Physio@Home: Exploring Visual Guidance and Feedback Techniques for Physiotherapy Exercises

Richard Tang[1]   **Xing-Dong Yang[1]**   **Scott Bateman[2]**   **Joaquim Jorge[3]**   **Anthony Tang[1]**

[1]University of Calgary
Alberta, Canada, T2N 1N4
{richard.tang, xdyang, tonyt}
@ucalgary.ca

[2]University of Prince Edward Island
Charlottetown, Canada, C1A 4P3
sbateman@upei.ca

[3]IST/INESC-ID/U Lisboa
Lisbon, Portugal
jaj@inesc-id.pt

ABSTRACT

Physiotherapy patients exercising at home alone are at risk of re-injury since they do not have corrective guidance from a therapist. To explore solutions to this problem, we designed Physio@Home, a prototype that guides people through pre-recorded physiotherapy exercises using real-time visual guides and multi-camera views. Our design addresses several aspects of corrective guidance, including: plane and range of movement, joint positions and angles, and extent of movement. We evaluated our design, comparing how closely people could follow exercise movements under various feedback conditions. Participants were most accurate when using our visual guide and multi-views. We provide suggestions for exercise guidance systems drawn from qualitative findings on visual feedback complexity.

Author Keywords

Physiotherapy; movement guidance; visualization; augmented reality;

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

As people age, they are more susceptible to joint and muscle injury. While improvements in healthcare mean that people can survive these injuries, many live longer while recovering from such injuries [20]. Physiotherapy is a post-injury rehabilitation activity that improves and restores physical function after injury or surgery [7, 20]. For example, after a dislocated shoulder is put back in place, a patient is taught several shoulder exercises to help restore strength and range-of-movement. In these sessions, the physiotherapist teaches and demonstrates exercises. A patient will then perform the exercise, and while they practice, the physiotherapist provides immediate feedback and correction to ensure the patient performs the exercise correctly.
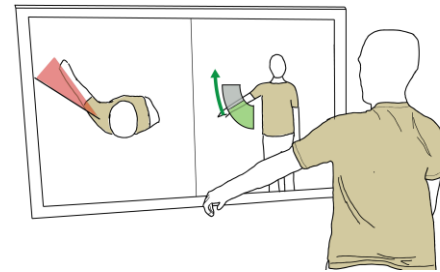
Figure 1: Our prototype shows front + top camera views of the exerciser with visualizations overlaid on their body.

This feedback ensures the patient will be able to recover physical functioning and avoid re-aggravating their injury [4]. In addition, the feedback can provide the patient with motivation to continue their exercises.

The problem is that patients also need to perform these exercises at home, without the guidance of a physiotherapist. Without guidance, patients risk performing their exercises incorrectly—at best without gaining restorative benefits, and at worst re-injuring themselves. Early technology explorations to help patients have used on-body sensors to provide movement and guidance feedback (e.g. [1, 9, 16, 29]). Yet, many of these initial explorations employ technologies that are expensive and unwieldy for home use.

The commoditization of computer vision technologies makes encumbrance-free, low cost body tracking a promising and viable alternative. Several researchers have explored the use of the Microsoft Kinect in this way—tracking motion, and then providing visual feedback to help teach and guide new movements (e.g. [3,21,25]). However, most of this prior work has focused on gross motor movements (e.g., dance), and in contexts forgiving to movement errors. In contrast, physiotherapy demands careful and controlled movements due to a patient's reduced mobility and the potential for re-injury. Given this focus, what are the kinds of feedback and guidance that physiotherapists are concerned with? How can we design effective visual feedback that allows people to take meaningful corrective actions during exercises?

Based on discussions with a physiotherapist, we derive five features of movement guidance and feedback. We realize these ideas in the design of Physio@Home, a prototype system for guiding and aiding proper performance of basic physical therapy exercises. As people perform exercises

(mimicking pre-recorded movements), the system provides several views with a real-time, on-screen 'Wedge' visualization which provides corrective guidance (Figure 1).

To evaluate our design approach, we conducted a study with 16 participants. Our results show that using the Wedge and multi-view camera allows people to perform exercises with the least amount of error compared to using video alone. While our strong results demonstrate the potential of our approach, we have identified several outstanding issues that need to be addressed in systems providing visual guidance for precise movements. These include the visual complexity of guides and avoiding information overload.

This work makes three contributions: first, we developed Physio@Home, a novel system that demonstrates the viability of visual feedback for guidance of physical therapy exercises in the home; second, we contribute a study of our system, which highlights the utility of visual feedback and multi-view for movement guidance, and finally, we identify a set of challenges for the design of similar systems.

RELATED WORK

Rehabilitation Systems Requiring Worn Sensors
There has been a recent push to develop rehabilitation solutions for treating patients at home. Examples include [4, 29], which use wearable kinematic sensors to track and visualize knee angles during exercises. These systems visualize the knee angle on a computer monitor to indicate whether the patient is correctly performing their exercises. Ayoade & Baillie [4] evaluated their system with patients undergoing knee rehabilitation over a six-week period and found that patients working with their system were recovering better than patients without. Pt Viz [2] also uses wearable sensors, but visualizes the correct knee angle directly on the wearable using lighted fabric.

BASE [9] was another kinematic sensor-based system that focused on supporting exercises at home. The system was designed for older adults, a population commonly requiring physiotherapy care, and focused on strength and balancing exercises. BASE displayed real-time feedback based on the alignment and angle of their legs.

Some rehabilitation systems were also designed as exergames to encourage and motivate patients. Alankus et al. [1] designed interactive games for stroke patients using webcams and Nintendo Wii remotes to track arm motions. Uzor and Baillie [24] implemented simple exergames for elderly patients using knee bending, sit-to-stand, and marching exercises and found that their exergames encouraged adherence to rehabilitation programs and would have the potential to improve mobility.

