Visuomotor Adaptation and Sensory Recalibration in Reversed Hand Movement Task

Jenny Lin\textsuperscript{1,}\textsuperscript{⋆} Yixin Zhu\textsuperscript{1,}\textsuperscript{⋆} James Kubricht\textsuperscript{2,} Song-Chun Zhu\textsuperscript{1} Hongjing Lu\textsuperscript{1,2}

\textsuperscript{1} Department of Statistics \quad \textsuperscript{2} Department of Psychology

\textsuperscript{⋆} Equal Contributors University of California, Los Angeles

Abstract

Visuomotor adaptation plays an important role in motor planning and execution. However, it remains unclear how sensorimotor transformations are recalibrated when visual and proprioceptive feedback are decoupled. To address this question, the present study asked participants to reach toward targets in a virtual reality (VR) environment. They were given visual feedback of their arm movements in VR that was either consistent (normal motion) with the virtual world or reflected (reversed motion) with respect to the left-right and vertical axes. Participants completed two normal motion experimental sessions, with a reversed motion session in between. While reaction time in the reversed motion session was longer than in the normal motion session, participants showed the learning improvement by completing trials in the second normal motion session faster than in the first. The reduction in reaction time was found to correlate with greater use of linear reaching trajectory strategies (measured using dynamic time warping) in the reversed and second normal motion sessions. This result appears consistent with linear motor movement planning guided by increased attention to visual feedback. Such strategical bias persisted into the second normal motion session. Participants in the reversed session were grouped into two clusters depending on their preference for proximal/distal and awkward/smooth motor movements. We found that participants who preferred distal-smooth movements produced more linear trajectories than those who preferred proximal-awkward movements.

Keywords: Virtual reality; motor planning; scene representation; visual misalignment

Introduction

Virtual Reality

Virtual reality (VR) technology provides an analog experience in a three-dimensional environment similar to that of the real world. In the real world, certain environmental factors and physical constraints are fixed and cannot be modified. However, VR allows researchers to design controlled virtual environments with ease and precision. In addition, modern advancements in VR tracking allow for accurate measurements of human body movements. Thus, task success, motor error and correspondence with candidate trajectories can be accessed directly.

Although previous studies in VR have focused primarily on hardware problems in order to improve user experience (Shotton et al., 2013; Weichert, Bachmann, Rudak, & Fisseler, 2013), simulation performance (e.g., Unreal Engine 4, Unity3d, and NVidia Flex), system integration (Lin et al., 2016; Shah, Dey, Lovett, & Kapoor, 2017), and locomotion in immersive experience (Bruder & Steinicke, 2014), recently efforts have been increasingly devoted to examining human perception and reasoning in virtual scenes (e.g., Azmandian, Hancock, Benko, Ofek, & Wilson, 2016; Mehra et al., 2016; Patney et al., 2016; Ye et al., 2017; Li, Liang, Quigley, Zhao, & Yu, 2017).

Motor Planning

The process of reaching toward an object in the environment involves minimizing the distance between the hand and target locations in the physical world (i.e., the hand and target states) over time. This is achieved by (1) planning a motor movement to achieve a desired task goal, (2) sending the associated motor commands to the arm, and (3) comparing observed sensory feedback to predicted sensory feedback to infer the current hand state and form subsequent motor commands (i.e., sensorimotor transformation; Battaglia-Mayer et al., 2014; Wolpert, 1997). The present study examined how reaching movements change in response to misaligned sensory feedback in a VR environment. Specifically, how do reaching trajectories change as visual and proprioceptive feedback are decoupled?

When visual and proprioceptive feedback are inconsistent, new mappings between visual and proprioceptive inputs are reestimated (Cressman & Henriques, 2009). Results from Cressman and Henriques’s (2009) study suggest that in addition to sensorimotor recalibration, visuomotor adaptation involves partial proprioceptive recalibration: i.e., humans “realign proprioceptive estimates of hand position to match visual estimates.” However, it has been demonstrated that visuomotor adaptation can occur in the absence of proprioceptive input, for example, in the case of deafferented individuals (Ingram et al., 2000; Miall & Cole, 2007). It is therefore possible that proprioceptive recalibration does not underlie visuomotor adaptation and that the two processes are independent from one another. This hypothesis is consistent with empirical results showing that humans curtail the contribution of proprioceptive input in the case of misaligned visual feedback (Bernier, Burle, Vidal, Hasbroucq, & Blouin, 2009; Wont & Henriques, 2009) when performing motor movements. Thus, when visual and proprioceptive feedback are inconsistent, people could reduce the contribution of proprioceptive information to the motor planning process and form new visuomotor transformations to achieve extrinsic goals.

