

Comparing Location Memory for 4 Sensory Modalities

Hendrik N.J. Schifferstein¹, Monique A.M. Smeets² and Albert Postma³

¹Department of Industrial Design, Delft University of Technology, Landbergstraat 15, 2628 CE Delft, The Netherlands, ²Clinical and Health Psychology, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, The Netherlands and ³Experimental Psychology, Helmholtz Institute, Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands

Correspondence to be sent to: Hendrik N.J. Schifferstein, Department of Industrial Design, Delft University of Technology, Landbergstraat 15, 2628 CE Delft, The Netherlands. e-mail: h.n.j.schifferstein@tudelft.nl

Accepted November 17, 2009

Abstract

Stimuli from all sensory modalities can be linked to places and thus might serve as navigation cues. We compared performance for 4 sensory modalities in a location memory task: Black-and-white drawings of free forms (vision), 1-s manipulated environmental sounds (audition), surface textures of natural and artificial materials (touch), and unfamiliar smells (olfaction) were presented in 10 cubes. In the learning stage, participants walked to a cube, opened it, and perceived its content. Subsequently, in a relocation task, they placed each stimulus back in its original location. Although the proportion of correct locations selected just failed to yield significant differences between the modalities, the proportion of stimuli placed in the vicinity of the correct location or on the correct side of the room was significantly higher for vision than for touch, olfaction, and audition. These outcomes suggest that approximate location memory is superior for vision compared with other sensory modalities.

Key words: cross-modal comparison, location memory, multisensory, object-location binding

Introduction

While people navigate through an environment (e.g., a building, a city, a natural landscape), they encounter many different sources of sensory information: People may notice the bright colors of a trash can, they may hear the sound coming from a drilling machine, they may smell a bakery, or they may notice the plants that tickle them when they accidentally touch them. The knowledge thus gathered can be used to develop a cognitive representation of an unknown environment based on a number of landmarks (Siegel and White 1975). Because all instances of sensory stimulation may leave a trace in memory, all types of sensory information can contribute to the construction of a spatial representation of the environment.

Up to now, the capacity of olfaction to function as a spatial sense has received only limited research attention. Porter et al. (2007) recently showed that humans can follow scent trails to determine the location of an odor source. However, the mechanisms that underlie this ability have not yet been fully elucidated. An early study (von Békésy 1964) suggested that both time delay and intensity differences between the 2 nostrils made it possible for people to localize an olfactory stimulus (i.e., directional smelling). However, later studies

suggested that internostriol intensity differences may depend on trigeminal stimulation rather than on olfactory stimulation (e.g., Kobal et al. 1989). In addition, head movements (Schneider and Schmidt 1967) and active sniffing (Porter et al. 2007; Frasnelli et al. 2009) have been suggested to contribute to odor localization ability.

In the present study, we investigate another potential spatial capacity of olfaction, that is, the capacity to remember specific locations of odor sources. The ability to remember locations is an important capacity for people and animals in general because it allows them to go back to locations where necessities of life can be found, such as shelter, water, food, and potential mates. In our study, participants were given a task in which they navigated through 3D space in order to put objects with distinctive smells back on their original location. The objects were all presented at the same height. Hence, object locations varied in 2 dimensions. We compared people's performance in relocating objects for olfaction to their performance for the sense of vision, audition, and touch, using stimulus sets that were similarly identifiable, intense, and complex.

Components of object location memory

Relocating objects in space is dependent on 3 different processes. First of all, participants have to recognize the different objects and should be able to distinguish between the various objects (object recognition). Second, they should be able to remember the different positions in space (positional encoding). Third, they should be able to indicate which object occupied which position (object–location binding) (Postma et al. 2008). According to Postma and De Haan (1996), the latter 2 processes are largely independent. These authors found that performance in an object-to-position assignment was dependent on capacity manipulations (number of stimuli) and verbal mediation, whereas positional encoding remained unaffected by these 2 factors. In the present experiment, we focused on the object processing and the object–location binding components.

Regarding object processing, we know of no studies that have directly compared object recognition over multiple sensory modalities. An important issue in comparative research is how to select comparable sets of stimuli because stimulus dimensions may vary considerably over modalities. In an attempt to assess how recognition performance is likely to differ between modalities, we have compared studies on identification performance for stimuli representative of everyday conditions hence applying the criterion of ecological validity. In fact, studies in which untrained observers were asked to identify common objects suggest that the highest identification performance is found for vision [near 100%] and touch [95–96%] (Klatzky et al. 1985; Klatzky, Loomis, et al. 1993). Product sounds and smells are, on average, accurately identified by 55% (range = 4–100%, standard deviation [SD] = 31%) (Ballas 1993) and 39% (range = 0–85%, SD = 24%) (Desor and Beauchamp 1974) of the participants, respectively. Schifferstein and Cleiren (2005) directly compared identification performance between 4 sensory modalities for a set of familiar, hand-size objects. They found that identification performance was comparable for vision and touch, intermediate for audition, and lowest for olfaction. However, the authors noted that vision is likely to outperform touch in many real-life situations because vision gathers information more rapidly than touch (Jones and O'Neil 1985) and can also easily process very large objects. As a consequence, vision has been found to direct exploration through touch (Heller 1982; Klatzky, Lederman, and Matula 1993). If these differences in identification performance approximate between-modality differences in stimulus recognition, they suggest that we either have to degrade all nonolfactory stimuli in order to keep recognition performance similar for all modalities or statistically control for object recognition differences when comparing the object–location binding scores.

The main question we try to answer is whether people connect a visual object or a picture easier to a location than a just heard sound or a particular smell. The literature on object–

location binding within the nonvisual modalities is extremely scarce. As regards the olfactory domain, Takahashi (2003) presented female participants with 5 unfamiliar odors in 2 different rooms. When asked to remember the odor only, the proportion of correct localization judgments was 0.56 and 0.58 in 2 studies, compared with a chance level of 0.50. When instructed to memorize the location as well, this proportion increased to 0.61, and after the rooms had been made more distinctive to 0.71. Degel and colleagues instructed participants to complete a variety of tests in a weakly scented room. Although participants were unaware of the scent, the degree of fit between odor and test room increased for participants who were unable to identify the odor (Degel and Köster 1999; Degel et al. 2001).

Klatzky et al. (2002) found that relocation performance in visual and auditory conditions did not differ in a virtual environment. Participants either saw a verbal label when they looked in the right direction (visual condition) or they heard the label being spoken from a loudspeaker when they faced the right direction (auditory condition). In the test stage, participants indicated the original direction of the stimulus by selecting 1 out of 5 possible directions in the horizontal plane. The functional equivalence for vision and 3D sound was replicated in a follow-up study (Klatzky et al. 2003), although in this study learning was faster for vision than for 3D sound, and vision tended to produce a more precise memory representation than sound.