These systems demonstrate the benefits of computer-based rehabilitation systems over traditional methods. They also emphasize the importance of measuring the angles of joints, and using interactive visual cues that directly show progress in an exercise. However, these systems are limited to sim-ple, coarse-grained movements such as bending a knee [4, 29] or aligning a patient's entire body [9]. Our work focuses on precision for more complex shoulder exercises.

Vision-based Systems for Rehabilitation
The Microsoft Kinect is a viable sensing platform capable of full-body and limb tracking. The Kinect also benefits from being readily available, easy to setup, and low-cost. For example, Huang [12] developed Kinerehab to track arm-based exercise movements. Similarly, Lee et al. [13] used the Kinect to track Tai Chi motions for physical rehabilitation. Rector et al's Eye-free Yoga [19] made use of the Kinect and audio feedback to develop an exergame for visually-impaired persons to match yoga postures. Similarly, MotionMA [25] uses the Kinect to focus on movement interpretation and feedback for performing repetitions. Work by Camporesi et al [6] studied how Kinect systems may be customized to improve treatment. Overall, rehabilitation systems have taken advantage of the Kinect's depth-sensing and skeletal tracking capabilities to develop easy to use and deploy systems for patients with a variety of conditions.

However, the Kinect is still subject to inaccurate tracking and skeleton placement. As noted in [12], the Kinect had difficulties tracking patients with walkers or wheelchairs. Work by Tao and Levin [23] found the optimal position of the Kinect to be between 1.45 and 1.75 meters in front of the user and 0.15 meters left or right. These area requirements may restrict their deployment in patient homes and introduces sources of error that must also be considered when used for rehabilitation. Despite these disadvantages, performance studies of the Kinect suggest that it has sufficient overall tolerance that it can be used for encouraging and maintaining exercise compliance [17].

Our project focuses on precise exercises that involve movements with a higher number of degrees-of-freedom. To eliminate tracking problems inherent with current iterations of the Microsoft Kinect, particularly with fine-grained rotation exercises, we use Vicon cameras. This allows us to design our system with a best-case motion tracker. We may soon have smaller, commodity-level sensors that near the Vicon's precision, without its extensive infrastructure.

Movement Instruction
The broader HCI community has also explored movement guidance. Early work explored techniques to teach gestures (e.g., for touchscreens), while more recent work has explored approaches for providing hints/guidance for full body movement. Various projects have explored different mechanisms for providing feedback, though in general, this work focuses on visual feedback. For example, OctoPocus [5] and ShadowGuides [10] are concerned with teaching gestures on touch screens. These systems present dynamic, real-time guides when the screen is touched to help teaching people gestures and movements.

Others have explored different ways of presenting this feedback—often in egocentric ways. For instance, rather

than relying on the interaction surface as the output space for the feedback, LightGuide [21] projects guidance on people's body (i.e., their hands). Here, the researchers studied how guides such as arrows and lines projected directly on the body can help guide arm motions. While effective, this projection approach may not be appropriate when body parts cannot be seen (e.g., the back of one's shoulder). Just Follow Me [28] taught movement using a head-mounted display and virtual reality. Just Follow Me uses ghostly arm outlines to convey movement instructions to the wearer. Similarly, White et al. [27] evaluated visual hints for instructing physical gestures with cards, while Henderson and Feiner [11] displayed arrows and visual guides directly on objects to complete assembly tasks.

These projects introduce the concepts of *feedback* and *feedforward*. Feedback is information conveyed during or after the execution of movement indicating what gestures the user is close to performing (e.g., OctoPocus) and the visualization of current posture (e.g., ShadowGuides). Feedforward provides information about how to complete future movements: OctoPocus realizes this through possible tracing paths, while ShadowGuides and Just Follow Me realize these as future hand poses and ghostly arm images, respectively. We drew on the richness of these past uses of feedback and feedforward information to advise the design of our visualizations for guiding precise arm movements.

Most close to our work, YouMove [3] is a full-body movement instruction system with personalized recording, annotation, and gradual learning using an augmented mirror. The feedback is presented in a "mirror" that allows people to see themselves. Their system taught movement in gradual steps by scoring movement similarity and removing guides as the performer becomes proficient. We build on this approach of an "augmented mirror" in three ways: first, we focus on finer-grained physio exercises (rather than gross movements like yoga and ballet); second, we focus on guiding movement through dynamic guides (as with [5, 10]) as opposed to teaching through repetition; and finally, our work explores the use of multi-view cameras for feedback.

DESIGN
Physiotherapists often provide patients with pamphlets [26] containing drawings of exercises, but it can difficult to understand these movements based only on a few snapshots. A better option is to refer patients to online videos that demonstrate complete exercises; however, this lacks the feedback necessary for precise movement guidance. Our goal was to design visual guidance that could be used at home to guide careful movements such as those required in physiotherapy. We envision such a system would be displayed on a television screen so that it functions like a mirror in their physiotherapist's office (e.g. [3]). Ideally, the system would not require on-body sensors or complex set-up, instead relying on low-cost cameras (e.g., Kinect).

We took an iterative approach to realize our final design. We worked with a practicing physiotherapist to learn the

basic exercises they teach patients, how to perform them correctly, and how they teach and correct their patients.

Example: shoulder abduction & adduction. For the purpose of this section, we describe one simple physiotherapy exercise: a shoulder abduction followed by adduction. This is a strengthening exercise that involves raising the arm (abducting) along the frontal plane (Figure 2) up to shoulder level while keeping elbow locked, and then lowering (adducting) the arm back down to the patient's side. This exercise may be performed with the arm being raised along the frontal plane, or 45 to 60 degrees from the patient's front. They are also often per-
formed with thumb facing upwards in order to work the muscles in the shoulder. Other exercises are described in our Evaluation.
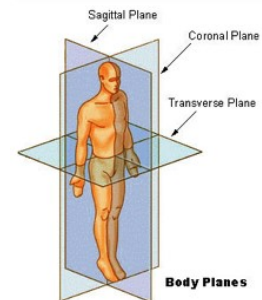
Characteristics of Guidance
Based on how physiotherapists train and correct exercise movements, we classified four guidance characteristics they communicate to patients: *plane (and range) of movement*, *maintaining position/angle*, *extent of movement*, and *movement speed*.


