Establishing Design Guidelines in Interactive Exercise Gaming: Preliminary Data from Two Posing Studies

Monica Zaczynski
Carleton University
Ottawa, Ontario
Monica.zaczynski@carleton.ca

Dr. Anthony Whitehead
Carleton University
Ottawa, Ontario
Anthony.whitehead@carleton.ca

ABSTRACT
Interactive gaming has demonstrated promise as a low-cost, at-home training and fitness instruction alternative. However, commercially available systems today are designed primarily for entertainment. As a result, the quality of instruction delivery and level of involvement may not meet the needs of a user.

This paper examines several issues including adapting for occlusion and lack of visibility; learning and orientation; and providing feedback to develop a set of design recommendations. While, the effect of visual delivery variation on performance was found to be inconclusive verbal delivery significantly improved performance. Additionally, verbal + haptic feedback produced the highest performance scores in comparison to other feedback types. With the additional support of qualitative data, these results, provide a strong foundation for future exploration of interactive training delivery that engages users, prevents injury and helps maintain fitness.

Author Keywords
Low-paced exercise training, Feedback, Haptic Feedback, Visual Delivery, Panning, Yoga, Wii Balance Board, Design Guidelines, Usability

ACM Classification Keywords
H.5.m. Information interfaces and presentation; Evaluation/methodology, User-centered design.

INTRODUCTION
Fitness and training games have become increasingly more prevalent for interactive gaming systems. They are more convenient than joining an overcrowded class while offering dedicated monitoring from the comfort of home. Interactive games can be designed for a range of activity types and can provide reporting and progress data if game software collects body measurement information effectively. Current commercially available systems are not necessarily designed for focused fitness but rather for entertainment, and as a result, the quality of instruction delivery and level of involvement falls short for the needs of users [8,16].

Though current research on exergaming for fitness and rehabilitation has focused on refining the systems hardware – body measuring technology – [7,11] or searching for an appropriate blend of entertainment and following a prescribed regimen – the actual game design – [2,5,15], it seems the quality of the activity execution has not been given due attention – the effective marrying of hardware and software – [1,5]. Likewise, the user’s self-efficacy associated with fully understanding what is expected is often overlooked. As much as a refined body measuring system and engaging game are important, paying attention to the users confidence in their own ability and understanding of expectations are essential to ensuring the level of interest and motivation that keep users motivated long term [3,13].

We will explore user understanding of pose presentation in low paced interactive training activities by examining elements that contribute to comprehension including occlusion and lack of visibility; learning and orientation and feedback. The goal of this paper is to document an ideal visual demonstration delivery that maximizes pose understanding and user self-efficacy, determine whether supplementary modalities are important for instruction, and determine if there is an ideal feedback delivery that promotes pose comprehension, confidence and motivation.

BACKGROUND & RELATED WORKS
Fitness instruction and yoga games are available for most interactive gaming systems today but are generally geared towards entertainment rather than accuracy. Games typically include a demonstration or training module preceding gameplay to teach users the action before they attempt to execute it themselves. Though this can mitigate the risk of injury, the quality of delivery can affect the comprehension of the action, which may increase the risk of injury. By allowing for a wide range of acceptable motion, interactive games accommodate for varying audience ability and evidently, understanding, at the expense of motion measurement precision. As a result, accepting a
wider range of error in the name of keeping the game fun and entertaining [6].

During execution, game feedback is limited to informing users of incorrect positioning rather than offering a way to correct. In some cases, because of the game’s method of motion capture, no feedback is provided at all. For example, at least three poses in the Nintendo’s Wii Balance Board Yoga game do not even use the balance board, thus not collecting any motion information or having capability of providing feedback. Though this may be an appropriate design for entertainment it is not necessarily useful or well suited for exercise training like Yoga, where precision and accuracy are important.

