Training Alchemy—Effectively Converting Traditional Training Content to Gold

Summer Rebensky, Adrian Flowers, Jonathan Reynolds, Cait Rizzardo, Peter Neubauer, & Kent Halverson
Aptima, Inc.
Woburn, MA
srebensky@aptima.com, aflowers@aptima.com, jreynolds@aptima.com, crizzardo@aptima.com, pneubauer@aptima.com, khalverson@aptima.com

ABSTRACT

Many institutions are looking to pivot their educational content and training methods to virtual reality (VR). A review of VR training research has demonstrated the ability to improve psychomotor skills, spatial tasks, and knowledge acquisition (Abich et al., 2021). However, training programs transitioned to VR can result in little to no learning gains compared to traditional methods—or in some cases worse outcomes as a result of VR technology design that does not align with or satisfy the task requirements (Howard et al., 2021). Furthermore, although research has demonstrated benefits in novice training environments, less has been done to determine if these methods can lead to mastery of the learning objectives (LOs; Fletcher et al., 2017). A focus on LOs and the learning strategies within VR will be key to ensuring the highest potential learning gains (Abich et al., 2021). The purpose of this paper is to describe a repeatable process that effectively links LOs to the most appropriate learning technology or environment (which may not be VR), aligns training needs with technological capabilities, and assists with decision making under budget constraints. Through each stage of the process, we provide example applications of the process to transition traditional content into VR. Establishing a standard, repeatable process is critical in ensuring that modern training resources are effectively leveraged to still achieve necessary training outcomes. The paper will provide guidance for considerations in implementing VR training and best practices for identifying when to use VR for training and education.

Keywords: Training Content Development, Layered Fidelity, Virtual Reality, Instructional Design

ABOUT THE AUTHORS

Dr. Summer Rebensky, Scientist, Aptima, Inc. has a background focusing on human performance, cognition, and training in emerging systems. Dr. Rebensky has previous experience as a research fellow as a part of the Air Force Research Laboratory's Gaming Research Integration for Learning Laboratory (GRILL) conducting research on drone operations and human-agent teaming utilizing game-based technology. Her other experience includes leading research efforts with the Air Force, Navy, and FAA out of the ATLAS Lab, human factors assessments of a trainer aircraft for the F-35 with Northrop Grumman and developing courseware with Raytheon. Her research experience involves leveraging VR, AR, and game-based technology to optimize human performance in training and operations. Additional research areas include individualizing learning through gamification and adaptive technologies, cybersecurity, human-agent teaming, and trust in AI. Dr. Rebensky received her BA in psychology, MS in aviation human factors, and PhD in aviation sciences with a focus in human factors from Florida Tech. US Citizen.

Dr. B. Adrian Flowers, Senior Research Engineer, Aptima, Inc. has a specialization in cognition and perception in virtual, mixed, and augmented reality (XR) environments. He has leveraged this expertise in two capacities: (1) mitigation of negative physiological events in real-time systems, and (2) enhancement of human perception using a full-body approach to human augmentation. In pursuit of these goals, he has designed and executed a number of human subjects studies and used the results to develop an expanding suite of XR capabilities. These include a psychophysical study on visual and cognitive degradation in augmented reality (AR) funded by C5ISR's Night Vision and Electronic Sensors Directorate (NVESD), which was used to develop an AR remote collaboration tool for Mission Command and a Department of Transportation-funded study on the use of multimodal sensory alerting to inform the design of future roadway alerting capabilities. Dr. Flowers also has extensive experience with real-time system design, from his work with Lockheed Martin, where he was the lead software engineer, designer, and team lead on the development of a highly-connected real-time system to support submarine training. He received his PhD from University of Rhode Island, where he led studies on the detection of physiological signatures of cybersickness in the Virtual Immersive Biometry-Enabled Environment Lab. Dr. Flowers also holds an MS in computer science from University of Rhode Island and a BS in computer science from Rutgers University. US Citizen.

Mr. Jonathan Reynolds, Associate Software Engineer, Aptima, Inc. has three semesters of experience developing software solutions for VR and AR technologies as an intern at NASA out of the Langley Research facility. He uses his experience and passion for VR and AR technologies to create impactful solutions to improve and solidify the human interaction with a virtual and/or augmented reality space. Mr. Reynolds received his BS from Shawnee State University in simulation/game design in programming. US Citizen.

Dr. Cait Rizzardo, Research Scientist, Aptima, Inc. has experience in performance assessment, evaluating simulation-based training programs, in spatial navigation research, and in developing testing protocols for research done in virtual environments. Currently, she supports several projects researching training effectiveness across a variety of audiences including pilots, maintainers, and ground and logistics operators for both the Air Force Research Laboratory (AFRL) Warfighter Interactions & Readiness Division and Air Education and Training Command (AETC). Dr. Rizzardo has assisted with and led projects involving spatial knowledge acquisition, cross-cultural linguistic norms, and display design and evaluation for after-action review. Dr. Rizzardo holds a PhD and MS in human factors/industrial and organizational psychology from Wright State University and a BA in psychology from Hollins University. US Citizen.

Mr. Peter Neubauer, Principal Software Engineer, Aptima, Inc. specializes in military training simulation systems, automated human performance assessment, and data analytics driving more effective and efficient training. He has more than 25 years of combined professional and academic experience creating innovative solutions for customers through expertise in technical project leadership, software design and architecture, software system integration, database systems, and real-time computer networking. He provides technical and project leadership for PETSTM, the Air Force’s automated human performance assessment, data storage, knowledge management, and analytics solution for live/virtual/constructive training. He is experienced with turning multi-modal, multi-format data into actionable data sets and applying big data techniques to deliver value to individuals, instructors, decision makers, and scientists. He works closely with Air Force pilots, research teams, and engineering teams to identify scientifically sound and technically sustainable solutions to emerging problems. Mr. Neubauer holds an MS in embedded systems and a BSE in computer systems engineering from Arizona State University. US Citizen.