The role of vision and touch in stimulus relocation tasks was compared most directly in the studies performed by Newell and colleagues, who created scenes of small, wooden shapes of familiar objects that could either be learned through vision or through touch. After learning a scene with 7 objects, 2 of the objects were exchanged, and the participants had to indicate which objects had changed positions. Newell et al. (2005) found that performance was similar in the visual and the haptic condition. Performance decreased when the orientation of the scene was changed by the experimenter between learning and test stage or when the participants had to switch between modalities (e.g., visual learning followed by a haptic test). For both modalities the effect of orientation change was partly offset when the participants actively moved to another location, while the scene remained unchanged (Pasqualotto et al. 2005). In the latter study, performance in the haptic conditions was slightly worse than in the visual conditions.

In conclusion, location memory has been investigated in separate studies for all 4 sensory modalities that are of interest here. Studies comparing object–location binding over multiple modalities have generally found little differences between the sensory modalities, although some have suggested a small advantage for vision.

The present study

The aim of the present study is to investigate whether humans are capable of relocating a smell, that is, whether they

remember the location where they previously experienced a smell, without the opportunity to follow a trace to that location. Furthermore, we investigate how this ability for olfaction compares with that of other sensory modalities, especially vision. To our knowledge, the present study is the first to compare object–location binding among 4 sensory modalities (olfaction, vision, audition, and touch) in a single study. First, participants explored 10 different locations in a room, where they opened a cube and perceived the stimulus inside. Later, they were asked to put the stimuli back on their original location.

The first dependent measure is the number of stimuli that is correctly relocated. From an evolutionary perspective, it is beneficial to be able to localize any important source of sensory stimulation, such as a certain sound (e.g., the laughter of a baby or the howl of a dangerous animal), smell (e.g., the baking of food or the fumes of a fire), or feel on the skin (e.g., a person touching you or a hot kettle hurting you). However, in all these examples, finding a direction or a rough body location is already sufficient to perform basic functions. Memorizing the exact source location is not always necessary to perform the basic functions described in these examples. To test relocation accuracy at a coarser level, we also recorded whether the stimulus was placed in the direct vicinity of the original location or on the right side of the room.

Furthermore, in a recognition task, participants indicated whether they recognized the stimuli from the relocation task or not, and we asked them to indicate whether their memory was detailed or general. This procedure was inspired by studies of the remember–know paradigm (Rajamaran 1993), which investigates the involvement of conscious memory in stimulus recognition by requiring participants to make a distinction between stimuli for which they have a conscious recollection of the previous encounter (remember) and stimuli which they just know to have encountered before, even though they do not have a conscious recollection (know). The results from this task were used to evaluate whether the stimuli in the 4 stimulus sets were recognized to approximately the same degree.

Methods

Stimulus selection

In order to make a plausible comparison between different sensory modalities, we needed stimulus sets that differed in perceptual modality but were otherwise equivalent for all 4 conditions (vision, audition, touch, and olfaction). Because many everyday smells generally tend to be hard to identify and provide only limited information on their sources (e.g., Schifferstein and Cleiren 2005), we reduced the amount of information in the stimuli for the other modalities. Therefore, we first performed 2 pilot studies in which we selected the stimuli for the main study.

The first selection criterion was that stimuli should be hard to identify and should elicit as few associations as possible;

otherwise, it would be relatively easy to label the stimuli verbally. In that case, the verbal labels might be remembered instead of the stimulus percepts. For the visual modality, there is evidence that people can remember the identity and location of pictures of meaningless, abstract objects that cannot be easily named (e.g., Nunn et al. 1998), and we assumed this to be the case for other sensory modalities as well. To optimize between-modality comparability, the second criterion was that the stimuli for each modality should differ on one, qualitative dimension only. All stimuli should be easily distinguishable in quality, whereas they should exhibit similar levels of complexity, intensity, and pleasantness.

We decided to use 10 stimuli per modality in the main study because people are able to store approximately 7 ± 2 chunks of information in working memory (Miller 1956). In a positional encoding task using sequential stimulus presentation, Igel and Harvey (1991) found that errors increased linearly with the number of stimuli (1–10). Furthermore, the use of hardly identifiable nonsense stimuli in all conditions is likely to make the task relatively complicated (e.g., Postma and De Haan 1996). Hence, by using 10 stimuli, all participants were likely to make some mistakes in all conditions.

The visual stimuli were 2-dimensional black-on-white single-line drawings of free forms, made with the use of mathematical equations (Davis 1975). The basis for each form is a circle, and the number and size of in- and outward folds vary. Each form had a diameter of approximately 13 cm and was presented on a 15×15 cm white background. The tactile stimuli were surfaces of 15×15 cm that differed in texture. Both natural and artificial materials were used. During the evaluation of the tactile stimuli, participants were unable to see the tactile stimuli in order to avoid interference from visual perception. For the auditory stimuli, we used sound fragments from www.findsounds.com. The sounds selected did not contain any melody, rhythm, or words. To make these sounds even more difficult to identify, we played them backward, and we changed the speed. The sounds lasted 1 s and were presented on Olympus voice recorders. The olfactory stimuli consisted of unfamiliar odorants provided by Quest, of which 0.5 mL was applied to Sorbarods (Ilacon Ltd) and stored in a cool room. A Sorbarod consisted of a closable cylinder (about 3.5 cm high and 2 cm in diameter) filled with an absorbent material (cellulose acetate). The experimental procedure in the olfactory condition of the pilot study deviated from the other 3 because participants waited for 25 s between responding to one stimulus and sampling the next one in order to avoid adaptation.

In the first pilot study, 40 participants ($N = 10$ per modality) from the participant pool used in the main study rated the stimuli for a single modality (30 stimuli for vision, touch, and audition; 20 stimuli for olfaction) on perceived intensity, complexity, and pleasantness on 5-point scales. In addition, they indicated whether they recognized the stimulus, and they reported the associations evoked by the stimulus. Stimuli with extreme mean ratings (<2.0 or >4.0) on intensity, complexity,

or pleasantness were deleted from the set. Subsequently, we selected 10 stimuli for the final set that were recognized by the smallest number of people and that elicited the smallest number of associations. In addition, 10 other stimuli were selected for use as distractors in the recognition task.

