Published in final edited form as:

Spat Cogn Comput. 2008 October ; 8(4): 333–356. doi:10.1080/13875860802286668.

Learning to Reach to Locations Encoded from Imaging Displays

Bing Wu^{a,b,*}, Roberta L. Klatzky^{a,c}, and George Stetten^{b,d}

^aDepartment of Psychology, Carnegie Mellon University, Pittsburgh, USA

^bRobotics Institute, Carnegie Mellon University, Pittsburgh, USA

^cHuman-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, USA

^dDepartment of Bioengineering, University of Pittsburgh, Pittsburgh, USA

Abstract

The present study investigated how people learn to correct errors in actions directed toward cognitively encoded spatial locations. Subjects inserted a stylus to reach a hidden target localized by means of ultrasound imaging and portrayed with a scaled graph. As was found previously (Wu et al., 2005), subjects initially underestimated the target location but corrected their responses when given training with feedback. Three experiments were conducted to examine whether the error correction occurred at (1) the mapping from the input to a mental representation of target location; (2) the mapping from the representation of target location to the intended insertion response, or (3) the mapping from intended response to action. Experiment 1 and Experiment 3 disconfirmed Mappings 1 and 3, respectively, by showing that training did not alter independent measures of target localization or the action of aiming. Experiment 2 showed that the output of Mapping 2, the planned response -- measured as the initial insertion angle -- was corrected over trials, and the correction magnitude predicted the response to a transfer stimulus with a new represented location.

Keywords

Action; Learning; Spatial representation; Cognitive mediation; Motor Planning

Introduction

Actions in everyday life are guided not only by perceptual cues, but also by more abstract information such as graphic illustrations or narrative text. Verbal estimates of distance from golf caddies may be used by players to plan and generate their swings. Scaled graphical diagrams are also commonly used in sports, for example, to direct basketball players' running patterns for a particular type of offense. Cognitive processes are involved in constructing spatial representations from such textual or graphical descriptions, which cannot be transduced by spatial-perceptual channels. We thus refer to such mental representations, and the actions they guide, as *cognitively mediated*.

Controlled experiments have demonstrated that internal spatial representations derived from cognitive processing, in whole or part, are sufficient to guide actions. Loomis, Klatzky and colleagues (Klatzky, Loomis et al., 2003; Loomis, Klatzky et al., 2002, 2007) showed that when people were asked to walk to a target location without vision, directly or indirectly, they performed essentially equivalently for perceptually encoded targets and for spatial language

*Corresponding Author: Bing Wu, Dept. of Psychology, Baker Hall, Carnegie Mellon University, Pittsburgh, PA 15213, phone: 412-268-4615, fax: 412-268-2798, email: bingwu@andrew.cmu.edu.

specifying the angle and distance to the target location. Simple scaled graphics have been found to provide sufficient information to support walking in the represented world (Richardson, Montello, & Hegarty, 1999). Our previous work has demonstrated that reaching actions in near space can be directed toward hidden targets that are viewed as a scaled graph obtained through ultrasound imaging (Wu, Klatzky, Shelton & Stetten, 2005; Klatzky, Wu, Shelton & Stetten, 2008).

Cognitively mediated input-response couplings, like perceptually guided actions, are subject to errors in spatial representation and motor output, which can be corrected by learning from feedback. We have demonstrated this process in previous studies, where people direct reaching movements toward targets that were imaged through ultrasound (Wu et al., 2005; Klatzky et al., 2008). As illustrated in Figure 1, subjects located a target using an ultrasound image in which an on-screen ruler indicated metric scale, then reported its location by pointing a stylus at it from multiple places. The convergence of these pointing responses was computed to determine the represented target location. Given targets located at depths from 3.5 cm to 6.5 cm, the represented locations, as determined from pointing responses, were systematically displaced upward relative to the true target locations. Moreover, when subjects first attempted to reach a target with the stylus, they initially aimed directly toward the represented location (i.e., their response was too shallow), and the stylus movement then proceeded along a direct line toward that same location (Wu et al., 2005). When feedback was provided, by showing the end position of the stylus relative to the target in the ultrasound image, subjects learned to center the stylus' terminal location, on average, directly on the target. Variability in localization as well as systematic error was reduced with practice (Klatzky et al., 2008).

In the present paper, we ask how people correct errors in actions that are directed to mental spatial locations encoded from graphic and symbolic inputs, like those used in ultrasound. This question is fundamental to furthering our understanding of cognitively mediated action. As was noted above, previous research has shown that linguistic or abstract symbolic designations of space can guide action equivalently to perceptual sources. It has been suggested that this equivalence arises at the level of an abstract, essentially amodal representation of space (Bryant, 1992; Klatzky et al., 2003; Loomis et al., 2007), possibly encoded in parietal cortex. We know, however, relatively little about this internal representation and its role in guiding action, particularly when it arises from cognitive sources.

One important issue has to do with the level of processing at which error correction in cognitively mediated action occurs. This is a question that has arisen in a variety of contexts, and it is clear that the decomposition into levels, as well as the locus of error correction, varies with the task. In reviewing the literature on adaptation to prism-induced sensorimotor discrepancies (Held, 1965; Kohler, 1962), for example, Redding and Wallace (2006) emphasized the mapping from internal spatial representations to motor control, which they suggested could adapt at peripheral or cognitive levels (realignment and recalibration, respectively). Models focusing on motor control (e.g., Bullock et al.'s [1992, 1993] DIRECT model) suggest that one should also consider adaptive learning within relatively late stages, where actions are controlled. Consistently with this idea, Goodbody and Wolpert (1998) suggested that when subjects learned to move to target locations within a distorting force field, adaptation could be localized at such a stage, namely, affecting dynamic control of motor execution by means of a forward model. Verwey and Heuer (2007) examined errors from the nonlinear visuomotor transformation involved in directing a screen cursor to a target location by means of a mouse and found it necessary to distinguish between stages of amplitude specification and trajectory generation.