Dr. Kent Halverson, Principal Scientist and Senior Director of the Training, Learning, and Readiness Division, Aptima, Inc. has research experience in training development, multi-modal performance measurement, organizational analysis, unobtrusive measurement, social network analysis, leadership development, training design and evaluation, survey development, and quantitative data analysis. Dr. Halverson manages a portfolio of research projects for a variety of DoD research agencies such as the Air Force Research Lab and the Defense Advanced Research Program Agency (DARPA), across a variety of operational domains (e.g., military pilots, infantry soldiers, intelligence analysts), levels of analysis (e.g., individuals, teams, human-machine teams, multi-team systems, organizations), and research perspectives (e.g., cognitive, behavioral, process, system). Dr. Halverson has held faculty positions at the Air Force Academy and the Air Force Institute of Technology as the Director of the Graduate Engineering Management Program. He holds a PhD in business management from University of Florida, an MS in civil engineering from the University of Illinois, and a BS in civil engineering from the US Air Force Academy. US Citizen.

Training Alchemy—Effectively Converting Traditional Training Content to Gold

Summer Rebensky, Adrian Flowers, Jonathan Reynolds, Cait Rizzardo, Peter Neubauer, & Kent Halverson

Aptima, Inc.

Woburn, MA

srebensky@aptima.com, aflowers@aptima.com, jreynolds@aptima.com, crizzardo@aptima.com, pneubauer@aptima.com, khalverson@aptima.com

INTRODUCTION

With the rise in virtual reality (VR), many program directors, program managers, instructional designers, engineers, and instructors wish to bring this modern capability into the classroom as an educational tool. Often, systems or technology are acquired and VR content developed before assessing whether the technology is the *right fit* for the learning objectives (LOs). Instructors may feel pressure to join the movement and transition their training content into VR worlds. When considering updating traditional classroom content for modern technology, instructional designers must not only envision the learning content in a new way, but also consider the stakeholders, developers, and end users. Some considerations include: (a) LOs, (b) perceptual requirements of learners to understand the content or task, (c) the measurement of higher level competencies or task proficiencies, (d) the integration with outside systems or learning management platforms, (e) perceptual affordances of learning technologies, (f) measurement and assessment technologies, and (g) the resources allotted not only for development but also for maintenance of the VR, training time, and physical space to set up the training equipment. When considering each of these elements, training content can be effectively moved to a VR environment while avoiding potential negative training outcomes.

Maintenance Training Next (MXN) was a significant achievement and example of how traditional training content can be moved to a VR environment--this effort resulted in a 50% increase in skill acquisition and proficiency attainment (Trabysh et al., 2021). The MXN course trained and allowed for skill rehearsal in VR and then trainees demonstrated those skills in hands-on evaluations. However, the MXN development process cannot be easily replicated for other domains or even other stages of maintainer training. The development time and costs for MXN were relatively high — Air Force staff and several contractor partners required over 2 years to develop a full initial iteration of the VR course for aircraft maintenance fundamentals. Additionally, though some skills transferred well, others were not supported by training-in-VR to the same degree. Instructors who rush to develop VR content may not see similar gains without following a standardized process that considers the wide variety of factors.

To optimally balance costs and outcomes for future courses in VR, course developers should ask themselves, "What are my training objectives and how is each objective best served by technology?" This paper aims to present a structured and repeatable process that can guide selecting and designing modern training systems in a way that optimizes learning and usage of newer technologies, without wasting time and money on translating training content to suboptimal modalities.

VR AND TRAINING

VR, specifically head-mounted VR headsets, is a technology medium that allows for increasingly immersive experiences and is becoming more accessible every day. Traditionally, computers and computer systems allow for basic simulations through 2D monitors or more "static" display options, leaving the user with limited interactions with the virtual environment. On the other hand, head-mounted VR devices not only allow for a more dynamic, stereoscopic display with users being visually engaged, but also allow users to look in all directions-- a naturalistic user interaction that fully immerses the learner like never before. Over the past few years, VR technology has become more prolific, more affordable, and better supported in the development environment, positioning the technology as a leading force in the simulation world. VR training has gained momentum due to its history of clear interactive capabilities as shown in VR reviews: (1) VR training can lead to improved knowledge retention particularly with respect to provided multimodal cueing or increased interactivity when compared to instructor-led training or PowerPoint presentations (Abich et al., 2021), (2) VR has also demonstrated effectiveness compared to less immersive technologies for cognitive skills, psychomotor skills, and affective skills (Jensen & Konradsen, 2018). Initiatives in medical and industrial domains have demonstrated benefits in utilizing VR as a tool for familiarization, safety procedures, or preplanning (Abich et al., 2021). Additional benefits to VR include increased confidence, motivation, and self-efficacy (Abich et al., 2021). Compared to classroom settings, trainees may find it easier to stay engaged and attentive in a VR environment (Sacks et al., 2013). Along with training benefits, VR is appealing for its fusion of the advantages of computer-based training (i.e., training events only have to be built once, digitally accessible and therefore transferable, unobtrusive and automated performance evaluations, associated cost savings) and live exercises (i.e., approximation of real-world spatial aspects and physical movements, demonstration of real-world results of actions). VR can provide immersive environments and, most importantly, interactive scenarios in which students can reinforce and apply their understanding of the content.

