

RESEARCH ARTICLE

Society for the Neural Control of Movement

Taking aim at the perceptual side of motor learning: exploring how explicit and implicit learning encode perceptual error information through depth vision

 Carlo Campagnoli,¹ Fulvio Domini,² and  Jordan A. Taylor¹

¹Department of Psychology, Princeton University, Princeton, New Jersey and ²Department of Cognitive, Linguistic and Psychological Sciences, Brown University, Providence, Rhode Island

Abstract

Motor learning in visuomotor adaptation tasks results from both explicit and implicit processes, each responding differently to an error signal. Although the motor output side of these processes has been extensively studied, the visual input side is relatively unknown. We investigated if and how depth perception affects the computation of error information by explicit and implicit motor learning. Two groups of participants made reaching movements to bring a virtual cursor to a target in the frontoparallel plane. The Delayed group was allowed to reaim and their feedback was delayed to emphasize explicit learning, whereas the clamped group received task-irrelevant clamped cursor feedback and continued to aim straight at the target to emphasize implicit adaptation. Both groups played this game in a highly detailed virtual environment (depth condition), leveraging a cover task of playing darts in a virtual tavern, and in an empty environment (no-depth condition). The delayed group showed an increase in error sensitivity under depth relative to no-depth. In contrast, the clamped group adapted to the same degree under both conditions. The movement kinematics of the delayed participants also changed under the depth condition, consistent with the target appearing more distant, unlike the Clamped group. A comparison of the delayed behavioral data with a perceptual task from the same individuals showed that the greater reaiming in the depth condition was consistent with an increase in the scaling of the error distance and size. These findings suggest that explicit and implicit learning processes may rely on different sources of perceptual information.

NEW & NOTEWORTHY We leveraged a classic sensorimotor adaptation task to perform a first systematic assessment of the role of perceptual cues in the estimation of an error signal in the 3-D space during motor learning. We crossed two conditions presenting different amounts of depth information, with two manipulations emphasizing explicit and implicit learning processes. Explicit learning responded to the visual conditions, consistent with perceptual reports, whereas implicit learning appeared to be independent of them.

depth perception; explicit learning; implicit learning; virtual reality; visuomotor rotation

INTRODUCTION

Adaptation is a fundamental process that helps maintain calibration of the motor system in a constantly changing environment. Although important progress has been made over the past 20 years to characterize the error sensitivity of the motor system (1–4), the input side of the adaptation process remains less explored.

One aspect of this issue concerns how visual error information gets encoded for motor learning in the first place. Most visuomotor adaptation studies have been constrained

to a two-dimensional (2-D) plane, such as that of a monitor or a flat frontoparallel surface, to display feedback of reaching (i.e., rotations), to limit any potential discrepancies between the visual representation and the physical world. However, most everyday life actions take place in a wider and three-dimensional (3-D) environment, where the relationship between the visual input and the physical world cannot be determined as straightforwardly. Depth information accurately is therefore critically important when reaching in 3-D space, both in terms of motor planning and error correction. This unveils an interesting puzzle.

First off, depth vision is far from being accurate and, in fact, it often leads to perceptual distortions of space (5–10), raising the question as to what information the motor system relies on when scaling error size information to adjust to visual perturbations in a 3-D environment. Second, motor learning is not a unitary process. Sensorimotor adaptation of goal-directed actions is known to be shaped by the interplay between explicit learning and implicit motor adaptation processes (11–13). Although both can be defined as error-minimization processes, numerous findings indicate that each process operates on different types of error information (14–37). It remains an open question as to whether explicit and implicit learning leverage different types of visual depth cues to estimate movement errors.

Here, we set out to explore if and how explicit and implicit processes incorporate multiple sources of depth information during sensorimotor adaptation, through a 3-D version of a visuomotor rotation task performed in virtual reality. Participants were asked to make horizontal, planar-reaching movements on a tabletop, starting from the center of the chest and moving outward in the forward direction. They aimed at a virtual target seen at varying distances in front of them at eye height. End point feedback was provided about the accuracy of the reach also on the frontoparallel plane, but it was manipulated so to appear either to the right or to the left relative to the hand's direction. To compensate for the perturbation, participants were required to adjust the direction of the forward reaching, either clockwise (for a leftward offset on the frontoparallel plane) or counterclockwise (for a rightward offset).

Explicit and implicit learnings were defined also in the context of the visuomotor rotation task, as the amount of reaiming consciously imparted by the participant, versus the amount of rotation developed involuntarily over time due to sensory-prediction errors. The two components were experimentally dissociated: one group of participants performed the task while receiving delayed end point feedback, which appears to weaken implicit adaptation and require explicit learning, another group received clamped end point feedback, which minimizes explicit learning to effectively isolate implicit adaptation (16, 38, 39).

To determine the potential influence of visual depth cues on explicit and implicit learning, the experiment had a 2×2 factorial design contrasting two viewing conditions against two types of end point feedback. For the viewing conditions, we tested a depth condition, where participants were immersed in a full scale virtual “tavern” scene, simulating a number of depth cues (perspective, texture, stereo disparity, objects of familiar size, etc.) common to the real world, and a no-depth condition, where participants were immersed in a nondescript gray environment in which the target was the only object visible.

To measure sensitivity to the error size, we varied the direction and magnitude of the rotation every two trials, as the experiment unfolded as a series of two-reaches sequences: participants in both groups (delayed and clamped) were instructed to aim straight at the target's bull's eye once (the “probe” reach), and then either reaim or keep aiming straight in the second trial (the “test” reach), respectively. The critical measure of error sensitivity was the degree of change from the probe reach to the test reach, with special interest in how

depth information could potentially influence error size interpretation by the explicit and implicit processes.

METHODS

Participants

Forty individuals participated in the experiment (15 females; age between 18 and 35 yr). They were equally divided into two groups of 20 for the motor task, whereas all 40 did the same perceptual task. The sample size for the motor task was determined through power analysis based on a previous study (40) that compared the slope of the error sensitivity function of both explicit and implicit learning. Given an effect size of 1.19, the minimum number of participants required to detect a difference between the two learnings through a two-sample *t* test was 19 per group ($\alpha = 0.05$, power = 0.95). The participants self-reported normal or corrected-to-normal vision. Handedness was measured with the Edinburgh Handedness Inventory. Participants received either a reimbursement of \$12/h or coursework credit for volunteering. The experiment was approved by the Institutional Review Board at Princeton University and each participant gave written informed consent before the experiment.

Apparatus and Procedure

Participants sat at a table with their head on a chinrest, which could be adjusted to a comfortable position such that the gaze pointed straight to the horizon. The visual scene was displayed through a head-mounted virtual reality system (HTC Vive) with a resolution of $1,440 \times 1,600$ pixel per eye with a 90 Hz refresh rate. The virtual scenes were modeled in Blender (www.blender.org) and rendered in the viewport of the Unity game engine (unity.com), where they were animated through custom C# routines. The experimenter first measured the participant's interpupillary distance (IPD) using a digital pupillometer, then helped her or him wear the VR system's headset, adjusting the spacing between the lenses to match the participant's IPD. Next, the participant performed a stereovision test using a custom program, followed by a short series of familiarization routines to get acquainted with the virtual environment, and finally the actual experiment began.

Viewing Conditions

The experiment included a perceptual and motor task. In both tasks, the target stimulus was a virtual dartboard and two visual conditions were tested: a no-depth condition, where the target was presented in isolation against a gray background, and a depth condition, where the target appeared at the end of a full-scale tavern's hall (Fig. 1A). Although both conditions were viewed in stereovision, they differed substantially in the strength and amount of available depth information. The uniform background of the no-depth condition did not allow the eyes to resolve binocular disparity across almost the entire visual field. The only exception was represented by the area subtended by the target, whose size was chosen such that, within that region, binocular disparity would be negligible. Under this condition, participants could determine distance by relying primarily