Here we consider reaching with a tool towards a target that is specified by a graphic image, hence requiring cognitive mediation, as shown in Figure 1. In this case, multiple possible loci

for error correction can again be considered. Our approach is motivated by Soechting and Flanders (1989; Flanders & Soechting, 1990), who described visually guided pointing as a transformation from a perceived target location to implementation of a planned arm orientation. Our model (Figure 2) describes the image-guided reaching task as a sequence of transformations between three levels: (1) a mental representation of the target location in space; (2) the plan of action with specific parameters of the movement; and (3) motor commands with kinematic and dynamic constraints on the execution of movement. Each transformation is conceived of as a mapping that is potentially subject to learning. Under Mapping 1, the symbolically scaled depth of the displayed target, D , is converted into an internal spatial representation D' . That target representation is then used to plan and guide the subsequent motor responses. Under Mapping 2, the represented target location is converted into an action plan; specifically, an insertion angle θ is designated for guiding the stylus from an entry to the represented target location. Mapping 3 converts the intention to act into motor commands, M , which drive muscles and generate sufficient force to produce the desired motion.

Errors could arise at any of these mappings, necessitating correction in the mapping function during learning. We thus distinguish among three loci for correction, each of which could be adjusted according to the feedback of knowledge of the final position of the stylus relative to the target, as shown in Figure 2. These loci resemble the stages defined for adaptation to perceptual distortions, except that the peripheral stage of sensory processing, which might be distorted by a prism in sensory-motor tasks, is replaced by a stage of cognitive encoding, which may exhibit systematic errors prior to training without any inducing mechanism being required.

These issues were addressed with the present task, in which subjects used real-time ultrasound imaging to locate a hidden target in a fluid-filled tank and then insert and guide a stylus to it as illustrated in Figure 1. Ultrasound imaging works on the same principle as sonar: A narrow ultrasound beam emerging from the transducer rapidly sweeps a plane along the axis of the transducer and reflects when encountering the target; the echoes carry information about the target's location relative to the transducer, which are received and converted into a real-time display of the scan plane on a LCD screen along with a metric for depth (in cm, see Figure 1). Several features of this display, which are characteristic of ultrasound in applied settings, require cognitive mediation. First, the ultrasound video, being displayed remotely on an LCD screen, is spatially displaced away from the scanning area emanating from the tip of the transducer. Second, the imaged plane is aligned with the transducer axes, but is mapped to the vertical and horizontal on the screen. Third, the scale of the display is arbitrary and changes with zoom, under control of the user. Thus in order to build up an internal spatial representation of target location from the ultrasound image, the cognitive processes needed include mental rotation, translation, and rescaling.

The action required for this task is to guide a stylus to the target. Insertion of the stylus proceeds in a plane lying perpendicular to the ultrasound scanning plane so that the subject is unable to see the stylus until it intersects the ultrasound image (Figure 1). As was noted above, we have found that subjects initially aim toward the represented target location and proceed along a direct trajectory toward it. They receive feedback only about the terminal position of the stylus relative to the target. The insertion of the stylus typically unfolds over a time-course on the order of 10–20 sec (Wu, et al., 2005), rather than being a ballistic aiming response (< 500 ms; Zelaznik, Shapiro & McClosky, 1981).

Our experimental strategy involved assessing how performance changed as a result of training on the insertion task with feedback over a series of trials, during which initial under-estimation error was corrected. Experiment 1 and Experiment 3 used a pre/post design, where differences between performance before and after training were used to infer whether learning altered a particular mapping. Experiment 2 examined the action trajectory over a series of trials with

feedback; then predicted aiming error in a new target context. The prediction used two parameters: the final aiming angle after training and the represented location of the new target, which was estimated from extant data.

Each of the experiments corresponded to one of the three mappings in Figure 2. Experiment 1 investigated whether the training led to changes in Mapping 1, the process of converting the display into a mental representation of target location. Before and after trials with the target-access task, subjects were asked to report the target location by means of a visual matching paradigm. If visible errors in inserting the stylus led to adjustments in represented target location, we would expect to see these representational changes reflected in subjects' reports, using the matching task, between the pre-learning and post-learning reports.

Experiment 2 and Experiment 3 built on the finding of Experiment 1, namely, that Mapping 1 was not the site of learning, implicating subsequent stages as the loci for error correction. Experiment 2 was designed to expose the changes in post-encoding mappings. In the experiment, subjects first learned the location of a fixed target from a series of insertions with feedback. After that, the apparent target depth in the ultrasound image was manipulated so as to increase the subjectively represented depth for the same target. Three hypotheses were considered as to how training might effect subjects' processing, with implications for differential performance: (a) Subjects might exhibit stimulus-specific learning, leading to no transfer of learning to new stimuli and thus no correction. Their performance then would be predicted entirely by the previously demonstrated representational error in Mapping 1, which would produce under-estimation. (b) Subjects could learn to persevere on a single response, even if they perceived the new target as at a different depth. Although this seems implausible, a priori, it would be objectively correct given our manipulation. Therefore we include it for completeness. (c) Subjects could exhibit a generalized remapping from represented locations to actions, leading them to direct the response to a location determined jointly by the error in Mapping 1 (an error that is known from Experiment 1 to be unaffected by training) and an adaptive correction toward generally deeper responses. This would produce over-estimation error, the magnitude of which would indicate how the correction was adjusted (if at all) when applied to new target representations.

Given confirmation of learning in mappings subsequent to Mapping 1, Experiment 3 was designed to test whether remapping occurred at the level of response execution (Mapping 3). Rejection of this hypothesis would give support to the assumption that the critical learning occurred in processes that convert the target representation to an intended response angle.

Experiment 1: Remapping from input display to represented depth

The purpose of the experiment was to examine whether training a cognitively mediated action produced a change in the mapping from the stimulus display to an internal spatial representation that guides action, which we call Mapping 1. Subjects reported the depth of ultrasound-imaged targets with a visual matching paradigm before and after training trials, during which they repeatedly guided a stylus toward a target and received feedback by seeing its terminal position in the ultrasound image. Our previous research indicated that systematic underestimation should occur in the representation of target depth (Wu et al., 2005; Klatzky et al., 2008), and as predicted, early in training subjects exhibited undershoot errors. With feedback for error correction, subjects improved their insertion performance very quickly. Assuming that the improvement in insertion performance was achieved by taking feedback into account to adjust Mapping 1 and so correct the target representation, subjects should be able to estimate the target depth more accurately after training. Underestimation error should at least be reduced for judging the target depth that had been trained in the insertion task, and the correction might also generalize to other depths.

