Extrinsic Cues Suppress the Encoding of Intrinsic Cues

Bhavin R. Sheth[1] and Shinsuke Shimojo[1,2]

Abstract

■ Remembering where objects are in space is fundamental to adaptive behavior. Little is known about how intact humans combine information from intrinsic (egocentric) and extrinsic (exocentric, allocentric, or landmark-based) coordinate systems to locate objects. Using a simple location estimation paradigm, this study finds that we mostly remember position in extrinsic coordinates. Intrinsic-coordinate-based mapping of space is less precise in the presence of landmarks or extrinsic cues than in their absence. Thus, not only do extrinsic frames of reference dominate internal representations of space, they suppress intrinsic-based representations as well. We speculate that this dominance–suppression hierarchy undercuts intersystem conflicts and underlies a single, undissociated spatial map in intact humans. ■

INTRODUCTION

Spatial information can be coded in one of two coordinate systems—extrinsic or intrinsic. In an extrinsic coordinate system (also called allocentric, exocentric, or object based), the location of a target is coded with respect to the locations of external objects in the scene. In an intrinsic coordinate system (also termed egocentric), target location is coded with respect to internal, egocentric cues. Because the cues are different, each coordinate system furnishes a slightly different view of the world. At issue is how these systems interact to code spatial information. Although there have been considerable psychophysical (Vetter, Goodbody, & Wolpert, 1999; Henriques, Klier, Smith, Lowy, & Crawford, 1998; McIntyre, Stratta, & Lacquanti, 1997; Enright, 1995; Flanders, Tillery, & Soechting, 1992), computational (Salinas & Abbott, 1995; Zipser & Andersen, 1988) and physiological (Batista, Buneo, Snyder, & Andersen, 1999; Clower et al., 1996) insights into the nature of the sensorimotor coordinate transformation and the coordinate system of the motor plan, little is known about how the two coordinate systems are combined to map position in humans.

One possibility is that observers linearly weigh the two coordinate systems (Pelz & Hayhoe, 1995; Matin & Fox, 1989), with the proviso that the more complex the scene, the larger is the reliance on the extrinsic system (Bridgeman & Graziano, 1989). It is likely though that there are factors other than scene complexity that affect how the two systems interact. Relative precision is one potential factor. The more precise a system is in mapping position, the more on will rely on it (Landy, Maloney, Johnston, & Young, 1995). Cue stability is another potential factor, as cues may shift position (which could occur if the cues themselves are moving, or are stationary but lie at different distances from a locomoting observer), making it nonoptimal for one to rely on cue locations while the target position is being internally recorded. Working memory and labor are also likely to be potential factors; the more memory- and labor-efficient system may well dominate the interaction. All of the aforementioned factors could interact. For example, the more precise system may be less stable or may utilize more memory. Thus, the interaction and integration of coordinate systems is not straightforward. One simple solution is to minimize the integration: Favor the extrinsic system and rely heavily upon it (Matin et al., 1982), but fall back on the intrinsic system when the former is unavailable.

Here, we demonstrate that when visible objects or landmarks are present in the visual scene, intact humans typically rely heavily on an extrinsic object-based system to record spatial information (see Matin et al., 1982, for the first demonstration of this). More important, we are the first to show that this reliance on extrinsic coordinates precludes a precise intrinsic-based representation. Relatedly, if no extrinsic cues or objects are present at the time of encoding, we observe precise intrinsic-based representation of space. Our data demonstrate a hierarchy in which an extrinsic-based spatial representation dominates and *suppresses* the formation and maintenance of an intrinsic-based representation. This hierarchical scheme precludes intersystem conflicts, which, in turn, precludes the existence of multiple dissociated spatial maps in intact humans.

RESULTS

Utilizing Extrinsic Cues: Does Their Presence Improve the Precision of Position Estimates?

In Experiment I, the task was to point, via a mouse cursor, to the remembered target location after some

[1]California Institute of Technology, [2]NTT Communication Science Laboratories

time (stimulus response interval, or SRI; also referred to as the setting period). On trials of the "with extrinsic cues" (EC+) condition, two clearly visible objects or landmarks were present over the entire duration of the trial (Figure 1B). Because internal proprioceptive cues are always available, both intrinsic and extrinsic systems contribute in retrieving the target from memory on EC+ trials.[1] On trials of the "without extrinsic cues" (EC−) condition, which were randomly interleaved with those of the EC+ condition, no objects other than the target were ever present (Figure 1A). It is reasonable to assume that target position is estimated less on the basis of extrinsic and somewhat more on the basis of intrinsic cues in the EC− condition in comparison with the EC+ condition. Finally, we varied the time between stimulus offset and the earliest possible onset of the pointing movement (SRI) to see whether response delay differentially affects localization on the EC+ and EC− conditions.

Extrinsic and intrinsic cues do not conflict on the EC+ trials, so potentially, one could integrate information from the extrinsic and intrinsic systems and reduce localization error in the process. Also, observers ($n = 4$) were free to move their eyes on the EC− trials, but were required to fixate the crosshair on the EC+ trials. Thus, whereas eye position (an intrinsic cue) was stable, hence reliable, on the EC+ trials, it was not necessarily so on the EC− trials. The above arguments imply that localization should be more precise on the EC+ trials than on the EC− trials. Figure 1C shows the group mean root mean square (rms) error as a function of SRI. Contrary to the prediction, target localization was just as precise on the EC− trials as on the EC+ trials. A repeated measures ANOVA revealed no significant main effect of condition, $F(1,27) = 0.61$, $MSE = 13.56$, SRI, $F(4,27) = 0.58$, $MSE = 12.97$, or their interaction, $F(4,27) = 1.07$, $MSE = 23.93$, $p = .39$.

One may contend that screen edges served as extrinsic references in both conditions, which might explain the lack of difference in precision. However, the room was dark, the screen edges were hardly visible, and none of our observers claimed to have consciously noticed them. Albeit, with dark adaptation, edges might become more visible, which by this reasoning would improve precision. The rapid phase of dark adaptation takes about 3 min (Hecht, Haig, & Chase, 1937) and, in that time, our observers ran about 10 trials. We compared (absolute) error magnitudes on these preadaptation trials with those on the remaining trials and found no difference (group means: pre: 0.50°, post: 0.52°, $p = .69$,

two-tailed unpaired t test). In addition, 95% of single-trial preadaptation error magnitudes were within the 95% confidence limits of the distribution of postadaptation error magnitudes. Thus, even if dark adaptation were to have made the screen edges more visible, it did not discernibly change observers' estimates of target position. Moreover, the screen edges were at least 11 cm from the target, a distance so large as to be ineffectual in shifting remembered target locations—even in room illumination (Sheth & Shimojo, 2001). Finally, both removing explicit extrafoveal extrinsic cues (Figure 2) and adding them in the setting period (Figure 3) significantly affected localization, indirectly arguing for the ineffectiveness of the screen edges as an extrafoveal extrinsic cue.[2,3] In sum, the data indicate that the lack of improved precision on the EC+ trials as compared to the EC− trials results, at least in part, from a failure of extrinsic and intrinsic representations to optimally integrate.