At the core of this issue is the body-measuring machine itself. The method and amount of body information collected determines the gaming system’s feedback capability. Commercially successful systems often use one or more types of sensors so that games can gather enough body information to create an engaging interaction or rapport with the user, however, this is not sufficient for providing constructive contextual feedback for the purpose of training. For example, the Nintendo Wii Balance board uses pressure sensors that are used to measure the user's center of balance and weight. Though balance is an important component to many poses and exercises in the Wii Fit game, it does not speak to the accuracy of the other elements of the pose, which are not collected. Two problems can occur here: firstly, since success is determined by only one measure – the weight distribution – pose accuracy, according to the game, may be achieved before the physical benefits from the other elements of the pose are attained, resulting in reduced effort by the user due to the game’s affirmation. Secondly, inaccuracy of the other pose elements that are not measured can undermine the exercise itself and even lead to injury.

Confidence in one’s ability is an important element to successfully complete a fitness routine [4,14]. However, due to limited body measurement collection, games offered for systems like the Wii Fit + Balance Board, Kinect, and PS-Move contribute to a patients’ false sense of confidence. An inaccurate pose may produce encouraging words and praise from the game rather than constructive feedback, giving users an inflated sense of understanding and ability.

Engagement and system accuracy of the body-measuring machine are both important aspects, however; literature that we have discussed suggests that the effectiveness of information delivery is crucial to providing a successful and engaging experience. An effective demonstration is often formed by the instructor’s experience; compiling elements that they found useful when learning it. When gaming is introduced, that element of ad hoc feedback from personal experience is lost and a host of new issues regarding visuospatial understanding and learning in a virtual 3D space surface. The aim of this study is to explore the variables that contribute to the quality of execution and user comprehension; this study will look at adapting for occlusion and lack of visibility, learning and orientation and providing feedback. We will examine each of these variables in detail next.

Adapting For Occlusion
For the purposes of this study, occlusion is defined as a visual obstruction that prevents the viewer from obtaining visual data required to complete an accurate mental image of the object being viewed. While in-person instruction has addressed occlusion on an individualized basis, training videos and interactive games to a certain extent have not. For example, Physiotherapists often provide patients with a sheet of printed exercises to perform at home. Like interactive pose systems, these 2-dimensional flat images try to illustrate the angle that best shows the body position, but occluded limbs or small body presentation errors may result in an incomplete mental image of the pose.

Designing a visual solution for providing multiple views or modalities without overwhelming the user and impairing their cognitive load can increase pose understanding and confidence and decrease the risk of injury.

Orientation & Self
Understanding and awareness of oneself with relationship to others is an essential part of imitation and kinesthetic learning [10]. Where body language and communication between self and instructor address these challenges in-person, a virtual instructor, virtual space and virtual avatar to represent self, produce new learning and interaction challenges. Maintaining the users’ sense of orientation of self and environment in comparison to their virtual avatar and the virtual environment is essential to learning through visual imitation and comparison to the instructor [12].

While gaming systems have made it possible to offer a strategic angle change when necessary, the seamlessness of the demonstration can also impact a users sense of orientation. Daems et. al found that when participants were asked to select like-poses, being shown two images at rotation angles greater than 30 degrees in difference consecutively, resulted in increased response times when compared to angles of lesser degree separations [4].

A commonly mistaken solution to visual misinterpretation is to add additional graphical elements to assist with clarification; however, adding more visual elements can increase the cognitive load and can make understanding even more difficult. This study will examine both performance and user confidence and preference to ensure design considers overall experience.

Feedback
We define feedback as any visual, verbal or physical criticism of an individual’s actions that lead them to achieve a target. While initial instruction usually demands the participant’s full attention, effective feedback has to
consider the interpretation factors described above while considering the user’s context – which in some cases, they may be mid-activity.

Providing constructive feedback and not just listing inaccuracies is something game-based instruction has yet to offer. Feedback can be divided into three requirements; the gathering of body information by the hardware sensors, the software’s interpretation of the information with appropriate response, and finally, how the response is delivered to the user through the game. While an ideal game design would marry the software and hardware capabilities effectively so that the system would be intelligently aware of the users’ activity from body measurement and would adapt modally to suit the users capability. In reality, constructive feedback is very difficult, if not impossible to provide if detailed body information is not collected.

While a significant amount of literature in exergaming and gaming for physiotherapy has focused on developing more suitable methods of collecting and interpreting body information by the system [11,17] exploration has been limited when it comes to delivering meaningful (contextually but also in an appropriate modality) and constructive criticism.