Method

Participants—Thirteen undergraduates from Carnegie Mellon University participated in the experiment with informed consent. They were naive regarding the purpose of the study. All were right-handed, and had normal or corrected-to-normal vision.

Stimuli—The stimuli were water tanks with two-tier lids of rigid plastic as shown in Figure 3a. The center part of each lid (4.0 cm in radius), which supported the ultrasound transducer, was stepped down by 1.6 cm. relative to the border and over-laid by a soft rubber layer that smoothed the contour. The indentation was clearly visible.

A set of four tanks was used in the depth-matching task, each containing three beads (1.0 cm in diameter) mounted at the target depths (3.5, 5.0 and 6.5 cm from the tank lid) and one dummy bead at random depth. The relative locations of the beads were fixed within a given tank but varied between tanks so as to create an unpredictable stimulus environment. In order to measure the subjects' perception of the depths of the beads using a matching paradigm, a measuring stick was placed alongside each tank. To prevent the subjects from remembering the previous answers in case a target depth was repeated, the measuring stick for each tank had its own unique labels and depth scale varying from 2.0 to 4.0 mm per lattice.

A different tank was used in the training task, inside which one bead (1.0 cm in diameter) was mounted at the depth of 5.0 cm. Two entry points for insertion were positioned on the lid along a radius of 5.0 cm relative to the point on the lid directly above the target, 1.0 cm from the edge of the indented portion. The task was to insert a stylus from one of these entry points to hit the target. Consequently, the entry points determined the desired insertion paths, which for both points had an elevation angle of 45°. The subjects alternated between the two entry points between trials; this alternation was meant to induce postural variation so that the subjects could not perform a purely biomechanical repetition from trial to trial.

Design and Procedure—Subjects were tested individually. The complete procedure consisted of a pre-test of judged target depths, then three learning blocks, followed by one post-test of judged target depths (See Figure 3b).

The pre-test and post-test measured subjects' estimates of target depths using a matching paradigm (see Figure 3a for illustration). The subject was instructed to begin each trial by holding an ultrasound transducer (Model: Pie-Medical 50S Tringa) upright and placing it over one of the four beads. To obtain a clear image of the target, he or she indented the transducer into the lid until it "bottomed out" against the stepped-down surface underneath. The total target depth was thus parsed into two contiguous segments: the displacement of the transducer relative to the unindented outer ring (1.6 cm in this experiment) and the depth from the transducer tip to the target (1.9, 3.4, or 4.9 cm for different target depths). The instruction specifically indicated that both depths had to be considered in order to locate the target: The first segment was directly perceived by vision and touch. The second was shown in the ultrasound image, accompanied by a metric scale that could change under user control; this segment was numerically read in centimeters and then mentally translated into a spatial representation by the subject with reference to a 1cm × 1cm grid placed on the table for use as a standard. The subject was required to estimate the depth of the target relative to the lid and then report a mark on the measuring stick that he or she thought was at the same depth as the target. After all targets in a tank had been tested, the next tank was introduced. The test order of tanks was counterbalanced across subjects. No feedback was given to the subject regarding the accuracy of his or her performance.

The pre-test was followed by the training phase, in which subjects performed repeated trials guiding a stylus to the same target. On each trial, the subject first placed the ultrasound

transducer over the target so as to obtain a clear image, and then judged the location of the target in 3D space and planned the insertion. He or she signaled the experimenter when ready, then immediately inserted the stylus as quickly as possible while maintaining as high a level of accuracy as possible. The stylus was inserted in a plane lying approximately perpendicular to the ultrasound scanning plane, as illustrated in Figure 1; hence the stylus was invisible in the ultrasound image until it reached the scanned area. This eliminated closed-loop control; subjects had to plan and direct the insertion using the initially judged location of the target.

On training insertion trials, feedback was provided by the point where the stylus penetrated the ultrasound imaging plane, allowing subjects to see both stylus and target and so infer how accurate their estimates of the target location were. If the target was successfully reached, the subject received visual confirmation of contact between the stylus and the target and also haptic feedback by pushing the target to demonstrate that the stylus had touched it. Upon failure, the subject was told to compare the visible stylus location in the ultrasound image to the target to understand why it was missed and to try to correct the error in the following trials. A total of eighteen training insertions were performed in three blocks, each consisting of six trials that were performed alternately from the two entry points. In addition, before beginning the training blocks, the participant practiced two insertions for familiarization with the procedure, using a different tank and target. Typically, the training took about half hour to complete, and the entire experiment -- pre-test, training, and post-test -- lasted about one hour.

Results and Discussion

We first consider learning during the training trials. Subjects' performance was measured by success rate, i.e., the percentage of trials in which the target was hit by the stylus. As shown in the inset figure in the upper left of Figure 3c, the subjects improved with practice (one-way repeated measures ANOVA with Session as the within-subjects factor: $F(2,24)=18.527$, $p<0.001$) and the mean success rate increased from 39.7% to 76.3%.

Next consider the pre- and post-tests, where subjects judged target depth by visual matching. Figure 3c plots the mean judged depth before and after training. Depths were generally underestimated, consistent with our previous findings (Wu et al, 2005; Klatzky et al, 2008). Importantly, depth judgments were almost identical before and after the training. A two-way repeated measures ANOVA (Pre/post test $\times$ Target depth) found no significant main effect for Training, $F(1,12)=0.748$, $p>0.4$, or its interaction with Target depth, $F(2,24)=0.287$, $p>0.7$. No improvement was shown even at the target depth of 5.0 cm, which had been used in the training trials (paired t-test: $t(12)=0.990$, $p>0.3$).

The results disconfirm the hypothesis that the improvement in insertion performance was due to changes in mapping from the image to the representation of the target location (Mapping 1). If learning refined the encoding of locations, it should be general enough to be applied to other spatial tasks involving the same representation. Thus we would also expect an improvement in subjects' judgments of target depths in this experiment. However, the pattern predicted on this basis was not observed.