Relying on Extrinsic Cues

Locations of the crosshair and the line varied from trial to trial of the EC+ condition, as did the distance between them and the target. Therefore, the extrinsic cues were not reliable from trial to trial of our task, which may have led observers to ignore them altogether, and thereby explain why the added presence of the extrinsic cues failed to improve localization. One way of testing this possibility is to abruptly extinguish the landmarks after target offset and see whether localization is affected. If landmarks are ignored, localization would be unaffected.

In Experiment 2, the same four observers participated as in Experiment 1. On trials of the "extrinsic cues remove" condition (EC+/−, Figure 2A), landmarks (the same as in Experiment 1) were present while the target was displayed, but were extinguished in the setting period before the cursor appeared. On trials of the EC+ condition (EC+, Figure 2B), landmarks were present throughout. Contrary to the above prediction, localization was significantly less precise if the landmarks were extinguished (EC+/−, Figure 2C, $F(1,27) = 4.78$, $MSE = 263.61$, $p = .04$)[4], suggesting that target position was encoded, at least in part, with respect to landmark locations.[5]

The data demonstrate more than just that. Although EC+ trials in the two experiments were identical, localization was significantly more precise, $F(1,27) = 8.03$, $MSE = 333.05$, $p < .01$, on the EC+ trials of Experiment 1 (Figure 1C, solid lines) than on the EC+ trials of

Figure 1. The effect of extrinsic cues on position estimates. (A) Sequence of screen displays (left) and timeline (right) on a trial of the EC− condition of Experiment 1 are given. As noted in the text, SRI is the time interval between mask offset and the reappearance of the cursor in the setting period. (B) Sequence of screen displays (left) and timeline (right) on a trial of the EC+ condition are shown. The crosshair and L are simultaneously presented and remain visible throughout target presentation and during the setting period. (C) Group mean rms error (ordinate) versus SRI (abscissa) on the EC− (filled squares, dotted lines) and EC+ conditions (open circles, solid lines) is plotted. Error bars represent one between-subjects ($n = 4$) standard error.
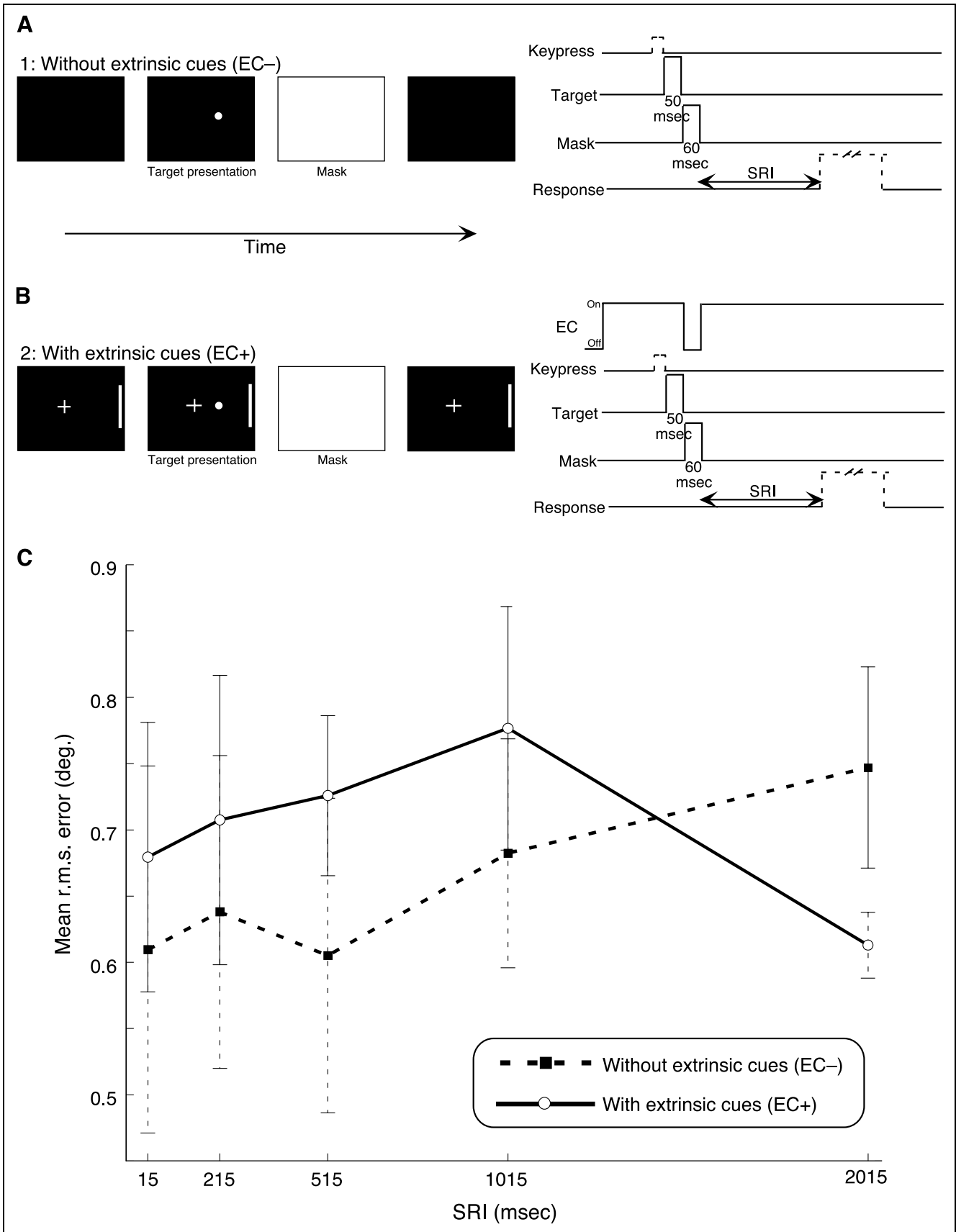
Experiment 2 (Figure 2C, solid lines).[6] The difference in error is surprising because extrinsic and intrinsic representations of space presumably integrate similarly on all EC+ trials. However, the trials with which the EC+ trials were intermixed in the two experiments were different, which might, in turn, have affected the relative reliability of the landmarks on the EC+ trials in the two experiments. If the landmarks were present during target presentation on Experiment 1 trials (EC+), they were also present in the setting period as well. In Experiment 2, in contrast, the landmarks were present during target presentation on all trials (EC+ and EC+/−), but were removed (with no advance warning) prior to localization on half of the trials (EC+/−). Thus, the comparatively lower reliability of the landmarks in Experiment 2 is likely to account for the larger error on Experiment 2 EC+ trials. That the decreased reliability of the landmarks was not offset by an enhanced reliance on the intrinsic representation (which is, under some circumstances, as precise, see Figures 1C and 4) indicates that extrinsic cues are not just used but relied upon to encode spatial information, and relied upon to the *exclusion* of intrinsic cues.