**DESIGN**

We evaluated the aforementioned factors in two experiments. The first experiment was divided into two sections. The first section focuses on visual demonstration delivery to examine the impact of occlusion from view and also orientation and self when angles and views are augmented during demonstration. The second section focuses on supplementary modalities and their impact on comprehension. Our second experiment examines the impact of various types of feedback, comparing current types of feedback with hypothetical contextual ones that explore different modalities and combinations of modalities. We will expand on these design goals next.

**STUDY 1: DEMONSTRATION DELIVERY**

**Methods**

Participants were described the purpose of the study before beginning and signed a consent form confirming they understood. They were then verbally asked a number of introductory experience and demographic questions which were recorded on paper. Each participant used a 15’ laptop to watch 2 sets of yoga pose demonstrations with five in each set, for a total of 10 different yoga poses.

Demonstration videos were created for the study using Maya animation software to the specifications of the visual delivery required for the study. Poses, verbal description and pose elements, were derived from Nintendo Wii’s Yoga game poses unless a particular pose was not available. In these cases a standard Bikram Yoga pose was used.

The conditions include:

**Set 1**

The first set of the demonstration delivery test focused on the visual delivery. Participants all saw the same poses in a random order matched with a random condition. All Participants saw all conditions. The conditions, illustrated in figure 5, include:

- Front: The most occluded view type.
- Front + Wide: The second begins with a frontal viewpoint then switches to the viewpoint that captures the widest visual area of the instructor – this view type offers an additional angle from the front angle to improve visualization of occluded areas.
- Front, Pan + Wide: The third begins with the frontal view, but will pan around the instructor to 40 degrees while they are performing the exercise. The viewpoint then switches to one that captures the widest visual area of the instructor – This view type offers a smoother transition between the same angles in the last view type.
- Front, Pan + Wide x1: Third but will offer an additional supporting viewpoint while the widest viewpoint is displayed – This view type reveals three angles throughout the demonstration, making occluded areas minimal.
- Front, Pan + Wide x2: This view type offers two additional images for a total of 4 angles throughout the demonstration making it the least likely to offer occluded areas. The amount of information displayed in this demonstration is what we anticipate will be a visual/cognitive peak.

Participants were able to watch the demonstration twice and were then asked to replicate the final pose using a flexible manikin. Both male and female manikins were available to participants. The final image of the demonstration was left on screen while the participant posed their manikin. To compensate for the lack of range of motion of the manikins and also to obtain a clearer idea of participant understanding of physical subtleties of the pose, participants were asked to verbally describe the elements of their posed manikin. With 6 years of regular Bikram Yoga practice, the researcher scored pose accuracy after each pose, based on a set of elements in the demonstrations. After each pose the researcher asked participants to verbally score their confidence in understanding, clarity of the demonstration and perceived pose accuracy on a 5-point Likert scale. Comments and criticism on experience were transcribed by the researcher using a laptop.

The second set of demonstration videos in the demonstration delivery test focused on verbal + visual vs. visual only delivery. Like the last set, participants were
shown a visual demonstration in a consistent visual delivery on a 15’ laptop. In this section, however, the poses were

provided in a random order with a random condition – with or without verbal description. Again, participants were able to watch the demonstrations twice and had to then pose and describe the manikin. They were able to see the final frame of the video while posing the manikin, just as in the first set. After each pose, they were scored by the researcher and asked the same questions as the previous set. Upon completing all 10 poses, participants completed a short exit survey about their preferences of visual and verbal delivery.

**Hypotheses**

We hypothesize that participants’ comprehension score will be best when offered a single, short but continuous, multi-view demonstration that will ensure any obstructed or occluded pose elements are clarified (H1). Suggesting that too many additional images will be visually overwhelming to participants (H2) and that they will score lowest when presented with only a front view of the pose (H3).

From strictly qualitative data we predict confidence ratings will follow pose accuracy scores. We expect participants will feel most confident in their understanding of poses offering a single, short but continuous, multi-view demonstration (H4). Conversely, we predict participants will feel least confident when provided only a single view of a pose (H5).

