Chapter 1
Sensory Contributions to Spatial Knowledge of Real and Virtual Environments

David Waller and Eric Hodgson

Abstract Most sensory systems are able to inform people about the spatial structure of their environment, their place in that environment, and their movement through it. We discuss these various sources of sensory information by dividing them into three general categories: external (vision, audition, somatosensory), internal (vestibular, kinesthetic) and efferent (efference copy, attention). Research on the roles of these sensory systems in the creation of environmental knowledge has shown, with few exceptions, that information from a single sensory modality is often sufficient for acquiring at least rudimentary knowledge of one's immediate environment and one's movement through it. After briefly discussing the ways in which sources of sensory information commonly covary in everyday life, we examine the types and quality of sensory information available from contemporary virtual environments, including desktop, CAVE, and HMD-based systems. Because none of these computer mediated systems is yet able to present a perfectly full and veridical sensory experience to its user, it is important for researchers and VE developers to understand the circumstances, tasks, and goals for which different sensory information sources are most critical. We review research on these topics, as well as research on how the omission, limitation, or distortion of different information sources may affect the perception and behavior of users. Finally, we discuss situations in which various types of virtual environment systems may be more or less useful.

Brian and Sarah set up their campsite at dusk. At dawn the next morning, Sarah decides to venture from the campsite briefly to explore the area. As she walks away from the camp, the sights and sounds of a nearby brook recede into the background. She feels her legs working, climbing uphill from the brook, her feet occasionally slipping backwards slightly, down the muddy trail. As an owl flies by, it captures her attention, and she turns her head quickly to watch it light in a nearby tree. Turning her head to admire the bird does not disorient Sarah; nor does it cause her to change

D. Waller (✉) · E. Hodgson
Department of Psychology, Miami University,
Psychology Building 218, Oxford, OH 45056, USA
e-mail: wallerda@muohio.edu

F. Steinicke et al. (eds.), *Human Walking in Virtual Environments*,
DOI: 10.1007/978-1-4419-8432-6_1,

her generally uphill course. After reaching the top of the hill, Sarah directs her gaze down toward a larger river running below her before deciding to move on back to her campsite. Even before her campsite comes into view, she can smell smoke coming from the campfire that Brian had just built.

This brief vignette illustrates the variety and the complexity of information that is available to people as they navigate through known and unknown places. In the story, Sarah uses a variety of online sensory information—visual, kinesthetic, vestibular, somatosensory, auditory, and even olfactory—to provide information about her environment and her place in it. She is also able to use offline information—internally stored knowledge—to determine her goals and destinations. Because nearly every sensory modality can contribute information about one's spatial disposition or about the spatial properties of one's environment, it is a challenge for scientists to understand how this multimodal influx of information is acquired, interpreted, combined, and acted upon. In this chapter, we will examine what is known about the different ways that people can take-in the spatial information that is available to them as they walk through real and computer-simulated (virtual) environments. In the first part of the chapter, we will briefly discuss what we know about the senses' contribution to spatial knowledge, including the circumstances under which each sense in isolation is necessary or sufficient for acquiring spatial knowledge and for enabling spatial behaviors. The modality-by-modality discussion, however, is only expository; it is important to realize that in most real situations, information flows simultaneously through all of the senses—that spatial information is multimodal, overlapping, and heavily redundant. Thus, in the second part of the chapter, we examine situations in which there are differing combinations of sensory information about space. In particular, we discuss the sensory contributions of today's computer-simulated environments, in which some sensory modalities have access to limited or imperfect information.

Before we can begin a meaningful discussion of incoming sensory information, we will do well to recognize that the issue of sensory contributions to spatial knowledge cannot be disentangled from conceptualizations of the type of knowledge that is produced. In general, the type of knowledge that one acquires about his or her environment depends critically on the goals that one has for interacting with the environment. On one hand, many goals are served exclusively by transient knowledge of one's location and orientation with respect to important objects in the immediate environment. In the situation above for example, Sarah did not need to recall the location of her campground in order to track and admire a nearby bird. Tasks such as speed, distance, and turn estimation, as well as online spatial updating (i.e., the ability to track one's location and orientation with respect to salient objects in the environment; see [62]) may involve a dynamic form of environmental knowledge that consists of little more than sensitivity to ongoing changes in immediately available information. Here, internal representations of the environment may be minimal—fragmented, sketchy, and schematic—or even nonexistent, because the environment itself provides easy access to necessary information. Such situations seem particularly well addressed by concepts that have developed out of the ecological approach to perception, such as information pickup, direct perception, and perception/action complementarity (see, for example, [31, 72, 107]).

On the other hand, people often rely on knowledge of remote locations to set future goals, to orient toward unseen locations, and to imagine themselves (and other people) in different places. Such knowledge consists of remembered information that is not perceptually available for immediate inspection. Although ecological approaches may be able to address and describe this type of knowledge (see for example, [41, 42]), it is generally more common in the literature to consider spatial memory as involving internal representations of environmental structure [29]. By this view, enduring spatial knowledge is the result of mental processing and storage of environmental and internal information. Most researchers accept a three-tiered description of enduring spatial knowledge that includes: (a) relatively rudimentary (and arguably nonspatial) memory for landmarks and features, (b) route knowledge of how places are interconnected, typically without regard to the metric distances and directions between them, and (c) configuration (or survey) knowledge of global metric relations among places in an environment. Of course, sensory contributions to any of these types of knowledge may be qualitatively different, which can complicate our discussion considerably.

Regardless of whether spatial knowledge enables action in the moment or is stored for subsequent use, it is clear that in order to be useful, spatial knowledge must coordinate the disposition of the knower with characteristics of the environment. Thus in laying-out the sensory bases for environmental cognition, we make a primary distinction between sensory information that provides either: (a) external information about the nature of one's environment, or (b) internal (or idiothetic) information about the status of one's body or effectors. After discussing these sources of sensory information, we turn briefly to a consideration of non-sensory information that is internally generated, such as attention allocation and other cognitive factors.

1.1 External Sensory Information

Vision provides a direct, rich, and precise source of spatial information, and it is undoubtedly the most researched sensory modality with respect to environmental knowledge and navigation. Like many other external senses, the visual system can provide detailed and useful information about the spatial layout of the immediate environment without significant bodily movement (see [24, 80] for a review of visual cues that enable apprehension of layout). The layout of an environment, including relative directions, distances, and scale can be accurately perceived and remembered from a stationary viewpoint (e.g., [92]), from brief glimpses of images of a spatial layout [118], and even from symbolic media such as maps [105] and gestures [3]. Frequently however, visual information about an environment comes from movement through it—from the optic flow generated from changes in one's position and heading. As one changes position (e.g., moving forward or backward) optic flow patterns radiate from a focus of expansion or contraction. As one changes orientation (e.g., rotating to one's left or right) flow patterns are laminar, moving across the retina in the direction opposite to the motion. Combinations of radial and laminar optic flow can signify relations among one's: (a) direction of travel (i.e., one's course), (b) head direction, and (c) gaze direction. In so doing, the online interpretation of optic flow provides information about the spatial relations in the environment that both are and are not dependent on one's viewpoint [31].

A substantial body of research (as well as informal observation) indicates that despite its availability and precision, visual information is often not necessary for performing many rudimentary spatial tasks or even for developing detailed knowledge of one's environment. The navigational abilities of congenitally blind people [34] attest to this idea. The experimental literature too is rife with demonstrations that visual information can be surprisingly unnecessary for accurate perception and memory of space, and researchers frequently conclude that internal senses (discussed below) are relatively more influential than vision for the acquisition of online environmental knowledge (e.g., [15]). On the other hand, it has been known for decades that in theory, visual information can be sufficient for performing a variety of spatial tasks [30, 31]. For online perceptual tasks such as speed [59], distance [28, 64], and heading [58] estimation, optic flow alone appears to be sufficient for enabling accurate spatial knowledge. However, when tasks require online tracking of one's rotational changes, visual information by itself tends to be much less sufficient. Klatzky et al. [55], for example, demonstrated that optic flow alone does not generally enable people to keep track of their orientation in space and that additional body-based information is required (see also [16, 76]). There is some evidence, however, that visual information can enable people to account for rotational changes, especially if it provides information about landmarks, and is available in a sufficiently large field of view [83, 85]. Finally, for memory-based tasks, even very briefly presented visual information can be sufficient for the acquisition of knowledge of spatial layout, including relative directions and distances between objects in a single scene [118]. Indeed, static visual information from photographs [4], as well as dynamic visual information from video [33] or desktop virtual environments [86, 89] can be sufficient for acquiring survey knowledge of relatively complex environments, although such knowledge may be slow to develop [111] and show especially large differences among users [87].