Experiment 3 provides further and more direct evidence that observers rely heavily on landmarks to encode space. On "EC move" trials, the landmarks physically shifted position after target presentation (range [1°, 2°] left or right of the original locations, mean shift = +1.5°) but prior to the setting period (Figure 3). As compared to estimates of the exact same target positions on "EC fixed" trials, estimates of target position on the EC move trials were shifted *in the same direction* as the shift in landmarks (mean shift in estimate = +1.02°, 68% of true mean shift magnitude). Position estimates were biased in the direction of cue shift on an overwhelming 75% of EC move trials ($p < .0001$, sign test). Thus, remembered target position was yoked to the locations of the landmarks, strongly suggesting that observers ($n = 7$) relied heavily on the landmarks to encode target position. In accord with this, rms error magnitudes were larger when landmarks shifted during the trial (mean = 1.92°) than when they did not (mean = 1.40°, $t(6) = 4.35, p < .005$).[7]

Consequences of Relying on Extrinsic Cues: Extrinsic Cues Suppress the Encoding of Intrinsic Cues

Arguably, the most telling signature of the reliance on the extrinsic representation of space is the detrimental effect it has on the intrinsic representation. On EC− trials of Experiment 1 and on EC+/− trials of Experiment 2, no landmarks were present during localization; hence, target position had to be retrieved largely from intrinsic cue-based memory. Landmarks were present on the EC+/− trials while the target was shown (and observers relied on them to encode target position). On EC− trials in contrast, no extrinsic cues were ever present, hence a representation based more on internal, egocentric cues was used to encode position. By comparing position estimates on EC− and EC+/− trials, we can address whether the reliance on extrinsic coordinates while encoding spatial information of the target has any consequences on encoding the same in intrinsic coordinates.

To normalize the data, we subtracted the error on the respective EC+ trials from the error on each condition.[8] If the fidelity of the intrinsic coordinate-based memory of target position was independent of whether or not target position was encoded in extrinsic coordinate-based memory, normalized error magnitudes on the EC− and EC+/− trials should be the same. However, observers were significantly less precise in retrieving target position from intrinsic coordinate-based memory if landmarks were present during target presentation (EC+ trials) than if they were absent (EC−trials) (Figure 4; $F(1,27) = 5.61, MSE = 421.28, p = .03$).[9] Thus, in representing target position on the basis of extrinsic cues, target position was encoded less precisely in intrinsic coordinates than was otherwise possible. Removal of extrinsic cues during retrieval exposed the fact that their presence caused intrinsic coordinate-based position information to not be recorded at all or to be recorded imprecisely. The *extrinsic coordinate*-based *degradation of intrinsic coordinate*-based *encoding* again indicates the extent to which space is nearly exclusively encoded in extrinsic coordinates in the intact brain.[8]

On EC− trials, observers could potentially fixate the target (although they were not required to). On EC+/− trials, observers fixated the crosshair at the time of target presentation, so the target was peripheral to the center of gaze. Arguably, localization was worse on the EC+/− trials compared with the EC− trials because pointing to a memorized peripheral point is less precise than to a locus on which gaze is fixated. We probed the role of fixation on extrinsic cue-dependent suppression of intrinsic cue-based mapping further. We let observers ($n = 3$) freely move their eyes throughout the trial on all trials of a separate pair of runs otherwise identical to Experiments 1 and 2. If the above argument was the underlying cause of the suppression of intrinsic cued based mapping, normalized errors on the EC− and EC+/− trials in the new set of runs should be similar. Contrary to the argument, observers continued to be significantly less precise in pointing to the remembered target location on EC+/− trials as compared to EC− trials (Figure 4, inset; $F(1,18) = 26.72, MSE = 463.34, p < .0001$).[10] Therefore, the sheer presence of extrinsic cues suppresses spatial information in intrinsic coordinates from being encoded and later maintained in memory. In sum, if reliable extrinsic cues are present during encoding, intact humans rely heavily on them to the exclusion of intrinsic cues in order to represent spatial information. Only in the absence of visible
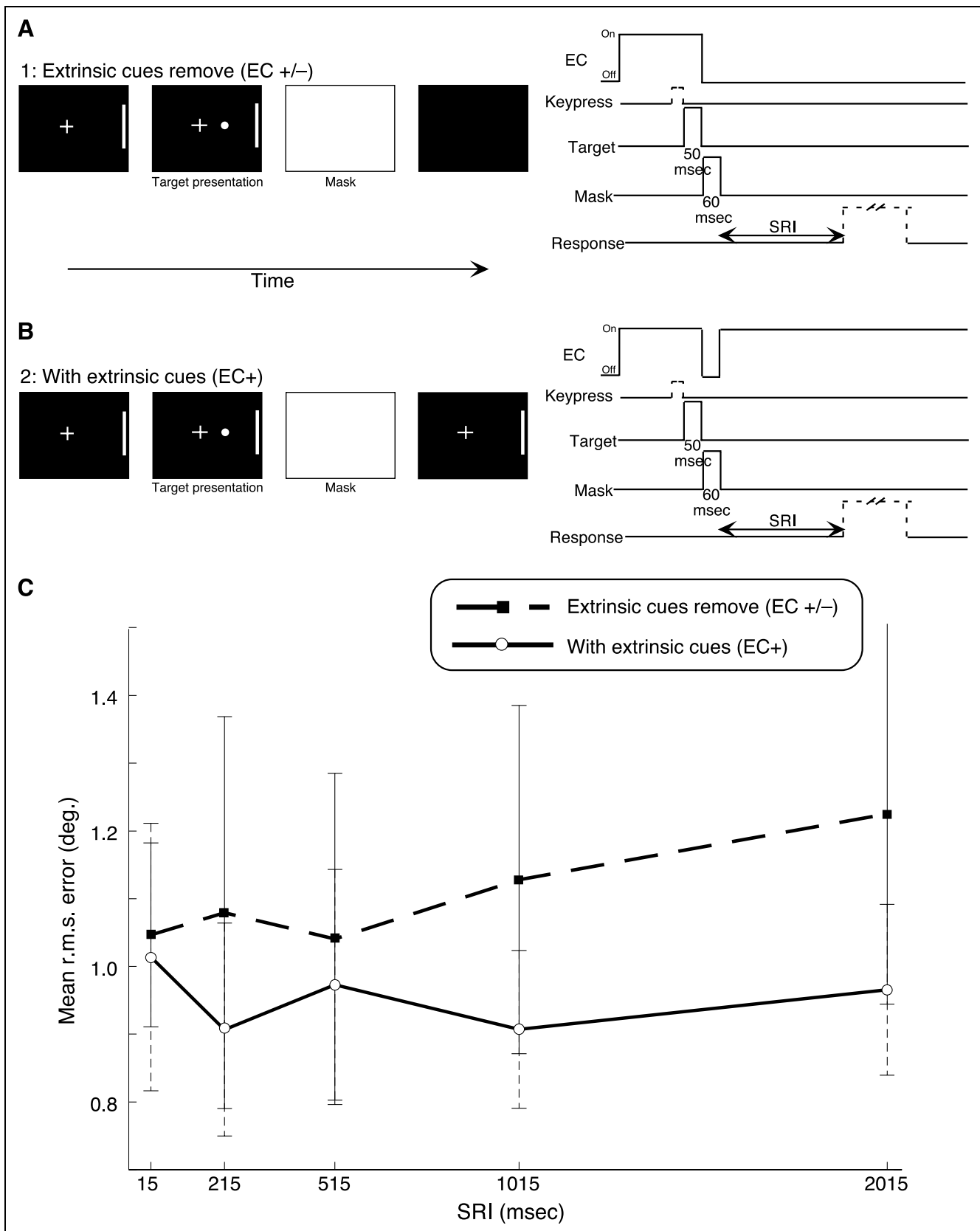
Figure 2. The effect on position estimates of removing extrinsic cues in the setting period. (A) Sequence of screen displays (left) and timeline (right) on a trial of the EC+/− condition of Experiment 2 are given. (B) Sequence of screen displays (left) and timeline (right) on a trial of the EC+ condition are shown. Stimulus sequence and timeline were identical to EC+ trials in Experiment 1 (see Figure 1B). (C) Group mean rms error versus SRI on the EC+/− (filled squares, dashed lines) and EC+ conditions (open circles, solid lines) is plotted. Error bars represent one *SEM*. The same observers participated as in Experiment 1.

