Learning movements from a virtual instructor

by
Jasper LaFortune

A THESIS

submitted to

Oregon State University

University Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in Computer Science
(Honors Scholar)

Presented May 16, 2016
Commencement June 2016

AN ABSTRACT OF THE THESIS OF

Jasper LaFortune for the degree of Honors Baccalaureate of Science in Computer Science presented on May 16, 2016. Title: Learning Movements from a Virtual Instructor .

Abstract approved:

____

Kristen Macuga

We examined the effects of perspective (first person versus third person) and immersion (immersive virtual reality versus nonimmersive video screen) on motor learning in order to assess the format of mental action representations. We also evaluated whether these effects were modulated by experience. Experienced dancers and novices practiced line dances by imitating a virtual instructor and then subsequently had to perform the dances from memory without an instructor present, following a delay. Accuracy for both practice and test trials was video coded. First person perspective formats led to better accuracy, immersive formats did not lead to better accuracy, and experienced dancers were more accurate than novices, but format did not interact with experience. These results suggest that during learning, individuals across experience levels represent complex actions in first person perspective, and that virtual instruction does not require immersion to be effective.

Key Words: Motor learning, action representation, immersive virtual reality, imitation, perspective

Corresponding e-mail address: lafortuj@oregonstate.edu

©Copyright by Jasper LaFortune
May 6, 2016
All Rights Reserved

Learning movements from a virtual instructor

by
Jasper LaFortune

A THESIS

submitted to

Oregon State University

University Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in Computer Science
(Honors Scholar)

Presented May 16, 2016
Commencement June 2016

Honors Baccalaureate of Science in Computer Science project of Jasper LaFortune presented on May 16, 2016.

APPROVED:

Kristen Macuga, Mentor, representing Psychology

Aurora Sherman, Committee Member, representing Psychology

Catherine Dark, Committee Member, representing Dance

Toni Doolen, Dean, University Honors College

I understand that my project will become part of the permanent collection of Oregon State University, University Honors College. My signature below authorizes release of my project to any reader upon request.

Jasper LaFortune, Author

The process of motor learning involves acquiring a new motor skill (Schmidt & Lee, 1988) and creating mental representations of the actions involved (Jeannerod, 1997). Knowing and understanding the conditions that optimize motor learning has been the subject of much research (e.g., Lee, Swinnen, & Serrien, 1994; Toussaint & Blandin, 2010). Such research can lead to an improved understanding of the mental representations that underlie motor learning (Berlucchi & Aglioti, 2010; Schack & Mechsner, 2006). Functional neuroimaging evidence suggests these representations are not general visual representations, but rather are specific to actions (Calvo-Merino, Grèzes, Glaser, Passingham, & Haggard, 2006). Researchers have examined how motor learning affects mental action representations (e.g., Frank, Land, & Schack, 2013), and how these representations in turn affect performance (e.g., Land, Volchenkov, Bläsing, & Schack, 2013). Besides advancing theoretical knowledge, studying the conditions that optimize motor learning and mental action representations can be used to improve rehabilitation of conditions such as stroke (Krakauer, 2006) and to improve coaching in sports (Schack, Essig, Frank, & Koester, 2014). By manipulating the conditions of motor learning, we can indirectly measure aspects of mental action representations (e.g., Toussaint & Blandin, 2010).

The recent rise in popularity of virtual reality (VR) (Reisinger, 2015) has sparked research interest in its potential to improve motor learning (e.g., Ribeiro-Papa et al., 2016). By emphasizing salient or relevant aspects of motor learning, VR systems can be useful for rehabilitation (Holden & Todorov, 2002; Krakauer, 2006). Indeed, research has found that VR systems can make effective treatments for stroke recovery (Henderson, Korner-Bitensky, & Levin, 2007), and for gait training in patients with Parkinson's disease (Mirelman et al., 2011). In addition, VR can be effective for motor training in general (Holden & Todorov, 2002), and also for specific applications, such as dance training (Eaves, Breslin, & Van Schaik, 2011). However, such research often confounds the various potentially beneficial aspects of VR. For example, VR offers the chance to interact realistically with a virtual environment (e.g., Mirelman et al., 2011). It can also offer augmented feedback, giving users additional information to help them correct mistakes (e.g., Eaves et al., 2011). We are interested in precisely which aspects of VR systems, among these and others, actually cause improvements in motor learning.

Previous research has covered many different conditions to optimize motor learning, including practice schedule (Lee et al., 1994), imagery (Toussaint & Blandin, 2010), context (Guadagnoli & Lee, 2004), and feedback (Eaves et al., 2011). We are interested in presentation format: what is the best way to present movements in order to promote motor learning? While research has considered the effects of presentation format on other tasks, such as navigation (Ruddle, Payne, & Jones, 1999; Shelton & McNamara, 2004; Waller, Hunt, & Knapp, 1998), its effects on motor learning have not been well established. In particular, we are interested in investigating the effect of perspective, or the orientation from which the instructor is presented. We are also interested in the effect of immersion, or moving in and feeling surrounded by the same real or virtual environment in which the instructor is presented.

Considerable research has studied the effects of perspective on motor tasks. Someone observing or imitating an action or body position must mentally transform the visual image of the model into their own perspective (Jackson, Meltzoff, & Decety, 2006). This task comes quickly in first person perspective (or "first person"), i.e. the model or instructor facing away from them (Steggemann, Engbert, & Weigelt, 2011). By contrast, it requires an additional spatial transformation in third person perspective (or "third person"), i.e. the model or instructor facing toward them (Jackson et al., 2006; Parsons, 1987). This extra perspective transformation increases latency (Steggemann et al., 2011), which we expect would be detrimental to learning. Further, first and third person perspectives seem to be processed differently by the brain. Observation of actions in first person activates different neural circuits than observation of actions in third person (David et al., 2006; Jackson et al., 2006). Imagery of actions from different perspectives evokes different mental representations (Stevens, 2005) and leads to different motor learning outcomes (White & Hardy, 1995) as well. The substantial degree of overlap between observation, imagery, and imitation of actions in the brain (Grafton & Hamilton, 2007; Grezes & Decety, 2001) suggests that such differences in observation and imagery may correspond to differences in imitation. Indeed, perspective of presentation does affect performance on simple imitation tasks (Vogt, Taylor, & Hopkins, 2003) and motor learning on tasks involving matching an exact position (U. Yang & Kim, 2002).

In spite of this research, the exact effects of perspective on motor learning are not known for several reasons. Although they overlap, the processes of observation, imagery, and imitation of a movement involve different neural circuits (Macuga & Frey, 2012), and thus conditions may affect each of these processes differently. Similarly, although motor performance and motor learning overlap, they are different, and in fact, conditions leading to better performance can lead to worse learning (Magill & Hall, 1990) when they require less cognitive effort (Lee et al., 1994). Additionally, motor learning results do not necessarily generalize across different task characteristics (Guadagnoli & Lee, 2004), and thus results for simple position-matching tasks may not generalize to learning more complex movements. Consequently, we are interested in the effect of perspective on motor learning of complex movements. By filling this gap in knowledge, we hope to inform best practices for motor instruction and to ascertain the perspective of the mental action representation used for motor learning.

The rising popularity of immersive virtual reality systems (Reisinger, 2015) has sparked interest in the effects of immersion (e.g., Gokeler et al., 2014; Lorenz et al., 2015). Immersive displays offer two potential advantages over nonimmersive ones. First, they display a virtual environment in stereoscopic 3D, the same way the real environment is normally seen. This leads the wearer to feel a sense of presence in the virtual environment (Slater, Lotto, Arnold, & Sánchez-Vives, 2009). Second, they are interactive, meaning that the viewpoint moves with the wearer's movements in the real environment. Early research found that immersive VR was ineffective for transfer of motor learning specifically (Kozak, Hancock, Arthur, & Chrysler, 1993) and was not associated with better task performance in general (Nash, Edwards, Thompson, & Barfield, 2000). However, this may have been specific to the tasks involved or the poor quality of early VR technology. Subsequent research showed immersive VR to be at least as effective as real-world training for a simple motor learning task (Rose et al., 2000). Furthermore, immersive VR was shown to be more effective than video instruction for motor learning of a complex action (Bailenson et al., 2008). However, this improvement was likely not due to the immersion itself. Rather, the authors attribute this benefit to the ability to see a stereoscopic virtual representation of oneself. VR systems used for training and rehabilitation stand to benefit from knowing whether these newly popular immersive displays are worth the expense. In addition, separating the effects of immersion from other factors can allow us to evaluate mental action representations. Do these representations include the immersive environment in which learning took place? Do they include the possibly interactive viewpoint from which learning took place? Existing research lacks the experimental control necessary to answer these questions.