In a preliminary experiment using these stimulus sets, we found that recognition ability still differed considerably between stimulus sets. Therefore, we performed a second pilot study. Starting out from the preselected stimulus sets, a panel of 5 sensory experts selected new stimulus candidates for inclusion in the sets. In addition, for each sensory modality a questionnaire was constructed with 10–18 items. Four of these items were general and were used for all modalities: “To what extent do you find how this [material feels / scent smells / sound sounds / shape looks] [pleasant / complex / familiar / associated with a well-known product]?” The other items were modality specific. For each sensory modality, the 10–20 preselected stimuli were rated on the 10–18 items by 15 students, using 7-point scales. For each sensory modality, a Principal Components Analysis (PCA) with Varimax rotation was performed on the responses. In this analysis, the judgments from a single participant for one product were used as cases.

To determine the final selection of the stimuli, the mean responses for each stimulus on the factors emerging from the PCA were determined. Because the tactual set contained only 10 stimuli and this set performed satisfactorily in the preliminary experiment, we tried to obtain stimulus sets with comparable patterns of variation over factors for the other modalities as well. This led to a final selection of 10 stimuli for each modality (see Appendix).

Stimulus presentation

Participants were introduced to the stimuli by walking to a stimulus at 1 of 10 different locations, opening a cube, and perceiving the stimulus. This task mimics the natural behavior of a person who explores an unfamiliar environment in which various stimuli can be discovered (e.g., browsing in a shop), while the person mainly uses visual and proprioceptive inputs to navigate through the environment. By using this procedure, the person's referential frame tends to be egocentric, in which the position of landmarks is relative to the changing position of the person walking. Participants' ability to remember these stimuli was later tested by having participants put the stimuli back on their original location. Although the stimuli differed between conditions, the way in which participants explored the environment (they move through the environment, they see the environment containing all the information on the spatial frame [walls, door, windows, floor, ceiling], and they open the cubes) in the learning phase and the way in which they relocated the cubes in the test phase were kept equal in all conditions. Hence, the involvement of the senses was similar over the 4 conditions, except for the experimental stimuli used. Stimulus duration was self-paced to approach ecological conditions.

All stimuli were presented in SensaCubes. The cubes were distributed approximately evenly over the space. The distances between all SensaCubes were 1.4 m or more. Five cubes were presented on each side of the room (left or right). On each side, they formed a diagonal front row with 3 cubes, and a diagonal back row composed of 2 cubes (Figure 1). All SensaCubes were presented at the same height (110 cm). They were rotated in such a way that it was easy to open them.

The SensaCubes were made from medium-density fiberboard (MDF). They were cubes measuring 20 × 17 × 17 cm (width × height × depth) that could be opened either from the front (tactile) or from the top (visual, auditory, olfactory) depending on the type of stimulus. Stimuli were attached to MDF pieces (17 × 15 cm) that were slid into the cubes at approximately 11 cm from the top of the cubes (Figure 2).

The visual and tactile stimuli were glued to the slides. In the tactile condition, visual perception was blocked by a piece of flexible rubber with a hole in the middle, through which the participant inserted the hand to feel the textures with their fingertips. Any sounds produced during tactile stimulus exploration were likely to be masked by sounds of touching the cube and the rubber cover. Each auditory stimulus (1 s) was recorded multiple times with a 2-s intermittent break between repeated stimuli on an Olympus voice recorder. In the olfactory condition, the Sorbarod with odorant was fitted into the slide. The Sorbarod was covered by a polystyrene beaker and a ring of flexible foam when the SensaCube was closed. As soon as the participant opened the SensaCube, the smell was released from the beaker.

Participants

Eighty participants, 48 men (17–35 years, mean age = 21.8 years) and 32 women (18–25 years, mean age = 20.4 years) took part in the study. They were undergraduate students from Delft University of Technology. Each participant took

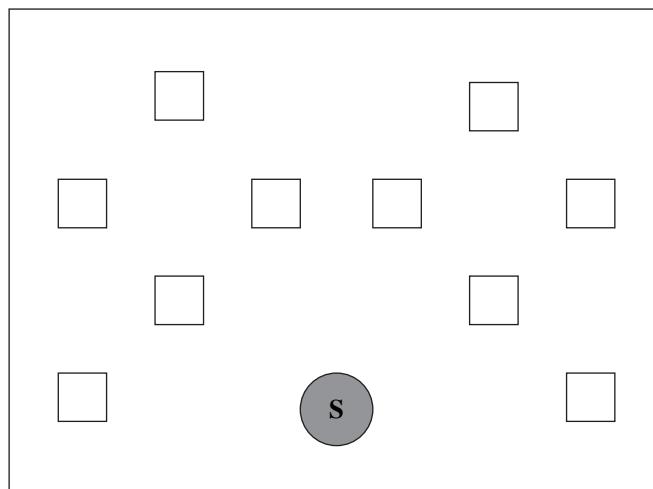


Figure 1 Spatial arrangement of the 10 stimuli in the room. S indicates the starting position.



Figure 2 Overview of the interior of a SensaCube in the olfactory condition, showing the Sorbarod attached to the slide and the plastic cup that covered the Sorbarod when the cube was closed.

part in 1 of the 4 unisensory conditions only ($N = 20$ per condition). The number of women per condition varied from 7 to 9. Participants were paid for their participation.

Procedure

Participants were taken to the marked starting point (S in Figure 1) in a rectangular room (about 5×7 m), where they faced 10 identical pillars with SensaCubes that each contained one stimulus. The stimuli were arranged in a front row (3 stimuli) and a back row (2 stimuli) on the left and right sides of the room, respectively.

Upon arrival, participants were informed how a SensaCube should be opened and how the stimulus should be explored. In the visual, auditory, and olfactory conditions, participants opened the cube from the top. In the visual condition, they just looked into the box. In the auditory condition, they pressed a green button to switch on the sound of the voice recorder and a red button to switch it off after listening to the sound. In the olfactory condition, they were instructed to approach the box carefully with the nose and to come closer only if they could not smell the odorant. In addition, they were instructed to smell the inside of their elbow after every

sample to neutralize their sense of smell. In the tactful condition the cube was opened at the front. Participants were instructed to use their preferred hand to evaluate all stimuli. Participants inserted their fingers through the opening in the rubber cover and moved their fingertips over the surface, without using their nails.

Subsequently, the experimenter instructed the participant to walk to a specific box by indicating its position in the room (e.g., front row, first box on the right). Depending on condition, they were instructed to carefully look at, listen to, sniff, or touch the stimulus in the SensaCube and, subsequently, to walk back to the starting point. After having returned to the starting point, the experimenter would indicate the position of the next box and so on. The sequence in which participants opened the cubes was randomized and differed between participants. After perceiving the stimuli in all cubes once, the procedure was repeated using a different presentation sequence so that all stimuli were sensed twice. On average, the learning phase took between 6.5 and 7 min for vision, audition, and smell and 8.5 min for touch.