Figure 2: Anatomical body planes. Coronal plane is referred to in paper as 'Frontal plane'

Characteristic	Realized in Physio@Home
Plane of movement *Path and direction that the user needs to move*	Wedge, Movement arc, Directional arrow
Range of movement *Start and finish of a movement*	Wedge
Maintaining position/angle *Joints to keep still/keep at a particular angle*	Nearest Arm
Extent of movement *Limits on error in movement*	Movement arc, Nearest arm, Top-down angle
Rate of movement	N/A

Table 1: Summary of guidance characteristics

We summarize these in Table 1 and explain them in the context of the shoulder exercise described previously.
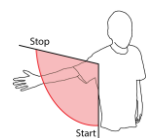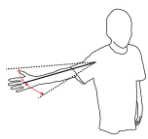
 Plane or Range of Movement. The *plane of movement* refers to the plane the body part will move along during the exercise. The *range* refers to the "start point" and "end point" of this movement. For instance, during non-angled shoulder abduction, the patient's arm moves up along the frontal plane (Figure 2), starting from resting position to where it is exactly aligned with the shoulder.
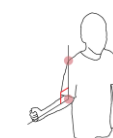
 Maintaining position or angle. For many exercises, certain joints need to be kept in either a fixed position, or at a fixed angle. In the case of a shoulder abduction/adduction, the arm must be kept straightened, and the shoulder kept level with the ground. Other exercises are

stricter—for example, with an external rotation exercise, the elbow needs to stay next to the body, and be bent at 90°.

Extent of movement. The *extent of movement* limits how a body part's motion can and should deviate from the plane of movement. For example, during an angled shoulder abduction, the arm must maintain its angle relative to the body's sagittal plane.
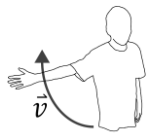
Rate of movement. This refers to how fast a body part must move. For some exercises, performing them slowly ensures the right muscles are being used. This characteristic applies to a variation of the shoulder adduction where the arm must travel slower as it returns to the patient's side. In many cases, the patient is free to proceed at their own pace.

Wedge Visualization

We designed a guide called the 'Wedge' to be displayed on-screen to convey an exercise's movement characteristics in real-time (Figure 3). We chose to use a dynamic visual guide instead of video so that the guide can be adjusted for a user's size and position and provide contextualized movement guidance while the patient is moving. The Wedge consists of several distinct parts: the Movement Arc, Directional Arrow, Nearest Arm, and Topdown Angle.
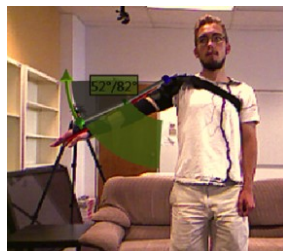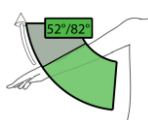

Figure 3: Wedge visualization. The dynamic visual guide is overlaid on top of a user's arm

Movement Arc. The central arc shape of the Wedge conveys the plane of movement for each part of an exercise. It is based on the motion of the arm with either the shoulder or elbow as the center of its radius and where the moving arm forms the shape of the arc.


The Movement Arc is divided into two parts: one section for the completed portion in green, and the other for the incomplete remainder of movement. As the user follows the plane, the green completed section grows to indicate progress, while the grey incomplete section shrinks to show how much of the movement still remains. This conveys both feedback and feedforward, and offers motivation for the user. In addition to the shape and fill of the Arc, we also provide a numeric angle indicator of their current and required arm angles to complete exercise.

Directional Arrow. We draw an arrow on the outside of the Movement Arc to show the direction the user must move in. Similar to the Movement Arc's feedforward, the Directional Arrow shows where to move to and how much of the movement is left as the stem shrinks with user progress.

Nearest Arm. We draw a red stick figure of the nearest correct arm from the exercise to the user's when they are in the wrong position. This guide provides feedback on the user's movements by letting them know if they are in the incorrect place and where they should be. When they are properly aligned, this guide disappears.
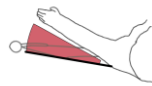
Topdown Angle. Similar to the Nearest Arm, we show a red arc in the top-down view when the user's arm is moving along a vertical plane and is not on the same angle. This arc grows and becomes more visible if the user is further away so their arm may maintain the required angle from their forward direction. This provides corrective feedback on the extent of the movement.

By design, the Wedge and its separate parts encapsulate all our movement characteristics except for rate of movement. We envision the latter being conveyed by animating sections of the arc and arrow to imply required movement speed. The Wedge uses simple visual elements to avoid screen clutter, and so that its components do not interfere with each other.

Multiple Camera Views

Most vision-based systems, particularly those using the Kinect, consist only of a single perspective view facing the user (e.g. [3,12,13,19,25]). While these systems are sufficient for exercise movements where the movement is perpendicular to the camera, depth perception is a challenge. Difficulty with understanding depth, combined with mirroring by the camera, a limited field-of-view, and angle of tilt, means users can have difficulty identifying if they need to move closer or further away from the camera.

To address this problem, we explored the use of multiple camera views. Multiple camera views are sometimes used in dance instruction [8] to show footwork from an alternate angle. PTMotions [18] developed an iOS app for viewing common physiotherapy exercises from front and side views for better viewing angles of posture. We extend this concept in Physio@Home (Figure 4) using the Wedge visualization by also adding real-time feedback. By using multiple views, an exercise movement may be clearer and easier to understand. For instance, a front-facing view may be better for movements on the frontal plane, while a top-down view may better for interpreting forward and backward depth.