Motor learning is a complex process, and as such the best conditions for it are not necessarily universal. Motor learning may depend on such factors as who is doing the learning and the characteristics of the motor task itself (Guadagnoli & Lee, 2004). Considerable research gives reason to believe that motor learning and its related processes and mental representations change with expertise. Motor experts have been shown to use different neural circuits than novices in the observation (Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2005; Calvo-Merino et al., 2006; Cross, Kraemer, Hamilton, Kelley, & Grafton, 2009;

Kirsch & Cross, 2015), imagery (Cross, Hamilton, & Grafton, 2006; Wei & Luo, 2010), and imitation (Vogt et al., 2007) of actions within their motor repertoire as compared to unfamiliar actions. Furthermore, motor expertise adapts mental action representations for better planning and comprehension of actions (J. Yang, 2015). As people gain motor expertise, their mental action representations change (Frank et al., 2013), becoming more hierarchically organized (Land et al., 2013; Schack & Mechsner, 2006). Additionally, certain practice conditions affect motor learning differently for motor experts than novices (Guadagnoli & Lee, 2004). Whether perspective and immersion are among these remains unknown. By filling this gap in knowledge, we hope to inform training applications aimed at expert audiences and suggest how the perspective and immersion of mental action representations change with motor expertise.

The present study investigates the effects of perspective and immersion on motor learning of complex actions. Based on the literature reviewed, we hypothesize that first person and immersive formats lead to improved motor learning over third person and nonimmersive formats, and that these effects are mediated by motor expertise, with experts learning better but being less affected by format than novices. We measure accuracy of line dances taught to experienced and novice dancers in first person, third person, immersive, and nonimmersive formats. We predict best learning in first person and immersive formats, better learning in experienced dancers than in novices, and smaller effects of format on experienced dancers than on novices.

Method

Participants

Informed consent was obtained from participants, who were 28 experienced dancers and 28 novices, 38 female and 18 male, aged 18 to 28, with a mean age of 20.5. Of the experienced dancers, 20 were female and 8 were male, and the mean age was 20.2. Of the novices, 18 were female and 10 were male, and the mean age was 20.7. Data from three additional participants was excluded because we discovered after their participation that they did not meet the selection criteria listed below.

Experienced dancers had taken a dance class for credit or had at least 20 hours of dance experience, as dance classes for credit involve approximately 20 hours of dance experience. Experienced dancers were recruited by in-class announcements in dance classes at Oregon State University that had at least one dance class as a prerequisite. They were offered extra credit in their dance classes for participating. Novices had never taken a dance class for credit and had less than 20 hours of lifetime dance experience. Novices were recruited from psychology classes at Oregon State University using the SONA online signup system, and were offered extra credit in their psychology classes for participating. All participants had normal or corrected-to-normal vision and were fluent in English. The experimental protocol was approved by the local Institutional Review Board (Study #7030).

Design

This experiment followed a 2 × 2 × 4 × 2 mixed factorial design. The first experimental factor was the perspective of instruction, with two levels: first person and third person. The second experimental factor was the immersion of instruction, with two levels: nonimmersive and immersive. The combinations of these two factors, shown in Figure 1, constituted four formats, Format A (first person, immersive), B (third person, immersive), C (first person, nonimmersive), and D (third person, nonimmersive). Participants completed a different dance in each of the four formats, necessitating a third experimental factor, the dance itself, with four dances: Dance W, X, Y, and Z. The formats were paired with different dances in order to diminish practice effects between conditions. The fourth quasi-experimental factor was the participant variable, dance experience, with two levels: experienced and novice.

A balanced Latin square design was employed for counterbalancing to control for potential format order effects. Four format orderings were used, with the property that every format appeared in every ordinal position exactly once, and followed each other format exactly once. The same balanced Latin square design was employed to counterbalance the effects of dance order. The four format orderings were paired with the four dance orderings, for 16 possible stimulus orderings, so that each dance was paired with every condition an equal number of times. Of the 16 stimulus orderings, 12 were completed by two participants from each group, and the remaining four were completed by one participant from each group. Condition orderings were fully counterbalanced, with every condition ordering being completed by seven participants from each group. Dance orderings were partially counterbalanced, with three of the dance orderings being completed by eight participants from each group, and the remaining one completed by four participants from each group.

Materials

The present study took place in the Cognition and Action in Real and Virtual Environments Laboratory, or CARVE Lab, at Oregon State University. The lab contained a 4.3 m by 5.7 m tracking space, in which participants learned and performed the line dances. During immersive conditions, participants wore a Sensics zSight stereoscopic head-mounted display (HMD), with dual 1280 x1024 SXGA OLED displays, one per eye. The graphics were updated at 60 Hz. The HMD, shown in Figure 2, provided an immersive, 60 degree diagonal field of view with 100% binocular overlap that displayed a virtual model of the lab, created with 3D modeling tools (Autodesk Maya) and a virtual reality development platform (Vizard) and rendered by a computer with an Intel Xeon quad-core E5-2603 0 processor clocked at 1.80 GHz and an nVidia GeForce GTX 660 Ti graphics card. An infrared video tracking system (Worldviz PPT-E), shown in Figure 2, and a three-axis orientation sensor (Inertial Labs OS3D) mounted on the HMD updated the participant’s position and orientation, allowing the user to move around the virtual lab model in real time as they physically moved around the actual tracking space. During nonimmersive conditions, participants wore a mockup of the head-mounted display, a modified bicycle helmet with mock cords trailing out the back, shown in Figure 3. The mockup served to minimize non-experimentally manipulated differences between nonimmersive and immersive conditions, such as the tethering, weight, and headmounting aspects of wearing an HMD. This also served the purpose of rendering video coders blind to condition. While wearing the mockup, participants watched the dances on an ASUS VE278Q, 27-inch, widescreen, 1920 x 1080 computer monitor, shown in Figure 1. The monitor was placed 4.25 m from the position where the participants started the dances at a height of 0.95 m, providing a 9 degree diagonal field of view. During all conditions, participants wore a modified eye mask that prevented them looking at their feet and made the HMD seem more immersive by blocking the participant's view of the actual lab. While performing the dances, participants were recorded on a Canon Vixia HF200 1080p video camera. The camera was placed on a tripod 1.2 m to the right of the computer monitor, at a height of 1.1 m.

The virtual lab model, shown in Figure 4, was a 3D model of the actual lab with similar dimensions and features. The instructor was a female virtual avatar approximately 1.8 meters tall, as shown in Figure 4. The experimental script took as input a stimulus file, specific to each participant ID, which defined the format and dance orderings. In all conditions, the instructor began at the front of the virtual lab model. The instructor was shifted in front of the participant, not superimposed over the participant, for better external validity and so that the perspectives varied only in the orientation of the instructor. In first person formats, the instructor faced forward, such that the instructor and participant were facing in the same direction, and performed the movements from that reference point. In third person formats, the instructor faced toward the participant, such that the instructor and participant were facing in opposite directions, and performed the same movements from that reference point. In all formats, participants were asked to perform the same movements on the same feet as the instructor.

I choreographed the dances and animated them from motion capture data. I performed the dances while wearing four infrared video trackers (Worldviz PPT-E), one on each foot and two on my head, and a three-axis orientation sensor (Inertial Labs OS3D). A virtual reality development platform (Vizard) recorded the data from these sensors and used its inverse kinematics engine to animate a virtual avatar approximating my movements. I choreographed the dances to be as similar as possible in difficulty, to minimize main effects of the dances themselves, but as different as possible in combinations of moves, to minimize carryover effects. However, some moves were necessarily repeated between dances. Additionally, I choreographed the dances to be easy enough that novices would not experience a floor effect, but difficult enough that experienced dancers would not experience a ceiling effect. Every dance was performed to an audio track of a voice reciting eight-counts along with a metronome at 90 beats per minute. Each dance lasted 20 counts, with one step per count and no syncopation. Step sheets of each dance can be found in Tables 1 – 4.

Procedure

Participants were run one at a time. When each participant arrived, the experimenter briefly explained the experiment, then asked them to read and sign an informed consent form. The experimenter showed and explained the equipment. First, the experimenter had the participant try on the face mask and explained that its purpose was to prevent participants from looking at their feet. Next, the experimenter helped the participant put on the HMD, and the participant completed a brief pre-training to familiarize themselves with the HMD. The experimenter asked the participant to turn and face the side wall in order to demonstrate that their view of the lab model updated as they moved around the actual lab. The experimenter then asked the participant to walk a few steps to the side and turn their upper body to face the instructor, in order to demonstrate that they would need to turn their upper body similarly while learning the dances. The experimenter then removed the HMD, showed the participant the mockup, and explained that its purpose was to minimize unintended differences between wearing the HMD and watching the monitor. Then, the participant tried on the mockup to ensure it would fit during the experiment.

After explaining the equipment, the experimenter explained how the participant would proceed through the experiment. They explained that the participant would learn four different line dances, and explained each of the formats in which they would learn the dances. They explained that regardless of format, the participant should perform the same movements on the same feet as the instructor. The experimenter then demonstrated stepping with the right foot and turning to the right while facing both away from and toward the participant, and checked for comprehension by seeing that the participant copied these movements correctly.

The experimenter then explained the phases of each dance. They explained that during the initial trial, the participant would stand and watch the dance one time through, imagining following along in their head. They explained that in the practice trials, the participant would follow along with instructor's movements in time, performing the dance. They explained that during the delay phase, the participant would sit and mentally rehearse the dance in their head for two minutes, without moving. Finally, they explained that for the testing trial, the participant would perform the dance from memory to the best of their ability.

