Mental motor imagery and the body schema: evidence for proprioceptive dominance

Jeffrey T. Shenton a, John Schwoebel b, H. Branch Coslett a, c

a University of Pennsylvania, 3 West Gates Building, 3400 Spruce St., Philadelphia, PA 19104, USA
b Division of Social & Behavioral Sciences, Psychology Program, Cazenovia College, Cazenovia, NY, USA
c Moss Rehabilitation Center, Philadelphia, PA, USA

Received 9 April 2004; received in revised form 11 June 2004; accepted 17 July 2004

Abstract

Previous studies have demonstrated that both visual and proprioceptive feedback influence motor control. The relative contributions of these sensory modalities to the on-line computation of body position—that is, the body schema—remain unclear. We report a study designed to explore the roles of vision and proprioception in motor planning. The task required subjects to judge if a pictured stimulus was a right or left hand; stimuli included pictures of a right or left hand in a palm up or palm down position and in six different angular rotations (0°, 60°, 120°, 180°, 240°, 300°). Each subject was tested with his/her right hand palm down and palm up. There were three conditions: a “control” condition (real hand in view), a “fake hand” condition (fake hand in view, real hand out of view), and a “proprioception” condition (no fake hand, real hand out of view). We found that proprioceptive input (that is, the subject’s “felt position”) had a significant influence on mental rotation whereas the visually perceived posture of the hand did not. We suggest that, at least under some circumstances, proprioceptive inflow may represent the dominant sensory input to the on-line representation of the body in space.

© 2004 Elsevier Ireland Ltd. All rights reserved.

Keywords: Motor imagery; Body schema; Motor planning; Vision; Proprioception

Execution of a motor act requires knowledge of the location of body parts in space. This information comes from two major sources. One source is efference copy information. In recent models of motor control, efference copy informs the “forward model” that in turn predicts the location of a body part in action [6,30]. The second source of information regarding body part position is provided by feedback from sensory systems. The integration of these two processes has been hypothesized to produce an on-line, real-time representation of the body position that we and others have termed the “body schema” [4,5,30].

There are two main sources of sensory inflow available to provide this on-line feedback to the CNS: vision and proprioception. The manner in which this sensory information is integrated remains unclear. There are several possibilities as to the relative importance of these sensory signals. Visually based location information, when it is available, may override proprioception. Alternatively, as visual information is not always available, proprioception may provide the primary source of information about body part location. Lastly, vision and proprioception may both be critical to the planning of action with their respective roles weighted according to the specifics of the planned action and the environment in which it is to be executed.

Single-cell recording studies in monkeys have shown that afferent sensory inputs are integrated at the neural level. Graziano [8] identified neurons in the monkey premotor cortex that responded both to a touch of the arm and to a visual target within reaching distance of that arm. When the monkey’s arm was approached by a visual target, activation was observed both when the arm was in the monkey’s view and when it was covered; activation was greater, however, in the former condition, suggesting that arm position was coded with respect to both vision and proprioception. Critically, the receptive fields of the neurons were altered by the
introduced a biologically plausible fake arm, and shifted independently of the felt position of the covered real arm. The shift of the receptive field, however, was not as large as was observed in response to the shift of a real arm moved congruently with proprioception, suggesting that these neurons were integrating information from both the visual and proprioceptive input of the real arm. In a second study, Graziano et al. [9] again used a fake arm paradigm to show that neurons in area 5 of the parietal lobe of the monkey responded most strongly to congruent visual and proprioceptive sensory information. Neurons in this area responded significantly to the introduction of a fake arm in any biologically plausible position, including palm up, while the real hand remained palm down. Neural activity was not significantly increased, however, when the handedness of the fake arm was incorrect (i.e., ipsilateral to the recording lobe) or when the fake arm was placed with the fingers pointing to the shoulder.

Studies in humans have suggested that vision and proprioception may interact in a complex manner in the planning of action. Using prisms to displace visual input of hand location before a reaching movement to a non-displaced target was initiated, Rossetti et al. [19] showed that the CNS integrated both visual and proprioceptive information in choosing target location. The bias in pointing was only 1/3 of the total angular shift of the prisms, demonstrating that the CNS integrated the conflicting information from kinesthetic space as well as visual space in constructing an initial, static hand posture for initiating movement to a target.