Chapters: Reducing Visual Clutter

In early iterations of Physio@Home, we noted that the Wedge, even with its simple visual elements, exhibited a

high level of visual clutter. To simplify what was being displayed, we decided to split exercises into a sequence of 'chapters', where each 'chapter' represents a single one-dimensional movement step. This reduces the amount of visual clutter on the screen. For example, the shoulder abduction/adduction exercise becomes two chapters: for the abduction chapter, it shows a Wedge with a directional arrow pointing up; for the adduction chapter, it shows the Wedge with a directional arrow pointing downward.

IMPLEMENTATION

Tracking. We implemented our prototype system using Vicon motion tracking cameras, the Proximity Toolkit [14], WPF, and the Helix3D toolkit. To track joints, users wear markers mounted on shoulder, elbow, and wrist support braces. The tracking system provides x, y, z coordinates for each joint in millimeters within the testing space. These coordinates are then used in the 3D viewport to place 3D elements in corresponding positions to the real world.

To set the 3D viewport camera, we applied Vicon markers to the corresponding RGB camera. We used the Proximity Toolkit to retrieve position and orientation and set the 3D camera properties to these values. We then manually adjust the pitch, yaw, and roll of the 3D camera to align its image with the RGB camera. We overlay the 3D viewport atop the video feed and aligned the Helix 3D and RGB cameras to appear as though they originate from the same location.

Vicon vs. Kinect. Our initial prototypes used the Microsoft Kinect depth cameras [22] (see [27] for more), but we opted to use the Vicon tracking system for this iteration due to the Kinect's inadequate tracking. We found that the Kinect had problems accurately placing joints when these were obscured by the user's body. For instance, when the user holds their hand perpendicular to the Kinect—either fully stretched out or with their elbow tucked against their side—the Kinect cannot see their elbow and misplaces it as it tries to track their skeleton. This problem complicated the placement of visual guides and error calculations. Smoothing data with filters and using multiple Kinects were inadequate; they both considerably slowed down the software.

To assess the accuracy, we conducted a small experiment where we completed five repetitions of the Elbow exercise (see below), recording the reported length of both forearm and bicep at 0.1s intervals. The Kinect produced lengths of high variability (bicep standard deviation = 10.8mm, forearm standard deviation = 28.9mm), whereas the Vicon pro-

Figure 4: Top (right) and front (left) camera views, as displayed on TV screen in front of user.

duced a far more stable result (bicep SD = 1.9mm, forearm SD = 0.7mm). Therefore, we chose the Vicon tracking system. We were interested in evaluating the extent to which the visualization would improve the accuracy in movement guidance. In addition, the Vicon tracker uses multiple cameras and could thereby track obscured joints more easily. We envision that over time, tracking will achieve Vicon-like accuracy with commodity-level costs.

Wedge Visualization. The Wedge visualization is rendered using the Helix3D toolkit. It consists of four visual elements: two pie slice elements that form the entire Movement Arc, one for percentage of chapter completed, and one for the chapter remainder; a pipe for the stem of the Directional Arrow; and, a cone for the Direction Arrow's head.

The first and last frames of the exercise chapter are used to create the start and end points for the Wedge's movement arc. To show percentage of the completed chapter, the user's live arm position is matched by vector to the most similar arm posture from the chapter.

Multi-view. We implemented our multi-view setup using two commercial RGB cameras mounted in front of and above the user. The top perspective was selected as the second view due to the exercises requiring more movements related to the transverse plane. In principle, it is possible to show more camera views; however, we limited this to first understand how a second view would be used.

Recording and Playback. Our system records movements captured by both the RGB and Vicon cameras at a rate of 60fps. For each frame, the raw images and x, y, z positions of the shoulder, elbow, and hand markers are captured.

Scaling. We transform recorded exercises to account for variations in a recorded participant's arm length and location to compute error. We first compute the length of the participant's bicep and forearm using the absolute Euclidean distances between the x, y, z positions of their shoulder, elbow, and wrist. We then iterate through each pre-recorded frame to compute normalized 3D vectors of the bicep and forearm, and then using the participant's shoulder as an origin, we multiply the bicep vector by the participant's bicep length and add it to their shoulder position to get a new elbow position from the exercise, now scaled to the participant's bicep length. We do the same with the forearm vector to get a scaled wrist position, and repeat over all frames of the recorded exercise until the exercise has been transformed and scaled to the participant.

Error metric. We implemented a tool to compare a user and exercise recording and compute average errors between them to show how closely the user was able to follow the exercise. Our tool uses the previously described scaling algorithm to scale and transform the exercise to the user's size and position. It then iterates through each recorded user frame; for each user frame, it searches through the scaled exercise frames to find a frame with the least error between elbow and wrist—the shoulder is excluded because our
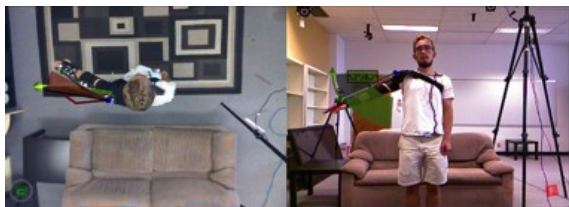
scaling algorithm uses the shoulder as the origin. This error is computed by absolute Euclidean distance in millimeters. This allows us to focus on how closely they could follow the exercise. Errors are accumulated and averaged by the number of user frames.

Authoring. We developed an authoring tool (similar to [3]) to encode playback instructions for each recorded exercise. This tool allowed us to seek frame-by-frame through a recorded file to mark groups of frames as 'chapters' in order to distinguish separate, single-dimension movements as independent parts of each exercise. Further, our tool allowed us to indicate which joints are to be kept stationary during an exercise, which is similar to physiotherapists highlighting arms and legs on diagrams of exercises.

EVALUATION

Our design process produced a rich, but relatively complex, visualization: the Wedge. The Wedge communicates several aspects of movement, and can be augmented by a secondary overhead camera view. We were interested in three specific questions: Does the Wedge help people to perform the exercises with increased accuracy? If so, are both the Wedge and multiple views necessary, or is one sufficient? Does the Wedge and multiple views perform differently for different types of exercises?