The above conclusion is also supported by the results of a previous learning-transfer experiment by Klatzky et al. (2008), as shown in Figure 4. In Experiment 2 of that paper, subjects first went through three sessions of training trials, in which they repeatedly performed insertions from one set of entry points. As in the present study, a significant improvement was found in their insertion performance across training. They then proceeded to conduct insertions to the same target but from a new, unpracticed entry point, which required a change in response angle. Note that because the target remained unchanged, if training had corrected the encoded target location, it should lead to insertion responses of similar accuracy across different entry points. However, subjects' success rate dropped dramatically with the new response, reverting to a

value close to close to untrained performance in Block 1. As an alternative, it seems more likely that the observed learning occurred at mappings from the encoded representation to the motor response. This hypothesis is examined in the following experiments.

Experiment 2: Learning in post-encoding mappings

This experiment examined the error correction at mappings from the encoded representation of location to the motor response (i.e., Mappings 2&3), by which a mental representation of target location was transformed to an aiming direction and then implemented in action. It was predicted that subjects would gradually correct the initial error, eventually aiming deeply enough to reach the target. After training, a change in the represented depth of the *same objective target* was induced. Re-aiming immediately after training was used to assess learning at combined Mappings 2&3, i.e., from the encoded representation to the observed response. As was noted in the introduction, if learning was specific to a particular target representation, no remapping would be transferred to and observed for new representations, leading to underestimation errors like those of a naïve subject. Alternatively, if learning produced a stereotyped motor pattern, the previous response angle should persevere despite the apparent change in the target depth, leading to correct responses. As a final alternative, changes in post-encoding mapping, if successfully generalized to the new target representation, predict over-estimation error. Transfer of a constant correction in elevation would mean that the error on the first post-training trial should be predictable from (a) the magnitude of the elevation change over the training trials and (b) the new represented depth, which was predicted from prior data. Graded generalization of the correction when applied to new targets would also be manifested in over-estimation, but of a magnitude smaller than the correction of the trained target.

Tests of these hypotheses require that the post-training target is the same as that used in training, but a change in its represented depth is induced. To do so, the experiment capitalized on a previous study (Wu et al., 2008), which measured the represented location of a single target with different lid indentations. The subjects were tested using two stimulus tanks. One had a 1.6-cm indentation in the center, like that used in the present Experiment 1, and revealed a mean representation error of -1.2 cm (Experiment 3 in Wu et al., 2008, where the minus sign denotes underestimation). The other tank had no indentation and produced a mean representation error of approximately -0.3 cm. In the present experiment, subjects were initially trained on a tank with a 1.6-cm indentation, and then transferred to a tank with a target at the same depth but without indentation on the lid. This post-training target should appear as deeper than the training target, although their objective depths were equal. According to the findings of Wu et al. (2008), the represented depth of the target should be increased by 0.9 cm (i.e., the difference between the errors previously observed with the two indentations) but still underestimated under the experimental conditions. The impact of this change in target representation on subjects' initial aiming is used to test the three hypotheses specified above. To reiterate, each hypothesis predicts a different pattern of performance. Specifically, target-specific learning predicts under-estimation; motor perseveration predicts accurate responses; and generalized post-encoding remapping predicts over-estimation error, of a magnitude that can be predicted from previous data.

Method

Participants—Sixteen naïve subjects with informed consent were tested. All had normal or corrected-to-normal vision.

Stimuli—The stimuli were tanks of opaque fluid, as illustrated in Figure 3a. Two tanks were used: the training tank was as in Experiment 1 with an indentation of 1.6 cm on the lid; the post-training tank was identical to the training tank except for having zero indentation. Inside the tanks, a bead (1.0 cm in diameter) was mounted 5.0 cm below the un-indented lid. On the

lids, two insertion entries were positioned along a radius of 5.0 cm. Therefore, the desired insertion path from each entry point to the target had an elevation angle of 45° with both tanks.

Design and Procedure—The design of the experiment is shown in Figure 5a. Four blocks of trials, three training blocks plus one post-training block, were conducted. After completing the training blocks, the subject proceeded immediately to the post-training block. Each of these blocks consisted of six trials that were performed alternately from the two entry points. The training and post-training trials used the same procedure as that used in the insertion trials in Experiment 1. One procedural difference between this and the previous experiment was that subjects received minimal training on the insertion task, so as to prevent them from gaining prior experience before starting the experimental trials. To familiarize them with the task, the experimenter first gave them a demonstration and then let them do only one insertion using a different tank and target. The first trial with the experimental target served as the pre-test measure.

At the beginning of each insertion trial, the subject first observed the target in an ultrasound image, estimated its location inside the tank, and then planned the insertion. The subject was instructed that once ready, he or she should insert the stylus as quickly as possible while maintaining as high a level of accuracy as possible. Once the target was reached, the subject was to push the target a little to prove that the stylus had really touched it. If the target was not reachable along the given trajectory, the subject was required to compare the stylus location in the ultrasound image to the target to estimate the error and to try to correct it in the following trials, as illustrated in Figure 1. The insertion trajectory was recorded by a magnetic tracker (miniBIRD 500; resolution: 0.1° in orientation) mounted on the stylus with a sampling rate of 103.3 Hz.

Results and Discussion

For analysis, the subject's insertion response was broken into three phases: (1) pre-insertion aiming, (2) insertion of the stylus to the point where it appeared in the ultrasound image, and (3) subsequent insertion using visual guidance from the ultrasound image. Typically, the subject inserted the stylus along the initial direction until it became visible in the image. After that point, the subject could adjust the insertion by seeing the relative position of the stylus to the target. For present purposes, only the accuracy of pre-insertion aiming was assessed, because it indicated how the subject planned to insert the needle. The plan was measured in terms of the stylus's initial aiming angles (azimuth and elevation) at the beginning of the insertion phase. The elevation error in aiming was of particular interest, because the experimental manipulation concerned target depth. Figure 5b plots the mean elevation error as a function of trial order. Of particular importance are the following trials: the first insertion in Block 1 at the beginning of training, the last trial of Block 3 at the end of training (learned state), and the first insertion after transfer in Block 4.