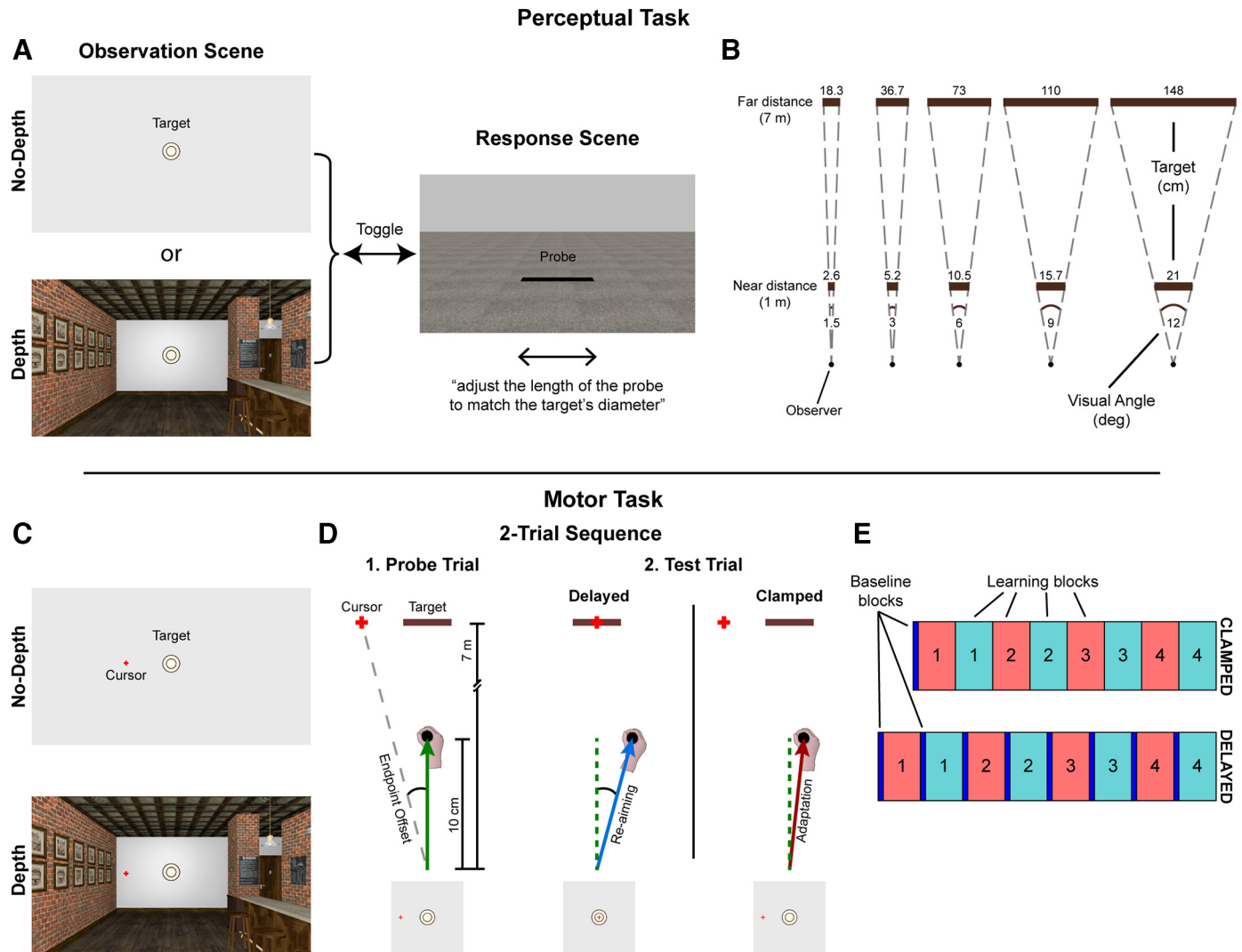


Figure 1. Methods of the study. **A:** in the perceptual task, participants toggled at will between an observation scene and a response scene. During observation, they saw a dartboard either in isolation (no-depth condition) or inside a tavern (depth condition). Only one condition was presented in each trial. In the response scene, they adjusted the length of a line to estimate the dartboard's diameter. **B:** bird's eye geometry of the stimulus set in the observation scene. In each trial, one out of 10 possible target objects was presented (brown rectangles), resulting from the combination of five sizes (in cm in the figure) and two egocentric distances (1 m or 7 m). The size and the distance were calculated such that the target subtended one of five possible visual angles (in deg). Each visual angle was therefore projected from either a near distance or a far distance. **C:** during the motor task, participants threw virtual darts at the dartboard by making small reaching movements and triggering the appearance of a small cross (Cursor), which provided feedback about the hand's direction. **D:** bird's eye geometry of the two-trial sequence executed in the motor task. The probe trial (1) came first: All participants were instructed to aim at the target's center by reaching 10 cm in the forward direction (green arrow), and a red cross appeared at the same depth of the target, showing the hand's direction plus an angular offset (see text for details). The feedback appeared either one second after movement completion (delayed group), or immediately following it (clamped group). The following test trial (2) was different between the groups: The delayed participants were instructed to reaim (blue arrow) so to counteract the offset and bring the red cross in the middle of the target; the clamped participants could not control the end point feedback's position, thus they were told to ignore it and keep aiming straight, which induced involuntary sensorimotor adaptation (red arrow). **E:** block design of the two groups of the motor task. The clamped group (top row) began with a few baseline trials at the beginning of the session, followed by eight consecutive blocks of learning trials, four under each viewing condition (depth or no-depth, in different colors). The delayed group (bottom row) performed baseline trials before each learning block. Both groups rested for 30 s between the completion of a learning block and the start of the next one.

on oculomotor information about the orientation of the eyes when fixating the stimulus. The target was presented at either the near distance of 1 m, where oculomotor signals are reliable, or at the far distance of 7 m, where the eyes are essentially parallel and ocular vergence is completely unreliable (41). The small visual angle subtended by the target also ensured that absolute distance information provided by vertical disparities was negligible (42). In sharp contrast, the

environment under the depth condition was enriched with a wide array of visual features, which ensured a reliable disparity field and also provided plenty of additional monocular cues to depth, including perspective lines, texture gradients, surface occlusions, and objects with familiar sizes. In both conditions, the target always appeared in the center of the visual field at eye height. Note in Fig. 1A that the central portion of the visual scene up to $\sim 30^\circ$ of visual angle (the wall

behind the dartboard) showed the same uniform gray background under both conditions. The size of this area was chosen so to ensure that 1) subjects could not have an advantage in the depth condition by being able to directly compare the target's location with a neighboring scene element or visual feature; 2) the end point feedback during the motor task would always appear against a neutral background (i.e., no depth information) in both conditions (see *Learning trials* paragraph in the *Motor Task* section).

Perceptual Task

To determine how depth vision influences the perception of size, participants performed a perceptual task in which they reported the apparent size of the target (the dartboard) while being immersed in either a depth-enriched or a depth-impoverted virtual environment (Fig. 1A). Each trial started with the "observation" scene, that is, with the target visible under one of the viewing conditions (depth or no-depth). The target was seen either at 1 m (near distance) or at 7 m (far distance) and subtended one of five visual angles: 1.5°, 3°, 6°, 9°, or 12°, corresponding approximately to a diameter of 2.6, 5.2, 10.5, 15.7 and 21 cm at the near distance, and of 18.3, 36.7, 73, 110, and 148 cm at the far distance (Fig. 1B). Through a keypress, participants could toggle between this scene and a second scene ("response"). The response scene showed a textured floor with a probe laying on its surface 4 m away from the observer. The highly detailed surface of the ground in the response scene was designed to provide participants with a reliable reference when reporting their estimates of distance and size from the observation scene. The probe's distance was chosen to be intermediate between the target's near and far distances. The probe was a black strip orthogonal to the line of sight. Although in the response scene, the participants used the keyboard to adjust the length of the probe until it matched the perceived diameter of the target in the observation scene. Participants were instructed to imagine that the target was illuminated by perfectly vertical sunlight, and that the probe was to represent the resulting shadow being cast on the floor. Using the keyboard, participants could switch back and forth between observation and response scenes as many times as they needed. Following a within-subject design, the task included five presentations of each of the five target's sizes at each distance in each condition, for a total of 100 trials.

Motor Task

To determine if depth vision affects how visual errors are encoded by the motor system, participants performed a visuomotor adaptation task, in which they were instructed to make a series of short reaching movements in depth to slide one of the HTC Vive controllers on a tabletop, starting from near the chest and moving in forward direction. They aimed these movements at a target (a virtual dartboard) presented at eye height and seen in the distance, while being immersed in either a depth-enriched or a depth-impoverted virtual environment. The controller was installed on a custom 3-D-printed support whose basis was covered in felt and glided on a glass surface, to minimize friction and auditory cues. The experimenter first calibrated the start position, then the participant donned the VR headset, rested

their head on the chinrest and practiced reaching out and returning to start position until the movement became smooth and natural.

The motor task was staged as a game of "darts," in which participants were instructed to make short rapid 10 cm reaching movements with the hand, as soon as the target appeared (median movement duration = 464 ms). When the hand was 10 cm away from the start position along the depth dimension, it triggered the appearance of a 1.5° wide cross on the screen (Fig. 1C). The cross was located at the same distance of the target and at the same height of the target's center (at eye height), whereas its lateral position provided end point feedback about the movement direction (Fig. 1D). No online feedback was given. The goal of the experiment was to study the influence of pictorial depth information, such as perspective and texture, during motor learning. As binocular disparity tends to override these cues altogether, and it is most reliable near the body (see also the results from the perceptual task in Fig. 2), in the motor task the target was always seen at the far distance (7 m) subtending a visual angle of 6° (diameter of 73 cm).

Experimental groups.

Although the perceptual task was identical for all participants, in the motor task half of the participants were assigned to a "clamped" group, presumed to be largely implicit, and the other half to a "delayed" group, presumed to be mostly explicit. Both groups performed eight blocks of trials back-to-back with a short 30-s break between each block. In a within-subject design, four blocks tested reaching under the depth condition, and four under the no-depth condition. For the delayed group, each block started with a short sequence of "baseline" trials followed by a long sequence of "learning" trials. Clamped participants performed a series of baseline trials only at the beginning of the first block.

Baseline trials.

During a baseline trial, the target appeared under one of the two viewing conditions and subjects aimed at its center by reaching forward and receiving veridical end point feedback about their movement's accuracy. For the clamped group, end point feedback was displayed immediately after the hand crossed a distance of 10 cm. For the delayed group, end point feedback was presented 1 s after the hand crossed 10 cm. For both groups, the feedback then remained visible for 1 s, after which the scene was cleared, and the subject could return to the start position to begin the next trial. A V-shaped metallic frame was installed at the start position to guide participants to the same initial location when returning.

Learning trials.

Each learning trial consisted of two reaching movements. In the first reaching, the "probe," participants of both groups were instructed to aim at the center of the target as accurately as possible, but the end point feedback was offset laterally by the computer by a given amount of degrees of visual angle. There were 11 offsets: 0°, ±1.5°, ±3°, ±6°, ±9°, or ±12°. For the clamped group, the position of the end point feedback was independent of the actual hand direction. For example, with a zero degree offset the end point feedback appeared exactly in the center of the target (i.e., no visual

error), regardless of where the subject actually reached with the hand. Similarly, when the offset was -9° , the end point feedback appeared 9° to the left of the target's center (i.e., bullseye), regardless of the participant's hand direction. Subjects in this group were instructed to ignore the end point feedback and always aim at the target's center.

For the delayed group, the end point feedback's position resulted from shifting the true hand direction by an amount corresponding to the offset size. For example, a 0° offset meant that the end point feedback was veridical (the actual direction of the reaching plus zero offset), and a $+9^\circ$ offset meant that the true hand direction was shifted 9° to the right. Subjects in this group were instructed to aim straight at the target's bullseye and pay attention to the end point feedback.