Several studies have addressed the relative roles of vision, proprioception, and tactile inputs in the calculation of the body schema under conditions where no action was required. For example, Botvinick and Cohen [1] found that gazing at a fake hand placed in close proximity to an unseen real hand caused subjects to report profound feelings of ownership of the false hand. When the fake hand was stroked congruently with the unseen real hand, subjects reported feeling the sensation at the location of the fake hand. The investigators further noted that subjects exhibited a bias to point in the direction of the fake hand while performing a hand-position matching task with their eyes closed. These findings were interpreted as an “intersensory bias” favoring the visual signal, and as an indication that visual position information could override proprioceptive information. Similarly, Pavani et al. [17] demonstrated that correct localization of a tactile stimulus on an unseen hand took substantially longer when visual information and tactile information were incongruent. Subjects also reported that they “felt” the stimulus at the position of the fake hand rather than their real hand. These effects were eliminated when the fake hands were moved such that the fingers were pointing toward the subject. This was interpreted as evidence that plausible visual input can modulate proprioceptive information.

Studies by van Beers et al. [24,25] performed a series of experiments in which subjects moved their (unseen) left hand underneath a table to indicate the position of a target on the table surface; in different conditions, target location was defined by vision, proprioception (the position of the right hand) or both vision and proprioception. They found that the variance in responses was smaller when both vision and proprioception were available. In a follow-up investigation [26], this group used vision-displacing prisms in a similar arm-matching paradigm. The best-fit model was one in which visual and proprioceptive information was integrated in a direction-dependent, rather than direction-independent way, indicating that the CNS integrated visual and proprioceptive information based on the accuracy of the unimodal signal. Vision was found to be more precise than proprioception in azimuth, while proprioception was more precise in the radial dimension (in depth). The findings were interpreted as strong evidence that the calculation of static body position in space relies on sensory integration (see also [27]).

The current study employs a mental motor imagery task to explore the relative contributions of vision and proprioception in motor planning. The use of our behavioral task is motivated by the observation that motor imagery is, in many critical respects, closely related to action. One line of evidence supporting a link between motor imagery and action comes from the work of Parsons [14]. Parsons reported that the time taken to execute a movement was highly correlated with the time taken either to explicitly or implicitly imagine the execution of the same movement. This led to the conclusion that a judgment of handedness for a pictured hand was mediated by mentally rotating one’s own limb to match the stimulus (see also [10] and [23]). Additional support for the claim that mental motor imagery relies on the same processes involved in action comes from studies involving functional imaging [11,15,16] and transcranial magnetic stimulation [7].

The current study was motivated by the desire to link work on action and motor planning [19,29] to investigations of the integration of multiple sensory inputs [1,17,24–27] in generating the body schema. More specifically, we sought to determine whether vision or proprioception was more important in coding the on-line position of the body in space during a motor planning task that did not require movement to a target. To this end, we used Parson’s hand laterality task and introduced a fake hand into the paradigm. Given the evidence from monkeys and man cited above demonstrating that a “fake” hand presented in a biologically plausible position may be incorporated into the on-line representation of the body in space, we reasoned that the coding of hand position for action may be determined by visual information; if this were the case, RTs in the fake hand condition would appear as if all mental rotations were initiated from the palm down position indicated by the fake hand, regardless of the position of the subject’s real hand. In contrast, if proprioception dominates the coding of hand posture, RT performance would be determined by the subject’s real hand posture, regardless of the position of the biologically plausible visual input.
Subjects included 54 right-handed individuals (31 female), with no history of neurologic or visual deficits. The mean age of participants was 21 years (range 17–31). The age of the participants did not differ between experimental groups, $F(2, 34) = 1.6, P < 0.212$. Subjects were randomly assigned to groups. All subjects gave their informed consent prior to testing, and were paid for their participation.

Stimuli were developed by Parsons [13,14,16], and included line drawings of a human hand in a total of 24 positions, 12 each of a left hand and a right hand. Both the left hand and the right hand were shown in six palm down and six palm up orientations, measured in degrees of angular rotation beginning with the fingers pointing away from the body (that is, straight up on the computer screen), and measured from medial to lateral orientations ($0^\circ$, $60^\circ$, $120^\circ$, $180^\circ$, $240^\circ$, and $300^\circ$).

Participants sat in a chair in front of a table, the horizontal surface of which was at mid-chest level. Stimuli were presented on a computer screen placed on the table in the subject's midline, approximately 2 ft from the subject's body. Stimuli were presented in four randomly ordered blocks. The first two blocks consisted of 48 trials each; they were considered practice and were not included in the analysis. The second two blocks consisted of 144 trials each. The task was self-paced, and in all conditions subjects were asked to respond as quickly and accurately as possible as to whether the stimulus was a right hand or a left hand. Subjects responded with their feet on one of two foot-switches; subjects responded with their right foot for a stimulus of a right hand, and with their left foot for a stimulus of a left hand. In all three conditions, subjects were instructed to refrain from moving their hands or body. The left hand was always situated palm down on the table, within the normal field of vision. In order to minimize visual differences between subjects' hands and the fake hand, subjects wore black gloves on both hands. The area from the neck to the wrists was covered with a cloth, so as to limit visual input to the table. A between-subject design was employed because we were unable to recruit a sufficient number of subjects who could return for three lengthy testing sessions. Eighteen subjects were tested in one of three different conditions: a control condition, a “fake hand” condition, and a “proprioception” condition.

