Media Psychology, 11:354–376, 2008

ISSN: 1521-3269 print/1532-785X online
DOI: 10.1080/15213260802285214

The Effect of Interactivity on Learning Physical Actions in Virtual Reality

JEREMY BAILENSON
Stanford University

KAYUR PATEL
University of Washington

ALEXIA NIELSEN
Stanford University

RUZENA BAJSCY
University of California at Berkeley

SANG–HACK JUNG
University of California at Berkeley

GREGORIJ KURILLO
University of California at Berkeley

Virtual reality (VR) offers new possibilities for learning, specifically for training individuals to perform physical movements such as physical therapy and exercise. The current article examines two aspects of VR that uniquely contribute to media interactivity: the ability to capture and review physical behavior and the ability to see one's avatar rendered in real time from third person points of view. In two studies, we utilized a state-of-the-art, image-based tele-immersive system, capable of tracking and rendering many degrees of freedom of human motion in real time. In Experiment 1, participants learned better in VR than in a video learning condition according to self-report measures, and the cause of the advantage was seeing one's avatar stereoscopically in the third person. In Experiment 2, we added a virtual mirror in the learning environment to further leverage the ability to see oneself from novel angles in real time. Participants learned better in VR than in video according to objective performance measures. Implications for learning via interactive digital media are discussed.

Address correspondence to Jeremy Bailenson, Department of Communication, Stanford, CA 94043. E-mail: bailenson@stanford.edu

Historically, virtual reality (VR) learning environments have been applied to a multitude of learning scenarios, from flight simulation (Hays, Jacobs, Prince, & Salas, 1992) to medical training (Berkley, Turkiyyah, Berg, Ganter, & Weghorst, 2004) to classroom learning (Pantelidis, 1993). One of the most exciting aspects of VR is its ability to leverage interactivity. Virtual systems offer a novel, flexible environment with affordances not possible from previous mediums like video and text (Blascovich et al., 2002). These virtual environments offer unique opportunities for learning on-demand (Trondsen & Vickery, 1997), customization and personalization (Kalyanaraman & Sundar, 2006), and feedback mechanisms (Lee & Nass, 2005). Previous research has shown that on-demand learning provides an advantage over face-to-face human interaction (Trondsen & Vickery, 1997). In a variety of contexts, VR offers possibilities to extend the notion of interactive learning in ways not possible through face-to-face interaction (see Bailenson et al., 2008, for a review of research on learning in VR).

The current studies measured the effects of learning physical tasks from a virtual system when compared to video, leveraging features such as three-dimensional depth cues, representations of the participant next to the instructor, and changes of scene angle not possible through traditional video representations.

INTERACTIVITY IN MEDIA

As Sundar and Nass (2000) point out, digital technology has drastically changed the way in which communication occurs; audiences, typically referred to as passive receivers, have now become more active in their media experience, often being referred to as "users." Conceptual definitions of interactivity typically emphasize three dimensions: technology, process, and user. Proponents of the technology dimension argue that interactivity is an affordance of technology (Steuer, 1992). Steuer defines interactivity as "the extent to which users can participate in modifying the form and content of a mediated environment in real time" (p. 84). This active modification of content is particularly salient in virtual worlds which function via the cycle of *tracking* (i.e., using sensing equipment to measure movements and behavior) and *rendering* (i.e., displaying a digital representation of the world to reflect the user's representation). In other words, using media such as the Internet, it is possible to have applications respond in a tailored manner to the way a user hits keys and moves the mouse. However, using VR, it is possible to have applications respond to tracking data on a much more sensitive basis—to the way a user moves, walks, gestures, and gazes. Consequently, VR is a medium that affords more interactivity than other media due to the richness of the potential behavioral tracking (Lanier, 2001).

Proponents of the process dimension argue that it is the mode of communication that influences interactivity, that is, one-way versus two-way communication (Rafaeli, 1988). In this way, VR is a media which can maximize the parameter of two-way communication as defined by Liu and Shrum (2002): the degree to which two parties in a media can deliver feedback to one another in real-time based on behavioral tracking data.

Finally, proponents of the user dimension contend that interactivity is a function of perception and is ultimately in the mind of the participant (Kiousis, 2002; Wu, 1999). The notion of perceived interactivity comes into play—a concept that hinges more on the apparent speed or responsiveness of a system rather than its actual interactive capabilities (Newhagen, Cordes, & Levy, 1995).

In all three views, feedback and contingency or expected contingency are the defining features of interactivity and can be uniquely examined in a VR learning environment (e.g., Cappella & Pelachaud, 2002; Skalski & Tamborini, 2007). In essence, the entire experience can be customized to fit the ideal learning needs of each participant, something that traditional training cannot offer.

VR AS A SOCIAL SCIENCE TOOL

Immersive VR perceptually surrounds the user, increasing his or her sense of presence or actually being within it. Consider a child's video game. If the child were to have special equipment that allowed her to take on the actual point of view of the main character of the video game, that is, to control that character's movements with her own movements such that the child is actually inside the video game, then she would be in an immersive simulation. In immersive VR, sensory information is more psychologically prominent and engaging than the sensory information gleaned from other types of media (Blascovich et al., 2002; Lanier, 2001).

There are two important features of VR that uniquely facilitate social science research. The first is tracking a user's movements, including body position, head direction, as well as facial expressions and gestures, thereby providing a wealth of information about where the user is focusing his attention, what he observes from that specific vantage point, and his reactions to the environment. The second is that the designer of the VR system has tremendous control over the user's experience and can alter the appearance and design of the virtual world to fit experimental goals, providing a wealth of real-time adjustments to specific user actions.

VR has been used to study many psychological processes such as perception (Loomis, Beall, Macuga, Kelly, & Smith, 2006), interpersonal distance (Bailenson, Blascovich, Beall, & Loomis, 2003), leadership (Hoyt & Blascovich, 2003) behavioral mimicry (Bailenson & Yee, 2005), violence (Persky & Blascovich, 2008), and obedience (Slater et al., 2006).

MOTION TRAINING IN VR

One research area in particular that leverages the affordances of VR is training individuals in physical motion-related tasks. For example, VR systems are being developed to assist in physical rehabilitation (Holden, 2005). Traditional rehabilitation requires repetitive motion, usually supervised by a human trainer, which can be arduous and boring for patients. Virtual systems have been developed as a way to create a more interactive rehabilitation experience, increasing patient motivation (Rizzo & Kim, 2005). In addition, systems can provide feedback on the patients' movements, which can reduce the errors made by patients during the motion process (Weiss, Rand, Katz, & Kizony, 2004). Athletics departments use VR to develop autonomous sports training systems; for example, one system analyzes images to capture participant posture during a golf swing in order to provide feedback (Smith & Lovell, 2003).