extrinsic cues do humans fall back more heavily on intrinsic coordinates.

DISCUSSION

Here, we explored how spatial information is stored in extrinsic (exocentric) and intrinsic (egocentric) referen-ces, and how the two representations interact. Extin-guishing (Figure 2) or displacing (Figure 3) extrinsic references prior to response profoundly impaired esti-mation of target position, demonstrating that extrinsic references were important in coding spatial informa-tion. Moreover, estimates were displaced in the same direction as the cues, and by nearly as large a magnitude
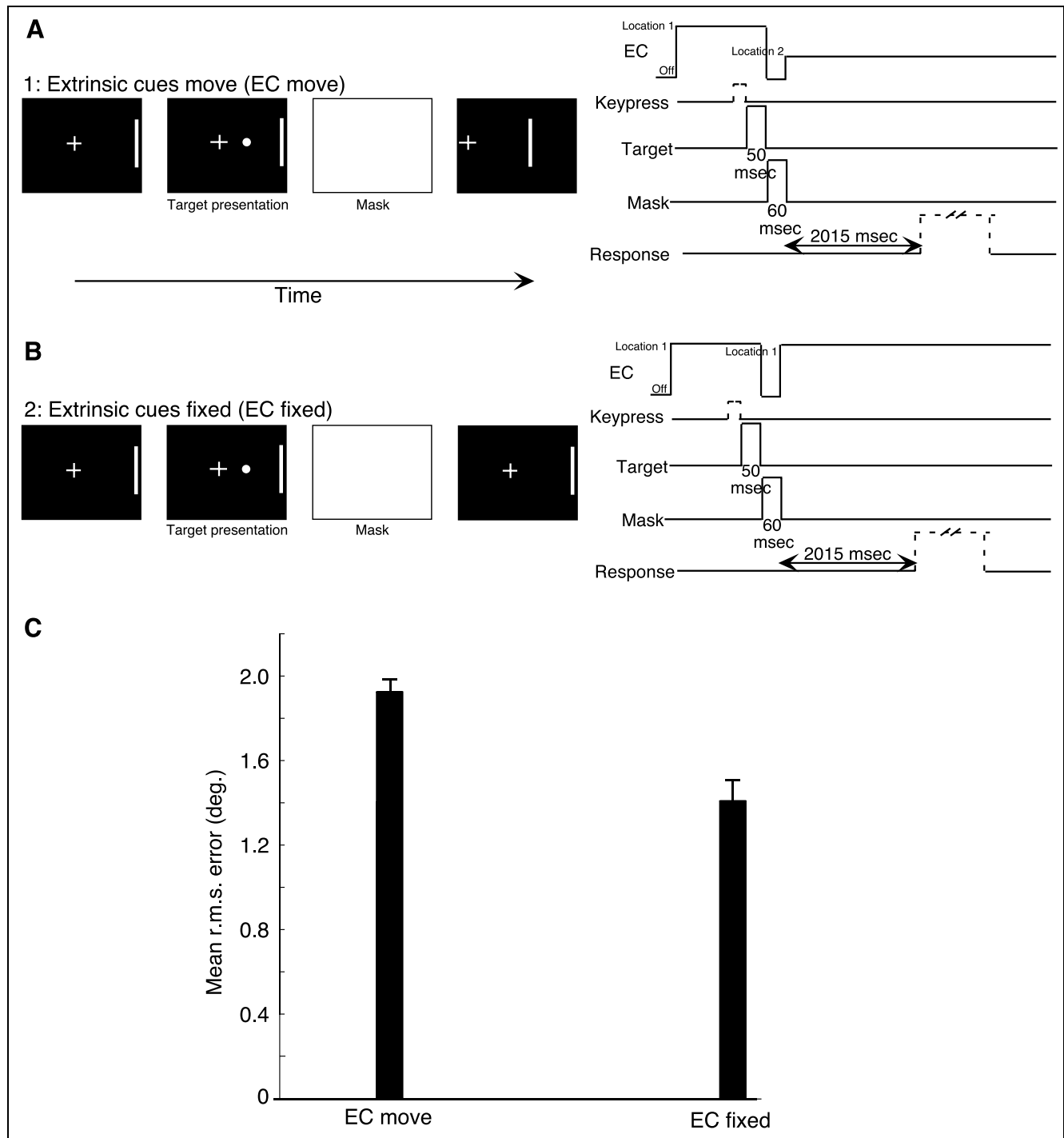