In the control condition both visual and proprioceptive input were manipulated simultaneously and congruently, by having the subject perform one test block with the right hand palm down on the table (in view), and the other block palm up. In the “fake hand” condition, the subject's right hand was placed under the cloth on the right thigh and a fake hand covered with a black glove was placed palm down in a natural orientation adjacent to their own palm-down left hand. One test block was performed with the right hand palm down on the thigh, and one block with the right hand palm up. The “proprioception” condition was identical in every respect to the “fake hand” condition, except that there was no fake hand present. Only the gloved left hand remained in view, and the position of the right hand (i.e., palm down and palm up) was again manipulated with the hand out of sight and resting on the right thigh. The order of right hand position (palm down or palm up) was counterbalanced within each of the three groups.

RTs were recorded. Overall, 11% of trials were deleted from the control condition, 8% were deleted from the fake hand condition, and 9% were deleted from the proprioception condition. Data from the left hand showed the anticipated effects of stimulus orientation, but no effect of experimental group. These data are not included here for space considerations.

An ANOVA was performed on the RT data in which group (control, fake hand, proprioception) served as a between-subject factor and subject hand position (palm down versus palm up), stimulus view (palm down versus palm up), and angular displacement (medial to lateral, $0^\circ$, $60^\circ$, $120^\circ$, $180^\circ$, $240^\circ$, and $300^\circ$) served as within-subject factors.

As illustrated in Figs. 1 and 2, main effects were observed for both stimulus view (palm down versus palm up; $F(1, 51) = 33.10, P < 0.0001$) and angular displacement, $F(5, 47) = 33.15, P < 0.0001$. In addition, there was an interaction between stimulus view and angular displacement, $F(5, 47) = 17.04, P < 0.0001$. Thus, faster RTs were observed in the palm down stimulus view condition at lateral orientations (i.e., $240^\circ$, $300^\circ$) than at medial orientations (i.e., $60^\circ$, $120^\circ$) and this difference approached significance, $t(53) = 1.68, P < 0.10$. In contrast, slower RTs were observed in the palm up stimulus view condition at lateral orientations than at medial orientations, $t(53) = 5.51, P < 0.0001$. These findings replicate previous reported observations by Parsons and co-workers [13,21] and suggest that subjects perform this task by rotating their own hand to match the stimulus hand. Further, the data...
suggest that performance is informed by biomechanical joint constraints which, for example, make it more difficult to move a palm up hand laterally than medially (see Fig. 2).

Second, and somewhat surprisingly, there was no main effect for group ($F(2, 51) = 0.58, P < 0.56$) (Figs. 3–5). In order to further examine the possibility of different patterns of performance across groups, difference scores for the critical manipulation of subject hand position were calculated for each subject (i.e., for the palm down stimulus view condition, we subtracted RTs for the subject palm down condition from the subject palm up condition). We then conducted a second ANOVA on these difference scores in order to examine whether the manipulations of visual input across the three groups resulted in different patterns of performance. Once again, there was no significant difference between groups for either the palm down stimulus view condition (the differences between palm up and palm down hand positions were 78 ms, 38 ms, and 63 ms for the control, fake hand, and proprioception conditions, respectively, $F(2, 51) = 0.23, P < 0.80$), or for the palm up stimulus view condition (99 ms, 116 ms, and 110 ms for the control, fake hand and proprioception conditions, $F(2, 51) = 0.02, P < 0.99$). Indeed, the effect size corresponding to these group differences was very small (eta = 0.069), suggesting that the visual manipulations across groups had little influence on the observed pattern of RTs.

Finally, we report analyses relevant to the influence of proprioception on RT performance. As shown in Fig. 6, a significant interaction between-subject hand position (palm down versus palm up) and stimulus view (palm down versus palm up) was noted ($F(1, 51) = 27.77, P < 0.0001$). For palm down stimuli, subjects were faster to respond when their right hand was palm down as compared to palm up, $t(53) = 2.53, P$...
subjects were faster to respond when their right hand was palm up as compared to palm down (1145 ms versus 1196 ms, t(35) = 2.74, P < 0.008). To more clearly assess the possible influence of proprioception on task performance, a similar analysis was performed in which only the data from the proprioception and fake hand conditions were included. This was done because the effects of vision and proprioception were confounded in the control condition. The results were consistent with the previous analysis, suggesting that the observed effect is due to proprioception alone. For palm down stimuli, subjects were faster to respond when their right hand was palm up as compared to palm down (1266 ms versus 1379 ms, t(35) = 2.19, P < 0.02). In the context of the above findings suggesting little influence of visual manipulations on performance, these findings suggest that mental rotations may be informed by proprioceptive information which specifies the position of the subjects’ hand. That is, proprioceptive information indicating the current position of the hand may be utilized in order to determine the starting point for the mental rotation that will align one’s hand with the orientation of the stimulus hand.