The medical field uses VR extensively for training physical motions of medical students (e.g., Seymour, 2005). One area in which VR applications may be helpful is training various types of full body motions. Previous research in this area has focused on Tai Chi in particular, largely due to its potential health benefits, its objectivity in terms of correct and incorrect performance, and its feasibility for use in immersive VR settings (i.e., slow motions that do not require huge spaces). Becker and Pentland (1996), in one of the first studies using Tai Chi in VR, created a tool to assist cancer patients in visualizing healthy blood cells by which Tai Chi moves were used as a basis for the patients' interaction with the system. The virtual environment was projected onto a large screen and computer vision tracked the participants' head and hand movements. The virtual teacher was also able to provide feedback to the students by comparing their movements to the expected motion and pointing out errors.

The only study to date that has measured learning from a virtual Tai Chi teacher was conducted by Chua et al. (2003). They taught participants Tai Chi using a skeleton-based rendering system that captured 42 degrees of freedom. Students viewed the virtual environment through a wireless head-mounted display (HMD). Participants were randomly assigned to learn four Tai Chi moves using five different learning modes. The control condition showed the teacher performing the moves directly in front of the student, simulating a physical world Tai Chi course. The other modes were four teachers surrounding the student, four teachers next to four images of the student side-by-side, superimposition of the student's avatar over a wireframe of teacher, and superimposition of a wireframe of the student over a stick figure of the teacher. In addition to self-report measures, the study used a computer algorithm to objectively evaluate participants' performance. For their objective measures, they found that none of the learning modes had a significant effect on participants' learning of the moves.

Although Chua and colleagues' (2003) experiment provided an important foundation to the study of immersive learning, the current work provides an approach that combines immersive technology with traditional social scientific methodologies. Consequently, while Chua et al. did not demonstrate improvements in learning, we use measurement tools sensitive enough to demonstrate the superiority of VR systems over other forms of learning media. There are four notable advancements of our study over Chua et al's.

First, because our system was a projection-based system, there was no cumbersome HMD to inhibit naturalistic body motions. We provided extremely high presence without having to drastically limit the users' movements. Second, the image-based reconstruction was not limited to a set number of measured degrees of freedom in the way that model-based, motion-capture animation systems are, so the user's physical motions were more realistic. Third, given that learning such a high-level system of movements is likely a gestalt phenomenon, we evaluated learning performance both via self-report and blind coder ratings. These methods allowed us to test how well participants learned the overall gestalt of the Tai Chi system of movements, as opposed to only analyzing the success of the learning on a micro, joint-by-joint basis. Finally, we not only measured performance while the virtual teacher was present, but also measured learning (i.e, how well the student performed the Tai Chi moves later on outside of the simulation) when the teacher was no longer present. In sum, the current research is unique both technologically in terms of tracking and rendering, as well as methodologically in terms of social scientific learning measurement.

OVERVIEW OF EXPERIMENTS

We examined some of the unique affordances of VR and learning by comparing Tai Chi students who learned from a three-dimensional digital teacher while in an immersive VR simulation to ones who learn Tai Chi from a video-like simulation. Two experiments were conducted. In both, participants learned three separate Tai-Chi moves from a recorded teacher; they were tested on the moves and given questionnaires on their learning experience. Our analysis strategy was to compare differences in learning and self-report across various experimental conditions.

In Experiment 1, we manipulated two variables, *media* (VR vs. video) and *review* (the ability to go back and watch oneself in the third person or the teacher between learning sessions or not). Given the previous research demonstrating the benefits of interactivity in nonimmersive contexts discussed above (e.g., Skalski, & Tamborini, 2007), we predicted that there would be better subjective and objective learning in interactive VR compared to video. Moreover, we predicted increasing the amount of interactivity and feedback by allowing participants to review their own behavior would increase learning, as previous research has demonstrated the value of self-review as a learning strategy (e.g., Trondson & Vickery, 1997).

Experiment 2[1] was similar to Experiment 1 except for two major differences. All participants had the ability to review their learning sessions and, in the VR condition, the degree of interactivity was increased by providing participants with two separate views of themselves in the third person (one in front of them in space; the other a mirror image) in real-time via adding a virtual mirror to the room. As in the first study, we predicted that participants would learn better in the VR condition than the video condition.

EXPERIMENT 1

Design

The experiment was a between-subjects 2 (*media*: video vs. VR) × 2 (*review*: review vs. no review) design. Participants were randomly assigned to each of the four conditions.

Apparatus

We utilized a fully immersive VR apparatus (Yang et al., 2005) that allows 360 degree, real-time motion capture of people and objects. (See Figure 1).

The apparatus is comprised of a large, rectangular "cage" measuring 3.2 meters (width) × 3.8 meters (depth) × 2.5 meters (height). Attached to this cage is a sea of 48 cameras that simultaneously captures images of the participant. In front of the apparatus is a computer array of 13 networked computers and two projectors used for the passive stereo display. In front of the projectors is a large (1.8 × 1.3 meters) projection screen. Off to the side. a lab technician operates the system from the operator control station.

The system itself is comprised of three main components: a) image processing, b) data transmission, and c) visualization.

The user enters the cage and stands in the middle of the floor, facing the projection screen. In the image processing stage, the sea of cameras captures images of the user. Twelve camera clusters are attached to the apparatus cage. Each cluster has four cameras, three grayscale cameras for stereo reconstruction and a color camera for texture capturing. In the first step, the scene of the background is captured for each camera to eliminate the views of the cage and room from the virtual image, creating a blank, black virtual space on the projection screen.

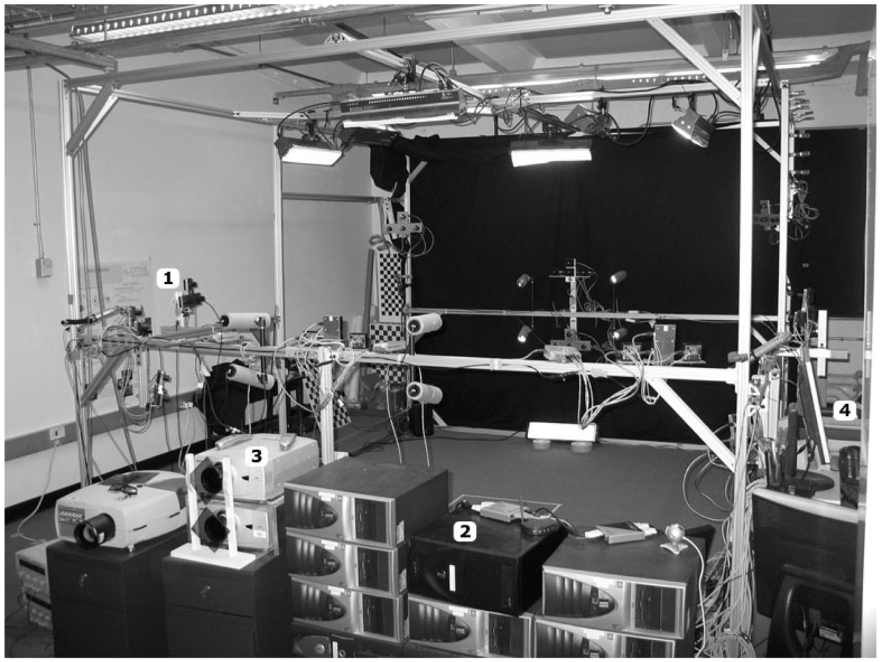

FIGURE 1 The tele-immersion system. Camera clusters (1) capture images of objects inside the cage. The computer array (2) processes the images to create a 360 degree representation. Dual projectors (3) project a polarized image to a projection screen (not shown.) The operator station (4) controls the system and provides the display for the review phase.

