Chapter 8
Representing 3D Space in Working Memory: Spatial Images from Vision, Hearing, Touch, and Language

Jack M. Loomis, Roberta L. Klatzky, and Nicholas A. Giudice

Abstract The chapter deals with a form of transient spatial representation referred to as a spatial image. Like a percept, it is externalized, scaled to the environment, and can appear in any direction about the observer. It transcends the concept of modality, as it can be based on inputs from the three spatial senses, from language, and from long-term memory. Evidence is presented that supports each of the claimed properties of the spatial image, showing that it is quite different from a visual image. Much of the evidence presented is based on spatial updating. A major concern is whether spatial images from different input modalities are functionally equivalent—that once instantiated in working memory, the spatial images from different modalities have the same functional characteristics with respect to subsequent processing, such as that involved in spatial updating. Going further, the research provides some evidence that spatial images are amodal (i.e., do not retain modality-specific features).

Keywords Amodal • Functional equivalence • Hearing • Language • Spatial image • Spatial updating • Touch • Vision • Working memory

---

J.M. Loomis (✉)
Department of Psychological and Brain Sciences, University of California,
Santa Barbara, CA 93106, USA
e-mail: Jmloomis99@gmail.com

R.L. Klatzky
Department of Psychology, Carnegie-Mellon University,
Pittsburgh, PA 15213, USA

N.A. Giudice
Spatial Informatics Program: School of Computing and Information Science,
University of Maine, Orono, ME 04469, USA

S. Lacey and R. Lawson (eds.), *Multisensory Imagery*,
DOI 10.1007/978-1-4614-5879-1_8,

8.1 Introduction

This chapter is concerned with what we call the "spatial image," a spatial representation that is relatively short-lived and, as such, resides within working memory. It plays an important role in the control of action in three-dimensional (3D) space when task-relevant perceptual information is no longer present (e.g., Tatler and Land 2011). Other researchers have used a wide variety of tasks, some dealing with action and others not, and different names to refer to the same or similar short-lived, action-related memory representations [e.g., egocentric model (Tatler and Land 2011), egocentric representation (Burgess 2008; Mou et al. 2004), sensorimotor location codes (May 2004; Kelly et al. 2007), spatial mental model (Taylor and Tversky 1992), and on-line transient representation (Waller and Hodgson 2006)]. All of these terms are meant to contrast with more enduring spatial representations in long-term memory (e.g., Amorim et al. 1997; Avraamides and Kelly 2008; Burgess 2006; Byrne et al. 2007; Easton and Sholl 1995; Huttenlocher et al. 1991; McNamara 2003; Mou et al. 2004; O'Keefe and Nadel 1978; Tolman 1948; Waller and Hodgson 2006; Wang and Spelke 2000). We have found it useful to refer to the underlying phenomenon as the spatial image, a term that is both more specialized and more evocative than the term "spatial representation."[1]

To give some idea of what we will be discussing, we ask the reader to engage in the following exercise, assisted by someone else. In your current location, look around and note three or four identifiable objects within a short walking distance. Close your eyes and begin walking forward. The other person will randomly choose which object is the goal and when you should turn. On command of the other person, turn toward the specified goal and walk a few steps in its direction. In opening your eyes, you can judge the accuracy with which you were spatially updating the location of the goal by noting how aligned your facing direction is with it. Similar exercises can be done with auditory targets and objects haptically sensed with a long pole. For most people, the spatial image representing the object locations is introspectively less vivid than visual imagery of familiar faces and familiar locations. In past research, we have used the term spatial image to refer to a single location (e.g., Loomis et al. 2002), but research indicates that the contents of spatial working memory can represent multiple point locations, simple paths, and oriented objects. Accordingly, we use spatial image to refer to the contents of spatial working memory, representing any of these possibilities. We hypothesize that even multiple surfaces, like those of a room, can be simultaneously represented in spatial working memory.

Our idea of the spatial image is closely connected to the more familiar idea of a percept (perceptual representation). Starting with the knowledge that perception

---

[1] Röder and Rösler (1998) used the term "spatial image" in connection with the study of internal scanning of images from both vision and touch. Because their study focused on 2D maps, their use of the term seems to refer to a concept narrower than ours. Conversely, in their review of behavioral and neural research on mental imagery, Struiksma et al. (2009) use "spatial image" as a more encompassing term than we do.

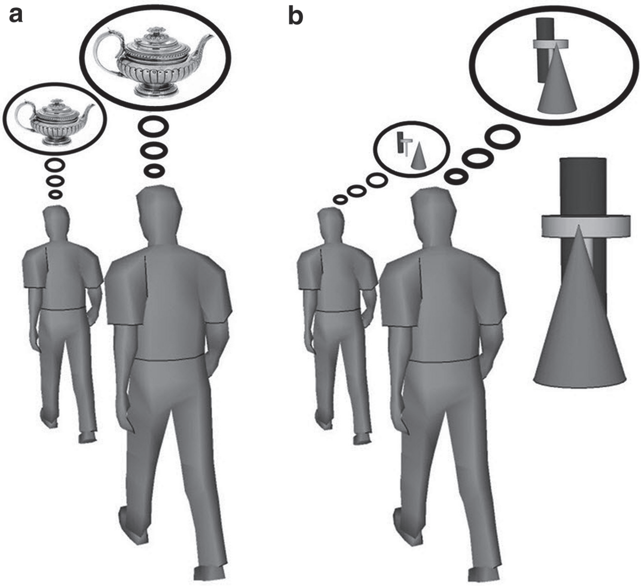

Fig. 8.1 Depiction of some differences between a visual image (**a**) and a spatial image (**b**). (**a**) As a person walks while imagining an object, the image remains more or less in front of the person and any changes in the visual image are not influenced by locomotion per se. (**b**) A person views a configuration of three objects and then closes the eyes. As the person then walks, the externalized spatial image of the objects corresponds spatially with percepts that would be there were the eyes still open. Accordingly, the objects represented within the spatial image vary in direction and exhibit parallax as the person walks forward

involves a causal chain involving sensory transduction, processing within the sensory pathway, and cortical brain activation, we accordingly adopt the view that the world we see, hear, and feel in everyday life is not the physical world itself but is instead a perceptual representation (e.g., Koch 2003; Lehar 2003; Loomis 1992; Russell 1948; Smythies 1994). For us, the percept and spatial image of object locations are similar representations in part because their referents appear external to the observer. Indeed, research supports the assumption that the spatial image produced by an external stimulus is spatially congruent with the resulting percept. Sometimes the perceptual location of an object is inaccurate with the result that the spatial image inherits this error (e.g., Philbeck et al. 1997).

We contrast the spatial image with a visual image that is depictive. This type of visual imagery is widely considered to retain properties of the visual percept such as color, texture, and shape; is picture-like in the sense of not exhibiting parallax; remains fixed within imaginal space as the person is walking or riding in a vehicle (Fig. 8.1a); and is experienced in the anterior half of the surrounding space. In our conception, the spatial image is experienced as external to the person's head and body and retains distance and direction information, such that as the observer spatially updates while translating, different components of the spatial image exhibit relative parallax (Fig. 8.1b). In addition, the spatial image is not confined to anterior space but can exist in all directions around the observer, much in the way that auditory percepts do. More importantly, the spatial image is fundamentally multisensory in origin, for it can be instantiated in spatial working memory by visual, auditory, and haptic stimulation, as well as by spatial language. Whether spatial images retain features of the sensory modality from which they arose or are amodal in nature is a current topic of investigation. Finally, spatial images can be instantiated by recall of spatial layout from long-term memory. We present evidence for these claims in a subsequent section.

Our discussion has focused on differences between the spatial image and a depictive visual image, because theories of mental imagery have predominantly been concerned with the visual modality. However, as the present volume demonstrates, there is also interest in modality-specific imagery relating to hearing, touch, and movement. In distinguishing these from the spatial image, it is useful to consider three criteria put forward to demonstrate modality-specific imagery (Stevenson and Case 2005): (a) the subjective experience resembles the percept; (b) the effects of imagery in the modality mimic those of perception; (c) memory-based images can interact with perception in the modality. As we will make clear in the sections that follow, the spatial image can support behaviors that arise from perceptual sources, like spatial updating, and a spatial image recalled from memory can interact with one formed from perception. Thus, because criteria (b) and (c) can be met by both the spatial image and a modality-specific image that conveys information about spatial layout, they do not provide a means of distinguishing the two. However, the spatial image *is* distinguished by the fact that its content is not specifically modal; for example, a spatial image formed from hearing does not convey modality-specific content such as timbre. Spatial images can integrate inputs from multiple perceptual and cognitive channels; they transcend modality-specific features while retaining the spatial information shared by the inputs.