Other external senses such as audition, olfaction, and the somatosenses (i.e., pain, pressure, vibration, and heat) can provide spatial information for both online and offline purposes; however, the information from these sources is typically not as rich as that provided by vision. Among these non-visual external senses, two are especially useful in generating environmental knowledge. First, audition can be used to localize noise-producing objects [117] and to sense the scale of a local environment [90]. Indeed, people can learn to echolocate well enough to aid nonvisual navigation [91, 98]. People can also gain accurate knowledge of an environment based solely on auditory information such as verbal directions [21], spatial language [63] or environmental descriptions [36]. Second, a very common—though rarely investigated—source of information about the external environment comes from somatosensory pressure receptors that provide information about: (a) acceleration, through one's points of contact with a surface of support, and (b) physical contact with objects in one's environment. When acquired actively, and in conjunction with kinesthetic information (discussed below) this latter source of somatosensory information is known as *haptics*, a topic that we discuss briefly below.

The other external senses likely have a negligible impact on the acquisition of environmental knowledge. In theory, one can pickup somatosensory information from external sources and use variations in the strength of these sources to navigate or to inform one about the environment. For example, direct sunlight may heat one side of a person more than another, allowing her to determine the direction of the sun and, by extension, her orientation. Similarly, the strength of vibrations under one's feet may be used to estimate one's proximity to a large machine on a factory floor. Indeed, the fact that vibratory devices have been constructed to signal directions for navigation [108] indicates that such information sources may be useful in some circumstances. However in most everyday situations, the role of these and other sensory systems (such as olfaction and taste) in environmental cognition is probably minimal and is certainly not well researched.

1.2 Internal Sensory Information

Idiothetic information about space derives from three principal sensory systems: (a) the vestibular system, which senses angular and linear acceleration of the head; (b) the kinesthetic/proprioceptive system, which provides information about the position, orientation, and movement of the musculature; and (c) the somatogravity system [66] that informs people about the direction of gravity. A fourth source of information, efference copy, is sometimes considered a source of idiothetic information [67]; however, we will defer discussion of efference copy to our discussion below of efferent information more generally. The vestibular and kinesthetic/proprioceptive systems have direct and important ties to spatial cognition, and we discuss them briefly below; however, we do not discuss further the somatogravity system, as it is little researched in humans, and its contributions to environmental knowledge are largely unknown.

The anatomical basis of the vestibular system is a set of structures in the inner ear—otoliths and semicircular canals—that sense linear and angular acceleration, respectively. In addition to supporting several postural and oculomotor reflexes, vestibular information provides information that can be doubly-integrated in order to determine linear or angular displacement. In this way, it is thought to provide a critical input to people's ability to dead reckon and to update (see [77]). In the experimental literature, vestibular information is typically isolated by passively transporting blindfolded participants; however, we note that such procedures do not completely isolate the vestibular sense, as they do not remove somatosensory information about inertial forces (e.g., the sensation of the wheelchair pressing against one's back). These experiments have demonstrated that vestibular information enables better than chance accuracy on online spatial tasks such as turn and distance estimation [7, 38, 116], as well as for more complex online tasks such as spatial updating. However, the accuracy and precision of vestibularly-acquired spatial information is typically rather low, and degrades rapidly as the knowledge derived from it becomes more complex. For example, Sholl [94] pushed blindfolded participants in a wheelchair over paths of varying complexity. Most participants were able to track simple two-segment paths with well-better than chance accuracy (see also [38]); however, participants were generally unable to keep track of complex trajectories (i.e., those involving more than three turns). With respect to the creation of relatively complex and enduring spatial knowledge, the sufficiency of vestibular information is largely unexplored in the literature, although Sholl's results would suggest that it is difficult to acquire complex route or survey knowledge on the basis of vestibular information alone. Among healthy adults, vestibular information may also be necessary for successful interaction with the environment, as sudden loss or disruption of it through pathology [20, 32, 75, 76] or experimental manipulation [101] can significantly impair performance on basic spatial tasks such as turn and distance estimation. However, people can generally adapt to a gradual progressive degradation of the vestibular system [57], and thus its full functioning is not always necessary for acquiring spatial knowledge. The role of the vestibular system as well as its relation to other idiothetic senses is reviewed in [57].

Although the terms proprioceptive and kinesthetic are often used interchangeably in the literature, we use the former to refer to information about a relatively static position or attitude of the musculature; whereas the latter refers to information about the movement of one's limbs or effectors. By this distinction, knowledge of the location of one's unseen hand, for example, may come from a proprioceptive sense; whereas the ability to brush one's teeth without a mirror would require kinesthetic information. A relatively small research literature has found that kinesthetic information is generally sufficient for acquiring knowledge of distances [67, 102] and orientation changes [8, 13, 52]. Indeed, compared to the vestibular and somatosensory components of idiothetic information, the proprioceptive component may enable relatively more accurate performance in a variety of online tasks, such as heading estimation [73, 103], turn estimation [53], distance estimation [67], and spatial updating [94]. As with vestibular information, the sufficiency of kinesthetic information for acquiring relatively sophisticated knowledge about spatial layout is under-researched. But the ability to acquire accurate spatial information about traveled distances [28], turns [83], routes [114], and spatial layout [86] from purely visual sources demonstrate that kinesthetic (and other sources of) information is not strictly necessary for the acquisition of spatial knowledge.

1.3 Efferent Sources of Information

Finally, we consider the contributions to spatial knowledge of four other internal sources of spatial information—efference copy, attention allocation, cognitive decision making, and mental transformations. Of course it is incorrect to consider these information sources as "sensory" inasmuch as they are not considered to carry afferent information from the peripheral to the central nervous system. Nonetheless, these information sources form a critical component of active engagement with one's environment and may have a strong influence on the nature and quality of spatial information that is available for acquiring environmental knowledge. It is worth noting that these sources of information are also tightly constrained and dictated by one's goals. Our brief discussion of these information sources draws heavily on the conceptual distinctions made by Chrastil and Warren [18] who used these concepts in their recent review of the extensive and complicated literature on active and passive contributions to spatial knowledge.

Efference copy [47] is a simultaneous record of the motor commands from the central nervous system to the musculature that enables organisms to account for the difference between external stimulation and the stimulation that arises as a consequence of their own actions. For example, as we discussed above, when one turns one's head clockwise, the visual system has access to laminar optic flow in a counterclockwise direction. Logically, such a pattern of optic flow could signal a counterclockwise rotation of the visible environment around a stationary viewpoint, rather than a rotation of a viewpoint in a stationary environment. However, people are typically able to distinguish these possibilities by accounting for the fact that motor commands actively produced a set of expected visual consequences (but see [43] for situations when people cannot). This knowledge of one's motor commands constitutes efference copy. We consider the efferent copy of motor commands to contain information about both the implicit or explicit intentions used to move in and interact with the environment, as well as information about the strength of these intentions. In this way, efference copy can be used to generate a set of expectations about the consequences of one's actions (see [12]). Discrepancies between the expected and the perceived consequences of a set of motor commands can indicate a need for perceptual recalibration or can be used as an indicator of the precision of one's intentions.