The experimenter then explained how each trial would proceed. A voice would count them in, then the dance would start. After 20 counts, or two eight-counts and four more counts, the dance would end, and they could return to the start position to await the next trial. The experimenter then explained briefly how accuracy would be scored. They told the participant to pay attention to the timing, direction, and size of the movements. The experimenter then demonstrated each of these concepts on a simple example move not present in the dance, stepping in place. The experimenter explained that the instructor would not be making any upper body movements, and that the participant should focus on lower body movements and turn their upper body as much as needed in order to see the instructor.

The experimenter then explained that the dances were designed to be difficult to learn, and that mistakes were okay and expected. The experimenter asked participants to do their best, and told them to continue dancing as well as they could if they made a mistake. After answering any questions, the experimenter then began the experiment.

Each dance proceeded as follows. The experimenter placed the HMD or mockup on the participant, ensuring that it was placed correctly and securely. The experimenter then explained that for the initial trial, the participant should stand and watch the dance one time through, imagining following along in their head without actually moving. The experimenter answered any questions, and then the virtual instructor performed the dance one time through. The experimenter then explained that during the following several trials, the participant should follow along with the instructor's movement in time, trying to match the instructor's steps. The experimenter reiterated that the participant should perform the same movement on the same feet as the instructor, then answered any questions. For each practice trial, the experimenter made sure the participant was in the start position, and then the dance started. While the participant performed the dance, the experimenter followed behind them, to manage the cable as well as for safety purposes. Another experimenter recorded video of the participant from the neck down. After all six practice trials, the HMD/mockup and mask were removed, and the experimenter told the participant to take a seat and mentally rehearse the dance in their head for a delay of two minutes prior to completing the testing trial, then started a timer. When the timer expired, the experimenter instructed the participant to perform the dance from memory to the best of their ability. During the testing trial, the participant could hear the counts of music as with previous trials, but could not see the virtual instructor, as they were not wearing the HMD, and the computer monitor was blacked out. The experimenter recorded video of the participant from the neck down for the testing trial. The participant then took a short break and proceeded to the next dance, for which the same observe, practice, test sequence was repeated. When all four dances were complete, the participant completed a questionnaire to determine information regarding demographics and previous dance experience.

Video Coding

We took great care to find a valid, reliable scheme for coding the accuracy of the line dances. Previous research has evaluated dance in various ways. The more subjective among them typically involve an expert panel of judges assigning a score for the entire dance on a Likert scale (e.g., Chatfield, 2009; Radell, Adame, & Cole, 2004). Although this approach may be valid, it gives only a small range of possible scores, and may measure aesthetic factors of the dance unrelated to the accuracy of the movement. More objective measures typically involve comparing each move in the dance against a model version, either automatically, using skeleton tracking (Alexiadis et al., 2011) or by hand, with trained video coders (Loke, Larssen, & Robertson, 2005; Warburton, Wilson, Lynch, & Cuykendall, 2013). These methods have excellent reliability and sensitivity, but operate on the assumption that participants are attempting the same move at the same time as the model. If a participant did the entire dance correctly, except two counts behind the instructor, they would receive the same score as if they did nothing at all. To better suit our purposes, I developed novel measures informed by the existing ones. To better ensure the validity of these novel measures, I interviewed four country line dance experts: two instructors, an instructor/judge, and a competitor/judge. I asked these experts about what kinds of mistakes they would look for in judging line dances, and what kinds of mistakes were most important for accuracy, as well as generally how they would approach the problem of judging line dance accuracy (L. Bryan, personal communication, October 18, 2015; C. Dark, personal communication, October 14, 2015; H. Skredsvig, personal communication, October 18, 2015; R. Buchholz, personal communication, October 18, 2015).

From this advice, I generated two different accuracy measures: an overall score, and a categorized score. The overall score consisted of deciding how many moves (out of ten) were performed acceptably for each dance, according to the coder's subjective judgment of acceptability. The categorized score consisted of giving scores out of two points for each of the 10 moves in each of four categories: order, weight, direction, and timing. I designed these two measures to try to optimize both objectivity and simplicity. The categorized score was quite objective, but also quite complex. The overall score was more subjective, but simpler.

In addition to myself, two other volunteers coded the videos for accuracy: one naïve coder with choreographed dance experience, and one naïve coder without choreographed dance experience. Both volunteer coders were naïve as to the purposes of the experiment. Though I was privy to the experimental hypotheses, I was blind to condition. I trained the other coders individually. First, I taught the coders each of the line dances so that they could perform the dances themselves. I also instructed coders to follow along to videos of the correct dances before they coded them. I also provided coders with a step sheet of each dance detailing every move, as shown in Tables 1 – 4. I walked the coders through the step-by-step coding protocol, then demonstrated how to apply the protocol to several example videos from pilot participants. Coders then coded three pilot participants on their own, checking that that their answers matched mine. After I was confident that the coders understood the coding scheme, each one coded the same random subsample of 21 participants, and we performed an interrater reliability test on this subsample. We chose to use Krippendorff's alpha, because it is a statistically normalized scale, can handle an arbitrary number of coders, and can handle ratio scale data (Hayes & Krippendorff, 2007). We found overall scores to be sufficiently reliable for tentative conclusions, at $\alpha = .7805$. However, we found overall scores on the last practice trial and testing trial to be sufficiently reliable for solid conclusions, at $\alpha = .8804$ and $\alpha = .8334$, respectively. We also found total categorized scores to be sufficiently reliable for solid conclusions, at $\alpha = 0.8232$. We therefore only used the last practice and testing trials in the testing of our hypotheses in order to be able to draw solid conclusions. After testing reliability, the naïve coder without choreographed dance experience proceeded to code the remaining participants. Beyond checking reliability, we did not include the results from the other two coders in our analysis.

For each of the two sets of participants (the reliability set and the remainder), coders proceeded in the following order. First, we coded the overall scores of all six practice trials and the testing trial on Dance W for all participants. Then, we did the same for Dance X, then Dance Y, then Dance Z. Next, we coded the categorized score of the testing trial on Dance W for all participants. Then, we did the same for Dance X, then Dance Y, then Dance Z. Thus, every trial had an overall score, and the testing trials additionally had a categorized score. The categorized score was coded last so that coders' more subjective opinions would not be a mere recall of earlier objective judgments.

For the overall score, coders watched the video closely, pausing and rewinding as necessary, and counted how many moves were performed acceptably. A move was defined as taking two counts of music, so each dance contained exactly ten moves. Coders were instructed to make a subjective judgment of acceptability, but base it only on weight, direction, and timing of the move performed as compared to the correct move. Coders were instructed to watch the video again to make sure they got the same number if they were not sure.

For the categorized score, coders broke down each dance into scores in four categories for each of the ten moves in the dance. We defined a move to be the two steps belonging to a pair of counts in the dance. We chose two steps as the smallest unit of a dance for two reasons. First, a single step is too difficult to identify. For example, if a participant steps forward, are they performing the first step of a walk-walk, or the first step of a step-kick? Considering two steps at a time is sufficient to differentiate these two moves, a necessary step in determining the accuracy. Second, four steps almost always contain two discrete moves. For example, a pivot-turn toe-tap takes four counts, and a participant might perform the pivot-turn perfectly and miss the toe-tap entirely. Considering two steps at a time allows coders to separate these two scores, for a more valid measure of accuracy.

From speaking with expert line dance judges and instructors, I identified four distinct categories of mistakes: order, weight, direction, and timing. An error of order involves forgetting a move or performing a move at the wrong time in the dance. An error of weight involves putting weight on the wrong foot for a step. An error of direction involves facing the wrong direction or moving the foot the wrong direction relative to the body for a step. An error of timing involves stepping off beat or not stepping on beat. For each pair of counts, coders gave a score out of two points in each of the four categories. In each category, two points meant perfect accuracy, one point meant one mistake, and zero points meant two or more mistakes or a "nothing move." Coders scored the categories in three separate sections. First, they determined which moves the participant was performing on which counts in order to determine the order score. Then, they determined the weight and direction scores according to the move the participant performed (not necessarily the intended move). Finally, they scored the timing according to the move the participant performed.

Before scoring the order category, the coder had to decide which moves the participant was performed on each pair of counts. The coder first indicated, for each pair of counts, any move from that dance that was recognizable as starting in those two counts. This could be multiple moves. For example, if a dance contained a move and that move's mirror image, and a participant performed the mirror image in place of the original, it could be recognized as either move. If no move from the present dance was recognizable, the coder indicated a “nothing move.”

By necessity, coders made a subjective judgment of whether a move was recognizable. A move was operationally defined to be recognizable if a verbal description of the participant’s movements matched a verbal description of the move with direction removed. The removal of direction was meant to capture the possibility of performing a mirror image of the move. In deciding recognizability, coders were instructed to ignore timing (other than assigning the move to a particular pair of counts). Coders were allowed to mute the video at this time in order to facilitate ignoring timing.

Once the coder had indicated all possible moves for the whole dance, they resolved discrepancies according to the following rules, in order. 1) If the correct move for a pair of counts was indicated as possible, choose that one. 2) A move can be performed at most once. If a move was possibly performed more than once, choose the time it was performed best. 3) Exactly one move must be assigned to each pair of counts. If more than one possible move was indicated for a pair of counts, choose the one that was closest to what the participant did. If no possible move was indicated, consider it a nothing move. 4) If, after these rules have been applied, there is still ambiguity, make a subjective judgment of what the participant intended to do.