8.2 A Conceptual Framework

The functional block diagram in Fig. 8.2 provides a conceptual framework for spatial images and their transformation during spatial updating. The spatial modalities of vision, hearing, and touch, each encode stimulus information from one or more locations and output spatial percepts. For each modality, the spatial image is assumed to be spatially congruent with the percept within representational space, but of much lower precision and complexity. When the stimulus terminates, subsequently resulting in termination of the percept, the spatial image continues to exist for some short duration. Spatial images also can be instantiated within spatial working memory by way of inputs from spatial language and from long-term memory.

The lower part of the diagram shows the subsystem that provides input to the spatial updating process. Real rotations and translations of the observer result in the perception of self-motion, which, as the result of path integration, yields estimates

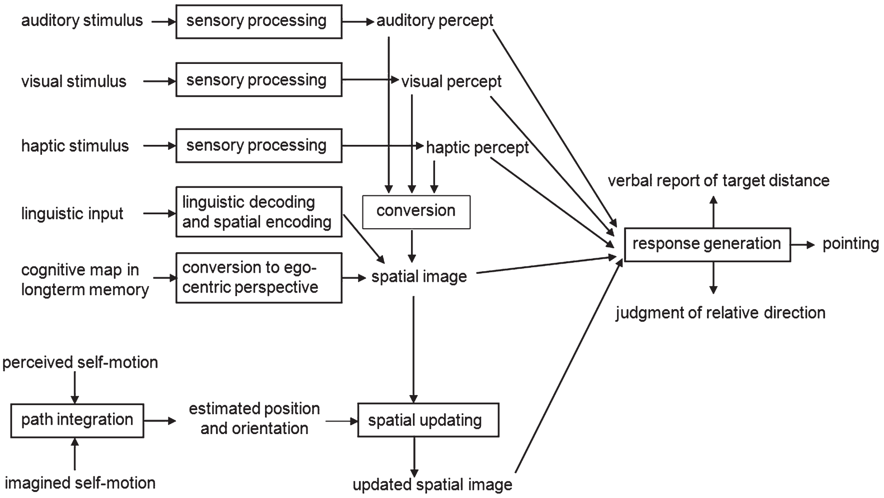

Fig. 8.2 Functional block diagram for the conceptual framework described in the text. Sensory inputs from vision, hearing, and touch give rise to percepts as well as spatially congruent spatial images. When the stimuli are removed, the percepts subsequently cease, but the spatial images remain. Spatial images can also be created by language and recalled from long-term memory. The lower section shows how both perceived and imagined self-motion can lead to a change in the observer's estimated position and orientation, which in turn can lead to spatial updating of the spatial image. The section on the right represents response generation. A wide variety of spatial judgments, several of which are shown, can be made on the basis of concurrent percepts or concurrent spatial images

of current position and orientation. The perception of self-motion is based on inputs such as optic flow, acoustic flow, inertial cues (e.g., vestibular), haptic cues, proprioceptive cues, and in the case of vehicular motion, feed-forward estimates of vehicular velocity based on an internal model of the dynamics of the vehicle (Loomis and Beall 2004). Another input to the path integration process is imagined self-motion, although it appears to be much weaker than perceived self-motion (Klatzky et al. 1998). The estimates of current position and orientation serve as input to the spatial updating process, which modifies the spatial image within spatial working memory (Byrne et al. 2007; Wiener et al. 2010). The updated spatial image provides estimates of the current locations and orientations of targets that were initially perceived. These estimates can be used to control locomotion relative to the targets. Not shown in the diagram are perceived changes in body posture during reaching and grasping that are signaled by proprioception and efference copy. These are involved in spatial updating at the scale of personal space.

The section of the diagram to the right deals with non-locomotor responses that can be executed based on percepts and/or spatial images. These responses include throwing balls at targets (perceived or updated), verbal estimates of distance and direction, and more complex judgments of the spatial layout of multiple targets, such as judgments of relative direction (JRDs).

8.3 Properties of the Spatial Image

8.3.1 Spatial Images Are Externalized like Percepts in 3D Representational Space

Like percepts from vision, hearing, and touch, spatial images are experienced as entities external to the head and body. This claim is supported by numerous spatial updating experiments in which the observer views an object location, closes the eyes, and then rotates and/or translates while mentally keeping track of the target's perceived location. Responses in such updating experiments include walking to the location of the updated target (e.g., Loomis et al. 1992; Philbeck and Loomis 1997; Rieser et al. 1990; Thomson 1983; Thompson et al. 2004), pointing toward it (e.g., Fukusima et al. 1997; Loomis et al. 1992; Riecke et al. 2005; Siegle et al. 2009), redirecting gaze toward it (Medendorp et al. 2003), throwing a beanbag (Sahm et al. 2005; Thomson 1983), walking and then gesturing with the hand (Ooi et al. 2001, 2006; Wu et al. 2004), and making a verbal report of its direction and distance (e.g., Klatzky et al. 2003). Auditory and haptic updating of single targets has also been demonstrated [e.g., audition (Ashmead et al. 1995; Loomis et al. 1998) and touch (Hollins and Kelley 1988)]. Figure 8.3 shows updating performance for visual and auditory targets situated 3 and 10 m away (Loomis et al. 1998). On some trials, observers walked directly to targets after viewing or hearing them, and on other trials, they were guided 5 m forward to a turn point, after which they walked unaided to the updated target locations. The near congruence of the centroids of the stopping points for direct and indirect paths for each target, especially for vision, indicates that updating is quite accurate for the average observer. The fact that the auditory responses were closer than the far targets and further than the near targets is consistent with the claim that the observers misperceived the distance of auditory targets and that the spatial image guiding the behavior inherited the perceptual error.

Figure 8.4 gives another example of spatial updating when perceptual errors are present (Ooi et al. 2001, 2006; Wu et al. 2004). In this case, when observers view a glowing target in an otherwise dark room, targets greater than 3 m away are perceived as closer but in the correct direction. When the target is positioned on the ground, the percept appears off the ground.

Most of the studies cited above involved updating a single target location, but research has also shown that people can update multiple target locations (e.g., Klatzky et al. 2003; Loarer and Savoyant 1991; Loomis et al. 1998; May 2004; Rieser 1989; Rieser and Rider 1991) as well as simple paths (e.g., Pham and Hicheur 2009). Figure 8.5 shows the mean walking trajectories from a fixed origin to oriented arrows on the ground plane, with visual guidance and without visual guidance following visual preview (Pham and Hicheur 2009). The figure indicates very similar walking trajectories with and without vision for three different targets. In another study, Loarer and Savoyant (1991) showed that after observers viewed several vertical columns at different distances and then walked with eyes closed,

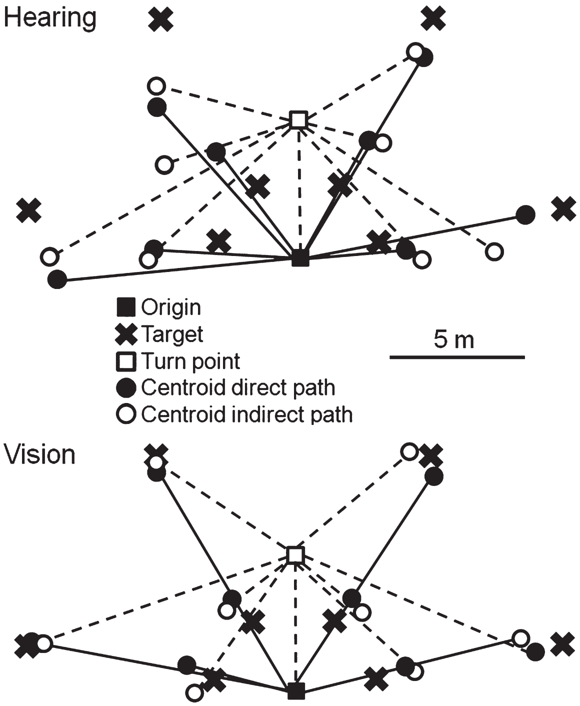

Fig. 8.3 Results of Experiment 3 of a study by Loomis et al. (1998). Observers were presented with auditory and visual targets at varying azimuths and at distances of either 3 or 10 m. On a given trial, after the target was presented, the observer attempted to walk without vision to its location either along a direct path or after being led forward 5 m to the turn point. The *open* and *closed circles* represent the centroids of the stopping points for the direct and indirect paths, respectively. The near congruence of the direct and indirect centroids indicates that spatial updating is quite accurate on average. Auditory targets were generally misperceived in distance, as indicated by the discrepancy between the target positions and the corresponding centroids. This is a modified version of Fig. 7 from Loomis et al. (1998) and is published here with permission of *Attention, Perception, & Psychophysics*

their responses indicated that the directional ordering of the columns changed as the observer approached, reflecting the parallax changes that would have been apparent were the eyes open. A similar conclusion is indicated by the results of Amorim et al. (1997), showing that observers are able to update both the location and orientation of an object.