Each camera cluster simultaneously captures images of the user. To aid in lighting, the cage is illuminated using eight professional photography lamps, which produce an even and diffuse light. Because dark colors tend to blend with the backdrop, users must wear brightly colored (preferably pastel) clothing to be seen well by the cameras. Each camera cluster then sends the image data through a FireWire connection to its own dedicated computer. Each computer processes the images using an image-based stereo reconstruction algorithm and creates a partial representation of the scene as a cloud of three-dimensional points.

In the data transmission stage, the individual computers send their representation of the scene to single rendering computer. In order to achieve a real-time representation of the user, these data must be sent simultaneously from all clusters with very little delay. The rendering computer combines over 75,000 three-dimensional points to create a real-time video representation of the user with a frame rate of about 10 frames per second.

In the visualization stage, the rendering computer produces a 360 degree, fully three-dimensional representation of the user (see Figure 2). The full representation of the user is projected onto a large screen. The system is capable of producing passive stereo projection using circular polarization.

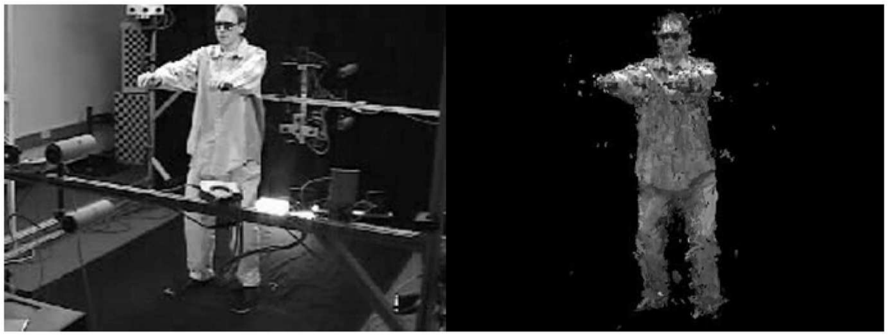

FIGURE 2 Projected image of participant in real life (left) and reconstructed in the virtual space (right).

Two slightly offset images are projected onto the screen by the two projectors through a polarizing filter. When viewed with polarized glasses, these dual images combine to create a three-dimensional image with depth cues.

The system is able to record the 360 degree scene as the user is moving in the virtual space. This recording can then be played back by the operator and viewed from any angle. The system also allows previously captured scenes to be played back simultaneously with real-time captured three-dimensional video. Users can see their real-time images inhabiting the same virtual space as the recorded representation of a user who was in the room in the past. This capability was essential for users in the VR review condition, so that they were able to review their own performance as well as the instructor's performance.

Materials

To capture the virtual model of the teacher, a real Tai Chi instructor volunteered to come into the lab and perform Tai Chi moves while being recorded by the VR apparatus. The moves were "brush knee twist," "part the wild horse's mane" and "throwing the loom." The instructor performed each of the three moves approximately 10 times, all recorded by the system. A team of three people, including the Tai Chi instructor, then chose the best instance of each of the three moves from the 10 samples.

During presentation of each of the three moves to the participants, the best instance was then put into an algorithm that played back the move four times consecutively, with a 10 second period of black screen in between each repetition of the move. The name of the move appeared in white lettering above the instructor when the move was being played back. During the 10 second break, the words "Rest" appeared above the black space for the first 5 seconds; "Get ready. Starting in 5 sec..." appeared for the last 5 seconds to warn the participant that the video was going to play again.

In the video condition, participants saw just the playback image of the instructor from a single, front-on camera angle (see Figure 3). Only one projector was used so that participants saw only a two-dimensional image of the instructor. In the VR condition, participants saw a real-time image of themselves next to the playback image of the instructor. Two projectors were used so that participants saw both the image of themselves and the instructor in three dimensions with depth cues. In addition, the system recorded the virtual model of participants, which would be used later for the review condition.

Two methods of playback used if a participant was assigned to a review condition. Participants in the video condition sat in front of a computer monitor and watched the video of the teacher alone performing the Tai Chi moves. They were able to rewind, fast forward, and pause the video using clickable options on the screen, much like the options that would be available watching a home exercise tape using VCR or DVD player. Participants in the VR condition sat in front of the same computer monitor. However, since they were in the VR condition and therefore saw an image of themselves next to the instructor during practices, they were able to watch a video of themselves performing the moves alongside the instructor. As in the video condition, participants were also able to rewind, fast forward, and pause. With the added benefit of 360 degree image capture with the VR system, participants were able to change the angle from which they viewed the scene by using the mouse. For example, during the recording playback, they could turn the image to view themselves and the instructor from the back, and then turn the view again to view themselves from the side.

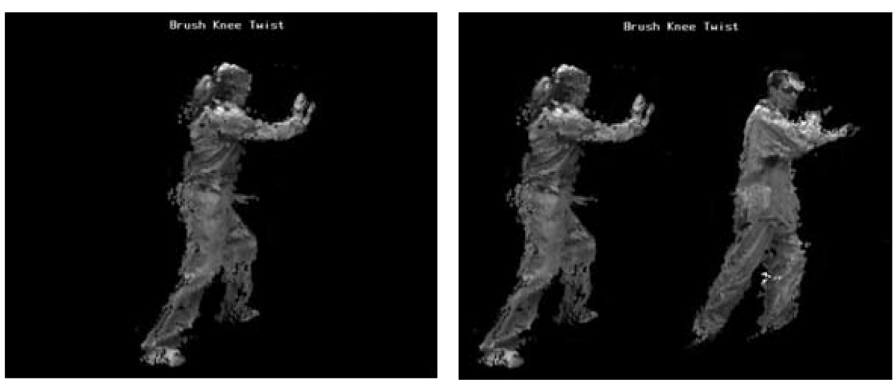

FIGURE 3 In the video condition (left), the participant sees the prerecorded video of the teacher alone on the screen. The image is flat and two dimensional to simulate the playback of a video recording. In the VR condition (right), the participant sees the prerecorded video of the teacher alongside a real-time representation of themselves. In the VR condition both representations appear to have three-dimensional depth cues, with the aid of polarized images from two projectors and polarized glasses.

