Check for updates

Transfer of Training from Virtual Reality and Augmented Reality: A Meta-Analysis Extended Abstract

Alexandra D. Kaplan[1], Jessica Cruit[1], Mica Endsley[2], Suzanne M. Beers[3], Ben D. Sawyer[1] and P.A. Hancock[1]
[1]University of Central Florida, Orlando, FL; [2]SA Technologies, Gold Canyon, Arizona; [3]MITRE Corporation, Colorado Springs, Colorado

Introduction

Training is required for humans to develop or further their skills. Many vital skills necessitate strict learning protocols and may be expensive and time consuming. So, any advancement in technology or science which might reduce the cost, either financial or temporal, will be of use to many people. For this reason, simulated training has gained acceptance as a means to increase training efficiency. Such training would include any technology utilizing mixed reality (MR), including both augmented reality (AR) and virtual reality (VR). These technologies can reduce some of the costs associated with training and can eliminate many risks by placing individuals in a simulation rather than in real-world dangerous situations. However, training in MR does not solve all the problems that plague current training methods. While simulation may be the solution to many issues, it comes with its own set of caveats.

One must consider whether mixed reality is in fact an effective medium for training. The most obvious complaint is the applicability of the training to execution of the task in the real world. The principle of encoding specificity indicates that when the learning environment is sufficiently different from the environment where that learning is measured, performance tends to suffer (Tulving & Thomson, 1973). This principle was further explored by Godden and Baddeley (1975) who found that scuba divers who memorized lists of words on dry land recalled those lists better above-rather than below-the surface of the ocean. This parallels many training situations and calls into question whether learning properly transfers between scenarios. One can extend the same caution to training in MR. The situations in which simulation-based training has the most benefit (risky, expensive, or unsafe conditions) also have the highest cost of failure. For these reasons, it is imperative that transfer of training from virtual reality be measured to determine whether the transfer is effective overall. A meta-analysis of the current empirical literature on the subject is a good way to begin to understand the effect of virtual training on performance.

Method

A literature search was conducted to find all published, peer-reviewed articles on the topic of training transfer from simulation- based training. Search strings included two terms; the primary term having to do with the vehicle of the virtual training and the second having to do with the training transfer outcome. Primary search terms used included “virtual reality,” “VR,” “augmented reality,” “mixed reality,” and “simulation.” Secondary terms included “encoding specificity,” “learning,” and “training.” Articles were accepted for inclusion if they examined training transfer from AR, VR, or MR, or real-world performance after virtual training, as a dependent variable. Articles were also required to meet the following criteria: 1) The article was from a peer-reviewed journal, a conference proceeding, a technical report or a dissertation, 2) The article contained original empirical data, 3) the study used virtual, augmented, or mixed reality training and real-world performance, and, 4) the study’s statistics included enough information to determine an effect size.

In all cases, the dependent variable was a performance measure after a simulated training protocol. The associated variables included age of the participants, difficulty of the task, and level of virtual immersiveness, which is here defined as the extent to which the virtual world is fully-realized. In these cases, virtual reality is more virtually immersive than augmented reality, and both are more immersive than an interactive video or a real-world control setting. A subset of the studies where immersiveness was a factor, compared training in a fully immersive virtual reality setting to training in a non-virtual control setting.

Results

Results, shown in Figure 1, indicate that virtual immersion, both overall and in the VR vs Control Group analysis, had no significant effect on post-training performance, since the associated confidence interval included zero. Additionally, task difficulty and age of participants did not have strong effects on the results of simulation-based training, but too few articles examined these topics for a confidence interval to be established.