Once the coder resolved any discrepancies, they derived the order score out of two points. If the recognized move was the intended move, the coder assigned two points (even if the recognized move was performed incorrectly, because that would be reflected in other categories). If the recognized move was a different move from the same dance, the coder assigned one point. If there was no move recognized from this dance, the coder assigned zero points to all categories for that pair of counts.

Next, the coder assigned weight and direction scores to each pair of counts according the recognized move (not necessarily the intended move). Coders were provided with step sheets, shown in Tables 1 – 4, that identified which foot should have weight, which direction the free foot should move, and which direction the participant should face for both steps of every move. Coders compared the participant's actual steps with the step sheet and counted the mistakes for each move. Finally, the coder assigned timing scores to each pair of counts according to the recognized move. Coders counted each time a participant did not step on beat or stepped off beat as one mistake. This resulted in a score out of 20 points for each category, for a total score out of 80 points. The sums were computed automatically, and not shown to coders, so that they would not be influenced by the total scores.

Results

We hypothesized that both experienced and novice dancers would learn better in first person than in third person and better in immersive than nonimmersive formats, but that experienced dancers would learn better and be less affected by both formats than novices. We tested each of these hypotheses with three separate 2 × 2 × 2 (Perspective [first person, third person] × Immersion [immersive, nonimmersive] × Experience [experienced, novice]) analyses of variance (ANOVAs). We performed one such ANOVA for each of the following measures coded by a naïve coder: last practice trial overall score (out of 10), testing trial overall score (out of 10), and testing trial total categorized score (out of 80). Thus, we analyzed one measure of performance (last practice trial) and two measures of learning (testing trial overall and total categorized scores). We also ran ANOVAs on each category of the categorized score, and found that all four category scores yielded results consistent with the total categorized score. No significant effects of dance order or condition order were found.

We hypothesized that first person perspective formats would improve performance and learning over third person perspective formats. We found a significant effect of perspective on all measures, with first person perspective formats resulting in better performance and learning than third person perspective formats. As seen in Figure 5, first person perspective formats (M = 7.839, SD = 1.325) were more accurate than third person perspective formats (M = 6.393, SD = 2.499) on the last practice trial, $F(1, 54) = 24.28$, $p < .001$, $\eta_p^2 = .310$. As seen in Figure 6, first person perspective formats (M = 6.027, SD = 2.291) were more accurate than third person perspective formats (M = 4.580, SD = 2.425) on the testing trial overall score, $F(1, 54) = 29.25$, $p < .001$, $\eta_p^2 = .351$, and first person perspective formats (M = 56.554, SD = 14.384) were more accurate than third person perspective formats (M = 47.804, SD = 15.962) on the testing trial categorized score, $F(1, 54) = 26.18$, $p < .001$, $\eta_p^2 = .326$, as seen in Figure 7.

We hypothesized that immersive formats would improve performance and learning over nonimmersive formats. We found no significant main effect of immersion on performance. However, we found significant main effects of immersion for one of the two measures, with immersive formats less accurate than nonimmersive formats. As seen in Figure 5, immersive formats (M = 7.080, SD = 1.909) were not significantly more or less accurate than nonimmersive formats (M = 7.152, SD = 1.839) on the last practice trial, $F(1, 54) = 0.10$, $p = .758$, $\eta_p^2 = .002$. As seen in Figure 6, immersive formats (M = 5.080, SD = 2.272) not significantly more or less accurate than nonimmersive formats (M = 5.527, SD = 2.383) on the testing trial overall score, $F(1, 54) = 3.23$, $p = .078$, $\eta_p^2 = .056$. Immersive formats (M = 50.250, SD = 15.150) were significantly less accurate than nonimmersive formats (M = 54.107, SD = 14.574) on the testing trial categorized score, $F(1, 54) = 6.95$, $p = .011$, $\eta_p^2 = .114$, as seen in Figure 7.

We hypothesized that experienced dancers would learn and perform better than novices. We found a significant main effect of experience on all measures, with experienced dancers performing better than novices. Experienced dancers (M = 7.759, SD = 1.534) were more accurate than novices (M = 6.473, SD = 1.558) on the last practice trial, $F(1, 54) = 9.68$, $p = .003$, $\eta_p^2 = .152$. Experienced dancers (M = 6.348, SD = 2.151) were more accurate than novices (M = 4.259, SD = 1.555) on the testing trial overall score, $F(1, 54) = 17.348$, $p < .001$, $\eta_p^2 = .243$, and experienced dancers (M = 59.768, SD = 12.871) were also more accurate than novices (M = 44.589, SD = 10.142) on the testing trial categorized score, $F(1, 54) = 24.02$, $p < .001$, $\eta_p^2 = .308$. We hypothesized that experienced dancers would be less affected by immersion and perspective than novices. However, we found no significant interactions between any combination of perspective, immersion, and experience on any measure of performance or learning.

In addition to testing our hypotheses with these three ANOVAs, we plotted how mean accuracy, as measured by the overall score, changed with successive trials for each condition, as seen in Figure 8. We then fit each of these to a power curve, and found these to fit the data very closely. Practice curves for first person were shifted up from those for third person, but immersive and nonimmersive curves were nearly indistinguishable. Best-fit equations and goodness of fit values can be seen in Table 5.

Discussion

Perspective

In line with our predictions, we found strong evidence that learning a line dance from a first person perspective, versus a third person perspective, improves the accuracy of both performance and learning. This supports the idea that motor learning of complex actions recruits a first person mental action representation, under the assumption that motor learning is best when presentation format matches the mental action representation used for learning. The effect sizes were considerable, making first person perspective the suggested format for training and rehabilitative applications. This result is consistent with prior findings that a first person perspective improves performance of simple motor tasks over a third person perspective (Vogt et al., 2003), suggesting that the advantage first person perspective presentation likely generalizes across tasks. It is also consistent with the prior finding that motor imagery evokes a motor representation in the first but not the third person perspective (Stevens, 2005). This consistency suggests that the mental action representations used in observation, imagery, and imitation (Grafton & Hamilton, 2007; Grezes & Decety, 2001; Macuga & Frey, 2012) overlap with learning as well. The additional spatial transformation required by the third person (Zacks, Mires, Tversky, & Hazeltine, 2000) decreased not only performance, but also learning. Although this is inconsistent with the finding that motor learning is best when conditions maximize the cognitive effort required (Lee et al., 1994), this effect has been shown not to generalize across all task variables (Guadagnoli & Lee, 2004). Our results indicate that the benefit of increased cognitive effort on motor learning do not generalize across different perspectives, likely because of the different mental representation evoked (Stevens, 2005).

Although the benefit of first person perspective for motor performance and learning was large and robust across measures and across the other variables in the present study, a few caveats should be mentioned that could limit interpretation of these results. First, it is worth noting that when instructors face their students, they could alternatively perform a horizontal mirror image of the movements so that they move in the same direction that the student should. By contrast, the present study required participants to perform the same movements as the instructor from the instructor's perspective, which requires an additional spatial transformation. Consequently, these results should not be interpreted as discouraging horizontal mirroring by instructors. We deliberately designed the stimuli this way, for better experimental control and because the theoretical implications for mental action representation depended on it. Second, although every participant demonstrated understanding before the trials that they were supposed to perform the same movements as the instructor from the instructor's perspective, a small number of participants (fewer than 10) mirrored the instructor until the experimenter corrected them after two or three trials, and a smaller number of those (fewer than five) continued mirroring anyway. Thus, the decreased accuracy of third person for this small minority of participants may have been due to some lack of comprehension rather than because of the different mental representation used. However, the benefit of first person held for the order category of the categorized score. As this category was not affected by mirroring, we are confident that these individuals did not significantly bias results.

Based on these results, we recommend future work investigating variations on the parameters of perspective. Dance instructors often use a variation of third person perspective in which the instructor mirrors movements horizontally, as mentioned above. Previous research has also studied a first person "ghost" perspective, in which a model of the instructor's avatar is superimposed over participants' own avatar (U. Yang & Kim, 2002). Comparing these variations to first person perspective as used in the present study could inform best practices for instructors and coaches.

Immersion

In contrast with our initial predictions, immersive formats were no different for performance and no different or slightly worse for learning than nonimmersive formats. This contradicts the idea that immersive virtual reality is beneficial for motor learning. These results necessitate explanation. First, the testing trial overall score showed no effect of immersion, while the testing trial total categorized score showed a significant effect. Given that the former was near significance, we argue that this inconsistency owes to the greater sensitivity of the latter. Possible overall scores ranged from 0 – 10 points, but the possible categorized scores ranged from 0 – 80 points, making the categorized scores more sensitive to small variations. This helps to explain why the effect of immersion on learning was only observed for one of the two testing measures. Second, we did not find an effect of immersion on performance. This difference likely owes to the additional transfer from learning in the virtual environment to testing in the real environment. In nonimmersive conditions, participants practiced and tested in the same (real) environment. By contrast, in immersive conditions, participants were required to transfer the movements from the virtual environment to the real one for testing. To limit the effects of this difference, the virtual lab model was designed to match the real lab as closely as possible. Still, this difference explains the moderate benefit of nonimmersive format for testing. Additionally, it suggests augmented reality as a preferable alternative to immersive VR, as it circumvents the need for this additional transfer.