We recruited 16 graduate students from a local university through email lists. Each study lasted an hour and participants were paid $20. We used a within-subjects design to evaluate both accuracy and subjective preference with four different combinations of the Wedge and the number of views (Interface condition): *single view with video playback* (VideoSingle), *single view with Wedge visualization* (WedgeSingle), *multiple views with video playback* (VideoMulti), and *multiple views with Wedge* (WedgeMulti). The conditions were presented in a Latin Square ordering to avoid bias.

Participants would complete the exercises while following an on-screen guide (either a video recording of the exercise being demonstrated or the Wedge visualization system). In the video conditions (i.e., the ones that did not use the Wedge visualization), participants would see a main video of themselves (like a mirror), with an inset video of the pre-recorded exercise, allowing them to mimic the exercise. In the Wedge visualization conditions, the visualization (based on the pre-recorded video) is overlaid atop the live video.

Procedure

Participants were introduced to the system with a short demonstration and, after being fitted with markers, completed a trial run of each condition. Participants were allowed to spend as much time as they needed to test and understand the visualization. During this phase, each participant was taught how to interpret the Wedge. Participants were also instructed not to move, turn or sway during the experiment to ensure accurate data collection.

Each participant provided 48 recorded exercise trials: 4 interfaces × 4 exercises × 3 trials. In some cases we recorded additional trials when tracking errors occurred with the Vicon system. The study concluded with a questionnaire and semi-open interview on their subjective preferences and experiences using the different conditions.

Exercises

Participants completed four real physiotherapy exercises. These four exercises help rebuild shoulder mobility after injury (e.g., a dislocated shoulder). Our study was designed to examine our systems under distinct and progressively more complex exercises to understand any potential limitations.

We focused on exercises relating to the shoulder, because participants can easily make the movements while standing. The shoulder is also a ball-socket joint (unlike, say, the knee) meaning that a wide range of movements and variation from a prescribed motion is possible (i.e., there is more possibility of error and, therefore, need for guidance). In addition to being an extremely common subject of rehab, the shoulder allowed us to control individual differences in physical abilities between participants: a person only needs to be able to stand and move their arm comfortably. For these reasons we felt the representative set of exercises in shoulder rehab would allow us to go in-depth with a single but potentially useful application of our system design, rather than focusing on a general motion feedback system.

Straight. Abduction of arm along the frontal plane up to shoulder level, followed by adduction of arm back to the participant's side. This is a simple frontal plane exercise.

Angled. Abduction of the arm at 45° from the frontal plane, followed by adduction back to the side. This is an angled variation of the *Straight* exercise, where interpreting the angle may be difficult.

Elbow. External rotation of forearm away from the center of the participant's body until 90° from the sagittal plane, followed by an internal rotation back to center. This exercise requires the participant to keep their elbow tucked against their side and is a difficult exercise to understand without depth cues (i.e., with just a frontal view).

Combo. Abduction of the arm along the frontal plane up to shoulder level, internal rotation of the arm until pointing forward, followed by an external rotation of the arm back to the frontal plane, and adduction of the arm back to the participant's side. This is a more complex exercise than the previous three, involving many components.

Performance measurements

We collected three performance measures: two distance error metrics (one for the hand and one for the elbow), and a measurement of the maximum angle of rotation achieved. We ignore speed as a measure because we are mainly interested in how closely participants can follow an exercise. The two error metrics captures how closely a participant

can follow a pre-recorded exercise delivered either by video or the Wedge system: one for the error from the hand and one from the elbow. For the Elbow exercise, we also recorded a separate metric—the maximum angle reached by participants during the external rotation. Because the Elbow exercise relies on a patient rotating outwards to their farthest extent, we were interested in evaluating how clearly our participants could interpret the required angle with the different interface conditions.

Data Analysis

Performance data were analyzed using 4×4 RM-ANOVA, with interface (VideoSingle, VideoMulti, WedgeSingle, WedgeMulti) and exercise (elbow, combo, angled, straight) as factors. Violations to sphericity used Greenhous-Geisser corrections to the degrees of freedom. Post-hoc tests used Bonneferoni corrections for multiple comparisons; only significant pairwise differences are reported. Post-hoc analysis was only performed to compare levels of the interface condition, as we were only interested in the performance of the different interfaces overall and within the different exercise conditions, and less interested in differences between different exercises. Subjective responses were analyzed using Friedman's test, and post hoc comparisons were done using the Wilcoxon signed-rank test. Before analysis, outlier trials were removed that were > 3 sd. away from the mean for any given exercise, this resulted in the removal of 24 of 1920 records (1.25%).

Results

We first present the performance results—including hand error, elbow error, and maximum rotation—we then present the analysis of subjective response data. We present our observations and the responses we received from participants during our semi-structured interviews in the discussion to help explain our results.

Performance Results

Hand Error: Across all exercises, the WedgeMulti had the lowest mean hand error improving error by ~1.7cms over the baseline VideoSingle (see Figure 5, left). While this difference is not large, it is reduced by the poor performance by all conditions for the elbow exercise, as larger difference can be seen in other exercises (e.g., WedgeMulti reduced error by 50% over VideoSingle in the angled exercise); see Figure 6. There was a significant main effect of both interface ($F_{3,45}$=20.15, $p<.001$) and exercise ($F_{1.77,26.54}$=7.012, p=.005) on hand error. Pairwise comparisons of interface showed that WedgeMulti had significantly
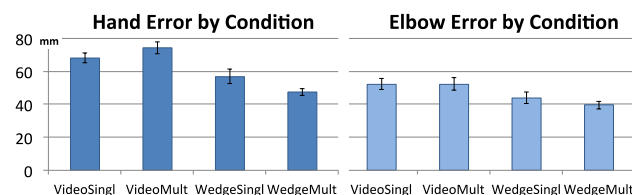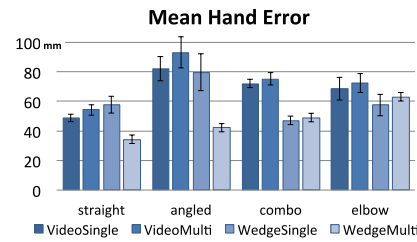

Figure 6: Mean hand error in mm (±SEM) for each interface grouped by exercise. Lower is better.

lower hand error than WedgeSingle ($p<.05$), VideoMulti ($p<.001$), and VideoSingle ($p<.001$). Hand error was lower for WedgeSingle than VideoSingle ($p<.05$).

