertently introducing distractions during task practice. For example, when the davit operator is performing a procedural checklist, there should be no requirements for them to temporarily pause the primary task in order to perform an ancillary activity, such as responding to a simulated radio call. This is particularly relevant when the training scenario requires the learner to deviate from a habitual routine (e.g., normally the learner is required to perform tasks A, B, and C in that specific order; but in this unique situation, this the learner is required to perform the tasks in A, C, B order.) If this were to happen, a poorly-timed disruption might cause the learner to forget Task B, because the completion of task C (normally the last task in the causal chain) is typically associated with the procedure being complete.

Seduction

A third pitfall is seduction. This occurs when enticing details in the scenario inadvertently guide the learner in the wrong direction. Seduction is defined as priming inappropriate existing knowledge (suggested by added visual cues, sounds, or words), which is then used to organize the incoming stimuli (R. Clark & Mayer, 2008). Therefore, it is imperative that instructional designers provide only the information pertinent to support the learning objectives.

A series of four experiments support the necessity of avoiding seductive details (Harp & Mayer, 1998). Evidence was found that students who read expository passages with interesting but irrelevant details recalled significantly fewer main ideas, and generated significantly fewer problem-solving transfer solutions than those who read the passages without the seductive details. Furthermore, they found that presenting the seductive details at the beginning of the passage exacerbated the negative effect. These findings reinforce the "less is more" sentiment, especially when the critical information is presented prior to beginning the training scenario, such as during a simulated pre-mission briefing.

Oftentimes, instructional developers include additional details to bolster the learning experience by helping the learner make meaningful connections. Given the risks, developers must vigilantly examine the information that they provide to ensure that the details directly support the learning objectives, and do not inadvertently seduce the learner. Examples include: pre-mission briefings that contain irrelevant or misleading intelligence, and; misleading visual or auditory prompts that do not accurately represent the full message that is being conveyed. However, it should be noted that an exception to avoiding this pitfall is if the learning objective is to help the learner differentiate the most critical information that is presented among a background of irrelevant details. The seduction pitfall can be applied to davit maintenance technician training in the following example. Before starting a simulated troubleshooting exercise to diagnose an oil spot on the deck, the learner checks the digital maintenance logs to see if the last technician had lubricated the bearings. Such information might cause the trainee to mistakenly believe that the oil spot was caused by the last maintenance activity, and not the result of a leaking motor. The seductive effect would further be exacerbated if the title of the most recent maintenance log entry – "Applied lubrication per the technical manual" – was not fully representative of the information contained in the log entry's full text – for example, if the technician lubricated the winch, which is located at a substantial distance away from the motor – because the full text is less likely to be read than the brief log entry title. Therefore, critical details – whether presented prior to the training scenario or during it – need to be carefully considered, lest they inadvertently seduce the learner down an unintended path.

Gamification without Consideration

The final pitfall to avoid is "gamification without consideration." Gamification is an effective learning method that has been shown to improve post-training performance by nearly one-third standard deviations (D. B. Clark et al., 2016), but when gamification elements are applied to training without consideration of the learning objectives, audience, domain environment, or the selected technology, it can negatively impact the training. There are seven major gamification principles, some of which include – clear performance metrics and feedback about one's progress toward the goal; providing the right level of challenge based on the learners' level of ability; supporting learner performance by providing initial assistance and then removing the assistance as the learner improves, and; supporting the learners' social needs by posting their high scores to leader boards. These four gamification elements involve computing measures of human performance during the training scenario and then using those measures for some training-related purpose, such as adapting the training content or motivating the learners to engaged in additional practice (Belanich, Orvis, & Mullin, 2004).

Gamification can be a "double-edged sword" if the elements are applied without explicitly considering the specific instructional objectives and/or audience characteristics. Therefore, it is important to consider whether the gamification intervention induces performance- or learning-related goals, respectively. For example, if learners are inadvertently rewarded for competing against one another – this often happens when a leader board is used to track the highest scores among a cadre of learners – they may attempt to "game the simulator" by selecting a single strategy that maximizes their points across all training scenarios, regardless of whether that strategy would be appropriate outside of the simulator. In such a case, the gamification will have induced a "performance goal orientation." By comparison, if the learners are purposely rewarded for improving upon their prior performance – for example by trying different strategies, and observing their effects on task performance – they will have likely learned a great deal during training. In such a case, the gamification would have induced a "learning goal orientation." With a learning orientated mindset, negative outcomes during training are viewed not as setbacks but rather as opportunities to grow and learn. By comparison, with a performance orientated mindset, errors during training are viewed as failures that are to be avoided. Previous research suggests that learning goals are associated with a variety of positive performance-related outcomes, including higher self-set goals, greater perseverance in the face of failure, and higher proficiency at the end of training (Payne, Youngcourt, & Beaubien, 2007; Utman, 1997). Preventing the gamification without consideration pitfall can be avoided by carefully considering the learning objectives and the target population. For example, if a davit maintenance technician must practice conducting a maintenance task, points can be given for faster completion times, but only if they do not make mistakes. This incentive could motivate the trainee, but place a higher emphasis on correctness.