Participants

Forty-one adult participants (20 males, 21 females) were recruited from the student population at a large West Coast university. Their mean age was 22.8 (*SD* = 6.7. min = 18, max = 56). Nineteen participants were Caucasian, 17 were Asian, two were Latino, two were Indian, and one was African American. All participants were currently enrolled in college or had received college degrees. Participants were randomly assigned to one of four conditions with approximately the same gender ratio in each condition. Cybersickness was not formally measured as not a single participant reported any symptoms during pilot testing or during the running of the two actual experiments.

Participants were recruited through department mailing lists and online advertisements, and were compensated for their time.

Procedure

Upon entering the lab, participants were asked to fill out a consent form and demographic survey (both on paper). After filling out these forms, the experimenter read the participants a brief overview of the various the phases of the task.

Participants were then provided with a light blue shirt and pants and were instructed to wear the clothes chosen to optimize the computer vision algorithm over their regular clothing. All participants were provided with the same blue shirt and pants, which they wore over their own clothes. Participants were also given polarized glasses to wear. These glasses were necessary for participants in the fully immersive condition to see the three-dimensional image produced by the polarized projection system. However, these glasses were given to participants in all conditions for increased experimental control; all participants were told that the glasses allowed them to better see the virtual instructor, even if this was not actually the case. The glasses did not affect the clarity of the image on the screen even in the video condition.

The experiment proceeded in five phases. In Training Phase 1, participants viewed the instructor performing the three separate Tai Chi moves. The display varied depending on condition, with the video condition participants seeing only the instructor and the VR condition participants seeing both the instructor and themselves. Participants were first reminded that they would later be tested on the moves and told to attempt to commit to memory as much of the moves as possible. For each move, the virtual instructor performed the move four times. In between each move, there was a 10 second period of rest where the screen was black. The participants were told to stand still and watch the instructor the first time the move was performed, so that participants would have a chance to get familiar with the movements without the distraction of trying to imitate them. Then for the remaining three repetitions, the participants were told to imitate the moves to the best of their ability. The name of each move was written in white lettering above virtual image. In addition, the experimenter verbally introduced the name of the move at the beginning of each series of moves, and reminded the participant of the name of the move at the end of each series. Participants' actions during this phase of the experiment were recorded using a digital video camera for later analysis.

In the Review Phase, participants were assigned to either the review condition or the no review condition. For the review condition, participants were seated at the operator station in front of a computer monitor. For the no review condition, participants were told that the system had to be recalibrated and were asked to wait for approximately the same amount of time as the review phase.

In Training Phase 2, participants repeated the task from Training Phase 1. They viewed the instructor performing the three Tai Chi moves four times each; they watched the first performance and then mimicked the teacher in the subsequent three performances. Participants were also reminded that they would be tested on the moves after the end of the third phase. Participants during this phase were also videotaped.

In the Testing Phase, participants were asked to recall and perform the Tai Chi moves they had learned without the aid of the virtual Tai Chi instructor. The experimenter called out the name of the move (in the same order in which the moves were learned), and asked the participant to perform the move. Again, the participants' moves were videotaped.

In the Questionnaire Phase, participants were asked to remove the provided clothing and polarized glasses. They then filled out a paper questionnaire containing questions about their perceptions of the virtual environment. Participants were then paid for their participation.

Measures

Subjective learning. Participants were administered a questionnaire after participating in the study (see the Appendix for the actual questions). There were 17 questions rated on Likert scale ranging from 1 to 5 with higher numbers indicating positive responses. These questions asked for participants' overall impressions of the instructor, the virtual environment, and the task. These questions were exploratory in nature and were designed to assess participants' overall impression of the experience. The remaining 10 open-ended questions asked for participants' feedback on the study, including which parts of the task they enjoyed and how the felt the experience could be improved. These open-ended questions were used to check for technological problems as well as to improve the procedure for the subsequent experiment. The Likert scale questions were averaged to provide an overall Positivity

Score of participants' experience in the environment (M = 3.11, SD = .63, min = 1.91, max = 4.71, Cronbach's α = .89). Because the reliability of the score was so high, and exploratory factor analyses did not demonstrate differences in conditions among the various factors, we only report data from the overall score without breaking the score down to individual factors.

Objective learning. The second measure was an objective behavioral measure of participants' ability to perform the Tai Chi moves during the task. The participants' performances during the task were recorded using a digital video camera. The recorded videos were analyzed separately by participant, phase, and movement. Each of the 41 participants performed the three moves three times during three phases for a total of 9 videos per participant or 369 videos.

At the beginning of the study, the Tai Chi teacher was recorded performing the individual Tai Chi moves using the digital video camera. Each move had its own video; these videos were used as a reference for grading the participants' performance during the task. The Tai Chi expert we recorded in the immersive system created a coder reference sheet and worked with the experimenters and coders to make sure there were clear definitions of differential performance on the moves. Each Tai Chi move was divided into eight distinct steps. The reference sheet contained detailed descriptions of each of the eight steps, including time codes on the corresponding video file to provide for greater consistency in coder evaluations.

Two blind coders watched and graded each video on 10 dimensions.[2] The coders graded each of the eight distinct steps for each move from the reference sheet on a seven-point scale from *very poor* (1) to *very well* (7). Intercoder reliability for each observation was high (Cronbach's α = .79). Consequently, the individual coder scores were averaged to create a mean score for the 10 dimensions graded for each video. The 10 dimensions were then averaged to create a mean score for each move. Moves for each phase were then averaged to create an overall performance score for each participant during each phase (M = 2.66, SD = .72, min = 1.21, max = 3.97).

Results and Discussion

Subjective learning. To test our hypotheses concerning differences among groups we ran an analysis of variance (ANOVA) with Media and Review as independent discreet variables and Positivity Score as the continuous dependent variable. All assumptions required for the ANOVA were met. There was a significant effect for the VR condition ($F[1, 37] = 9.679$, $p = .004$, partial Eta-Squared = .21) such that participants rated VR more positively than video. Neither the effect of review ($F[1,37] = 1.21$, $p = .28$, partial Eta-Squared = .03) nor the interaction between review and 3D ($F[1, 37] = .71$, $p = .41$, partial Eta-Squared = .02) were significant. (See the Appendix for the means for each question.)