Copyright 2019 by Human Factors and Ergonomics Society. DOI 10.1177/1071181319631207

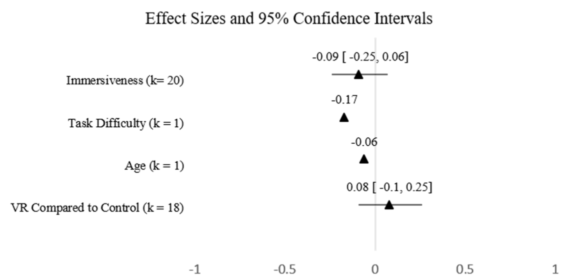

Figure 1: Calculated Effect Sizes

Discussion

The fact that the zero could not be excluded from any of the present confidence intervals appears to indicate that we see equivalence between VR/AR training and traditional instructional techniques. If we are optimistic about these results, we can note that the use of virtual, mixed, and augmented realities provides at least an equivalent experience to what is normally used for instruction. However, simulation has been held out to offer superior training capacities, and so the case for this at present is at best "not proven."

One hopeful aspect of the present findings is that the safety and cost benefits of VR/AR remain, even as the training which accrues does not differ from other, real-world exposures. Indeed, if safety and affordability are important motivational forces, as they often are in training, then VR/AR provides this advantage and at least equivalent transfer levels. However, it is important for the authors here to note that while overall the difference between training mechanisms was null, in some cases, participants performed worse following virtual-reality training. The difference in tasks may explain the disparate effects of simulation. For example, one study utilized virtual reality to teach balance to stroke patients (Lee, Kim & Lee, 2015). Results of this particular study seemed to indicate that for a large number of participants, performance had declined in the follow-up assessment. On the other hand, a study where participants learned math showed that scores were consistently higher after the training intervention (Bier, Ouellett & Belleville, 2018). While each study had its own results, and such results were varied enough to prevent any deep investigation of sub-moderators, it did appear that the majority of the cases where people performed poorly after the training, involved some sort of physical activity.

While it may appear that the results of this meta-analysis are inconclusive, they actually highlight several important lacunae and disparities in the extant literature that should be addressed if any meaningful progress is to be secured from future examinations (meta or otherwise). The first concerns individual differences; soldiers, for example, are a very different population from elderly stroke victims. Studies included in this analysis reported on both of these populations and the intrinsic assumption of homogeneity is obviously in threat of being fractured by such sampling variations. The second vital question concerns the technology that is being employed. Clearly, one example here can be that a wide range in quality of virtual reality headsets was employed. Thus, there may be questions concerning the *physical ergonomics* of such systems; do they fit? Are they comfortable? How long can they be worn? These questions emerge independent of other technical issues such as display fidelity and display size (Hancock, Sawyer, & Stafford, 2015). Although many of the reported methodology sections do not provide sufficient evidence, we suspect that the simulation environments in the reported studies were likely of very different quality overall. The final area of discrepancy that we point out here is the differing tasks involved. For example, the varying nature of the complexity of the task appears to interact with the precise specification of the training protocol (Wulf & Shea, 2002). Thus, there will be occasions in which extrapolation from the assimilation of simple laboratory-confined skills will not extend to that of complex, real-world skills. In fact, this may even lead to negative transfer; an eventuality to be avoided for any practically-oriented agency.

References:

*contact first author for list of articles used in analysis

Bier, B., Ouellet, É., & Belleville, S. (2018). Computerized attentional training and transfer with virtual reality: Effect of age and training type. *Neuropsychology, 32 (5),* 597.

Godden, D.R., & Baddeley, A.D. (1975). Context□ dependent memory in two natural environments: On land and underwater. *British Journal of Psychology,* 66 (3), 325-331.

Hancock, P.A., Sawyer, B.D., & Stafford, S. (2015). The effects of display size on performance. *Ergonomic*s, *58* (3), 337-354.

Lee, H.Y., Kim, Y.L., & Lee, S.M. (2015). Effects of virtual reality-based training and task-oriented training on balance performance in stroke patients. *Journal of Physical Therapy Science, 27 (6),* 1883-1888.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80 (5), 352-37

Wulf, G., & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin & Review*, *9* (2), 185-211.