In the second reaching movement of the learning trial, the "test," the end point feedback's behavior was the same as in the probe (fixed for the clamped participants and shifted for the delayed participants), but the instructions differed between groups. The clamped group was instructed to aim again at the target's bullseye and keep ignoring the end point feedback, whereas the delayed group was instructed to counteract the offset observed in the probe reach, such as to bring the (shifted) end point feedback as close as possible to hitting the target's center. Like in the baseline trials, end point feedback on learning trials appeared either instantaneously (clamped) or with a 1-s delay (delayed) and remained visible for 1 s. To ensure that the participants kept aiming at the target's center, every trial (i.e., every sequence of a probe and a test reach) the target's horizontal position was extracted from a uniform distribution within the interval $[-1.5, 1.5]^\circ$ (the computer-controlled end point feedback of the clamped group was also shifted accordingly). This manipulation also prevented that participants could exploit the same absolute target location to estimate the position of the end point feedback.

Experimental Design

Participants in both groups performed 264 learning trials. Each learning trial comprised two reaching movements (a probe trial followed by a test trial)—2 viewing conditions \times 11 offsets \times 12 repetitions (amounting to 528 reachings)—distributed across eight blocks (four blocks per visual condition).

The order of the viewing conditions (depth and no-depth) over the course of the eight blocks was counterbalanced between subjects, as well as the order of the tasks (perceptual task first vs. motor task first). The delayed group performed an additional 14 baseline trials at the beginning of each block to emphasize reaiming in the subsequent learning block. For the opposite reason, namely, to avoid canceling implicit adaptation, the clamped group did only 10 baseline trials at the beginning of the experiment (Fig. 1E).

Data Analysis

The responses of the perceptual task and the kinematic data were analyzed in R (43). For the perceptual task, the main dependent variable was the adjusted size of the probe object in the Response scene. An exploratory data analysis was first performed on a subject-by-subject basis for the

detection of possible outliers, defined as those adjustments which exceeded 1.5 times the IQR of the total distribution of the responses for that individual. These observations were excluded, amounting to 0.9% of the entire data set.

For the motor task, the primary dependent variable was the lateral position of the hand when it reached 10 cm in depth from the start position, which the program recorded in real time and used to project the end point feedback. From that value, the instantaneous hand direction was calculated as the angular deviation from the forward direction (positive values for clockwise deviations). In addition, the 2-D positional data of the controller (x and z coordinates of the reaching movement) was also examined. First it was checked for missing frames, on a subject-by-subject and trial-by-trial basis. A Savitzky-Golay filter of fourth order was then applied to the trajectory data and used to compute velocity and acceleration. The onset of the reaching movement was chosen to be the first frame of the longest sequence of frames in which the velocity's z component (the velocity in the forward direction) was continuously greater than 2 cm/s. Within the same sequence, the end of the movement was marked by the first frame in which the z velocity dropped below 2 cm/s. From the processed trajectory data, we extracted a number of dependent variables, including velocity and acceleration maxima and their times of occurrence.

RESULTS

The analysis of the results was divided into three steps: first, we examined the perceptual and motor responses separately, then we compared the performances together to study if the two tasks shared a common substrate for the processing of depth cues.

Perceptual Task

The perceptual task measured how size perception is influenced by the amount of visual depth cues present in the environment. In it, participants adjusted the length of a line to match the perceived diameter of a circular target object (a dartboard) which was seen either inside a highly detailed room (depth condition) or in an empty space (no-depth condition). Every individual did the same perceptual task regardless of the motor task group they were assigned to.

On average, participants increased the visual size of the probe in a semilinear fashion as the visual angle subtended by the target increased, as shown in Fig. 2. Participants matched the angular size of the near stimuli with relatively high accuracy (white circles), whereas they showed a bias to overestimate the visual size of the stimuli at the far distance (filled circles). Moreover, this overestimation bias for the far targets was greater under the depth condition relative to no-depth.

The angular size of the probe's adjusted length, matching the perceived angular size of the target, was submitted to an omnibus univariate $5 \times 2 \times 2$ ANOVA, crossing three factors: the target's simulated size, the viewing distance, and the viewing condition. All three main effects were significant (target size: $F_{4,156} = 558.96$, $P < 0.001$, $\eta_p^2 = 0.94$; viewing distance: $F_{1,39} = 252.17$, $P < 0.001$, $\eta_p^2 = 0.87$; visual condition: $F_{1,39} = 39.64$, $P < 0.001$, $\eta_p^2 = 0.50$), as well as all the interactions (viewing distance \times target size: $F_{4,156} = 30.32$, $P <$

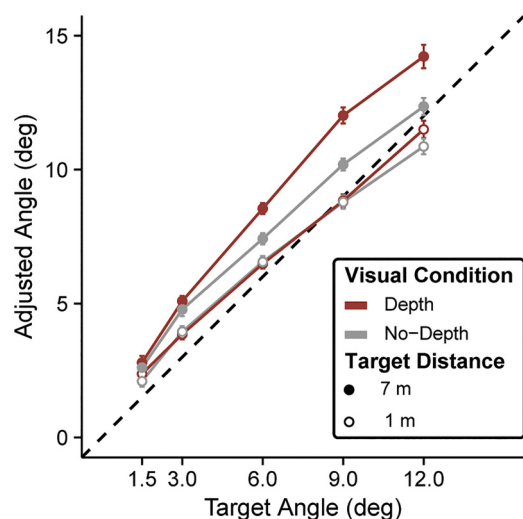


Figure 2. Adjusted visual angle of the probe in the perceptual task as function of the stimuli's visual angle. Adjustments of the far (7 m) and the near (1 m) stimuli are represented by filled circles and open circles, respectively. The dashed line represents the projection of the stimuli at the probe distance (4 m). Red data points and solid lines identify the responses under the depth condition, while the responses under the no-depth condition are displayed in gray. Error bars represent the within-subject 95% CI. CI, confidence interval.

0.001, $\eta_p^2 = 0.44$; viewing distance \times visual condition: $F_{1,39} = 56.62$, $P < 0.001$, $\eta_p^2 = 0.59$; target size \times visual condition: $F_{4,156} = 13.10$, $P < 0.001$, $\eta_p^2 = 0.25$; viewing distance \times target size \times visual condition: $F_{4,156} = 7.72$, $P < 0.001$, $\eta_p^2 = 0.17$.

The five target's angular sizes were easily distinguishable from one another, as confirmed by the Holm-corrected post hoc comparison of all pairwise differences between the adjusted sizes. Furthermore, in agreement with the psychophysics literature (44, 45), the just-noticeable-difference decreased for larger stimuli, namely, the same size interval appeared smaller between two large stimuli than between two small stimuli. This led to a gradual departure from linearity in the responses as a function of the visual angle subtended by the target (Fig. 2). The critical test here was to see whether the different depth contexts would affect this size estimation and how.

The main effect of the viewing condition showed that the size estimates were significantly greater when the target was seen under the depth condition than in isolation ($F_{1,39} = 39.64$, $P < 0.001$). Figure 2 shows that this effect was primarily driven by the interaction with the fixation distance, as it was significant at 7 m ($F_{1,39} = 71.14$, $P < 0.001$; solid dots), but not at 1 m ($F_{1,39} = 2.47$, $P = 0.12$; open circles). This was expected, since at 7 m the pictorial cues (perspective, texture, relative size, etc.) dominated size perception over binocular disparity, which instead decays rapidly with the square of the distance (46). As a result, when the dartboard was seen at the far distance in the richly detailed scene of the depth condition, it appeared compellingly larger (because seemingly more distant) than when seen at the same distance against the flat background of the no-depth condition. For this reason, 7 m was also the distance selected for the motor task.

The reverse side of the interaction between depth cues and fixation distance was that, under the depth condition,

the difference between the perceived size of the far objects and that of the near objects was greater relative to no-depth (Fig. 2, solid dots vs. open circles for each color). The abundance of depth cues in the former condition allowed participants to distinguish between the two fixation distances more clearly, resulting in more separate size estimates at near and at far distances. This pattern is in accordance with a known compression/expansion mechanism affecting the perceived visual space which depends on the amount of depth cues available in the environment (8).

In summary, size estimates were systematically greater when the target was seen in the tavern scene than when seen against a uniform background, and this effect was maximal at the farthest fixation distance as anticipated.

Motor Task

In the motor task, participants performed a visuomotor rotation task aiming at a target seen at eye height from a distance of 7 m. To focus selectively on explicit and implicit components of motor learning, they were divided between a delayed group and a clamped group, respectively. To study if and how the two processes are affected by changes in the depth information in the environment, both groups did the motor task under a depth and a no-depth visual condition. Throughout this task, all participants performed three types of trials (baseline, probe, and test). The direction of their reaching movements, which we will refer to as hand angle, was measured as the angle between the instantaneous direction of the hand when it had traveled 10 cm away from the body and the forward direction.

All participants aimed straight during the baseline trials, namely, with an average hand angle not different from zero (delayed: $-0.004^\circ \pm 0.007^\circ$, $t_{19} = -0.65$, $P = 0.52$; clamped: $0.03^\circ \pm 0.04^\circ$, $t_{19} = 0.73$, $P = 0.47$). In the learning blocks, the average hand angle of the probe trial was also zero in both groups, as expected (delayed: $0.06^\circ \pm 0.1^\circ$, $t_{19} = 0.64$, $P = 0.53$; clamped: $-0.14^\circ \pm 0.13^\circ$, $t_{19} = -1.12$, $P = 0.28$). Although both groups completed these movements in the same amount of time (delayed: 510 ± 40 ms; clamped: 460 ± 40 ms; $t_{38} = 0.83$, $P = 0.41$), clamped group exhibited shorter RTs than delayed group (clamped: 364 ± 44 ms; delayed: 642 ± 61 ms; $t_{38} = 3.69$, $P < 0.01$), as the latter group actively tried to counter the visual rotation. Finally, in the test trials, the hand angle was again on average zero (delayed: $0.18^\circ \pm 0.16^\circ$, $t_{19} = 1.13$, $P = 0.27$; clamped: $-0.2^\circ \pm 0.14^\circ$, $t_{19} = -1.4$, $P = 0.18$) as the distribution of the offsets was symmetrical around the straight direction. Moreover, as the test trials alternated with the probe trials during the learning phase of the experiment, their timing followed the same pattern (same movement time in both groups, shorter RTs for the clamped group).