Objective learning. In order to test our hypotheses concerning differences among groups in regards to learning we ran a multivariate analysis of variance with the mean score for each of the three phases (Training 1, Training 2, and Test) as the continuous dependent variables and Media and Review as discreet independent variables. All assumptions required for the ANOVA were met (All Fs < 1.3, ps > .25, *partial Eta Squares* <.05).

Summary

In the current study, participants reported that they learned better, enjoyed the experience more, and thought the virtual teacher was more credible in VR than in video. However, this difference only emerged in self-report questionnaires, not in the behavioral data we observed of the participants actually performing the Tai Chi moves, and regardless, the ability to review seemed to make little difference on either measure.

Previous research on interactive media (Burgoon et al., 2000) demonstrated that adding interactive channels are only helpful if those channels directly relate to task performance by increasing personal involvement, creating a mutual connection between the user and the interface. In our VR condition, we may not have leveraged the utility of seeing oneself in the third person enough to create differences in actual learning. Consequently, in Experiment 2 we decided to increase the utility of the interactive feature by adding a second way of seeing oneself in the third person. By offering multiple angles of the self in real-time, it should be possible for the learners to better assess the quality of their movements and to compare the differences between their own actions and those of the teacher. As Liu and Shrum (2002) point out, providing more feedback is an essential aspect of interactivity; by allowing people to see two separate versions of themselves simultaneously, they gain more feedback, specifically more information about how their body is moving via different angles in real time. Consequently, we predicted that we would observe better objective and subjective learning in VR with the increased interactivity channel than in video.

EXPERIMENT TWO

Design

We manipulated one variable, Media. In the video condition, the student only saw two-dimensional images of the environment and of the virtual teacher, and only could see the teacher from a single, front-on camera angle. In the VR condition, participants saw four stereoscopic human representations, an image of themselves rendered in the third person and the teacher from behind, as well as a reflection of both those images in a virtual mirror. (See

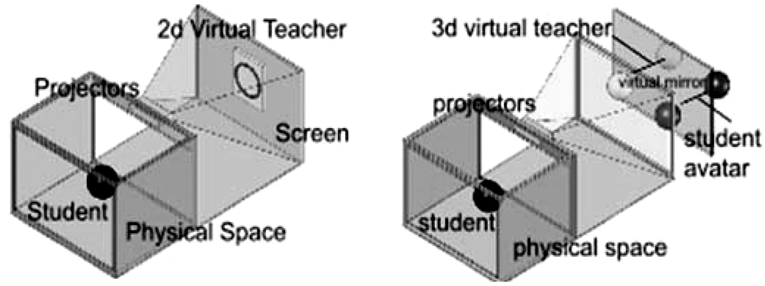

FIGURE 4 Diagrams of the two conditions. The position of the student in physical space and the teacher in virtual space is depicted in the video condition (left), The position of the student in physical space and both student and teacher in virtual space are depicted in the VR condition (right).

Figure 4). For the VR condition, we chose a mirror rather than arbitrary figures duplicated in space to make the interaction as natural as possible for the participants. Furthermore, the ability to duplicate images from different vantages is an inherent affordance of our system, and the inclusion of a mirror has been shown to be beneficial when learning physical motion (Sewall, Reeve, & Day, 1988).

Participants

Twenty-four undergraduate students participated in the study and were compensated for their time. Their mean age was 21.2 ($SD = 2.20$. min = 18, max = 27). Fourteen participants were Caucasian, seven were Asian, one was Latino, one was Indian, and one was African American. All participants were currently enrolled in college or had received college degrees. Participants were randomly assigned with the constraint that there were six subjects of each gender in each condition.

Procedure

The procedure was identical to Experiment 1, except that all participants were able to review their performance and during each phase participants practiced the move three times as opposed to four in the previous study. Also, the method used to fast forward and rewind in the review phase was a slightly different technological procedure in that the participant using the keyboard to control the viewpoint or alternatively instructing the experimenter how to control the viewpoint.

Results and Discussion

Subjective learning. We computed a similar Positivity Score as in Experiment 1 ($M = 3.83$, $SD = 1.19$, min = 2.82, max = 4.88, Cronbach's

$\alpha = .87$). To test hypotheses between groups, we ran a t test with Media as the discreet independent variable and Positivity Score as the continuous dependent variable. All assumptions required for the ANOVA were met. The difference was not significant, $t(22) = .26$, $p < .80$, with video receiving a mean of 2.31 ($SD = .44$) and VR receiving a mean of 2.37 ($SD = .65$). To further explore the data, we also broke down the subject groups by gender, and analyzed recomputed the dependent variable by various factor analyses. None of the analyses approached statistical significance.

Objective learning. Videos were coded in the exact same way as in Experiment 1. Each measure was rated on a 7-point scale with lower numbers indicating worse performance. For each of the participants, there were three recorded phases with three moves each for a total of 9 videos per participant. Of all total videos, there were 4 missing or corrupted videos and 11 videos where the participant chose not to perform the task. The coder interreliability was moderate (Cronbach's $\alpha = .77$).[3] We then averaged the ratings from the two coders such that for a given video we had only 10 scores that indicated the mean of the two coders. We then averaged the scores into a single learning measure for each video ($M = 2.90$, $SD = 1.06$, min = 1.37, max = 5.43, Cronbach's $\alpha = .98$). Table 1 indicates the scores by condition and phase. Participants in the virtual condition consistently outperformed participants in the video condition, especially during the crucial trial of testing which was our strongest measure of actual learning.

GENERAL DISCUSSION

Summary

The current studies examined the affordances of using VR assist in the learning of physical tasks. Experiment 1 demonstrated an advantage for VR over video in terms of subjective ratings. Experiment 2 demonstrated

TABLE 1 The Mean and Standard Deviation Values for Participants Ratings in Each Condition Across All Tests.

	Condition		Significance	
	Video (*SD*)	VR (*SD*)	t	p
Training 1	2.06 (0.60)	2.76 (1.22)	1.78	0.05
Training 2	2.48 (0.91)	3.35 (1.21)	1.97	0.03
Testing	2.96 (0.88)	3.78 (1.52)	1.63	0.06
Mean Score	2.50 (0.74)	3.33 (1.28)	1.86	0.04

The p value (one tailed due to our a priori, directional prediction) and t scores show students learned better in the VR condition. All assumptions required for the t tests were met.

objective improvements in learning from the VR condition compared to the video condition.