The fact that we found no benefit of immersion for performance or learning is surprising. The growing popularity of immersive displays (Reisinger, 2015) would suggest that the interactivity and sense of presence they offer are worth the expense. Indeed, prior research has found immersive VR to be effective for motor learning and rehabilitation (e.g., Bailenson et al., 2008; Gokeler et al., 2014; Rose et al., 2000). However, such research conflated immersion with several other potentially beneficial aspects of VR. For example, VR can provide the benefit of additional visual feedback by showing an individual their own avatar (Bailenson et al., 2008) and how it compares to a model (U. Yang & Kim, 2002). Visual feedback does not require immersion to be beneficial for motor learning (Eaves et al., 2011), suggesting that the benefits of immersive VR may be due to feedback and other factors. We isolated the effects of immersion, and found that immersive displays do not confer a motor learning advantage. Our results indicate that motor training applications do not stand to gain from investing in immersive displays. This is surprising, as the qualities offered by immersive displays seem beneficial at their face. However, our results suggest these qualities are irrelevant to the action representations they evoke.

Several possible limitations on the external validity of the present study provide recommendations for future work. First, the (virtual) environment in which the instructor taught the dances closely matched the (real) environment in which participants learned the dances for nonimmersive conditions. This is not generally true when learning movements from a video. Immersion could be beneficial when it affords the chance to learn a movement in an environment similar (identical) to that of the instructor. We recommend future work investigating the effects of immersion on motor learning when the virtual environment does not match the real one. Second, the dances involved in the present study involved minimal turning, each facing only two walls. Immersion might be more beneficial for learning movements involving more turns, which require keeping track of which direction the instructor is facing. Finally, the head-mounted display itself may have been perceived as somewhat cumbersome. Immersion may become more beneficial for motor learning as future developments enable more lightweight, wireless VR technology. Based on these factors, we limit our interpretation thus: immersive VR, in its current state, does not improve motor learning for dances involving minimal turning.

Experience

In line with our predictions, experienced dancers learned and performed better than novices. This supports the idea that mental action representations are stronger in motor experts than novices. However, in contrast with our predictions, perspective and immersion had the same effects on experienced dancers as novices. Mental action representations do not seem to be more flexible in motor experts than novices. All differences between experienced dancers and novices were large in magnitude and robust across measures and conditions, suggesting that the samples were appropriately selected. We can conclude that the perspective and immersion of motor instruction do not need to be tailored to varying levels of motor experience. We can additionally conclude that whatever differences exist between the mental action representations of motor experts and novices, perspective and immersion do not seem to be among them.

Although we made every effort to control for extraneous variables, the possibility remains that experience was confounded with other factors. First, by necessity, experienced dancers were compensated differently than novices. Whereas we offered experienced dancers extra credit in their dance classes, we offered novices extra credit in their psychology classes. Because these dance classes were offered for fewer credits than the psychology classes, experienced dancers may have been less extrinsically motivated to participate, and more intrinsically motivated to do their best on the task. Still, we took care to make sure that all participants were given extra credit for participation, to minimize this difference as much as possible. Second, also by necessity, experienced dancers were recruited differently than novices. Whereas experienced dancers were recruited via in-class announcements, novices were recruited via an online sign-up system. Signing up thus required more effort from experienced dancers, who as a result may have been more motivated to participate. However, all participants were likely motivated to do well by the mere fact that dancing badly in front of two observers is embarrassing, as we think our participants would confirm.

Based on these results, we recommend future work investigating how experience and other individual differences affect other conditions related to motor learning. We noticed a small number of participants improving between the final practice trial and the testing trial, suggesting a unique strategy or ability during mental rehearsal. This suggests that a comparison of motor learning across mental imagery ability might yield interesting results. We also noticed some participants wanting a chance to practice without the instructor before the testing trial. This begs the question of how individual differences such as expertise interact with effects of practice schedule and mental vs. physical practice.

Practice Curves

In line with theoretical models of motor learning, mean accuracy increased with successive trials in high accordance with the power law of practice (Anderson, 1981). This suggests that our stimuli and coding methods were appropriate for their intention of measuring motor learning. Consistent with results from the last practice trial, practice curves for first person conditions were shifted pronouncedly up from those for third person conditions, but curves for immersive and nonimmersive conditions were nearly identical. This consistency suggests that results from the last trial apply to performance in general on this task. As such, we can conclude that first person is better than third person and immersive formats are no better than nonimmersive formats for motor learning.

Several factors limit the confidence of these conclusions. First, reliability was heterogeneous across trials, with earlier trials less reliable than later ones. As such, the practice curves may make a less accurate model for earlier trials than for later trials. However, they were made using results from a naïve coder, so this effect should not lead to confirmation bias. Second, due to time constraints, we did not code a categorized score for the practice trials. This measure might have had the sensitivity to pick up smaller differences in performance that the overall score could not distinguish. However, the results from the total categorized score were largely consistent with the overall score. As such, we have reason to expect that any effects on performance not detected by the overall score would be small. To be certain, though, we recommend future work using our more sensitive measure on practice trials.

General Discussion

The present study examined the roles of perspective and immersion on motor performance and learning using four video-coded line dances. We developed a novel video coding scheme that was simple enough to be carried out by a trained novice, objective enough to be reliable, efficient enough to be carried out on hundreds of videos, and sensitive enough to detect even small differences in accuracy. We verified that the choreographed dances were similar enough in complexity to eliminate any confounding dance-specific effects, while also different enough in content to prevent any carryover effects. We also managed to make them difficult enough to prevent ceiling effect in experienced dancers, but not so difficult as to create a floor effect (or undue embarrassment) in novices. After checking and eliminating these potential variables as confounds, we were able to assess how different instructional formats influence motor performance and learning for novice as well as experienced dancers.

Several factors mediate the benefits of this methodology. First, we have only shown this coding scheme to be reliable for this particular set of data. It may not be suitable for different dances or types of movement. Further, it may not be suitable for different ranges of performance. It tops out when performance is essentially accurate, even if technique or musicality are lacking. On the other end, it sees poor reliability when accuracy is so low as to make it difficult to discern one move from another. Second, the coding scheme is time-intensive. The categorized score and overall score together demand about six to ten minutes to code about 20 seconds of dancing. Third, it is difficult. The overall score inherently involves binning accuracy of moves at a subjective threshold deemed acceptable. If different coders have different thresholds, they can give widely different overall scores. Despite these limitations, we found this methodology effective for the present study and recommend its use for future work.

The benefits of this research are several. First, prior research on the mental representation of action has primarily used brain imaging methods (e.g., Buccino et al., 2004; Chaminade, Meltzoff, & Decety, 2005) or behavioral measures of how individuals mentally organize actions into a hierarchy (Frank et al., 2013; Schack & Mechsner, 2006). The present study provides a novel behavioral measure of mental action representation format. Second, the use of line dances provides a high degree of external validity to the study of motor learning. Findings about learning of line dances have the potential to generalize to observation, imagery, and imitation of complex movements.

References

Alexiadis, D. S., Kelly, P., Daras, P., O'Connor, N. E., Boubekeur, T., & Moussa, M. B. (2011). *Evaluating a dancer's performance using kinect-based skeleton tracking.* Paper presented at the Proceedings of the 19th ACM international conference on Multimedia.

Anderson, J. R. (1981). *Cognitive skills and their acquisition*: Psychology Press.

Bailenson, J., Patel, K., Nielsen, A., Bajscy, R., Jung, S.-H., & Kurillo, G. (2008). The effect of interactivity on learning physical actions in virtual reality. *Media Psychology, 11*(3), 354-376.

Berlucchi, G., & Aglioti, S. M. (2010). The body in the brain revisited. *Experimental Brain Research, 200*(1), 25-35. doi:10.1007/s00221-009-1970-7

Buccino, G., Vogt, S., Ritzl, A., Fink, G. R., Zilles, K., Freund, H. J., & Rizzolatti, G. (2004). Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. *Neuron, 42*(2), 323-334.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: an FMRI study with expert dancers. *Cerebral Cortex, 15*(8), 1243-1249. doi:10.1093/cercor/bhi007

Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology, 16*(19), 1905-1910. doi:10.1016/j.cub.2006.07.065

Chaminade, T., Meltzoff, A. N., & Decety, J. (2005). An fMRI study of imitation: action representation and body schema. *Neuropsychologia, 43*(1), 115-127. doi:10.1016/j.neuropsychologia.2004.04.026

Chatfield, S. J. (2009). A test for evaluating proficiency in dance. *Journal of Dance Medicine & Science, 13*(4), 108-114.

Cross, E. S., Hamilton, A. F., & Grafton, S. T. (2006). Building a motor simulation de novo: observation of dance by dancers. *Neuroimage, 31*(3), 1257-1267. doi:10.1016/j.neuroimage.2006.01.033

Cross, E. S., Kraemer, D. J., Hamilton, A. F., Kelley, W. M., & Grafton, S. T. (2009). Sensitivity of the action observation network to physical and observational learning. *Cerebral Cortex, 19*(2), 315-326. doi:10.1093/cercor/bhn083

David, N., Bewernick, B. H., Cohen, M. X., Newen, A., Lux, S., Fink, G. R., . . . Vogeley, K. (2006). Neural representations of self versus other: visual-spatial perspective taking and agency in a virtual ball-tossing game. *Journal of Cognitive Neuroscience, 18*(6), 898-910. doi:10.1162/jocn.2006.18.6.898

Eaves, D. L., Breslin, G., & Van Schaik, P. (2011). The short-term effects of real-time virtual reality feedback on motor learning in dance. *Presence: Teleoperators and Virtual Environments, 20*(1), 62-77.