Other experiments have shown updating with pure observer rotation (e.g., Farrell and Robertson 1998; May 2004; Presson and Montello 1994; Riecke et al. 2005; Rieser 1989; Waller and Hodgson 2006), but these are less diagnostic about externalization of the spatial image than are translational tasks. With translation, the location of the updated spatial image changes in its direction during travel as a function of its distance, and updating research shows that the response is exquisitely sensitive to target distance. In contrast, observer rotation induces changes in direction that are

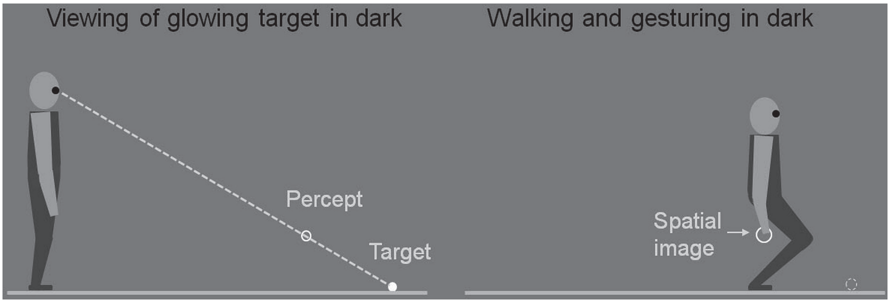

Fig. 8.4 Depiction of spatial updating of a target on the ground. When a dimly glowing target light is placed on the ground in a dark room more than 2 m away, it is perceived as closer than it actually is. Ooi et al. (2001, 2006; Wu et al., 2004) used a novel response in which the observer walked forward in the dark and then gestured with the hand to indicate the remembered location (spatial image) of the target. The results showed that the indicated location was closer than the target but along the same line of direction as viewed from the origin, indicating correct visual perception of direction but under-perception of distance. These results are consistent with a spatial image being formed at a location congruent with the percept and accurate updating of the spatial image during locomotion

independent of distance so rotational updating tasks are scale independent and can be performed more easily without an externalized spatial representation.

With many of the translational updating tasks involving a single target, there is an obvious alternative to the hypothesis of an externalized spatial image: execution of a preprogrammed action. The idea is that while viewing a target location or path to follow, an observer preprograms a motor response and upon closing the eyes executes the response. While some of the reported studies might involve this strategy, other experiments strongly support the hypothesis that observers update a spatial image by showing that observers can modify their actions in the midst of updating (e.g., Farrell and Thomson 1999; Fukusima et al. 1997; Loomis et al. 1998; Philbeck et al. 1997; Thomson 1983). To illustrate, we briefly describe the results of Fukusima et al. (1997), using a task similar to the exercise described at the beginning of the chapter. After viewing a single target, the observer walked along an oblique path; then, on instruction, turned and walked toward the target. Because performance was accurate even when a preprogrammed response was precluded by the observer's not knowing when the turn would occur, the hypothesis of updating with respect to an externalized spatial image is supported.

In theory, spatial images can arise from any activity that gives rise to a percept. Accordingly, beyond spatial images being associated with normal visual, auditory, and haptic perception, we would expect them to arise from specialized forms of perception like feeling targets with a cane or probe, echolocation based on reflected sound, and tactile and auditory substitution of vision. The most interesting cases are those in which perception based on short periods of sensing is followed by actions revealing perceptual localization of the targets, for it is in these cases that spatial working memory and putative spatial images are implicated. Bennett et al. (2011) conducted an experiment in which observers felt vertical poles of different heights

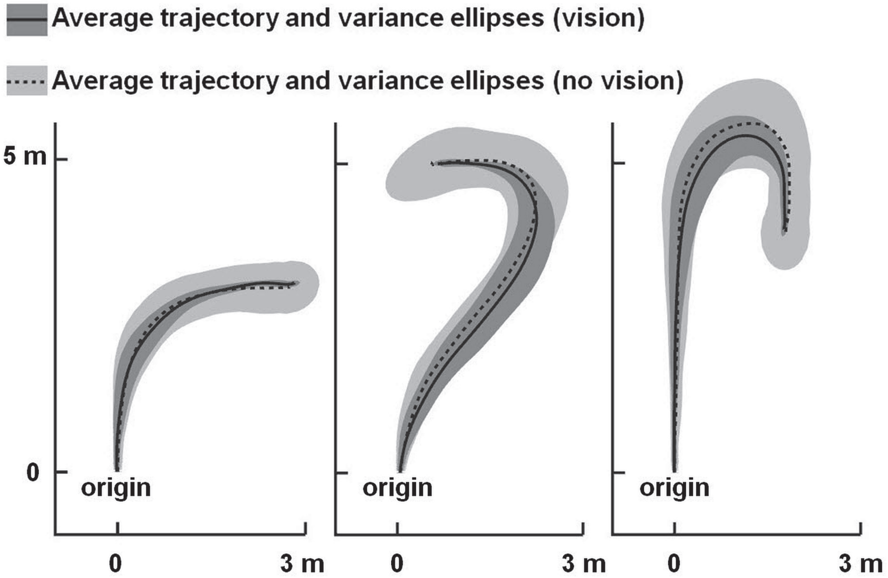

Fig. 8.5 Illustrative results from a study by Pham and Hicheur (2009) in which observers first viewed arrows placed on the ground plane at different distances and directions from the origin and then walked with or without vision so as to proceed along the length of the *arrow* (not depicted) and stop right at the tip of the *arrowhead*. The three panels show the responses to three of the many stimuli used. The average walking trajectories were very similar for the vision and no vision conditions. As expected, variability decreased in the vision condition as the observers neared the end of the *arrow*; whereas, variability remained high in the no vision condition. Other analyses showed that the velocity profiles were also very similar for the vision and no vision conditions. This figure is a modification of Fig. 3 from Pham and Hicheur (2009) and is published with permission of the American Physiological Society

at different distances and directions. Sensing was performed with extended touch using a 1-m probe, a 0.5-m probe, the bare hand, and vision, with the short probe and hand conditions requiring the observer to step forward to contact the target. Immediately afterward, the target was removed, and observers sidestepped from the origin and then moved forward and gestured with the hand to indicate its location. With this measurement procedure, which is based on triangulation, performance was very accurate in all four conditions, indicating mediation by an externalized spatial image corresponding to the target location.

There has been considerable research on echolocation and tactile and auditory substitution of vision, but the tasks employed so far allow for concurrent sensing of the targets while making judgments and thus are not dependent on spatial working memory [e.g., echolocation (Gordon and Rosenblum 2004; Hughes 2001; Teng et al. 2012) and sensory substitution (Auvray and Myin 2009; Chebat et al. 2011)]. Echolocation would, though, be another good way to test for localization mediated by a spatial image. For example, a large, reflecting target could be presented and the observer would sense its location using echolocation. After its removal, the observer would attempt to walk to its location.

8.3.2 *Spatial Images Exist in All Directions*

Unlike depictive visual images, which likely appear only in directions forward of the head, spatial images exist in all directions. Evolutionarily, this makes sense, for once a spatial image has been formed from visual input, if it is to be useful for action, it needs to continue to represent the same environmental locations despite rotations and translations of the head. Furthermore, because hearing and touch give rise to percepts in all directions about the head and body, the resulting spatial images must be omnidirectional. There is an abundance of evidence supporting the omnidirectionality of spatial images—a large number of studies show that people can update locations in all directions around the body during rotations and translations (e.g., Easton and Sholl 1995; Farrell and Robertson 1998; Giudice et al. 2011; Loomis et al. 1998; May 2004; Mou et al. 2004; Presson and Montello 1994; Riecke et al. 2005; Rieser 1989; Rieser et al. 1986; Waller and Hodgson 2006; Waller et al. 2002). With this evidence that spatial images exist and can be updated in all directions, a more interesting question is whether updating performance is better in front than behind. Horn and Loomis (2004) conducted an experiment to examine this question by comparing performance on a task in which the previously viewed target was either in front of or behind the observer during the updating phase. The observer viewed a target at one of various locations in an open field and then turned to face or face away from the target with eyes closed. The observer then sidestepped several meters and attempted to face the updated target. Two performance measures (mean signed angular error and within-observer variability) showed no reliable differences between updating in front and behind, and the third (mean absolute angular error) showed only slightly poorer performance behind (14.8° behind vs. 12.6° in front). This direct comparison of updating in front and behind shows that updating performance is performed well in both directions with minimal differences between them.

8.4 Functional Equivalence and Amodality

The above mentioned research indicates that spatial images based on visual, auditory, and haptic input can be updated. Other research has established that people can form spatial representations from linguistic descriptions of a scene and make spatial judgments similar to those produced while viewing or recalling that scene (e.g., Avraamides 2003; Avraamides and Kelly 2010; Bryant et al. 1992; Denis and Cocude 1989; De Vega and Rodrigo 2001; Franklin and Tversky 1990; Shelton and McNamara 2004; Struiksma et al. 2009; Taylor and Tversky 1992; Zwaan and Radvansky 1998). In a similar vein, Lyon and Gunzelmann (2011) found that visual depiction of movement along a 3D path conveyed with a first-person perspective and verbal description of movement along the same path resulted in nearly identical spatial judgments about the path. These research findings suggest that language can give rise to spatial images. Stronger evidence comes from experiments showing that spatial updating can be performed with respect to targets specified by language

(Klatzky et al. 2003; Loomis et al. 2002). Thus, the evidence supports the conceptual framework presented earlier (Fig. 8.2) in which the senses and language, all can give rise to spatial images (see also Bryant 1997).