Although visual delivery is important it is often not enough on its own to gain a comprehensive understanding of an exercise, particularly in situations where the user must avert their gaze to complete the pose correctly. Under these conditions we anticipate that participants would score higher and also rate confidence higher if they were provided with verbal description of the pose in addition to visual (H6, H7). Conversely, they would score lower when only provided visual demonstration (no verbal) (H8).

**Participants**

20 participants were recruited to complete the Demonstration Delivery portion (10 male, 10 female). The mean age of participants was 30 years old with a standard deviation of 7. On average, participants reported to be in above average physical health.

**Results**

We began by looking for an effect of visual demonstration on pose accuracy. In this case, we predicted that the single, short but continuous, multi-view demonstration or Front, Pan + Wide would elicit improved pose accuracy. Figure 1 presents the mean pose accuracy scores and standard deviation across visual demonstration types. A repeated measure ANOVA did not reveal a significant relationship between pose score and visual demonstration conditions, \( F(5,95) = .842, p > .05 \), partial \( \eta^2 = .042 \).

A paired-samples T-Test was performed on pose scores for verbal and non-verbal conditions. We discover that there is, in fact, a statistically significant effect of verbal conditions on pose score, \( t(19) = 2.132, p = .046, d = 0.48 \), where the verbal condition improved performance by 11%. It is clear from figure 2 that participants performed better with verbal description than without.

**User Feedback**

A repeated measures ANOVA analysis did not reveal a significant relationship between confidence and visual demonstration conditions, \( F(5,95) = .842, p > .05 \), or perceived pose accuracy and visual demonstration, \( F(4, 76) = 1.121, p > .05 \). However, when examining the mean Likert scores of participant-reported demonstration quality, we can see that the Front view type was scored much lower than the remaining types. In support of this observation, analysis using the Greenhouse-Geisser correction revealed a significant effect of demonstration quality preference on visual demonstration, \( F(2.853, 54.205) = 3.361, p < .027, \) partial \( \eta^2 = .150 \). Post-hoc analysis with a Boneferroni adjustment further clarify the significant differences between the Front view and the other view types, suggesting participants found the Front view of lesser quality.
We then examined the responses to post-pose questions on confidence in understanding, quality of demonstration and perceived pose accuracy for the verbal section. A paired-samples T-Test was performed on each of the Likert questions with verbal vs. non-verbal conditions. We discover that there is a statistically significant effect of verbal conditions on confidence, \( t(19) = 3.356, p = .003, d = 0.75 \) and also quality of demonstration, \( t(19) = 2.396, p = .027, d = 0.54 \). This suggests that participants, not only prefer being provided verbal description to support a demonstration, they are more confident in their understanding when provided with it. We did not find any significant effect of verbal conditions on perceived performance.

Following analysis of collected quantitative data, we began exploring individual participant responses. These responses were transcribed and coded using a Grounded Theory method.

Categories were combined to form groups of ‘negative comments’ and ‘positive comments’. From figure 3 we can see that the front view produced the highest number of negative comments and, in fact, received significantly more comments relating to the visual inadequacy of the demonstration, specifically. Panning view with 1 additional angle received the least number of negative comments. Conversely, the panning view with 2 additional angles received the most amount of positive comments further supporting the notion that a panning view with additional angles is the most preferred type of visual delivery for 3D virtual demonstration.

At the end of the study, participants were asked to rate the most helpful type of visual feedback and the least helpful. The panning view with 1 additional angle was found to be the most helpful with a mean rating of 3.55, STD = 1.05. The front view was the least helpful with a mean rating of 1.35, STD = 0.99. Participants deemed the wide view angle – the most visible surface area of the instructor – more important than the other visual elements. A repeated measure ANOVA analysis revealed a statistically significant effect of demonstration deliveries, \( F(4, 64) = 5.325, p < .001, \eta^2 = .250 \). Post-hoc analysis with a Bonferroni adjustment revealed that participants significantly preferred the wide view compared to the front pan wide x2 (1.059 (95% CI, 1.156 to 1.961), \( p = .015 \)). While additional views were, on average, less helpful than other visual elements. The exit questionnaire also further supported the trend that verbal delivery was very important to participants, receiving a mean rating of 4.65 on a Likert scale of helpfulness, STD = 0.67.