In this study, participants reached toward targets in a virtual environment where their hand movements were shown to be either consistent or reversed (in the vertical and left-right axes) with respect to virtual movement. If proprioceptive inputs are ignored (perhaps due to their unreliability in the reversed movement environment), participants should rely more heavily on visual inputs when planning and executing movements. Moreover, we expect participants will adapt to the reversed environment by constructing and implementing new visuomotor mappings. Although we pre-
In opposite directions (i.e., in both the vertical and left-right axes), the virtual controller in VR was moved in the virtual environment. To generate the visual display of in size and shape were used to represent controller position. 3D meshes which matched the Vive motion controllers were recorded and analyzed.

Experiments

In the present study, we examined whether humans can adapt to environments where visual estimates of objects’ positions are inconsistent with proprioceptive input. Participants interacted with virtual targets using two motion controllers in a VR application, where the movement of the virtual controller either matched the motion of the physical controller or was flipped on certain axes (both vertical and left-right). Participants were instructed to touch a series of virtual targets with the virtual controllers and then return to a neutral pose in between targets. Response time and arm movement trajectories were recorded and analyzed.

Participants and Apparatus

A total of 20 participants (10 female and 10 male) participated in the study. Participants were graduate students at the University of California, Los Angeles. The average age of participants was 22.8 years old with a standard deviation of 2.67. All participants had normal or corrected-to-normal vision. Of the 20 participants, 16 had never interacted with VR technology prior to participating in the experiment.

The VR system integrated Unreal Engine 4 with an HTC Vive headset and two motion controllers, one held in each hand. 3D meshes which matched the Vive motion controllers in size and shape were used to represent controller position in the virtual environment. To generate the visual display of reversed movement, the virtual controller in VR was moved in opposite directions (i.e., in both the vertical and left-right axes) to the physical displacement of the controller moved by human participants in the real world (Fig. 1). Participants began the experiment by moving their hands into a neutral pose where the physical and virtual controllers were aligned to the same position. Movement along the depth axis (i.e., forward vs. backward) was not reversed.

The targets were cyan capsules of 20 cm height and diameter. We chose cyan as the color of the targets in order to ensure the targets would be visible against the background of the environment. The targets began glowing when touched by a controller, providing visual feedback to the subject indicating whether they had successfully touched the target. The color of the targets did not change between experimental sessions. To ensure that for any given target location each participant reached approximately the same distance, we required that the participants assume a neutral pose before the next target was spawned. We define the neutral pose as follows:

At the beginning of each testing block, participants were told to hold both controllers in front of them at waist level with their elbows held loosely at their sides. Participants were allowed to adjust their pose until they were comfortable, but were informed that they needed to be able to comfortably reach forward, up-down and side-to-side from this pose. Participants then started an experimental block by pressing the trigger button on the bottom of either motion controllers. A transparent rectangular prism was spawned such that its center was located at the midpoint of the two controllers. This rectangular prism defined each participant’s neutral zone, and we considered the participant to be in a neutral pose when both controllers overlapped with the neutral zone for an uninterrupted 0.5 seconds. In order to provide feedback to the user about whether they were in a neutral pose, the neutral zone changed color to reflect how many controllers overlapped with it: black for zero controllers, grey for one, and green for two. The neutral zone only changed color when the participant needed to enter a neutral pose, and otherwise remained green.

Response time was defined as the duration between the initial spawning of the target to when it was deactivated. Trajectory was defined as the three-dimensional movement of the controllers over this time period. For a video demonstrating the experimental setup, please see https://vimeo.com/216580864.

Procedure

The experiment was conducted in a quiet office, and all physical obstacles were removed from the testing area. Participants remained standing and stationary for the duration of the experiment. They received a warning signal if they moved near the boundaries of the virtual environment.