A major part of the research done by our group has been concerned with whether the spatial images formed from vision, hearing, touch, and language exhibit "functional equivalence" (Loomis and Klatzky 2007). This refers to the hypothesis that once sensory or linguistic inputs are encoded as a spatial image, subsequent image-mediated behaviors depend only on the properties of that image, and not the source modality. Our research, described below, has demonstrated performance that is equivalent, or nearly so, across a range of spatial tasks involving different sensory modalities.

Assuming that functional equivalence holds, there are three interpretations in terms of underlying mechanisms (Giudice et al. 2011). The first of these, the separate-but-equal hypothesis, posits that equivalent spatial behavior across different inputs arises from modality-specific spatial representations that are isomorphic. Spatial isomorphism is not itself sufficient; what is further required are processes, either modality specific or modality general, that support equivalent processing outcomes. For example, there might be different mechanisms for calculating the Euclidean distance within auditory and visual representations, but if the representations are spatially isomorphic and the processes do not differ in accuracy, functional equivalence is guaranteed. A fundamental problem with this hypothesis, however, is that it suffers from a lack of explanatory power, as it offers no general principle by which modality-specific images would result in equivalent performance across modalities.

The second interpretation, the common-recoding hypothesis, postulates that inputs from multiple modalities are recoded into a single, modality-specific representational format. For example, all spatial inputs could be converted into 3D visual representations in memory (Newell et al. 2005).

The third interpretation, the amodal hypothesis, postulates that functional equivalence arises when information from all modalities converge onto a common spatial image that does not retain any modality-specific information (for a related term, metamodal, see Pascual-Leone and Hamilton 2001). Bryant (1997) proposed essentially this hypothesis with his idea of a spatial representation system (SRS), which provides a common format for the different input modalities of vision, hearing, touch, and language (see also Struiksma et al. 2009).

8.4.1 *Evidence for Functional Equivalence in Spatial Updating*

Spatial updating is a good way of testing for functional equivalence. In our work, we have shown near functional equivalence for updating tasks using combinations of vision, hearing, touch, and spatial language. In comparisons between different sensory modalities, especially with the goal of addressing functional equivalence, effort must be taken to match the perceptual representations by adjusting for known sensory-specific encoding biases (e.g., Klatzky et al. 2003). For example, hearing often results in greater errors in distance perception compared to vision (e.g., Fig. 8.3). Failure to take this into account makes it difficult to interpret whether differences in test performance are due to differing perceptual errors during encoding or to fundamental differences in the spatial representation.

When we began this line of research, we wished to know if people could form a spatial image from a simple utterance specifying direction and distance, successfully update that image while walking, and perform with the same level of accuracy as with input from vision and hearing. This prediction was tested and supported in two studies (Klatzky et al. 2003; Loomis et al. 2002). In the 2002 study, locations were specified by language or hearing, after which observers immediately attempted to walk without vision to the specified location; in the 2003 study, observers learned multiple locations specified by vision, hearing, or language and later recalled one of the learned locations prior to walking to it without vision. We discuss only the latter study, because it adjusted for differences in distance encoding between vision, hearing, and language. There were two experiments, one involving the use of a pointer and verbal estimates to indicate the estimated target locations and the other involving blind walking to the estimated target locations. Because an analysis by Loomis and Philbeck (2008) showed that verbal reports are biased toward underestimation, we focus on the latter experiment.

Figure 8.6 gives the spatial layouts for the three modalities used in the experiment. The vision and language targets were at the same nominal locations, ranging in distance from 0.9 to 3.7 m. Because of the tendency for indoor auditory targets to be perceived as closer than they were physically, auditory stimuli were presented using loudspeakers placed at slightly larger distances, as shown, so as to produce perceptual locations close to those of the visual condition. In the learning phase, observers in the hearing condition heard synthetic speech labels (e.g., “baby,” “horse”) presented by the loudspeakers. In the language condition, observers heard synthetic speech giving the target coordinates followed by the label. In the vision condition, observers saw labels presented at eye level. Observers learned the target locations and then, when prompted with a label, attempted to report their directions and distances using a pointer and verbal report, respectively. The learning phase terminated when accuracy of both pointing and distance reports met stringent criteria. In the test phase, observers responded to each target label by walking either directly to the target or walking indirectly to the target after being passively guided to the turn point, either in front of or to the side of the origin. Of interest was the amount of updating error indicated by a difference in terminal locations for the direct and indirect paths.

Figure 8.6 gives the centroids of the terminal points for direct and indirect walks. The small separations between direct and indirect centroids in all six panels of Fig. 8.6 indicate that updating performance was good for all conditions. Language produced slightly larger updating errors than the two perceptual modalities, which did not reliably differ. The experiment demonstrated that spatial images can be formed from language as well as from auditory and visual perception, that spatial updating occurs for all modalities, and that the spatial images of vision and hearing exhibit functional equivalence in spatial updating, with those of language exhibiting near functional equivalence.

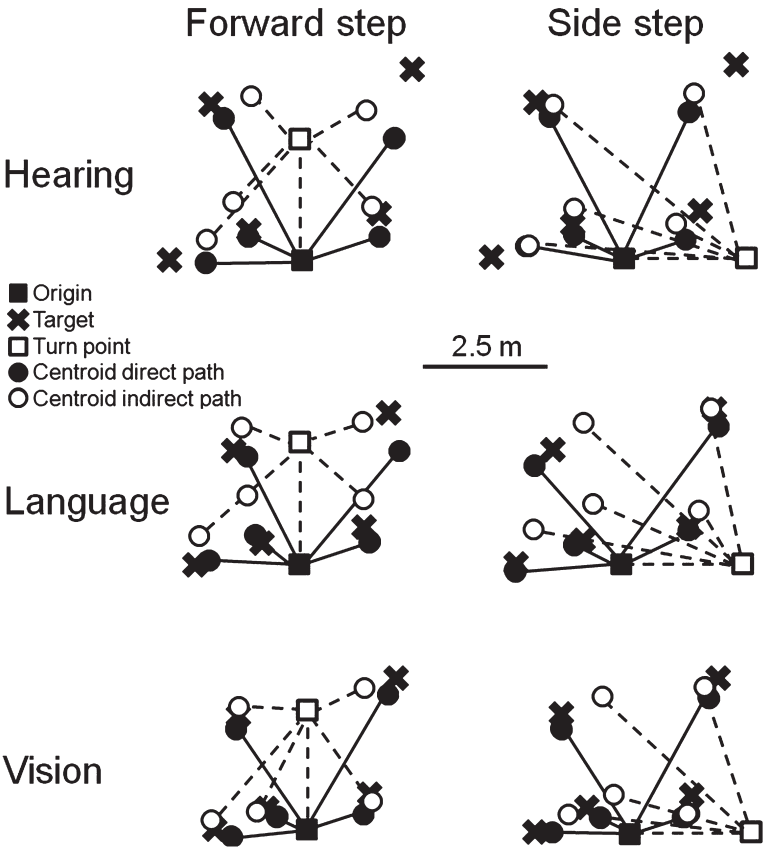

Fig. 8.6 Partial results of Experiment 2 of Klatzky et al. (2003). Observers were presented with targets specified by hearing, vision, or spatial language. The targets (X) varied in distance and direction. The auditory stimuli were placed further away to compensate for the expected underperception of distance by hearing. Observers learned the locations of three or five targets. During testing, the observer recalled the location of a specified target and then attempted to walk without vision to its location, proceeding (1) along a direct path, (2) along an indirect path after being led forward 2.5 m, or (3) along an indirect path after being led 2.5 m to the right (sidestepping). The *open* and *closed circles* represent the centroids of the stopping points for the direct and indirect paths, respectively. The near congruence of the direct and indirect centroids indicates that spatial updating is quite accurate on average in all conditions. This figure is based on Figs. 4 and 5 from Klatzky et al. (2003) and is published here with the permission of *Experimental Brain Research*

8.4.2 Evidence for Functional Equivalence in Allocentric Judgments

The above studies evaluated functional equivalence in connection with egocentric judgments within the context of spatial updating, but the spatial image concept is by no means limited to that context. Another way to test for functional equivalence of spatial images is to examine patterns of error and response times in other forms of spatial judgment. The study by Avraamides et al. (2004) evaluated functional equivalence of spatial images encoded from vision and spatial language, this time in connection with allocentric judgments in which observers indicated the relative distance and direction from one target to another. Observers learned indoor layouts of four targets in both vision and language conditions to an accuracy criterion. The allocentric reports then followed; on each trial, two targets were identified, and the observer reported the inter-target distance and direction. Distance was reported verbally, and direction was reported by aligning a pointer with the direction from one target to the other. Strong evidence of functional equivalence was obtained from the correlated patterns of response latencies, distance errors, and direction errors.