Frank, C., Land, W. M., & Schack, T. (2013). Mental representation and learning: the influence of practice on the development of mental representation structure in complex action. *Psychology of Sport and Exercise, 14*(3), 353-361.

Gokeler, A., Bisschop, M., Myer, G. D., Benjaminse, A., Dijkstra, P. U., van Keeken, H. G., . . . Otten, E. (2014). Immersive virtual reality improves movement patterns in patients after ACL reconstruction: implications for enhanced criteria-based return-to-sport rehabilitation. *Knee Surgery, Sports Traumatology, Arthroscopy*. doi:10.1007/s00167-014-3374-x

Grafton, S. T., & Hamilton, A. F. d. C. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science, 26*(4), 590-616.

Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Human Brain Mapping, 12*(1), 1-19.

Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior, 36*(2), 212-224. doi:10.3200/JMBR.36.2.212-224

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*(1), 77-89.

Henderson, A., Korner-Bitensky, N., & Levin, M. (2007). Virtual reality in stroke rehabilitation: a systematic review of its effectiveness for upper limb motor recovery. *Topics in Stroke Rehabilitation, 14*(2), 52-61.

Holden, M. K., & Todorov, E. (2002). Use of virtual environments in motor learning and rehabilitation. In K. M. Stanney (Ed.), *Handbook of Virtual Environments: Design, Implementation, and Applications* (pp. 999-1026). Cambridge, MA: Lawrence Elbraum Associates.

Jackson, P. L., Meltzoff, A. N., & Decety, J. (2006). Neural circuits involved in imitation and perspective-taking. *Neuroimage, 31*(1), 429-439.

Jeannerod, M. (1997). *The cognitive neuroscience of action*: Blackwell Publishing.

Kirsch, L. P., & Cross, E. S. (2015). Additive Routes to Action Learning: Layering Experience Shapes Engagement of the Action Observation Network. *Cerebral Cortex, 25*(12), 4799-4811. doi:10.1093/cercor/bhv167

Kozak, J., Hancock, P., Arthur, E., & Chrysler, S. (1993). Transfer of training from virtual reality. *Ergonomics, 36*(7), 777-784.

Krakauer, J. W. (2006). Motor learning: its relevance to stroke recovery and neurorehabilitation. *Current Opinion in Neurology, 19*(1), 84-90.

Land, W. M., Volchenkov, D., Bläsing, B. E., & Schack, T. (2013). From action representation to action execution: exploring the links between cognitive and biomechanical levels of motor control. *Frontiers in Computational Neuroscience, 7*(127), 1-14. doi:10.3389/fncom.2013.00127

Lee, T. D., Swinnen, S. P., & Serrien, D. J. (1994). Cognitive effort and motor learning. *Quest, 46*(3), 328-344.

Loke, L., Larssen, A. T., & Robertson, T. (2005). *Labanotation for design of movement-based interaction.* Paper presented at the Proceedings of the second Australasian conference on Interactive entertainment.

Lorenz, M., Busch, M., Rentzos, L., Tscheligi, M., Klimant, P., & Frohlich, P. (2015). *I'm There! The influence of virtual reality and mixed reality environments combined with two different navigation methods on presence.* Paper presented at the Virtual Reality (VR), 2015 IEEE.

Macuga, K. L., & Frey, S. H. (2012). Neural representations involved in observed, imagined, and imitated actions are dissociable and hierarchically organized. *Neuroimage, 59*(3), 2798-2807.

Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human Movement Science, 9*(3), 241-289.

Mirelman, A., Maidan, I., Herman, T., Deutsch, J. E., Giladi, N., & Hausdorff, J. M. (2011). Virtual reality for gait training: can it induce motor learning to enhance complex walking and reduce fall risk in patients with Parkinson's disease? *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 66*(2), 234-240.

Nash, E. B., Edwards, G. W., Thompson, J. A., & Barfield, W. (2000). A review of presence and performance in virtual environments. *International Journal of human-computer Interaction, 12*(1), 1-41.

Parsons, L. M. (1987). Imagined spatial transformation of one's body. *Journal of Experimental Psychology: General, 116*(2), 172-191.

Radell, S. A., Adame, D. D., & Cole, S. P. (2004). The impact of mirrors on body image and classroom performance in female college ballet dancers. *Journal of Dance Medicine & Science, 8*(2), 47-52.

Reisinger, D. (2015). Why virtual reality is about to go mainstream. *Fortune, October*.

Ribeiro-Papa, D. C., Massetti, T., Crocetta, T. B., Menezes, L. D. C. d., Antunes, T. P. C., Bezerra, Í. M. P., & Monteiro, C. B. d. M. (2016). Motor learning through virtual reality in elderly-a systematic review. *Medical Express, 3*(2).

Rose, F. D., Attree, E. A., Brooks, B. M., Parslow, D. M., Penn, P. R., & Ambihaipahan, N. (2000). Training in virtual environments: transfer to real world tasks and equivalence to real task training. *Ergonomics, 43*(4), 494-511. doi:10.1080/001401300184378

Ruddle, R. A., Payne, S. J., & Jones, D. M. (1999). Navigating large-scale virtual environments: what differences occur between helmet-mounted and desk-top displays? *Presence: Teleoperators and Virtual Environments*, *8*(2), 157-168.

Schack, T., Essig, K., Frank, C., & Koester, D. (2014). Mental representation and motor imagery training. *Frontiers in Human Neuroscience, 8*(328), 1-10. doi:10.3389/fnhum.2014.00328

Schack, T., & Mechsner, F. (2006). Representation of motor skills in human long-term memory. *Neuroscience Letters, 391*(3), 77-81. doi:10.1016/j.neulet.2005.10.009

Schmidt, R. A., & Lee, T. (1988). *Motor control and learning*: Human Kinetics.

Shelton, A. L., & McNamara, T. P. (2004). Orientation and perspective dependence in route and survey learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(1), 158-170. doi:10.1037/0278-7393.30.1.158

Slater, M., Lotto, B., Arnold, M. M., & Sánchez-Vives, M. V. (2009). How we experience immersive virtual environments: the concept of presence and its measurement. *Anuario de Psicología, 40*, 193-210.

Steggemann, Y., Engbert, K., & Weigelt, M. (2011). Selective effects of motor expertise in mental body rotation tasks: Comparing object-based and perspective transformations. *Brain and Cognition, 76*(1), 97-105.

Stevens, J. A. (2005). Interference effects demonstrate distinct roles for visual and motor imagery during the mental representation of human action. *Cognition, 95*(3), 329-350.

Toussaint, L., & Blandin, Y. (2010). On the role of imagery modalities on motor learning. *Journal of Sports Sciences, 28*(5), 497-504.

Vogt, S., Buccino, G., Wohlschläger, A. M., Canessa, N., Shah, N. J., Zilles, K., . . . Fink, G. R. (2007). Prefrontal involvement in imitation learning of hand actions: effects of practice and expertise. *Neuroimage, 37*(4), 1371-1383. doi:10.1016/j.neuroimage.2007.07.005

Vogt, S., Taylor, P., & Hopkins, B. (2003). Visuomotor priming by pictures of hand postures: perspective matters. *Neuropsychologia, 41*(8), 941-951.

Waller, D., Hunt, E., & Knapp, D. (1998). The transfer of spatial knowledge in virtual environment training. *Presence: Teleoperators and Virtual Environments, 7*(2), 129-143.

Warburton, E. C., Wilson, M., Lynch, M., & Cuykendall, S. (2013). The cognitive benefits of movement reduction evidence from dance marking. *Psychological Science, 24*(9), 1732-1739.

Wei, G., & Luo, J. (2010). Sport expert's motor imagery: functional imaging of professional motor skills and simple motor skills. *Brain Research, 1341*, 52-62. doi:10.1016/j.brainres.2009.08.014

White, A., & Hardy, L. (1995). Use of different imagery perspectives on the learning and performance of different motor skills. *British Journal of Psychology, 86 (Pt 2)*, 169-180.

Yang, J. (2015). The influence of motor expertise on the brain activity of motor task performance: A meta-analysis of functional magnetic resonance imaging studies. *Cognitive, Affective, and Behavioral Neuroscience, 15*(2), 381-394. doi:10.3758/s13415-014-0329-0

Yang, U., & Kim, G. J. (2002). Implementation and evaluation of “just follow me”: An immersive, VR-based, motion-training system. *Presence: Teleoperators and Virtual Environments, 11*(3), 304-323.

Zacks, J. M., Mires, J., Tversky, B., & Hazeltine, E. (2000). Mental spatial transformations of objects and perspective. *Spatial Cognition and Computation, 2*(4), 315-332.