There was a significant interaction effect between interface and exercise ($F_{4.05,60.69}$=6.026, $p<.001$) for hand error. Within all exercises pairwise comparisons showed that WedgeMulti had consistently lower hand error than VideoMulti ($p<.001$), VideoSingle ($p<.001$), and WedgeSingle ($p<.05$, not combo*)*, with the exception of the elbow exercise, where no differences were observed. For combo, WedgeSingle had significantly lower hand error than both VideoMulti ($p<.001$) and VideoSingle ($p<.001$).

Elbow Error: Overall exercises, WedgeMulti had the lowest mean hand error improving error by ~1 cm over the baseline VideoSingle (see Figure 5, right). However, again this number was reduced by performance in the elbow exercise (see Figure 7). There was a significant main effect of interface ($F_{3,45}$=9.895, $p<.001$) on elbow error. However, there was no effect observed for exercise ($F_{1.5,23.01}$=2.073, $p>.05$) on elbow error. Pairwise comparisons again showed that WedgeMulti had significantly lower elbow error than VideoMulti ($p<.005$) and VideoSingle ($p<.001$), but no other pairwise differences were observed. See Figure 7.

There was an interaction effect objected between interface and exercise for elbow error ($F_{9,135}$=2.091, $p<.05$). Pairwise comparisons within the groups show that WedgeMulti had significantly lower elbow error than VideoMulti ($p<.05$, for combo and straight), VideoSingle ($p<.05$, for angled, combo and straight), WedgeSingle ($p<.05$, straight). The only other pairwise differences observed was for WedgeSingle, which had significantly lower elbow error than VideoMulti ($p<.005$) and VideoSingle ($p<.05$), but just during the combo exercise.
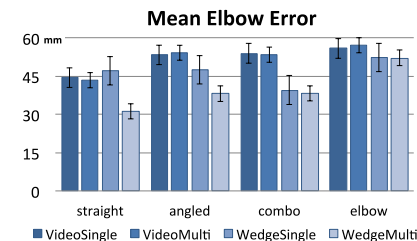

Figure 5: Performance by Condition (Left) Mean Hand Error ±SEM (Right) Mean Elbow Error ±SEM. Lower is better.

Figure 7: Mean elbow error in mm (±SEM) for each interface grouped by exercise. Lower is better.
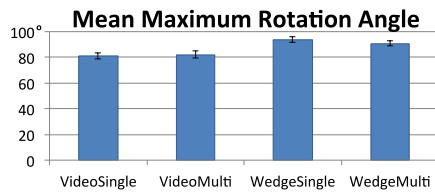
Figure 8: Mean max rotation angle ±SEM. Higher is better.

Rotation Angle: Figure 8 presents the mean maximum rotation angles obtained by participants. Analysis showed a significant effect of interface on max rotation angle during the elbow exercise ($F_{3,45}$=13.285, p<.001). Pairwise comparison showed that participants made significantly higher rotations with WedgeMulti than VideoSingle (p<.05), and that WedgeSingle had higher rotation than both VideoSingle (p<.005) and VideoMulti (p<.005).

Subjective Response Results
At the end of the experiment participants were asked to rank each condition on two criteria. First, participants ranked the interfaces on how accurate they felt the interface allowed them to be. Second, they also ranked the interfaces based on their subjective preference. The mean ranks can be seen in Figure 9, where 4 is ranked highest, and 1 is ranked lowest. We also asked participants their favorite visualization (video / Wedge) and view (single / multiple).

Subjective Accuracy: Analysis found a significant effect of interface on accuracy rankings (χ^2(3)=22.754, p<.001). Pairwise comparisons showed participants felt they were more accurate with WedgeMulti than VideoMulti (z=-2.61, p<.01) and VideoSingle (z=-3.20, p<.001). Participants also felt they were more accurate with WedgeSingle than VideoSingle (z=-2.83, p<.005) and with VideoMulti than VideoSingle (z=-2.97, p<.005).

Subjective Preference: Participants were split on their most preferred methods. The rankings in Figure 9 show no clearly preferred method. There was no effect of interface observed on preference ranking (χ^2(3)=5.0, p>.05).

Preferences for Visualization and Views: 14 participants responded to which group of interfaces they preferred, 9 chose the Wedge visualization interfaces (WedgeSingle and WedgeMulti) and 5 chose the simpler video-only interfaces (VideoSingle and VideoMulti). Eleven indicated they preferred using multiple views, while 5 selected a single view.

Limitations
Our recruited participants were local graduate students rather than on-going physiotherapy patients or seniors. We
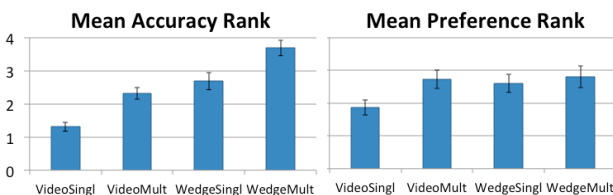

Figure 9: Subjective Rankings (Left) Mean Accuracy ±SEM (Right) Mean Preference ±SEM; 4 is best.

selected this participant pool due to the early state of our prototype and difficulties recruiting and working with these populations. This is notable limitation, as these participants were healthy and may not have the same appreciation for the stricter requirements for physiotherapy patients. However, 5 out of 16 participants had prior physiotherapy experience, and we discuss their experiences below.