More recently, we investigated functional equivalence of spatial images built up from touch and vision between blind and sighted observers, this time at the scale of a table top (Giudice et al. 2011). Observers learned simple route maps (four points connected by three straight-line segments) by either seeing or feeling them. After learning, they performed judgments of relative direction (JRDs). These experiments exploited a well-known phenomenon in the spatial cognition literature known as the "alignment effect"—spatial representations are accessed more rapidly and more accurately when responses require imagining the environment from the same (aligned) orientation as at learning than from other (misaligned) orientations (e.g., Waller et al. 2002). The studies showed that alignment biases that have been commonly demonstrated for visual map learning (e.g., Levine et al. 1982) also occur with haptic map learning. Importantly, spatially updating the learned map in working memory while moving around it can induce a shift in the alignment effect, such that what was easy before movement becomes difficult after movement and vice versa (Harrison 2007; Waller et al. 2002). In the second experiment of this study, such a movement condition was included. Figure 8.7 shows the pattern of response errors in two of the conditions, rotation around the map and a "stay" condition in which the observer remained in the initial orientation. Alignment for the JRDs was defined not with respect to the learned perspective, but with respect to the spatial image in working memory, such that when the observer moved around to the opposite perspective view of the physical map, the perspective of the spatial image was correspondingly altered. The results in Fig. 8.7 clearly reveal a shift in the alignment effect for both touch and vision, confirming the results of Waller et al. (2002) using a different definition of alignment after movement. More importantly, response latencies and errors were remarkably similar for touch and vision over all twelve test conditions of the first two experiments, further confirming the pattern shown here for just four of the conditions and providing strong evidence for functional equivalence.

A third experiment replicated these findings with ten blind individuals, who performed the same task as the sighted observers but with haptic maps only. The results were like those of the haptic conditions with the sighted, demonstrating that (1) similar alignment biases occur with the blind, (2) accurate haptic updating is possible in this population, and (3) the equivalence shown in the previous work between haptic and visual learning cannot be attributed to recoding of the haptic information into a visual form. The highly similar pattern of results between modalities across all testing conditions and between blind and sighted observers

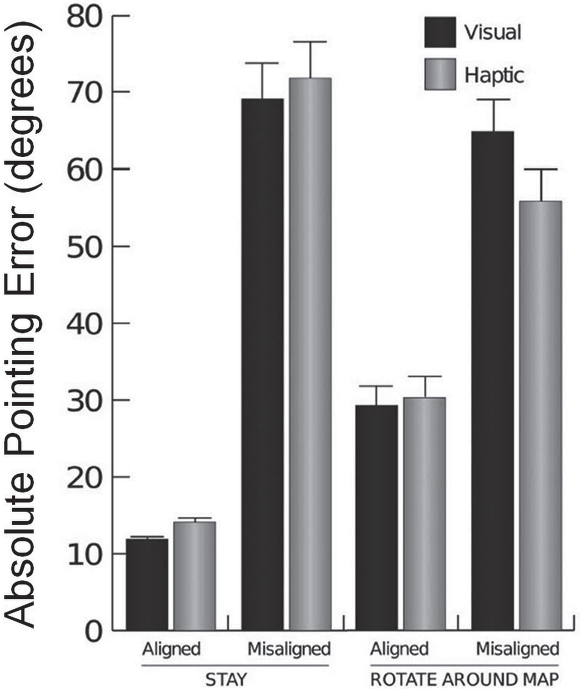

Fig. 8.7 Partial results of Experiment 2 in a study by Giudice et al. (2011). During the learning phase, observers viewed or felt maps with 3 linear segments. Their task was to remember the spatial layout of the four numbered vertices. During an intervening phase with eyes closed, observers remained in place (Stay) or walked around to the other side of the map (rotate around map). They were then to imagine standing at one vertex of the map (e.g., "2") facing another (e.g., "4") and then were asked to rotate physically to face a third vertex (e.g., "1" or "3"). Absolute pointing errors were very similar for touch and vision across the four conditions here as well as eight other conditions in this and another experiment, strongly supporting the claim of functional equivalence of spatial images derived from visual and haptic input. Note that aligned/misaligned in the rotate around map condition was defined with respect to the spatial image after rotation and not with respect to the learned perspective. Error bars are standard errors of the mean. This figure is a modification of Fig. 4 from Giudice et al. (2011) and is published here with permission of the American Psychological Association

supports the amodal hypothesis that information from all modalities converges onto a common spatial image that transcends any one spatial modality.

8.4.3 Evidence for Functional Equivalence in Learning Bimodal Layouts

We have postulated that functional equivalence is possible because people are acting on an amodal spatial image, but as discussed earlier, the common-recoding and separate-but-equal hypotheses are possible alternative explanations. The equivalent performance of blind and sighted observers described above (Giudice et al. 2011) provides evidence against the recoding hypothesis. A further study (Giudice et al. 2009) specifically addressed the separate-but-equal hypothesis. This claims that different inputs lead to sensory-specific representations that are isomorphic, or nearly so, and that support similar behavior. Two experiments were conducted to investigate whether learning of interspersed haptic and visual object layouts build up into a unitary spatial image, independent of the encoding modality, or whether the individual input modalities are preserved in the ensuing representation. Bimodal and unimodal layouts were designed to be isomorphic; the question is whether there would be a cost of switching between the components of a bimodal spatial image, as would be expected from segregation by modality.

Observers were presented with a bimodal layout of six objects displayed on a circular platform that surrounded them. Half of the objects were felt and half were seen using a spatially congruent virtual environment. Importantly, in Experiment 1, the haptic and visual layouts were learned in isolation, whereas in Experiment 2, they were learned as a single interspersed layout. After learning all target locations to criterion, observers were tested on their ability to make judgments of relative direction between target pairs. As the critical manipulation, the two targets were either in the same modality (e.g., both visual) or a different modality (e.g., visual and haptic).

Results from Experiment 1, which temporally separated the haptic and visual objects during learning, showed clear non-equivalence. That is, trials where the start and end object came from different modalities produced responses that were significantly slower and less accurate than for pure visual or pure haptic trials. Thus, there was a switching cost for trials that require relating locations across two separate spatial images. These results suggest that temporal segregation of the layouts at learning led to distinct spatial images for the two modalities. By contrast, results from Experiment 2, where an integrated bimodal layout was learned, provide support for the formation and accessing of an amodal spatial image. Specifically, the response latencies and pointing errors did not differ between intra- and intermodal trials.

8.4.4 Summary of Research on Functional Equivalence

In this section, we have presented evidence from several tasks to demonstrate functional equivalence across input modalities. Strong similarities were demonstrated not only across the senses of vision, hearing, and touch but also the cognitively mediated modality of spatial language. What these inputs have in common is the space from which they originate. There is great efficiency in a cognitive architecture for which the same higher-level processes can be deployed, regardless of the input channel that is providing information about the surrounding space (Bryant 1997). The variety of tasks across which equivalence has been demonstrated should also be emphasized. They range from direct walking to targets to judgments of spatial layout under assumptions of imagined movement. We have not only demonstrated that functional equivalence was the norm across multiple studies, paradigms, and observer groups, but our studies also presented evidence against competitors to the amodal hypothesis we favor, namely, separate isomorphic representations and common recoding.

8.5 Instantiating Spatial Images from Long-Term Memory

As mentioned in the Sect. 8.1, models of spatial representation commonly distinguish between short-term and long-term memory representations for spatial content (Amorim et al. 1997; Avraamides and Kelly 2008; Burgess 2006; Byrne et al. 2007; Easton and Sholl 1995; Huttenlocher et al. 1991; McNamara 2003; Mou et al. 2004; Waller and Hodgson 2006; Wang and Spelke 2000). Byrne et al. (2007) presented a computational model that traces bidirectional processing connections between the two forms of storage. Spatial learning corresponds to the transfer of perceptual information into a more enduring representation in the long-term store. Conversely, a layout retrieved from long-term memory can augment or complete a representation encoded perceptually into working memory, as long as the frames of reference can be co-registered.

Formation of a spatial image in working memory, suitable for spatial updating and derived from information in long-term memory, has been demonstrated in several studies. Rieser et al. (1994) found that children as young as 3 1/2 years could recall their classroom while at another location and then, while walking and turning, update the mental representation of the classroom as if they were there.