**Discussion**

Though we predicted a single, short but continuous, multi-view (pan) demonstration – Front, Pan + Wide – would perform best, statistical analysis found no significant difference in pose accuracy across each of the visual demonstration deliveries (H1, H2, H3). While not significant, the increase in mean pose accuracy from 1 additional image to 2 may indicate that we have yet to reach a peak of cognitive load. For the purposes of the experiment, we felt that two additional angles to the main demonstration would suffice as an example of excessive information but performance did not decrease (H2). In that case, further research exploring greater amounts of visual information to determine where the peak of cognitive load lies.

From questionnaire feedback and verbal account we can tell that participants were overall most positive about the panning view with 1 additional angle but most confident with 2 additional angles. While this is not what we expected (H4), in the absence of feedback, the additional views may have given participants a sense of confidence by offering additional points of reference to self-check their own pose accuracy.

The front view demonstration type provided a highly occluded perspective of the pose. While a high number of participants noted the poor visual quality of this feedback type this was not reflected in the pose accuracy scores. The front view was not found to have a significant effect on pose accuracy or confidence. Additionally, participants reported almost double the amount of negative commentary on the front view than any other visual delivery, specifically, more than double the amount of comments regarding the visual inadequacy. This seems to support the notion that, in the absence of feedback, the amount of reference points to self-check one’s pose accuracy can impact the confidence and self-efficacy of a user. Future research in interactive gaming demonstration might assess the impact of varying amounts of visual information, including more than 2 additional angles, on performance and cognitive load and what that means for user confidence.

We found that some of our hypotheses were not confirmed. Participant scores were not found to have a significant difference between self-reported confidence in

![Figure 3 Grouped Negative & Positive Comments for Visual Deliveries](image-url)
understanding (H5) or with perceived performance ratings (H4).

Verbal description to supplement visual demonstration was found to have a significant effect on pose accuracy and also confidence in understanding compared to visual demonstration with no verbal description (H6, H7, H8). From verbal account, participants noted that they were better able to visualize how to get into the pose, where they should feel a stretch and how to correctly position themselves better with verbal description than without. Without description participants said they were unsure about where they should feel the stretch and overall did not feel confident in their understanding. This was also reflected in high scoring of verbal delivery in the exit questionnaires.

**STUDY 2: FEEDBACK DELIVERY**

**Methods**

During the Feedback study, participants were tested on comprehension of yoga poses delivered by Nintendo’s Wii Fit Yoga game and Balance Board. The intent of this experiment is to find an improved method of feedback delivery that can prevent injury during exercise in an at-home physiotherapy environment. Different types of feedback were assessed including visual, verbal, haptic and combinations of feedback.

Like the demonstration delivery study, participants were asked a number of introductory experience and demographic questions by the researcher. To complete the experiment, participant’s interacted with Nintendo’s Wii Fit Yoga fitness game and balance board. Six yoga poses were selected from the game and were delivered in demonstration mode (rather than play mode) in a random order. The researcher used the Wiimote to select the appropriate pose. The participants watched the demonstration once and executed the pose during the second demonstration along with the 2D instructor. During execution, participants received a randomly chosen feedback condition to help correct any pose errors. The feedback conditions included:

- **No feedback:** As a baseline.
- **Mirror Feedback:** To determine the effectiveness and value of a participant’s own mirrored body as comparative feedback. Yoga studios and physiotherapy clinics often use mirrors to observe positioning and posture. Similarly, some gaming systems use on-screen avatars to offer users a method of visually comparing their position to the virtual instructor. The Mirror used was a long rectangular full body mirror. It was placed in front of the participant careful to not obstruct the demonstration.
- **Wii Yoga’s built-in graphic and verbal feedback:** Mechanism measured by the balance board. To evaluate the effect of the combination of graphical and verbal feedback driven by limited body measurement information on participant performance and confidence.
- **Custom verbal feedback:** Delivered by the researcher that is detail-focused. Simulates the effect of a richer body measuring system and evaluate the importance of more comprehensive body measurement information on performance and confidence.
- **Haptic feedback:** Simulated by tapping participants on an incorrectly positioned area. This condition is meant to imitate the assistive touch of a physiotherapist putting a patient into the correct position and evaluate the effect of an alternate supporting modality, such as touch, on performance in a gaming environment.
- **Haptic and verbal feedback:** Combining the last two conditions to simulate the modalities utilized at an in-person session.

