The Quarterly Journal of Experimental Psychology

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Frames of reference in spatial span
Paolo Bernardis a b c & Tim Shallice a d
a Cognitive Neuroscience Sector, SISSA-ISAS (International School for Advanced Studies), Trieste, Italy
b Dipartimento di Psicologia, Università degli Studi di Trieste, Italy
c B.R.A.I.N. Centre for Neuroscience, Trieste, Italy
d Institute of Cognitive Neuroscience, University College London, London, UK

Available online: 13 Jul 2011

To cite this article: Paolo Bernardis & Tim Shallice (2011): Frames of reference in spatial span, The Quarterly Journal of Experimental Psychology, 64:12, 2438-2454

To link to this article: http://dx.doi.org/10.1080/17470218.2011.604202

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Frames of reference in spatial span

Paolo Bernardis1,2,3 and Tim Shallice1,4

1Cognitive Neuroscience Sector, SISSA-ISAS (International School for Advanced Studies), Trieste, Italy
2Dipartimento di Psicologia, Università degli Studi di Trieste, Italy
3B.R.A.I.N. Centre for Neuroscience, Trieste, Italy
4Institute of Cognitive Neuroscience, University College London, London, UK

In four experiments, a computerized Corsi-like paradigm was used to assess which of the many reference frames are used in visuospatial short-term memory. By varying the relative orientation (slanted +/−45° or in an upright position) of the head and the displays, we modulate the utility of the allocentric, egocentric (eye- and head-centred), and template-centred reference frames. The results of all experiments showed the crucial importance of the gravitational allocentric reference frames while using visuospatial short-term memory to retain a spatial sequence of elements. The results also provide some support for a mental rotation process involved in recognition following angular displacement of a multi-item display.

Keywords: Visual short-term memory; Spatial; Corsi test; Mental rotation; Span.

Imagine you are in the middle of a draughts or checkers game. It is your turn, and you are planning several moves ahead. A successful play requires that you remember all the pieces you will move and all possible positions that can occur given your opponent moves. This is a common type of cognitive task that requires visuospatial short-term memory, the ability to remember the locations of objects in the environment. This ability requires the capacity to localize the objects in space and to code their individual positions using an initially stationary reference position related to one, but preferably more, frames of reference (FoR). Traditionally, distinctions have been drawn between allocentric FoRs, ones based on the environment, and egocentric FoRs, ones based on the subject’s body (for a review see Jolicoeur & Humphrey, 1998; Klatzky, 1998). Put simply, the most crucial difference between the egocentric and the allocentric FoRs is the relation to the perceiver, who, in the egocentric FoR, is the origin of the coordinate system with respect to which locations are represented. In the allocentric FoR, on the other hand, points are located within a framework external to the perceiver and independent of his or her position.

Different subclasses can be made within these two FoRs. Traditionally, in the egocentric FoR we consider representations related to particular parts of the perceiver’s body as subclasses and so

Correspondence should be addressed to Paolo Bernardis, Dipartimento di Psicologia, Università degli Studi di Trieste, Via S. Anastasio, 12, 34134 Trieste Italy. E-mail: paolobernardis@units.it
The research was supported by two PRIN grants “Studio dell’evoluzione temporale dell’intenzionalita’ motoria nel lobo frontale” and “Studio neuropsicologico della rappresentazione e della pianificazione delle azioni nello spazio e nel tempo”.

2438 © 2011 The Experimental Psychology Society
http://www.psypress.com/qjep
http://dx.doi.org/10.1080/17470218.2011.604202
refer to retinotopic, head- or trunk-centred FoRs. A new subclass has recently been proposed, which is termed a template-centred frame of reference (Avons, 2007; McNamara, 2003; Mou & McNamara, 2002). In a template-centred FoR an object’s position is encoded with respect to the other objects in near space. It may be considered a subclass of the allocentric FoRs, but it differs from most allocentric FoRs in being intrinsic to a local set of objects and independent both of the subject and of general environmental frameworks like those provided by the ground or the walls of a room. It is plausible that it could be used in short-term memory, as it can be critical in object identification (Hinton, 1979; Hinton & Parsons, 1988).

A second very different allocentric FoR is that provided by the gravitational field. This, too, is known to be critical in object recognition. Rock showed that object recognition, no matter its retinal orientation, was better when carried out in the normal environmental upright position than in an alternative orientation (Rock & Heimer, 1957); moreover, in another study with ambiguous figures, observers chose the solution compatible with an upright orientation (Rock, 1956). Corballis, Zbrodoff, and Roldan (1976), and McMullen and Jolicoeur (1990) found similar results; they showed that observers judging the orientation of alphanumeric characters use gravitational coordinates rather than retinal ones.