An important issue is whether the spatial image is degraded by storage in long-term memory. That long-term storage could introduce noise into the spatial image is indicated by studies showing a loss in precision of perceptual traces during memory storage (e.g., Amorim et al. 1997; Huttenlocher et al. 1991; Waller and Hodgson 2006). It is also possible that systematic bias might be introduced by storage in memory. Giudice et al. (in press) tested the effects of long-term storage in an experiment where composite spatial images were formed, combining locations retrieved from long-term memory with locations perceptually encoded from the same environment. The observer first learned a set of three targets (the LTM set), by viewing each one several times under dim illumination from a constant vantage point. The perceived locations of the targets were then measured by having the observer walk to each one without vision. The observer then left the room and took part in a mental rotation task intended to disrupt working memory. After returning to the room, the observers reported the locations of the LTM targets by a second round of blind walking. They then stepped sideways to a new vantage point, from which they learned three new targets (the WM set). Finally, the observers made judgments of relative direction (JRDs: imagine facing X, point to Z) involving two WM and two LTM targets or one WM and one LTM target.

The results indicated, first, that there was no shift in the reported positions of the LTM targets from the first to the second blind walking test. Thus, storage of the spatial image over a period of minutes did not produce systematic bias in the memory representation, and retrieval did not reimpose encoding biases. Second, in the JRD task, there was no effect of memory status (LTM vs. WM targets) on the time to make the judgments or on systematic error (which was, in any case, low). One effect of LTM storage was observed, however: The absolute error in the JRD task was greater, when the judged pair involved an LTM target, as compared to pairs with WM targets only. Similarly, pointing to single LTM targets showed greater absolute error than to WM targets. Thus, while memory storage neither added systematic bias nor precluded integration with targets recently encoded into working memory, it did apparently reduce the spatial precision of the remembered target location.

8.6 Memory-Load Demands on the Spatial Image

The spatial image, being one type of content within spatial working memory, is presumably subject to capacity limitations. A question of particular interest is whether this form of representation degrades as more locations are simultaneously represented. While several experiments have examined this issue, of necessity they measure the effects of memory load in the context of some other task, particularly spatial updating, which may by itself be subject to load effects. Given the resulting bias toward effects of memory demands, it is all the more impressive that the spatial image, at least within the context of spatial updating, appears to be unaffected by the number of locations stored within the range of 1–6, possibly more.

In general, the relevant experiments have a baseline condition with no updating and another in which observers translate, rotate, or both before responding. If there is no effect of memory load on the baseline condition, which can occur with small load or when targets are learned to a common criterion, the updating condition can be examined in isolation. Otherwise, a measure of updating error has been used; this subtracts the load effect in the baseline condition from the updating condition to assess the additional error attributable to updating per se.

In a study involving both adults and children, Rieser and Rider (1991) found no effect of the number of visual targets (1–5) on constant or variable errors in pointing without vision from a new location. In an fMRI study done using virtual reality, Wolbers et al. (2008) found that the error and latency effects attributable to updating over a simulated forward translation did not vary reliably over 1–4 targets. In a recently completed experiment by our group, observers walked without vision to the location of a target that had been previously viewed. Target locations ranged from 1.5 to 2.25 m in distance and varied in direction. During the observation interval, 1, 3, or 6 targets, identified by colored lights in a dimly lit room, were presented. During the response phase, observers began walking forward and were informed of the color of the goal target, at which time the observers turned and walked the rest of the way to the goal target. The centroids of the individual stopping points were all close to the targets (mean of 14 cm). More important is the precision of the responses, represented by the variability of the stopping points from the corresponding centroids. For 1, 3, and 6 targets, the mean distances were 28, 26, and 30 cm, respectively, revealing little tendency for updating precision to decrease with number of targets. One study that did report an effect of memory load on performance in a study of updating in virtual reality with 1–3 target locations had procedural differences that make comparisons with other studies difficult (Wang et al. 2006).

Still larger numbers of targets have been investigated in updating tasks, with equivocal results. Harrison (2007) examined the effect of updating under rotations with 4–8 targets and found that set size affected pointing latency but not absolute error. Harrison's study was based on one by Hodgson and Waller (2006) with up to 15 targets. No memory-load effect on updating error after rotation was observed; however, this result was taken as evidence for a distinction between on-line and off-line spatial updating (Amorim et al. 1997). The argument is that as the spatial image becomes too complex, target locations are off-loaded to long-term memory. Updating then incorporates a process of memory retrieval as well as the change of egocentric coordinates with locomotion. An effect of number of stored locations was expected for on-line updating, under the assumption of limited spatial working memory, but no such effect was expected for off-line. However, the studies reported initially in this section suggest that effects of memory load on the spatial image, up to several locations, are not evidenced, even when on-line updating takes place. Beyond that point, the possibility of off-loading to long-term memory makes its capacity difficult to measure.

8.7 Are Spatial Images, as Considered Here, Synonymous with Classical Spatial Imagery?

In the voluminous literature on mental imagery, there is evidence of at least two distinct forms, visual imagery and spatial imagery, and their associated neural systems (e.g., Farah et al. 1988; Hegarty and Kozhevnikov 1999; Knauff 2009; Kozhevnikov et al. 2005, 2002; Mazard et al. 2004; Motes et al. 2008). Visual imagery retains visual features, such as color, is pictorial, and can contain lots of detail. Spatial imagery, in contrast, is coarse, more abstract, three-dimensional, and capable of representing objects undergoing motion. Some of the most compelling evidence for the distinction comes from recent research by Kozhevnikov and her colleagues (Blazhenkova and Kozhevnikov 2010; Blajenkova et al. 2006; Kozhevnikov et al. 2005); in their work, they focus on visualization ability and find support for two corresponding types of ability, object visualization and spatial visualization. The support consists of systematic differences in self-report, performance on different behavioral tasks, and psychometric evidence relating to choice of career (Blajenkova et al. 2006). Recently, Lacey et al. (2011) found that object and spatial dimensions of imagery can be observed in haptic and multisensory representations as well. Clearly, the spatial image concept that is central to our chapter has affinities to spatial imagery and spatial visualization in the mental imagery literature. However, despite this, we are hesitant to identify spatial imagery in the classical sense with the spatial image as discussed here. Our focus has been a form of image that is externalized like a percept and scaled to the environment and, thus, can serve as a goal for action in space (see also Byrne et al. 2007). Spatial imagery in the classical sense, like visual imagery, can be manipulated through active imagination. Because it is less strongly tied to particular objects in the surrounding environment, it appears to have greater flexibility than the spatial image as defined here. Like visual imagery, spatial imagery as generally conceived can be imagined at different scales, in different directions, and can undergo rigid motion (Shepard and Metzler 1971).

8.8 Neural Substrate of the Spatial Image

Our theoretical model and the behavioral research we have described places clear constraints on the possible neural substrate of the spatial image: (a) It can be based on inputs from multiple sensory modalities, spatial language and long-term memory; (b) it represents space in egocentric coordinates; (c) it provides a basis (spatial updating) for guiding action when perceptual information is temporarily unavailable.

These features are generally consistent with the posterior parietal cortex (PPC), which has long been noted for its involvement in spatial attention and perceptually directed action (see Milner and Goodale 2008). PPC is multimodal; it has been characterized as an integration area for visual, somatosensory, auditory, and proprioceptive signals (Andersen 1997). Cognitively mediated spatial processes have also been implicated within PPC (Farah et al. 1988; Kosslyn and Thompson 2003). The PPC in primates is part of a network for transforming visual inputs into motor responses and likely plays a similar role in humans (e.g., Chang et al. 2009; Fernandez-Ruiz et al. 2007).

Byrne et al. (2007) have developed a neuro-computational model that satisfies most of the above constraints; in their work, they specifically pointed to the precuneus, the posterior medial portion of the parietal lobe (Cavanna and Trimble 2006), as a likely site for spatial working memory, which they call the *parietal window*. The content of the parietal window, which is synonymous with the spatial image, is described as a spatial map that is head-centered and egocentric and that represents the locations of visible landmarks and objects derived either from perception or memory. Their model contrasted the parietal window with the function of medial temporal areas, which provide an allocentric map.

Wolbers et al. (2008) specifically implicated the precuneus as the site for spatial updating. In their experiments using virtual reality, observers first learned the locations of one to four objects on the visual ground plane. In the delay phase that came next, observers experienced the objects as either remaining stationary (control condition) or moving forward visually, which elicited updating. After the delay phase, observers indicated the direction of the specified object. In the search for candidate brain loci, the critical signature for updating was sensitivity to both the presence of translation and an effect of the number of objects. Only the precuneus fulfilled these requirements after ruling out activation due to spatial motor planning. However, the relative contributions of the PPC and MT areas to spatial updating continue to be a matter of ongoing discussion (Wiener et al. 2010).