Other efferent sources of information have origins in "higher-level" cognition and include factors such as constraints about how and where to allocate attention, decisions about how and where to navigate, and the ability to transform spatial information. Chrastil and Warren [18] point out that these internal sources of information are often confounded or conflated in studies that examine the relative effects of active versus passive navigation on the acquisition of spatial knowledge. However, because virtual environments (VEs) enable yoked playback of others' interactive experiences, computer simulations have recently provided researchers a helpful methodological tool for teasing some of these influences apart. For example, Christou and Bülthoff ([19], Experiment 3) asked participants to explore a model of a complex attic space on a desktop virtual reality system and to familiarize themselves with the environment and its contents. One group of participants navigated through the environment actively, by manipulating a trackball that controlled the position and orientation of their simulated viewpoint. A second group of participants had access to the same visual information as the first, but passively viewed playbacks of the explorations made by matched participants in the other condition. Differences between these groups on the ability to recognize novel and previously seen views were minimal, not statistically reliable, and in some cases numerically favored the passive learners. The researchers concluded that the acquisition of spatial knowledge was probably influenced more by the relative amounts of attention deployed to either the navigational interface or to the learning of the environment than by active decision making about where and how to navigate. Similar conclusions were reached by Chrastil and Warren in their review of the broader literature, who noted that differences between active and passive learning of environments are probably more heavily influenced by the allocation of attention than by the ability to choose one's own route through the environment.

1.4 Relative Influence of External and Internal Sensory Information

To this point, we have discussed different sources of spatial information in isolation. Yet we have also noted that in normal real-world situations, multiple sources of sensory and efferent information are mutually available, and that these sources typically provide overlapping and redundant information. Thus, in understanding the contributions of sensory systems in real-world environments, it is often more relevant and practical to consider the relative influences of these sources when many are present, rather than considering the degree to which isolated sources of sensory information are necessary or sufficient for the acquisition of spatial knowledge. On the other hand, users of VE systems often do not have access to one or more sensory sources of information. Thus, in understanding the contributions of sensory systems in VEs, it is especially relevant to consider the consequences to spatial knowledge acquisition when some sources of sensory information are absent, degraded, or uninformative. These two topics (i.e., overlapping sensory information in real world environments and degraded sensory sources in virtual environments) comprise the outline for the final section of this chapter. After a brief discussion of the relative contributions of different sources of sensory information in real-world environments, we close by summarizing what is known about the impact of unavailable or degraded sources of sensory information that comes with various types of VE systems.

1.4.1 Sensory Contributions in the Real World

As we have seen in many cases, separate sensory sources are frequently naturally coupled, and thus can be extremely difficult to isolate experimentally. For example, it is difficult for experimenters to dissociate kinesthetic information about neck movement from vestibular information about the head's rotation. As a result of these naturally conflated sensory sources, researchers commonly combine various sources of sensory or efferent information into broader categories. For example, many researchers (ourselves included) refer to *body-based* information to refer to the amalgam of vestibular, kinesthetic, and efferent information that is primarily associated with self-movement. Others may refer to the amalgamated concept *inertial information* to refer to the combination of vestibular, somatosensory, and sometimes visual information that indicates acceleration through space. Additionally, the concept of *podokinesthetic* (or *podokinetic*) information is sometimes used in the literature to refer to the grouping of somatosensory, kinesthetic, and efferent information available about foot contact with the ground during locomotion. Finally, *haptic* information [60] derives from the combination of efferent, kinesthetic and somatosensory information that arises during the active manipulation of objects. In these ways, groups of commonly occurring sensory sources can be isolated and examined for their combined contribution to spatial knowledge. It is worth noting that the advent of virtual reality technology has made much of this research relatively easy to do, although visual-idiothetic dissociations using mirrors [70] or prism glasses [39, 99] have also historically been able to address many of these questions.

A growing body of research has examined the degree to which the presence of these amalgamated sources of information facilitates the acquisition or enhances the quality of spatial knowledge when it and others are available. Probably most of this work has focused on the relative influences of body-based information when vision is available [54, 55, 73, 94, 102, 103]. With respect to online tasks that require transient knowledge, the literature generally shows that body-based senses do facilitate the acquisition of accurate spatial knowledge of one's environment [15, 53, 116]. Indeed, as mentioned above, there is some evidence that especially with respect to acquiring accurate information about turns and orientation, body-based senses may be necessary [7, 55, 88].

The relative contributions of visual and body-based senses in tasks and environments that require sophisticated enduring knowledge of space, however, is less clear, and only a handful of studies have addressed this issue. The few that have tended to conclude that body-based senses contribute minimally to environmental learning. Waller et al. [113] examined the degree to which inertial information facilitates the visual system in acquiring configural knowledge of a large environment by manipulating the amount of valid inertial information available from a 1.6 km car trip through a real-world neighborhood. Three matched groups of participants received visual information by viewing the trip through a head-mounted display (HMD). These three groups differed on the presence and quality of inertial (i.e., vestibular and somatosensory) information: one group had access to valid inertial cues, one group had access to invalid inertial cues, and a final group had access to no inertial cues. The results clearly showed no differences among these groups in their (fairly accurate) memory of the layout of the environment. Interestingly however, in a closely-related follow-up study, Waller et al. ([112]; see also [110]) showed that the additional idiothetic cues provided by proprioception and efference copy did make a small but significant improvement in the accuracy of spatial knowledge acquired from vision. The primary conclusions from all of these studies, however, has been that the role of idiothetic information in forming an enduring representation of large environments is quite small—generally much less than the facilitative effect of idiothetic on transient knowledge of local environments.

Finally, it is worth noting that questions about the relative influence of various sensory systems (and combinations of sensory information) on spatial knowledge are closely related to the question of how and whether spatial information from different sources is integrated or combined into a coherent and unitary percept of space. Recent research has provided empirical support for the combination of visual information with auditory (e.g., [1]), vestibular (e.g., [14]), and haptic information [26]. Most of this work is consistent with a Bayesian model of sensory integration in which the weights assigned to various sources of information are determined by estimates of the source's relative precision [10, 17, 68].

Having discussed the influence of various sensory systems to the acquisition of knowledge of real world environments, we now turn our attention to the spatial information available to users of VEs. As suggested earlier, a primary difference between the sensory contributions of real and virtual environments is that it is typical for the latter to involve the degradation or elimination of various sources of sensory information that are commonly available in the real world.

1.4.2 Sensory Contributions in Virtual Environments

As VE systems have proliferated over the last several decades, a variety of interfaces have emerged, with similarly diverse combinations of spatial information available to users. For example, users of the 1950s-era Sensorama [40]—a widely recognized precursor to modern VE systems—sat in front of a stationary display that presented moving stereoscopic images, but lacked any accompanying head or neck motion. The Sensorama's seat was capable of moving and thus provided a bit of sensorimotor information, but the users' feet and legs remained relatively stationary, lacking proprioceptive and kinesthetic input. Surround audio provided pre-programmed spatial auditory information, and fans created the illusion of wind blowing from a certain direction across the users' skin, which could be used to indicate the direction of one's movement. Even synthetic smells could inform the user of being near one location versus another. Of course the Sensorama was not a true VE system inasmuch as users could not actively navigate or interact with their environment, but rather passively took in a pre-programmed experience. Nonetheless, this early system captures the types of sensory limitations and tradeoffs that are inherent in all VE designs. In this final section of our chapter, we discuss three fundamentally different types of VE systems: (a) desktop; (b) CAVE (Cave Automatic Virtual Environment); and (c) HMD-based systems. Examples of these types of systems are depicted in Fig. 1.1. For each general type of system, we summarize the types of sensory information that are available as well as which sources are absent, degraded, or limited, and what is known about performance differences as a result of these limitations on sensory information. Finally, general use cases under which each of these systems excels or suffers are considered.