An important attempt to study the role of FoRs in visuospatial short-term memory was made by Avons (2007) with a modified version of the Corsi Block Tapping Test, the visuospatial short-term memory test sensitive to right posterior lesions (Berch, 1998; De Renzi & Nichelli, 1975; Milner, 1971). The classical clinical version of the Corsi Test is composed of 9 rigid cubes attached to a board, which is placed on a table; the position of the cubes is fixed. Patients have to reproduce the sequence of cubes that the examiner has just touched. When the Corsi test is computerized, the board with the cubes is represented as seen from above, and the solid cubes are replaced by filled squares randomly arranged within a rectangular surrounding frame. This computerised version, which is displayed vertically, can be seen as a plan view of the Corsi Apparatus and can be imagined as a template (Avons, 2007). Avons assessed the ability of subjects to remember a 7-element spatial sequence when the whole display of elements moved to a new spatial location during the encoding or retrieval phase. The study involved only translation operations, when all square elements of the Corsi-like configuration moved together to a new spatial position during each step of the sequence presentation (encoding phase) or during sequence retrieval. Participants performed more poorly when a translation of the display occurred during the encoding phase. However, there was no problem if it occurred at retrieval. Avons concluded that retrieval of short-term memory for locations appears to be based on a translation-independent representation that uses a template-centred description of object spatial positions. However, translation during presentation is disruptive. Moreover, in the encoding phase egocentric cues could play a role.

The manipulation and methodology used by Avons did not allow him to distinguish between the two main FoRs, the allocentric environmental and the egocentric one (named the extrinsic ones by Avons as the positions of the squares are encoded with respect to an origin external to the stimulus), but only between the extrinsic ones and the template-centred one (a reference frame intrinsic to the display, using Avons’ terminology). This was a limitation in Avons’s methodology because the coding with respect to the two extrinsic reference frames could well produce different effects in visuospatial memory. Indeed, in the conclusions of his paper he suggests a possible role of egocentric mechanisms during the encoding phase, but implicitly excludes a role for allocentric mechanisms. However, this interpretation seems too strong, as a translation operation preserves coding with respect to one type of allocentric FoR—namely, with respect to the gravitational (or vertical or inertial) reference frame.

The purpose of this work is to investigate how the different frames of reference are used to structure representations in visuospatial short-term memory systems (Baddeley & Hitch, 1974; Logie, 1995).
or, more specifically, in the visual cache, one of the two components of the visuospatial sketchpad in Logie’s (1995) version of the Baddeley–Hitch working memory model. In our study we followed Avons in using a modified version of the Corsi test to assess spatial span in healthy participants. However, we go beyond the Avons study in using a combination of transformations of the head and display orientations. In particular to determine the relative contribution of the different types of frames of reference available, we manipulated the orientation of the display (the set of elements presented on the monitor) and the orientation of participants’ heads, by either a clockwise or a counterclockwise rotation. This method permitted us to separate the contribution of the main FoRs probably involved: the allocentric (gravitational) FoR, the egocentric FoR (eye- and head-centred FoR), and the template-centred FoR. Given that our method involved only head rotations and not body rotations, we were not able to separate the trunk-centred FoR from the gravitational one. Note that in the rest of the paper when we use the general term egocentric FoR, we mean the eye- and head-centred FoR, not the trunk-centred FoR.

Successful performance in this type of task involves the ability to perform spatial transformations like mental rotation on visual stimuli (in our case the displays). Since Shepard and colleagues (Cooper & Shepard, 1973; Sheppard & Metzler, 1971) demonstrated the possibility of studying mental imagery scientifically, many studies have been made of rotation. Traditionally, such a type of transformation has been studied with single objects but also for single positions in a frame (Bricolo, Shallice, Priftis, & Meneghello, 2000) and in a few cases for sets of objects (e.g. Wraga, Creem, & Proffitt, 1999). Given that the use of rotation was to extend the findings of Avons, in the current experiments we studied mental rotation in the context of a standard type of visual-spatial short-term memory design.

The combination of a change of the display orientation and/or the head orientation during the retrieval phase allows one to contrast different types of FoR. The simplest contrast involves three different conditions. In the first condition, in which there is no rotation either of the display or of the head, all three types of reference frame (egocentric, allocentric gravitational, and allocentric template-centred) could potentially be used to encode and retrieve the spatial position of the sequence. We label this condition as stationary since both the display and the head remain in the same position in both the encoding and retrieval phases. In a second condition, which requires a rotation of both the display and the head in the same direction, two reference frames—namely the egocentric (eye- and head-centred) and the allocentric template-centred—guarantee a correct encoding and retrieval of the sequence. We label this condition HD-rotated since both the display and the head are rotated in the retrieval phase. In the third condition, which requires only the rotation of the display, the allocentric template-centred reference frame appears the optimal one to ensure the correct encoding and retrieval of the sequence. We label this condition D-rotated, since only the display is rotated in the retrieval phase. We investigated how the ability to retrieve the sequence is influenced by the reference systems that are potentially effective. Comparison of the stationary and HD-rotated conditions would point to the importance of allocentric gravitational FoR and that between the HD-rotated and D-rotated to that of egocentric (eye- and head-centred) FoR, FoRs for which any possible specific effects could not be distinguished in the study of Avons.