While we did not predict this disparate pattern of inconsistent findings across dependent variables, the dissociation between self-report and behavioral findings is not uncommon. In the current studies, inserting the virtual mirror into the VR condition, while increasing interactivity and causing better two-way communication between the simulation and the user, may not have been a pleasant experience for the user. In other words, learning can sometimes be like "bad medicine"; when the learning process is more difficult in the short-term, the results can be more productive in the long run (Schwartz, Varma, & Martin, in press). Moreover, many educational theorists point to the "tacit" nature of learning (Polanyi, 1967), in which there is a strong dissociation between explicit and implicit learning processes. Finally, other studies have also demonstrated this dissociation of effectiveness and popularity in virtual simulations—maximizing interactivity is not always well received in self-report, even when it improves performance (Bailenson, Beall, Blascovich, Loomis, & Turk, 2005). In fact, there is ample research dedicated to the discordance among self report measures and behavioral measures when measuring behavior in virtual reality (Bailenson et al., 2004, Bailenson, Swinth, et al., 2005; Slater, 2005), concluding that neither self-report nor behavioral measures are sufficient, and that only by examining a host of measures can one assess virtual behavior. While doing so does not always offer the clearest picture, it does allow a more thorough examination of the research area.

Limitations and Future Work

On a theoretical level, one of the major limitations of the current set of studies is the confounding of multiple factors. The goal of the current work was to attempt to demonstrate an advantage in learning for VR over other types of traditional learning media such as video. Given that goal, we manipulated variables by maximizing differences between conditions. As a consequence, we do not know which extra channels contributed are responsible for the effect. In Experiment 1, we did not find any effect of the ability to review; in Experiment 2, we demonstrated learning advantages when the ability to review was held constant, so the contributing affordance was likely not review. In both studies, the VR condition was seen stereoscopically, while the video condition was not. Thus, we confounded two additional channels of information—stereovision with the ability to see oneself in the third person in real-time. However, given that Experiment 2, which allowed participants to see themselves in real-time from *two* angles, showed actual learning differences while the first study did not would leave us to believe it is the ability to see oneself in real-time from multiple angles which increases learning.

Furthermore, in Experiment 2, we did not include a condition in which the video participants also got to see themselves in a mirror. Given that many workout gyms and training facilities have mirrors, this would have been a useful control condition. Consequently, we may have set the scales in favor of the VR condition. However, on this first initial pass we wanted to ensure that we could actually show differences between the two media conditions, especially given that previous work failed to find any differences at all (e.g., Chua et al., 2003). In future work, we plan on further parsing out these differences.

Moreover, that the objective measure was computed by coders who were not themselves tai chi experts could be problematic, as their assessment of the quality of the moves may not have been highly accurate. Consequently, there may be noise in the data that is keeping the effect size smaller than it would be if the coders were better versed in Tai Chi. In addition, participants in the video condition wore the same glasses as in the VR condition in order to maintain experimental control. However, given that normal watchers of videos do not wear cumbersome glasses, this does somewhat reduce the ability to generalize from the study. If had we found consistent differences both in subjective and objective measures across the two studies, this would be better evidence for our hypotheses than finding subjective differences in the one study and objective differences in the other.

On a technological level, one of the major difficulties with this study was the limitations of current image-based rendering systems. The reconstruction quality of the Tai Chi teacher's image was the main complaint among participants during debriefing. Participants found it difficult to see many of the smaller movements, such as hand position, in the reconstructed image of the teacher, which may have contributed to generally lower objective performance scores. In addition, in the VR condition, there was a delay between participants' movements and the system's reconstruction of their movements, which caused some frustration for participants who were trying to compare their own image to the instructor's. The system has this latency due to network transfer speeds and processing times. Finally, the current system has a variable frame rate, with a maximum of 10 frames per second. This frame rate is slow compared to standards such as television, which is shown at 30 frames per second. The slower frame rate creates more jagged movements. These issues will mostly likely lessen as computer technology continues to improve.

One suggestion for future work is to provide automated feedback on subjects' performance from the system. An algorithm could be implemented that could compare the participant's movements to the instructor's and point out errors, either visually, verbally or during the review phase. In addition, in the current studies, participants had to look forward at the projection screen during the entire experiment in order to see the instructor. A more immersive display, such as an HMD or a projection system that surrounds the participant like a CAVE (Cruz-Neira, Sandin, & DeFanti, 1993), would allow participants to turn their heads, possibly increasing mobility and degree of immersion, which could both improve learning.

Conclusions

In the current studies, we have demonstrated the incremental effect of interactivity cues on learning. Moreover, we have provided an encouraging glimpse of the potential of using immersive VR for teaching physical actions—physical therapy, choreography, training work applications, and myriad other applications. As interactivity and realism of our environment increases, we hope to bridge the gap between learning from a virtual teacher in a virtual environment and learning from face-to-face interaction with a real teacher. In the current work, we have demonstrated that immersive VR provides better learning of physical movements than a two-dimensional video. As technology and our understanding of how to leverage the interactive aspects of that technology improves, we should see larger gains in learning from VR.

NOTES

1. Portions of Experiment 2 were presented at a conference previously (Patel & Colleagues, 2006).
2. Several subjects experienced technical difficulties while participating in the experiment, including computer freezes and network failures. These subjects were removed from the objective performance analysis, as their learning was disrupted by unplanned breaks. In total, nine subjects were removed.
3. The reliability was initially low (Cronbach's α =.55). After exploratory analysis, we found that three new coding questions in particular that we added to the second study were causing low reliability (ones which asked about the general impression of the moves as opposed to the step-by-step accuracy). Consequently, we removed the three questions and examined the remaining 10 to compute a score similar to the first study. The reliability improved to 0.77. We then computed the reported *t* tests with both all 13 questions and only the 10 questions. The patterns of statistical significance of the hypotheses did not change.

REFERENCES

Authors. (2006). The effects of fully immersive virtual reality on the learning of physical tasks. *International Workshop on Presence*.

Bailenson, J. N., Aharoni, E., Beall, A. C., Guadagno, R. E., Dimov, A., & Blascovich, J. (2004). Comparing behavioral and self-report measures of embodied agents' social presence in immersive virtual environments. *Proceedings of the 7th Annual International Workshop on PRESENCE*, October 13–15, Valencia, Spain.

Bailenson, J. N., Beall, A. C., Blascovich, J., Loomis, J., & Turk, M. (2005). Transformed social interaction, augmented gaze, and social influence in immersive virtual environments. *Human Communication Research, 31,* 511–537.

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin, 29*, 1–15.

Bailenson, J. N., Swinth, K. R., Hoyt, C. L., Persky, S., Dimov, A., & Blascovich, J. (2005). The independent and interactive effects of embodied agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *PRESENCE: Teleoperators and Virtual Environments, 14,* 379–393.

Bailenson, J. N., & Yee, N. (2005). Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science, 16,* 814–819.

Bailenson, J. N., Yee, N., Blascovich, J., Beall, A. C., Lundblad, N., & Jin, M. (2008). The use of immersive virtual reality in the learning sciences: Digital transformations of teachers, students, and social context. *The Journal of the Learning Sciences, 17*, 102–141.

