A Framework for the Testing and Validation of Simulated Environments in Experimentation and Training

David J. Harris[1]*, Jonathan M. Bird[2], Philip A. Smart[2], Mark R. Wilson[1] and Samuel J. Vine[1]

[1] *School of Sport and Health Sciences, University of Exeter, Exeter, United Kingdom,* [2] *Centre for Simulation, Analytics and Modelling, University of Exeter Business School, Exeter, United Kingdom*

New computer technologies, like virtual reality (VR), have created opportunities to study human behavior and train skills in novel ways. VR holds significant promise for maximizing the efficiency and effectiveness of skill learning in a variety of settings (e.g., sport, medicine, safety-critical industries) through immersive learning and augmentation of existing training methods. In many cases the adoption of VR for training has, however, preceded rigorous testing and validation of the simulation tool. In order for VR to be implemented successfully for both training and psychological experimentation it is necessary to first establish whether the simulation captures fundamental features of the real task and environment, and elicits realistic behaviors. Unfortunately evaluation of VR environments too often confuses presentation and function, and relies on superficial visual features that are not the key determinants of successful training outcomes. Therefore evidence-based methods of establishing the fidelity and validity of VR environments are required. To this end, we outline a taxonomy of the subtypes of fidelity and validity, and propose a variety of practical methods for testing and validating VR training simulations. Ultimately, a successful VR environment is one that enables transfer of learning to the real-world. We propose that key elements of psychological, affective and ergonomic fidelity, are the real determinants of successful transfer. By adopting an evidence-based approach to VR simulation design and testing it is possible to develop valid environments that allow the potential of VR training to be maximized.

Keywords: fidelity, presence, training, transfer, validity, virtual reality

INTRODUCTION

How real is virtual reality? This question raises weighty metaphysical issues, but it also poses a very practical challenge for scientists seeking to use virtual reality (VR) technologies for experimentation and training. A simulation aims to reproduce some aspects of a task (e.g., perceptual information and behavioral constraints) without reproducing others (e.g., danger and cost; Gray, 2002; Stoffregen et al., 2003). Consequently, understanding the degree of concordance between the simulated environment and the corresponding real-world task is essential for the successful application of VR, in both the psychology lab and in the field. While this is a challenging endeavor on its own, there is also considerable confusion within cognitive science about terms like *fidelity, validity, immersion* and *presence* and how VR environments should be evaluated. For

example, environments are often judged to be "high fidelity" if they provide a detailed, realistic visual scene, but the superficial appearance may have little relationship with functionality, especially in the context of VR for education and training (Carruth, 2017). In effect, the distinction between presentation and function is often overlooked.

For VR to be implemented more effectively as a training tool, greater conceptual clarity is imperative, and more rigorous ways of testing and validating simulations must be developed. In this review, we aim to address some of this confusion by, firstly, addressing some conceptual issues and outlining a taxonomy of fidelity and validity, and secondly, by proposing evidence-based methods for establishing fidelity and validity during simulation design. Here we particularly focus on VR for training perceptual-motor skills, such as for applications to sport, surgery, rehabilitation and the military – the kind of active skills which may be particularly suited to VR training (Jensen and Konradsen, 2018). However, these principles may also apply to many uses of VR as a training tool, and as such, we hope to provide a framework for those seeking to develop more effective, evidence-based VR simulations.

IMMERSION AND PRESENCE

Immersive VR is an alternate world composed of computer-generated sounds and images with which the user can interact using their sensorimotor abilities (Burdea and Coiffet, 2003; Slater and Sanchez-Vives, 2016). The proliferation of technologies for both using and creating augmented reality (AR), mixed reality (MR), and VR experiences has led to rapid adoption of VR as a training tool within human factors (Gavish et al., 2015), sport (Bird, 2019), rehabilitation (Levac et al., 2019), and surgery (Hashimoto et al., 2018). The lure of new technologies for training is sufficiently great that these methods have been applied before there is a foundational understanding of how to optimally implement VR training (albeit with some successes, e.g., Calatayud et al., 2010; Gray, 2017). Particular issues yet to be addressed include: the determinants of effective transfer of training (Rose et al., 2000; Rosalie and Müller, 2012); the requisite levels of fidelity and validity and how to test them (Gray, 2019); and the effect of VR on basic cognitive and perceptual processes (Harris et al., 2019d).

As a means of reducing the aforementioned confusion surrounding key terminology, we adopt Slater and Sanchez-Vives (2016) definition of *immersion* as the technical capability of a system that allows a user to perceive the virtual environment through natural sensorimotor contingencies. While immersion is an objective feature of the input provided to the user, the subjective experience that is created of actually being inside the virtual environment is termed *presence* (Baños et al., 2004; Bowman and McMahan, 2007). Slater (2009) and Slater and Sanchez-Vives (2016) suggest that there are two important components to the experience of presence; *place illusion* (the illusion of "being there" in the virtual environment), and *plausibility* (that the depicted scenario is really occurring). A consequence of place illusion and plausibility is that users behave in VR as they would do in similar circumstances in reality (Slater and Sanchez-Vives, 2016), which is of paramount importance for VR training. Despite users knowing that the virtual environment is fictitious (Stoffregen et al., 2003), researchers have suggested that presence can prompt users to feel anxious near illusory drops (Meehan et al., 2002), maintain social norms with virtual others (Sanz et al., 2015), and exhibit stress when forced to cause harm to avatars (Slater et al., 2006).

For training and experimentation purposes the virtual environment needs to be only as "real" as is required for achieving the desired learning outcome, be that training perceptual-motor skills (Tirp et al., 2015), habituating to stress inducing stimuli (Botella et al., 2015) or studying sensorimotor processes (Buckingham, 2019). However, differing target populations may need to be engaged in different ways to produce similar learning outcomes. For VR to be effective in a training context, there must be a correspondence between key elements of the real and virtual tasks that are functional for task learning. Other elements such as graphical realism are often inconsequential in comparison (Dahlstrom et al., 2009). For instance, when using VR to study the perceptual information that informs catching a ball, there is no requirement that the ball looks realistic, the scene is highly detailed, or the task is particularly immersive (see e.g., Fink et al., 2009; Zaal and Bootsma, 2011). Nonetheless, realistic kinematic and depth information pertaining to the ball are necessities. It is elements of the simulation such as these that determine the *fidelity* and *validity* of virtual environments. In the remainder of the article we outline various types of fidelity and validity, how they can be assessed, and how they contribute to effective transfer of training (see **Table 1**).