Although VR training has demonstrated potential positive learning outcomes, there are other instances that demonstrate VR may not be as effective across domains and particular skills. As lower-fidelity virtual technologies continue to advance in interactive capabilities (computer-based environments, tablet training, etc.) the effectiveness of VR training compared to lower-cost and more accessible mid-fidelity systems has begun to taper (Howard et al., 2021). Additionally, there are mixed results as to whether VR can reduce training time, with some training applications resulting in longer training times and others leading to shorter training and certification times (Abich et al., 2021). Research has also demonstrated VR as less effective compared to lower immersion technologies due to technology limitations or challenges, cybersickness, or distractions from learning in the VR system (Jensen & Konradsen, 2018). In addition, utilizing VR in non-interactive ways (e.g., reading pdf, watching videos) will not lead to learning gains simply by placing content within VR (Abich et al., 2019). More research is needed on a long-term scale, after initial evaluations, extending to the operational field to determine if the training transfer to real-world operational settings (Woon et al., 2021) generates a sufficient return on investment in VR.

A lack of research on mastery level attainment leaves open questions as to whether VR training platforms can train beyond the novice level (Fletcher et al., 2017). Ensuring mastery level learning will be particularly important as programs transition entire curricula to VR with insufficient consideration as to whether VR can fully support and optimize comprehension of that content to achieve the LOs. Other meta-analytic reviews by Kaplan et al. (2021) have found that VR training can lead to no learning gains compared to traditional learning. The results from empirical studies range from demonstrating no gains to some gains in VR training. The results are also mixed based on whether the LO was physical or cognitive, with more consistent findings for physical applications. Utilizing VR in building and worksite construction training has demonstrated improved technical work, even one month later. However, the same advantages over classroom training were not observed for safety protocols (Sacks et al., 2013). Many of these evaluations compare VR training to traditional instruction or no instruction. Instructional system design as a process may lack consideration for the trainee's perceptual needs and the correct mapping to the fidelity of emerging technologies (Stacy et al., 2017). The lack of evaluation on whether the *content* put into VR or the *way* it was designed makes it difficult to determine if VR is ineffective in training a particular skill or if the ineffectiveness was due to the way that training scenario was designed. The range of potential influencing factors and the uncertainty regarding VR training effectiveness necessitates a robust process for designing VR training to ensure effectiveness. We posit that the key to effective VR training is mapping the LOs to key stimuli for the trainee, ensuring a task-technology fit, and then developing content and assessments that effectively achieve the LOs and resulting performance.

THE SECRET TRANSFORMATION OF BASIC TRAINING TO VR TRAINING GOLD

The philosophy of alchemy aims to turn lead or copper into rich metals such as gold. Similarly, the transformation of course content into the perfect training solution can feel impossible to obtain. The allure of VR can be very appealing, and many managers are vulnerable to adopting VR when it is sold as a panacea or if it is perceived as being in fashion, (Abrahamson, 1991). Fortunately, by using systematic approaches, the process of bringing traditional learning content into modern technologies can be effective. Without using an approach rooted in LOs and training needs, one will inevitably end up with an ineffective, overengineered, or over-budget solution—much like the failed material mixtures in alchemy. The following sections detail how to put the end goal—improved training— first, building the goal into each step of the process. Figure 1 presents the overall process.