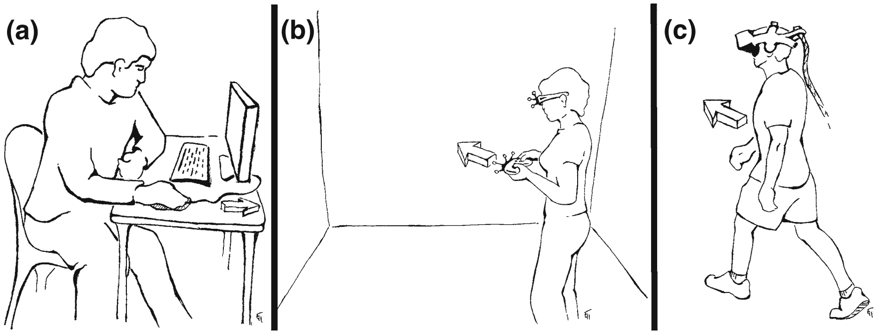

Fig. 1.1 **a** A user of a typical desktop VE system, which offers comparatively impoverished sensory information to users. The primary navigation interface in this case is the mouse and keyboard. Somatosensory, ideothetic, and inertial sensory information about virtual movements are not available, as the user's body is stationary. Visual field-of-view is also typically limited in desktop displays. **b** A user in a typical CAVE system, which visually surrounds her. The user is free to turn and look about and she can move freely over a small area. The primary navigation interface in this case is a wand-type device. CAVEs offer relatively wide and physically-accurate visual field-of-view and typically feature surround-sound audio, but sacrifice veridical somatosensory, idiothetic, and inertial sensory feedback over large ranges. **c** A user in a typical HMD system. The primary navigation interface is the user, who can walk and turn naturally. Range of navigation is limited in some systems by the cable length or, if the rendering computer is worn, the size of the tracking area. Visual field-of-view is typically limited in HMDs

1.4.2.1 Desktop Virtual Environments

Classic Desktop Systems

In a traditional desktop VE (Fig. 1.1a), a stationary user is seated in front of a display and uses a set of arbitrary controls to navigate through a virtual space (e.g., press "w" to move forward). This can include simulations that are presented on desktop computers, laptops, or on a television using a gaming console. While such simulations can contain accurate positional audio and optic flow that inform users about the simulated movement of the depicted viewpoint, the simulation presents nearly all other spatial senses with information that does not match. The user's idiothetic senses will report—quite accurately—that she is sitting still even though the images on the screen simulate her movement. Likewise, traditional efferent information about the user's intended movements (e.g., "walk 2 m forward") are replaced with information about intended button presses and joystick movements that map to unknown units of translation and rotation. This leaves the user not only in a state of sensory conflict, but also removes sensory-motor feedback that can be useful for accurate path integration [16, 76]. Above, we noted that idiothetic information seems to be particularly important for tracking rotational changes (e.g., [55]), and as such, it can logically be expected that users of desktop VE will have relative difficulty with updating their current heading and integrating multiple legs of a journey to keep track of their orientation within a virtual world. This may lead to users becoming lost and disoriented more often, which is perhaps the reason that mini-maps and virtual compasses are so prevalent in desktop game environments.

Given the difficulty of spatial updating in a desktop VE, it might reasonably be expected that users would have difficulty integrating their experiences into a coherent survey representation or cognitive map of the VE. The evidence here is less clear. While there is evidence that people may be able to form survey representations from desktop VEs in some cases (e.g., [86, 89, 111]), there are also examples of users failing to piece together relatively simple spatial layouts after desktop navigation. For example, Stanton et al. [97] had participants explore two opposite legs of a virtual kite-shaped maze on a desktop VE system. The task required participants to travel from corner 1–2 and back, and later from corner 3–4 and back without ever traveling between these two routes. Despite repeated practice locating each of the four corners and the presence of distinct visual landmarks with which to locate these places and paths relative to each other, participants were unable to take novel shortcuts at above-chance levels in subsequent testing. It is unclear whether this same result would have been found in VEs that offer wide fields of view (such as CAVES, discussed below) or that incorporate body-based senses (such as HMD-based systems, discussed below).

Desktop + Motion Controller

Many modern gaming systems and some desktop simulation systems allow the user to interact with simulations using naturalistic motions. The best known of these systems are perhaps the Nintendo Wii and Microsoft Kinect, both of which can be used either on their respective game consoles or on a computer. Such systems leverage the idea of including a user's body into the desktop simulation experience. Indeed, Microsoft marketed the Kinect system with the slogan, "You are the controller." These interfaces are still in their infancy, making it hard to draw firm conclusions about the impact that they have on the user's ability to form accurate spatial knowledge. There is also substantial variation in the way that users interact with various devices both between platforms and between different simulations on the same platform. For example, different simulations of the Wii might use gestures from one of the user's hands, posture information from a balance board, traditional button-presses, or some combination of the above. Similarly, the Kinect performs full-body tracking of the user that could be directly mapped onto the movements of one's avatar, processed to extract pre-defined control gestures, or merely recorded and post-processed. Other accessories, like the Wii balance board, Dance Revolution dance pad, or even the early Nintendo Power Pad could be leveraged to implement a walking-in-place navigation interface.

While it is difficult to make specific claims about the impact of motion controls on spatial knowledge, there are some generalities that can be drawn from these types of interaction. First, including one's body into the user interface of a desktop simulation should improve spatial sensing and navigation insomuch as the movements pertain to the user's navigation. For example, leaning left to steer leftward or to initiate a turn provides more accurate efferent information about leftward movement than, say, pushing the left arrow key. It also differentially tenses muscles in the left and right legs to indicate leftward movement and engages the posture that one might adopt when leaning into a leftward turn to counteract the impending inertia. Spatial knowledge acquisition may or may not be helped, however, by other physical actions such as waving in a defined pattern to call-up a menu or even flicking one's wrist upward to jump—an action typically performed with one's legs. Indeed, this is an area ripe for further research. It is well known that spatial congruence between controls and actions can facilitate users' behavior (e.g., [79, 93]). In the previous example, flicking one's wrist upward to jump would certainly be preferable to some other arbitrary motion—say, rotating one's wrist leftward—given that the upward motion is congruent with an upward jumping action. Whereas the research is unequivocal in favoring spatially mapped controls relative to arbitrary controls, it is unclear whether well-mapped controls are as beneficial to users as the actual motions they replace. In other words, flicking one's wrist upward to jump is certainly better than rotating it leftward, but is it as good as actually jumping? From a perceptual standpoint, it seems unlikely. Users of these systems are still largely stationary or confined to a small area with a fixed toward-the-screen orientation, so some deficits in spatial knowledge are still likely.

Uses for Desktop VEs

While navigation in a desktop VE may be the most limiting from a multi-sensory perspective, there are compelling reasons to use desktop systems, and perhaps even several uses for which it is ideal. Desktop VEs are clearly superior to CAVEs and HMDs on many non-spatial dimensions, including system cost, convenience, and availability to a wide audience. Thus, one might be willing to trade off the fidelity of spatial perception and learning in favor of other factors. Additionally, many computer simulations require little navigation or environmental knowledge. Specifically, when simulating a single virtual object or a small number of objects—a new product design, for example—it may more intuitive to present the VE as a desktop simulation. Users can sit at a desk or table and interact with the virtual object on that table in a similar manner to which they might sit and interact with a physical object on the table. If a VE does not require users to travel, or it is relatively unimportant for them to stay oriented or update their movements, then a desktop VE may be best suited to that particular task.

Conversely, for simulations in which navigation is crucial or in which accurate spatial perception and learning is desired, a desktop VE is likely to be a poor choice. For example, a training simulation designed to teach elite police forces to raid a building would be best implemented in a system that enabled users to physically move, crouch, look around, and remain keenly aware of their position and heading.

1.4.2.2 CAVE Systems

Classic CAVE Systems

A *CAVE* [23] consists of multiple large projected surfaces—traditionally at least three walls that surround a user, and sometimes a floor, ceiling, or fourth wall—that form a booth in which one or more users stand while interacting with the VE (Fig. 1.1b). CAVE systems differ from traditional desktop VEs in that users do not merely sit and view a simulation on a small screen, but rather step physically into the simulated space because the projected surfaces surround them. Such systems thus are able to engage kinesthetic, vestibular, somatosensory, and efferent systems. For example, instead of pressing a button to rotate the virtual viewpoint to one's left, a user need only turn his head or shift his feet. Perhaps most strikingly, however, CAVE displays are generally much more visually immersive than desktop VEs and can present more complete optic flow to inform users about their movements. Access to a wide field of view, and particularly peripheral vision, has been shown to be beneficial for gauging the speed [106] and accuracy of navigation [2, 95]. While typical desktop displays present a simulated view frustum[1] with a horizontal field of view (HFOV) of around 60°, a CAVE display typically displays more than 180° HFOV with three walls or a full 360° HFOV with four walls. Likewise, systems with a floor and/or ceiling will offer significantly improved vertical field of view (FOV). Indeed, a six-walled CAVE is able to simulate a full field of optical flow perfectly regardless of the user's facing direction, although a user's actual viewing angle may be artificially limited if the system requires users to wear shutter glasses. These wide FOVs not only allow peripheral visual information to be displayed, but enable it to be displayed to a physically appropriate location on the user's retina. It is well known that reducing the field of view, for example in an HMD [2] or a desktop VE with a physically matched viewing angle [5], can reduce performance in spatial tasks. Indeed, when normally-sighted users wear a reduced-FOV HMD, they effectively reproduce the inaccuracies in knowledge of spatial layout that are typical of patients who have peripheral field loss [27].