TRANSFER OF TRAINING

The capacity to effectively apply and adapt learning in the face of constant environmental variation (i.e., transfer) is fundamental to many human activities (Rosalie and Müller, 2012). Transfer of training occurs when prior experiences in a particular context can be adapted to similar or dissimilar contexts (Barnett and Ceci, 2002). Ultimately, the test of a successful VR training simulation is the degree to which skills learned in the virtual environment can be applied to the real-world. Classical theories of learning, like Thorndike's (1906) identical elements theory (later developed into Singley and Anderson's (1989) identical productions model) support the notion that successful transfer is contingent on the coincidence of stimulus or response elements in learning and transfer contexts[1], suggesting that only near transfer may be possible. For instance, accurate size estimation of geometric shapes is dependent on specific learning with objects of similar size and shape (Woodworth and Thorndike, 1901).

The foremost competing paradigm to similarity-based transfer is principle-based transfer theory (Judd, 1908), which focuses on the coherence of principles, rules, or laws between settings, irrespective of superficial contextual variation. This approach

---

[1]Singley and Anderson (1989) describe the key component for transfer as similarity of the production rule, which specifies the action that must be produced in a certain situation, taking the form of an IF-THEN rule.

proposes that learning can more easily be generalized, provided equivalent principles or rules are present in learning and transfer contexts. Achieving far transfer of learning that is generalized across domains which are only loosely related to each other is, however, notoriously difficult to achieve (Sala and Gobet, 2017). Nonetheless, VR training generally does not aim to achieve domain general improvements. Instead, as VR aims to recreate the performance environment, near transfer between tightly coupled domains is the goal, such as is common across human learning. The challenge facing the field of VR training is to establish whether a simulation is realistic enough, and how to enhance the aspects of realism that really matter for effective transfer. We believe that a better understanding of fidelity and validity in the design and testing of VR environments can help to meet this challenge.

VALIDITY

In a general sense, validity is the extent to which a test, model, measurement, simulation, or other reproduction provides an accurate representation of its real equivalent. For example, a valid measurement truly represents the underlying phenomenon it claims to measure. Similarly, a valid simulation is one that is an accurate representation of the target task, within the context of the learning objectives and the target population. This is not to say that the simulation is *the same*. A simulation aims to capture key features of the real task and environment, rather than exactly emulate or imitate it. A number of types of validity are discussed in relation to measurement methods in Psychology (e.g., criterion and concurrent validity) but there are two primary types to consider in simulation design; *face validity* and *construct validity* (see **Table 1**).

Face Validity

Face validity is the subjective view users have of how realistic a simulation is. Accordingly, face validity may be an important contributor toward perceptions of plausibility (Slater, 2009). Face validity is often highly dependent on the superficial visual features of the simulation (see section "Physical Fidelity" below) but is also influenced by structural and functional aspects, such

as how user input relates to actions. Consequently, the design of the simulation and the technical capabilities of the system (i.e., immersion) are important determinants of face validity. Within a learning context, face validity is important in one sense, and irrelevant in another. It can be important because it correlates with up-take and is often needed to achieve buy-in, which can derail training if not achieved. Conversely, it is also irrelevant because it likely has no correlation with actual learning. A simulation can have face validity and be useless; a simulation can have no face validity at all and be an excellent training tool (Dahlstrom et al., 2009). As discussed, theories of transfer propose that a coincidence of stimulus and response elements, or underlying principles between the practice and target tasks is required for transfer of learning. Face validity is unlikely to be a good indicator of whether any of these conditions are met.

Assessment of face validity often relies on participant reports and verbal feedback about whether the simulation is a good representation of the real task, either formally or informally (e.g., Bright et al., 2012). Collecting participant feedback regarding face validity is a commonly used approach in the field of surgical simulation. In this context, the opinions of expert surgeons are often sought about how the simulation looks and feels and is an important part of simulation validation (Sankaranarayanan et al., 2016; Roberts et al., 2017). A similar approach is also used in the development and testing of aircraft simulations with expert pilots (Perfect et al., 2014). In many other contexts face validity is not explicitly tested, but remains an implicit factor in simulation adoption. Consequently, face validity may well be a hurdle to overcome, but not a major contributor to training success.

Construct Validity

Construct validity exists in a more objective sense than face validity, and is the extent to which the simulation provides an accurate representation of the real task. As a result, it is crucial for achieving transfer of learning. Many simulations are used to track learning, or to index proficiency on a task, which depends on some level of functional similarity between the simulation and the real task. A simulation with good construct validity should be sensitive to variation in performance between individuals

TABLE 1 | Summary of validity and fidelity terminology.

Term	Description	How to test
Face validity	Does the simulation look and feel realistic?	Self-reports from users concerning plausibility
Construct validity	Does the simulation provide an accurate representation of real task performance?	Ability of the simulation to distinguish real-world experts from novices and track improvements
Physical fidelity	Is there a high degree of detail and realism in the physical elements of the simulation?	Participant reports of realism and measures of presence (both self-report and psychophysiology)
Psychological fidelity	Does the simulation accurately represent the perceptual and cognitive features of the real task?	Measurement and comparison of mental effort, gaze behavior, neural activity etc., between real and virtual tasks
Affective/emotional fidelity	Does the simulation elicit emotional responses (e.g., stress or fear) in a similar way to the real task?	Self-reported experiences of users or online monitoring of psychophysiological indices of affect
Ergonomic/biomechanical fidelity	Does the simulation elicit realistic motor movements?	Assessing the realism of VR movement parameters through motion tracking, and comparing amplitude, speed, inter-joint coordination etc., with real actions

(e.g., real-world novices and experts) and within individuals (e.g., learners developing over time), as this would indicate a coherence of principles, rules or stimulus and response elements between the real and VR task. These fundamental similarities are likely required for transfer of training.

Predictive validity refers to the ability to reliably predict future performance outcomes and is closely related to construct validity. If the simulated task replicates some core aspect of the real skill (i.e., good construct validity), then simulation performance is more likely to accurately predict future real-world performance[2]. One possible application of VR simulations is to serve as a tool for recruitment or selection (e.g., Moglia et al., 2014). Evaluating the aptitude of trainees in VR is particularly appealing when assessment on the real task would be impractical or dangerous, such as in military or surgical settings. For such purposes, predictive validity must be established to ensure that selection decisions based on simulation performance are reliable.

There are clear opportunities for testing construct validity through expert versus novice comparisons and sensitivity to practice-induced improvements. This validation method has previously been adopted by Harris et al. (2019a) when comparing putting performance of novice and elite-level golfers in a VR golf simulation, and by Bright et al. (2012) when comparing tissue resected when using a minimally invasive surgical simulation. Not only did the real-world experts resect more tissue than the novices in the simulation, but the simulation was sensitive to practice induced improvements in the novices (Bright et al., 2012, 2014). Measuring inter- and intra-individual variation is an effective method for demonstrating construct validity, but to achieve it in the first place requires an understanding of the competencies and skills of the to-be-trained task, and ensuring that the rules, interactions and criteria of the real-task have one-to-one counterparts in the simulation.