Practice Session First, participants familiarized themselves with the VR headset and motion controllers. Participants were given a demonstration of the neutral position and told to move both of their controllers to the indicated locations. After participants confirmed that they were capable of comfortably performing the required range of movements from their neutral pose, they were informed that both response time and movement trajectories would be recorded. Prior to the testing session, participants completed a practice session with five
Participants completed nine blocks consisting of ten trials each. The first three blocks (N1 session) were completed with normal movements. The subsequent three blocks (R1 session) were completed with reversed movements. Participants were informed along which axes arm movement would be reversed (i.e., the left-right and up-down directions). The last three blocks (N2 session) were completed with normal movements once again. Participants were given breaks between blocks to rest their arms. After indicating that they were ready to continue, participants proceeded to the subsequent block.

At the start of each block, the virtual meshes were aligned with the locations of the physical controllers. Each participant completed the same nine blocks. Target locations were evenly distributed throughout an 80 \times 80 cm region located 35 cm in front of the neutral zone. The order of the target positions within each block was randomized between participants.

**Results**

**Response Time Analysis**

As expected, participants showed much longer response times (RT) in the reversed-motion condition than in the normal-motion condition. There was a four-fold increase between median RT for the N1 relative to the R1 session. Interestingly, upon returning to normal movement in the N2 session, participants showed a 20.1% improvement in response time compared to the N1 session (t(600) = 7.07, p < .001; see Fig. 2a). Moreover, response times in the three blocks of the N2 session displayed a decrease in the N2 session that was not evident in the N1 session (Fig. 2b), indicating a learning effect that was not present in the N1 session (b = −0.0067[−0.0101, −0.0033]), indicating a learning effect that was not present in the N1 session (b = −0.0067[−0.0101, −0.0033]).

**Trajectory Analysis**

Next, trajectory analysis was performed to further quantify human performance relative to candidate trajectories. We define the baseline trajectory as the shortest linear path between the hand start position and the target location. All trajectories were interpolated to 500 3D points to account for variation in trajectory length. Dynamic Time Warping (DTW) was then utilized to determine the minimum distance mapping between the ideal and behavioral trajectories. DTW is a distance measure algorithm that has been used extensively in the speech recognition community (e.g., Berndt & Clifford, 1994). By estimating a non-linear mapping between two time-dependent sequences, DTW provides a numerical representation of the similarity between any pair of spatiotemporal sequences. Other communities including robotics and biology have also adopted and modified this algorithm for various signal-comparison applications.

The DTW trajectory distance measure revealed closer correspondence to baseline trajectories in the N2 session compared to the N1 session (Fig. 3a), suggesting a learning effect through practice. There was also a clear decrease in DTW distance across the three blocks within the N2 session that was not evident in the N1 session (Fig. 3b), suggesting humans moved their arms more linearly (i.e., closer to the baseline linear trajectory) upon return to the normal motion environment. To rule out the possibility that the increasing linear movements in the N2 session was due to familiarization with the VR system, we performed a linear regression on the median trajectory difference among participants (Fig. 4). Although there is no noticeable trend in the N1 session, performance in the N2 session shows a strong improvement that falls well outside the 95% confidence region for N1. Moreover, the slope in N2 was approximately equal to that in the R1 session, although the regression coefficient in the R1 ses-

![Figure 2: Response time analysis for the normal- (N1 and N2) and reversed-motion (R1) trial sessions. Red horizontal lines indicate median response times. The bottom and top edges of the blue boxes indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points that were not considered outliers, and red ‘+’ symbols indicate outliers. (a) Session median times to reach targets: 1.29, 5.71 and 1.03 seconds. (b) Block median times to reach targets: 1.30, 1.22, 1.34, 6.10, 5.68, 5.36, 1.13, 1.06 and 0.94 seconds.](image1)

![Figure 3: Trajectory analysis using DTW to quantify the discrepancy between human reaching behavior and a linear motion trajectory – the straight line between the hand’s starting position and target location. Median distance scores for each session: 505.62, 613.24, and 483.55 cm. Block medians: 501.03, 525.70, 496.62, 671.75, 648.25, 582.34, 513.57, 490.28, and 435.84 cm.](image2)
Figure 4: Linear regression results using median DTW distance among 20 subjects across 90 trials divided into 3 sessions and 9 blocks. Red dashed lines represent 95% confidence intervals for the regression coefficient estimates. Slopes in the three sessions are 0.216, −3.810, and −4.436. Session is more uncertain: i.e., the confidence interval of the R1 slope is greater than that of N2. This suggests a large degree of within-group variability, which is further explored in the following sections.