Valid and realistic body-based sensory information in CAVEs, however, is limited to head, neck, and eye movements. Users cannot typically walk or travel more than a few steps in any direction (but see below), placing real limits on the amount of podokinesthetic sensory information that can be acquired during navigation. As mentioned above, the lack of incorporation of body-based senses—particularly for navigational tasks requiring knowledge of orientation changes—can impede learning and produce less accurate spatial knowledge ([88]; see also [16]).

[1] Desktop VEs are rarely constructed to match the physical viewing angle (the visual extent taken up by the monitor) and the virtual field of view (the amount of the VE that is visible). A user's physical viewing angle can vary greatly depending on her distance from the display, and it has been shown both that users do not notice this discrepancy [56] and that a wider field of view is advantageous in many tasks (e.g., [5]).

CAVE + Locomotion

In some CAVE systems, users are able to navigate through a computer simulated environment by making walking motions rather than by using a joystick or other controls. This may be done either by tracking the motion of a user's legs while stepping in place or while implementing some forms of redirected walking (e.g., [82]) or by having users walk on a traditional treadmill (e.g., [69]) or omnidirectional treadmill [22, 25, 96] or even while making stepping or leaning motions on a Wii balance board (e.g., [44]). But as we discussed above with current motion controllers for desktop systems, this type of interface increases the *involvement* of idiothetic and efferent systems, but does not necessarily provide *accurate* idiothetic and efferent information. Of course vestibular information in such systems also tends to be quite limited. The sensation of stepping in place is not the same as walking forward, for example. As discussed below, this limitation is not unique to CAVEs, but also applies to any VE system with limited movement range. Tethered HMDs, single-screen projections, and motion-controlled desktop or gaming platforms (e.g., Wii, Kinect) can all be configured to permit the user some small area in which to move, but inherently prevent users from having true, natural mobility through the VE. Users are aware at some level that large movements are not possible within the physically available space, and this knowledge can interfere with even well-simulated motion. When studying illusory motion (vection) in the laboratory, for example, informing participants that their chair is incapable of moving can delay the onset of vection [61]. Alternatively, seating participants on a rotating stool that *might* rotate, raising their feet from the floor, and providing subtle vibrations to the stool to haptically simulate rotation can facilitate even weak vection effects such as auditory vection [84].

Uses for CAVEs

While CAVEs offer only limited range of motion, they are by far the most visually immersive type of VE system currently available. These attributes make a CAVE ideal for applications in which users do not need to move more than a few steps or in which peripheral vision is crucial. A CAVE is ideal for any type of vehicle simulator, in which the user remains seated in a cockpit and uses realistic controls to operate the virtual vehicle. Although even in this case, vestibular and somatosensory inputs would not be accurately simulated without some type of motion platform to generate acceleration forces. A CAVE would also be ideal to train users who will subsequently be operating in a confined space, such as medics gathered around a patient, astronauts working inside a space station capsule, an aircraft technician learning to repair a virtual jet engine, or a machine operator learning to run a dangerous press on a factory floor. Like desktop VEs, CAVEs are also well suited to visualizing small spaces such as a prototype shelf layout in a retail store, or single virtual objects such as a sculpture or an engine design. In these cases, extended foot navigation is unnecessary. Indeed, it has become increasingly common for large manufacturers to use CAVEs in the design process of creating automobiles, farm equipment, airplanes, jet engines, and other devices.

On the other hand, CAVEs are not particularly ideal when a large amount of ambulatory navigation is necessary. Navigation in these cases is conceptually similar to operating a glass elevator floating in the VE—a user can walk a few steps back and forth within the elevator, but ultimately must use some other controls to move the elevator through space. CAVEs that incorporate a natural walking interface, such as an omni-directional treadmill, are more suited to large-scale navigation tasks but are relatively rare. Stepping in place in a traditional CAVE can also approximate natural walking to a degree, but is a suboptimal solution relative to using a treadmill system or walking in an HMD.

1.4.2.3 HMD-Based Systems

Classic HMD-Based System

HMD-based VEs allow users to wear the virtual display and carry it with them during movement (Fig. 1.1c). Unlike CAVEs or desktop VEs, which are in a fixed location and require the user to remain in a relatively fixed space, HMDs can allow users to roam freely around a larger area. From the perspective of providing or simulating sensory information about space, this is an excellent way to simulate spatial sensory information accurately. Naturalistic navigation through virtual spaces is accomplished by physically walking, turning one's body, or looking around. All of the spatial senses including vision, audition, proprioception, vestibular, and efference can access veridical information about one's movements and current state in the simulated environment, with the exception that peripheral visual information is often lacking or absent due to the decreased FOV. This is particularly true relative to CAVEs, which can offer a FOV equal to that of natural viewing, minus any obstruction of shutter glasses. However with HMDs, the loss of visual information in the periphery can vary widely depending on the specific device being used (e.g., FOVs ranging from below 30° to 150° or more). Despite offering a view frustum similar to that of a desktop VE display, HMDs can create a stronger sense of visual immersion by allowing users to turn their heads and look around, giving the impression that the virtual world truly surrounds them, and also by using blinders that occlude a user's view of the surrounding physical environment. This prevents the visual experience for users of desktop VEs in which perceived movement in the optic flow of the VE conflicts with perceived stability of the world that surrounds the monitor.

Despite the potential of HMDs to provide natural walking and complete sensory feedback for navigation, this benefit has been limited historically by a lack of space available for navigation. While it is relatively easy to create a very large VE, the HMD systems that portray them have typically been confined to a small physical area due to (a) the size of one's VE facility, such as a small laboratory room, (b) the range and capabilities of available motion tracking equipment, (c) the fixed-length cable that tethers the HMD to a rendering computer, or often (d) a combination of these factors. The result, of course, is that the utility of full body-based sensory input during navigation becomes much like users of a CAVE—full body-based sensory input is available, as long as the user does not take more than a few steps. These space constrains have been addressed in much the same manner as CAVEs. While users may control their orientation by turning their head or body, linear movements are relegated to a joystick control or accomplished by walking in place (e.g., [35, 104]).

Altered HMD-Systems

A variety of methods have been employed to allow users to navigate through HMD-based VEs in a way that provides sensory information from all modalities. Perhaps the simplest of these methods is to increase the available physical space and untether the user. If sufficiently large physical space is available (e.g., a gymnasium or airplane hangar), recent advances in motion tracking technology have made it is possible to track user motion over much larger volumes with sufficient resolution and accuracy. Optical tracking systems, for example, are now available with high resolutions that can differentiate between distal objects, fast update rates, and on-board graphics processing capabilities. Software advances have also made it possible to chain a large number of cameras together. Likewise, it is possible to untether the HMD by either transmitting rendered images to the user wirelessly (e.g., [65]) or by having the user wear a high-powered but portable rendering computer (e.g., [109]). Our own large-scale HMD facility [109] for example, uses these approaches and provides users more than 1,100 m^2 of tracked space in which to walk. It is even possible to situate users in large, open, outdoor spaces while wearing rendering and tracking equipment in order to simulate very large VEs without sacrificing naturalistic navigation and full idiothetic sensory feedback by employing a combination of inertial and GPS position tracking (e.g., [6]).

An alternative, but much more common approach has been to employ specialized hardware navigation interfaces. Indeed, a wide range of devices has been created to permit navigation in a tethered HMD, including omnidirectional treadmills [25], roller skates [49], unicycles [78], stepping platforms [51], robotic floor tiles [50], or discs of ball bearings [48]. Many of these navigation interfaces can be used interchangeably with a CAVE or HMD display, with the important consideration that an HMD will occlude users' view of the navigation device. Hollerbach [46] has written a review of many such devices, along with their advantages and shortcomings for virtual navigation, and the implications of permitting (as in a CAVE), or not permitting (as in an HMD), immersed users to view the navigation interface device. In such a system, natural gaits are possible, providing accurate proprioceptive and efference information to the user. Inertial information, however, will be in conflict in such cases as the user remains relatively stationary in the treadmill's center. Consider the recently developed Virtusphere [65], for example. Untethered users walk inside of a hollow sphere that sits on a base of rollers. Because the sphere has its own mass, it will not stop, start, or change directions with a high degree of responsiveness, and users must essentially recalibrate their movements to account for the movement of the surface under their feet.