After having been presented with all 10 stimuli twice, participants waited outside in the hall to give the experimenter the opportunity to prepare for the test phase. Then, the participants were informed that their task would be to place each stimulus back on its original location. They were deliberately not informed about this task beforehand to avoid activating any strategies that would optimize memorization, such as trying to label each stimulus with a name. After returning to the starting point, participants were presented with a stimulus in a SensaCube and were instructed to place the SensaCube back on its original location. After the participant had placed the cube back on 1 of the 10 pillars, they were asked to give a number on a 7-point scale indicating the degree of certainty with which they had placed the stimulus back in its original location. Subsequently, the participant picked up the cube and gave it back to the experimenter at the starting point. Then, the experimenter handed over the next cube to the participant, with the instruction to place this cube back and so on. Note that this procedure allowed participants to put multiple cubes back in the same location. The sequence of stimuli during the localization task differed from the sequence in which stimuli had been perceived during the learning task. On average, the test phase took between 5.5 and 6 min for vision, audition, and smell and 7 min for touch.

Finally, participants performed a recognition task in a separate room. In the recognition task, participants were presented with 20 stimuli, 10 targets and 10 distractors of a single modality. Participants indicated whether they 1) recognized the stimulus from the previous task and had a detailed, conscious recollection of the moment they had perceived it, possibly supplemented with associations, 2) recognized the stimulus but only in a general way without any specific details of the previous encounter, or 3) did not recognize the stimulus from the previous task.

Data analysis

The responses in the recognition task were analyzed using signal detection theory measures (Macmillan and Creelman 1991) based on the “no” versus “yes” (general and detailed) responses in order to differentiate between perceptual sensitivity and response bias. In this context, perceptual sensitivity refers to the participant’s ability to distinguish between “old” stimuli previously encountered during the experiment and “new” stimuli not previously encountered. Response bias, on the other hand, refers to the tendency to favor one response over the other. Because we asked our participants to indicate whether they had a detailed, general, or no recollection, we could classify either all the detailed and general recollections or only the detailed recollections as “yes” responses. Although these 2 approaches yielded different values of $P(\text{hits})$, A' , and B'' , the 2 calculations yielded similar outcomes when comparing the different modalities because the number of general recollections did not differ between conditions. Therefore, we only report the calculations here based on the total number of detailed and general recollections.

We calculated A' as a nonparametric measure of perceptual sensitivity from the proportion of hits (H = the proportion of “yes” responses when the stimulus had been presented in the previous task) and the proportion of false alarms (F = the proportion of “yes” responses when the stimulus had not been presented before) using the following formulas (Grier 1971; Aaronson and Watts 1987):

$$A' = 0.5 + [(H-F)(1+H-F)]/[4H(1-F)] \text{ for } H > F \text{ and}$$

$$A' = 0.5 - [(F-H)(1+F-H)]/[4F(1-H)] \text{ for } H < F.$$

We calculated B'' as the nonparametric measure of response bias using

$$B'' = [H(1-H) - F(1-F)]/[H(1-H) + F(1-F)] \text{ for } H > F \text{ and}$$

$$B'' = [F(1-F) - H(1-H)]/[H(1-H) + F(1-F)] \text{ for } H < F.$$

A' represents the area under the Receiver Operating Characteristic (ROC) curve; it grows with increasing perceptual sensitivity and its value can vary from 0 to 1. B'' reflects the degree to which an ROC curve tends to be asymmetrical with respect to its negative diagonal; it varies from -1 (extreme bias to say “yes”) to +1 (extreme bias to say “no”). The proportion of hits, the proportion of false alarms, A' , and B'' were calculated per participant. These variables were all subjected to univariate analysis of variance (ANOVA).

For the localization task, we calculated the proportion of correctly relocated cubes for each stimulus (P_{loc}). In addition, we recorded whether the cube was placed back on a spatially neighboring location, just one place away from the correct location (P_{near}). We also recorded the number of times a cube was placed on the correct side of the room: left or right (P_{side}). The latter measure presupposes that participants might be

able to pick the correct side of the room if they had only a rough idea about the correct location. Subsequently, we tested with univariate ANOVA for P_{loc} , P_{near} , and P_{side} whether there were significant differences between the 4 modalities. First, this analysis was performed on the values of P calculated from all responses. Second, the analysis was performed for values of P for trials in which the stimulus yielded a detailed recollection in the recognition task. Post hoc *t*-tests with Bonferroni correction were used for paired comparisons.

We expected that relocation accuracy would be higher if participants were more certain about their response. Therefore, we determined the mean certainty ratings for correct and incorrect responses for the various measures of P . In addition, we checked whether certainty ratings were also related to reporting detailed, general, or no recollection in the recognition task.

Results

Stimulus recognition

Because our approach suggests that the unisensory stimuli in the 4 sets need to be recognizable to the same degree in order to evaluate the results of the relocation task, we first analyzed the data of the recognition task. Table 1 shows the mean values for the proportions of hits, the proportions of false alarms, A' , and B'' in the 4 conditions.

One-way ANOVAs showed significant modality differences for the first 3 variables ($P < 0.001$) but not for B'' ($P > 0.20$). Apparently, stimulus recognition for audition was worse than for all other senses, as shown by the low value of A' (0.49) resulting from both a relatively low proportion of hits (0.53) and a large proportion of false alarms (0.54). The proportion of hits was highest for vision, which was associated with the largest value for A' (0.89). For these 2 variables, the means for vision did not differ significantly from the means for touch and smell. The proportion of false alarms, however, was significantly smaller for vision than for all 3 other modalities. The differences in proportions of hits were not due to differences in the number of general

Table 1 Means of the signal detection parameters calculated from the individual recognition data

Modality	<i>N</i>	$P(\text{detail})^a$	$P(\text{general})^b$	$P(\text{hits})^c$	$P(\text{false alarms})$	A'	B''
Vision	20	0.40 ^d	0.41	0.81 ^d	0.16 ^d	0.89 ^d	0.02
Touch	20	0.35 ^{d,e}	0.38	0.73 ^d	0.29 ^e	0.79 ^d	-0.06
Audition	20	0.18 ^e	0.36	0.53 ^e	0.54 ^f	0.49 ^e	-0.06
Smell	20	0.37 ^d	0.39	0.75 ^d	0.31 ^e	0.80 ^d	-0.05

^a $P(\text{detail})$ = proportion of detailed recollections for target stimuli.

^b $P(\text{general})$ = proportion of general recollections for target stimuli.

^c $P(\text{hits}) = P(\text{detail}) + P(\text{general})$.