EXPERIMENT 1

In the first experiment we manipulated the position of the head and display in order to produce the three different conditions in which different FoRs could be used with differing effectiveness. One additional factor that may influence the ability to

---

1 With the label allocentric gravitational FoR we refer to a reference system in which gravitational force is the main, but not necessarily the only, reference frame with respect to which objects are encoded.
operate on individual elements and to extract an allocentric template-centred frame of reference is the number of elements composing the display. In particular the larger the number of elements in the whole display, the higher is the probability of potential confusion between individual elements of the sequence in wrong, albeit close, positions. Moreover, a small number of elements is easier to retain than a large number of elements; this may make it easier to extract allocentric template-centred frames of reference. Given these premises, then, as stated above, we expected to obtain that displays with a small number of elements would be less handicapped by rotation operations.

Method

Participants

Fourteen students from the University of Trieste participated in this experiment. They were paid € 4, or they received course credits.

Apparatus

The experiment was displayed on a 19″ LCD monitor with a resolution of 1280 × 1024 pixels. The participant’s head rested on a chinrest. Attached to the chinrest were two firm and slightly cushioned surfaces slanted by 45°, adjoining the chinrest just on the left and right sides of the chin. The distance of the participant from the screen was 50 cm; the monitor subtended 28° of visual angle. The room was dimly illuminated.

Materials

The display consisted of 9 or 13 elements (light brown squares sides of 1.7° of visual angle) placed randomly on a black background and oriented in the same direction. The position of all elements changed randomly on each trial. At the centre of the display was a red triangle (1° of visual angle in height) indicating the orientation of the display. Each spatial sequence consisted of 5 elements chosen randomly without replacement from the set of 9 or 13.

Procedure

At the beginning of each trial, the whole display was presented for 1 s, then the first location of the sequence appeared. Each location was indicated by a change of the element colour to light green and was displayed for 1 s. The last location was followed by a 2-s blank screen during which a grating indicating the orientation of the head in the next retrieval phase was shown. The grating was composed of 7 thin red lines on a black background and was oriented vertically or slanted by ± 45°. Participants were asked to align their head in accordance with the orientation of the grating. When the grating was slanted, they had to place their cheek on the 45° slanted surface before the end of the 2-s interval. Next, the display appeared again in its retrieval position and remained stationary until the participant had finished responding. Participants responded by indicating the remembered sequence elements in order by moving the mouse and clicking with it. At the end of each trial participants realigned their head to the upright position. All participants were undergraduate students experienced in the use of a mouse; the motor task of pointing with a mouse was therefore natural to them and had been used very frequently by all of them. The presentation of each trial was self-paced to maximize attention levels.

The different combinations of the head and display position, which determine how many FoRs are available, generated three different conditions: stationary, HD-rotated, and D-rotated. We use the more general label FoR to name this factor. In the stationary condition (see Figure 1), participants kept their head vertical during both the encoding and the retrieval phases; the grating lines were oriented vertically. In the HD-rotated condition (see Figure 1), participants kept their head vertical during the encoding phase but rotated it during the interval; the grating lines were slanted by ± 45 degrees. The display was presented rotated in the retrieval phase, in accordance with the head orientation. In the D-rotated condition (see Figure 1), participants kept the head
vertical during both the encoding and retrieval phases; but the display was presented rotated in the retrieval phase. Consequently, the head and the display did not have the same orientation in the two phases.

In summary we had 3 FoR conditions (stationary, HD-rotated, and D-rotated), 2 directions of head movement (towards left and right), 2 different number of display elements (9 or 13), and 8 trials of each. In addition, the HD-rotated and D-rotated conditions were split into two directions of movement (the display elements rotated towards the left or the right). In the HD-rotated condition both the head and the display rotated in the same direction. Each subject participated in all conditions. The total number of trials for each subject was 80 divided into 6 blocks, plus 2 practice trials before each block of trials. Each block involved only one condition and one number (9 or 13) of display elements. For each number (9 or 13) of display elements, the 40 trials were divided into 8 trials for the stationary condition and 16 trials each for both the HD-rotated and D-rotated conditions (8 leftward and 8 rightward). Participants were able to relax for one or two minutes between each of the 6 blocks. The order of blocks was randomized across subjects.

To take into account the compensatory ocular counterroll that occurs when one rotates the head (Schworm, Ygge, Pansell, & Lennerstrand, 2002), we adjusted the display rotation by subtracting the number of degrees corresponding to the oculomotor compensation. Given that a 45° rotation of the head leads to 10° compensation of the eyes, we rotated the displays by 35°. This adjustment was necessary to ensure a perfect alignment between the retina and the display in the HD-rotated condition.

Analysis

The dependent variable was the proportion of elements correctly remembered, taking into account both position and order. We submitted the data to a within-subject analysis of variance in which the factors were the FoR (3 levels), the number of elements in the display (2 levels, 9 or 13 elements), and the serial position. To reduce the number of degrees of freedom, we summed results in neighbouring positions (1,2 vs. 4,5) so reducing the 5 serial positions to 2. Given that no recency effects were found in any of the experiments, the mathematical description of the 5-point serial position curve with only two points fit well the general trend of the serial position curve.