As an alternative to these hardware-based solutions, several researchers have undertaken the development of software-based systems to allow physical navigation through VEs that are larger than the available physical space. Such techniques include combinations of redirected walking [81], motion compression (e.g., [71]), and reorienting users who approach the space's limits (e.g., [74]). These techniques, which we will refer to in aggregate as redirected walking, work by subtly altering the virtual display in an effort to induce a change in the user's movements. For example, by slowly rotating the VE leftward about a user's viewpoint while she is attempting to walk forward, she will—without realizing it—adjust her course to the left. When done correctly, a user can be induced essentially to walk in circles and thus to use a relatively small physical space in navigating a much larger virtual one. Again, these techniques introduce non-veridical idiothetic sensory input that conflicts with the visual and auditory sensory input that users perceive regarding their movements. However, these conflicts are typically designed to be below consciously detectable thresholds and seem not to impact spatial learning [45] or increase motion sickness [100].

Uses for HMD Systems

Like other VE systems, there are situations for which an HMD-based system may be more or less appropriate. If it is important for users to be able to navigate by means of natural ambulation—for example, when familiarizing users with a large-scale environment—then an un-tethered HMD would be particularly appropriate. Additionally, HMDs are ideal for occluding users' view of the surrounding environment, thereby removing visual distractions and increasing the sense of immersion. For example, if a user is seated on a motion platform (e.g., [9]), it may be ideal to remove it from view so that the resulting motion appears to be coming from a virtual source and not a visible physical source.

Conversely, HMDs are not ideal for situations in which it would be advantageous for users to see their own hands, feet, or body within the VE—an inherent feature of CAVEs. Similarly, multiple users can gather around a desktop VE or step into the same CAVE to share a virtual space. Because a user wearing an HMD is essentially blindfolded to the physical world, any additional users and the user's own body must be simulated within the VE, which can require substantial additional resources in terms of motion tracking, communication, and computational load. Finally, HMDs are adequate but perhaps not ideal at rendering close booth-sized environments that require little or no navigation, such as a cockpit or driving simulator. As mentioned above, the visual expanse of a CAVE makes it very well suited to these situations. Because an HMD typically restricts the user's FOV, the same scene will require substantially more head movement to apprehend in an HMD relative to a CAVE, which probably adversely affects the quality of acquired spatial knowledge [27].

1.5 Conclusion

In this chapter, we have enumerated and briefly described the contributions of three external (visual, auditory, and somatosensory), two internal (vestibular and kinesthetic), and two efferent (efference copy, and attention allocation) sources of information that collectively and routinely inform humans about their place in the environment. With few exceptions, each of these information sources in isolation can be sufficient for acquiring at least rudimentary knowledge of the environment and one's movement through it. At the same time, because of our flexibility in handling various combinations of spatial information, none appears to be truly necessary, even for generating relatively abstract and sophisticated knowledge about one's environment. Thus it is a complex endeavor for researchers to determine the relative contributions of different sources of sensory information when—as is typical in real world situations—most or all are differentially available. Nonetheless, a great deal of exciting contemporary research has begun to describe and understand how different sources of information are acquired, weighted, and combined in the creation and expression of our environmental knowledge.

These issues have very real consequences for users of virtual environments. Currently all VE systems must make some sacrifices on the information that they provide to users. Desktop systems and CAVES eliminate or degrade many body-based sources of spatial information, while HMD systems limit the quality of visual information and (typically) the range of full-body movement. Research on the relative importance of these different sources of sensory information is critical for making intelligent decisions about how to use and choose among the many different options for VE systems. The issue of conflicting or degraded sensory information in VEs has additional practical importance when one notes that discrepancies among sensory information are widely believed to underlie the onset of simulator sickness (e.g., [37, 115]) and are also thought to impact user's sense of "presence" and "immersion" in the VE [11].

Perhaps equally important for the application of sensory psychology to VE technology is an appreciation of the role of efferent information—particularly attention allocation—to the acquisition of knowledge of one's environment. The literature is generally consistent with the notion that attention allocation is perhaps the strongest contributor to spatial knowledge among efferent sources [18]. Navigation interfaces for VEs that demand a user's limited attentional resources are thus especially likely to impede one's ability to learn about their environment. More generally, because no VE is yet able to present a perfectly full and veridical sensory experience to its user, it is important for researchers and VE developers to understand the circumstances, tasks, and goals for which these numerous information sources are most critical.

References

1. Alais D, Burr D (2004) No direction-specific bimodal facilitation for audiovisual motion detection. Cogn Brain Res 19:185–194
2. Alfano PL, Michel GF (1990) Restricting the field of view: perceptual and performance effects. Percept Motor Skills 70:25–45
3. Alibali MW (2005) Gesture in spatial cognition: expressing, communicating and thinking about spatial information. Spatial Cogn Comput 5:307–331
4. Allen GL, Siegel AW, Rosinski R (1978) The role of perceptual context in structuring spatial knowledge. J Exp Psychol Hum Learn Mem 4:617–630
5. Arsenault R, Ware C (2002) Frustum view angle, observer view angle and VE navigation. In: Vidal CA, Kimer C (eds) Proceedings of the V Simposio de Realidade Virtual, Brazilian Computer Society, pp 15–25
6. Bachmann E, Calusdian J, Hodgson E, Yun X, Zmuda M (2012) Going anywhere anywhere—creating a low cost portable immersive VE system. Manuscript submitted for publication
7. Bakker NH, Werkhoven PJ, Passenier PO (1999) The effects of proprioceptive and visual feedback on geographical orientation in virtual environments. Presence Teleoperators Virtual Environ 8:36–53
8. Becker W, Nasios G, Raab S, Jürgens R (2002) Fusion of vestibular and podokinesthetic information during self-turning towards instructed targets. Exp Brain Res 144:458–474
9. Berger DR, Schulte-Pelkum J, Bülthoff HH (2010) Simulating believable forward accelerations on a Steward motion platform. ACM Trans Appl Percept 7(1):1–27 (Article 5)
10. Besson P, Richiardi J, Bourdin C, Bringoux L, Mestre DR, Vercher J (2010) Bayesian networks and information theory for audio-visual perception modeling. Biol Cybern 103:213–226
11. Biocca F, Kim J, Choi Y (2000) Visual touch in virtual environments: an exploratory study of presence, multimodal interfaces, and cross-modal sensory illusions. Presence Teleoperators Virtual Environ 10:247–265
12. Blakemore SJ (2003) Deluding the motor system. Conscious Cogn 12:647–655
13. Bruggeman H, Piuneu VS, Rieser JJ (2009) Biomechanical versus inertial information: stable individual differences in perception of self-rotation. J Exp Psychol Hum Percept Perform 35:1472–1480
14. Butler JS, Smith ST, Campos JL, Bülthoff HH (2010) Bayesian integration of visual and vestibular signals for heading. J Vis 10:1–13
15. Campos JL, Byrne P, Sun HJ (2010) The brain weights body-based cues higher than vision when estimating walked distances. Eur J Neurosci 31:1889–1898
16. Chance SS, Gaunet F, Beall AC, Loomis JM (1998) Locomotion mode affects the updating of objects encountered during travel: the contribution of vestibular and proprioceptive inputs to path integration. Presence Teleoperators Virtual Environ 7:168–178
17. Cheng K, Shettleworth SJ, Huttenlocher J, Rieser JJ (2007) Bayesian integration of spatial information. Psychol Bull 133:625–637
18. Chrastil ER, Warren WH (2012) Active and passive contributions to spatial learning. Psychon Bull Rev 19:1–23
19. Christou CG, Bülthoff HH (1999) View dependence in scene recognition after active learning. Mem Cogn 27:996–1007
20. Cohen HS (2000) Vestibular disorders and impaired path integration along a linear trajectory. J Vestib Res 10:7–15
21. Couclelis H (1996) Verbal directions for way-finding: space, cognition, and language. In: Portugali J (ed) The construction of cognitive maps. Kluwer, Dordrecht, The Netherlands, pp 133–153
22. Cruz-Neira C, Reiners D, Springer JP (2010) An affordable surround-screen virtual reality display. Soc Inf Disp 18:836–843
23. Cruz-Neira C, Sandin DJ, DeFanti TA, Kenyon RV, Hart JC (1992) The cave: audio visual experience automatic virtual environment. Commun ACM 35:64–72