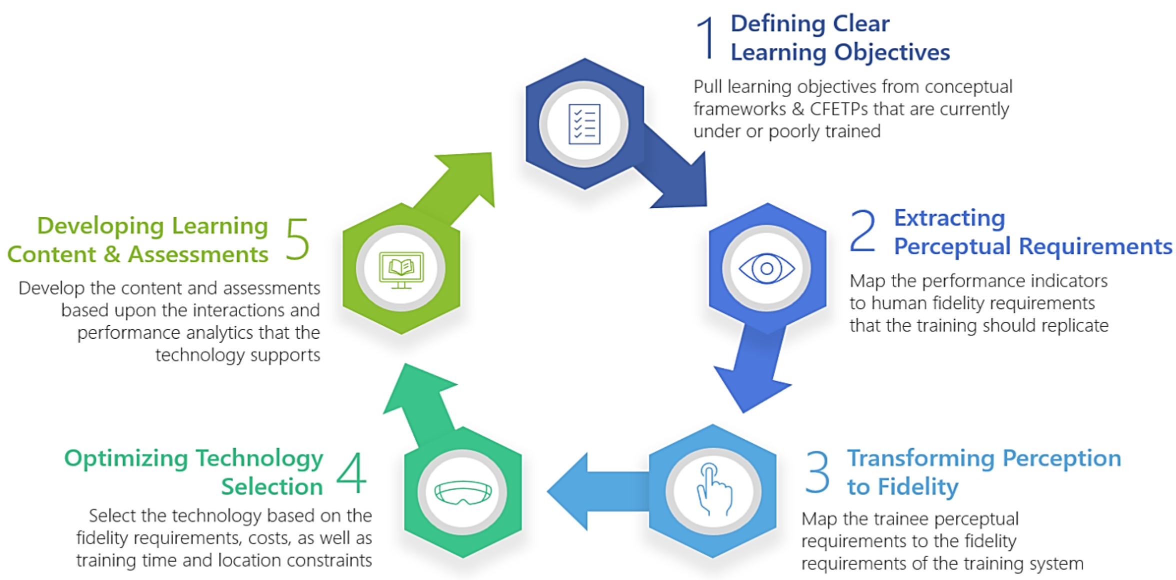

Figure 1. Transitioning Traditional Content to VR

Step 1: Defining Clear LOs

With the growth of VR in commercial, military, and consumer domains, the desire to capitalize on the gains of the immersive platform is expanding to new potential training and educational domains. There is one common pitfall in this approach. Selecting a technology before identifying training areas with deficiencies is entirely putting the cart before the horse. By starting with the training goals, we can ensure the content, technology, and assessments support achieving our desired learning outcomes. Instructional designers can leverage many existing materials to determine which LOs to target. For example, a Career Field Education and Training Plan (CFETP) or Specialty Training Standard (STS) can provide actionable LOs. CFETP examples for a structural mechanic include skills such as using shop equipment like drill presses and band saws, classifying damage to metal structures, and opening and closing discrepancies in the Integrated Maintenance Data System (Department of the Air Force, 2019). Extracting key skills, tasking, or knowledge that a trainee must acquire to successfully perform their job allows one to begin dissecting what the training content must include. In some domains, competency frameworks or models have been developed to capture what skills are necessary for a position. The competency model for military nursing includes combat casualty care skills, regulations, patient safety, nutrition support of the wounded, casualty identification and assessment, and many other skills (Ma et al., 2021). For Army leadership, some of the higher-level competencies include facilitating ongoing development, maintaining relevant cultural awareness, remove work barriers, or designate, clarify, and deconflict roles, and building trust (Horey, et al., 2004). An argument can be made that not all of these skills can be trained using a textbook or classroom environments. Similarly, it should be argued that not all of these skills are ideal to train in VR.

Which skills should be considered for translating into VR training? Many hands-on and knowledge-based LOs can be achieved within VR training that may be insufficiently trained through other means. Training content must build on itself to reinforce acquired information, introduce new, often complex, information, and address both high-level competencies and task proficiency. If addressing high-level competencies, training content may be more focused on ensuring foundational understanding, whereas with training focused on task proficiency, training content may focus on repeated scenario-based assessment under varying conditions. Educational frameworks that classify the spectrum of cognitive skills required for learning, such as Bloom's Taxonomy, can be leveraged to determine the spectrum of training goals and thus, help determine the fidelity and technological requirements to ensure those goals are met. Bloom's Taxonomy labels and organizes lower- to higher-level cognitive skills: knowledge, comprehension, application, analysis, synthesis, and evaluation (Bloom, 1956). *Knowledge* is the foundational cognitive skill and is demonstrated by retention of specific, discrete pieces of information; multiple-choice questions are often used to assess someone's knowledge. The demonstration of this skill and its assessment do not provide evidence of *Comprehension*, a higher-level cognitive skill demonstrating the understanding of what information means within a context. Trainees who acquire a level of comprehension can exercise that skill across various situations, *Applying* their knowledge. These lower-level skills are incorporated into a trainee's ability to then *Analyze*, *Synthesize* and *Evaluate*, activities that involve critically appraising their application and the results, and which represent the culmination of learning. These levels of Bloom's taxonomy can be mapped to the required knowledge, skills, and abilities found in publications such as CFETPs. Although there are clear gains for utilizing VR for the apply, analyze, synthesize, and evaluate stages, there may be additional gains developing for knowledge and comprehension within VR. VR can provide simulated scenarios which allow trainees to apply the skills they've learned in realistic scenarios. However, VR could also provide the opportunity to explore concepts and relationships in a self-guided manner which can provide a strong foundation for later stages of Bloom's taxonomy. There is no clear stage of learning models in which VR is preferrable, but they can be utilized to determine if VR meets the stages of the framework more-so than other technologies or solutions.

Instructional designers can consider a task or required knowledge and begin to determine whether the training should be focused on competencies or task proficiency. In training focused on competencies, the goal is to ensure the trainee has foundational knowledge, skills, or abilities to complete their day-to-day job. In task proficiency training, the goal is to ensure that the knowledge, skills, or abilities are at the required level of expertise (Wolters et al., 2014). Once a clear LO is derived, we can then determine what is necessary to include in the learning content. In the absence of identifying the details of what the trainee *really* needs, the effort is at risk of overengineering the training content or developing training that does not ultimately lead to positive training outcomes. Training focused on task proficiency may also include assessing skill decay and ensuring a trainee remains at the acceptable level over time. Depending on the goals and the style of training, it may require different approaches to the training itself. For example, a high-level competency LO may only require that a trainee knows how to complete a task, whereas a task proficiency LO may require that a trainee can compete the task within an allotted amount of time. Fortunately, with CFETPs, the level of competency for each knowledge, skill, ability, and other characteristics (KSAOs) are specified. For example, task knowledge levels identified in CFETPs for maintenance include: (a) "can name parts, tools, and simple facts about the task", to (b) "can determine step by step procedures for doing the task", to (c) "can identify why and when the task must be done and why each step is needed", to (d) "can predict, isolate, and resolve problems about the task (Advanced Theory)" (Department of the Air Force, 2019). Some CFETP-listed task levels may provide accurate descriptions whereas others may require reworking or higher-level abstraction to translate them into actionable LOs for a training program. An abstracted LO from a CFETP for maintainers may then be, "can determine step by step procedures for classifying damage for metal structures".

Hands-on LOs could include performing a maintenance procedure on an aircraft, suturing a wound, or briefing a unit on a mission plan. Many knowledge-related tasks can also be addressed in VR within the same domains such as planning and prioritizing the maintenance tasks of the day for a maintenance crew, determining how to triage patients, or learning the resources to access to draft a mission plan. An LO with a Bloom's level of knowledge may only require a trainee to be able to recite a piece of information, which can be achieved in many ways to assess competency. Within a nursing competency framework, an LO may be, "be able to apply nutritional support of the wounded under a variety of situations". An LO with a focus on task-proficiency may be "remove mechanical corrosion at a particular rate per square inch". With higher level competency LOs, VR may provide more opportunities to explore potential combinations and be utilized as a supplemental training aid. With task proficiency LOs, VR can allow a trainee to act out a procedure in scenario-based training to assess task proficiency. In this instance, a skill derived from the CFETP may be to remove corrosion, and VR can provide opportunities to inexperienced trainees for increased repetitions to hone their skills. Note that pre-existing LOs not developed using a systematic process, such as proposed in this paper, or developed without the support of personnel trained in job analysis, may not be clearly defined or appropriately worded. Consequently, LOs should be sufficiently overhauled to allow mapping to perceptual requirements. In addition, as missions, tasks, or equipment change over time, LOs should be iteratively refined or updated. The time spent in this step will provide a solid foundation for the rest of the process. Once you have developed clearly defined LOs, the next step is determining the perceptual requirements (see Figure 1).