FIDELITY

Fidelity is the extent to which a simulation recreates the real-world system, in terms of its appearance but also the affective states, cognitions or behaviors it elicits from its users (Perfect et al., 2014; Gray, 2019). To achieve construct validity (and effective transfer) it is necessary to ensure there is a sufficient level of fidelity in relevant aspects of the simulation. For instance, when implementing VR as a training tool for perceptual-motor skills it may be important to assess fidelity for eliciting affective states like stress, for directing attention to relevant information, and allowing for movements representative of the real skill. However, the fidelity of the simulation must be assessed in relation to training goals. Gray (2002) propose that a simulation can only be considered "high-fidelity" in relation to the research question being asked, and the same holds true for training. A simulation that is developed to train a motor skill (e.g., a golf swing) would be required to elicit realistic actions, but realistic ergonomics might be irrelevant for a simulation that is developed to enhance proficiency of a purely cognitive task.

---

[2]Similarly, if a test has good construct validity it is likely to show *test-retest reliability*, because it is accurately measuring aspects of the real skill.

Physical Fidelity

Physical fidelity refers to the level of realism provided by the physical elements of the simulation; primarily visual information (including field of view) as the principal sensory modality in VR, but also realistic behavior of objects, adherence to the normal laws of physics, and level of functionality. As is the case for face validity, the physical fidelity of the environment is likely to be important for eliciting a feeling of presence in the participant, and in particular creating the illusion of plausibility (Slater, 2009), which will depend heavily on the immersion of the technology. If basic elements of physical fidelity are low, such as allowing the user to walk through walls, the illusion will be broken. When the term "high fidelity" is used in relation to simulations, it is generally in reference to high physical detail, but as we outline below, realistic sights and sounds are just one element of a high-fidelity simulation.

While high graphical realism is likely to increase motivation to engage with simulation training and adds to the "wow factor" of VR (a highly positive, but superficial response), it is unlikely to be what creates effective transfer. For instance, in the context of sport, high graphical realism would seem to be important, but sporting performance is dependent on the efficient use of only a subset of the available perceptual information (Davids et al., 2005), making much of the detail irrelevant for training. One instance where physical realism may be important is in using VR to acclimatize performers to a particular environment or for a VR equivalent of mental imagery or visualization (e.g., Sorrentino et al., 2005). Other than adherence to normal physical laws, high fidelity physical features will often not be the stimulus-response correspondences or underlying principles that determine transfer, as outlined in theories of transfer. Nonetheless, very low physical fidelity may still be a barrier to effective training if it lowers motivation to use the simulation.

One method for assessing physical fidelity is through either direct participant reports or measurements of presence (e.g., Harris et al., 2019a). As presence is a result of achieving a sufficient level of fidelity, this is an indirect measure, but high levels of presence would indicate that the physical realism is sufficient to make the simulation believable (i.e., plausibility) and induce the feeling of "being there" (i.e., place illusion). Presence seems to be important for increased engagement in virtual training (Stanney et al., 2003), and can be measured either through self-report (Usoh et al., 2000) or online using psychophysiological indices like eye-movements, electroencephalography (EEG) and heart rate (Jennett et al., 2008; Nacke and Lindley, 2008). Whether higher levels of presence are beneficial for learning beyond the effect of enhanced engagement remains to be established (Dahlstrom et al., 2009; Fowler, 2015; Gray, 2019), and is an important question for future research.

Psychological Fidelity

Psychological fidelity is the degree to which a simulation replicates the perceptual-cognitive demands of the real task (Gray, 2019). For instance, a high-fidelity driving simulation should require the participant to attend to similar areas of the scene (e.g., other cars, pedestrians, street signs) and demand a

similar level of attentiveness and effort as if they were engaged in real-world driving. Accordingly, practice in this simulation should also lead to the development of psychological skills germane to real driving, such as learning to attend to the most informative areas of the road and predict the behavior of other traffic. Developing psychological skills like these would likely support real-world transfer.

Particular considerations for psychological fidelity include determining whether individuals: exhibit similar gaze behavior in real and virtual tasks (e.g., Vine et al., 2014); use similar perceptual information to control their actions (e.g., Bideau et al., 2010); and experience similar levels of cognitive demand (e.g., Harris et al., 2019a). Achieving place illusion will support these aspects of fidelity. For applications to domains like sport, surgery, and the military that place demands on perceptual-cognitive skills, psychological fidelity may be one of the most important prerequisites for developing an effective simulation. Encouragingly, perceptual-cognitive skills have been shown to be transferable between closely related sports (Abernethy et al., 2005; Causer and Ford, 2014), supporting the notion that VR environments with good psychological fidelity should elicit positive transfer. While a number of studies have shown overall performance benefits as a result of VR training, few have directly addressed the development of perceptual-cognitive skills, such as the control of attention or anticipation (Tirp et al., 2015).

There are a number of ways to test psychological fidelity, including comparisons of mental effort, gaze behavior or neural activity between real and VR contexts. For instance, comparisons of gaze behavior between real and simulated surgery have previously been used to validate surgical simulations. For example, Vine et al. (2014) found that, in comparison to the VR task, during the real operation expert surgeons made more frequent, shorter duration fixations indicative of a less efficient visual control strategy. The authors suggested that the additional auditory and visual distractions, as well as the stress of the real operation, were responsible for the differences. Findings such as these highlight how many factors contribute to psychological fidelity in VR.

Evaluations of psychological fidelity can also address the mental and physical demands of the virtual task and compare them to the real task. For instance, Harris et al. (2019c) compared cognitive demands between real and virtual versions of a block stacking game ("Jenga") using a self-report measure of task load, the SIM-TLX. Similarly, Frederiksen et al. (2020) compared cognitive load between head mounted immersive VR and standard screen presentation on a surgical training simulation, finding cognitive load (indexed by secondary task reaction time) to be significantly elevated in the immersive VR condition. Cognitive load is particularly relevant in the context of education and training as an optimal level of load is important for successful learning outcomes (Kirschner, 2002). If VR imposes additional load it could pose a challenge for training. Cognitive load could also be related to elements of user experience, like presence, although this was not the case in the aforementioned study of Harris et al. (2019c). In summary, assessments of psychological fidelity may require a combination of approaches, as well as an understanding of the perceptual-cognitive skills that are responsible for expert performance in the given task.