Means with the same superscripts (d,e,f) were not significantly different in post hoc paired comparisons between modalities with Bonferroni adjustment ($P < 0.05$).

memories but differences in the number of detailed memories, for which audition differed significantly from vision and olfaction; the difference between audition and touch just failed to reach significance ($P = 0.054$).

Given that differences in A' were not statistically significant between vision, touch, and olfaction, the raw results of the relocation task for these 3 modalities can be compared directly in order to evaluate whether object–location binding processes are equivalent. For the auditory modality, we expect that performance in the stimulus relocation task will be relatively low because stimulus recognition was worse than for the other modalities. To enable us to compare object–location binding for all 4 modalities, we will also calculate the outcomes of the relocation task only for those stimuli for which the individual participant reported a detailed recollection in the recognition task.

Stimulus–position binding

For each stimulus, we calculated the proportion of trials in which they were correctly relocated. P_{loc} in the relocation task (Figure 3) differed significantly from chance level (0.10) for all sensory modalities (2-tailed t -tests, $P < 0.05$). This indicates that participants were able to perform the experimental task in all conditions to some degree.

The proportions of correct answers for each stimulus (P_{loc} , P_{side} , and P_{near}) were used as dependent variables in univariate ANOVAs with modality as the independent variable. Although the means for vision tended to be higher than those for the other modalities for all 3 measures (Figure 3), the number of correctly chosen locations P_{loc} yielded no significant effect of modality ($F_{3,36} = 1.1$, $P > 0.20$, $\eta^2 = 0.08$). However, the cruder indexes of correct responses did yield significant differences: The number of stimuli that was placed correctly on the left versus right side of the room P_{side} ($F_{3,36} = 4.1$, $P < 0.05$, $\eta^2 = 0.26$) as well as the number of stimuli that was put in the direct vicinity of the correct location P_{near} ($F_{3,36} = 3.5$, $P < 0.05$, $\eta^2 = 0.23$) showed a significant effect of modality. Post hoc tests with Bonferroni adjustment showed that for P_{side} the visual modality outper-

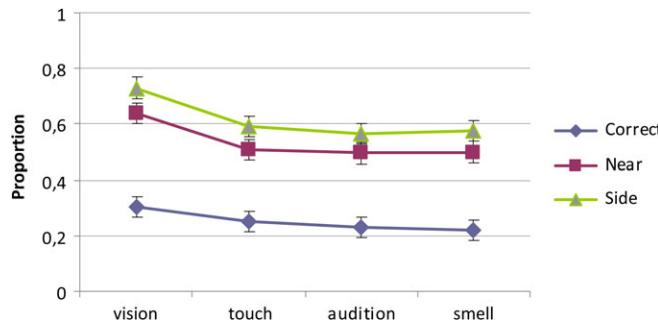


Figure 3 Mean proportion of correct answers (\pm standard error) in the relocation task, using responses for all stimuli. The criterion was the exact correct location (correct), the vicinity of the location (near), or the side of the room (side).

formed the auditory and olfactory modalities ($P < 0.05$) and also tended to outperform the tactile modality ($P = 0.080$). For P_{near} , these differences only showed a tendency toward significance: audition ($P = 0.054$), olfaction ($P = 0.068$), and touch ($P = 0.108$). All differences between audition, olfaction, and touch did not reach significance ($P > 0.20$).

Because stimulus recognition is likely to increase the number of correctly selected locations, we performed this analysis again only for those stimuli that produced a detailed recollection in the recognition task. We expect the modality effects to be more pronounced in this analysis. Because 4 of the 10 auditory stimuli yielded no detailed recollections for any of the participants, we have only 6 observations for audition in this analysis compared with 10 for the other modalities. Again, the means for vision tended to be higher than those for touch, audition, and smell (Figure 4). Nevertheless, P_{loc} just failed to yield a significant effect of modality ($F_{3,32} = 2.2$, $P = 0.109$, $\eta^2 = 0.17$). However, P_{side} ($F_{3,32} = 6.1$, $P < 0.01$, $\eta^2 = 0.37$) as well as P_{near} ($F_{3,32} = 3.9$, $P < 0.05$, $\eta^2 = 0.27$) showed a significant effect of modality. Post hoc tests showed now somewhat more convincingly that for P_{side} the visual modality outperformed the auditory ($P < 0.05$), olfactory ($P < 0.01$), and tactile ($P < 0.05$) modalities. For P_{near} , again all the corresponding differences showed a tendency toward significance: audition ($P = 0.053$), olfaction ($P = 0.071$), and touch ($P = 0.070$). The differences between audition, olfaction, and touch did not reach significance ($P > 0.20$).

Relationships between tasks

For each relocation response ($N = 800$), participants gave a certainty judgment. This measure was likely to be related to the accuracy of the response. Indeed, certainty ratings tended to be higher for correct responses than for incorrect responses ($P_{\text{loc}} = 4.70$, $N = 201$ vs. 3.43 , $N = 599$, $P < 0.001$ in 2-tailed t -test for independent samples; $P_{\text{side}} = 4.02$, $N = 492$ vs. 3.31 , $N = 308$, $P < 0.001$; $P_{\text{near}} = 4.17$, $N = 429$ vs. 3.26 , $N = 371$, $P < 0.001$). We also checked whether the certainty response during the relocation task was higher for stimuli that were correctly chosen in the recognition task. Indeed, stimuli

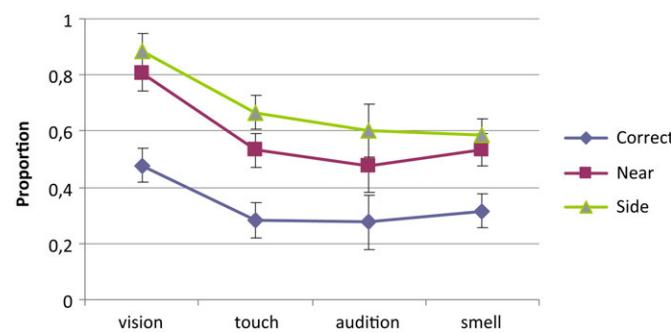


Figure 4 Mean proportion of correct answers (\pm standard error) in the relocation task, using only responses for stimuli with a detailed recollection. The criterion was the exact correct location (correct), the vicinity of the location (near), or the side of the room (side).

that evoked a detailed recollection on average produced higher certainty ratings in the relocation task (4.44, $N = 257$) than stimuli that evoked a general (3.43, $N = 305$) or no recollection (3.42, $N = 238$). In a between-participant univariate ANOVA with the certainty ratings as dependent variable, the ratings for the detailed memories were significantly higher than those for the other 2 response types ($P < 0.001$), which did not differ significantly from each other ($P > 0.20$) in post hoc tests with Bonferroni correction.