The main variable we used to quantify learning throughout the whole study was the individual sensitivity to the error size, computed through a subject-by-subject linear regression of the hand angle as a function of the offset. This analysis returned an average slope of 0.96 ($t_{19} = 23.98$, $P < 0.01$) for the delayed group (mean $R^2 = 0.97$), and an average slope of 0.04 ($t_{19} = 3.88$, $P < 0.01$) for the clamped group (mean $R^2 = 0.14$) (Fig. 3). Though smaller than the effect size of the explicit learning, the significant scaling with the error size exhibited by the clamped group confirms previous

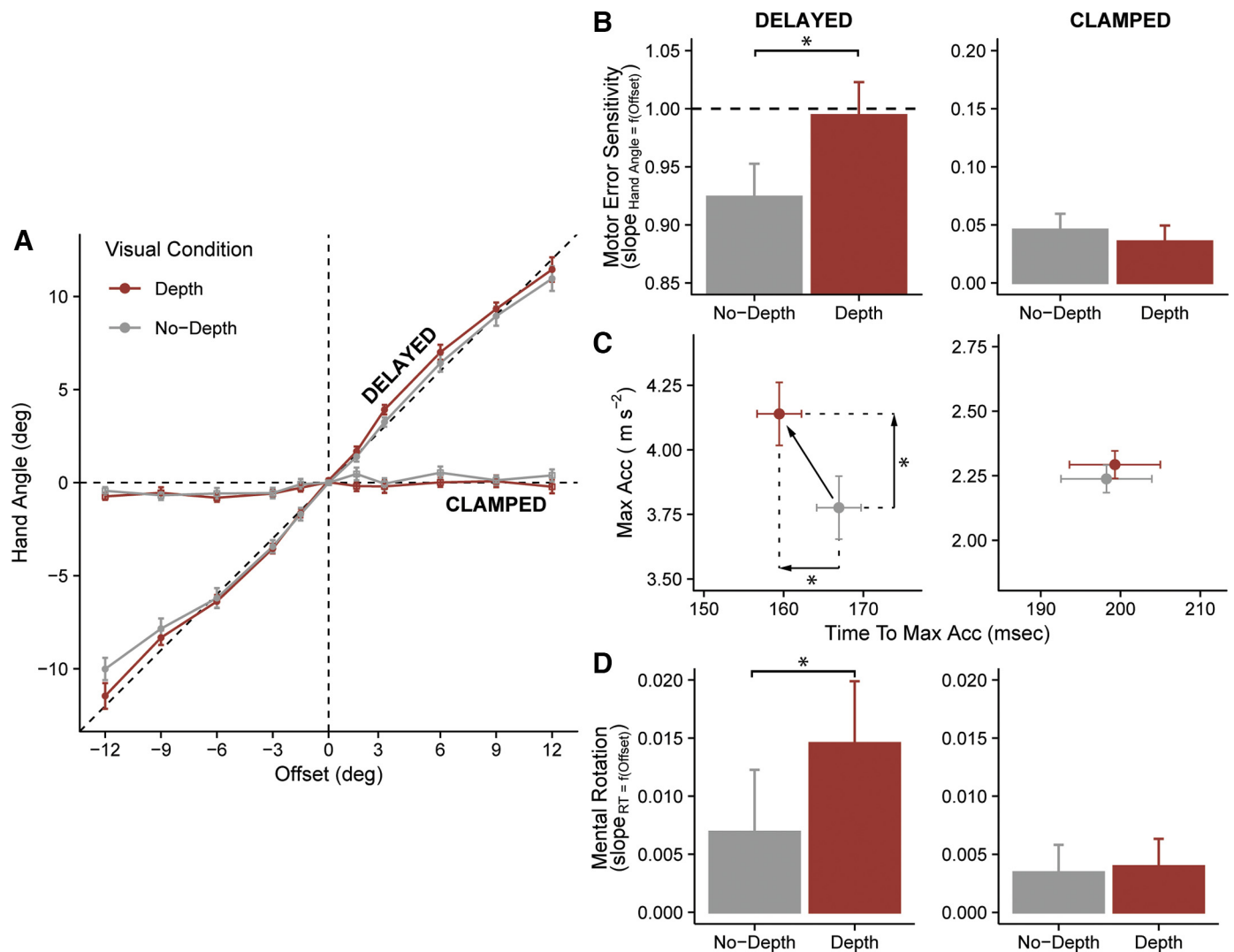


Figure 3. *A*: hand angle of the test trials as a function of the offset presented in the probe trial, for the delayed group (circles) and the clamped group (squares). Error bars represent the within-subject SE of the mean in *A*, *B*, *C* and the between-subject SE of the mean in *D*. Across all panels, the same color code is used to indicate the visual conditions (depth in red and no-depth in gray). *B*, *C*, and *D* show additional statistics about the performance of the two groups—left, delayed; right, clamped. The aspect ratio is the same across each row of panels. Significant differences ($P < 0.05$) are marked with an asterisk. *B*: increase in error sensitivity under the depth condition relative to no-depth (for delayed but not for clamped), expressed through the slope of the linear regression relating the hand angle to the offset size; *C*: anticipation (x-axis) and increase in magnitude (y-axis) of the maximum acceleration (Max Acc) under depth relative to no-depth (for delayed but not for clamped), compatible with a movement aimed at a farther target in the former condition. *D*: net modulation of the slope relating the reaction times to the (estimated) offset size, evidence of mental rotation: increase under depth relative to no-depth (for delayed but not for clamped).

evidence that the motor system can adapt to perturbations of the visuomotor mapping that vary as frequently as every other trial (40, 47–51). Moreover, the clamped group showed learning despite that certain features of the task could have potentially reduced its magnitude, such as receiving end point rather than continuous feedback (52, 53), and the fact that the experimental setup required converting the error estimate from the frontoparallel plane (that of the visual input) to the transversal plane (that of the hand).

Next, a $2 \times 6 \times 2$ mixed ANOVA was performed on the hand angle of the test trials, crossing one between-subjects factor (group) with two within-subject factors (offset size and visual condition). All three main effects were significant (group: $F_{1,38} = 444.95$, $P < 0.01$, $\eta_p^2 = 0.92$; offset size: $F_{5,190} = 408.42$, $P < 0.01$, $\eta_p^2 = 0.92$; visual conditions: $F_{1,38} =$

4.56, $P = 0.04$, $\eta_p^2 = 0.11$), as well as the interactions between group and offset size ($F_{5,190} = 360.35$, $P < 0.01$, $\eta_p^2 = 0.91$) and between group and visual condition ($F_{1,38} = 13.96$, $P < 0.01$, $\eta_p^2 = 0.27$).

The most critical finding of this task, the main effect of the visual conditions, was caused by the depth condition yielding larger compensations than the no-depth condition (Fig. 3A). This behavior mirrored the bias in the estimation of the target's size in the perceptual task: relative to the non-descript environment, participants showed greater compensations when reaching in the tavern scene.

More precisely, this main effect was driven by its interaction with the group factor, as the visual condition had a significant effect on the performance of the delayed group ($F_{1,38} = 17.24$, $P < 0.01$) although not for the clamped group

($F_{1,38} = 1.28$, $P = 0.26$). This difference between the responses of the two groups in the test trials can also be viewed in terms of error sensitivity, calculated as the slope of the linear regression between the hand angle and the offset size (Fig. 3B). For the delayed participants, switching to the tavern environment caused sensitivity to significantly increase by 0.08 ($t_{19} = 3.5$, $P < 0.01$) relative to the nondescript environment, reaching a slope value of 0.99. For comparison, no significant change could be found between conditions for the clamped group ($t_{19} = -0.94$, $P = 0.82$).

Although we cannot rule out that the small magnitude of the implicit adaptation could have prevented a depth-related effect to be detectable, the hand angle data could also hint to a genuine dissociation between explicit and implicit processes with respect to how depth information is integrated for the encoding of error signals. The analysis of two other behavioral markers provided partial evidence in favor of the latter possibility, suggesting that indeed depth vision selectively affected the movement plan of the delayed group only.

First, participants in the delayed group exhibited faster movements when reaching under the depth condition than under the no-depth condition (Fig. 3C), consistent with them aiming at a target that appeared more distant in the former case, and in agreement with the results of the perceptual task. The peak acceleration of the test reach was greater ($t_{19} = 2.1$, $P = 0.02$) and also occurred earlier in time ($t_{19} = -2.21$, $P = 0.02$). In contrast, the clamped group showed no change between conditions neither in the peak acceleration's magnitude ($t_{19} = 0.73$, $P = 0.24$) nor in its time of occurrence ($t_{19} = -0.38$, $P = 0.35$).

Second, based on the previous results we hypothesized that the reaction times would be greater under the depth condition than under no-depth, due to increased mental rotation required to compute a perceptually bigger reaiming angle (28). To test this prediction, we regressed the reaction times on the offset size (absolute value) for each subject, using the resulting slope as a measure of mental rotation. To calculate the net mental rotation in the test trial, we subtracted the slope of the probe trial's fit from that of the test trial's fit (Fig. 3D). We calculated two such values for each subject, one for each visual condition. We found a marginally significant effect of the visual condition in the delayed group ($t_{19} = 2.12$, $P = 0.047$), and no effect in the clamped group ($t_{19} = 0.32$, $P = 0.75$).