Affective Fidelity

There has been considerable interest in VR applications for training tasks that are too dangerous to rehearse in the real-world (e.g., critical incidents in heavy industry) or for acclimatizing trainees to the high levels of stress they are likely to face in the field (e.g., defense and security). For these purposes, as well as for applications like treating anxiety disorders, a high level of affective fidelity is required (Moghimi et al., 2016). Affective or emotional fidelity requires the simulation to elicit a realistic emotional response in the user, such as fear, stress or excitement. The success of Virtual Reality Exposure Therapy (VRET), where exposure-based treatments for anxiety disorders are implemented in VR, indicate that realistic emotional responses can be achieved (Krijn et al., 2004; Morina et al., 2015). Similarly, Chirico and Gaggioli (2019) found that VR scenes elicited a range of emotions in a similar manner to the real thing. This realistic affective reaction relies on achieving the illusion of plausibility discussed by Slater (2009) and a sense of presence (Diemer et al., 2015), or no emotional response will occur. Familiarization with the emotion of anxiety can improve subsequent performance when anxious (Saunders et al., 1996). Hence, VR environments capable of eliciting some degree of emotional response may have significant benefits for preparing performers for pressurized environments (Pallavicini et al., 2016).

Affective fidelity can be easily assessed through self-report or psychophysiological measurement during the VR experience and compared to the real task. The clearest implementation of assessing affect in VR is in measuring stress during threatening VR experiences like Slater's Milgram obedience (Slater et al., 2006) and the illusory pit room experiments (Meehan et al., 2002). Online measurements of cardiovascular activity (Ćosić et al., 2010), skin conductance (Meehan et al., 2002) and EEG alpha power (Brouwer et al., 2011) have all been employed as objective, online measures of stress induced by VR experiences. While strong stress responses have been elicited for phobic stimuli in VR, creating stress responses similar to those that will be experienced during high level sport, during complex surgical procedures or by defense and security personnel may be more challenging.

Ergonomic and Biomechanical Fidelity

Ergonomic and biomechanical fidelity is the degree to which the VR environment allows for and promotes realistic movement patterns in the user, making the immersion of the system a major determinant of ergonomic fidelity. Despite advances in VR technology (e.g., haptic gloves, exoskeleton suits, muscle stimulation), the provision of realistic haptic information in VR remains a major challenge (Lopes et al., 2017). Haptic information is important for developing motor control, but if it is unavailable in VR, movements learned or performed in a simulation may differ from those learned in the real-world. Harris et al. (2019d) describe how the combined effect of artificial creation of depth (e.g., vergence-accommodation

conflict; Kramida, 2016) and lack of end-point haptic information (Whitwell et al., 2015; Wijeyaratnam et al., 2019) may combine to force users toward a more deliberate mode of action control that is unlike real perceptual-motor skills. Further work is required to explore these potential limitations to learning and performing actions in VR, but the difference between real tennis shots and those performed in Nintendo Wii tennis provide a stark illustration of how widely action patterns can diverge. Biomechanical fidelity is a particular issue for applications to rehabilitation and sport, where low levels of fidelity could be actively disruptive if suboptimal motor patterns are reinforced in VR.

Assessment of biomechanical fidelity relies on motion tracking, and comparisons of movement amplitude, speed and inter-joint coordination between real and virtual environments. A number of attempts have been made to test the biomechanical and ergonomic features of VR environments, which have generally highlighted the difficulty in achieving this type of fidelity. Bufton et al. (2014) found that compared to real table tennis, participants in three different virtual table tennis games produced larger and faster movements when hitting the virtual ball. Similarly, in a simple reaching and grasping task Magdalon et al. (2011) found reaches were slower with wider grip apertures, and Covaci et al. (2015) found basketball shots made in a virtual environment had lower ball speed, higher height of ball release, and higher basket entry angle, compared with real basketball.

An important action to consider is walking, as VR is being widely adopted for studying and retraining gait (e.g., Young et al., 2011), but requires a VR-linked treadmill to allow users to move any appreciable distance. Work has already demonstrated biomechanical differences between VR and real walking (Mohler et al., 2007; Janeh et al., 2017), but the divergences from conventional movements could be related to the stationary locomotion methods (i.e., treadmills) rather than fundamental issues with traversing in the virtual environment. These differences seem to be a current limitation rather than prohibitive of gait retraining applications, and may well improve with technological advances (i.e., greater immersion).

PRACTICAL ISSUES

In addition to establishing that a simulation is a valid recreation of the target task and is of sufficient fidelity to enable transfer of training, there are a number of practical concerns for effective implementation of VR training which may moderate the effectiveness of the simulation (see **Figure 1**), some of which we outline below (Carruth, 2017).

Pedagogical Viability
An important pedagogical consideration is how design intentions correspond with training requirements. The importance of clear learning outcomes for designing and evaluating training has been well documented in the education literature (Kraiger et al., 1993). Successful training outcomes depend upon being able to articulate those aims for learners and design and create a

simulation that supports them. A subsequent pedagogical issue is whether the simulation can be effectively adopted within a curriculum or training program to provide real learning benefits. Pedagogical viability is most often a pragmatic issue (e.g., *Where is the simulation? How is it used? How much does it cost?*), but can also be a conceptual and theoretical issue, according to your view of learning. The simulation is only a tool, and like any other must be appropriately periodized within the wider curriculum to achieve benefits. Important elements of instructional design should be considered to maximize the impact of simulation based training, such as levels of complexity and specificity in learning objectives, scaffolding of learning and evaluation of training (Kirschner, 2002; Carruth, 2017; Jensen and Konradsen, 2018).
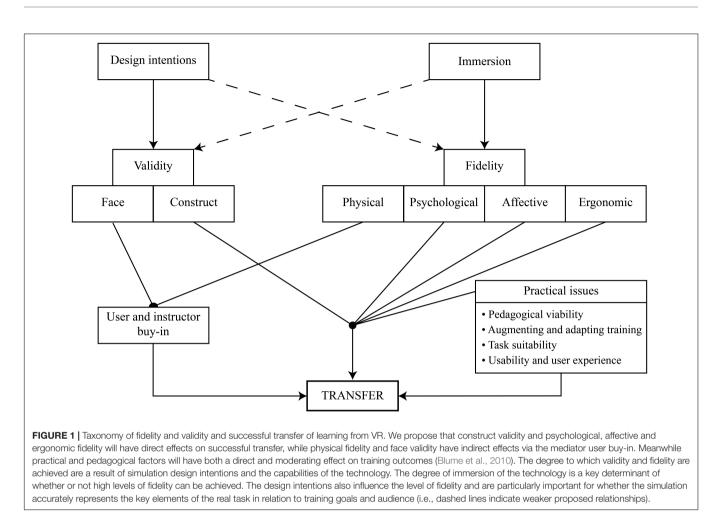
Augmenting and Adapting Training
One of the most compelling reasons to use VR for training is the possibility to augment and improve on existing practices with new methods, rather than just replacing them. Gray (2017) illustrates this well in the case of a baseball batting simulation. Virtual batting practice was found to outperform real batting practice, but only when the virtual version provided task demands constantly matched to the skill of the user. The complete control of the training space afforded by VR allows environmental constraints to be manipulated to improve skill acquisition (Renshaw et al., 2009). Guiding information, such as cues to important information or eye movement patterns of experts, can also be added to speed learning (Janelle et al., 2003; Vine et al., 2013). Other approaches include adaptive VR, which modifies the simulation to suit either the current performance level or psychophysiological state of the user (Moghim et al., 2015; Vaughan et al., 2016). An effective implementation of VR training makes full use of these possibilities.