24. Cutting JE, Vishton PM (1995) Perceiving layout: the integration, relative dominance, and contextual use of different information about depth. In: Epstein W, Rogers S (eds) Handbook of perception and cognition: Vol. 5: perception of space and motion. Academic Press, New York
25. Darken RP, Cockayne WR, Carmein D (1997) The omni-directional treadmill: a locomotion device for virtual worlds. In: Proceedings of ACM symposium on user interface software and technology, pp 213–221
26. Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415:429–433
27. Fortenbaugh FC, Hicks JC, Turano KA (2008) The effect of peripheral visual field loss on representations of space: evidence for distortion and adaptation. Investig Ophthalmol Visual Sci 49:2765–2772
28. Frenz H, Bremmer F, Lappe M (2003) Discrimination of travel distance from 'situated' optic flow. Vis Res 43:2173–2183
29. Gallistel CR (1990) The organization of learning. MIT, Cambridge
30. Gibson JJ (1958) Visually controlled locomotion and visual orientation in animals. Br J Psychol 49:182–194
31. Gibson JJ (1979) The ecological approach to visual perception. Houghton Mifflin, Boston
32. Glasauer S, Amorim MA, Vitte E, Berthoz A (1994) Goal directed linear locomotion in normal and labyrinthine-defective subjects. Exp Brain Res 98:323–335
33. Goldin SE, Thorndyke PW (1982) Simulating navigation for spatial knowledge acquisition. Hum Factors 24:457–471
34. Golledge R, Klatzky R, Loomis J (1996) Cognitive mapping and wayfinding by adults without vision. In: Portugali J (ed) The construction of cognitive maps. Kluwer, The Netherlands, pp 215–246
35. Grant SC, Magee LE (1998) Contributions of proprioception to navigation in virtual environments. Hum Factors 40:489–497
36. Guidice NA, Bakdash JZ, Legge GE (2007) Wayfinding with words: spatial learning and navigation using dynamically updated verbal descriptions. Psychol Res 71:347–358
37. Harm DL (2002) Motion sickness neurophysiology, physiological correlates, and treatment. In: Stanney KM (ed) Handbook of virtual environments: design, implementation, and applications. Erlbaum, Mahwah, pp 637–661
38. Harris LR, Jenkin M, Zikovitz DC (2000) Visual and non-visual cues in the perception of linear self-motion. Exp Brain Res 135:12–21
39. Hay JC, Pick HL Jr (1965) Visual capture produced by prism spectacles. Psychon Sci 2:215–216
40. Heilig M (1992) El cine de futro: the cinema of the future. Presence Teleoperators Virtual Environ 1:279–294 (Original work published 1955)
41. Heft H (1996) The ecological approach to navigation: A Gibsonian perspective. In: Portugali J (ed) The construction of cognitive maps. Kluwer Academic Publishers, Dordrect, pp 105–132
42. Heft H (1997) The relevance of Gibson's ecological approach for environment-behavior studies. In: Moore GT, Marans RW (eds) Advances in environment, behavior, and design, vol 4. Plenum, New York, pp 71–108
43. Hettinger LJ (2002) Illusory self-motion in virtual environments. In: Stanney KM (ed) Handbook of Virtual Environments. Erlbaum, New Jersey, pp 471–492
44. Hilsendeger A, Brandauer S, Tolksdorf J, Fröhlich C (2009) Navigation in virtual reality with the wii balanceboard^{TM}. In: GI-Workshop Virtuelle und Erweiterte Realität, ACM
45. Hodgson E, Bachmann E, Waller D (2011) Redirected walking to explore virtual environments: assessing the potential for spatial interference. ACM Trans Appl Percept 8, Article 22
46. Hollerbach JM (2002) Locomotion interfaces. In: Stanney KM (ed) Handbook of virtual environments: design, implementation, and applications. Erlbaum, New Jersey, pp 239–254
47. von Holst E, Mittlestaedt H (1950) Das reafferenz princip: (Wedlselwirkungen zwischen Zentrainervensystem und Peripherie.) Die Naturwissenschften, 37, 464–476. Translated in:

Dodwell PC, ed (1971) Perceptual processing: stimulus equivalence and pattern recognition (pp. 41–72). Appleton-Century-Crofts, New York
48. Huang J, Chiu W, Lin Y, Tsai M, Bai H, Tai C, Gau C, Lee H (2000) The gait sensing disc: a compact locomotion device for the virtual environment. In: Proceedings of the international conference in central Europe on computer graphics, visualization and interactive digital media. Pilsen, Czech Republic
49. Iwata H (2000) Locomotion interface for virtual environments. In: Hollerbach J, Koditshek D (eds) Robotics research: the ninth international symposium. Springer-Verlag, London, pp 275–282
50. Iwata H, Yano H, Fukushima H, Noma H (2005) CirculaFloor. IEEE Comput Graph Appl 25:64–67
51. Iwata I, Yano H, Nakaizumi F (2001) Gait master: a versatile locomotion interface for uneven virtual terrain. In: Proceedings of IEEE virtual reality conference, pp. 131–137
52. Jürgens R, Becker W (2006) Perception of angular displacement without landmarks: evidence for Bayesian fusion of vestibular, optokinetic, podokinesthetic, and cognitive information. Exp Brain Res 174:528–543
53. Jürgens R, Boß T, Becker W (1999) Estimation of self-turning during active and passive rotation in the dark. Exp Brain Res 128:491–504
54. Kearns MJ, Warren WH, Duchon AP, Tarr MJ (2002) Path integration from optic flow and body senses in a homing task. Perception 31:349–374
55. Klatzky RL, Loomis JM, Beall AC, Chance SS, Golledge RG (1998) Updating an egocentric spatial representation during real, imagined, and virtual locomotion. Psychol Sci 9:293–298
56. Kubovy M (1986) The psychology of linear perspective and renaissance art. Cambridge University Press, Cambridge
57. Lackner JP, DiZio P (2005) Vestibular, proprioceptive, and haptic contributions to spatial orientation. Ann Rev Psychol 56:115–147
58. Lappe M, Bremmer F, van den Berg AV (1999) Perception of self-motion from visual flow. Trends Cogn Sci 3:329–336
59. Larish JF, Flach JM (1990) Sources of optical information useful for perception of speed of rectilinear self-motion. J Exp Psychol Hum Percept Perform 16:295–302
60. Lederman SJ, Klatzky RL (2009) Haptic perception: a tutorial. Atten Percept Psychophys 71:1439–1459
61. Lepecq J, Giannopulu I, Baudonnière P (1995) Cognitive effects on visually induced body motion in children. Perception 24:435–449
62. Loomis JM, Klatzky RL, Golledge RG, Philbeck JW (1999) Human navigation by path integration. In: Golledge RG (ed) Wayfinding: cognitive mapping and other spatial processes. Johns Hopkins, Baltimore, pp 125–151
63. Loomis JM, Klatzky RL, Avraamides M, Lippa Y, Golledge RG (2007) Functional equivalence of spatial images produced by perception and spatial language. In: Mast F, Jäncke L (eds) Spatial processing in navigation, imagery, and perception. Springer, New York, pp 29–48
64. May M, Klatzky RL (2000) Path integration while ignoring irrelevant movement. J Exp Psychol Learn Mem Cogn 26:169–186
65. Medina E, Fruland R, Weghorst S (2008) Virtusphere: walking in a human size VR "hamster ball". In: Proceedings of the human factors and ergonomics society, pp 2102–2106
66. Mittelstaedt H (1996) Somatic graviception. Biol Psychol 42:53–74
67. Mittelstaedt ML, Mittelstaedt H (2001) Idiothetic navigation in humans: estimation of path length. Exp Brain Res 139:318–332
68. Nardini M, Jones P, Bedford R, Braddick O (2008) Development of cue integration in human navigation. Curr Biol 18:689–693
69. Neider MB, Gaspar JG, McCarley JS, Crowell J, Kaczmarski H, Kramer AF (2011) Walking and talking: dual-task effects on street crossing behavior in older adults. Psychology and Aging (Advance online publication)
70. Nielsen TI (1963) Volition: a new experimental approach. Scand J Psychol 4:225–230