Consider the very first insertion at the beginning of training. The response angle was significantly undershot in aiming ($t(15)=4.383$, $p<0.001$). The mean elevation error was $-6.2^\circ \pm 1.4^\circ$ (where the minus sign denotes aiming shallower than the target). This error value was not significantly different ($t(15)=0.891$, $p>0.3$) from the data of Wu et al. (2008, Experiment 3), where the judged depth (estimated by triangulation from pointing) was 3.83 ± 0.15 cm for the same target under the same experimental settings by a different group of twenty-four subjects, and the corresponding angular error was $-7.5^\circ \pm 1.1^\circ$. The aiming error and its congruence with previous pointing data indicate that the insertion in the present task was planned in response to the judged target location, if no additional correction or feedback was provided.

The undershoot error was quickly detected and eventually corrected across trials. At the end of the training (the last trial of Block 3), the mean elevation error in aiming was only $-0.4^\circ \pm 0.9^\circ$, which did not significantly differ from zero ($t(15)=0.498$, $p>0.6$).

An overshoot error occurred in the first post-training trial of Block 4, where the new target representation was induced ($5.3^\circ \pm 0.7^\circ$, $t(15)=7.271$, $p<0.001$) by removing the lid indentation. This violates the first hypothesis, namely, that there would be no generalization after learning. If the training-induced adaptation in Block 1–3 did not generalize, the subject would perform as if untrained. We can estimate the error that would be produced by an untrained subject with the new tank, by means of data from our previous study with the same experimental settings (i.e., target depth 5.0 cm and un-indented lid; Experiment 3 in Wu et al. 2008; see Figure 4c). There the target depth was found to be under-estimated as 4.66 ± 0.19 cm. In short, the hypothesis that learning failed to transfer would lead to undershoot rather than overshoot errors as were observed. We can also rule out the hypothesis of response perseveration, which predicts no error with the new target representation.

We now turn to the hypothesis that subjects adjusted mappings subsequent to encoding and generalized the remapping to the new target representation. In fact, the amount of overshooting was close to the prediction from that hypothesis, based on a combination of the training-induced correction and the error in estimation of target location. Given that the error in the first insertion in Block 1 was -6.2° , a full correction would be 6.2° after training. One must also consider, however, the error in the representation of target depth with the post-training display. According to our previous data (Wu et al., 2008; see Figure 4c) the average representation error was -0.34 cm for a target depth of 5.0 cm, producing an angular error of -2.0° in the insertion direction. By summing the amount of correction during learning and the error in represented depth, we would expect an overshoot error of 4.2° . A t-test compared this prediction with the data (5.3° error) and found the difference was not significant ($t(15)=1.572$, $p>0.1$). This suggests that not only did the elevation correction from training transfer to the new target representation, but there is no evidence for a reduction in its magnitude.

To summarize, the training resulted in a correction that generalized to a new mental representation of target location. After training, insertions to a new target representation were found to be planned in response to that representation and also to be influenced by the training-induced adaptation, consistently with applying a constant bias. In addition, given that the observed overshoot was actually greater than the prediction, there is no evidence that the adaptation was reduced in the post-training trials.

Experiment 3: Remapping from intended response to action

Experiment 1 showed that training produced significant changes in subjects' insertion responses and ruled out Mapping 1, from the input display to an internal spatial representation, as the locus of learning. Experiment 2 confirmed remapping from the encoded spatial representation to the response, but it did not specify which mapping stage subsequent to encoding was altered by learning. As illustrated in Figure 2, Mapping 2 converts the spatial representation of target location, together with the starting point of the insertion, to a planned insertion angle; Mapping 3 then generates and executes the motor commands to direct the needle along that angle to the target. Experiment 3 tested the hypothesis that training altered Mapping 3. Before and after a series of training trials, subjects were tested in an orientation-matching task, which required them to hold and rotate a stylus with their unseen hand in order to replicate a visually specified orientation. This task uses the explicit representation of a desired orientation to elicit a plan and hence isolates Mapping 3, from plan to execution of the orienting response. If Mapping 3 was altered, a difference would be expected between subjects' responses before and after training. Specifically, if the training changed Mapping 3 so as to

compensate for the under-estimation exhibited in encoding, the responses after training should be remapped so as to over-estimate the required response angle.

Method

Participants—Fifteen naïve subjects with informed consent were tested. All were right-handed and had normal or corrected-to-normal vision.

Stimuli—Seven orientations were tested in the orientation-matching task used as a pre-test and post-test, ranging from 15° to 60° with a step of 7.5° . They were presented to the subject in a random order as graphs with a thick solid line rotated upward by the desired angle, relative to a horizontal reference line. Each was tested with three repetitions. In addition, mixed in with 21 test trials, there were four dummy trials of 0° and 90° . The stimuli for the training task were as in Experiment 1.

Design and Procedure—The design was the same as Experiment 1 (Figure 6a); that is, the phases were a pre-test, a series of training trials with the insertion task using the indented lid, and a post-test. The difference was the task in the pre-test and post-test phases, which required matching a visual angle with a wrist rotation. On each trial of the pre- and post-test task, the subject sat in front of a wood board, on which the stimulus graph was placed. He or she positioned the right arm on an armrest next to the seat, holding a stylus in the right hand. A large board ($1.0\text{ m} \times 2.0\text{ m}$) was used to occlude view of the hand, as depicted in Figure 6a. Subjects were asked to rotate the stylus and replicate the orientation of the stimulus line (i.e., to make the stylus parallel to the line shown in the stimulus graph). They were encouraged to be as accurate as possible; no time limit was set. Once ready, the subject signaled the experimenter, who pressed a computer key to record the response angle taken from the same magnetic tracker as in Experiment 2 mounted on the stylus; no feedback was given. Before the pre-test, subjects completed two practice trials at 0° and 90° to familiarize them with the procedure.