After forming new visuomotor mappings in the R1 session, participants’ movement trajectories became increasingly linear: i.e., closer to the baseline trajectory. If participants began relying on visual feedback when constructing and revising their motor plans (i.e., proprioceptive inputs were suppressed), we would expect them to execute linear movement paths. The increasingly linear motor movements over the course of the R1 session are consistent with this prediction. Interestingly, reliance on visual inputs appeared to persist in the following N2 session when proprioceptive and visual information were recoupled. We predict that with further exposure to the normal-motion environment, the linearity of participants’ reaching patterns would return to the level measured in the N1 session.

Possible Planning Models in Reversed Motion Blocks

While the shortest linear path between two points is the most direct trajectory, it is not necessarily the most optimal reaching strategy: e.g., due to mechanical limb constraints. To examine this, we used DTW to compare against other candidate trajectories to assess their potential as possible movement strategies. One possible alternative strategy is to consider each axis independently in order to plan motor movements in the reversed motion condition. To examine this alternative strategy, human trajectories were compared to all six possible axis decompositions (Fig. 5a) using DTW. While some participants did demonstrate paths that were more similar to various axis decompositions, participants’ trajectories were generally more similar to the shortest linear path (Fig. 3b), indicating that most participants were not considering each axis independently.

Another observation of participants’ trajectories is that they were noisy, especially during the reversed motion session. Since participants were instructed to reach a set of given targets, their movements were goal-directed and partially guided. We compared participants’ trajectories with predictions from a guided random walk model Pearson (1905).

Figure 5: (a) Six possible axis decompositions were generated by computing the shortest path along each axis. (b) Human trajectories were compared against all six axis decompositions using DTW, and the minimum value was reported. Session medians: 736.72, 734.38, and 661.90 cm. (c) 10 of the guided random walks generated between the given start and end point. (d) Human trajectories were compared against 100 guided random walks using DTW, and the most similar value was reported. Block medians: 315.03, 307.15, 300.95, 522.34, 515.89, 452.21, 310.25, 278.97, and 247.46 cm.

Given a starting point, a set of 100 proposed moves were generated within a 5 cm radius. Next, the model computed the distance between each of the proposed movements and the end point. A movement was then chosen from two options: 1) the shortest distance with probability .2, or 2) randomly chosen movement among the 100 (random) proposed movements with probability .8. Finally, after approximately a few hundred iterations, the guided random walk model converged and reached the end point, as shown in Fig. 5c. Measured by DTW, human movement trajectories were found to be more similar to the guided random walks not only during the reversed-motion session but also during both normal-motion sessions (Fig. 5d). The fit of the model predicted trajectories to human performance across all the three sessions suggests that participants’ motor movements were goal-directed but executed with inherent motor noise.

Movement Strategies in Reversed Motion Blocks

In the normal-motion sessions, participants consistently used both arms to perform the reaching task, while favoring the controller closest to the target. In the reversed-motion session, however, a variety of strategies emerged. Some participants predominantly used one hand regardless of the location of the target relative to their neutral zone. Others favored the hand that was furthest from the target. Thus, we further examined the distribution of participants’ reaching strategies.

In certain experimental trials, touching the target with the nearest hand required the participant to reach across their body while looking in the opposite direction, due to the reversed axes. This pose is physically difficult to accomplish. In contrast, the participant could reach for the target with their opposite hand, resulting in a pose that was physically comfortable. However, this would require the participant to use the hand that was physically furthest away, which is highly nonintuitive (Fig. 6). The cost to execute a path is thus dependent on not just proximity but also kinesthetic ease of execution.