71. Nitzshe N, Hanebeck UD, Schmidt G (2004) Motion compression for telepresent walking in large target environments. Presence Teleoperators Virtual Environ 13:44–60
72. Noë A (2004) Action in perception. MIT Press, Boston
73. Ohmi M (1996) Egocentric perception through interaction among many sensory systems. Cogn Brain Res 5:87–96
74. Peck TC, Whitton MC, Fuchs H (2009) Evaluation of reorientation techniques for walking in large virtual environments. IEEE Trans Vis Comput Graph 15:121–127
75. Péruch P, Borel L, Magnan J, Lacour M (2005) Direction and distance deficits in path integration after unilateral vestibular loss depends on task complexity. Cogn Brain Res 25:862–872
76. Péruch P, May M, Wartenberg F (1997) Homing in virtual environments: effects of field of view and path layout. Perception 26:301–311
77. Potegal M (1982) Vestibular and neostriatal contributions to spatial orientation. In: Potegal M (ed) Spatial abilities development and physiological foundations. Academic Press, New York, pp 361–387
78. Pratt DR, Barham PT, Locke J, Zyda MJ, Eastman B, Moore T et al (1994) Insertion of an articulated human into a networked virtual environment. In: Proceedings of the 1994 AI, simulation, and planning in high autonomy systems conference. University of Florida, Gainesville
79. Prinz W (1997) Perception and action planning. Eur J Cogn Psychol 9:129–154
80. Proffitt DR (2006) Distance perception. Curr Dir Psychol Sci 15:131–135
81. Razzaque S (2005) Redirected walking. Doctoral dissertation, University of North Carolina, Chapel Hill
82. Razzaque S, Swapp D, Slater M, Whitton MC, Steed A (2002) Redirected walking in place. In: Muller S, Stuzlinger W (eds) Proceedings of the eurographics workshop on virtual environments. Eurographics Association, pp 123–130
83. Riecke BE, Cunningham DW, Bülthoff HH (2007) Spatial updating in virtual reality: the sufficiency of visual information. Psychol Res 71:298–313
84. Riecke B, Feuereissen D, Rieser JJ (2009) Auditory self-motion simulation is facilitated by haptic and vibrational cues suggesting the possibility of actual motion. ACM Trans Appl Percept 6, Article 20
85. Riecke BE, van Veen HAHC, Bülthoff HH (2002) Visual homing is possible without landmarks: a path integration study in virtual reality. Presence Teleoperators Virtual Environ 11:443–473
86. Rossano MJ, West SO, Robertson TJ, Wayne MC, Chase RB (1999) The acquisition of route and survey knowledge from computer models. J Environ Psychol 19:101–115
87. Ruddle RA (2001) Navigation: am I really lost or virtually there? In: Harris D (ed) Engineering psychology and cognitive ergonomics, vol 6. Ashgate, Burlington, pp 135–142
88. Ruddle RA, Lessels S (2009) The benefits of using a walking interface to navigate virtual environments. ACM Trans Comput Hum Interact 16:1–18
89. Ruddle RA, Payne SJ, Jones DM (1997) Navigating buildings in "desk-top" virtual environments: experimental investigations using extended navigational experience. J Exp Psychol Appl 3:143–159
90. Sandvad J (1999) Auditory perception of reverberant surroundings. J Acoust Soc Am 105:1193
91. Schenkman BN, Nilsson ME (2010) Human echolocation: blind and sighted persons' ability to detect sounds recorded in the presence of a reflected object. Perception 39:483–501
92. Shelton AL, McNamara TP (1997) Multiple views of spatial memory. Psychon Bull Rev 4:102–106
93. Shin YK, Proctor RW, Capaldi EJ (2010) A review of contemporary ideomotor theory. Psychol Bull 136:943–974
94. Sholl MJ (1989) The relation between horizontality and rod-and-frame and vestibular navigational performance. J Exp Psychol Learn Mem Cogn 15:110–125
95. Sholl MJ (1996) From visual information to cognitive maps. In: Portugali J (ed) The construction of cognitive maps. Kluwer, Dordrecht, pp 157–186

96. Souman JL, Giordano PR, Schwaiger M, Frissen I, Thümmel T, Ulbrich H, De Luca A, Bülthoff HH, Ernst MO (2011) CyberWalk: enabling unconstrained omnidirectional walking through virtual environments. ACM Trans Appl Percept 8, Article 25
97. Stanton DEB, Wilson PN, Foreman N (2003) Human shortcut performance in a computer-simulated maze: a comparative study. Spatial Cogn Comput 3:315–329
98. Stoffregen TA, Pittenger JB (1995) Human echolocation as a basic form of perception and action. Ecol Psychol 7:181–216
99. Stratton G (1896) Some preliminary experiments on vision without inversion of the retinal image. Psychol Rev 3:341–360
100. Steinicke F, Bruder G, Jerald J, Frenz H, Lappe M (2009) Estimation of detection thresholds for redirected walking techniques. IEEE Trans Vis Comput Graph 16(1):17–27
101. St. George RJ, Fitzpatrick RC (2011) The sense of self-motion, orientation and balance explored by vestibular stimulation. J Physiol 589:807–813
102. Sun HJ, Campos JL, Chan GSW (2004) Multisensory integration in the estimation of relative path length. Exp Brain Res 154:246–254
103. Telford L, Howard IP, Ohmi M (1995) Heading judgments during active and passive self-motion. Exp Brain Res 104:502–510
104. Templeman JN, Denbrook PS, Sibert LE (1999) Virtual locomotion: walking in place through virtual environments. Presence Teleoperators Virtual Environ 8:598–619
105. Thorndyke PW, Hayes-Roth B (1982) Differences in spatial knowledge acquired from maps and navigation. Cogn Psychol 14:560–589
106. Toet A, Jansen SEM, Delleman NJ (2007) Effects of field-of-view restrictions on speed and accuracy of manoeuvring. Percept Motor Skills 105:1245–1256
107. Turvey MT, Carello C (1986) The ecological approach to perceiving-acting: a pictorial essay. Acta Psychol 63:133–155
108. Van Erp JBF, Van Veen HAHC, Jansen C, Dobbins T (2005) Waypoint navigation with a vibrotactile waist belt. ACM Trans Appl Percept 2:106–117
109. Waller D, Bachmann E, Hodgson E, Beall AC (2007) The HIVE: A Huge Immersive Virtual Environment for research in spatial cognition. Behav Res Methods 39:835–843
110. Waller D, Greenauer N (2007) The role of body-based sensory information in the acquisition of enduring spatial representations. Psychol Res 71:322–332
111. Waller D, Hunt E, Knapp D (1998) The transfer of spatial knowledge in virtual environment training. Presence Teleoperators Virtual Environ 7:129–143
112. Waller D, Loomis JM, Haun DBM (2004) Body-based senses enhance knowledge of directions in large-scale environments. Psychon Bull Rev 11:157–163
113. Waller D, Loomis JM, Steck S (2003) Inertial cues do not enhance knowledge of environmental layout. Psychon Bull Rev 10:987–993
114. Witmer BG, Bailey JH, Knerr BW, Parsons KC (1996) Virtual spaces and real world places: transfer of route knowledge. Int J Hum Comput Stud 45:413–428
115. Yardley L (1992) Motion sickness and perception: a reappraisal of the sensory conflict approach. Br J Psychol 83:449–471
116. Yardley L, Higgins M (1998) Spatial updating during rotation: the role of vestibular information and mental activity. J Vestib Res 8:435–442
117. Yost WA (2001) Auditory localization and scene perception. In: Goldstein EB (ed) Blackwell handbook of perception. Blackwell Publishers, Malden, pp 437–468
118. Zhang H, Mou W, McNamara TP (2011) Spatial updating according to the intrinsic reference direction of a briefly viewed layout. Cognition 119:419–429