Task Suitability
The attraction of new technologies for training makes it easy to fall into the trap of overusing them. VR may allow more personalized and more accessible training in many instances, but is unlikely to be the best option when real-world practice is available. Even in the light of rapidly advancing VR technologies, the specificity of training principle highlights that to improve at, say catching a ball, there is one thing above all others that is likely to provide the greatest benefit – just catching a ball. For fine sensorimotor skills in particular, the unusual perceptual information in VR (lack of haptics and conflicting cues to depth) means that VR is not "real" enough to compete with real practice (Harris et al., 2019b). Hence, VR can be useful when the skill could otherwise not be practiced in the real-world (for practical or safety reasons), or when training can be *improved* in VR, but the rationale for using VR over other methods should be clear.

Usability and User Experience
Much like face validity, user experience and usability is not a primary factor in training effectiveness, but may be a barrier to implementation and uptake. An otherwise high-fidelity simulation can be derailed if the user experience is poor. Issues like interaction and navigation in VR pose considerable technological challenges and if the solutions that

FIGURE 1 | Taxonomy of fidelity and validity and successful transfer of learning from VR. We propose that construct validity and psychological, affective and ergonomic fidelity will have direct effects on successful transfer, while physical fidelity and face validity have indirect effects via the mediator user buy-in. Meanwhile practical and pedagogical factors will have both a direct and moderating effect on training outcomes (Blume et al., 2010). The degree to which validity and fidelity are achieved are a result of simulation design intentions and the capabilities of the technology. The degree of immersion of the technology is a key determinant of whether or not high levels of fidelity can be achieved. The design intentions also influence the level of fidelity and are particularly important for whether the simulation accurately represents the key elements of the real task in relation to training goals and audience (i.e., dashed lines indicate weaker proposed relationships).

are implemented make using the VR tool difficult, excessively complicated or unpleasant, then trainees are unlikely to engage with the simulation (Sutcliffe and Kaur, 2000; Gray, 2002). Questionnaires are typically administered as a means of assessing usability and are often bespoke to the simulation (Sutcliffe and Kaur, 2000).

CONCLUSION

In this review we have discussed some of the challenges of validating VR simulations for applications to training. Ultimately, a VR training environment is judged on its ability to elicit positive transfer to the real-world. To achieve this goal, the VR environment needs to be just real enough to develop new skills that can be applied to real tasks. While many realistic behaviors may require the participant to believe they are present in the virtual environment (i.e., place illusion), and that the events are really happening (i.e., plausibility), the perception of presence may not always be the primary consideration when developing effective training simulations[3]. While the concepts of immersion and presence are often

used to determine realism, we have suggested expanding this into a typology of fidelity (see **Figure 1**). Factors such as psychological, affective or ergonomic fidelity are likely to be the more important determinants of effective transfer and are important to evaluate during simulation design. Consequently, researchers are encouraged to address the factors that drive realism in different contexts (e.g., *What is the contribution of each of the constructs in the typology?*) and to explore the extent to which specific markers of fidelity impact upon performance outcomes.

Given the speed of recent technological development it is pertinent to consider what the future of simulated training might look like. One approach that is poised to assume a major role in simulated training is AR and MR, as AR through mobile phone and tablet displays moves into fully immersive headsets, such as the Magic Leap 1 and Microsoft's HoloLens. AR and MR overlay virtual information on the physical world, which allows real-world training scenarios to be furnished with additional information and guiding cues. AR and MR were not addressed in this framework, but may pose new issues for designing and evaluating training, as fidelity and validity issues may primarily relate to how well the virtual assets are perceived to assimilate with physical ones, and how physical actions interact with virtual assets. Additionally the issue of

---

[3]Presence does appear to correlate with performance in VR, but it is unknown if the same holds true for learning (Stevens and Kincaid, 2015).

presence in VR is somewhat avoided, provided that virtual assets are accepted as part of the physical world. Research on AR and MR in training is in its infancy (e.g., see Gavish et al., 2015; Palmarini et al., 2018; Vovk et al., 2018), and further work is needed to explore these issues surrounding testing and validation.

We have provided a number of suggestions for how fidelity and validity can be assessed in VR but also simulated training more generally, and have emphasized the importance of addressing the right markers for the intended training purpose. The potential for VR training is huge, but this field could be hampered by a lack of evidence-based testing and injudicious application. VR technology will continue to develop, driven by the huge gaming market, but a fundamental understanding of the principles underpinning effective design for training purposes and an evidence-based approach to testing will be key to the success of VR training across many domains.

AUTHOR CONTRIBUTIONS

DH, JB, PS, MW, and SV contributed to the design, development, and writing of the manuscript.

FUNDING

REFERENCES

Abernethy, B., Baker, J., and Côté, J. (2005). Transfer of pattern recall skills may contribute to the development of sport expertise. *Appl. Cogn. Psychol.* 19, 705–718. doi: 10.1002/acp.1102

Baños, R. M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., and Rey, B. (2004). Immersion and emotion: their impact on the sense of presence. *Cyberpsychol. Behav.* 7, 734–741. doi: 10.1089/cpb.2004.7.734

Barnett, S. M., and Ceci, S. J. (2002). When and where do we apply what we learn?: a taxonomy for far transfer. *Psychol. Bull.* 128, 612–637. doi: 10.1037/0033-2909.128.4.612

Bideau, B., Kulpa, R., Vignais, N., Brault, S., Multon, F., and Craig, C. (2010). Using virtual reality to analyze sports performance. *IEEE Comput. Graph. Appl.* 30, 14–21. doi: 10.1109/MCG.2009.134