Feedback was based on the scored elements directly from the demonstration and was provided for the duration of the onscreen instruction – about 1 minute. Participants received feedback only when position was incorrect. Pose accuracy was scored by the researcher, based on a set of elements in the demonstrations. Participants then verbally scored their confidence in pose understanding before feedback and after receiving feedback as well as perceived pose accuracy on a 5-point Likert scale. Comments and criticism on experience were transcribed by the researcher using a laptop. Participants were not informed of their pose accuracy scores. Upon completing all 6 poses, participants completed a short exit survey about their preferences of feedback type. Participants were reminded that if a condition wasn't clear that they could be shown an example to assist in answering the questions.

**Hypotheses**

As we have discussed, feedback is an essential part of fully grasping a pose or activity. Using the mentioned modes of feedback, we anticipate that participants would perform best (H9) and feel most confident in their pose understanding (H10) when provided with customized verbal and haptic feedback that replicated feedback delivery offered by a physiotherapist. We also predict that because the Wii offers feedback only for a particular aspect of the pose, balance, and not necessarily the primary objective of the pose, Wii feedback with balance board would improve confidence but not performance score (H11).

**Participants**

A different group of 20 participants were recruited for the Feedback section (11 male, 9 female) of the study. The mean age of participants was 29 years old with a standard deviation of 5. On average, participants reported to be in above average physical health.
Results

We analyzed the collected measures by performing a repeated measures analysis of variance (ANOVA) using feedback types (6) x performance scores.

We began by looking for an effect of feedback on pose accuracy. In this case, we predicted that the customized combination of verbal feedback and haptic feedback would illicit improved pose accuracy. Figure 4 presents the mean pose accuracy scores and standard deviation across feedback types. Here, we can already see a positive increase in pose accuracy for the three customized feedback types. Mauchly’s test of Sphericity was applied to assess variance and covariance. This test was found to be significant \( \chi^2(14) = 42.436, p = .000, W=.084 \) suggesting that the assumption of sphericity was violated and groups do not demonstrate homogeneity. The Greenhouse-Geisser correction revealed a statistically significant effect of feedback conditions on pose accuracy, \( F(2.900, 55.107) = 10.231, p < .0005, \text{partial } \eta^2 = .350 \). Post-hoc analysis with a Bonferroni adjustment revealed that performance significantly increased from no feedback to verbal feedback (.156 (95% CI, .004 to .307), \( p = .041 \)), and from no feedback to haptic feedback (.169 (95% CI, .009 to .329), \( p = .032 \)), and finally, the largest increase in performance from no feedback to the combination of verbal and haptic feedback (.184 (95% CI, .032 to .336), \( p = .010 \)). Neither mirror feedback (-.061 (95% CI, -.294 to .173), \( p = 1.00 \)), nor Wii graphic + verbal feedback (-.048 (95% CI, -.274 to .179), \( p = 1.00 \)), had a significant increase in performance from no feedback.

User Feedback

We then looked at the Likert scores participants provided after each pose to determine if there is an effect of feedback on confidence and on perceived pose accuracy. Here, we predicted that the customized combination of verbal feedback and haptic feedback would improve confidence and also perceived pose accuracy, along side performance.

We began by examining confidence. Figure 4 presents the mean Likert scores and standard deviation for confidence in understanding after feedback across each feedback type. There appears to be a lot of variability, but we can see the

Discussion

As predicted, participants scored higher on pose accuracy with more customized feedback. While custom verbal and haptic feedback, individually had a significant relationship with pose accuracy, haptic + verbal feedback received the highest mean score with 95% and a standard deviation of 0.090 (H9). Participant reports found that this was not only the preferred feedback type but was also found to make participants the most confident in their understanding of the pose. These findings demonstrate that a more customized type of feedback is necessary to ensure not only performance accuracy, which is essential for an at-home rehabilitation system, but also users’ confidence in their
understanding and their ability. Statistical analysis also demonstrated that haptic + verbal feedback significantly increased participant confidence in understanding compared to no feedback (H10).