Our system was also limited due to the placement of our markers. While our current Vicon setup were able to reliably track our wearable marker patterns, they still required participants to keep the markers visible at all times to ensure tracking would not be lost.

DISCUSSION
Our results show that both the Wedge visualization and multiple views may be needed in combination to improve guidance. Our study highlight four main findings: the Wedge visualization with multiple views performed consistently as the most accurate technique; no technique performed better than any other for the elbow exercise based on our error metrics; both Wedge conditions improved the ability to perform the rotation movements found in the elbow exercise; and despite performing the best and participants feeling that they were most accurate with the Wedge, participants were split on which visualization they prefer.

We discuss these findings below with observations from the study and participant comments from the semi-structured interview to help explain the results.

Why was the Wedge with multi-view the most accurate?
The Wedge interface with multiple views (front-on and top-down) was the most accurate in terms of both hand and elbow errors (for 3 of 4 exercises). Participants effectively interpreted the information from the Wedge, and multiple views added benefit. The required angle for participants' abduction/adduction movements of the exercises were clearly conveyed by the top-down view available with the multi-view version, while the Wedge's Topdown Angle and Nearest Arm guides provided the necessary information to better allow participants to keep their arm aligned.

Visual guidance from the Wedge resulted in participants stopping as soon as any corrective guides appeared, realigning themselves, and then resuming the movement. These actions ensured they would stay on the correct path at all times. Without corrective guides (in video condition), participants had no direct indication of how far off their movements were, and would continue through the exercise.

Why didn't the Wedge perform better in the Elbow exercise?
The Wedge did not perform better than the other interface conditions for the Elbow exercise due to our current implementation. While the Wedge shows the required vertical angle for the other exercises via its Topdown Angle guide, it does not provide an analogue for the front view. This meant participants did not have sufficient feedback for maintaining their horizontal angle. As a result, the Wedge did not perform any better than video.

We also suspect that our mechanism for tracking the elbow (trackers affixed to participants' forearms closest the elbow) caused problems. During rotation, the markers would sometimes shift position, resulting in a potentially misleading change in the visualization of arm position. Because of this problem and the fact that this exercise is mainly focused on rotation of the forearm, we believe evaluating participant performance for the Elbow exercise is best done by the maximum angle of rotation.

Why did the Wedge help for rotation movements?

While Wedge errors for the Elbow exercise were not much lower than the Video, participants were able to rotate roughly 10 degrees farther during the exercise using either Wedge conditions. Using only the Video conditions, the participant can see that they must rotate outwards during the exercise, but not how far they must go. This often resulted in participants stopping early. Both Wedge conditions showed the fully required extent of the rotation; the Movement Arc in the top-down view showed how much rotation is still required, and the Directional Arrow in both views showed when movement in a direction was needed.

Why did people not prefer the Wedge?

While the Wedge conditions were rated the most accurate, preferential rankings were split due to their difficulty and complexity. Even though some participants were able to follow movements more accurately by performing short 'micro-corrections' whenever a guide appeared, these corrections required noticeable attention on their part. The guides appear whenever there is the slightest misalignment and do not disappear until it is corrected, leading to comments that the Wedge was "too strict." Still, some participants felt this could be a potential benefit for a physiotherapy patient, as it would force them to follow exercise movements carefully and pay close attention to the feedback.

A related complaint about the Wedge was that its Nearest Arm guide felt misleading. While this guide helped participants see when their movement was incorrect, it did not tell them how to correct themselves. This resulted in several participants trying to align themselves with the Nearest Arm, but finding it difficult to get their arms in the correct position. The five participants preferring Video felt that it was more straightforward. Participants reported being more comfortable following video because it allowed a more fluid movement—more fun than the stricter Wedge.

How did physiotherapy patients feel about the Wedge?

Five participants had prior physiotherapy experience. Their feedback was consistent: all rankings indicated that Wedge with multiple views as the most accurate; furthermore, three preferred WedgeMulti, one preferred video, and one liked every technique. Two of these participants also indicated that all techniques could have a role in an effective physiotherapy: video provides an easy-to-understand demonstration of the movement that would introduce a patient to a new exercise and help them adjust to a learning curve, while the Wedge would be helpful later in follow-up ses-

sions to understand the finer-grained movement characteristics. This feedback is important as exercises are sometimes painful to perform correctly—the feedback provides reassurance that things are moving forward (P14). Finally, these participants were notably more forgiving of momentary tracking and visualization errors by the Wedge, because it helped overall with accuracy.

Challenges for Design

Our findings build on prior work by demonstrating that on-screen visualizations *can* work for high-precision movements, e.g. in physiotherapy. Our process has helped us to uncover several design implications and challenges for systems that guide or correct movement and exercises through visual feedback overlaid on live video (e.g. [3]).

Reducing Visual Complexity. Even with the simple movements that we were exploring in our work, participants were sometimes overwhelmed with the complexity and amount of information presented. Given the guidance characteristics outlined in Table 1, this should probably not be too surprising—it is possible to perform exercises incorrectly in many different ways. Enabling all aspects of the Wedge visualization at once means that there may be simply too many things to attend to at once.

Balancing visual complexity with sufficient guidance is the central challenge to our approach. In our work, we explored one way to do this, by separating different parts of an exercise into segments, and presenting only the parts of the visualization that were relevant given where the participant was in the exercise. Yet a scaffolding approach might be appropriate too: for instance, providing only the wedge part of the visual guide until a person has mastered the basic motion—once this is accomplished, then showing visualizations that illustrate the extents, and so forth. This would only show visualizations piecemeal so as to not overwhelm the user. Another approach might be to show only corrective feedback visualizations (e.g. Nearest Arm) once a person had already mastered the exercise.