8.9 Relevance for Assistive Technology for Blind People

A major challenge for blind people when navigating is a lack of access to information about the environment. When compensatory nonvisual information is provided about the environment, many of the challenges that spatial processing presents to the blind can be mitigated. For example, obstacle avoidance during travel and learning about the layout of objects within the nearby environment have long been facilitated by natural echolocation and the use of a long cane or dog guide. Improvement of these skills has been the goal of developers of ultrasonic obstacle avoiders, GPS-based navigation devices, and sensory substitution devices using tactile and auditory displays (e.g., Giudice and Legge 2008; Levesque 2009). The amodal spatial-image hypothesis that is supported by the work described in this chapter suggests that assistive technologies for the blind should build on capacities for spatial processing that blind people share with sighted people (Loomis et al. 2012). For reviews of research on spatial cognition in blind and sighted people, see Cattaneo and Vecchi (2011) and Struiksma et al. (2009).

Acknowledgments Preparation of this chapter was supported by NIH grant R01EY016817 awarded to the three authors. Roberta Klatzky also acknowledges the support of the Alexander von Humboldt Foundation. The authors thank Rebecca Lawson, Simon Lacey, Maria Kozhevnikov, Marios Avraamides, and Jonathan Kelly for helpful comments on earlier versions of the manuscript.

References

Amorim MA, Glasauer S, Corpinot K, Berthoz A (1997) Updating an object's orientation and location during nonvisual navigation: a comparison between two processing modes. Atten Percept Psychophys 59:404–418

Andersen RA (1997) Multimodal integration for the representation of space in the posterior parietal cortex. Philos Trans R Soc Lond B Biol Sci 352:1421–1428

Ashmead DH, DeFord LD, Northington A (1995) Contribution of listeners' approaching motion to auditory distance perception. J Exp Psychol Hum Percept Perform 21:239–256

Auvray M, Myin E (2009) Perception with compensatory devices: from sensory substitution to sensorimotor extension. Cogn Sci 33:1036–1058

Avraamides MN (2003) Spatial updating of environments described in texts. Cogn Psychol 47:402–431

Avraamides MN, Kelly JW (2008) Multiple systems of spatial memory and action. Cogn Process 9:93–106

Avraamides MN, Kelly JW (2010) Multiple systems of spatial memory: evidence from described scenes. J Exp Psychol Learn Mem Cogn 36:635–645

Avraamides MN, Loomis JM, Klatzky RL, Golledge RG (2004) Functional equivalence of spatial representations derived from vision and language: evidence from allocentric judgments. J Exp Psychol Learn Mem Cogn 30:801–814

Bennett CR, Giudice NA, Klatzky RL, Loomis JM (2011) Spatial images developed through extended touch: comparing updating performance between haptic and visual learning. Poster presented at the 52nd Annual Psychonomics Society Meeting, Seattle, WA, November 2011

Blajenkova O, Kozhevnikov M, Motes M (2006) Object-spatial imagery: a new self-report imagery questionnaire. Appl Cogn Psychol 20:239–263

Blazhenkova O, Kozhevnikov M (2010) Visual-object ability: a new dimension of non-verbal intelligence. Cognition 117:276–301

Bryant DJ (1997) Representing space in language and perception. Mind Lang 12:239–264

Bryant DJ, Tversky B, Franklin N (1992) Internal and external spatial frameworks for representing described scenes. J Mem Lang 31:74–98

Burgess N (2006) Spatial memory: how egocentric and allocentric combine. Trends Cogn Sci 10:551–557

Burgess N (2008) Spatial cognition and the brain. Ann N Y Acad Sci 1124:77–97

Byrne P, Becker S, Burgess N (2007) Remembering the past and imagining the future: a neural model of spatial memory and imagery. Psychol Rev 114:340–375

Cattaneo Z, Vecchi T (2011) Blind Vision: The Neuroscience of Visual Impairment. MIT Press, Cambridge, MA

Cavanna AE, Trimble MR (2006) The precuneus: a review of its functional anatomy and behavioral correlates. Brain 129:564–583

Chang SW, Papadimitriou C, Snyder LH (2009) Using a compound gain field to compute a reach plan. Neuron 64:744–755

Chebat DR, Schneider F, Kupers R, Ptito M (2011) Navigation with a sensory substitution device in congenitally blind individuals. Neuroreport 22:342–347

Denis M, Cocude M (1989) Scanning visual images generated from verbal descriptions. Eur J Cogn Psychol 1:293–307

De Vega M, Rodrigo MJ (2001) Updating spatial layouts mediated by pointing and labelling under physical and imaginary rotation. Eur J Cogn Psychol 13:369–393

Easton RD, Sholl MJ (1995) Object–array structure, frames of reference, and retrieval of spatial knowledge. J Exp Psychol Learn Mem Cogn 21:483–500

Farah MJ, Hammond KM, Levine DN, Calvanio R (1988) Visual and spatial mental imagery: dissociable systems of representation. Cogn Psychol 20:439–462

Farrell MJ, Robertson IH (1998) Mental rotation and the automatic updating of body-centered spatial relationships. J Exp Psychol Learn Mem Cogn 24:227–233

Farrell MJ, Thomson JA (1999) On-line updating of spatial information during locomotion without vision. J Mot Behav 3:39–53

Fernandez-Ruiz J, Goltz HH, DeSouza JFX, Vilis T, Crawford JD (2007) Human parietal "reach region" primarily encodes intrinsic visual direction, not extrinsic movement direction, in a visual–motor dissociation task. Cereb Cortex 17:2283–2292

Franklin N, Tversky B (1990) Searching imagined environments. J Exp Psychol Gen 119:63–76

Fukusima SS, Loomis JM, Da Silva JA (1997) Visual perception of egocentric distance as assessed by triangulation. J Exp Psychol Hum Percept Perform 23:86–100

Giudice NA, Betty MR, Loomis JM (2011) Functional equivalence of spatial images from touch and vision: evidence from spatial updating in blind and sighted individuals. J Exp Psychol Learn Mem Cogn 37:621–634

Giudice NA, Klatzky RL, Loomis JM (2009) Evidence for amodal representations after bimodal learning: integration of haptic-visual layouts into a common spatial image. Spat Cogn Comput 9:287–304

Giudice NA, Klatzky RL, Bennett CR, Loomis JM (in press) Combining locations from working memory and long-term memory into a common spatial image. Spat Cogn Comput

Giudice NA, Legge GE (2008) Blind navigation and the role of technology. In: Helal A, Mokhtari M, Abdulrazak B (eds) Engineering Handbook of Smart Technology for Aging, Disability, and Independence. Wiley, New York

Gordon MS, Rosenblum LD (2004) Perception of sound-obstructing surfaces using body-scaled judgments. Ecol Psychol 16:87–113

Harrison AM (2007) Reversal of the Alignment Effect: Influence of Visualization and Spatial Set Size. In: Proceedings of the 29th Annual Conference of the Cognitive Science Society. Curran Associates, Red Hook, NY

Hegarty M, Kozhevnikov M (1999) Types of visual-spatial representations and mathematical problem solving. J Educ Psychol 91:684–689

Hodgson E, Waller D (2006) Lack of set size effects in spatial updating: evidence for offline updating. J Exp Psychol Learn Mem Cogn 32:854–866

Hollins M, Kelley EK (1988) Spatial updating in blind and sighted people. Percept Psychophys 43:380–388

Horn DL, Loomis JM (2004) Spatial updating of targets in front and behind. Paidéia 14:75–81

Hughes B (2001) Active artificial echolocation and the nonvisual perception of aperture passability. Hum Mov Sci 20:371–400

Huttenlocher J, Hedges LV, Duncan S (1991) Categories and particulars: prototype effects in estimating spatial location. Psychol Rev 98:352–376

Kelly JW, Avraamides M, Loomis JM (2007) Sensorimotor alignment effects in the learning environment and in novel environments. J Exp Psychol Learn Mem Cogn 33:1092–1107

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, Beall AC, Chance SS, Golledge RG (1998) Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. Psychol Sci 9:293–298

Knauff M (2009) A neuro-cognitive theory of deductive relational reasoning with mental models and visual images. Spat Cogn Comput 9:109–137

Koch C (2003) The Quest for Consciousness. Roberts and Co., Englewood, CO

Kosslyn SM, Thompson WL (2003) When is early visual cortex activated during visual mental imagery? Psychol Bull 129:723–746

Kozhevnikov M, Hegarty M, Mayer R (2002) Revising the cognitive style dimension: evidence for two types of visualizers. Cogn Instr 20:47–77

Kozhevnikov M, Kosslyn S, Shephard J (2005) Spatial versus object visualizers: a new characterization of visual cognitive style. Mem Cognit 33:710–726

Lacey S, Lin JB, Sathian K (2011) Object and spatial imagery dimensions in visuo-haptic representations. Exp Brain Res 213:267–273

Lehar S (2003) The World in your Head: A Gestalt View of the Mechanism of Conscious experience. Erlbaum, Mahwah, NJ

Levesque V (2009) Virtual display of tactile graphics and Braille by lateral skin deformation. Unpublished manuscript, Department of Electrical and Computer Engineering, McGill University, Montreal, Canada. Available online at http://vlevesque.com/papers/Levesque-PhD.pdf