While mean Likert scores in confidence after receiving feedback for Wii graphic + verbal, custom verbal and custom verbal + haptic feedback were the same, it should be noted that mean pose accuracy scores for poses that received Wii feedback was observably lower than other feedback types at 72%, SD=0.20 (H11). This seems to suggest that the Wii feedback makes users over confident in their actions even when they are not performing poses correctly. Users do not receive the benefits of the pose while not being aware of their inaccuracies and may lead to injury if over confidence leads to over exertion.

This trend was also found in verbal accounts where Wii feedback had the most commentary regarding feedback inadequacy, but also received the highest amount of reports for confidence. We speculate this is due to participants not being clear on how to interpret the system’s feedback but then feeling confident in pose completion once the system congratulated them on a job well done. Here is an example of participants comments on the inadequacy of Wii feedback:

“I was confident in what I had to do but the feedback was unclear. I didn't know what to do to my body get myself into the correct position.” - Participant 15

In the categorized verbal responses we can see that physical limitation was noted by participants a number of times (24). While this was not under review in our testing it is our belief that users shouldn’t ever feel as though they cannot perform an action, particularly in a training scenario when the system should be better geared to a user’s capabilities. The system should provide appropriate alternatives that still capture the major benefit of the pose. In a similar vein, research on the field of exergaming has explored adapting a game’s difficulty to a user’s heart rate [9]. Future research might explore the effect of pose alternatives and detection of user ability on confidence and pose accuracy.

**CONCLUSIONS**

In this paper, we examined elements that contribute to pose comprehension in a virtual 3D environment including occlusion and lack of visibility, learning and orientation, and feedback, in an effort to better understand what is necessary to achieve a high level of performance, confidence and motivation.

We evaluated these factors in two experiments. The first experiment was divided into two sections. The first section focused on visual demonstration delivery and determined that a panning view, which includes an angle that captures the greatest visible surface area of the instructor and additional angles is preferred though didn't greatly impact performance. The second section focused on supplementary modalities and established that verbal delivery improved performance and increased user confidence while being a favoured element in demonstration. Our second experiment established that more customized and contextual modes of feedback are necessary to improve feedback and increase user comprehension and confidence. While addressing our design goals, these evaluations offer a direction and set of guidelines for designing for training and fitness applications.

**DESIGN RECOMMENDATIONS**

The evaluation of these factors can be summarized into the following design recommendations for future research and design of interactive gaming systems for fitness and training:

1. **Multiple Visual Angle Demonstration:** When demonstrating a pose, we recommend that users be provided with multiple angles of a demonstration to help them resolve the visual information they require to understand a pose. Additional angles can provide more points of reference for the user to check their understanding or self-positioning. Panning can offer a seamless range of view without imposing a cognitive increase with a visual angle switch. Although there were no significant differences between pose accuracy or confidence and visual delivery methods, the inclusion of a wide view that incorporated the largest amount of visible surface area of the instructor, a panning view and also simultaneous additional angles were found to be preferred elements by users.

2. **Multi-modal Demonstration:** We observed that participants required verbal or clear descriptive language about what their body should be doing, where to feel the stretch, when to breathe etc., in order to complete their mental model of the pose but also to feel confident in their understanding on the pose being executed. Descriptive verbal delivery was found to have a significant relationship with pose accuracy and confidence in understanding. Instruction without verbal description was found to feel inadequate and lead to a lack in confidence. Some participants expressed the importance of analogies to help relate to the pose more effectively. This can help connect with users and can contribute to engagement.

3. **Clear Orientation Environment:** When providing a visual demonstration we recommend that instruction begins by establishing a clear orientation environment. This may include the ability to face the instructor away or towards the user or simple visual cues like a yoga mat. This allows the user to mentally clarify the relationship between their own body and the instructor. We observed that several users would miss an entire demonstration because they were cognitively engrossed in interpreting which hand or foot to use. Establishing a
clear orientation environment becomes especially important with complicated poses.