Results and Discussion

The matched orientations in the pre-test and post-test, averaged across subjects, are shown in Figure 6b. Subjects' pre- and post-training responses were very similar and exhibited a common pattern: orientations were over-estimated at small angles, for example, 10° to 30° , while large angles ($>45^\circ$) were matched relatively accurately. A two-way repeated measures ANOVA, with factors of stage of test (pre- or post-training) and orientation (7 levels) found no significant main effect of stage, $F(1,14)=0.340$, $p>0.5$, and no significant interaction, $F(6,84)=1.491$, $p>0.2$.

Considering the training trials, the success rate again improved significantly across training sessions (one-way repeated measures ANOVA: $F(2,28)=23.791$, $p<0.001$). If this improvement was produced by remapping from the intended response angle to the action, so as to compensate for encoding errors, the training-induced adaptation should be observed in the orientation-matching task in the form of post-test overestimations. However, the comparison of the pre- and post-test found no change.

It might be argued that the failure to observe transfer arises because the orientation-matching task used in the pre- and post-tests differs substantially from the aiming action trained in the learning trials. The plausibility of this interpretation is heightened by the fact that the end goals of the pre/post and training tasks differ, which could translate into different motor commands. We acknowledge this possibility; however, we would argue that the pre/post task isolates an essential component of the training task, namely, the process of Mapping 3: from intended angle of response to the motor achievement of that response. Isolating Mapping 3 for the pre-

and post-tests necessitated differences from the training task, including providing an explicit representation of the target orientation and withholding visual feedback. It should also be noted that the use of multiple insertion points, and therefore multiple required trajectories, in training was intended to make learning more general than a stereotyped motor response, which might be more dependent on task similarity for cross-task interference to occur. In short, under the assumption that the pre/post task instantiates a mapping from intended angle to execution, as does the training task, the experiment includes what is required to test the hypothesis. On this basis, re-mapping at the level of Mapping 3 is contra-indicated by the results of Experiment 3.

General Discussion

The present study demonstrates that cognitively mediated action, like perceptually guided action, is subject to correction, as feedback about errors is assimilated. The experiments examined the learning of an image-guided reaching task and found evidence localizing the observed error correction to re-mapping from a represented target location to the action aimed at that location. The results rule out two other mappings that might, a priori, have been assumed to mediate the observed learning, namely, from the stimulus display to a mental spatial representation that guides action, and from intended action to motor outcome.

Unchanged mapping from input to representation

Unlike a visually-guided pointing response to a target in 3D space, the reaching action examined in this study was directed by a spatial representation created through cognitive mediation. Subjects first used real-time ultrasound imaging to locate a hidden target. They then aimed and inserted a stylus to the target location along a trajectory that was invisible until the stylus intersected the ultrasound scanning plane (Figure 1). The task thus can be broken down into the following specific subtasks: localizing the target, planning the insertion accordingly, and generating motor commands. In the context of image-guided action, localizing the target corresponds to building a mental representation of the target location in 3D space, using the depth information in the image relative to the transducer. In Experiment 1 we found little effect of training on this subtask. If learning had altered the process of encoding target locations, one might have expected carry-over of improvement to subjects' estimation of target depth. Yet in contrast to the considerable learning of the insertion response across the training trials, no change was found when judgments of target location were directly assessed with a matching paradigm before vs. after training (Figure 3).

A similar point was made in the context of perceptually encoded environments by Richardson & Waller (2005, Experiment 2). Their subjects walked while blindfolded to locations in an immersive virtual environment, either directly or indirectly. Error-corrective feedback was provided only for the direct-walking task. If the feedback affected location encoding generally, improvements should also be found with indirect walking. However, although subjects' accuracy improved by 42% in direct walking after training, there was only a 13% change in indirect walking. This finding indicates that the improvement did not occur at the level of representing target location.

Prism-adaptation studies also provide examples where adaptive recalibration fails to occur at the general level of encoding a target representation. It has often been demonstrated that inter-manual transfer of adaptive effects is absent or very small. For example, Martin et al. (1996) asked subjects to use one arm to make underhand or overhand throws with prisms. Prism adaptation occurred in the throwing arm, but did not affect throwing with the untrained arm; moreover, there was no transfer of training across different throwing postures with the same arm. This suggests that rather than being perceptual in origin, the adaptation was narrowly applied to the motor response.

In short, a variety of studies using perceptual encoding, but differing in the nature of encoding and the scale of the response, indicate that feedback about final results of actions benefits a mapping other than that from the stimulus to a representation of its location. The present Experiment 1 extends this pattern to representations that are encoded from a scaled graphical display, using cognitive mediation.

Unchanged mapping from intended action to motor outcome

Using a procedure analogous to Experiment 1, but now evaluating the change in motor outcome after training, Experiment 3 evaluated whether error feedback altered Mapping 3, from the intended action to the action itself. Motor-level learning has been found, for example, when people repeatedly perform the task of lifting an object with a constant weight (Johansson & Westling, 1988). The grip and load forces during the lift-off are gradually adjusted to the object, so that when a new object is introduced, there is a carry-over in the form of inappropriately large force for a lighter object or small force for a heavier one. However, this form of adaptation has been attributed to a memory system at the sensorimotor level (Johansson, 1991), which seems unlikely to play a role in the present paradigm where cognitive mediation plays a critical role.

Accordingly, Experiment 3 found that training corrected subjects' insertion responses, but did not change their responses to match a visually specified orientation on the post-training trials. That is, the response pattern was altered when made in response to the training stimulus, but not to another input that called for the same type of action. This supports the conclusion that the training altered the intended action and not its instantiation in the motor system.

Recalibration of the representation-action coupling

While Experiment 1 and Experiment 3 indicated that Mappings 1 and 3 are unchanged by feedback, Experiment 2 provides evidence that what is remapped with training in our task is the link from spatial representation to action, which takes the encoded location as an input and plans the reaching action in terms of a corresponding aiming angle. The experiment further indicated that the learning generalized without reduction across an approximate 35% shift in represented depth.