SUMMARY AND CONCLUSIONS

Despite the substantial amount of resources that have been invested in acquiring and applying extended reality technology for use in training, many such efforts have been divorced from the larger science of learning (Cook et al., 2013; Zendejas et al., 2013). This is unfortunate, but not unexpected. It seems to be the norm, rather than the exception, that the training community-at-large engages in a recurring cycle of initial hype followed by subsequent disillusionment every time that a new technology becomes mainstream.

The purpose of this practitioner-based paper was to identify five validated, holistic instructional approaches that are well-suited for use during scenario-based, extended reality training. In retrospect, these five instructional approaches represent the tip of the proverbial iceberg. Based on informal discussions that we have had with colleagues across the learning sciences, we believe that there are likely to be very few validated instructional techniques that cannot be faithfully represented using extended reality methods. One merely needs to leverage these new technologies in creative ways, for example by: highlighting critical environmental cues (visual orientation); systematically ordering the training scenarios (desirable differences, contrasting cases); introducing secondary tasks (peripheral detection), and; including external stressors (stress exposure training), among others.

We have also attempted to show how the indiscriminate application of technology can inadvertently hinder the learning process. This is perhaps one of the most critical reasons for using well-tested instructional approaches: these methods have been systematically designed such that all of the instructional design features are mutually-reinforcing, and therefore don't inadvertently send the learner "mixed messages" or steer them down the wrong path. It is our sincere hope that this paper helps to begin a dialogue among technologists, instructional designers, and learning scientists – with the ultimate goal of ending the "hype cycle" so that we as a field can stand on the shoulders of giants. We must never forget that those who we serve – the warfighters, first responders, and clinicians, among others – pay the price for our failed attempts to continually reinvent the wheel.

ACKNOWLEDGEMENTS

The views expressed in this paper are solely those of the authors. The authors wish to thank Dr. Ray Perez (Office of Naval Research), Mr. Neal Jefferis and Mr. Charles Arrants (Cape Henry Associates), Mr. Kelly Schneider (PEO Carriers), and Ms. Anastacia MacAllister (Iowa State University) for helping to shape our ideas regarding extended reality training.

REFERENCES

Appelman, R. (2005). Designing experiential modes: A key focus for immersive learning environments. *TechTrends, 49*(3), 64-74.

Bacca, J., Baldiris, S., Fabregat, R., & Graf, S. (2014). Augmented reality trends in education: A systematic review of research and applications. *Journal of Educational Technology & Society, 17*(4), 133-149.

Barron, B. J., Schwartz, D. L., Vye, N. J., Moore, A., Petrosino, A., Zech, L., & Bransford, J. D. (1998). Doing with understanding: Lessons from research on problem-and project-based learning. *Journal of the Learning Sciences, 7*(3-4), 271-311.

Belanich, J., Orvis, K., & Mullin, L. (2004). Training game design characteristics that promote instruction and motivation. In *Proceedings of the 2014 Interservice/Industry Training, Simulation and Education Conference.* Paper No. 1526. Arlington, VA: National Training and Simulation Association.

Bjork, E., & Bjork, R. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. Gernsbacher, R. Pew, L. Hough, & J. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56-64). New York: Worth.

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research, 86*(1), 79-122.

Clark, R., & Mayer, R. (2008). Applying the coherence principle: Adding interesting material can hurt learning. In *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning* (pp. 133-154). San Francisco, CA: Pfeiffer.

Consorti, F., Mancuso, R., Nocioni, M., & Piccolo, A. (2012). Efficacy of virtual patients in medical education: A meta-analysis of randomized studies. *Computers & Education, 59*(3), 1001-1008.

Cook, D. A., Erwin, P. J., & Triola, M. M. (2010). Computerized virtual patients in health professions education: A systematic review and meta-analysis. *Academic Medicine, 85*(10), 1589-1602.

Cook, D. A., Hamstra, S. J., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., . . . Hatala, R. (2013). Comparative effectiveness of instructional design features in simulation-based education: Systematic review and meta-analysis. *Medical Teacher, 35*(1), e867-e898.

Dreyfus, S., & Dreyfus, H. (1980). *A five-stage model of the mental activities involved in directed skill acquisition* Technical Report No. ORC 80-2. Operations Research Center, University of California - Berkeley. Berkeley, CA.

Driskell, J. E., & Johnston, J. H. (1998). Stress exposure training. In J. Cannon-Bowers and E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 191-217). Washington, DC: American Psychological Association.

Driskell, J. E., Johnston, J. H., & Salas, E. (2001). Does stress training generalize to novel settings? *Human Factors, 43*(1), 99-110.

Driskell, J. E., Salas, E., & Johnston, J. (2001). Stress management: Individual and team training. In E. Salas, C. A. Bowers, & E. Edens (Eds.), *Improving teamwork in organizations: Applications of resource management training* (pp. 55-72). Mahwah, NJ: Erlbaum.