Coupling between Perception and Motor Learning

Following separate analyses, the performances in the perceptual and in the motor tasks were then directly compared to test the hypothesis that the results in both tasks stemmed from a common mechanism underlying the processing of depth cues. The rationale for this comparison was that, if such mechanism truly existed, then it should develop similarly over time in both tasks, and also be detectable at a subject-by-subject level. This analysis, which was carried on the delayed data only since the effect size of the clamped group was too small, found evidence in support of both points.

To look at the evolution of the delayed data over time, the motor results were first broken down into each learning block. As per the experimental design, the depth and no-

depth environments were each presented in four successive learning blocks across the motor task (Fig. 1E), enabling us to compare how the error size sensitivity varied on a block-by-block basis in each visual condition.

The data set for the analysis of the temporal evolution of the perceptual responses was slightly different, since the perceptual task consisted of a unique block. In here we looked at how the responses changed on a trial-by-trial basis. For each subject we split the data set into sequences of trials in which the same visual condition was presented consecutively due to randomization. We were able to obtain a full sample size for strings up to a length of four trials.

We first looked at the temporal evolution of the responses in the two tasks.

In the motor data, we found that the greater error sensitivity in the depth condition relative to the no-depth condition was caused by a deterioration of the performance under no-depth (Fig. 4A). We fitted the error size sensitivity as function of the number of blocks for each condition, and found that it was accurate [intercept = 1.02, $t_{\text{difference from 1}(19)} = 0.44$, $P = 0.66$] and constant (slope = -0.003 , $t_{19} = -0.14$, $P = 0.89$) under depth, whereas it decreased significantly under no-depth (slope = -0.04 , $t_{19} = -2.66$, $P < 0.01$).

Similar to the motor results, longer sequences of perceptual size judgments made under the same visual condition correlated with a stronger effect of depth information and vice versa. To compute this effect, for each subject we fitted two linear models of the average adjustments as a function of the number of consecutive trials, one for the depth and another for the no-depth condition. We found a significant interaction between the trials sequence's length and the visual condition, such that the no-depth adjustments became significantly smaller for longer sequences of consecutive trials ($t_{39} = -1.88$, $P = 0.03$), whereas the depth adjustments remained constant ($t_{39} = 0.39$, $P = 0.70$; Fig. 4B), attributable to size adaptation (see DISCUSSION). Though these results are only indirectly confirmatory, as they show a correlation between a block-by-block effect with a trial-by-trial effect, they nonetheless reveal the signature of the same mechanism affecting, on a different scale, both the perceptual and the motor results.

Second, we compared the motor results with the perceptual results at a subject-by-subject level. This comparison

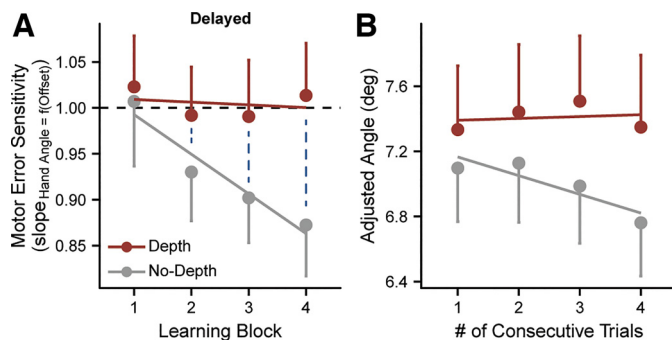


Figure 4. A: motor task: block-by-block evolution of the motor error sensitivity for each visual condition (delayed group). B: perceptual task: average estimate of the target size as function of the number of trials consecutively under the same visual conditions. All error bars show the within-subject means \pm SE.

was driven by the observation of a systematic bias in the motor results. As the top row of Fig. 5A shows, the error size sensitivity under the no-depth condition showed a maximum, nearly significant drop in the second block ($t_{19} = -1.47$, $P = 0.078$), which is evidence of a sharp underestimation of the target's distance relative to the first block. This decrease then abated in the remaining third and fourth blocks (gray data). For comparison, under the depth condition (red data), the transition from the first block to the second block was accompanied by a much smaller drop in error sensitivity ($P = 0.27$).

Although at first blush this too could be viewed as completely spurious, a careful observation of the movement kinematics revealed the presence of the exact same pattern (Fig. 5A, center and bottom rows). Under the no-depth condition, both the hand's maximum velocity and maximum acceleration underwent a dramatic reduction after the first transition from the first block to the second block (max velocity: $t_{19} = -3.33$, $P < 0.01$; max acceleration: $t_{19} = -2.60$, $P < 0.01$), consistent with the reduction of the target's estimated distance suggested by the analysis of the error sensitivity. Moreover, this means that the error was lawfully underestimated in turn, since the offset size scaled with the viewing distance. In the remaining blocks of the motor task, the kinematic profile of the hand then remained relatively stable, again mirroring the error sensitivity: both maxima showed

no significant reduction under the depth condition (max velocity: $P = 0.20$; max acceleration: $P = 0.20$).

Note in the graphs in Fig. 5A that, although there was a general reduction of all behavioral indices in the second block (all trended downward), likely due to habituation to the visual scene, this very reduction was always systematically greater in the no-depth condition than the depth condition.

All these observations suggested that the first shift between blocks under different visual conditions was the most critical, in terms of when the depth cues had their most visible impact on error sensitivity. To verify this conclusion, we repeated the previous correlation analysis between the motor and the perceptual shift in error sensitivity (Δ slope) only this time we subset each learning block and compared it with the same individuals' perceptual results. This resulted in four correlation statistics (Pearson's r), comparing the same Δ slope_{perceptual} with the Δ slope_{motor} from the first, second, third, and fourth motor learning blocks. Consistent with the behavioral observations, we found that the perceptual data showed a selective and strong positive correlation with the motor data of the second block ($r_{\text{block2}} = 0.59$, $P = 0.02$)—that is, after the first shift between visual conditions but not after the following shifts ($P_{\text{block1}} = 0.25$, $P_{\text{block3}} = 1$; $P_{\text{block4}} = 1$; all P values were Bonferroni-corrected).

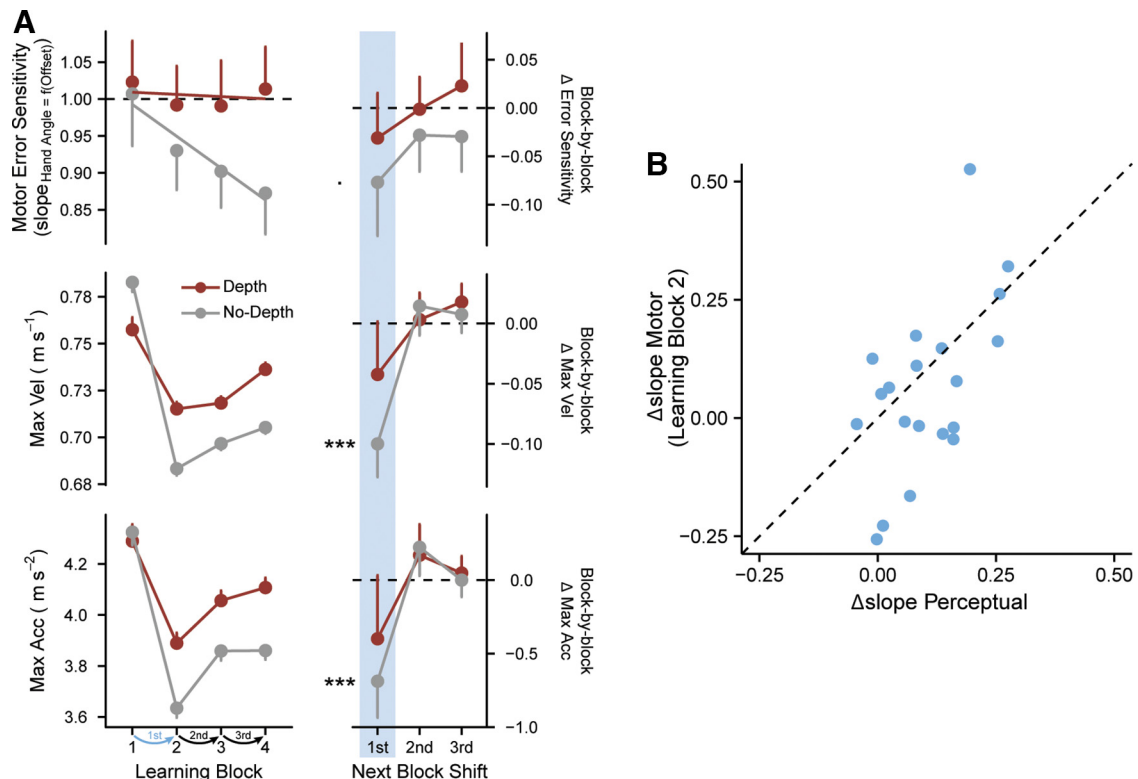


Figure 5. A: group-level effects of depth information on motor learning (delayed participants). *Left*, evolution of three behavioral indexes [*top*, motor error sensitivity, same as in Fig. 4A; *middle*, maximum velocity (Max Vel); *bottom*, maximum acceleration (Max Acc)] over the course of the learning blocks of each visual condition (depth in red, no-depth in gray). *Right*, block-by-block difference of each variable (change relative to the previous block): most of the motor learning was determined after shifting to the second block. Significant differences from zero are marked with \cdot ($P < 0.1$) and *** ($P < 0.01$). Error bars show the within-subject means \pm SE. B: individual-level effect of depth information on the perception-action coupling. Comparison between the effect of the visual conditions on the error sensitivity (Δ slope) in the motor task (y -axis) and in the perceptual task (x -axis) during learning block #2. Each point represents an individual score.