Bird, J. M. (2019). The use of virtual reality head-mounted displays within applied sport psychology. *J. Sport Psychol. Act.* 1–14. doi: 10.1080/21520704.2018.1563573

Blume, B. D., Ford, J. K., Baldwin, T. T., and Huang, J. L. (2010). Transfer of training: a meta-analytic review. *J. Manage.* 36, 1065–1105. doi: 10.1177/0149206309352880

Botella, C., Serrano, B., Baños, R. M., and Garcia-Palacios, A. (2015). Virtual reality exposure-based therapy for the treatment of post-traumatic stress disorder: a review of its efficacy, the adequacy of the treatment protocol, and its acceptability. *Neuropsychiatr. Dis. Treat.* 11, 2533–2545. doi: 10.2147/NDT.S89542

Bowman, D. A., and McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer* 40, 36–43. doi: 10.1109/MC.2007.257

Bright, E., Vine, S., Wilson, M. R., Masters, R. S. W., and McGrath, J. S. (2012). Face validity, construct validity and training benefits of a virtual reality turp simulator. *Int. J. Surg.* 10, 163–166. doi: 10.1016/j.ijsu.2012.02.012

Bright, E., Vine, S. J., Dutton, T., Wilson, M. R., and McGrath, J. S. (2014). Visual control strategies of surgeons: a novel method of establishing the construct validity of a transurethral resection of the prostate surgical simulator. *J. Surg. Educ.* 71, 434–439. doi: 10.1016/j.jsurg.2013.11.006

Brouwer, A.-M., Neerincx, M. A., and Kallen, V. (2011). EEG alpha asymmetry, heart rate variability and cortisol in response to virtual reality induces stress. *J. Cyberther. Rehabil.* 4, 21–34.

Buckingham, G. (2019). Examining the size–weight illusion with visuo-haptic conflict in immersive virtual reality. *Q. J. Exp. Psychol.* 72, 2168–2175. doi: 10.1177/1747021819835808

Bufton, A., Campbell, A., Howie, E., and Straker, L. (2014). A comparison of the upper limb movement kinematics utilized by children playing virtual and real table tennis. *Hum. Mov. Sci.* 38, 84–93. doi: 10.1016/j.humov.2014.08.004

Burdea, G. C., and Coiffet, P. (2003). *Virtual Reality Technology.* New York, NY: John Wiley & Sons.

Calatayud, D., Arora, S., Aggarwal, R., Kruglikova, I., Schulze, S., Funch-Jensen, P., et al. (2010). Warm-up in a virtual reality environment improves performance in the operating room. *Ann. Surg.* 251, 1181–1185. doi: 10.1097/SLA.0b013e3181deb630

Carruth, D. W. (2017). "Virtual reality for education and workforce training," in *Proceedings of the 2017 15th International Conference on Emerging ELearning Technologies and Applications (ICETA)*, Stary smokovec, 1–6. doi: 10.1109/ICETA.2017.8102472

Causer, J., and Ford, P. R. (2014). "Decisions, decisions, decisions": transfer and specificity of decision-making skill between sports. *Cogn. Process.* 15, 385–389. doi: 10.1007/s10339-014-0598-0

Chirico, A., and Gaggioli, A. (2019). When virtual feels real: comparing emotional responses and presence in virtual and natural environments. *Cyberpsychol. Behav. Soc. Netw.* 22, 220–226. doi: 10.1089/cyber.2018.0393

Ćosić, K., Popović, S., Kukolja, D., Horvat, M., and Dropuljić, B. (2010). Physiology-driven adaptive virtual reality stimulation for prevention and treatment of stress related disorders. *Cyberpsychol. Behav. Soc. Netw.* 13, 73–78. doi: 10.1089/cyber.2009.0260

Covaci, A., Olivier, A.-H., and Multon, F. (2015). Visual perspective and feedback guidance for VR free-throw training. *IEEE Comput. Graph. Appl.* 35, 55–65. doi: 10.1109/MCG.2015.95

Dahlstrom, N., Dekker, S., Winsen, R., and Nyce, J. (2009). Fidelity and validity of simulator training. *Theor. Issues in Ergon. Sci.* 10, 305–314. doi: 10.1080/14639220802368864

Davids, K., Williams, A. M., and Williams, J. G. (2005). *Visual Perception and Action in Sport.* London: Routledge. doi: 10.4324/9780203979952

Diemer, J., Alpers, G. W., Peperkorn, H. M., Shiban, Y., and Mühlberger, A. (2015). The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Front. Psychol.* 6:26. doi: 10.3389/fpsyg.2015.00026

Fink, P. W., Foo, P. S., and Warren, W. H. (2009). Catching fly balls in virtual reality: a critical test of the outfielder problem. *J. Vis.* 9:14. doi: 10.1167/9.13.14

Fowler, C. (2015). Virtual reality and learning: Where is the pedagogy? *Br. J. Educ. Technol.* 46, 412–422. doi: 10.1111/bjet.12135

Frederiksen, J. G., Sørensen, S. M. D., Konge, L., Svendsen, M. B. S., Nobel-Jørgensen, M., Bjerrum, F., et al. (2020). Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: a randomized trial. *Surg. Endosc.* 34, 1244–1252. doi: 10.1007/s00464-019-06887-8

Gavish, N., Gutiérrez, T., Webel, S., Rodríguez, J., Peveri, M., Bockholt, U., et al. (2015). Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. *Interact. Learn. Environ.* 23, 778–798. doi: 10.1080/10494820.2013.815221

Gray, R. (2017). Transfer of training from virtual to real baseball batting. *Front. Psychol.* 8:2183. doi: 10.3389/fpsyg.2017.02183

Gray, R. (2019). "Virtual environments and their role in developing perceptual-cognitive skills in sports," in *Anticipation and Decision Making in Sport*, eds A. M. Williams and R. C. Jackson (Abingdon: Routledge). doi: 10.4324/9781315146270-19

Gray, W. D. (2002). Simulated task environments: the role of high-fidelity simulations, scaled worlds, synthetic environments, and laboratory tasks in basic and applied cognitive research. *Cogn. Sci. Q.* 2, 205–227.