It is particularly interesting to contrast the malleability of the two mappings that lead to a specification of the response location but precede action per se. Why, one may ask, is error correction localized at Mapping 2, from mental representation to intended action, rather than eliciting a change in Mapping 1, i.e., from the stimulus input to the mental representation itself? One possibility is that Mapping 2 is intrinsically labile. According to Schmidt's (e.g., 1975; 1988) theory, the nature of motor learning makes it particularly malleable. Performers of a habitual action are said to acquire a motor program, which can be freely parameterized (within biomechanical constraints). Learning then corresponds to the tuning of a parameter, which can be quite rapid. Another possibility, rather than plasticity in Mapping 2, is that Mapping 1 is intrinsically resistant to change, forcing adaptation at another point in order to reduce error. In other words, one could attribute the site of remapping in terms of readiness to change Mapping 2 on the output side or inflexibility on the input side.

The nature of image mediation may itself direct the re-calibration process toward Mapping 2 instead of Mapping 1. In the typical perceptual paradigm, error in target encoding is signified by a discrepancy between two distinct sensory systems or cues that provide the same spatial coordinate values (Wallach, 1968). If one of these can be determined to be correct, it will specify the required remapping. For example, if a person wearing prisms tries to point in a target direction, the kinesthetic feedback from movement is aligned with the motor plan but not with the visible outcome, and the discrepancy can be attributed to perceptual encoding.

Remapping then occurs to reduce or eliminate the inconsistency. That process, however, relies on the ability to detect the source of the misinformation. In the present image-guided paradigm, information about the response outcome is placed in arbitrary coordinates in the video display along with the target location, separating it from the motor system that produced the response. There is no direct sensory-motor coupling that allows for validation of the response. From the image, the viewer cannot determine whether the error stemmed from misperception of the target location or an inaccurate action.

Further Issues

As shown in Experiment 2, the calibration of Mapping 2 effectively generalized to a new represented depth without detectable moderation of the correction magnitude. The present paradigm could not assess generalization across a broad range of targets, given the limits on the physical apparatus. It would be useful, however, to determine how the mapping was recalibrated over a broader range, in order to assess the form of the generalization function.

In a recent study addressing a similar question, Mon-Williams and Bingham (2007) investigated the recalibration of direct reaching by distorting the visual or haptic feedback. Although the feedback was provided at one particular distance, the induced recalibration was observed across the whole reaching space. Specifically, the linear relationship between stimulus distance and reach distance was altered in terms of both slope and bias, which were shown to be calibrated independently. Considering the training effect on Mapping 2 in the present studies, given that the initial insertions were usually undershot, subjects could make an adjustment by adding a constant bias or proportional increase to the planned insertion. Although the results from the single new location in Experiment 2 are consistent with the constant-bias prediction, the data could also be fit by a proportional correction, given the measurement precision. Future work with a virtual ultrasound display (Shelton, 2007) will allow us to assess more data points without the limitations of real water tanks. However, the rapidity of re-learning observed in Experiment 3 suggests that it will be difficult to assess the generality of remapping in a within-subjects design.

A question that remains unanswered is the level at which Mapping 2 arises when the input is encoded with cognitive mediational processes. Does the mapping from cognitively mediated location produce an intention to act at a cognitive level, or is it more deeply embedded in the perception-action system? At an extreme, one could conceive of the intention as verbal ("when it looks 4 cm deep, aim for 5 cm deep"). The present data cannot definitely determine the level at which adaptation occurred, but whatever that level, the task requires that it be translated into spatially directed action. As described above, the insertion task relies on a spatial representation and not on ongoing visual feedback. Hence, learning at a purely cognitive level does not explain how the changed cognitive representation mediates a change in spatial action.

Another argument against purely cognitive learning comes from evidence that broad generalization after adaptive training is characteristic of change at relatively peripheral levels, and narrower generalization is more central (e.g., Bedford, 1993). Applying this rubric to the present data, the degree of transfer to a new represented location that was observed in Experiment 2 suggests that the change in mapping from representation to action was not simply mediated by a cognitive rule with a narrow range of application. Again, it would be useful to have a broader range of post-training locations to further address this question.

In any case, the present paper provides a provocative extension of an old problem: How do people recalibrate systems when presented with errors in spatially directed action? We have shown that considering this question in regard to cognitively mediated representations of space is a fruitful line of investigation. Remapping of spatially directed action appears to be a fundamental process that applies to representations arising from cognitive as well as perceptual

sources. This commonality supports previous studies, cited in the introduction, pointing to the existence of an amodal spatial representation that functions equivalently across a variety of input pathways.

Acknowledgement

This work is supported by grants from NIH (R01-EB00860 & R21-EB007721) and NSF (0308096). Parts of this work were presented at the 2007 Annual Meeting of the Vision Sciences Society.