Levine M, Jankovic IN, Palij M (1982) Principles of spatial problem solving. J Exp Psychol Gen 111:157–175

Lyon DR, Gunzelmann GM (2011) Functional equivalence and spatial path memory. Q J Exp Psychol 64:2081–2087

Loarer E, Savoyant A (1991) Visual Imagery in Locomotor Movement Without Vision. In: Logie RH, Denis M (eds) Mental Images in Human Cognition. Elsevier, The Hague

Loomis JM (1992) Distal attribution and presence. Presence (Camb) 1:113–119

Loomis JM, Beall AC (2004) Model-based control of perception/action. In: Vaina L, Beardsley S, Rushton S (eds) Optic Flow and Beyond. Kluwer Academic Publishers, Boston, MA

Loomis JM, Da Silva JA, Fujita N, Fukusima SS (1992) Visual space perception and visually directed action. J Exp Psychol Hum Percept Perform 18:906–921

Loomis JM, Klatzky RL (2007) Functional equivalence of spatial representations from vision, touch, and hearing: relevance for sensory substitution. In: Rieser JJ, Ashmead DH, Ebner FF, Corn AL (eds) Blindness and Brain Plasticity in Navigation and Object Perception. Lawrence Erlbaum Associates, New York

Loomis JM, Klatzky RL, Giudice NA (2012) Sensory substitution of vision: importance of perceptual and cognitive processing. In: Manduchi R, Kurniawan S (eds) Assistive Technology for Blindness and Low Vision. CRC, Boca Raton, FL

Loomis JM, Klatzky RL, Philbeck JW, Golledge RG (1998) Assessing auditory distance perception using perceptually directed action. Percept Psychophys 60:966–980

Loomis JM, Lippa Y, Klatzky RL, Golledge RG (2002) Spatial updating of locations specified by 3-D sound and spatial language. J Exp Psychol Learn Mem Cogn 28:335–345

Loomis JM, Philbeck JW (2008) Measuring spatial perception with spatial updating and action. In: Klatzky RL, Behrmann M, MacWhinney B (eds) Embodiment, Ego-Space, and Action. Taylor & Francis, New York

May M (2004) Imaginal perspective switches in remembered environments: transformation versus interference accounts. Cogn Psychol 48:163–206

Mazard A, Tzourio-Mazoyer N, Crivello F, Mazoyer B, Mellet E (2004) A PET meta-analysis of object and spatial imagery. Eur J Cogn Psychol 16:673–695

McNamara TP (2003) How are the locations of objects in the environment represented in memory? In: Freksa C, Brauer W, Habel C, Wender K (eds) Spatial Cognition III: Routes and Navigation, Human Memory and Learning, Spatial Representation and Spatial Reasoning. Springer, Berlin

Medendorp WP, Tweed DB, Crawford JD (2003) Motion parallax is computed in the updating of human spatial memory. J Neurosci 23:8135–8142

Milner AD, Goodale MA (2008) Two visual systems re-viewed. Neuropsychologia 46:774–785

Motes MA, Malach R, Kozhevnikov M (2008) Object-processing neural efficiency differentiates object from spatial visualizers. Neuroreport 19:1727–1731

Mou W, McNamara TP, Valiquette CM, Rump B (2004) Allocentric and egocentric updating of spatial memories. J Exp Psychol Learn Mem Cogn 30:142–157

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

O'Keefe J, Nadel L (1978) The Hippocampus as a Cognitive Map. Oxford University Press, Oxford

Ooi TL, Wu B, He ZJJ (2001) Distance determined by the angular declination below the horizon. Nature 414:197–200

Ooi TL, Wu B, He ZJ (2006) Perceptual space in the dark affected by the intrinsic bias of the visual system. Perception 35:605–624

Pascual-Leone A, Hamilton R (2001) The metamodal organization of the brain. In: Casanova C, Ptito M (eds) Vision: from Neurons to Cognition (vol 134, Ch 27: Progress in Brain Research). Elsevier, Amsterdam

Pham QC, Hicheur H (2009) On the open-loop and feedback processes that underlie the formation of trajectories during visual and nonvisual locomotion in humans. J Neurophysiol 102:2800–2815

Philbeck JW, Loomis JM (1997) Comparison of two indicators of visually perceived egocentric distance under full-cue and reduced-cue conditions. J Exp Psychol Hum Percept Perform 23:72–85

Philbeck JW, Loomis JM, Beall AC (1997) Visually perceived location is an invariant in the control of action. Percept Psychophys 59:601–612

Presson CC, Montello DR (1994) Updating after rotational and translational body movements: coordinate structure of perspective space. Perception 23:1447–1455

Riecke BE, von der Heyde M, Bülthoff HH (2005) Visual cues can be sufficient for triggering automatic, reflex-like spatial updating. ACM Trans Appl Percept 2:183–215

Rieser JJ (1989) Access to knowledge of spatial structure at novel points of observation. J Exp Psychol Learn Mem Cogn 15:1157–1165

Rieser JJ, Ashmead DH, Talor CR, Youngquist GA (1990) Visual perception and the guidance of locomotion without vision to previously seen targets. Perception 19:675–689

Rieser JJ, Garing AE, Young MF (1994) Imagery, action, and young children's spatial orientation: it's not being there that counts, it's what one has in mind. Child Dev 65:1262–1278

Rieser JJ, Guth DA, Hill EW (1986) Sensitivity to perspective structure while walking without vision. Perception 15:173–188

Rieser JJ, Rider E (1991) Young children's spatial orientation with respect to multiple targets when walking without vision. Dev Psychol 27:97–107

Röder B, Rösler F (1998) Visual input does not facilitate the scanning of spatial images. J Mental Imagery 22:165–182

Russell B (1948) Human Knowledge: Its Scope and Limits. Simon and Schuster, New York

Sahm CS, Creem-Regehr SH, Thompson WB, Willemsen P (2005) Throwing versus walking as indicators of distance perception in similar real and virtual environments. ACM Trans Appl Percept 2:35–45

Shelton AL, McNamara TP (2004) Spatial memory and perspective taking. Mem Cognit 32:416–426

Shepard RN, Metzler J (1971) Mental rotation of three-dimensional objects. Science 171:701–703

Siegle JH, Campos JL, Mohler BJ, Loomis JM, Bülthoff HH (2009) Measurement of instantaneous perceived self-motion using continuous pointing. Exp Brain Res 195:429–444

Smythies JR (1994) The walls of Plato's Cave. The Science and Philosophy of Brain, Consciousness, and Perception. Avebury, Aldershot

Stevenson RJ, Case TI (2005) Olfactory imagery: a review. Psychon Bull Rev 12:244–264

Struiksma ME, Noordzij ML, Postma A (2009) What is the link between language and spatial images? Behavioral and neural findings in blind and sighted individuals. Acta Psychol 132:145–156

Tatler BW, Land MF (2011) Vision and the representation of the surroundings in spatial memory. Philos Trans R Soc B 366:596–610

Taylor HA, Tversky B (1992) Spatial mental models derived from survey and route descriptions. J Mem Lang 31:261–282

Teng S, Puri A, Whitney D (2012) Ultrafine spatial acuity of blind expert human echolocators. Exp Brain Res 216(4):483–488

Thomson JA (1983) Is continuous visual monitoring necessary in visually guided locomotion? J Exp Psychol Hum Percept Perform 9:427–443

Thompson WB, Willemsen P, Gooch AA, Creem-Regehr SH, Loomis JM, Beall AC (2004) Does the quality of the computer graphics matter when judging distances in visually immersive environments. Presence (Camb) 13:560–571

Tolman EC (1948) Cognitive maps in rats and men. Psychol Rev 55:189–208

Waller D, Hodgson E (2006) Transient and enduring spatial representations under disorientation and self-rotation. Mem Cognit 32:867–882

Waller D, Montello DR, Richardson AE, Hegarty M (2002) Orientation specificity and spatial updating of memories for layouts. J Exp Psychol Learn Mem Cogn 28:1051–1063

Wang RF, Spelke ES (2000) Updating egocentric representations in human navigation. Cognition 77:215–250

Wang RF, Crowell JA, Simons DJ, Irwin DE, Kramer AF, Ambinder MS et al (2006) Spatial updating relies on an egocentric representation of space: effects of the number of objects. Psychon Bull Rev 13:281–286

Wiener JM, Berthoz A, Wolbers T (2010) Dissociable cognitive mechanisms underlying human path integration. Exp Brain Res 208:61–71

Wolbers T, Hegarty M, Büchel C, Loomis JM (2008) Spatial updating: How the brain keeps track of changing object locations during observer motion. Nat Neurosci 11:1223–1230

Wu B, Ooi TL, He ZJ (2004) Perceiving distances accurately by a directional process of integrating ground information. Nature 428:73–77

Zwaan RA, Radvansky GA (1998) Situation models in language comprehension and memory. Psychol Bull 123:162–185