Step 2: Characterizing Perceptual Requirements

A critical question in the design of VR training content, is what content needs to be developed and how trainees should interact with that content. Humans interpret and interact with the world through six bodily senses, taste, vision, olfaction, hearing, touch, and proprioception. Beyond physical perception, there are also cognitive processes and mechanisms employed in real-world operations that can be lost in a simulated training scenario (e.g., state-based learning effects, social response behaviors). Exact replication of physical and cognitive perception within VR is not feasible, so developers must select a subset of perceptual requirements to simulate at a level that maximizes the efficacy of training. To identify which perceptual requirements should be satisfied, an outcome-driven development model can enable developers to map training outcomes to underlying perceptual requirements.

A relevant framework that maps human perceptual domains to training indicators and outcomes is the Layered Fidelity Framework (Stacy, Walwanis, Wiggins, & Bolton, 2017). This framework discretizes human physical, perceptual, and cognitive capabilities into Pedagogical, Social, Cognitive, Intuitive, Perceptual & Motor, and Sensory & Muscular domains. Each of these perceptual domains is associated with exemplar outcomes as shown in Table 1.

Using this framework, instructors can identify the relative importance of each of the human fidelity layers with respect to stimulating the learner to achieve the LO. For example, early studies on VR for maintenance training found that when trying to minimize procedural training errors, it was most effective to match the pedagogical methodology of VR training to the real-world use-case, whereas task completion time, as a result of visuomotor execution (e.g., muscle memory), was reduced by instead focusing their efforts in Motor & Muscular domains (Hochmitz & Yuviler-Gavish, 2011). This framework enables instructors to define training requirements, by identifying the perceptual domains associated with required LOs.

Table 1. Layered Fidelity Framework

Human Fidelity Layer	Associated Performance Indicators
Pedagogical	Sequencing of training experiences to match competency dependencies Time to meet training criterion
Social	Trainee collaborative skills and understanding of interaction dynamics
Cognitive	Working memory limit effects; Problem-solving skills
Intuitive	Ability to predict subsequent events; Response time; On-task accuracy
Perceptual & Motor	Attentional (and inattentional) effects; Effects of context on perception; Effects of knowledge on perceptions
Sensory & Muscular	Psychophysical effects

Conversely, if training requirements are not based on learning outcomes, the result can be catastrophic. In 2018, a Boeing 737 MAX aircraft crashed, killing everyone on board. The investigation for this incident revealed that because of the 737 MAX's similarities to the 737, Boeing did not feel that it was necessary to make a simulator to explicitly emulate the 737 MAX and had their pilots train on existing 737 simulators. One of the key differences between the 737 MAX and 737, is that the 737 MAX made use of the maneuvering characteristics augmentation system (MCAS), designed to reduce stalling by automatically assisting to keep the aircraft level when its angle of attack is deemed irregular. This assistance was provided by lowering the nose of the aircraft and pushing the pilot's yoke in the down direction. However, because the simulators in use were for the 737 and not the 737 MAX, they did not simulate intervention by the MCAS, or train pilots on the steps required to interact with the MCAS (e.g., how to understand when it is active or disable it when necessary). The investigation found that during the incident, the aircraft was rapidly going up and down before it crashed, leading investigators to determine that pilot was likely fighting with the MCAS system, which may have erroneously activated, before the accident occurred. In response, Boeing posted a bulletin notice explaining the failure with the MCAS system, but still did not explicitly train pilots to recognize, understand, or respond to MCAS failures. In 2019, a second 737 MAX crashed, once again killing everyone on board, and once again, the aircraft was found to be rapidly moving up and down before the accident as if the pilot were fighting with the MCAS system (Johnston & Harris, 2019).

Boeing felt that the existing 737 training processes were sufficient to make pilots proficient to use the 737 MAX. These simulators provided trainees a realistic graphical environment to practice standard procedures and enabled them to practice usage of the aircraft in normal operating conditions. However, this methodology failed these pilots in three key domains: (1) there were no *pedagogical* processes in place to formally educate and evaluate understanding of the MCAS system, (2) the simulator did not simulate the *perceptual & motor* characteristics of an MCAS associated emergency, which meant that pilots were not able to recognize the indicators that MCAS was active and malfunctioning, and (3) the trainees were unable to practice the critical *sensory & motor* responses necessary to mitigate an MCAS emergency, and defaulted to attempting to level out the aircraft, which is what their training had taught them do to when the aircraft tilts downward as they experienced.

After the second accident, every 737 MAX in use across 41 countries was immediately grounded. With properly and systematically mapped LOs to perceptual requirements and thus training systems, Boeing would have recognized the new perceptual requirements to train to the 737 MAX and identified ideal training systems for these gaps. Similarly, not defining what needs to be a part of the training experience in VR, particularly when VR serves to replace live or hands-on training, could have unintended consequences; inevitable incidents will likely occur in operations after training with VR that does not encompass the trainee's perceptual needs. With clearly defined perceptual requirements, the next step of the process is determining how the perceptual needs translate to technological capabilities (see Figure 1).