We examined the interplay between the two constraints (i.e., proximity and ease of motor execution) in planning motor movements. Criteria were defined as follows: a trajectory
is considered proximal if a participant uses the hand initially closest to the target, and considered distal if he uses the hand initially furthest from the target. The trajectory is considered awkward if it requires reaching across the body’s center and smooth if it does not. These criteria result in four different trajectory categories: proximal-smooth, proximal-awkward, distal-smooth, and distal-awkward (See Fig. 7). In the normal-motion sessions, participants strongly favored the proximal-smooth strategy, with the distal-awkward strategy occurring only in a few selected trials where the target was close to the mid-line. In the reversed motion session, participants demonstrated all three strategies except the distal-awkward.

We performed k-mean clustering on participants’ trajectories in the reversed-motion session and found that two stable clusters emerged. Cluster size was split evenly at ten participants each, indicating that half of the participants were more likely to use the proximal-awkward strategy and the other half were more likely to use the distal-smooth strategy. The former group favored visual proximity: i.e., they attempted to reach the target using the hand that was closest to the target. The latter group favored smooth motion: i.e., they used learned associations between proprioceptive feedback and visual movement to predict which hand choice would result in the least awkward pose. In this case, participants were required to imagine the potential trajectories and associated proprioceptive feedback to plan their movement. These findings suggest that humans adopt different strategies to cope with the novel task in the reversed motion session by focusing on either spatial proximity for efficiency or smooth motion to avoid impossible or awkward poses.

A linear regression analysis was performed on DTW measurements after separating participants into the two groups as shown in Fig. 9. It is clear that the pose-focused participants demonstrated greater improvement compared to proximity-focused participants, although this learning effect did not persist in the subsequent normal-motion session.

**Discussion**

When planning motor movement according to misaligned visual feedback, proprioceptive feedback has been shown to be suppressed while attention to visual information is enhanced. We hypothesized that in the case of reversed virtual feedback, target-directed reaching movements would rely primarily on visual feedback and thus accord with candidate linear trajectories. This prediction is confirmed by participants in the reversed-motion session using only a single hand, which arguably arises due to the relative ease of forming new visuo-motor mappings with a single arm compared to both arms simultaneously. We found that participants in the reversed-motion session (R1) exhibited a preference for linear trajectories, which agrees with increasing suppression of using proprioceptive information to guide motor movements. Interestingly, this increasing linear preference—and corresponding reliance on newly formed visuo-motor mappings—persisted into the second normal motion session (N2) although it was not observed in the first normal session (N1). We predict that this bias toward linear movement strategies would diminish with
further exposure to the normal-motion environment, as traditional sensorimotor mappings utilizing proprioceptive information are employed.

However, the main finding of the present study could have resulted from increased familiarity with the VR system and environment. Thus, a follow-up study to this experiment is to establish a second control condition where each of the three experimental sessions involve normal motion. If performance does not vary across the three normal sessions, the finding that reversed motion increases preference toward visually-guided, linear motor movements would be strengthened. Additionally, movement in the virtual world was reversed on two axes (vertical and left-right) in the present study. Future work should examine how performance changes when a single axis—or different pairs of axes—are flipped. Moreover, would exposure to one reversed axis improve performance under a second (different) reversed axis?

Tactile signals are an important cue for planning and executing object interactions (Johansson & Flanagan, 2009). One of the major disadvantages with current commercial VR products is that tactile feedback is missing in the virtual world. In the present study, we compensated for the lack of tactile feedback by using additional visual cues to indicate successful reach events; however this does not change the fact that a significant source of feedback is missing. For future studies it would be worth providing a haptic signal through the controller’s actuators or using a tactile data glove to administrate more fine-grained feedback. We predict that implementing haptic feedback to the current experimental method would inhibit suppression of proprioceptive information and consequently interfere with the formation of new visuomotor mappings.

Future work should also examine sensorimotor recalibration in more complicated tasks than the present reaching movements: e.g., stacking blocks or completing towers of Hanoi problems. In these tasks, cognitive resources are devoted to planning a sequence of motor movements, which may yield strong interference to the visuomotor adaptation process and provide a unique window to study the interplay between motor planning and reasoning.

Acknowledgments The work reported herein was supported by DARPA XAI grant N66001-17-2-4029, DARPA SIMPLEX grant N66001-15-C-4035, ONR MURI grant N00014-16-1-2007, NSF grant BCS-1353391, and a NSF Graduate Research Fellowship.

References