In this and all subsequent analyses, the Greenhouse–Geisser correction was applied where sphericity was violated, but we report the original uncorrected degrees of freedom. The Scheffe method was used ($p < .05$) to test post-hoc differences.

Results

In Figure 2 we plot the proportion of squares correctly recalled for each FoR condition and serial position. We observed three main results. The first was the difference between the FoR conditions: performance in the stationary condition...
was much better than that in the other two conditions for both display sizes, but performance in the two conditions HD-rotated and D-rotated was very similar. The second was the effect of the number of elements in the display: increasing the number of elements from 9 to 13 produced a decrease in accuracy. The third was the serial position effect, which consisted of a general decline across serial positions with no recency effect. These results were confirmed by the ANOVA. There was a significant effect of all three factors: FoR, $F(2, 26) = 54.15, p < .001$, number of elements, $F(1, 13) = 47.75, p < .001$, and serial position, $F(1, 13) = 45.06, p < .001$. Post-hoc analysis revealed that the stationary condition was different from both the HD-rotated ($p < .001$) and D-rotated ($p < .001$) ones, but that the HD-rotated and D-rotated conditions did not differ.

**Discussion**

Surprisingly, we found no accuracy differences between the HD-rotated and the D-rotated conditions. Only the head position was changed between these two conditions. In the HD-rotated condition the head was aligned with the display in both the encoding and the retrieval phases, and therefore both egocentric and allocentric template-centred FoRs were available (see Footnote 2). However, in the D-rotated condition the head was misaligned during the retrieval phase, and therefore of the two only the allocentric template-centred FoR was available. In other words, given that the performance was the same, it seems that in this task the visuospatial short-term memory system does not use positional information relative to the egocentric FoR relative to the head and/or on the eye. The other important result is the strong effect of allocentric cues, which increase the performance significantly. This effect could be due to gravitational and probably other environmental cues. In addition, if, as proposed by Avons (2007), the allocentric template-related FoR was the only FoR being used, then it is difficult to explain the large difference between the stationary condition on the one hand, and the HD-rotated and D-rotated conditions on the other. To confirm these observations, we decided to control and remove as far as possible all external position cues, so as to place greater emphasis on both the gravitationally defined FoR and the egocentric FoR, if it were available. The shape of the elements is a factor that generates a strong template FoR-related cue; when the elements are squares, the main orientation of the display is immediately perceived, and this could have increased the
importance of an allocentric template-centred FoR. In contrast, when the shape is circular, the utility of such allocentric template FoR-related cues is weakened. In Experiment 2 two different element shapes were used, a square and a circle, the second of which should weaken the power of the template-related FoR. To check whether this difference had an empirical effect, we used both types of shapes, square and circle.

In addition, we had to exclude a potential source of bias that could be present in the HD-rotated and D-rotated conditions, which arises from the need for participants to change head orientation from one direction to the other. In both these conditions participants first responded to 8 consecutive leftward trials in a block and then 8 consecutive rightward trials, or vice-versa. We noted that a difficulty could occur in the middle of the block. When participants had to move the head in one direction instead of the other, only 2 s are allowed to identify the direction of the indicator arrow and make the movement. It is plausible that observers need more time to adjust to a new direction of rotation. This bias could have affected the use of the eye/head egocentric FoR and therefore helped to hide any difference between the HD-rotated condition, where the head had to be rotated, and the D-rotated condition, where no head rotation was needed. We removed this possible artefact in the following experiment by modifying the experimental design.

The difference in performance between 9 and 13 elements confirmed our hypothesis: the accuracy decreased with the increase in the number of elements. This result was similar to those of Gunzelmann and Anderson (2006). In a visual scene with realistic 3D objects, participants had to identify the position of a target among several others randomly positioned. The target position was indicated on a differently oriented 2D map. Accuracy decreased with an increase in the number of elements (distractors) near the target. Gunzelmann and Anderson explained this result as an effect of the complexity of the description necessary to represent the spatial relations between the target and the surrounding elements, which becomes more complex as the number increases. In our case we can hypothesize a similar type of explanation of the way the position of the sequence elements was coded. The larger number of elements would have produced more difficult coding of each element of the sequence and so less accurate performance.

Comparing our results visually with those of Avons (2007), we obtained a superior performance on average, probably because we used sequences of 5 elements rather than 7. We tested the effect of this factor further in an additional experiment (see Appendix). This additional experiment confirmed the expected sequence length effect, showing a lowering of the serial position curve with 7 elements compared with 5. With 7 elements, we obtained serial position curves very similar to those of the Avons (2007) control conditions, except in one respect. He found a consistent recency effect, which we did not find.

EXPERIMENT 2

Experiment 2 was mainly aimed at removing all biases and external position cues except perceived gravity to emphasize the allocentric gravitational FoR and the egocentric FoR. We displayed elements through a circular hole, hiding the monitor frame and corners, which could be used as anchors to code elements position. Circles and squares were used. As in Experiment 1, we manipulated the position of the head/display. After these modifications perceived gravity and element shape (when the items were squares) were the only accessible allocentric cues; therefore the FoRs now available (see Note 2) in the stationary condition were the allocentric gravitational, the egocentric, and the allocentric template-centred one. As a main result we expected that by reducing the utility of external allocentric position cues (monitor frame and corners and the square-shaped elements), we would increase the importance of the eye/head egocentric ones. In addition, square-shaped elements would be expected to produce a general increase in performance compared to circular-shaped elements if template FoR-related cues is of critical importance.
Method

Participants
Twelve students (divided into 2 groups of 6) from the University of Trieste participated in this experiment. They were paid €4, or they received course credits.