Step 3: Transforming Perception to Fidelity

After the perceptual requirements necessary to support training have been identified, the ways in which training technologies can satisfy the perceptual requirements must then be determined (i.e., their fidelity). If fidelity specifications are underspecified, trainees may not receive the necessary stimuli; conversely, over-specifying fidelity requirements can lead to unnecessary demands on processing power, significant and extended cost expansion, and increased development time. To accurately determine fidelity requirements, instructors must determine the LOs necessary for trainees within each perceptual domain. For example, during an inspection, a maintainer must be able to recognize and respond to a component that has malfunctioned. Once the maintainer has identified that the component has malfunctioned (perceptual & motor), they must immediately resolve any risk caused by the component (sensory & muscular), then systematically check associated systems to assess the cause of the error and resolve the root issue (pedagogical). While working through this repair process, the maintainer must monitor the state of other components in parallel in an information dense environment (cognitive), increasing the difficulty of the task. Table 2 provides an example breakdown of the domain-specific fidelity requirements for this example scenario.

Table 2. Maintenance Repair Fidelity Breakdown

Fidelity domain	Trainee expectations	System requirements
Perceptual & motor	• Visual recognition of necessary components • Visual recognition of alerts associated with component malfunction • Audio recognition of alerts associated with component malfunction	• Replication of the shape and distinguishing visual characteristics of platform components and alerts • Replication of associated auditory alerts
Sensory & muscular	• None	• None
Pedagogical	• Understand steps to resolve component malfunction • Understand where to seek out or find information on unexpected occurrences	• Scaffolded skill training • Audio feedback for error correction
Cognitive	• Identify component malfunction in information and sensory dense environment • Identify rarely occurring component malfunction	• Replication of stimuli for sensory interference • Replication of working memory requirements • Non-uniform error elicitation

In the example requirement breakdown in Table 2, it can be noted that there are no sensory & muscular requirements listed. This is because the maintainer in this example is not required to develop muscle memory for this maintenance procedure. If a developer were to instead implement a training simulation for this scenario with an increased focus on sensory & muscular fidelity (e.g., haptics, temperature, kinematics) without accurate representation of cognitive requirements, not only could the cost of the simulation increase exponentially, but the resultant system would leave trainees unable to apply their training in a high-stress environment where it is needed most. Once ideal technology requirements have been identified, the next step is to determine which technology best suits the requirements and constraints of the training (see Figure 1).

Step 4: Optimizing Technology Selection

VR as a full training platform may not be beneficial, but VR technology could target specific types of training (Abich et al., 2021). Meta-analytic reviews of VR training programs have demonstrated task-technology fit as a strong moderator of VR training effectiveness (Howard et al., 2021). Training requirements can be determined at the task level. For example, camera-guided surgery, in which a surgeon's real-world task is to watch a computer monitor, affords a computer-based training system. On the other hand, military enemy identification and firearm skills tasks could benefit from being trained in a cave automatic virtual environment (CAVE) simulator that can more appropriately map to those relevant real-world interactions (Howard et al., 2021). Once the LOs and perceptual requirements for training tasks have been determined, the level of fidelity necessary to train may or may not require VR capabilities

The benefits of a lower-fidelity task-technology fit can be seen in the aviation industry's implementation of personal computer aviation training devices (PCATDs). PCATDs have demonstrated beneficial training transfer effectiveness ratios validating it as a substitute environment when compared to live training (Taylor et al., 2004). In this instance, PCATDs are used to train some of the crucial skills but not all. If there is a clear split between LOs that afford computer-based training and others that afford VR, then it may be beneficial to build out training across both platforms with the aspects that require VR training as the only parts within VR. If training cognitive processes, lower fidelity VR training has demonstrated to be more effective than highly immersive VR due to additional interaction challenges and cyber sickness (Woon et al, 2021). Now, as VR training enters the pilot training domain, additional considerations include the cost of VR, VR sickness, perceptual distortion, and potentially interaction mismatches that would occur if using generic VR controllers compared to flight controls. Additionally, for maintenance tasks that require a level of understanding related to the torque required, it may be crucial to continue to conduct the training on the actual system to some degree. When we consider the specific task which we are aiming to train then we can determine the task-technology fit. By putting LOs first, we can more accurately determine the fit of the technology.

If VR affordances make it a favorable option for training, one must then consider the *type* of VR to utilize. A computer-based VR headset requires connection to a VR-capable computer, whereas a standalone VR headset (also known as "wireless" headsets) can be utilized independent of a computer. At this stage of the training development, one will want to consider the constraints of the training location and available or targeted duration for training, and budgetary constraints for equipment. Each solution has its benefits and limitations. Computer-based VR, being connected to a PC, often has access to more computing power than standalone VR resulting in more freedom for higher fidelity and demanding simulation environments. Computer-based VR can include base stations (also known as "lighthouses") that allow for more accurate tracking with low latency compared to the inside-out tracking system in a standalone headset. Systems with base stations are ideal for VR tasks that require the user to make finite movements. Base station tracking also requires a larger physical space to set the tracking system up and can generally cost more than the average standalone solution.

A standalone headset offers lower cost and more freedom to be utilized anywhere, allowing training to be completed indoors, outdoors, and during operations. Limitations do exist relative to battery life, limiting the length of training sessions; however, newer solutions for standalone headsets are beginning to incorporate swappable batteries, accommodating longer training times. Budget, time, and location constraints are factors that can be evaluated once the training needs are identified to ensure the technology has the necessary features without overinvesting in capabilities that exceed the training needs. After technologies have been selected, an instructional designer can work to develop content and assessments that are best suited to the selected technology in the next step (see Figure 1).