Becker, D., & Pentland, A. (1996). Using a virtual environment to teach cancer patients Tai Chi. *Proceedings of International Conference on Automatic Face and Gesture Recognition*.

Berkley, J., Turkiyyah, G., Berg, D., Ganter, M., & Weghorst, S. (2004). Real-time finite element modeling for surgery simulation: An application to virtual suturing. *IEEE Transactions on Visualization and Computer Graphics*, *10*(3), 314–325.

Blascovich, J., Loomis, J., Beall, A., Swinth, K., Hoyt, C., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry, 13*, 103–124.

Burgoon, J. K., Bonito, J. A., Bengtsson, B., Cederberg, C., Lundberg, M., & Allspach, L. (2000). Interactivity in human–computer interaction: A study of credibility, understanding and influence. *Computers in Human Behaviour 16*, 553–574.

Cappella, J., & Pelachaud, C. (2002). Rules for responsive robots: Using human interactions to build virtual interactions. In A. L. Vangelisti, H. T. Reis, & M. Fitzpatrick (Eds.), *Stability and change in relationships* (pp. 325–354). Cambridge, MA: Cambridge University Press.

Chi, E. H., Song, J., & Corbin, G. (2004). "Killer App" of wearable computing: wireless force sensing body protectors for martial arts. *Proceedings of UIST 2004*, 277–285. Killington, VT: IEEE.

Chua, P. T., Crivella, R., Daly, B., Hu, N., Schaaf, R., Ventura, D., et al. (2003). Training for physical tasks in virtual reality environments: Tai chi. *Proceedings of IEEE Virtual Reality*, 87–94.

Cruz-Neira, C., Sandin, D. J., & DeFanti, T. A. (1993). Surroundscreen projection-based virtual reality: The design and implementation of the cave. *Proceedings of SIGGRAPH '93: Computer Graphics*, 135–142.

Hays, R., Jacobs, J., Prince, C., & Salas, E. (1992). Flight simulator training effectiveness: A meta-analysis. *Military Psychology*, *4*(2), 63–74.

Holden, M. K. (2005). Virtual environments for motor rehabilitation: Review. *Cyberpsychology & Behavior*, *8*(3), 187–211.

Hoyt, C. & Blascovich, J. (2003). Transformational and transactional leadership in virtual and physical environments. *Small Group Research, 34,* 678–715.

Kalyanaraman, S., & Sundar, S. S. (2006). The psychological appeal of personalized online content in Web portals: Does customization affect attitudes and behavior? *Journal of Communication, 56*, 110–132.

Kiousis, S. (2002). Interactivity: A concept explication. *New Media & Society, 4(3),* 355–383.

Lanier, J. (2001, April). Virtually there. *Scientific American*, 66–75.

Lee, K. M., & Nass, C. (2005). Social-psychological origins of feelings of presence: Creating social presence with machine-generated voices. *Media Psychology, 7,* 31–45.

Liu, Y., & Shrum, L. J. (2003). What is interactivity and is it always such a good thing?: Implications of definition, person and situation for the influence of interactivity on advertising effectiveness. *Journal of Advertising, 31* (4), 53–64.

Loomis, J. M., Beall, A. C., Macuga, K. L., Kelly, J. W., & Smith, R. S. (2006). Visual control of action without retinal optic flow. *Psychological Science*, *17*, 214–221.

Newhagen, J., Cordes, J., & Levy, M. (1995). Nightly@nbc.com: Audience scope and the perception of interactivity in viewer mail on the Internet. *Journal of Communication*, *45*, 164–175.

Pantelidis, V. S. (1993). Virtual reality in the classroom. In J. Hirschbuhl (Ed.), *Computer studies: Computers in education* (pp. 175–179). Guilford, CT: The Dushkin Publishing Group. (Reprinted in *Educational Technology*, *33*(4), 23–27).

Patel, K., Bailenson, J. N., Hack-Jung, S., Diankov, R., & Bajcsy, R. (2006, August). The effects of fully immersive virtual reality on the learning of physical tasks. *Proceedings of PRESENCE 2006: The 9th Annual International Workshop on Presence*. August 24–26, Cleveland, Ohio.

Polanyi, M. (1967). *The tacit dimension*. New York: Anchor Books.

Persksy, S., & Blascovich, J. (2008). Immersive virtual video game play and presence: Influences on aggressive feelings and behavior. *Presence: Teleoperators and Virtual Environments*, *17*, 57–72.

Rafaeli, S. (1988). Interactivity: From new media to communication. In R. P. Hawkins, J. M. Wiemann, & S. Pingree (Eds.), Adv*ancing communication science: Merging mass and interpersonal processes* (pp. 110–134). Beverly Hills, CA: Sage.

Rizzo, A., & Kim, G. J. (2005). A SWOT analysis of the field of virtual-reality rehabilitation and therapy, *Presence: Teleoperators and Virtual Environments*, *14*(2), 119–146.

Schwartz, D. L., Varma, S., & Martin, L. (in press). Dynamic transfer and innovation. In S. Vosniadou (Ed.), *Handbook of conceptual change*. Mahwah, NJ.

Sewall, L. P., Reeve, G., & Day, R. A. (1988). Effect of concurrent visual feedback on acquisition of a weightlifting skill. *Perceptual Motor Skills, 67*, 715–718.

Seymour, N. E. (2005). Virtual reality in general surgical training. *European Surgery*, *37*(5), 298–303.

Seymour, N. E., Gallagher, A. G., Roman, S. A., et al. (2002). Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of Surgery*, *236*(4), 458–463.

Skalski, P., & Tamborini, R. (2007). The role of social presence in interactive agent-based persuasion. *Media Psychology, 10(3)*, 385–413.

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*(1): e39. doi:10.1371/journal.pone.0000039.

Slater, M. (2004) How colorful was your day? Why questionnaires cannot assess presence in virtual environments, *Presence: Teleoperators and Virtual Environments 13*(4), 484–493.

Smith, A. W. B., & Lovell, B. C. (2003). Autonomous sports training from visual cues. In *Proceedings of the Third Australian and New Zealand Conference on Intelligent Information Systems* (pp. 279–284). December 10–12, 2003, Sydney, Australia.

Snyder, S. S., & Nass, C. (2000). Source orientation in human-computer interaction: Programmer, networker, or independent social actor. Communication Research, *27*(6), 683–703.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, *42*(4), 73–93.

Trondsen, E., & Vickery, K. (1997). Learning on demand. *Journal of Knowledge Management*, *1*(3), 169–180.

Weiss, P. L., Rand, D., Katz, N., & Kizony, R. (2004). Video capture virtual reality as a flexible and effective rehabilitation tool. *Journal of Neuroengineering & Rehabilitation*, *1*(1), 12.