General discussion

People may remember the location of an object by remembering its feel, smell, sound, or appearance or any combination of these. The present study set out to compare location memory performance for olfaction with performance for vision, touch, and audition, using unisensory stimuli only. For 3 of the 4 stimulus sets we created, we obtained similar performance in an object recognition task. Equivalence of performance in the recognition task seems an important prerequisite for further analyses because stimuli that are not recognized are unlikely to be relocated correctly. Therefore, in pilot studies, we selected a set of 10 stimuli for each modality that were hard to identify. To obtain these stimuli, we had to reduce the complexity of visual stimuli dramatically, and we used very abstract sounds.

When relocation performance for the different modalities was compared, no differences were found for the number of correctly located stimuli. However, vision outperformed the other modalities when cruder measures of location performance were used (correct side of room or neighboring location). When the analysis was repeated only for stimuli for which participants had a detailed recollection, this effect was somewhat enhanced. This suggests that vision may have an advantage over the other modalities in stimulus relocation tasks. Below, we will further discuss the stimulus recognition task, the stimulus relocation task, and possible practical implications of our findings.

Stimulus recognition

In order to indicate the correct location of a stimulus, the participant should be able to recognize the stimulus and to distinguish it from the other stimuli. If the stimulus is not recognized, it is unlikely that the correct location can be determined. As expected, the present data show that there is a clear relationship between stimulus recognition and relocation performance at the level of separate responses: Correct responses in the relocation task tend to coincide with more detailed recollections in the recognition task for a particular stimulus.

In the present study, we tried to obtain sets of stimuli that would be recognized to the same extent. Nevertheless, the recognition task showed that the auditory stimuli yielded lower values of A' than the other three modalities. A possible methodological complication may be that performance during the recognition task also depends on the characteristics of the distractor stimuli used. Distractor stimuli that resemble

target stimuli are likely to increase the number of false alarms in the recognition task. Although we used an extensive, careful procedure to select the target stimuli for the 4 sensory modalities, it is possible that in the auditory condition distractor stimuli were confused more often with targets than in the other conditions. This could explain why A' in the recognition test was comparatively low for audition, whereas relocation performance was similar to olfaction and touch. Fortunately, we were able to correct for any differences in recognition performance between conditions, by calculating the proportion of correct relocation judgments only for stimuli that yielded a detailed recollection in the recognition task.

Remembering stimulus location

Our study indicates that people are able to link abstract, relatively meaningless sensory stimuli to locations in a space. The number of stimuli that is returned to their exact original location does not differ significantly between the 4 different sensory modalities. For approximate location memory (in the vicinity of the original location or on the same side of the room), however, the analyses suggest that vision outperforms the other modalities.

We wish to argue that vision may have an advantage over other modalities in stimulus relocation tasks because it is also the modality that tends to be used for representing spatial information internally (in visual images) as well as externally (in maps) (Kosslyn 1994). Possibly, vision can more easily gain access to the spatial representation of the room layout in memory. When it comes to the representation of space, people are capable of constructing mental models that contain information about spatial relations and distances between objects (Taylor and Tversky 1992; Bestgen and Dupont 2003). Under experimental conditions, representations of space can be acquired through inputs from various sensory modalities (Shelton and McNamara 2001; Yamamoto and Shelton 2005). However, among sighted individuals, space seems to be primarily represented in visual terms, and the visual system seems especially suitable for dealing with spatial information (O'Connor and Hermelin 1978). Hence, despite the fact that all modalities may contribute to the construction of a mental representation of a space containing different objects, we can thus argue that this representation may generally be accessed most easily in a visual manner (e.g., see Kosslyn 1994, p. 13; Spivey and Geng 2001; Reisberg et al. 2003).

Furthermore, in perceptual tasks, the visual modality appears to be the most suitable for making fine-grained distinctions between spatial locations (e.g., Pick et al. 1969; Warren and Cleaves 1971; Warren et al. 1981). According to the "modality appropriateness" hypothesis (Freides 1974; Welch and Warren 1980), subjects weigh modality inputs according to their relative unimodal performance capabilities for a specific task. Therefore, if the visual modality is better or faster in determining the location of an object under naturalistic conditions, people will tend to rely mainly on the

visual modality for performing this task, as has been demonstrated in many empirical studies (e.g., Warren et al. 1981; Pavani et al. 2000). Even though people's attention may be drawn to an object by perceiving a particular sound or smell or by accidental touching, the visual modality will probably be used subsequently to scan the environment and to determine the exact location of the source.

Possibly, vision outperforms the other modalities in object-location binding tasks because people are most experienced in performing visual relocation tasks. For instance, it may be more common for people to determine where they have last seen a specific animal rather than where they have last heard a particular song. However, whether object-location binding improves with practice is an empirical question that awaits further study, in particular for the nonvisual modalities. Hasher and Zacks (1979) have suggested that object positions are coded automatically into memory due to their ecological significance. In that case, experimental practice in object relocation tasks would be unlikely to improve relocation performance, because it would not add significantly to everyday location memory performance. However, Hasher and Zacks's hypothesis has not been supported consistently (see Postma et al. 2008 for a review).

In our learning task and the relocation task people navigated through the environment using mainly visual and proprioceptive information. There were two reasons for using a uniform visuoproprioceptive spatial coding for stimuli in all modalities. One was to keep the grain of the spatial information the same over all modalities. The other was that in the natural world vision and proprioception are also the prime sources of spatial information. However, the involvement of vision in the navigation part of the task may have provided an advantage during the relocation of the stimuli.

In the present study, participants were not explicitly instructed to memorize the location of the stimuli they perceived because we wanted to mimic a situation in which people encounter unfamiliar stimuli in their environment, which they need to retrace later under unknown conditions. In principle, our results might differ from those obtained in experiments in which explicit memorization instructions were used. However, because location information may be automatically encoded together with information on other attributes when a person pays attention to an object (Hasher and Zacks 1979; Treisman 1998), it may not be necessary to instruct people explicitly to pay attention to the object's location. In fact, several studies have found that the explicit instruction to remember objects locations did not improve object location accuracy (Pezdek and Evans 1979; Ellis 1991; Köhler et al. 2001). Therefore, we expect that explicit instructions to memorize stimulus locations would have yielded similar results.

Possible practical applications

The finding that people can remember different types of landmarks may open up possibilities to design systems that help people to navigate in large buildings or through the streets of

a town. In addition to the graphical wayfinding systems that are typically used in hospitals, airports, and the like (e.g., Mijksenaar 1997), we can think of systems that are (partly) based on sounds, feels, or smells. This may be especially interesting to accommodate people with visual handicaps (e.g., Herssens and Heylighen 2008) or limited cognitive capabilities (e.g., Pagliano 2001). In the latter case, providing sensory information through multiple modalities may help to communicate essential navigation information. For instance, in the Dutch nursing home Elderhoeve for the demented elderly, important crossings in navigation routes in the home that all used to be very similar in appearance, were redesigned starting out from different themes (e.g., silence, rain, theater, beach, art, and music) that were implemented by using clearly distinguishable colors, shapes, smells, and sounds in order to facilitate navigation (Falck and Schaffelaars 1999).