Figure 3. The effect of displacing extrinsic cues in the setting period. (A) Sequence of screen displays (left) and timeline (right) on a trial of the EC "move" condition of Experiment 3 are given. The SRI was 2015 msec. (B) Sequence of screen displays (left) and timeline (right) on a trial of the EC "fixed" condition are shown. (C) Bar graph comparing group mean rms errors on the two conditions.
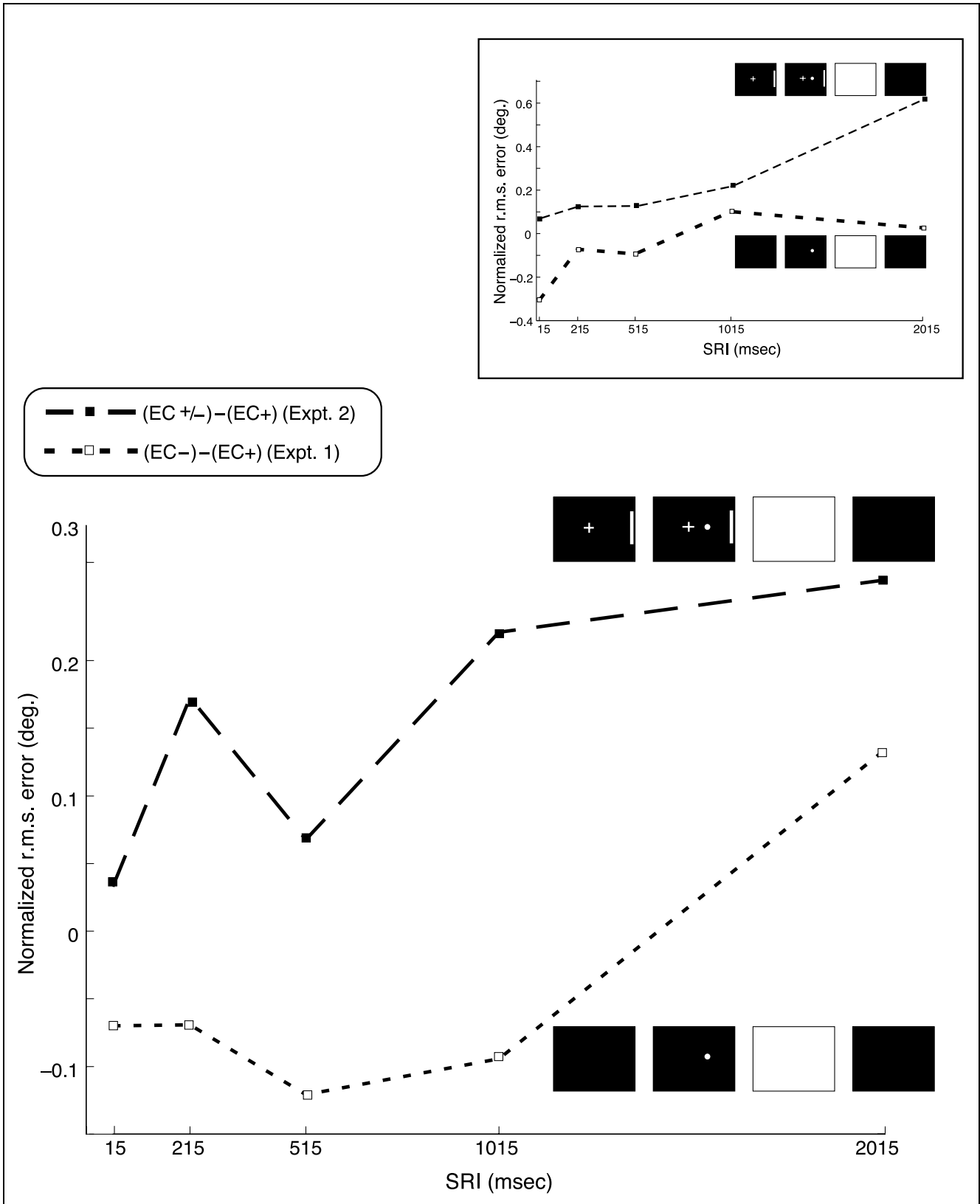
Figure 4. The effect of extrinsic cues on the encoding of space in intrinsic coordinates. Normalized (see text) group mean rms error versus SRI on the EC+/− (filled squares, dashed lines) and EC− conditions (open squares, dotted lines) is plotted. The inset shows normalized data from a similar pair of runs as Experiments 1 and 2, with the exception that, on all trials, observers were free to move their eyes anywhere on the screen.

(Experiment 3), further demonstrating the reliance on the extrinsic system to record position (see also Pelz & Hayhoe, 1995; Schlag & Schlag-Rey, 1995; Matin et al., 1982). The representation in intrinsic coordinates did not offset this loss in precision, signifying the extent to which extrinsic cues dominate and how this dominance corrupts intrinsic cue-based encoding. The latter (and most critical) point was demonstrated directly by demonstrating that position information was encoded far less precisely in intrinsic coordinates if extrinsic cues were present during encoding than if they were absent (Figure 4). Theoretically, position should be encoded separately in each coordinate system; the disappearance of extrinsic cues prior to retrieval and a concomitant inability to use the extrinsic system should not affect the precision of the position information encoded in the intrinsic system. That this was not the case (Figure 4) argues rather that factors other than precision may also help determine how information is encoded.

Next, we will discuss some issues concerning our experimental choices and their limitations. Finally, we will discuss the study's implications.

Varying the SRI

Varying the time between stimulus offset and the earliest possible onset of the pointing movement (SRI) had no major effect on our results. Neither the effect of SRI alone nor its interaction with trial type was significant in any of our experiments. That the eye and body can remain stable if desired over short time intervals (1 sec or less), so intrinsic cues may not be necessarily worse than stable extrinsic cues over these short delays. However, over longer time intervals, random body twitches, drifts in eye position, or overt movements of either body or eye do occur, all of which enhance noise in the intrinsic system. Nonetheless, comparisons of intrinsic and extrinsic systems over longer time intervals than the ones studied do not seem very relevant, as the high throughput rates of sensory data in real life mean that decisions about where to point and where to move are usually made quickly (1 sec or less).

Modes of Motor Response: Mouse Cursor versus Finger

In our experiments, the observers' eyes viewed a vertical screen while their hand rested in a horizontal plane on top of a mousepad. Although adapting to this visuomotor transformation is relatively easy (Vetter et al., 1999), using a mouse-guided pointer is not the norm. In contrast, using a finger as a pointer is contrast. However, there are limitations to using a finger. A visible finger is a powerful, familiar, three-dimensional external frame of reference that would compromise trials (e.g., EC– trial) where extrinsic cues are presumed absent. An invisible finger, on the other hand, might lead to noisier esti-

mates, which, in turn, might blur out any potential differences between the different conditions. A second limitation is that it would be difficult to obtain the same precision with finger because the finger is an object more than one degree wide. Thus, while it would be preferable to while confirm with a finger in the future, we tentatively claim that our findings are not limited by the use of a cursor.