Harris, D., Buckingham, G., Wilson, M., Brookes, J., Mushtaq, F., Mon-Williams, M., et al. (2019a). Testing the fidelity and validity of a virtual reality golf putting simulator. *PsyarXiv* [Preprint]. doi: 10.31234/osf.io/j2txe

Harris, D., Buckingham, G., Wilson, M. R., Brookes, J., Mushtaq, F., Mon-Williams, M., et al. (2019b). Testing the effects of virtual reality on visuomotor skills. *PsyarXiv* [Preprint]. doi: 10.31234/osf.io/yke6c

Harris, D., Wilson, M., and Vine, S. (2019c). Development and validation of a simulation workload measure: the simulation task load index (SIM-TLX). *Virtual Real.* 1–10. doi: 10.1007/s00268-011-1141-4

Harris, D. J., Buckingham, G., Wilson, M. R., and Vine, S. J. (2019d). Virtually the same? How impaired sensory information in virtual reality may disrupt vision for action. *Exp. Brain Res.* 237, 2761–2766. doi: 10.1007/s00221-019-05642-8

Hashimoto, D. A., Petrusa, E., Phitayakorn, R., Valle, C., Casey, B., and Gee, D. (2018). A proficiency-based virtual reality endoscopy curriculum improves performance on the fundamentals of endoscopic surgery examination. *Surg. Endosc.* 32, 1397–1404. doi: 10.1007/s00464-017-5821-5

Janeh, O., Langbehn, E., Steinicke, F., Bruder, G., Gulberti, A., and Poetter-Nerger, M. (2017). Walking in virtual reality: effects of manipulated visual self-motion on walking biomechanics. *ACM Trans. Appl. Percept.* 14:12. doi: 10.1145/3022731

Janelle, C. M., Champenoy, J. D., Coombes, S. A., and Mousseau, M. B. (2003). Mechanisms of attentional cueing during observational learning to facilitate motor skill acquisition. *J. Sports Sci.* 21, 825–838. doi: 10.1080/0264041031000140310

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., et al. (2008). Measuring and defining the experience of immersion in games. *Int. J. Hum. Comput. Stud.* 66, 641–661. doi: 10.1016/j.ijhcs.2008.04.004

Jensen, L., and Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Educ. Inf. Technol.* 23, 1515–1529. doi: 10.1007/s10639-017-9676-0

Judd, C. H. (1908). The relation of special training to general intelligence. *Educ. Rev.* 36, 28–42.

Kirschner, P. A. (2002). Cognitive load theory: implications of cognitive load theory on the design of learning. *Learn. Inst.* 12, 1–10. doi: 10.1016/S0959-4752(01)00014-7

Kraiger, K., Ford, J. K., and Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *J. Appl. Psychol.* 78, 311–328. doi: 10.1037/0021-9010.78.2.311

Kramida, G. (2016). Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE Trans. Vis. Comput. Graph.* 22, 1912–1931. doi: 10.1109/TVCG.2015.2473855

Krijn, M., Emmelkamp, P. M. G., Olafsson, R. P., and Biemond, R. (2004). Virtual reality exposure therapy of anxiety disorders: a review. *Clin. Psychol. Rev.* 24, 259–281. doi: 10.1016/j.cpr.2004.04.001

Levac, D. E., Huber, M. E., and Sternad, D. (2019). Learning and transfer of complex motor skills in virtual reality: a perspective review. *J. Neuroeng. Rehabil.* 16:121. doi: 10.1186/s12984-019-0587-8

Lopes, P., You, S., Cheng, L.-P., Marwecki, S., and Baudisch, P. (2017). "Providing haptics to walls & heavy objects in virtual reality by means of electrical muscle stimulation," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, Denver, CO, 1471–1482. doi: 10.1145/3025453.3025600

Magdalon, E. C., Michaelsen, S. M., Quevedo, A. A., and Levin, M. F. (2011). Comparison of grasping movements made by healthy subjects in a 3-dimensional immersive virtual versus physical environment. *Acta Psychol.* 138, 126–134. doi: 10.1016/j.actpsy.2011.05.015

Meehan, M., Insko, B., Whitton, M., and Brooks, F. P. Jr. (2002). "Physiological measures of presence in stressful virtual environments," in *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques*, San Antonio, TX, 645–652. doi: 10.1145/566570.566630

Moghim, M., Stone, R., Rotshtein, P., and Cooke, N. (2015). "Adaptive virtual environments: A physiological feedback HCI system concept," in *Proceedings of the 2015 7th Computer Science and Electronic Engineering Conference (CEEC)*, Colchester, 123–128. doi: 10.1109/CEEC.2015.7332711

Moghimi, M., Stone, R., Rotshtein, P., and Cooke, N. (2016). Influencing human affective responses to dynamic virtual environments. *Presence* 25, 81–107. doi: 10.1162/pres_a_00249

Moglia, A., Ferrari, V., Morelli, L., Melfi, F., Ferrari, M., Mosca, F., et al. (2014). Distribution of innate ability for surgery amongst medical students assessed by an advanced virtual reality surgical simulator. *Surg. Endosc.* 28, 1830–1837. doi: 10.1007/s00464-013-3393-6

Mohler, B. J., Campos, J. L., Weyel, M. B., and Bulthoff, H. H. (2007). "Gait parameters while walking in a head-mounted display virtual environment and the real world," in *Proceedings of the IPT-EGV Symposium*, eds B. Fröhlich, R. Blach, and R. van Liere (Aire-la-Ville: Eurographics Association), 1–4.

Morina, N., Ijntema, H., Meyerbröker, K., and Emmelkamp, P. M. G. (2015). Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. *Behav. Res. Ther.* 74, 18–24. doi: 10.1016/j.brat.2015.08.010

Nacke, L., and Lindley, C. A. (2008). "Flow and immersion in first-person shooters: measuring the player's gameplay experience," in *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, Ithaca, NY, 81–88. doi: 10.1145/1496984.1496998

Pallavicini, F., Argenton, L., Toniazzi, N., Aceti, L., and Mantovani, F. (2016). Virtual reality applications for stress management training in the military. *Aerosp. Med. Hum. Perform.* 87, 1021–1030. doi: 10.3357/AMHP.4596.2016

Palmarini, R., Erkoyuncu, J. A., Roy, R., and Torabmostaedi, H. (2018). A systematic review of augmented reality applications in maintenance. *Robot. Comput. Integr. Manuf.* 49, 215–228. doi: 10.1016/j.rcim.2017.06.002

Perfect, P., Timson, E., White, M. D., Padfield, G. D., Erdos, R., and Gubbels, A. W. (2014). A rating scale for the subjective assessment of simulation fidelity. *Aeronaut. J.* 118, 953–974. doi: 10.1017/S0001924000009635

Renshaw, I., Davids, K. W., Shuttleworth, R., and Chow, J. Y. (2009). Insights from ecological psychology and dynamical systems theory can underpin a philosophy of coaching. *Int. J. Sport Psychol.* 40, 540–602.