Apparatus and materials
These were the same as in Experiment 1 except for the following modifications: the shape of the elements was square or circular (measuring 1.7° of visual angle in diameter), the total number of display elements was 9, the length of the sequence was 5 elements, the illumination of the room was totally dark, and the monitor frame was hidden by black stiff cardboard with a circular aperture in the centre. This aperture permitted a view of the whole display; the surface of the black cardboard filled the entire visual field. The monitor brightness was set to the minimum value to prevent indirect illumination of the surrounding environment.

Procedure
Procedure was the same as in the Experiment 1, except for the presence of two groups of participants. In each group, participants rotated the head in only one direction throughout the experiment when in an HD-condition. This was toward the left in the first group and toward the right in the second group. All participants were tested with both circular and square elements in all three FoR conditions. The presentation of each stimulus in each condition was repeated 16 times. In total there were 96 trials (16 repetitions of 2 shapes in 3 FoR conditions) divided into 6 blocks (2 shapes × 3 FoR conditions) randomly assigned to each participant, and 2 groups; 6 participants were assigned to the leftward rotation group, and 6 to the rightward rotation group.

Analysis
As in Experiment 1, we used the proportion of elements correctly remembered as dependent variables, the 5 serial positions falling into two levels. The other within-subjects factors were the FoR (3 levels) and element shape (2 levels). The head direction rotation was treated as a between-subjects factor.

Results
The general pattern of results was similar to that of Experiment 1. In Figure 3 we plot the mean accuracy (calculated as the proportion of elements correctly remembered, taking into account both

![Figure 3](image-url)
position and order) for each FoR condition and serial position. In an analogous fashion to Experiment 1, we observed effects of the FoR factor and serial position. The results for the circular and square elements were very similar. These findings were confirmed by an ANOVA involving the following factors: FoR, element shape, and serial position. Significant effects were found of the within-group factors: FoR, $F(2, 20) = 17.178$, $p < .001$, serial position, $F(1, 10) = 6.559$, $p = .028$. There was no effect of element shape, $F(1, 10) = 0.003$. There was also an interaction of FoR × Serial position, $F(2, 20) = 6.327$, $p = .016$; the difference between the three FoR conditions was more pronounced for the later serial positions. However, post-hoc analyses of the Condition × Serial position interaction showed that both in the earlier and the later serial positions, the stationary condition was different from both the HD-rotated condition (early serial positions: $p < .001$; late serial positions: $p < .001$) and the D-rotated condition (early serial positions: $p = .007$; late serial positions: $p < .001$), but the HD-rotated and D-rotated conditions did not differ.

Discussion

Experiment 2 did not confirm the expected element shape effect in that there was no difference in performance whether square elements or circular elements were used. More critically, the results of Experiment 2 confirmed the main effect of Experiment 1. Even when all major allocentric cues other than perceived gravity were removed, head- and eye-centred egocentric cues do not seem to be used in the task; apparently they are not useful for encoding and retrieving the spatial position of stimulus elements. Thus even when nearly all major allocentric cues are removed, the difference in performance between the stationary and the two rotated conditions means that template-related cues cannot be the only form of coding used. Instead, the allocentric gravitational FoR plays an important role in this type of task.

Before accepting such a strong conclusion, we need to exclude alternative interpretations. For this reason Experiment 3 was designed to rule out other biases and alternative explanations. There is a second potential bias related to head rotations. It is conceivable that the position of the head could create difficulties independently of the display. Perhaps having your head rotated produces a general reduction in performance. Encoding a sequence of elements or retrieving a previously learned sequence could be more difficult when the head is rotated toward one shoulder. This could be simply because one is not used to keeping the head rotated or, for instance, because it could create discomfort for the neck muscles and so lead to attentional problems. It could also be due to a reduced ability to identify shapes in that orientation. Although it is well known that rotation of the head in the picture plane does not influence recognition of familiar shapes (Corballis, Zbrodoff, Shetzer, & Butler, 1978; Eley, 1982; Jolicoeur, 1985; McMullen & Jolicoeur, 1992; Simion, Bagnara, Roncato, & Umiltà, 1982), when novel objects of the same complexity as letters are used, responses to stimuli rotated with respect to the encoding view were slower and less accurate (Tarr & Pinker, 1989). Our experiment used novel configurations of shapes and so may be subject to a similar problem. To avoid these potential biases, in Experiment 3 we modified the experimental design and added a condition in which the head rotates but the display does not.