In general, compliant movements (defined as a movement constrained by external contact, such as those using a hand-held cursor, mouse cursor, or manipulandum) involve a different control strategy than unconstrained movements, and the relative contribution of proprioception to reach control is different for constrained and unconstrained movements (Desmurget, Jordan, Prablanc, & Jeannerod, 1997). This has profound implications for generalizing theories of motor control based on compliant movements to natural, unconstrained movements. However, differences between constrained and unconstrained movements, i.e., factors relevant at the motor level, are tangential to the perception and memory of external object position—the focus of the present study. In general, issues pertaining to arm movement control, such as the interaction of vision and proprioception in estimating limb position, or the role of feedforward motor commands, are orthogonal to the issues addressed here, which pertain to spatial perception and memory.

Why Rely on an Extrinsic Coordinate System: Implications for Computation

There are intriguing implications of our study for theoretical models of space representation. Typical cue integration schemes assume that different cues are mutually independent (e.g., Andy et al., 1995), which is not true here, as we saw that the presence of extrinsic cues degraded the information coded in the intrinsic system (Figure 4). Some other cue integration scheme is needed to better explain our findings.

The heavier reliance on the extrinsic system, even though, in our hands, it was no more precise than the intrinsic system (Figure 1) implies that factors other than precision also contribute towards this reliance. An a priori belief that an extrinsic system is precise over longer periods than an intrinsic system could be a factor for why humans rely on the extrinsic system. Another possibility is that the extrinsic system is more efficient on memory and labor than an intrinsic system. Unlike external objects, the eyes, head, and body often change position over short periods of time. A viable intrinsic coordinate system has to either keep multiple records of past body and eye positions from target encoding to the present time, which takes memory, or continually update the records, which is laborious. In the extrinsic coordinate system, in contrast, the position of a target is obtained from its distances from different external

markers or objects. Assuming that the positions of objects in the world do not usually change over short periods of time, there is no need to continually update target–cue distances, or keep multiple records of past target–cue distances in memory. Under the assumption of a stable world, the world is the memory store (Ballard, Hayhoe, Li, & Whitehead, 1992; O'Regan, 1992; O'Regan et al., 1999). That the memory of target position shifts in lockstep with shifts in marker locations and, therefore, is yoked to the locations of these markers in the world (Experiment 3) accords with the assumption of a stable world with stable objects. Thus, more economical memory (and labor) use could be one a priori factor of many why intact humans rely on the extrinsic system.

Implications for Neurophysiology and Neuropsychology

The idea of a latent intrinsic coordinate system has implications for physiology. Neurons in areas LIP (Snyder, Grieve, Brotchie, & Andersen, 1998), VIP (Duhamel, Bremer, BenHamed, & Graf, 1997), 7a (Snyder et al., 1998) and the superior colliculus (Stuphorn, Bauswein, & Hoffman, 2000) of the monkey brain encode position in body-part referenced, object-independent coordinates. Recordings in these studies have been performed in highly impoverished visual environments lacking any visible landmarks. In light of this, neurons in the aforementioned brain areas may represent the unmasking (or development through overtraining) of a latent capacity in primates to map space in egocentric coordinates in the absence of a landmark-based map of space. It would be interesting to compare how the responses of these neurons if a landmark-based map of space were present (Dassonville, Schlag, & Schlag-Rey, 1995, propose that neurophysiologic studies of saccadic eye movement and localization may have ignored the possibility that the brain may be using exocentric cues to calculate the dimensions of the saccade).

The idea of a latent intrinsic coordinate system applies in the case of patient DF who was unable to discriminate between objects on the basis of shape, yet was able to make skilled movements that varied as a function of the object's shape (Goodale, Milner, Jakobson, & Carey, 1991; Goodale et al., 1994). It was found that the early part of DF's reaching movement was no better than her verbal performance (Murphy, Racicot, & Goodale, 1996). However, by using transient, nonvisual, proprioceptive information about the configuration of her hand and fingers (Murphy et al., 1996, themselves cite this as a possibility), DF was able to redirect her reaches *in midflight* to the correct target. Thus, DF's arguably intact visuomotor skills can be explained by the fact that, lacking unambiguous extrinsic, visual cues, she learned over the years since her brain injury to also use intrinsic cues to help represent space (Tillery, Flanders, & Soechting, 1991)—cues that go largely unused in intact subjects. Despite large differences in our tasks and theirs, our idea of a dominant visual, extrinsic cue-based system and unmasking of a latent nonvisual, intrinsic system under visually impoverish conditions may provide the glimmer of an explanation for DF's intact visuomotor performance.

To conclude, intact humans represent space primarily in cost-effective extrinsic coordinates, which are usually reliable and more precise in the long term. In doing so, we suppress representations mapped in intrinsic, egocentric spatial coordinates. However, if extrinsic cues are absent or unseen while encoding space, intact humans fall back on intrinsic coordinates that can sometimes be as precise as extrinsic coordinates in the short term. Thus, there is a partial hierarchy of coordinate systems in the representation of spatial information in the human brain.

METHODS

All stimuli were presented on a Sony Trinitron monitor (75 Hz refresh; 37.5 × 30-cm viewing area) under control of a MAC PowerPC running MATLAB (Mathworks Inc.) and Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The observer sat comfortably in a chair 57 cm in front of the screen, with the head immobilized in a chin rest (Handaya Co., Japan). Viewing was binocular. Room lights were turned off (room illumination < .01 cd/m^2). A computer mouse rested on a horizontal pad placed on the table about halfway between the screen and the observer. The cursor driven by the mouse was a small ''+'' on the screen. The cursor crosshair had a resolution of 3.5 arcmin. In the response stage, the observer dragged the cursor to the remembered location of the target. A mouse button press confirmed the observer's choice. The coordinates of the cursor at the time of the button press and those of the true target location on each trial were recorded. There was little dark adaptation prior to an experimental run (<< 1 min).