In our paradigm, hand position defined by proprioception, rather than vision, influenced the coding of hand position during mental motor imagery. Before considering the causes and implications of these data it is critical to note that this pattern of performance is not likely to be an artifact of specific experimental factors such as the visual appearance and location of the fake hand. One might argue that subjects did not incorporate the fake hand into the body schema because the stimulus lacked adequate verisimilitude or proximity to be processed as part of the body schema. This is unlikely for several reasons. Other investigations involving human subjects [1,17,18] have reliably reported fake limb attribution to the body using similar stimuli (fake hands, rubber gloves) and a similar proximity to the real hand (< 10 cm). Additionally, investigations with primates [9] have shown that the presence of a fake arm significantly increases neuronal activation in bimodal neurons located in parietal area 5 of the monkey.

regardless of its position relative to the felt position of the real arm, as long as the position was biologically plausible (i.e., palm up position while the real hand was palm down).

One possible explanation for the failure of the fake hand to be incorporated into the body schema is that our paradigm lacks an active cue that may be necessary to promote visual attention and thereby drive bimodal integration of the isolated afferent visual and proprioceptive signals. While “visual capture” (the attribution of a felt position to a seen position) using prisms [29] is a well-documented phenomenon, and studies using fake hands [1,17] have reliably replicated this phenomenon, in everyday life vision of a relevant body part will not always be available (i.e., when visual attention is directed other than actively toward the hands). This was the case in our paradigm: while the real left and fake right hand were placed in the visual field of the subject, they were simultaneously asked to complete a mental motor imagery task and to respond with their feet. One might suggest that structures analogous to those isolated by Graziano [8,9] (sites of converging visual, proprioceptive, and tactile inputs) require that attention—either in the sense of sensory processing or motor preparation—be directed to the hand to modulate the proprioceptive signal. In the absence of engaged visual attention or intention to act, a proprioception-dominated representation of body position in space may predominate. This suggests that proprioception may represent the baseline coding of body position in space, and that vision is recruited only on an as-needed basis.

Consistent with this hypothesis, Welch et al. [29] reported that “prism adaptation” (a gradual repositioning of proprioceptive information that persists after prism removal), was greatest when subjects moved the limbs over extended periods. When no limb movement was allowed, shifts of the hand’s felt position onto the prism-shifted visual position were transient and ceased as soon as the eyes were closed. Likewise, the “intersensory bias” favoring vision reported by Botvinick and Cohen [1] depended crucially on active bimodal stimulation. While the visual and proprioceptive input was kept static (as in our paradigm), the subjective attribution of the felt position to the seen position was achieved only with active attention to a congruent cue. Also of note is that while prisms alone have not been demonstrated to be of substantial benefit for subjects with neglect, prisms in conjunction with repetitive reaching to visual targets ameliorate neglect [20].

Mon-Williams et al. [12] provided precedent for the view that proprioceptive postural information may represent the baseline code of the body schema. In their study prisms of variable angular displacement were used in a fingertip-matching paradigm to assess the dominance of either vision or proprioception under either full cue or reduced cue conditions. Subjects were asked to match both the displaced seen position and then the felt position of one hand to their second, unseen hand underneath a thin table. When the experiment was performed under normally lighted background conditions, matching in the “felt” condition mapped closely onto the performance in the “seen” condition, indicating that
visual capture to the displaced, seen position had taken place. When the experiment was repeated in a darkened room with the scene position of the hand represented by an LED attached to the fingertip, the seen and felt matching performance dissociated: while matching in the “seen” condition reflected the proportional displacement of the prism (as it had in the full-cue condition), matching in the “felt” position was unaffected by prism displacement. Subjects also reported that the felt position seemed to appear displaced from the seen position, and that the fingertip seemed to “float” in space. Kinesthetic position was thus shown to dominate visual information as a function of the coherence and richness of the background environment, even when visual information of hand position was readily available.

Finally, the distinction between “extrinsic” and “intrins-ic” egocentric coordinate systems for motor planning may be relevant to our findings. We [22] and others [2,3,28] have argued for a distinction between movements made to a target in space as coded in an “extrinsic egocentric” coordinate system and movements that serve to alter the relationship between body parts as coded in an “intrins-ic” egocentric coordinate system. On this account, movements that require active movement toward a visual target (such as reaching to a fork) are likely to be coded in with respect to a visually based, “extrinsic” coordinate system. In contrast, rotating the hand to achieve a desired position may involve a proprioception-based, “intrin-sic” coordinate system. Consistent with this view, Rossetti et al. [19] showed a visual modulation on the coding of body position was readily available.

References