This analysis outlined the importance of the first transition from one condition block to the other, which appeared to mark a crucial moment during explicit motor learning. In the second learning block, each delayed participant recalibrated their reaching movements according to their own perceptual rescaling of the scene. This is visible in the scatterplot of Fig. 5B if, for example, a participant estimated the target's size in approximately the same way under either visual condition, showing a weak effect of the depth information on perception—that is, a low $\Delta\text{slope}_{\text{perceptual}}$ —they also showed a weak effect of the depth information on action—a low $\Delta\text{slope}_{\text{motor}}$. Vice versa, a participant who showed a clear improvement in the estimation of the target's size under the depth condition relative to no-depth—resulting in a high $\Delta\text{slope}_{\text{perceptual}}$ —also scaled their reaches more accurately under the same visual condition (high $\Delta\text{slope}_{\text{motor}}$).

Overall, the comparison between the perceptual and the motor data uncovered compelling parallelisms between the two tasks. Although these similarities must be taken with caution due to the post-hoc nature of the analysis, they are remarkably consistent with the results of the contrasts planned a priori. In conclusion, although more future exploration on this issue is necessary, the present data strongly supports a direct connection between the perceptual analysis of the visual scene and the behavior exposed by explicit motor learning.

DISCUSSION

Behavioral and neuroanatomical research on motor adaptation have uncovered evidence of both explicit and implicit error-minimization mechanisms, which drive the human ability to adapt to changes in the environment. Both processes are known to react to error signals differently from one another (18, 28, 53, 54). Regardless of these behavioral differences, any incoming source of error information requires translating a proximal sensory stimulation into a representation that is useful to update the motor output.

Theoretically, the translation step poses a problem, because it happens in the perceptual domain and perception does not always represent reality in a veridical fashion (7, 55–57). In practice, the effects of potentially inaccurate perception-action coupling have not been a concern for most of the previous investigations, which have probed mostly planar movements with feedback seen on a flat surface or a frontoparallel setup. However, in the natural environment motor actions are embedded in a 3-D space, and they frequently consist of movements aimed at multiple distances where even small perceptual changes can lead to major mistakes. Here, we sought to investigate the role of depth information in motor learning using virtual reality to render a 3-D version of a classic visuomotor rotation task in depth.

Participants performed outward reaching movements aiming at the center of a target located at a simulated distance of 7 m and seen on the frontoparallel plane at eye height. They did so under two visual conditions: A depth condition, where the environment was enriched with retinal and extraretinal depth cues, and a no-depth condition, where the environment offered no reliable distance cues except for oculomotor convergence. We crossed these visual conditions with two end point feedback conditions, delayed

feedback and “clamped” feedback, which have been shown to emphasize learning by explicit and implicit processes, respectively (16, 20, 23, 38, 39, 58–60).

The learning task consisted of two reaching movements in sequence, a probe trial and a test trial. In the first trial (probe) participants aimed straight at the target's bullseye and witnessed the perturbation, according to their feedback condition. In the second trial (test), the delayed group was allowed to reaim to compensate for the offset, whereas the clamped group was instructed to ignore the feedback and keep aiming straight. Finally, to assess the effect of the visual conditions on size estimation, all participants did a perceptual task where they judged the size of the target under the depth and the no-depth conditions.

Both groups of participants exhibited sensitivity to the error size in the test trial, although to different extents. Delay participants made ample and mostly accurate corrections, whereas the clamped participants exhibited a small subconscious compensatory bias proportional to the size of the visual offset. The different error sensitivity between groups was expected, since explicit re-aiming and implicit adaptation have been previously shown to lead to separate forms of learning, each with specific characteristics (14, 16, 17, 24). Albeit small, the significant modulation exhibited by the clamped group confirms previous results that implicit learning can develop after as little as two trials (40, 47–51), and also converges with recent evidence showing serial dependencies between saccade adaptation and spatial localization (61).

More importantly, delayed participants showed a systematic modulation of their error sensitivity driven by the depth condition, such that their re-aiming direction was smaller under no-depth relative to depth and correlated with a reduction in the movement's speed and acceleration. Albeit being relatively small, the decrease in error sensitivity was present for both rightward and leftward errors of the cursor, thus somewhat acting as an internal replication, and we expect that the same increase would be more pronounced with stronger depth cues and at larger distances. This effect of the depth cues suggests that, during explicit motor learning, the error size (i.e., the amount of lateral shift of the end point from the target's center) was estimated as a lateral offset. That is, by scaling its retinal projection by some estimate of the fixation distance to return a length estimate (as opposed to an angular estimate, which is independent of the fixation distance). This would explain why the hand direction was susceptible to changes in the context: as per the perceptual results, perceived size was smaller in the no-depth environment relative to depth, yielding less reaiming, because the overwhelming lack of depth cues caused the flattening of the perceptual space (62) and made the objects in front of the observer look closer and thus smaller. The opposite happened when the scene was abundant with depth cues.

In contrast to the delay participants, the error sensitivity function of the clamped group, as well as their movement kinematics, appeared to be independent of the visual conditions. This selective effect of depth vision on the explicit but not on the implicit learning is a finding that bears further investigation. There are reasons to believe that this null effect may just be an absence of evidence rather than

evidence of absence. The effects of the depth cues might have been too small to be detected, given how flat the error sensitivity function was for the clamped group. Or maybe the depth cues were concentrated too much in the periphery of the visual field relative to the area where the clamped feedback appeared. At this stage, there are merits to both accounts. On the one hand, a common mechanism of depth analysis shared by both explicit and implicit learnings would be in principle a parsimonious account. On the other hand, there is potential support in the literature to the hypothesis that explicit and implicit processing of depth information could feed from different encodings of the depth signals.

Explicit-Egocentric versus Implicit-Retinitopic Encoding of 3-D Error Information

A possible explanation as to why implicit learning appeared to resist the effect of the depth cues is that adaptation may rely on a different, lower level form of spatial processing, such as, for example, expressing the error size retinotopically as the angular distance between the projections of the end point and of the target's center. In principle, such an encoding mechanism is effectively more stable and resistant to contextual changes, because it constructs an error signal based on local features taken out of their context. The downside of its specificity is a lack of flexibility, for example if changes in the context suddenly render those same features as irrelevant for the action's goal.

This description of how implicit learning may function in 3-D space fits well with the behavioral results of the previous studies conducted in 2-D, which characterize sensorimotor adaptation as stereotyped, slow and largely independent of the context or task demands (14, 16). Notably, it also fits well with recent neuroimaging studies showing that the cerebellum contains multiple topographic maps of the visual field that are retinotopically organized (63–65). These findings could suggest that the computation of an error signal is somehow expressed in retinal coordinates. Another similarity with the existing literature is that implicit learning appears to be more insensitive compared to explicit learning (16, 18, 19, 23). Indeed, a retinotopic mechanism for localizing error information is computationally more parsimonious than scaling retinal projections in egocentric coordinates.

An egocentric-versus-retinotopic dichotomy could be the hint to a possible specialization of the motor learning components with respect to depth information, where implicit adaptation contributes maximally within the near space whereas explicit learning handles errors in the far space. Retinal vision is especially important in the peripersonal space, where things are usually manipulated. Within reachable distance, retinal disparity is by far the most reliable of the depth cues (66) and most common manipulatory actions can be thought of as alignment tasks. However, since disparity decays rapidly with the square of the distance, monocular pictorial cues constitute the predominant signals for the interpretation of the 3-D space beyond ~ 2 m (46). The interaction between distance and visual condition found in this perceptual task is a classic signature of this competition between cues. When the stimulus was at the near distance of 1 m, binocular disparity was strong enough to allow observers to detect that the target did not change in size (nor in

egocentric position) between the depth and the no-depth scenes. Conversely, binocular disparity was negligible when the target was at the far distance of 7 m, thus size (and position) estimation was entirely driven by the richness in the pictorial cues.

It is possible that sensorimotor adaptation would selectively respond to errors in the peripersonal space because small corrections that are meaningful near the body correspond to enormous overcompensations at large distances, which would require a counter-intervention by the explicit system. At far distances, the competition between the two types of learning would trigger a positive feedback loop that would not converge towards a stable state. Moreover, a near/far division of labor is consistent with previous work showing that implicit learning is most efficient with small errors that are within the range of the motor noise, whereas it saturates with larger perturbations that are instead handled by the explicit learning (67).

If explicit and implicit learning were equally competitive at large distances, their interaction would be very problematic for the motor system, since even small errors would result in great instability. Furthermore, the fact that the movement kinematics of the clamped group did not reflect changes in the apparent distance of the target whatsoever (as opposed to the delayed group) is another evidence in favor of a true null effect of depth. Considering the many analogies between the data collected here and the previous literature, the present study hints that implicit learning might be more concerned with the direction of a perturbation, while explicit learning with its magnitude (16–18; but see Ref. 23). Recent data showed that implicit learning could decrease its contribution over time, akin to an inhibition mechanism, when the motor system faces task demands that make sensorimotor adaptation counterproductive to the performance (68, 69).

Perceptual Biases in VR

The perceptual results suggest that judgments of error size were affected by multiple sources of bias distorting the estimation of the egocentric distance of the stimuli. First, stereoscopic devices such as VR head-mounted displays (HMDs) are known for causing a perceptual compression of the simulated environment along the depth dimension, due to the mismatch between the accommodative distance of the lenses and the fixation distance produced by the rotation of the eyes, known as vergence-accommodation conflict (70–75). In our perceptual task specifically, this may have been in part responsible for the asymmetry between the error estimations at the near distance of 1 m versus at 7 m. Interestingly though, the perceptual data were decisively pulled toward the specific distance of 4 m (Fig. 2, identity line), precisely the probe's distance. Regardless of the potential effect of the vergence-accommodation conflict, any visual compression caused by it would be common to both the depth and no-depth conditions and, as such, it cannot account for any relative differences between them.