As regards the application of olfaction in human navigation, we would like to refer to Haque (2004), who describes an interactive smell system that allows for the three-dimensional placement of fragrances. The system creates dynamic olfactory zones and boundaries by using computer-controlled fragrance dispensers and careful air control that enable parts of the space to be selectively scented without dispersing the fragrances through the entire space. In this space without physical boundaries, people can explore an invisible smell environment, in which smells travel slowly through space in straight lines, until people mix the fragrances through the movements of their bodies.

Conclusion

The present study shows that people can learn to remember specific locations through their sense of smell. In fact, the ability to remember locations by smell seems comparable to the ability for audition and touch, provided that the stimuli offer as little information and are recognizable to the same degree as the olfactory ones. Only vision retains a small advantage over the other modalities under these conditions. This suggests that object-location binding under natural conditions could be mainly determined by the informative value of a stimulus, and not by the modality through which it is perceived.

Funding

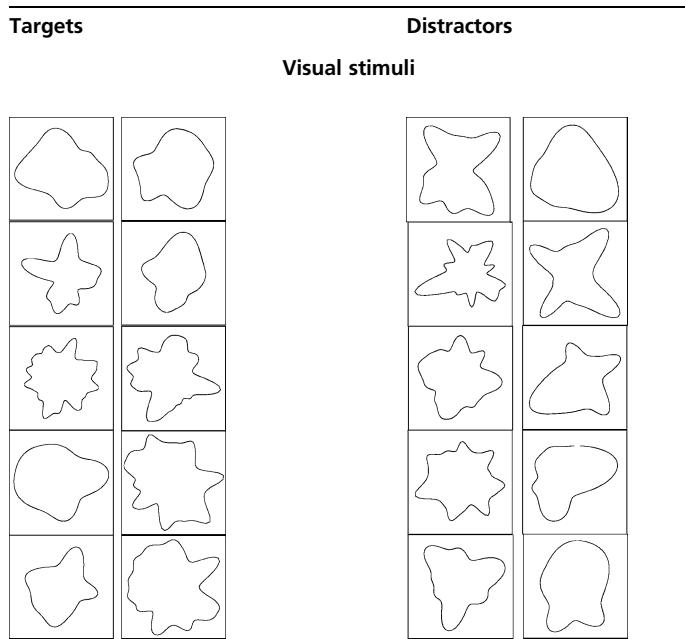
This research was supported by MAGW VIDI grants 452-02-028 and 452-03-334 of the Netherlands Organization for Scientific Research (N.W.O.).

Acknowledgements

The authors thank I.T.C. Hooge, B.M. de Kroon, R. Luxen, and R. van de Velden for technical support; Maarten Streefkerk, Sella Masselink, Mirte van Deursen, and Eva van Eijk for their contribution to several pilot studies and the current experiment; and Cees van Biezen and Peter Traas from Quest International for providing the odorants and the Sorbarods. Maarten Streefkerk is especially acknowledged for designing the SensaCubes.

Appendix

The 40 stimuli used in the relocation task (target) and the additional stimuli used as foils in the recognition task (distractor)



Tactual stimuli

Parasol fabric	Thin foam sheet
Acrylic sheet	Felt
Suede	Hairy cow skin
Wire mesh	Artificial fur
Tempex (isolation material)	Resin sheet (upholstery)
Tissue paper	Foam back of linoleum/vinyl
Isolation material covered with aluminum foil	Carpet (door mat)
Polypropylene	Corduroy
Rubber (back of door mat)	1-cm thick mattress foam
Polystyrene	Cork

Auditory stimuli

Bat	Kiss
Bottle cork popping	Hammer
Starting car	Plane
Lion	Rain II
Rain	Thunder
Chainsaw	Bubbles
Briefcase	Blender
Fireworks	Doctor
Telephone dialing	Train
Duck	Alarm

Table Continued

Olfactory stimuli	
Methyl nonyl ketone	Alcohol C8
Petiole 10%	Ethyl hexanoate
Jasmopyrane forte	Methyl anthranilate 10%
Elinthal 1%	Phenyl ethyl methyl ether 10%
Styrrallyl acetate	Phenyl acetaldehyde dimethyl acetal 10%
Cyclamal 10%	Para cresyl methyl ether
Cytronellyl acetate	Nopylacetate
Ionone alpha	Alicate
Gyrane	Anther
Florocyclene	Pelargene

References

Aaronson D, Watts B. 1987. Extensions of Grier's computational formulas for A' and B" to below-chance performance. *Psychol Bull.* 102:439–442.

Ballas JA. 1993. Common factors in the identification of an assortment of brief everyday sounds. *J Exp Psychol Hum Percept Perform.* 19:250–267.

Bestgen Y, Dupont V. 2003. The construction of spatial situation models during reading. *Psychol Res.* 67:209–218.

Davis RG. 1975. Acquisition of verbal associations to olfactory stimuli of varying familiarity and to abstract visual stimuli. *J Exp Psychol Hum Learn Mem.* 104:134–142.

Degei J, Köster EP. 1999. Odors: implicit memory and performance effects. *Chem Senses.* 24:317–325.

Degei J, Piper D, Köster EP. 2001. Implicit learning and implicit memory for odors: the influence of odor identification and retention time. *Chem Senses.* 26:267–280.

Desor JA, Beauchamp GK. 1974. The human capacity to transmit olfactory information. *Percept Psychophys.* 16:551–556.

Ellis N. 1991. 'Word meaning and the links between the verbal system and modalities of perception and imagery' or 'In verbal memory the eyes see vividly, but ears only faintly hear, fingers barely feel and the nose doesn't know'. In: Logie RH, Denis M, editors. *Mental images in human cognition.* Amsterdam (The Netherlands): North-Holland. p. 313–329.

Falck M, Schaffelaars D. 1999. *Geur & ontwerp (smell & design).* Eindhoven (The Netherlands): [ZOO] Producties.

Frasnelli J, Charbonneau G, Collignon O, Lepore F. 2009. Odor localization and sniffing. *Chem Senses.* 34:139–144.

Freides D. 1974. Human information processing and sensory modality: cross-modal functions, information complexity, memory, and deficit. *Psychol Bull.* 81:284–310.