EXPERIMENT 3

In this experiment we added a fourth condition, in which the head rotated but the display did not. In this last situation the encoding/retrieval of the sequence could be successfully achieved with respect to both an allocentric FoR, which is in this case the allocentric gravitational reference frame or the template-centred FoR. In Experiments 1 and 2, the comparison between the stationary and the other conditions suggested that the allocentric and, in particular, the gravitational FoR is very important for good performance in the retrieval phase. The similarities in performance between the HD-rotated and the D-rotated conditions suggest at most a minor role of the egocentric
FoR. For these reasons we expected good performance in the fourth condition, which is more similar to that observed in the stationary condition.

In addition, we manipulated the initial head starting position, which could be vertical or slanted, to control potential biases related to the unusual neck position. We did not expect to obtain effects of the initial head starting position, given that a comfortable surface supported the head when it was slanted.

Method

Participants
Twelve students (subdivided into 2 groups) from the University of Trieste participated in this experiment. They were paid €4, or they received course credits.

Apparatus and materials
They were the same as in Experiment 2, except for the modification that the shape of the elements was circular.

Procedure
This was the same as in Experiment 2, except for the inclusion of the fourth condition, and the two different head starting positions (vertical or slanted). The four possible combinations deriving from the modification of the head and display orientation are summarized as follows:

1. stationary condition (see Figure 1): the orientation of both the head and the display were kept identical (both vertical or slanted) in the encoding and in the retrieval phases;
2. HD-rotated condition (see Figure 1): the orientation of both the head (H) and the display (D) was changed in the retrieval phase;
3. D-rotated condition (see Figure 1): the orientation of the display (D), but not the head orientation, was changed in the retrieval phase;
4. H-rotated condition (see Figure 4): the orientation of the head (H), but not the display orientation, was changed in the retrieval phase.

In each condition participants started with the head oriented vertically or slanted. The presentation of each stimulus in each condition was repeated 16 times. In total there were 80 trials (10 trials × 2 starting head orientation × 4 FoR conditions) divided into 8 blocks randomly assigned to each participant, and 2 groups; 6 participants were assigned to the leftward rotation group, and 6 to the rightward rotation group.

Analysis
This was carried out as in the previous experiments, except that we divided the FoR factor into two nested factors: head orientation change (2 levels, stationary plus D-rotated conditions vs. HD-rotated plus H-rotated conditions: see Figure 4), and display orientation change (2 levels, stationary plus H-rotated conditions vs. HD-rotated plus D-rotated conditions: see Figure 4). The other within-subject factor was the starting head orientation (2 levels, namely vertical or slanted). Since no effects were found in the previous experiment, the direction of the rotation was not included in the analysis.
Results

In Figure 5 we give the mean accuracy for each FoR condition and serial position. As in the previous experiments, we found that the results of the four conditions fell into two levels—stationary and H-rotated being better, and HD-rotated and D-rotated worse. In an ANOVA involving the following factors: head orientation change, display orientation change, starting position, and serial position, significant effects were found of the within-group factors: display orientation change, \( F(1, 11) = 35.025, p < .001 \), serial position, \( F(1, 11) = 24.81, p < .001 \). There were no effects of the factors head orientation change, \( F(1, 11) = 0.01 \), or starting position, \( F(1, 11) = 0.062 \). There was an interaction of Head orientation change \( \times \) Display orientation change, \( F(1, 11) = 5.633, p = .037 \), and one of Head orientation change \( \times \) Display orientation change \( \times \) Serial position, \( F(1, 11) = 6.301, p = .029 \). Post-hoc analyses to test differences in the interaction of Head orientation change \( \times \) Display orientation change \( \times \) Serial position revealed that in the early serial positions the stationary condition was different from both the HD-rotated (\( p < .001 \)) and the D-rotated (\( p < .001 \)) conditions, and also that the H-rotated condition was different from both the HD-rotated (\( p = .008 \)) and the D-rotated (\( p < .001 \)) conditions.

Discussion

The equivalence in accuracy between the HD-rotated and D-rotated conditions obtained in the previous experiments has been replicated. In this experiment we analysed separately the effect of head orientation change (see Figure 6, third panel) from the effect of the display orientation change (see Figure 6, fourth panel). Only the second of these possible changes—in display orientation— influenced accuracy of responding; the main difference between the unchanged conditions (stationary plus H-rotated) and changed ones (HD-rotated plus D-rotated) is that the allocentric gravitational FoR can be used (see Footnote 1) in the former, but not in the latter case. Therefore this result confirms that a crucial role is played by the allocentric gravitational FoR in this particular spatial span task. By contrast, change or lack of change in the position of the head (stationary plus D-rotated versus H-rotated plus HD-rotated) did not produce a difference in spatial span performance. Facilitating or not the use of an eye/head egocentric FoR (stationary plus HD-rotated versus H-rotated plus HD-rotated) had only a minor effect in later serial positions, when allocentric cues are weaker.

The current findings mean that other potential biases, like attentional effects or effects due to an unusual or uncomfortable neck position (starting position see Figure 6, first and second panel), can also be excluded as having played a causal role in the patterns of results obtained previously, which were held to produce equivalent accuracy in the HD-rotated and the D-rotated conditions.