The near-far asymmetry found in the perceptual task suggests a second source of bias, a phenomenon previously termed “visuospatial compression” by which the perceptual space tends to compress along the depth dimension relative to the physical space, such that near and far objects appear

compressed towards a common intermediate distance (8). Perceived size, which is lawfully related to perceived distance, warps accordingly: For example, a small coin near the body and the moon in the sky are judged as more similar in size if viewed such that their images subtend the same area on the retina and are surrounded by uninformative background, because their retinal projections tend to be scaled by a common distance (76, 77). Compared with the veridical simulated values (in cm), participants matched a larger size to the near target and a smaller size to the far target, consistent with the two stimuli appearing closer in depth and thus more similar in diameter. Figure 2 shows the signature of this effect: the near and far estimates are drawn towards the dashed line representing the target's projection at the probe's distance of 4 m. This perceptual bias was promoted by three experimental conditions: first, the adjustments were made at the same intermediate distance, which caused near and far images to be rescaled by a similar factor; second, the near and the far targets subtended the same exact retinal size to begin with, flattening their two estimates toward a common value; third, and most critical for the study, the overwhelming lack of depth information in the surroundings of the target under the no-depth condition prevented the visual system from contrasting the perceptual compression of space.

The depth condition was specifically designed to measure the degree to which retinal depth cues allowed observers to disambiguate how far and how big the target actually was. Under this condition, in addition to ocular vergence, participants could rely on plenty of pictorial cues to assess the target's egocentric distance. As expected, participants adjusted the near and the far targets as if they appeared more dissimilar in size relative to no-depth ($F_{1,39} = 233.65$, $P < 0.001$), compatible with them also appearing more distant from one another.

Conclusions

As the interaction between the two learning systems is multifaceted and complex, this investigation has just scratched the surface of the problem. Future work will need to address how depth cues are incorporated in detail and how explicit and implicit systems work in cooperation or in competition with each other.

ACKNOWLEDGMENTS

We thank Cameron Hayes and Chandra Greenberg for help with data collection.

GRANTS

This work was supported by the National Science Foundation under Grant No. 1827550.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

C.C., F.D., and J.A.T. conceived and designed research; C.C. performed experiments; C.C., F.D., and J.A.T. analyzed data; C.C.,

F.D., and J.A.T. interpreted results of experiments; C.C. prepared figures; C.C. drafted manuscript; C.C., F.D., and J.A.T. edited and revised manuscript; C.C., F.D., and J.A.T. approved final version of manuscript.

REFERENCES

1. Thoroughman KA, Shadmehr R. Learning of action through adaptive combination of motor primitives. *Nature* 407: 742–747, 2000. doi:10.1038/35037588.
2. Körding KP, Wolpert DM. Bayesian integration in sensorimotor learning. *Nature* 427: 244–247, 2004. doi:10.1038/nature02169.
3. Cheng S, Sabes PN. Modeling sensorimotor learning with linear dynamical systems. *Neural Comput* 18: 760–793, 2006. doi:10.1162/089976606775774651.
4. He K, Liang Y, Abdollahi F, Fisher Bittmann M, Körding K, Wei K. The statistical determinants of the speed of motor learning. *PLoS Comput Biol* 12: e1005023, 2016. doi:10.1371/journal.pcbi.1005023.
5. Glennerster A, Rogers BJ, Bradshaw MF. Stereoscopic depth constancy depends on the subject's task. *Vision Res* 36: 3441–3456, 1996. doi:10.1016/0042-6989(96)00090-9.
6. Bradshaw MF, Parton AD, Glennerster A. The task-dependent use of binocular disparity and motion parallax information. *Vision Res* 40: 3725–3734, 2000. doi:10.1016/S0042-6989(00)00214-5.
7. Johnston EB. Systematic distortions of shape from stereopsis. *Vision Res* 31: 1351–1360, 1991. doi:10.1016/0042-6989(91)90056-B.
8. Campagnoli C, Croom S, Domini F. Stereovision for action reflects our perceptual experience of distance and depth. *J Vis* 17: 21, 2017. doi:10.1167/17.9.21.
9. Campagnoli C, Domini F. Does depth-cue combination yield identical biases in perception and grasping? *J Exp Psychol Hum Percept Perform* 45: 659–680, 2019. doi:10.1037/xhp0000636.
10. Volcic R, Fantoni C, Caudek C, Assad JA, Domini F. Visuomotor adaptation changes stereoscopic depth perception and tactile discrimination. *J Neurosci* 33: 17081–17088, 2013. doi:10.1523/JNEUROSCI.2936-13.2013.
11. Taylor JA, Ivry RB. The role of strategies in motor learning. *Ann N Y Acad Sci* 1251: 1–12, 2012. doi:10.1111/j.1749-6632.2011.06430.x.
12. Huberdeau DM, Krakauer JW, Haith AM. Dual-process decomposition in human sensorimotor adaptation. *Curr Opin Neurobiol* 33: 71–77, 2015. doi:10.1016/j.conb.2015.03.003.
13. McDougall SD, Ivry RB, Taylor JA. Taking aim at the cognitive side of learning in sensorimotor adaptation tasks. *Trends Cogn Sci* 20: 535–544, 2016. doi:10.1016/j.tics.2016.05.002.
14. Taylor JA, Ivry RB. Flexible cognitive strategies during motor learning. *PLoS Comput Biol* 7: e1001096, 2011. doi:10.1371/journal.pcbi.1001096.
15. Leow L-A, Marinovic W, Riek S, Carroll TJ. Cerebellar anodal tDCS increases implicit visuomotor remapping when strategic re-aiming is suppressed (Preprint). *bioRxiv* 091397, 2016. <https://doi.org/10.1101/091397>.
16. Morehead JR, Taylor JA, Parvin DE, Ivry RB. Characteristics of implicit sensorimotor adaptation revealed by task-irrelevant clamped feedback. *J Cogn Neurosci* 29: 1061–1074, 2017. doi:10.1162/jocn_a_01108.
17. Butcher PA, Taylor JA. Decomposition of a sensory prediction error signal for visuomotor adaptation. *J Exp Psychol Hum Percept Perform* 44: 176–194, 2018. doi:10.1037/xhp0000440.
18. Bond KM, Taylor JA. Flexible explicit but rigid implicit learning in a visuomotor adaptation task. *J Neurophysiol* 113: 3836–3849, 2015. doi:10.1152/jn.00009.2015.
19. Bond KM, Taylor JA. Structural learning in a visuomotor adaptation task is explicitly accessible. *eNeuro* 4: ENEURO.0122-17.2017, 2017. doi:10.1523/ENEURO.0122-17.2017.
20. Poh E, Taylor JA. Generalization via superposition: combined effects of mixed reference frame representations for explicit and implicit learning in a visuomotor adaptation task. *J Neurophysiol* 121: 1953–1966, 2019. doi:10.1152/jn.00624.2018.
21. Schween R, Taylor JA, Hegele M. Plan-based generalization shapes local implicit adaptation to opposing visuomotor transformations. *J Neurophysiol* 120: 2775–2787, 2018. doi:10.1152/jn.00451.2018.