Evans, J. S. B. T., & Stanovich, K. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives in Psychological Science, 8*(3), 223-241.

Fowlkes, J. E., Norman, J. W., Schatz, S., & Stagl, K. C. (2009). Contrasting cases: A strategy for advanced learning using simulation-based training. In *Proceedings of the Human Factors and Ergonomics Society* (pp. 1935-1938). Santa Monica, CA: Human Factors and Ergonomics Society.

Gartner Inc. (2017). Gartner Hype Cycle for Emerging Technologies, 2017. In *https://www.gartner.com/smarterwithgartner/top-trends-in-the-gartner-hype-cycle-for-emerging-technologies-2017/*.

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review, 62*(1), 32.

Hagemann, N., Strauss, B., & Cañal-Bruland, R. (2006). Training perceptual skill by orienting visual attention. *Journal of Sport and Exercise Psychology, 28*(2), 143-158.

Haque, S., & Srinivasan, S. (2006). A meta-analysis of the training effectiveness of virtual reality surgical simulators. *IEEE Transactions on Information Technology in Biomedicine, 10*(1), 51-58.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90*(3), 414-434.

Issenberg, S. B., Ringsted, C., Østergaard, D., & Dieckmann, P. (2011). Setting a research agenda for simulation-based healthcare education: A synthesis of the outcome from an Utstein style meeting. *Simulation in Healthcare, 6*(3), 155-167.

Jahn, G., Oehme, A., Krems, J. F., & Gelau, C. (2005). Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. *Transportation Research Part F: Traffic Psychology and Behaviour, 8*(3), 255-275.

Kerr, R., & Booth, B. (1978). Specific and varied practice of motor skill. *Perceptual and Motor Skills, 46*(2), 395-401.

Kirkman, M. A., Ahmed, M., Albert, A. F., Wilson, M. H., Nandi, D., & Sevdalis, N. (2014). The use of simulation in neurosurgical education and training: A systematic review. *Journal of Neurosurgery, 121*(2), 228-246.

Loukopoulis, L. D., Dismukes, R. K., & Barshi, I. (2009). *The multitasking myth: Handling complexity in real-world operations*. Aldershot, UK: Ashgate.

Magill, F. A. (1998). Knowledge is more than we can talk about: Implicit learning in motor skill acquisition. *Research Quarterly for Exercise and Sport, 69*(2), 104-110.

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education, 70*, 29-40.

Nagendran, M., Gurusamy, K. S., Aggarwal, R., Loizidou, M., & Davidson, B. R. (2013). *Virtual reality training for surgical trainees in laparoscopic surgery*. Retrieved from http://cochranelibrary-wiley.com/doi/10.1002/14651858.CD006575.pub3/full

Payne, S., Youngcourt, S., & Beaubien, J. (2007). A meta-analytic examination of the goal orientation nomological net. *Journal of Applied Psychology, 92*(1), 128-150.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3-25.

Radu, S. (2018). The future of war may be virtual. *U.S. News & World Report*. Retrieved from https://www.usnews.com/news/best-countries/articles/2018-03-20/the-us-military-wants-to-lead-the-innovation-game-in-vr

Ritz, L. T. (2015). *Teaching with CAVE virtual reality systems: Instructional design strategies that promote adequate cognitive load for learners.* Unpublished Master's thesis. University of Wyoming, Laramie, WY. Retrieved from http://repository.uwyo.edu/plan_b/3

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*(1), 20-27.

Santos, M. E. C., Chen, A., Taketomi, T., Yamamoto, G., Miyazaki, J., & Kato, H. (2014). Augmented reality learning experiences: Survey of prototype design and evaluation. *IEEE Transactions on learning technologies, 7*(1), 38-56.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*(4), 207217.

Schwartz, D. L., & Bransford, J. (1998). A time for telling. *Cognition and Instruction, 16*, 475-522.

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology, 103*(4), 759.

Schwartz, D. L., Tsang, J. M., & Blair, K. P. (2016). *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them*. New York: Norton.

Shepherd, I. (2017). *A conceptual framework for simulation in healthcare education.* Unpublished Doctoral dissertation. Victoria University, Mebourne, Australia. Retrieved from http://vuir.vu.edu.au/35047/

Suss, J., & Ward, P. (2013). Investigating perceptual anticipation in a naturalistic task using a temporal occlusion paradigm: A method for determining optimal occlusion points. In *Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting* (pp. 304-308). Santa Monica, CA: Human Factors and Ergonomics Society.

Utman, C. H. (1997). Performance effects of motivational state: A meta-analysis. *Personality and Social Psychology Review, 1*(2), 170-182.

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research, 34*(3), 229-243.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors, 50*(3), 449-455.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89-100.

Zendejas, B., Brydges, R., Hamstra, S. J., & Cook, D. A. (2013). State of the evidence on simulation-based training for laparoscopic surgery: A systematic review. *Annals of Surgery, 257*(4), 586-593.