Experiment 1

Trials of the following two conditions were randomly intermixed. (1) Without extrinsic cues (EC−; Figure 1A): A blank screen appeared and the observer had to press a key to continue the trial. The target (0.4° diameter, 10.6 cd/m^2, its horizontal coordinate was randomly chosen from a uniform distribution of 0° to 8°, as measured from the center of the screen) was shown for a brief period (50-msec duration). The target was always presented at eye level, and was 15° on average from the screen edges. Following target offset, the entire screen flashed on (60-msec duration) and then off . At a variable time (SRI: 15, 215, 515, 1015, or 2015 msec) after the screen flash or mask, the cursor appeared at the bottom center of the screen and the observer had to use it to

estimate the target position with it (response stage). (2) With extrinsic cues (EC+; Figure 1B): Target parameters were the same as in the EC− condition. Two landmarks were present from the trial's outset. A crosshair ("+") served as a fixation point (FP). Each of its two orthogonal arms was 0.3° long. The intersection of the arms was 1.8 arcmin in diameter. The luminance of each pixel of the crosshair was 2 cd/m². The horizontal coordinate of the crosshair was randomly chosen from a uniform probability distribution of range 0° to 4° from the center of the screen, with the provision that it had to be at least 0.25° from the target. The vertical line ("L," luminance: 53.3 cd/m²) was 2.0° long, 0.09° wide, and was displayed at eye level. Its horizontal coordinate was from a uniform probability distribution of range 0° to 4° from the center of the screen, with the provision that it had to be at least 0.25° from both the target and the crosshair. The crosshair and L could be on the same side of the target or on opposite sides. After achieving stable fixation, the observer ($n = 4$) pressed a key to continue. As in the previous condition, the target appeared for 50 msec, followed by the 60-msec-long mask. After the mask, the crosshair and L reappeared in the same respective locations as before, and the observer had to fixate the crosshair in the localization setting period. Target location was chosen from the same set of 30 randomly chosen locations over all SRI (5) and condition (2) pairs ($10 = 5 \times 2$). The locations were ordered differently for each SRI and condition pair in a random manner, and then randomly interleaved in the task. The task typically took around 110 min altogether and was divided into two sessions, each consisting of 150 trials (300 trials total per observer).

Experiment 2

Trials of the following two conditions were randomly intermixed. (1) Extrinsic cues remove, or EC+/−; Figure 2A): The crosshair and L appeared at the beginning of the trial. The observer pressed a key after fixating on the crosshair. The target was then shown for 50 msec. Stimulus parameters defining the target, crosshair, and L, their range of locations, and so on were the same as those in Experiment 1. After target offset, the screen flashed (60 msec). After the screen flash, the landmarks did not reappear. The observer was free to move her eyes during the setting period. The cursor appeared on a blank screen, and the observer had to estimate the target position as before. (2) EC+ (Figure 2B): Identical to EC+ trials of Experiment 1. There were 300 trials total ($30 \times 5 \times 2$) for each observer (same four observers as in Experiment 1).

Experiment 3

Stimulus parameters defining the target, crosshair, and vertical line (L) were identical to those in previous experiments. (1) Extrinsic cues move (EC move; Figure 3A): Following the screen flash or mask, the crosshair and L were redisplayed at a new pair of locations 1° to 2° from their locations before the mask (mean displacement: 1.5°). The observer had to refixate the new crosshair location. The cursor appeared 2015 msec after the mask and the reappearance of the landmarks in the new locations, and the observer could then point to the location of the extinguished target with the cursor. (2) Extrinsic and intrinsic cues (EC+; Figure 3B): The condition was identical to similar conditions in Experiments 1 and 2, except that the SRI was a constant 2015 msec in this experiment. There were 60 (30×2) randomly interleaved trials for each observer ($n = 7$).

Analysis: Precision Measure (rms Error)

The target and landmarks were always displayed at eye level. Only the horizontal coordinates of the true and estimated target locations were used in the analysis. On each trial, we subtracted the true horizontal location of the target from the observer's estimate and squared this value. Single-trial errors thus computed were summed across all trials of the same condition, then divided by the total number of trials of the given condition, and finally the square root was taken of this mean error. This calculation was done separately for each individual observer, and the group mean rms error shown is the group mean of all the individual rms errors.

Notes

1. There was a brief (60 msec) interval between target offset and the setting period, during which a blank, intensely bright field was flashed. The onset of the blank field swamps the ongoing signal and serves as a mask that destroyed the retinal afterimage of the target (O'Regan, Rensink, & Clark, 1999). In a control experiment, we compared the performance of seven observers with and without the mask in the intervening 60-msec period, and found that the mask changed error magnitudes by nearly identical amounts on both EC+ and EC− conditions, $t(6) < 1, p > .65$.

2. In other experiments (two controls, cf. footnote 3), we did find a statistically significant difference in localization error between the EC− and EC+ conditions. Therefore, screen edges, even if visible, were ineffective as an extrafoveal extrinsic frame of reference in the EC− condition.

3. In two separate experiments, we varied the initial hand position (center or 15 cm right of body midline), and initial cursor position (bottom center of screen or 10° right) across trials. Neither initial hand nor cursor position (three observers each) had any effect on the precision of target position estimates (F values < 1). See Vindras, Desmurget, Prablanc,

and Viviani (1998) and Desmurget, Rossetti, Jordan, Meckler, and Prablanc (1997) for studies on the influence of hand position on reaching accuracy.

4. The main effect of SRI, $F(4,27) = 0.41$, $MSE = 22.72$, and the interaction between SRI and condition, $F(4,27) = 0.67$, $MSE = 36.88$, were not significant.

5. There was no consistent pattern of error bias toward the landmarks in the EC+ trial data of Experiments 1 and 2 as compared with trials of the interleaved condition. A case in point, estimates of two of four observers on the EC+ trials in Experiment 1 were biased away from the two landmarks, estimates of one observer were biased toward them, and the last showed no tendency in either direction (see Kerzel, 2002; Sheth & Shimojo, 2001; Hubbard & Ruppel, 1999; Musseler, Heijden, Mahmud, Deubel & Ertsey, 1999; Mateeff & Gourevich, 1983, for further discussions on this issue).

6. The cursor, despite being a screen marker, could not have biased the task toward an external encoding strategy as it appeared after the target location was encoded (c.f footnote 8). (A supposedly nonscreen marker, e.g., a finger, is visible close to the screen, and is a screen markers in effect, therefore; Moreover, if the presence of the cursor as an external marker were solely responsible for the bias in strategy, the presence of explicit external marker should not have altered the bias at all. This was not the case, however (see Figures 2 and 4). Thus, the reliance on an extrinsic encoding strategy is a property of the observer, and is genuine.

7. Our motivation though is to study what happens if observers attend to, and not ignore the landmarks. This is the rationale for the indirect design: Intermixing each condition with a condition in which landmarks remain after target offset (EC+ trials) forces observers from ignoring the landmarks. Alternately, one can intermix trials of EC−, EC+/−, and EC+ conditions in the same session. We did so in a separate experiment (2-sec SRI), and found that observers ($n = 5$) were significantly less precise ($p < .05$, paired) on the EC+/− trials than on the interleaved EC− trials, consistent with the results shown in Figure 4.

8. Extrinsic cues degrade intrinsic-cue-based encoding, but once encoded in intrinsic coordinates, the sudden appearance of extrinsic cues or landmarks in the setting period has no effect whatsoever on localization. One experimental condition was EC−. In the other condition, landmarks first appeared in the setting period after target presentation, and stayed on until the observer's ($n = 4$) response. There was no difference ($p > .5$) in estimates of target position between the two conditions.