22. **Schween R, Langsdorf L, Taylor JA, Hegele M.** How different effectors and action effects modulate the formation of separate motor memories. *Sci Rep* 9: 17040, 2019. doi:10.1038/s41598-019-53543-1.
23. **Kim HE, Morehead JR, Parvin DE, Moazzezi R, Ivry RB.** Invariant errors reveal limitations in motor correction rather than constraints on error sensitivity. *Commun Biol* 1: 19, 2018. doi:10.1038/s42003-018-0021-y.
24. **Mazzoni P, Krakauer JW.** An implicit plan overrides an explicit strategy during visuomotor adaptation. *J Neurosci* 26: 3642–3645, 2006. doi:10.1523/JNEUROSCI.5317-05.2006.
25. **Slachevsky A, Pillon B, Fournier P, Pradat-Diehl P, Jeannerod M, Dubois B.** Preserved adjustment but impaired awareness in a sensory-motor conflict following prefrontal lesions. *J Cogn Neurosci* 13: 332–340, 2001. doi:10.1162/08999290151137386.
26. **Slachevsky A, Pillon B, Fournier P, Renié L, Levy R, Jeannerod M, Dubois B.** The prefrontal cortex and conscious monitoring of action: an experimental study. *Neuropsychologia* 41: 655–665, 2003. doi:10.1016/s0028-3932(02)00225-7.
27. **Taylor JA, Ivry RB.** Cerebellar and prefrontal cortex contributions to adaptation, strategies, and reinforcement learning. *Prog Brain Res* 210: 217–253, 2014. doi:10.1016/B978-0-444-63356-9.00009-1.
28. **McDougle SD, Taylor JA.** Dissociable cognitive strategies for sensorimotor learning. *Nat Commun* 10: 40, 2019. doi:10.1038/s41467-018-07941-0.
29. **de Brouwer AJ, Albaghdadi M, Flanagan JR, Gallivan JP.** Using gaze behavior to parcellate the explicit and implicit contributions to visuomotor learning. *J Neurophysiol* 120: 1602–1615, 2018. doi:10.1152/jn.00113.2018.
30. **Anguera JA, Reuter-Lorenz PA, Willingham DT, Seidler RD.** Contributions of spatial working memory to visuomotor learning. *J Cogn Neurosci* 22: 1917–1930, 2010. doi:10.1162/jocn.2009.21351.
31. **Christou AI, Miall RC, McNab F, Galea JM.** Individual differences in explicit and implicit visuomotor learning and working memory capacity. *Sci Rep* 6: 1–13, 2016. doi:10.1038/srep36633.
32. **Holland P, Codol O, Oxley E, Taylor M, Hamshire E, Joseph S, Huffer L, Galea JM.** Domain-specific working memory, but not dopamine-related genetic variability, shapes reward-based motor learning. *J Neurosci* 39: 9383–9396, 2019. doi:10.1523/JNEUROSCI.0583-19.2019.
33. **Miall RC, Wolpert DM.** Forward models for physiological motor control. *Neural Netw* 9: 1265–1279, 1996. doi:10.1016/s0893-6080(96)00035-4.
34. **Martin TA, Keating JG, Goodkin HP, Bastian AJ, Thach WT.** Throwing while looking through prisms: I. Focal olivocerebellar lesions impair adaptation. *Brain* 119: 1183–1198, 1996. doi:10.1093/brain/119.4.1183.
35. **Tseng YW, Diedrichsen J, Krakauer JW, Shadmehr R, Bastian AJ.** Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *J Neurophysiol* 98: 54–62, 2007. doi:10.1152/jn.00266.2007.
36. **Schlerf JE, Xu J, Klempfuss NM, Griffiths TL, Ivry RB.** Individuals with cerebellar degeneration show similar adaptation deficits with large and small visuomotor errors. *J Neurophysiol* 109: 1164–1173, 2013. doi:10.1152/jn.00654.2011.
37. **Butcher PA, Ivry RB, Kuo S-H, Rydz D, Krakauer JW, Taylor JA.** The cerebellum does more than sensory prediction error-based learning in sensorimotor adaptation tasks. *J Neurophysiol* 118: 1622–1636, 2017. doi:10.1152/jn.00451.2017.
38. **Brudner SN, Kethidi N, Graepner D, Ivry RB, Taylor JA.** Delayed feedback during sensorimotor learning selectively disrupts adaptation but not strategy use. *J Neurophysiol* 115: 1499–1511, 2016. doi:10.1152/jn.00066.2015.
39. **Schween R, Hegele M.** Feedback delay attenuates implicit but facilitates explicit adjustments to a visuomotor rotation. *Neurobiol Learn Mem* 140: 124–133, 2017. doi:10.1016/j.nlm.2017.02.015.
40. **Hutter SA, Taylor JA.** Relative sensitivity of explicit reaiming and implicit motor adaptation. *J Neurophysiol* 120: 2640–2648, 2018. doi:10.1152/jn.00283.2018.
41. **Howard IP.** *Perceiving in Depth, Volume 1: Basic Mechanisms* (1st ed.). Oxford University Press, 2012, p. 672.
42. **Gillam B, Chambers D, Lawergren B.** The role of vertical disparity in the scaling of stereoscopic depth perception: an empirical and theoretical study. *Percept Psychophys* 44: 473–483, 1988. doi:10.3758/bf03210433.
43. **R Core Team.** R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2020. <https://www.R-project.org/>.
44. **Oyama T.** Perceived size and perceived distance in stereoscopic vision and an analysis of their causal relations. *Percept Psychophys* 16: 175–181, 1974. doi:10.3758/BF03203271.
45. **Ross H, Plug C.** *The Mystery of The Moon Illusion: Exploring Size Perception*. Oxford, UK: Oxford University Press, 2002, p. 288.
46. **Howard IP, Rogers BJ.** *Seeing in Depth, Volume 2: Depth Perception*. Toronto, ON, Canada: University of Toronto Press, 2002.
47. **Fine MS, Thoroughman KA.** Motor adaptation to single force pulses: sensitive to direction but insensitive to within-movement pulse placement and magnitude. *J Neurophysiol* 96: 710–720, 2006. doi:10.1152/jn.00215.2006.
48. **Fine MS, Thoroughman KA.** Trial-by-trial transformation of error into sensorimotor adaptation changes with environmental dynamics. *J Neurophysiol* 98: 1392–1404, 2007. doi:10.1152/jn.00196.2007.
49. **Wei K, Körding K.** Relevance of error: what drives motor adaptation? *J Neurophysiol* 101: 655–664, 2009. doi:10.1152/jn.90545.2008.
50. **Marko MK, Haith AM, Harran MD, Shadmehr R.** Sensitivity to prediction error in reach adaptation. *J Neurophysiol* 108: 1752–1763, 2012. doi:10.1152/jn.00177.2012.
51. **Semrau JA, Daitch AL, Thoroughman KA.** Environmental experience within and across testing days determines the strength of human visuomotor adaptation. *Exp Brain Res* 216: 409–418, 2012. doi:10.1007/s00221-011-2945-z.
52. **Hinder MR, Tresilian JR, Riek S, Carson RG.** The contribution of visual feedback to visuomotor adaptation: how much and when? *Brain Res* 1197: 123–134, 2008. doi:10.1016/j.brainres.2007.12.067.
53. **Taylor JA, Krakauer JW, Ivry RB.** Explicit and implicit contributions to learning in a sensorimotor adaptation task. *J Neurosci* 34: 3023–3032, 2014. doi:10.1523/JNEUROSCI.3619-13.2014.
54. **McDougle SD, Bond KM, Taylor JA.** Explicit and implicit processes constitute the fast and slow processes of sensorimotor learning. *J Neurosci* 35: 9568–9579, 2015. doi:10.1523/JNEUROSCI.3619-13.2014.
55. **Todd JT, Tittle JS, Norman JF.** Distortions of three-dimensional space in the perceptual analysis of motion and stereo. *Perception* 24: 75–86, 1995. doi:10.1068/p240075.
56. **Bingham GP, Zaal F, Robin D, Shull JA.** Distortions in definite distance and shape perception as measured by reaching without and with haptic feedback. *J Exp Psychol Hum Percept Perform* 26: 1436–1460, 2000. doi:10.1037/0096-1523.26.4.1436.
57. **Todd JT, Norman JF.** The visual perception of 3-D shape from multiple cues: Are observers capable of perceiving metric structure? *Percept Psychophys* 65: 31–47, 2003. doi:10.3758/bf03194781.
58. **Kitazawa S, Kohn T, Uka T.** Effects of delayed visual information on the rate and amount of prism adaptation in the human. *J Neurosci* 15: 7644–7652, 1995. doi:10.1523/JNEUROSCI.15-11-07644.1995.
59. **Honda T, Hirashima M, Nozaki D.** Adaptation to visual feedback delay influences visuomotor learning. *PLoS One* 7: e37900, 2012. doi:10.1371/journal.pone.0037900.
60. **Kim HE, Parvin DE, Ivry RB.** The influence of task outcome on implicit motor learning. *eLife* 8: e39882, 2019. doi:10.7554/eLife.39882.
61. **Cont C, Zimmermann E.** The motor representation of sensory experience. *Curr Biol* 31: 1029–1036.e2, 2021. doi:10.1016/j.cub.2020.11.032.
62. **Foley JM.** Binocular distance perception. *Psychol Rev* 87: 411–434, 1980. doi:10.1037/0033-295X.87.5.411.
63. **Buckner RL, Krienen FM, Castellanos A, Diaz JC, Yeo BT.** The organization of the human cerebellum estimated by intrinsic functional connectivity. *J Neurophysiol* 106: 2322–2345, 2011. doi:10.1152/jn.00339.2011.
64. **van Es DM, van der Zwaag W, Knapen T.** Topographic maps of visual space in the human cerebellum. *Curr Biol* 29: 1689–1694.e3, 2019. doi:10.1016/j.cub.2019.04.012.
65. **Winawer J, Curtis C.** Cerebellar function: multiple topographic maps of visual space. *Curr Biol* 29: R699–R702, 2019. doi:10.1016/j.cub.2019.06.005.
66. **Howard IP, Rogers BJ.** *Binocular Vision and Stereopsis*. New York: Oxford University Press, 1995.
67. **Miyamoto YR, Wang S, Smith MA.** Implicit adaptation compensates for erratic explicit strategy in human motor learning. *Nat Neurosci* 23: 443–455, 2020. doi:10.1038/s41593-020-0600-3.

68. **Avraham G, Morehead JR, Kim HE, Ivry RB.** Re-exposure to a sensorimotor perturbation produces opposite effects on explicit and implicit learning processes (Preprint). *bioRxiv*, 2021. <https://doi.org/10.1101/2020.07.16.205609>.
69. **Albert ST, Jang J, Sheahan HR, Teunissen L, Vandevoorde K, Herzfeld DJ, Shadmehr R.** An implicit memory of errors limits human sensorimotor adaptation. *Nat Hum Behav*: 1–15, 2021. doi:10.1038/s41562-020-01036-x.
70. **Wann JP, Rushton S, Mon-Williams M.** Natural problems for stereoscopic depth perception in virtual environments. *Vision Res* 35: 2731–2736, 1995. doi:10.1016/0042-6989(95)00018-U.
71. **Bingham GP, Bradley A, Bailey M, Vinner R.** Accommodation, occlusion, and disparity matching are used to guide reaching: a comparison of actual versus virtual environments. *J Exp Psychol Hum Percept Perform* 27: 1314–1334, 2001. doi:10.1037//0096-1523.27.6.1314.
72. **Hoffman DM, Girshick AR, Akeley K, Banks MS.** Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *J Vis* 8: 33.1–33.30, 2008. doi:10.1167/8.3.33.
73. **Domini F, Shah R, Caudek C.** Do we perceive a flattened world on the monitor screen? *Acta Psychol (Amst)* 138: 359–366, 2011. doi:10.1016/j.actpsy.2011.07.007.
74. **Scarfe P, Glennerster A.** The science behind virtual reality displays. *Annu Rev Vis Sci* 5: 529–547, 2019. doi:10.1146/annurev-vision-091718-014942.
75. **Campagnoli C, Hung B, Domini F.** Explicit and implicit depth-cue integration: evidence of systematic biases with real objects (Preprint). *bioRxiv*, 2021. <https://doi.org/10.1101/2021.03.19.436171>.
76. **Gillinsky AS.** Perceived size and distance in visual space. *Psychol Rev* 58: 460–482, 1951. doi:10.1037/h0061505.
77. **Gogel WC.** The sensing of retinal size. *Vision Res* 9: 1079–1094, 1969. doi:10.1016/0042-6989(69)90049-2.