Wu, G. (1999). Perceived interactivity and attitude toward website. In M. S. Roberts (Ed.), *Proceedings from the 1999 American Academy of Advertising* (pp. 254–262). Gainesville, FL: University of Florida.

Yang, Z., Liang, J., Yu, B., Cui, Y., Nahrstedt, K., Jung, S. H., & Bajcsy, R. (2005, December). TEEVE: Tele-immersive environment for everybody. In *Proceedings of IEEE International Symposium of Multimedia*, (pp. 112–119). December 12–14, 2005, Irvine, CA.

APPENDIX

Mean and Standard Deviation for Questions in Experiment 1

	3D review (N = 11)		3D no review (N = 10)		2D review (N = 10)		2D no review (N = 9)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1. To what extent did you have the sense that you were in the same place as the instructor? (from *a very small extent* [1] to *a very large extent* [5])	2.82	1.17	2.90	1.29	2.00	0.82	2.40	0.84
2. How would you rate your awareness of the instructor's intentions/wishes in this task? (from *low awareness* [1] to *high awareness* [5])	2.91	1.22	3.10	1.10	2.20	0.79	2.90	0.74
3. Did you find this task pleasant or unpleasant? (from *very pleasant* [1] to *very unpleasant* [5])	3.82	0.98	3.80	0.79	3.50	0.71	3.70	0.67
4. Did you experience this task as something that you did together/jointly with the instructor, or as something you did on your own/separately)? (from *a very large extent on my own* [1] to *a very large extent together* [5])	3.64	1.12	3.70	1.25	2.80	1.03	2.70	0.95
5. How easy or difficult was this task? (from *very difficult* [1] to *very easy* [5])	2.18	0.75	2.80	1.14	2.50	0.53	2.80	0.92
6. How easy or difficult was it to move around in the environment? (from *difficult* [1] to *very easy* [5])	4.27	0.65	4.50	0.71	3.60	1.07	4.40	0.52
7. How personal was your experience in the learning environment? (from *personal* [1] to *impersonal* [5])	3.36	0.92	3.10	0.99	2.30	0.95	2.90	1.29
8. How social was your experience in the learning environment? (from *social* [1] to *antisocial* [5])	3.09	0.70	3.00	1.33	2.10	0.88	2.50	1.18
9. How lively was your experience in the learning environment? (from *lively* [1] to *lifeless* [5])	3.64	1.21	3.90	0.57	2.80	1.03	3.10	0.99
10. How pleasant was your experience in the learning environment? (from *pleasant* [1] to *unpleasant* [5])	3.82	1.25	4.10	0.88	3.50	0.85	3.80	1.14
11. Did you find the instructor (the Tai Chi teacher) to be close or distant? (from *close* [1] to *distant* [5])	3.10	0.88	2.80	1.55	2.30	1.34	2.30	0.95
12. Did you find the instructor (the Tai Chi teacher) to be responsive or unresponsive? (from *responsive* [1] to *unresponsive* [5])	2.00	0.77	2.40	1.58	1.60	0.70	2.50	0.85
13. Did you find the instructor (the Tai Chi teacher) to be active or passive? (from *active* [1] to *passive* [5])	3.18	1.40	3.00	1.41	2.00	1.15	2.10	1.45
14. Did you find the instructor (the Tai Chi teacher) to be warm or cold? (from *warm* [1] to *cold* [5])	3.00	0.89	3.00	1.25	2.50	1.18	3.10	0.74
15. Did you find the instructor (the Tai Chi teacher) to be helpful or unhelpful? (from helpful [1] to unhelpful [5])	3.91	0.83	3.60	1.26	2.90	0.88	3.20	1.23
16. Did you find the instructor (the Tai Chi teacher) to be realistic or fake? (from *realistic* [1] to *fake* [5])	3.36	1.21	3.70	1.42	2.40	1.51	2.00	0.67
17. Did you find the instructor (the Tai Chi teacher) to be an expert or a novice? (from an *expert* [1] to a *novice* [5])	4.91	0.30	4.40	0.70	4.00	0.67	4.50	0.53

Mean and Standard Deviation for Questions in Experiment 2

	3D/review (N = 12)		2D/review (N = 12)	
	Mean	SD	Mean	SD
1. To what extent did you have the sense that you were in the same place as the instructor? (from *a very small extent* [1] to *a very large extent* [5])	2.67	1.23	2.50	1.24
2. How would you rate your awareness of the instructor's intentions/wishes in this task? (from *low awareness* [1] to *high awareness* [5])	2.42	0.79	2.83	1.03
3. Did you find this task pleasant or unpleasant? (from *very pleasant* [1] to *very unpleasant* [5])	3.33	0.65	3.42	0.67
4. Did you experience this task as something that you did together/jointly with the instructor, or as something you did on your own/separately)? (from *a very large extent on my own* [1] to *a very large extent together* [5])	3.17	1.53	3.00	1.35
5. How easy or difficult was this task? (from *very difficult* [1] to *very easy* [5])	2.67	0.78	2.50	1.09
6. How easy or difficult was it to move around in the environment? (from *difficult* [1] to *very easy* [5])	3.75	1.06	3.92	1.00
7. How personal was your experience in the learning environment? (from *personal* [1] to *impersonal* [5])	3.75	1.48	3.42	0.90
8. How social was your experience in the learning environment? (from *social* [1] to *antisocial* [5])	4.08	0.79	3.50	0.80
9. How lively was your experience in the learning environment? (from *lively* [1] to *lifeless* [5])	4.00	0.85	3.50	0.67
10. How pleasant was your experience in the learning environment? (from *pleasant* [1] to *unpleasant* [5])	3.00	1.13	2.67	0.65
11. Did you find the instructor (the Tai Chi teacher) to be close or distant? (from *close* [1] to *distant* [5])	3.83	1.11	3.83	0.94
12. Did you find the instructor (the Tai Chi teacher) to be responsive or unresponsive? (from *responsive* [1] to *unresponsive* [5])	4.50	0.80	4.25	0.97
13. Did you find the instructor (the Tai Chi teacher) to be active or passive? (from *active* [1] to *passive* [5])	3.50	1.24	3.67	1.07
14. Did you find the instructor (the Tai Chi teacher) to be warm or cold? (from *warm* [1] to *cold* [5])	3.75	0.87	3.42	0.90
15. Did you find the instructor (the Tai Chi teacher) to be helpful or unhelpful? (from *helpful* [1] to *unhelpful* [5])	3.17	1.03	3.17	1.19
16. Did you find the instructor (the Tai Chi teacher) to be realistic or fake? (from *realistic* [1] to *fake* [5])	3.58	1.24	3.33	1.15
17. Did you find the instructor (the Tai Chi teacher) to be an expert or a novice? (from *an expert* [1] to *a novice* [5])	2.08	1.08	2.42	0.79