4. **Verbal Delivery as Supplementary**: Relating to the previous recommendation, when demonstrating a pose, we suggest that verbal delivery remain a supplementary modality to visual delivery. We observed that the use of direction (i.e., left and right) in description often lead to confusion. Visual delivery should be used as a point of reference or imitation while verbal delivery provides analogy, unobservable detail and establishes rapport with the user. Haptic feedback can also assist with directional instruction.

5. **Custom, Contextual feedback**: Though we can really only evaluate currently available fitness and training games that offer feedback, this feedback was found to be inappropriate for more performance focused activities like training and rehabilitation. When offering feedback for these situations, we recommend providing users with customized, contextual feedback to help them resolve errors. Custom, contextual feedback offers a direct solution to inaccuracy rather than just demonstrating inaccuracy and allowing the user to determine a solution. Custom, contextual feedback types, including verbal, haptic and a combination of these, were found to have a significant relationship with pose performance and were also highly preferred to commercially available, Wii system feedback or self-regulated modes of feedback such as a mirror or avatar of one’s self.

6. **Multi-Modal Feedback**: Speaking further to the point of contextual feedback, we recommend that feedback be provided in a modality that compliments the orientation of the action. It is likely the user is mid-action when receiving feedback, in these cases, the system should adapt it is feedback modality to accommodate the user – provide haptic or verbal feedback if the user’s eyes are averted, for example. It was noted that participants rated modalities that were ill-matched to the pose worse than more appropriate ones. Participants did not even take advantage of the feedback provided because they weren’t able to see it from the position they were in. In some cases, participants performed as though they had no feedback at all if they were provided with an ill-matched modality and even said they did not use or notice the feedback.

7. **Understand & Adapt For Capability**: Especially in training, the system’s awareness of the user’s capabilities is essential to ensuring poses are performed accurately and without injury. For this reason we recommend that the system be aware of user ability and provide alternative actions that still address the benefit of the pose without overexerting or discouraging the user. During analysis we found that participants often commented on their physical limitations when attempting the poses. Not feeling capable or able to complete a pose can lead to noncompliance. Providing participants with alternatives, that still address the gross benefit of the action, yet still offer a challenge, may make users feel confident while still performing the pose correctly, without risk of injury and working towards recovery.

While we recognize that many of these recommendations are not currently realizable and may still require a lot of technical exploration, we feel that by exploring user needs and submitting possible solutions we can begin the iterative journey to a well-designed application that exceeds the needs of the user.

**LIMITATIONS & FUTURE WORK**

This paper only addresses some of the many factors that contribute to action comprehension. The goal of this paper was to begin to build a foundational set of design guidelines for further research in the area where none exists.

We acknowledge the sample sizes of all studies were very small, limiting adequate power of these research studies. As a result many of these observations aren’t fully supported statistically, however, we believe that qualitative feedback can, at a minimum, drive further investigation.

In the visual delivery section we chose a very general set of delivery types to reveal patterns that could be later explored. We feel we accomplished this. While we did not find a significant relationship between the deliveries we chose and pose accuracy future research in interactive gaming demonstration might assess the impact of varying amounts of visual information, including more than 2 additional angles, on performance and cognitive load and what that means for user confidence.

While we did find a significant effect of verbal + haptic feedback, suggesting the use of a haptic feedback system has implications. Haptic systems are typically comprised of a series of wearable bands with sensors and feedback mechanisms that need to be attached to the body in specific locations to ensure effective data collection and feedback. The challenge of repeatedly applying the haptic system is not only tedious but also difficult to perform accurately each time. Additionally, an ideal haptic solution would also consider ease of donning the system without assistance since users may be applying it themselves. Future exploration of the feasibility of haptic feedback to support interactive training would need to consider the context as well as performance and user confidence.

Beyond elements of these studies, exploration of learning styles and adapting for different user needs could also enrich the application of guidelines for rehabilitation in gaming systems.
REFERENCES