References

- Bedford FL. Perceptual and cognitive spatial learning. *Journal of Experimental Psychology: Human Perception and Performance* 1993;19:517–530. [PubMed: 8331313]
- Bryant DJ. A spatial representation system in humans. *Psychology* 1992;3–16.
- Bullock, D.; Grossberg, S.; Guenther, FH. A self-organizing neural network model for redundant sensory-motor control, motor equivalence and tool use; *IJCNN Proceedings, IV*; 1992. p. 91-96.
- Bullock D, Grossberg S, Guenther FH. A self-organizing neural model of motor equivalent reaching and tool use by a multijoint arm. *Journal of Cognitive Neuroscience* 1993;5:408–435.
- Flanders M, Soechting JF. Parcellation of sensorimotor transformations for arm movements. *Journal of Neuroscience* 1990;10:2420–2427. [PubMed: 2376780]
- Goodbody SJ, Wolpert DM. Temporal and amplitude generalization in motor learning. *Journal of Neurophysiology* 1998;79:1825–1838. [PubMed: 9535951]
- Held R. Plasticity in sensory-motor systems. *Scientific American* 1965;213:84–94. [PubMed: 5828465]
- Johansson, RS. How is grasping modified by somatosensory input?. In: Humphrey, DR.; Freund, H-J., editors. *Motor control: Concepts and issues*. New York: John Wiley & Sons; 1991. p. 331-355.
- Johansson RS, Westling G. Coordinated isometric muscle commands adequately and erroneously programmed for the weight during lifting tasks with precision grip. *Experimental Brain Research* 1988;56:550–564.
- Klatzky RL, Lippa Y, Loomis JM, Golledge RG. Encoding, learning and spatial updating of multiple object locations specified by 3-D sound, spatial language, and vision. *Experimental Brain Research* 2003;149:48–61.
- Klatzky RL, Wu B, Shelton D, Stetten G. Effectiveness of augmented-reality visualization vs. cognitive mediation for learning actions in near space. *ACM Transactions on Applied Perception* 2008;5:1–23.
- Kohler I. Experiments with goggles. *Scientific American* 1962;206:62–86. [PubMed: 14457820]
- Loomis, JM.; Klatzky, R.; Avraamides, M.; Lippa, Y.; Golledge, R. Functional equivalence of spatial images produced by perception and spatial language. In: Mast, F.; Jäncke, L., editors. *Spatial processing in navigation, imagery, and perception*. New York: Springer; 2007. p. 29-48.
- Loomis JM, Lippa Y, Klatzky RL, Golledge RG. Spatial updating of locations specified by 3-D sound and spatial language. *Journal of Experimental Psychology: Human Learning, Memory, and Cognition* 2002;28:335–345.
- Martin TA, Keating JG, Goodkin HP, Bastian AJ, Thach WT. Throwing while looking through prisms II. Specificity and storage of multiple gaze-throw calibrations. *Brain* 1996;119:1199–1211. [PubMed: 8813283]
- Mon-Williams M, Bingham GP. Calibrating reach distance to visual targets. *Journal of Experimental Psychology: Human Perception and Performance* 2007;33:645–656. [PubMed: 17563227]
- Richardson AE, Montello D, Hegarty M. Spatial knowledge acquisition from maps, and from navigation in real and virtual environments. *Memory & Cognition* 1999;27:741–750.
- Richardson AR, Waller D. The effect of feedback training on distance estimation in virtual environments. *Applied Cognitive Psychology* 2005;19:1089–1108.
- Redding GM, Wallace B. Generalization of Prism Adaptation. *Journal of Experimental Psychology: Human Perception and Performance* 2006;32:1006–1022. [PubMed: 16846294]
- Schmidt RA. A schema theory of discrete motor skill learning. *Psychological Review* 1975;82:225–260.

- Schmidt, RA. Motor Control and Learning: A Behavioral Emphasis. 2nd ed.. Champaign, IL: Human Kinetics; 1988.
- Shelton, D. Unpublished doctoral dissertation. Carnegie Mellon University; 2007. Virtual tomographic reflection for psychophysical analysis of the Sonic Flashlight.
- Soechting JF, Flanders M. Sensorimotor representations for pointing to targets in three-dimensional space. *Journal of Neurophysiology* 1989;62:582–594. [PubMed: 2769349]
- Verwey WB, Heuer H. Nonlinear visuomotor transformations: Locus and modularity. *Quarterly Journal of Experimental Psychology* 2007;60:1629–1659.
- Wallach, H. Adaptation based on cue discrepancy. In: Held, R.; Taub, E.; Freeman, SJ., editors. *Neuropsychology of Spatially Oriented Behavior*. Dorsey Press; 1968. p. 209-229.
- Wu B, Klatzky RL, Shelton D, Stetten G. Mental concatenation of perceptually and cognitively specified depth to represent locations in near space. *Experimental Brain Research* 2008;184:295–305.
- Wu B, Klatzky RL, Shelton D, Stetten G. Psychophysical evaluation of in-situ ultrasound visualization. *IEEE Transactions on Visualization and Computer Graphics (Special Issue: Haptics, Virtual and Augmented Reality)* 2005;11:684–699.
- Zelaznik HN, Shapiro DC, McClosky D. Effects of a secondary task on the accuracy of single aiming movements. *Journal of Experimental Psychology: Human Perception AND Performance* 1981;7:1007–1018. [PubMed: 6457105]

Figure 1.

Systematic underestimation in the image-mediated judgments of target location. Hidden targets were observed through ultrasound images that provided a numerical metric for depth. Subjects reported the judged location of a target in 3D space by pointing a stylus at it. Significant underestimation was observed in the perceived target location (Wu et al., 2005).

Figure 2.

The image-guided reaching task characterized as a sequence of mapping operations: (1) localization of the target in space using the target depth displayed in the image relative to the imaging position; (2) transforming the represented target location to a planned aiming orientation; and (3) motor implementation of the aiming response. Learning from feedback may change one or more of these mappings.

Figure 3.

(a) Illustration of the experimental set-up used to assess subjects' judgments of target location in Experiment 1. (b) Design of Experiment 1. The pre-test and post-test measured the participant's estimation of target depth relative to the lid using a matching paradigm. The training consisted of three sessions and the task was to direct a stylus into a target with ultrasound guidance. (c) Mean judged depth as a function of the target depth in the pre- and post-tests. The inset figure shows the average success-rate of insertions across three training blocks. Error bars represent ± 1 standard error.

Figure 4.

Mean success rate of insertions across training and transfer blocks. Subjects practiced stylus insertions from one set of entry points in Block #1–3, then proceeded immediately to conduct insertions toward the same target but from different entry points in Block #4. The open and fill circles in the illustration of Block #4 show the relative locations of different entry points. Error bars represent ± 1 standard error. (Adapted from Figure 6 of Klatzky et al, 2008, reprinted with permission of ACM)

Figure 5.

(a) Design of Experiment 2. The indentation of the probe area was 1.6 cm in the training trials (Block #1–3) and zero in the post-training test (Block #4). The removal of indentation would induce a change in the representation of target location, although the target was same (Wu et al, 2008). (b) Mean elevation error in the initial aiming of the needle across the training (*indented*) and post-training (*non-indented*) trials. The minus (plus) values denote aiming toward a place shallower (deeper) than the target. Error bars represent ± 1 standard error.

Figure 6.

(a) Design of Experiment 3. Subjects were tested in an orientation-matching task in the pre-test and post-test phases. The training was same as Experiment 1. (b) Mean matched orientation as a function of the target orientation in the pre- and post-tests. The inset figure shows the success-rate of insertions across the training blocks. Error bars represent ± 1 standard error.