Grier JB. 1971. Nonparametric indexes for sensitivity and bias: computing formulas. *Psychol Bull.* 75:424–429.

Haque U. 17–19 November 2004. The choreography of sensations: three case studies of responsive environment interfaces. In: Thwaites H, editor.

10th International Conference on Virtual Systems and Multimedia. Softopia, Ogaki City, Japan: Amsterdam (The Netherlands): IOS Press.

Hasher L, Zacks RT. 1979. Automatic and effortful processes in memory. *J Exp Psychol Gen.* 108:356–388.

Heller MA. 1982. Visual and tactile texture perception: intersensory cooperation. *Percept Psychophys.* 31:339–344.

Herssens J, Heylighen A. 2008. Haptics and vision in architecture. In: Lucas R, Mair G, editors. *Sensory Urbanism Proceedings 2008*. Edinburgh (UK): The Flaneur Press. p. 102–112.

Igel A, Harvey LO. 1991. Spatial distortions in visual perception. *Gestalt Theory.* 13:210–231.

Jones B, O'Neil S. 1985. Combining vision and touch in texture perception. *Percept Psychophys.* 37:66–72.

Klatzky RL, Lederman SJ, Matula DE. 1993. Haptic exploration in the presence of vision. *J Exp Psychol Hum Percept Perform.* 19:726–743.

Klatzky RL, Lederman SJ, Metzger VA. 1985. Identifying objects by touch: an "expert system". *Percept Psychophys.* 37:299–302.

Klatzky RL, Lippa Y, Loomis JM, Golledge RG. 2002. Learning directions of objects specified by vision, spatial language, or auditory spatial language. *Learn Memory.* 9:364–367.

Klatzky RL, Lippa Y, Loomis JM, Golledge RG. 2003. Encoding, learning, and spatial updating of multiple object locations specified by 3-D sound, spatial language, and vision. *Exp Brain Res.* 149:48–61.

Klatzky RL, Loomis JM, Lederman SJ, Wake H, Fujita N. 1993. Haptic identification of objects and their depictions. *Percept Psychophys.* 54:170–178.

Kobal G, van Toller S, Hummel T. 1989. Is there directional smelling? *Experientia.* 45:130–132.

Köhler S, Moscovitch M, Melo B. 2001. Episodic memory for object location versus episodic memory for object identity: do they rely on distinct encoding processes? *Mem Cognit.* 29:948–959.

Kosslyn SM. 1994. Image and brain: the resolution of the imagery debate. Cambridge (MA): MIT Press.

Macmillan NA, Creelman CD. 1991. Detection theory: a user's guide. Cambridge (UK): Cambridge University Press.

Mijksenaar P. 1997. Visual function: an introduction to information design. Rotterdam (The Netherlands): 010 Publishers.

Miller GA. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol Rev.* 63:81–97.

Newell FN, Woods AT, Mernagh M, Bühlhoff HH. 2005. Visual, haptic and crossmodal recognition of scenes. *Exp Brain Res.* 161: 233–242.

Nunn JA, Polkey CE, Morris RG. 1998. Selective spatial memory impairment after right unilateral temporal lobectomy. *Neuropsychologia.* 36: 837–848.

O'Connor N, Hermelin B. 1978. Seeing and hearing and space and time. London (UK): Academic Press.

Pagliano P. 2001. Using a multisensory environment, a practical guide for teachers. London (UK): David Fulton Publishers.

Pasqualotto A, Finucane CM, Newell FN. 2005. Visual and haptic representations of scenes are updated with observer movement. *Exp Brain Res.* 166:481–488.

Pavani F, Spence C, Driver J. 2000. Visual capture of touch: out-of-the-body experiences with rubber gloves. *Psychol Sci.* 11:353–359.

Pezdek K, Evans GW. 1979. Visual and verbal memory for objects and their spatial locations. *J Exp Psychol Hum Learn Mem.* 5: 360–373.

Pick HL, Warren DH, Hay JC. 1969. Sensory conflict in judgments of spatial direction. *Percept Psychophys.* 6:203–205.

Porter J, Craven B, Khan RM, Chang SJ, Kang I, Judkewicz B, et al. 2007. Mechanisms of scent-tracking in humans. *Nat Neurosci.* 10:27–29.

Postma A, De Haan EHF. 1996. What was where? Memory for object locations. *Q J Exp Psychol.* 49A:178–199.

Postma A, Kessels RPC, van Asselen M. 2008. How the brain remembers and forgets where things are: the neurocognition of object–location memory. *Neurosci Biobehav Rev.* 32:1339–1345.

Rajamaran S. 1993. Remembering and knowing: two means of access to the personal past. *Mem Cognit.* 21:89–102.

Reisberg D, Pearson DG, Kosslyn SM. 2003. Intuitions and introspections about imagery: the role of imagery experience in shaping an investigator's theoretical views. *Appl Cogn Psychol.* 17: 147–160.

Schifferstein HNJ, Cleiren MPH. 2005. Capturing product experiences: a split-modality approach. *Acta Psychol.* 118:293–318.

Schneider BA, Schmidt CE. 1967. Dependency of olfactory localization on non-olfactory cues. *Physiol Behav.* 2:305–309.

Shelton AL, McNamara TP. 2001. Visual memories from nonvisual experiences. *Psychol Sci.* 12:343–347.

Siegel AW, White SH. 1975. The development of spatial representations of large-scale environments. In: Reese HW, editor. *Advances in child development and behavior.* Vol. 10. New York: Academic Press. p. 9–55.

Spivey MJ, Geng JJ. 2001. Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychol Res.* 65: 235–241.

Takahashi M. 2003. Recognition of odors and identification of sources. *Am J Psychol.* 114:527–542.

Taylor HA, Tversky B. 1992. Spatial mental models derived from survey and route descriptions. *J Mem Lang.* 31:261–292.

Treisman A. 1998. Feature binding, attention and object perception. *Philos Trans R Soc Lond B Biol Sci.* 353:1295–1306.

von Békésy G. 1964. Olfactory analogue to directional hearing. *J Appl Physiol.* 19:369–373.

Warren DH, Cleaves W. 1971. Visual-proprioceptive interaction under large amounts of conflict. *J Exp Psychol.* 90:206–214.

Warren DH, Welch RB, McCarthy TJ. 1981. The role of visual-auditory "compellingness" in the ventriloquism effect: implications for transitivity among the spatial senses. *Percept Psychophys.* 30: 557–564.

Welch RB, Warren DH. 1980. Immediate perceptual response to intersensory discrepancy. *Psychol Bull.* 88:638–667.

Yamamoto N, Shelton AL. 2005. Visual and proprioceptive representations in spatial memory. *Mem Cognit.* 33:140–150.