Step 5: Developing Learning Content and Assessments

Once technology is selected, the instructional designer can begin to develop training content and assessments that are ideal for that technology. Similar to the first step of this process, leveraging educational frameworks such as Bloom's taxonomy can help guide the content development. If the LOs are determined to be of the *Analysis* level of Bloom's taxonomy, then we may expect the trainee to be able to distinguish between the components of very similar procedures. In a traditional training setting, analysis may include being able to write about each process, correctly mapping word-bank options to the similar procedures or knowing when to enact one procedure over another. In a VR training setting, the trainee may be able to also *Apply* the procedures (e.g., treating a wildfire versus a gas fire), immediately reinforcing or correcting their previously acquired knowledge. Not all procedures and their tasks require a full, high-fidelity environment to be practiced; instructional designers should consider the individual necessary cues driving the learning process per LO to determine the degree to which training content should be included in VR. Offering content in lower fidelity training should be accepted and encouraged as it increases the efficiency and cost of training resources so that we can apply higher fidelity resources to the skills that necessitate those systems.

Assessments will then need to be selected based on the training content. Some approaches include system-based assessment (that is, the machine performing assessments without instructor in-the-loop), instructor assessments (e.g., grade sheets), or self-report assessments. For training components being translated into VR, existing assessments embedded in the traditional training content may or may not fit into the VR training effectively. VR training systems have the ability to use many unobtrusive metrics for knowledge, procedure, accuracy, or speed. VR simulations have been able to discriminate between experienced and inexperienced users in operating room tasks (Haque & Srinivasan, 2006). However, interactive assessments can greatly increase technical costs and should be evaluated by potential value gained (e.g., if assessing knowledge, utilizing multiple choice questions with a reference image as opposed to selecting directly on the image can greatly reduce development costs). Meta-analytic reviews suggest that VR training programs may be better at measuring and evaluating behaviors as opposed to learning gains (Howard et al., 2021). Therefore, it may be beneficial to invest interactive assessment centered around behavior-related skills as opposed to knowledge assessment. Utilizing original resources such as CFETPs can allow one to determine if the LO goals require the trainee to demonstrate knowledge or behaviors. Additionally, by incorporating system-based measures or linking instructor-based ratings, researchers may be able to address the gap of VR training effectiveness in operations by storing and assessing effectiveness over time. The data provided by VR training can facilitate training effectiveness studies in the background.

The instructional designer will also want to consider in what way the VR training best contributes to the overall training. The needs, in terms of content and assessment, will depend upon whether the expectation for the training is to *supplement* or *replace* traditional training methods. If the VR training is meant to supplement traditional training content, providing an interactive scenario with knowledge checks throughout may be sufficient. In the medical domain VR training has been shown to be effective when self-guided for procedural learning particularly for short durations (Woon et al., 2021). If the VR training is meant to replace traditional training content and assessment, one will want to ensure that the assessment methods truly capture higher level competencies or task proficiency and not VR savviness. For example, developing interactive scenarios that require the trainee to pick up and assemble objects within an allotted time may only assess the trainee's efficiency at using VR controls to complete a task. The assessments incorporated within training-in-VR are of great concern, especially as VR continues to enter domains that aim to utilize it as a certification platform. If VR does not afford the method of assessment needed to reliably certify trainees, a re-evaluation of technology at step 4 may be needed.

Each piece of training content and assessment selected should be weighed against the allotted training time and development budget available. This approach can allow development time to be prioritized for key components of the training scenario (e.g., spending development time on realistic interaction for operating a fire extinguisher, using lower cost development resources such as a textbox for explaining the chemistry of a fire extinguisher).

THE PROCESS IN PRACTICE

Each step of this process produces knowledge and data justifying development decisions that ensure training content is linked to the LOs. In many other processes, unstructured notes attempt to capture this data. It is therefore difficult to query this knowledge thus, limiting the ability to revise training as requirements evolve. For example, new VR technology with lower cost, longer battery life, or higher fidelity might revise technology selection. LOs will change over time, too. CFETPs undergo updates every few years, and training must be reevaluated and updated accordingly. Additionally, budgets and organizational priorities may change. Data supporting each step of the process must be in a well-structured accessible format; otherwise, analyses may be needlessly repeated and insights from prior work may be lost. Organizations may waste resources creating training rather than delivering the most effective and efficient training. To combat this, each step of the process must have structured, machine-readable artifacts and supporting tools capturing the knowledge necessary to sustain the process. Structured knowledge also creates opportunities for future machine-aided decision making about training development, for machine generation of training content, and for precision learning leveraging the linkages from assessments to underlying LOs. For example, our implementation of the process outlined in this paper expedites VR content authoring with automatic generation of dialog and assessment objects from training scripts. If designed well, learning strategies and approaches could be transitioned to standardized VR learning templates to expedite development across training lessons or courses. Defining the VR standard learning methods could lead to rapid, extensible, and familiar methods for learning—much like traditional classrooms utilize such as PowerPoint lectures and Kahoot! With proper knowledge and data structuring, content such as traditional rubrics can be fed into decision support tools that define the perceptual and resulting fidelity requirements. Technology analysis of alternative tools can then be used to determine if one technology best supports the LOs. Then automatic generation suites can transform content and assessments into VR courses which can be repeatable and extensible across lessons or domains.

Figure 1 demonstrates the flow of the iterative process outlined in this paper. The process should be applied for new training programs and when existing training programs update CFETPs or aim to improve the training program based on current ineffective methods. Technologies like VR are exciting. It is easy for instructors or other staff to seek out VR without first assessing the training value. Popular press uses phrases like “fear of missing out”, “shiny object syndrome”, and “buzzword compliance” to express the tendency for people to pursue something new and exciting primarily because it is new and exciting. A systematic process with a well-defined workflow encourages decision makers to think critically and use data to make better decisions for his or her organization.