Tables

Counts	Move	Description	Weight 1st Count	Weight Change	Facing 1st Count	Facing 2nd Count	Feet Movement 1st Count	Feet Movement 2nd Count
First "1, 2"	1	Toe Tap Right	Left	Change	Front	Front	Forward	Together (Back)
First "3, 4"	2	Toe Tap Left	Right	Change	Front	Front	Forward	Together (Back)
First "5, 6"	3	Box Step Forward	Right	Change	Front	Front	Forward	Forward
First "7, 8"	4	Box Step Back	Right	Change	Front	Front	Back	Back
Second "1, 2"	5	Rock Step Back	Right	Change	Front	Front	Back	In Place
Second "3, 4"	6	Pivot Turn to the Left	Right	Change	Front	Left	In Place OR Right	In Place
Second "5, 6"	7	Rock Step Forward	Right	Change	Left	Left	Forward	In Place
Second "7, 8"	8	Walk Back	Right	Change	Left	Left	Back	Back
Third "1, 2"	9	Rock Step Back	Right	Change	Left	Left	Back	In Place
Third "3, 4"	10	Step Kick OR Walk Walk	Right	No Change OR Change	Left	Left	Forward	Forward

Table 1. Step sheet for Dance W. Coders scored order, weight, and direction according to the step sheet.

Counts	Move	Description	Weight 1st Count	Weight Change	Facing 1st Count	Facing 2nd Count	Feet Movement 1st Count	Feet Movement 2nd Count
First "1, 2"	1	Step Touch Right	Right	No Change	Front	Front	Right	Together (Right)
First "3, 4"	2	Step Touch Left	Left	No Change	Front	Front	Left	Together (Left)
First "5, 6"	3	Grapevine Right	Right	Change	Front	Front	Right	Right Behind
First "7, 8"	4	Step Touch Right	Right	No Change	Front	Front	Right	Together (Right)
Second "1, 2"	5	Step Turn Scuff	Left	No Change	Front OR Left	Left	Left OR (Forward IF Facing Left)	Forward
Second "3, 4"	6	Rock Step Forward	Right	Change	Left	Left OR Front	In Place	In Place
Second "5, 6"	7	Step Touch Right	Right	No Change	Front	Front	Right	Together (Right)
Second "7, 8"	8	Toe Tap Left	Right	No Change	Front	Front	Left	Together (Right)
Third "1, 2"	9	Grapevine Left	Left	Change	Front	Front	Left	Left Behind
Third "3, 4"	10	Step Turn Scuff	Left	No Change	Front OR Left	Left	Left OR (Forward IF Facing Left)	Forward

Table 2. Step sheet for Dance X. Coders scored order, weight, and direction according to the step sheet.

Counts	Move	Description	Weight 1st Count	Weight Change	Facing 1st Count	Facing 2nd Count	Feet Movement 1st Count	Feet Movement 2nd Count
First "1, 2"	1	Step Touch Right	Right	No Change	Front	Front	Right	Together (Right)
First "3, 4"	2	Toe Tap Left	Right	Change	Front	Front	Left	Together (Right)
First "5, 6"	3	Step Touch Right	Right	No Change	Front	Front	Right	Together (Right)
First "7, 8"	4	Toe Tap Left	Right	Change	Front	Front	Left	Together (Right)
Second "1, 2"	5	Pivot Turn to the Left	Right	Change	Front	Left	In Place OR Right	In Place
Second "3, 4"	6	Toe Tap Forward Back	Left	No Change	Left	Left	Forward	Back
Second "5, 6"	7	Walk Forward	Right	Change	Left	Left	Forward	Forward
Second "7, 8"	8	Toe Tap Forward Back	Left	No Change	Left	Left	Forward	Back
Third "1, 2"	9	Rock Step Forward	Right	Change	Left	Left OR Front	Forward	In Place
Third "3, 4"	10	Step Touch Right	Right	No Change OR Change	Front	Front	Right	Together (Right)

Table 3. Step sheet for Dance Y. Coders scored order, weight, and direction according to the step sheet.

Counts	Move	Description	Weight 1st Count	Weight Change	Facing 1st Count	Facing 2nd Count	Feet Movement 1st Count	Feet Movement 2nd Count
First "1, 2"	1	Step Kick Forward	Right	No Change	Front	Front	Forward	Forward
First "3, 4"	2	Step Kick Back OR Step Step Back	Left	No Change OR Change	Front	Front	Back	Back
First "5*, 6"	3	Rock Step Forward	Right	Change	Front	Front	Forward	In Place
First "7, 8"	4	Rock Step Back	Right	Change	Front	Front	Back	In Place
Second "1, 2"	5	Grapevine Right	Right	Change	Front	Front	Right	Right Behind
Second "3, 4"	6	Step Touch Right	Right	No Change	Front	Front	Right	Together (Right)
Second "5, 6"	7	Toe Tap Left	Right	Change	Front	Front	Left	Together (Right)
Second "7, 8"	8	Pivot Turn to the Left	Right	Change	Front	Left	In Place OR Right	In Place
Third "1, 2"	9	Walk Forward	Right	Change	Left	Left	Forward	Forward
Third "3, 4"	10	Step Kick	Right	No Change	Left	Left	Forward	Forward

Table 4. Step sheet for Dance Z. Coders scored order, weight, and direction according to the step sheet.

	Equation	Goodness of fit
Immersive First	$y = 5.8392x^{0.1778}$	$R^2 = 0.9723$
Nonimmersive First	$y = 6.3648x^{0.1393}$	$R^2 = 0.9119$
Immersive Third	$y = 3.4976x^{0.3448}$	$R^2 = 0.9928$
Nonimmersive Third	$y = 3.7375x^{0.3374}$	$R^2 = 0.9634$

Table 5. Power-fit mean practice curve equations and goodness of fit values for each condition. Mean accuracy closely followed the power law of practice. Conditions are collapsed across experience.

Figures

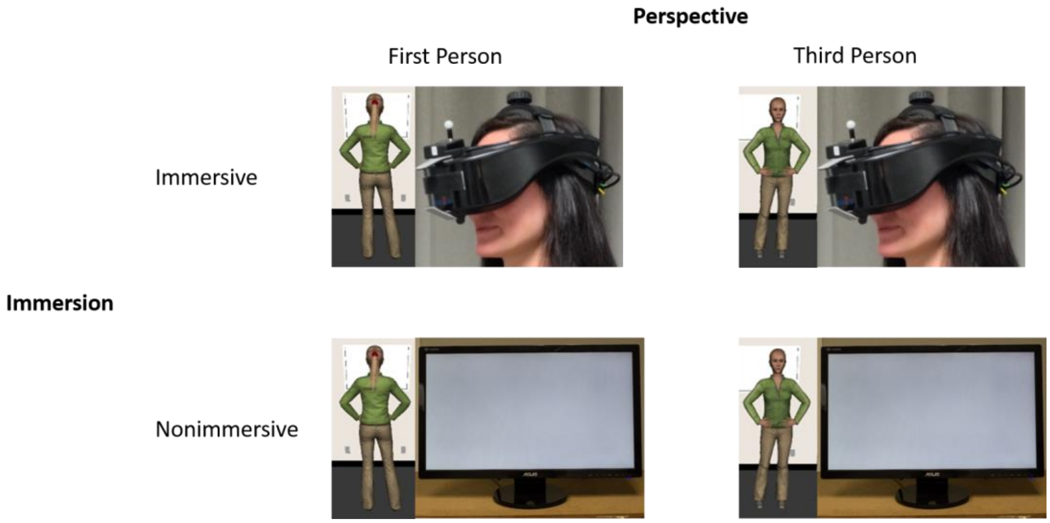

Figure 1. Confusion matrix of instruction formats used for the experimental design. Perspective (first person versus third person) was crossed with immersion (immersive versus nonimmersive) to make four formats. In first person, the instructor faced forward; in third person, the instructor faced toward the participant; in immersive formats, the instructor was presented on an HMD; in nonimmersive formats, the instructor was presented on a computer monitor.

Figure 2. The Sensics zSight HMD worn by participants for immersive conditions, along with the infrared trackers mounted on the HMD. These gave its position and orientation, allowing the wearer to move around the virtual lab model as they moved around the actual tracking space. A researcher followed behind the participant, managing the cable and ensuring safety.

Figure 3. The mockup HMD worn by participants for nonimmersive conditions. This controlled for the non-experimentally manipulated differences between the immersive and nonimmersive conditions, including the additional weight on the head and the social presence of a researcher holding the cable, as well as rendering video coders blind to condition.

Figure 4. The virtual instructor avatar in the lab model, shown here in first person. The lab model was designed to match the real lab as closely as possible.