9. There was no significant effect of SRI, $F(4,27) = 1.25$, $MSE = 94.12$, $p = .31$, and no significant interaction between the two variables, $F(4,27) = 0.30$, $MSE = 22.23$.

10. There was a significant effect of SRI, $F(4,18) = 4.20$, $MSE = 72.76$, $p < .025$, owing to the general trend toward larger error with longer SRI (Figure 4, inset), and a significant SRI × Condition interaction, $F(4,18) = 4.94$, $MSE = 85.68$, $p < .01$, owing to the monotonic increase in error with SRI on the EC+/− trials and an asymptotic increase in error with SRI on the EC− trials.

REFERENCES

Ballard, D. H., Hayhoe, M. M., Li, F., & Whitehead, S. D. (1992). Hand-eye coordination during sequential tasks. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 29,* 331–338.

Batista, A. P., Buneo, C. A., Snyder, L. H., & Andersen, R. A. (1999). Reach plans in eye-centered coordinates. *Science, 285,* 257–260.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 443–446.

Bridgeman, B., & Graziano, J. A. (1989). Effect of context and efference copy on visual straight ahead. *Vision Research, 29,* 179–1736.

Clower, D. M., Hoffman, J. M., Votaw, J. R., Faber, T. L., Woods, R. P., & Alexander, G. E. (1996). Role of posterior parietal cortex in the recalibration of visually guided reaching. *Nature, 383,* 618–621.

Dassonville, P., Schalg, J., & Schlag-Rey, M. (1995). The use of egocentric and exocentric location cues in saccadic programming. *Vision Research, 35,* 2191–2199.

Desmurget, M., Jordan, M., Prablanc, C., & Jeannerod, M. (1997). Constrained and unconstrained movements involve different control strategies. *Journal of Neurophysiology, 77,* 1644–1650.

Desmurget, M., Rossetti, Y., Jordan, M., Meckler, C., & Prablanc, C. (1997). Viewing the hand prior to movement improves accuracy of pointing performed toward the unseen contralateral hand. *Experimental Brain Research, 115,* 180–186.

Duhamel, J.-R., Bremer, F., BenHamed, S., & Graf, W. (1997). Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature, 389,* 845–848.

Enright, J. T. (1995). The non-visual impact of eye orientation on eye–hand coordination. *Vision Research, 35,* 1611–1618.

Flanders, M., Tillery, S. I. H., & Soechting, J. F. (1992). Early stages in a sensorimotor transformation. *Behavioral and Brain Sciences, 15,* 309–362.

Goodale, M. A., Meenan, J. P., Bulthoff, H. H., Nicolle, D. A., Murphy, K. J., & Racicot, C. I. (1994). Separate neural pathways for the visual analysis of object shape in perception and prehension. *Current Biology, 349,* 604–610.

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature, 349,* 154–156.

Hecht, S., Haig, C., & Chase, A. M. (1937). The influence of light-adaptation on subsequent dark-adaptation of the eye. *Journal of General Physiology, 20,* 813–850.

Henriques, D. Y. P., Klier, E. M., Smith, M. A., Lowy, D., & Crawford, J. D. (1998). Gaze-centered remapping of remembered visual space in an open-loop pointing task. *Journal of Neuroscience, 18,* 1583–1594.

Hubbard, T. L., & Ruppel, S. E. (1999). Representational momentum and the landmark attraction effect. *Canadian Journal of Experimental Psychology, 53,* 242–256.

Kerzel, D. (2002). Memory for the position of stationary objects: Disentangling foveal bias and memory averaging. *Vision Research, 42,* 159–167.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35,* 389–412.

Mateeff, S., & Gourevich, A. (1983). Peripheral vision and perceived visual direction. *Biological Cybernetics, 49,* 111–118.

Matin, L., & Fox, C. R. (1989). Visually perceived eye level and perceived elevation of objects: Linear additive influences from visual field pitch and from gravity. *Vision Research, 29,* 315–324.

Matin, L., Picoult, E., Stevens, J. K., Edwards, M. W., Jr., Young, D., & MacArthur, R. (1982). Oculoparalytic illusion: Visual-field dependent spatial mislocalizations by humans partially paralyzed with curare. *Science, 216,* 198–201.

McIntyre, J., Stratta, F., & Lacquanti, F. (1997). Viewer-centered frame of reference for pointing to memorized targets in three-dimensional space. *Journal of Neurophysiology, 78,* 1601–1618.

Murphy, K. J., Racicot, C. I., & Goodale, M. A. (1996). The use

of visuomotor cues as a strategy for making perceptual judgments in a patient with visual form agnosia. *Neuropsychology, 10,* 396–401.

Musseler, J., Van der Heijden, A. H. C., Mahmud, S. H., Deubel, H., & Ertsey, S. (1999). Relative mislocalization of briefly presented stimuli in the retinal periphery. *Perception & Psychophysics, 61,* 1646–1661.

O'Regan, J. K. (1992). Solving the ''real'' mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology, 46,* 461–488.

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of ''mudsplashes.'' *Nature, 398,* 34.

Pelli, D. G. (1997). The Video toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Pelz, J. B., & Hayhoe, M. H. (1995). The role of exocentric reference frames in the perception of visual direction. *Vision Research, 35,* 2267–2275.

Salinas, E., & Abbott, L. F. (1995). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Journal of Neuroscience, 15,* 6461–6474.

Schlag, J., & Schlag-Rey, M. (1995). Illusory localization of stimuli flashed in the dark before saccades. *Vision Research, 35,* 2347–2357.

Sheth, B. R., & Shimojo, S. (2001). Compression of space in visual memory. *Vision Research, 41,* 329–341.

Snyder, L. H., Grieve, K. L., Brotchie, P., & Andersen, R. A. (1998). Separate body- and world-referenced representations of visual space in parietal cortex. *Nature, 285,* 887–891.

Stuphorn, V., Bauswein, E., & Hoffman, K. P. (2000). Neurons in the primate superior colliculus coding for arm movements in gaze-related coordinates. *Journal of Neurophysiology, 83,* 1283–1299.

Tillery, S. I. H., Flanders, M., & Soechting, J. F. (1991). A coordinate system for the synthesis of visual and kinesthetic information. *Journal of Neuroscience, 11,* 770–778.

Vetter, P., Goodbody, S. J., & Wolpert, D. M. (1999). Evidence for an eye-centered spherical representation of the visuomotor map. *Journal of Neurophysiology, 81,* 935–939.

Vindras, P., Desmurget, M., Prablanc, C., & Viviani, P. (1998). Pointing errors reflect biases in the perception of the initial hand position. *Journal of Neurophysiology, 79,* 3290–3294.

Zipser, D., & Andersen, R. A. (1988). Reach plans in eye-centered coordinates. *Nature, 331,* 679–684.