CONCLUSION

As we look to use VR in more training domains, ensuring the implementation goals are aligned with the LOs will be fundamental for successful integration. We must start to recognize that VR training is not one size fits all. First, a real dissection of the learning content and mapping of its objectives to the strengths of VR are necessary to ensure accurate and effective training in VR. Second, instructional designers must identify what sensory and perceptual information would be necessary to achieve the LOs. From there, the perceptual requirements can be transformed into fidelity requirements of the system and technology that include those fidelity requirements can be considered. Based upon technology costs, training time limitations, location constraints, and other logistical concerns can be weighed against one another to select the technology that allows a goldilocks fit. Finally, based upon the affordances of the technology, the LOs, and potential methods for assessing performance, the content and assessments of the learning technology can be developed. This approach also allows for the design of tools that facilitate each stage of the process—like streamlined integration of existing training materials. As our training technologies advance into more complex and multifaceted environments, tools to help SMEs extract LOs and facilitate the entire process will also be needed. Programs across military, education, and commercial applications are looking to bring modern and novel technologies into learning and training to improve the effectiveness. However, LOs as the foundation for training design will allow the design of training systems to support the trainee fully. Utilizing the process outlined within this paper can ensure that VR continues to add value to trainee proficiency and training effectiveness.

ACKNOWLEDGEMENTS

The authors would like to thank the 2G Logistics Planners instructor community for providing invaluable domain expertise and training content during this research endeavor. Additionally, the authors would like to acknowledge the infrastructure and support of Dr. Jesse Johnson and the AETC A9/Det 23 team, as well as their overall vision for pairing training and technology.

REFERENCES

Abich, J., Parker, J., Murphy, J. S., Eudy, M. (2021). A review of the evidence for training effectiveness with virtual reality technology. *Virtual Reality, 25.* 919-933. https://doi.org/10.1007/s10055-020-00498-8

Abrahamson, E. (1999). Managerial fads and fashions: The diffusion and rejection of innovations. *The Academy of Management Review, 16*(3). 586-612. https://doi.org/10.2307/258919

Bloom, B. S. (1956). Taxonomy of educational objectives: The classification of educational goals. New York, NY: Longmans, Green.

Department of the Air Force. (2019). AFSC 2A7XS Low observable aircraft structural maintenance career field education and training plan. *Headquarters US Air Force.*

Fletcher, J. D., Belanich, J., Moses, F., Fehr, A., Moss, J. (2017). Effectiveness of augmented reality & augmented virtuality. In Syme, G., Hatton MacDonald, D., Fulton, B. and Piantadosi, J. (eds) MODSIM2017, 22nd International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2017, pp. 894–900. ISBN: 978-0-9872143-7-9. www.modsimworld.org/papers/2017/Effectiveness_of_AR_and_VR.pdf

Haque, S., & Srinivasan, S. (2021). A meta-analysis of the training effectiveness of virtual reality surgical simulators. *IEEE Transactions on Information Technology in Biomedicine, 10*(1). 51-58.

Horey, J., Fallesen, J., Morath, R., Cronin, B., Cassella, R., Franks, W., & Smith, J. (2004). Competency based future leadership requirements. *United States Army Research Institute for the Behavioral and Social Sciences.*

Howard, M. C. (2021). A meta-analysis of virtual reality training programs. *Computers in Human Behavior, 121. 1-16. https://doi.org/10.1016/j.chb.2021.106808*

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies, 23*. 1515-1529. https://doi.org/10.1007/s10639-017-9676-0

May, H., Cheng, Z., Tan, J., & Lou, Y. (2021). Competency development for military nursing: Developing the competency framework for global health. *Research Square.* https://doi.org/10.21203/rs.3.rs-468341/v2

Sacks, R., Perlman, A., & Barak, R. (2013). Construction safety training using immersive virtual reality. *Construction Management and Economics, 31*(9). 1005-1017. https://doi.org/10.1080/01446193.2013.828844

Stacy, W., Walwanis, M., Wiggins, S., & Bolton, A. (2013). Layered fidelity: An approach to characterizing training environments. In *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), Orlando, FL.*

Taylor, H. L., Talleur, D. A., Rantanen, E. M., & Emanuel, T. W. (2004). The effectiveness of a personal computer aviation training device, a flight training device, and an airplane in conducting instrument proficiency checks. *Aviation Human Factors Division Institute of Aviation.*

Trabysh, V., Feldhacker, T. J., Rizzardo, C., Johnson, J., Tripp, L., Halverson, K., & Gonzalez, A. (2021, July). I Feel the Need for Speed: Empirical Evidence of the Effectiveness of VR Training Technology on Knowledge and Skill Acquisition. In *International Conference on Applied Human Factors and Ergonomics* (pp. 19-26). Springer, Cham.

Wolters, H. M., O'Shea, P. G., Ford, L. A., Fleisher, M. S., Adeniyi, M. A., Conzelman, C. E., & Webster, R. J. (2014). Identifying and training brigade command competencies. *Military Psychology, 26*(4), 278-291. 10.1037/mil0000052

Woon, A. P. N, Mok, W. Q., Chieng, Y. J. S., Zhang, H. M., Ramos, P., Mustadi, H. B., & Lau, Y. (2021). Effectiveness of virtual reality training in improving knowledge among nursing students: A systematic review, a meta-analysis and meta-regression. *Nurse Education Today, 98.* https://doi.org/10.1016/j.nedt.2020.104655