GENERAL DISCUSSION

The main result, which emerged from all experiments involving this Corsi-like task, is the lack
of any major eye/head egocentric FoR effects in the encoding and retrieval of a visuospatial sequence. In those situations where the system could potentially take advantage of the use of an eye/head egocentric as well as a template-centred FoR (HD-rotated condition), there was little or no difference from when a template-centred FoR alone was available (D-rotated condition). The same happened in the H-rotated condition, where the system made little use of the eye/head egocentric FoR relying instead on allocentric ones.

We are not arguing that the egocentric FoR plays no role in such tasks. Indeed, the relevance of the egocentric frame of reference is crucial in everyday situations where a computation of the spatial positions of the body and objects in the environment is required, and where there is an active interaction of the subject with the surrounding objects. And there are many studies that demonstrate the crucial importance of the egocentric FoR, in addition to the allocentric one, in spatial memory (for a review see Burgess, 2006). However, all these studies used a task requiring the specific positioning of the body and or the effectors in the environment for a correct solution. Burgess (2006) proposed that there are some factors like the amount of self-motion between encoding and retrieval, or the spatial features of the environment, which determine the weight of the egocentric and allocentric FoR in the mnestic trace.

In the current paradigm the use of an eye/head egocentric FoR is permitted (at least in the stationary, HD-rotated and H-rotated conditions) but not required in order to encode and retrieve the sequence. Moreover, in an experiment with conditions somewhat similar to our paradigm, Mou, Xiao, and McNamara (2008) investigated spatial memory of a briefly viewed layout and found results in line with ours. Participants saw an array of five objects on a table and, after a short delay, indicated whether the target object had been moved. Their results indicated that spatial memory uses representations of inter-object spatial relations and an environmentally defined reference frame instead of an egocentric reference frame.

From another perspective the important role of the inertial or gravitational FoR has emerges very clearly from these experiments. Our results confirm and extend those of Hock and Sullivan (1981) in the long-term memory domain. These authors investigated how strongly alternative spatial reference systems (gravitational vs. egocentric) were involved in intentional and incidental learning. Incidental learning was achieved by asking participants to give an aesthetic judgment on several polygons that were presented in an encoding phase; intentional learning was obtained by asking subjects explicitly to memorize the polygons explicitly in the encoding phase. The results showed that the gravitational FoR is the prevalent reference system used by participants when the instructions emphasize long-term memory encoding. By contrast, when participants were instructed to judge the aesthetic aspects of the polygons, the egocentric was the FoR of choice. In a similar
fashion, our results suggest that in the short-term memory domain, too, when mnestic abilities are explicitly required, the gravitational FoR is much more important than the eye/head egocentric one, which plays a minor role only when allocentric cues are especially weak.

In our introduction we pointed out that the combinations of head/display rotations used (stationary, HD-rotated D-rotated and H-rotated) would not discriminate entirely between the three main Frames of Reference that could be effective. In the posterior parietal cortex (PPC) there is evidence for neurons that encode spatial positions in different coordinate frames, each having different representations, such as an eye-(or retina-) centred map, a head-centred map, and possibly a trunk-centred map (Xing & Andersen, 2000). In all of our head/display rotation combinations, any trunk-centred egocentric FoR independent of gravity would also always be aligned with the allocentric gravitational FoR. Its possible role remains to be investigated. However, it should be noted that trunk position has been found not to influence spatial attention in normal subjects (Grubb & Reed, 2002).

Mental rotations

Ever since they were initially investigated (Cooper & Shepard, 1973; Shepard & Metzler, 1971), there has been controversy concerning the mechanisms underlying mental rotation. It has been standard to assume that they involve analogue-type transformations of spatial representation, but alternative interpretations have been used to explain similar results without the need to make analogue transformations (Pylyshyn, 1981, 2003; see also Bricolo et al., 2000). Whatever the underlying process, if we rotate a set of elements around its centre in an experimental paradigm like the so-called Corsi Test (Milner, 1971), one can reproduce the sequence after the rotation only if one can make a one-to-one mapping between the encoded and the transformed set of elements. This will require the ability to extract a frame of reference and to use that to match the encoded and the transformed state since the individual elements are not individually distinguishable and to make multiple computations on each element separately would be impracticable.

Another fundamental ability to succeed in such a task is the perception of the objects’ orientation and form. In particular, the retinal orientation and the assignment of cardinal direction, both upright and left/right directions are critical (Rock, 1974; Rock, Di Vita & Barbeito, 1981). The perception of uprightness is influenced most by two factors: the surrounding environment and the gravitational (inertial) reference frame (Attneave & Olson, 1967). Most investigators who have dealt with the problem of orientation in form perception indicate that the gravitational FoR could have a crucial role both in encoding and recognition process.