Semantics of Visual Illustration. Although we articulate the aspects of movement guidance that need some sort of visualization feedback (Table 1), we did not thoroughly explore the space of visual illustration. While our visual guides are somewhat rudimentary, they were designed strictly to address each aspect of the movement guidance framework that we outlined. Nevertheless, a lot of work has explored how illustrative techniques from comics can be used to suggest, convey, and communicate movement (e.g. [15]). Exploring this rich vocabulary may lead to more concise and intuitive visualizations.

Depth Perception and Capture. We also saw that participants had a hard time interpreting depth easily from video (and to a lesser extent, our visualization). This is perhaps a consequence of the 2D camera capture and the 2D representation/rendering on the screen of the video. In practice, we use head motion to gain a better understanding of depth

when, for example, looking in a mirror. It may be possible that combining head tracking with depth video capture will allow people to better understand what is going on in regard to the playback video and/or the visualization.

Camera Placement. A single video feedback window is insufficient to accurately perceive movement—both one's own movement, and that of the on-screen guide. The exception to this is when the movement happens strictly on the plane perpendicular to the camera. In the general case, multiple cameras—particularly when one is perpendicular to movement—are useful to improve the accuracy of movement. From a hardware standpoint, this is not far-fetched: most modern phones come equipped with powerful cameras that could be used. One trade-off is that multiple cameras add to the visual clutter and complexity of the scene.

CONCLUSIONS

Physiotherapy patients exercising at home do not have the benefit of guidance and feedback, and there is a strong possibility of re-injury with incorrect exercise movements. Physio@Home explores the use of a dynamic on-screen movement guide called the Wedge and multiple camera views to guide movements, with attention to supporting depth perception and precision. We found that participants performed exercises with the least error using the Wedge and multi-camera views. From this, we identified several characteristics required for accurate movement guidance, and challenges for exercise guidance systems. Physiotherapy services will continue to be in high demand as the population ages. With increasingly capable and inexpensive motion tracking cameras on their way, the concepts from Physio@Home will be able to help meet the needs of physiotherapy patients in the future.

ACKNOWLEDGEMENTS

REFERENCES

1. Alankus, G., et al. (2010) Towards Customizable Games for Stroke Rehabilitation. *CHI*, 2113-2122.
2. Ananthanarayan, S., et al. (2013) Pt Viz: towards a wearable device for visualizing knee rehabilitation exercises. *CHI*, 1247-50.
3. Anderson, F., et al. (2013) YouMove: enhancing movement training with an augmented reality mirror. *UIST*, 311-320.
4. Ayoade, M. & Baillie, L. (2014) A novel knee rehabilitation system for the home. *CHI*, 2521-2530.
5. Bau, O., Mackay, W.(2008)OctoPocus: a dynamic guide for learning gesture-based command sets. *UIST*, 37-46.
6. Campresi, C., et al. (2014). Motion Parameterization and Adaptation Strategies for Virtual Therapists. *IVA*. 99-108.
7. Canadian Physiotherapy Association. http://www.physiotherapy.ca/About-Physiotherapy
8. Diddlyi. Online Irish Dance Workshops. http://www.diddlyi.com
9. Doyle, J., et al. (2010). BASE - An interactive technology solution to deliver balance and strength exercises to older adults. *PervasiveHealth*, 1-5.
10. Freeman, D., et al. (2009) ShadowGuides: visualizations for in-situ learning of multi-touch and whole-hand gestures. *ITS*, 165-172.
11. Henderson, S.J., & Feiner, S.K. (2011). Augmented reality in the psychomotor phase of a procedural task. *ISMAR*, 191-200.
12. Huang, J-D. (2011) Kinerehab: a kinect-based system for physical rehabilitation: a pilot study for young adults with motor disabilities. *ASSETS*, 319-320.
13. Lee, J-D., Hseih, C-H., & Lin, T-Y. (2014) A Kinect-based Tai Chi exercises evaluation system for physical rehabilitation. *ICCE*, 177-178.
14. Marquardt, N., et al. (2011) The Proximity Toolkit: Prototyping Proxemic Interactions. *UIST*, 315-326.
15. McCloud, S. (1993) Understanding Comics.
16. Nicolau, H., et al. (2013). Computer-assisted rehabilitation: towards effective evaluation. IJCPS, 11-26.
17. Nixon, M.E., et al. (2013) Quantitative evaluation of the Microsoft KinectTM for use in an upper extremity virtual rehabilitation environment. *ICVR*, 222-228.
18. PTMotions. Program Viwer. http://ptmotions.com/ptm_pv.html
19. Rector, K., et al. (2013) Eyes-free yoga: an exergame using depth cameras for blind & low vision exercise. *ASSETS*. 8 pages.
20. Services Canada. (2013) Physiotherapists. http://www.servicecanada.gc.ca/eng/qc/job_futures/statistics/3142.shtml
21. Sodhi, R., et al. (2012) LightGuide: projected visualizations for hand movement guidance. *CHI*, 179-188.
22. Tang, R., et al. (2014) Physio@Home: Design explorations to support movement guidance. *CHI EA,* 1654-1656
23. Tao, G., et al. (2013) Evaluation of Kinect skeletal tracking in a virtual reality rehabilitation system for upper limb hemiparesis. *ICVR*, 164-165.
24. Uzor, S. & Baillie, L. (2014). Investigating the long-term use of exergames in the home with elderly fallers. *CHI*, 2813-2822.
25. Velloso, E., et al. (2013) MotionMA: motion modelling and analysis by demonstration. *CHI*, 1309-1318.
26. Vines, A. (2013) Rehabilitation after shoulder dislocation: information for patients. Prepared for UK National Health Services, Oxford Radcliffe Hospitals
27. White, S., et al. (2007) Visual Hints for Tangible Gestures in Augmented Reality. *ISMAR,*1-4.
28. Yang U. & Kim, G-J. Implementation and evaluation of "just follow me": an immersive, VR-based, motion-training system. Presence, vol. 11, issue 3, 304-323
29. Yeh, S-C., et al. (2012) A lower limb fracture postoperative-guided interactive rehabilitation training system and its effectiveness analysis. Healthcom, 149-54