Roberts, P. G., Guyver, P., Baldwin, M., Akhtar, K., Alvand, A., Price, A. J., et al. (2017). Validation of the updated ArthroS simulator: face and construct validity of a passive haptic virtual reality simulator with novel performance metrics. *Knee Surg. Sports Traumatol. Arthrosc.* 25, 616–625. doi: 10.1007/s00167-016-4114-1

Rosalie, S. M., and Müller, S. (2012). A model for the transfer of perceptual-motor skill learning in human behaviors. *Res. Q. Exerc. Sport* 83, 413–421. doi: 10.1080/02701367.2012.10599876

Rose, F. D., Attree, E. A., Brooks, B. M., Parslow, D. M., and Penn, P. R. (2000). Training in virtual environments: transfer to real world tasks and equivalence to real task training. *Ergonomics* 43, 494–511. doi: 10.1080/001401300184378

Sala, G., and Gobet, F. (2017). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Curr. Dir. Psychol. Sci.* 26, 515–520. doi: 10.1177/0963721417712760

Sankaranarayanan, G., Li, B., Manser, K., Jones, S. B., Jones, D. B., Schwaitzberg, S., et al. (2016). Face and construct validation of a next generation virtual reality (Gen2-VR©) surgical simulator. *Surg. Endosc.* 30, 979–985. doi: 10.1007/s00464-015-4278-7

Sanz, F. A., Olivier, A., Bruder, G., Pettré, J., and Lécuyer, A. (2015). "Virtual proxemics: locomotion in the presence of obstacles in large immersive projection environments," in *Proceedings of the 2015 IEEE Virtual Reality (VR)*, Bonn, 75–80. doi: 10.1109/VR.2015.7223327

Saunders, T., Driskell, J. E., Johnston, J. H., and Salas, E. (1996). The effect of stress inoculation training on anxiety and performance. *J. Occup. Health Psychol.* 1, 170–186. doi: 10.1037/1076-8998.1.2.170

Singley, M. K., and Anderson, J. R. (1989). *The Transfer of Cognitive Skill*. Cambridge, MA: Harvard University Press.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 3549–3557. doi: 10.1098/rstb.2009.0138

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., et al. (2006). A virtual reprise of the Stanley Milgram obedience experiments. *PLoS One* 1:e39. doi: 10.1371/journal.pone.0000039

Slater, M., and Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Front. Robot. AI* 3:74. doi: 10.3389/frobt.2016.00074

Sorrentino, R. M., Levy, R., Katz, L., and Peng, X. (2005). Virtual visualization: preparation for the Olympic Games long-track speed skating. *Int. J. Comput. Sci. Sport* 4, 40–45.

Stanney, K. M., Mollaghasemi, M., Reeves, L., Breaux, R., and Graeber, D. A. (2003). Usability engineering of virtual environments (VEs): identifying multiple criteria that drive effective VE system design. *Int. J. Hum. Comput. Stud.* 58, 447–481. doi: 10.1016/S1071-5819(03)00015-6

Stevens, J. A., and Kincaid, J. P. (2015). The relationship between presence and performance in virtual simulation training. *Open J. Model. Simul.* 03, 41–48. doi: 10.4236/ojmsi.2015.32005

Stoffregen, T. A., Bardy, B. G., Smart, L. J., and Pagulayan, R. (2003). "On the nature and evaluation of fidelity in virtual environments," in *Virtual and Adaptive Environments: Applications, Implications and Human Performance Issues*, eds L. J. Hettinger and M. W. Haas (Mahwah, NJ: Lawrence Erlbaum Associates), 111–128. doi: 10.1201/9781410608888.ch6

Sutcliffe, A. G., and Kaur, K. D. (2000). Evaluating the usability of virtual reality user interfaces. *Behav. Inf. Technol.* 19, 415–426. doi: 10.1080/014492900750 052679

Thorndike, E. L. (1906). *Principles of Teaching*. New York, NY: Seiler.

Tirp, J., Steingröver, C., Wattie, N., Baker, J., and Schorer, J. (2015). Virtual realities as optimal learning environments in sport – A transfer study of virtual and real dart throwing. *Psychol. Test Assess. Model.* 57, 57–69.

Usoh, M., Catena, E., Arman, S., and Slater, M. (2000). Using presence questionnaires in reality. *Presence* 9, 497–503. doi: 10.1162/105474600566989

Vaughan, N., Gabrys, B., and Dubey, V. N. (2016). An overview of self-adaptive technologies within virtual reality training. *Comput. Sci. Rev.* 22, 65–87. doi: 10.1016/j.cosrev.2016.09.001

Vine, S. J., Chaytor, R. J., McGrath, J. S., Masters, R. S. W., and Wilson, M. R. (2013). Gaze training improves the retention and transfer of laparoscopic technical skills in novices. *Surg. Endosc.* 27, 3205–3213. doi: 10.1007/s00464-013-2893-8

Vine, S. J., McGrath, J. S., Bright, E., Dutton, T., Clark, J., and Wilson, M. R. (2014). Assessing visual control during simulated and live operations: Gathering evidence for the content validity of simulation using eye movement metrics. *Surg. Endosc.* 28, 1788–1793. doi: 10.1007/s00464-013-3387-4

Vovk, A., Wild, F., Guest, W., and Kuula, T. (2018). "Simulator sickness in augmented reality training using the Microsoft HoloLens," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Vol. 209, Montreal, QC, 1–9. doi: 10.1145/3173574.3173783

Whitwell, R. L., Ganel, T., Byrne, C. M., and Goodale, M. A. (2015). Real-time vision, tactile cues, and visual form agnosia: Removing haptic feedback from a "natural" grasping task induces pantomime-like grasps. *Front. Hum. Neurosci.* 9:216. doi: 10.3389/fnhum.2015.00216

Wijeyaratnam, D. O., Chua, R., and Cressman, E. K. (2019). Going offline: differences in the contributions of movement control processes when reaching in a typical versus novel environment. *Exp. Brain Res.* 237, 1431–1444. doi: 10.1007/s00221-019-05515-0

Woodworth, R. S., and Thorndike, E. L. (1901). The influence of improvement in one mental function upon the efficiency of other functions. (I). *Psychol. Rev.* 8, 247–261. doi: 10.1037/h0074898

Young, W., Ferguson, S., Brault, S., and Craig, C. (2011). Assessing and training standing balance in older adults: a novel approach using the 'Nintendo Wii' Balance Board. *Gait Posture* 33, 303–305. doi: 10.1016/j.gaitpost.2010.10.089

Zaal, F. T. J. M., and Bootsma, R. J. (2011). Virtual reality as a tool for the study of perception-action: the case of running to catch fly balls. *Presence* 20, 93–103. doi: 10.1162/pres_a_00037

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer JF declared a past collaboration, with one of the authors, MW, to the handling Editor.