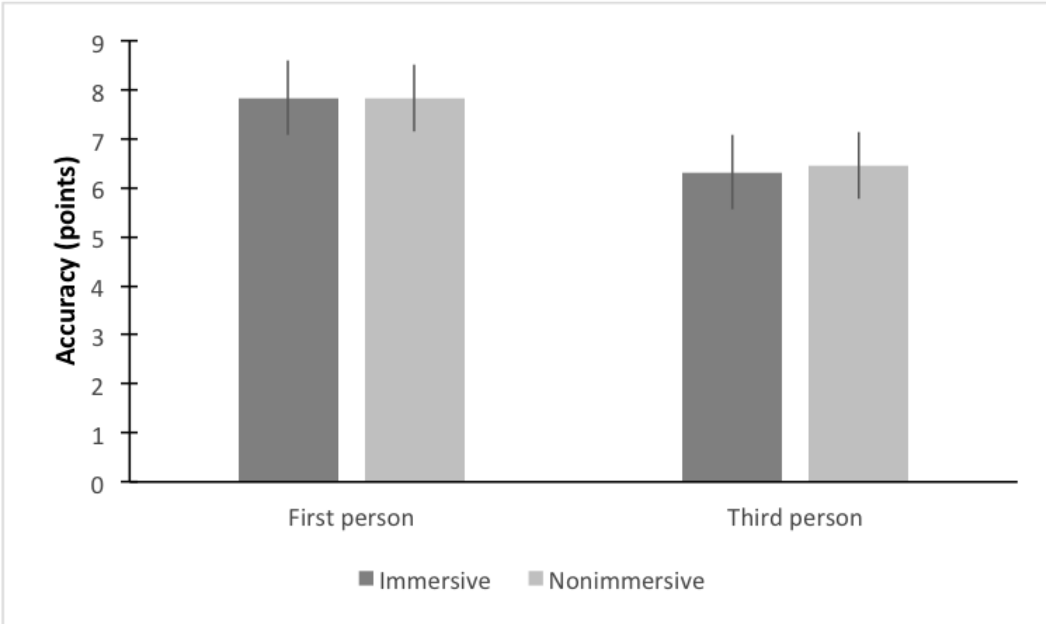

Figure 5. Mean overall scores on the last practice trial for each condition. First person led to significantly better accuracy than third person, and immersion did not lead to significant differences in accuracy. Conditions are collapsed across experience, and error bars represent standard error.

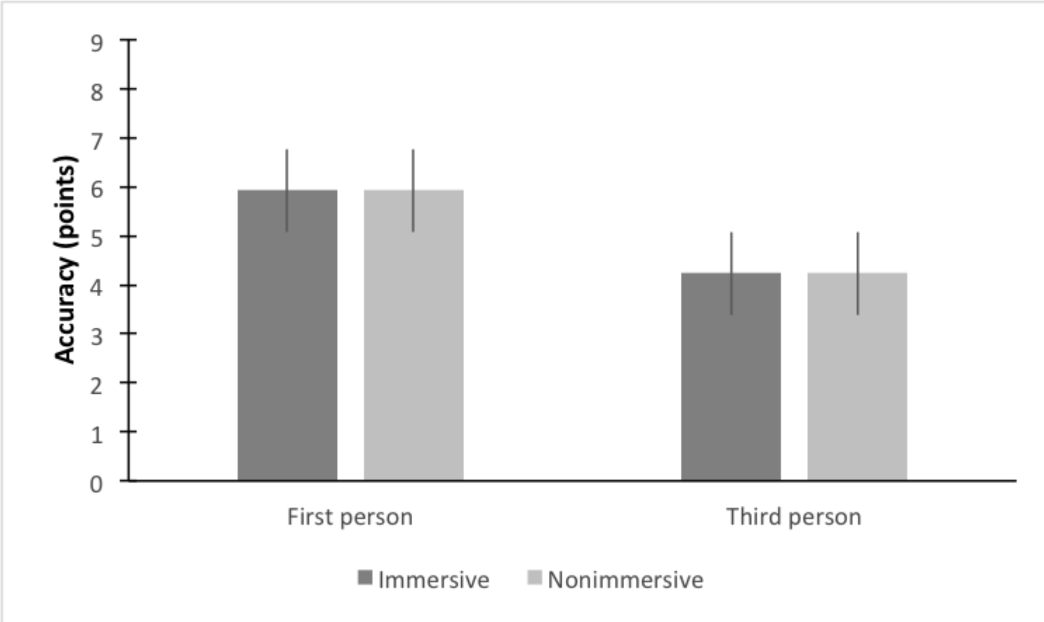

Figure 6. Mean overall scores on the testing trial for each condition. First person led to significantly better accuracy than third person, and immersion did not lead to significant differences in accuracy. Conditions are collapsed across experience, and error bars represent standard error.

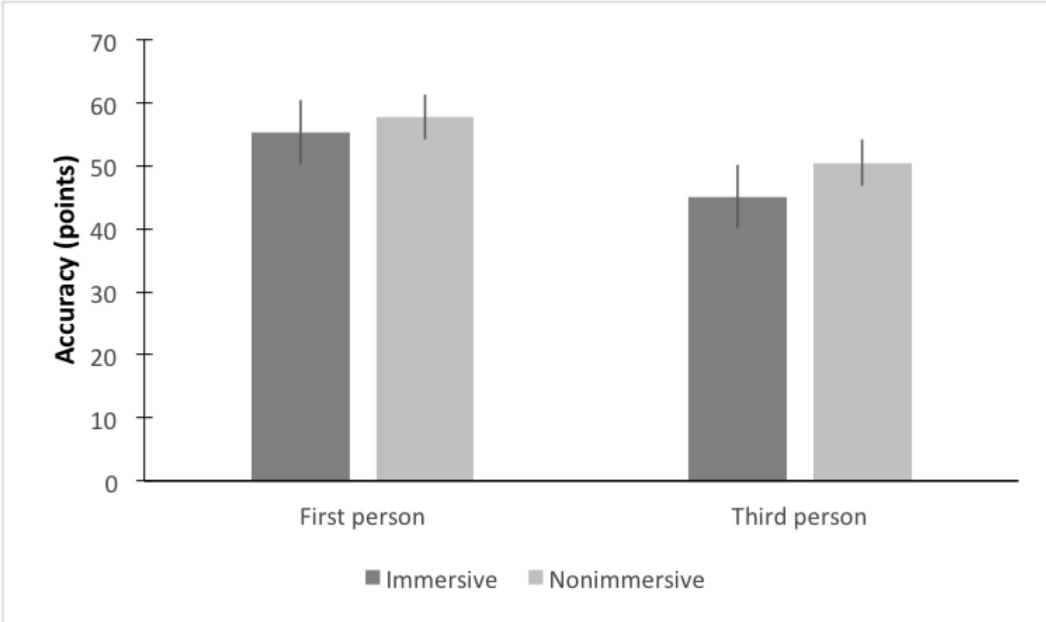

Figure 7. Mean total categorized scores on the testing trial for each condition. First person led to significantly better accuracy than third person, and nonimmersive formats led to significantly better accuracy than immersive formats. Conditions are collapsed across experience, and error bars represent standard error.

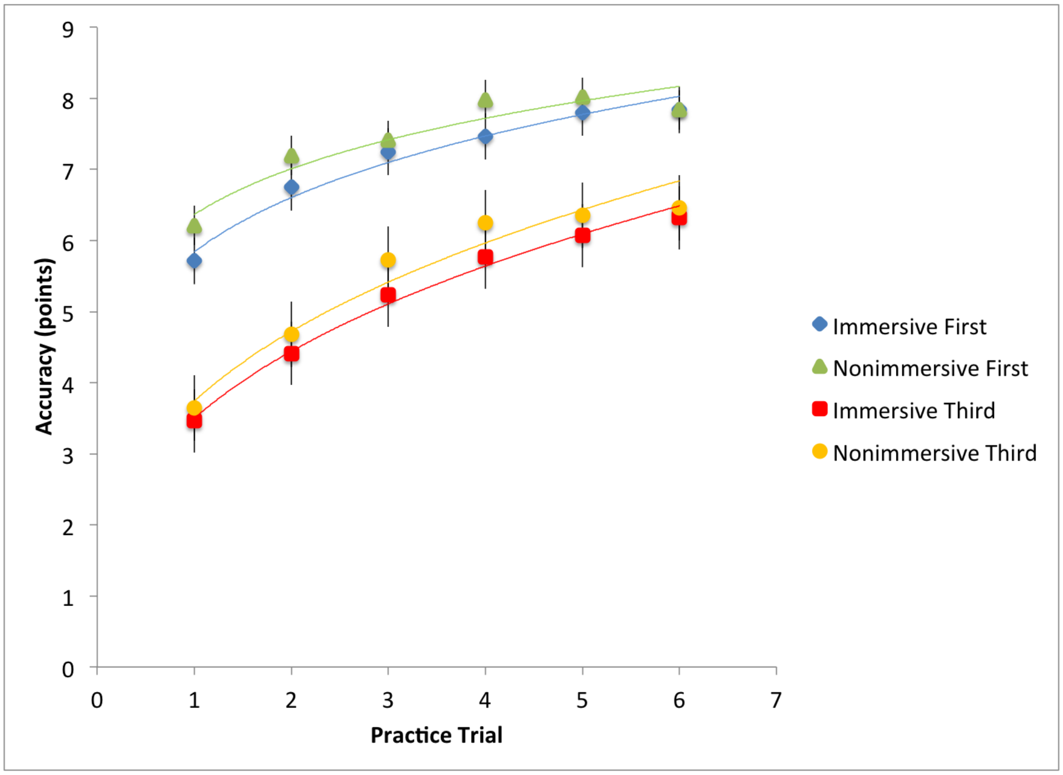

Figure 8. Best-fit mean practice curves for each condition. Mean accuracy increased with successive trials according to the power law of practice. Conditions are collapsed across experience, and error bars represent standard error.