Returning to our findings, in the HD-rotated and D-rotated conditions, which are characterized by different orientations of the display between the encoding and the retrieval phases, a mental rotation is required to reproduce the sequence. Presumably the mental rotation is of the overall shape of the elements of the display related to a principal axis. This presupposition receives support from the different latencies of the first response. In Figure 7 we plot the mean latencies of the first move and the average time of the following moves, according to condition and experiment. The latencies of the first move of the conditions HD-rotated and D-rotated are clearly longer (range: 1.5 to 2 s) than those of stationary (1 s) and H-rotated conditions (1.1 s), and this is true for all four experiments. By contrast, the average latencies of the subsequent moves are similar across all conditions and for all experiments (range: 0.8 to 1.1 s). The latency difference support the idea that a brief initial time is spent to rotate the display, before the participant starts to point with the mouse to the element sequence. The solution is achieved in three steps: first a mental rotation (performed only in the HD-rotated and D-rotated conditions), presumably of the principal axis of the display, which is necessary to realign it, then retrieval of the memory trace relative to this rotated axis, and then the sequence reproduction.

A possible alternative to a Shepard-style analogue mental rotation is that an allocentric template-centred FoR relies on extraction of principal
axes. Then one must extract the same principal axes from the rotated display, but this is more difficult and thus takes time. On either of these possibilities—the mental rotation or more difficult extraction of the principal axes—the process must be operating in the allocentric gravitational domain, not the egocentric one.

**CONCLUSION**

Two main conclusions can be drawn from these experiments. When one has to retain a spatio-temporal sequence of objects, the most reliable FoR that can be used is an allocentric system and, in particular, one that is based on gravitational cues. No matter what the orientation of the head and eyes, the positions of objects in space are coded by integrating information from local (the allocentric template-centred FoR), and general (the allocentric gravitational FoR) reference systems. This result confirms those obtained Friederici and Levelt (1990), who studied the ability to assign positions in space in the absence (and presence) of gravitational information. They found that under weightless conditions during inflight zero gravity test participants adopt head-retinal vertical as a frame of reference, but in 1 g conditions the gravitational vertical is used whenever possible.

In contrast to the results of Avons (2007), our findings on rotation exclude any major effect of the use of egocentric cues, computed from an eye- or head-centred coding, in visual short-term memory, and emphasize the crucial role of gravitational ones. Avons used a translation operation on spatial relations. This type of manipulation does not permit one to separate allocentric and egocentric cues. By contrast, the combination of head/display rotation allows the effects of allocentric and egocentric (when based only on a eye- and head-centred coding) cues to be separated and shows allocentric gravitational cues to be critical.

Original manuscript received 25 October 2010
Accepted revision received 17 April 2011
First published online 18 August 2011
REFERENCES


**APPENDIX**

**Additional experiment**

To provide a comparison with Avons’ study (2007), we increased the length of the sequence to contain 7 elements. We expected that this would lead to a serial position curve similar to that reported by him.

**Method**

Participants

Twelve students (divided into 2 groups) from the University of Trieste participated in this experiment. They were paid €4, or they received course credits.

**Apparatus and materials**

These were the same as in Experiment 2, except for the following modifications: the shape of the elements was only circular; the length of the sequence was 5 or 7 elements.

**Procedure**

The same as in Experiment 2. In total, there were 96 trials (16 repetitions of 2 number of elements in 3 FoR conditions) divided into 6 blocks (2 shapes × 3 FoR conditions) randomly assigned to each participant, and 2 groups; 6 participants were assigned to the leftward rotation group, and 6 to the rightward rotation group.

**Analysis**

As in Experiments 1–3 we used the proportion of elements correctly remembered as the dependent variables, with the 5 and 7 serial positions falling into two levels, (1 + 2 vs. 4 + 5, and 1 +

**Figure A1.** Accuracy of serial spatial recall in the three conditions. The left-hand graph shows findings with 5 elements, the right-hand one with 7 elements.
2 + 3 vs. 5 + 6 + 7). The within-subjects factors were the FoR (3 levels) and the sequence length (2 levels). The head direction rotation was treated as a between-subjects factor.

**Results**

The general pattern of results was similar to that of Experiment 1. In Figure A1 we plot the mean accuracy (calculated as the proportion of elements correctly remembered, taking into account both position and order) for each FoR condition and serial position. In an analogous fashion to Experiment 1, we observed effects of the FoR factor, the sequence length, and serial position. These results were confirmed by the ANOVA. There were significant effects of the factors: FoR, $F(2, 20) = 34.09, p < .001$, sequence length, $F(1, 11) = 39.22, p = .0008$, and serial position, $F(1, 11) = 26.98, p = .002$. Post-hoc analysis again revealed differences between the stationary condition and the HD-rotated ($p < .001$) and D-rotated ($p < .001$) ones, but not between the HD-rotated and D-rotated conditions.

**Discussion**

The additional experiment confirmed the expected sequence length effect by showing a lower serial position curve with 7 elements compared with 5. With 5 elements, we obtained serial position curves very similar to those of the Avons (2007) control conditions, except in one respect. He found a consistent recency effect, which we did not find. More critically, the results of this experiment confirmed the main effect of Experiment 1. Even when all major allocentric cues other than perceived gravity were removed, head- and eye-centred egocentric cues do not seem to be used in the task; apparently they are not used in order to encode and retrieve the spatial position of stimulus elements.