2

The Impact of Multimodal-Multisensory Learning on Human Performance and Brain Activation Patterns

Karin H. James, Sophia Vinci-Booher, Felipe Munoz-Rubke

2.1 Introduction

The human brain is inherently a ***multimodal-multisensory*** dynamic learning system. All information that is processed by the brain must first be encoded through sensory systems and this sensory input can only be attained through motor movement. Although each sensory modality processes different signals from the environment in qualitatively different ways (e.g., sound waves, light waves, pressure, etc.), these signals are transduced into a common language in the brain. The signals are then associated and combined to produce our phenomenology of a coherent world. Therefore, the brain processes a seemingly unlimited amount of multisensory information for the purpose of interacting with the world. This interaction with the world, through the body, is multimodal. The body allows one to affect the environment through multiple motor movements (hand movements, locomotion, speech, gestures, etc.). These various actions, in turn, shape the multisensory input that the brain will subsequently receive. The ***feedforward-feedback loop*** that occurs every millisecond among sensory and motor systems is a reflection of these multisensory and multimodal interactions among the brain, body, and environment. As an aid to comprehension, readers are referred to this chapter's Focus Questions and to the Glossary for a definition of terminology.

In the following, we begin by delving deeper into how sensory signals are transduced in the brain and how multimodal activity shapes signal processing. We then provide samples of research that have demonstrated that multimodal interactions with the world, through action, facilitate learning. An overview of research on performance measured by overt behavioral responses in adult and developing populations is followed by examples of research on the effects that multimodal learning has on brain plasticity in adults and children. Together, the behavioral and neuroimaging literature underscore the importance of learning through multimodal-multisensory interactions throughout human development.

2.2 The Multimodal-Multisensory Body

The ultimate utility of a sensory mechanism is to convey information to an organism in the service of eliciting environmentally appropriate action. An interesting question arises in consideration of the inherently multisensory nature of behavior:

How is it that the human perceptual system provides us with seamless experiences of objects and events in our environment?

The difficulty in answering this question lies in one's conception of the role of the human perceptual system. Approaching this question as a ***constructivist*** would lead to a major impasse: How it is that the brain is able to infer meaning from sensory input and translate among sensory modalities, given that these signals have little fidelity to the environmental stimulus by which they were evoked? Further, how are the signals combined given that the signal of one sense is not directly comparable to the signal of another? This impasse is referred to as the ***binding problem*** and is a logical outcome of a constructivist approach to the role of the human perceptual system [Bahrick and Lickliter 2002]. If each sensory modality is transduced into a unique neuronal firing pattern, then the only way to infer the appropriate response to that particular set of sensory input is to effectively combine them into a unified percept. On the other hand, recent theories of perceptual development portray the human perceptual system as a multimodal system that responds to unisensory and multisensory inputs with differential weighting on ***modality-specific stimulus properties*** and ***amodal stimulus properties***, respectively [Stein and Rowland 2011, Stein et al. 2014]. Formally, this theory is referred to as ***intersensory redundancy*** (see Figure 2.1).

Intersensory redundancy is based upon the observation that an organism and its environment are structured such that the sensory systems of an active perceiver will experience certain consistencies and inconsistencies. The detection of these consistencies is related to the ability of an object or event to produce sensory stimuli in more than one sensory system in a spatiotemporally consistent manner. The detection of inconsistencies is related to the inability of an object or event to produce

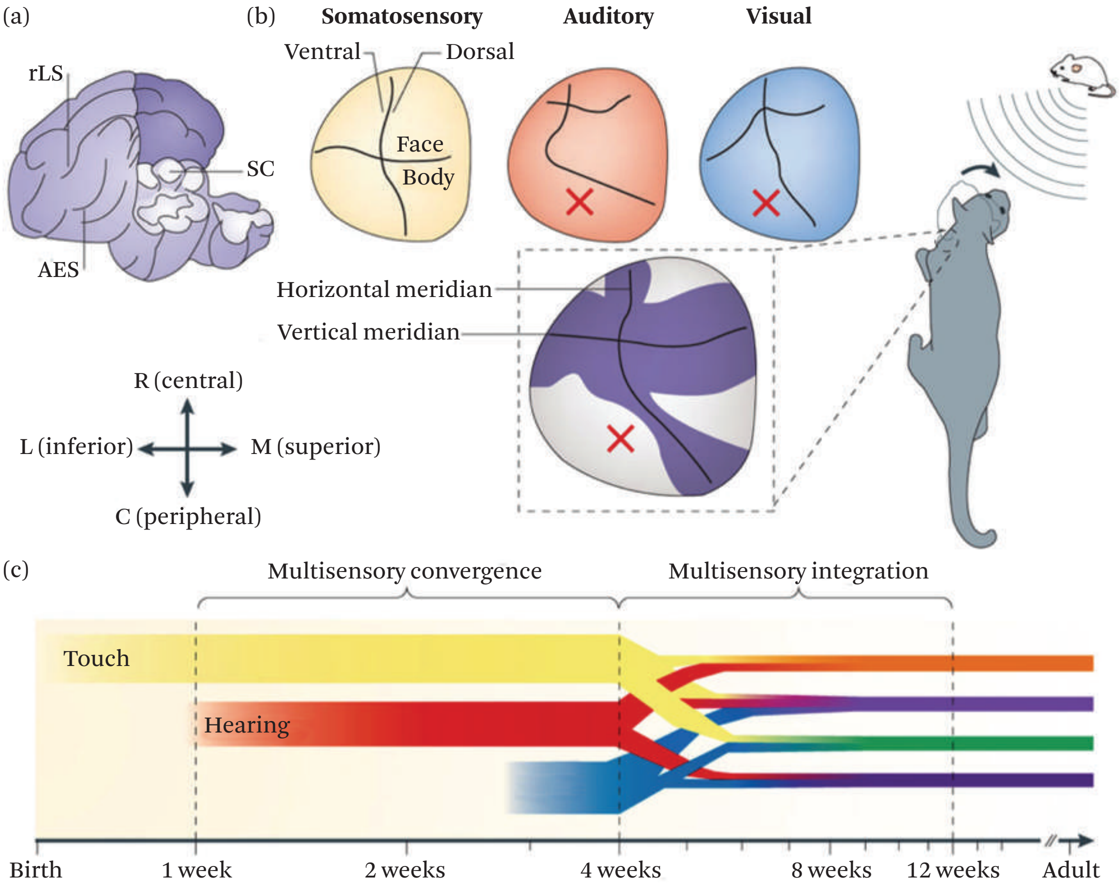

Figure 2.1 **A. Diagram of superior colliculus of the cat with associated cortical connections.** B. The three sensory representations in the superior colliculus (visual, auditory and somatosensory) are organized in an overlapping topographical map. Space is coded anterior-posterior. And stimuli in space are coded by multiple modalities. In adults, this leads to enhancements of neuronal activity and an increase in behavioral responses. C. In early development, the unisensory inputs converge to produce multisensory neurons but these neurons cannot yet integrate their multisensory cross-modal inputs. This latter function only develops around 4 weeks of age. (From Stein et al. [2014])

such sensory stimulation. Whether or not an object or event is capable of producing spatiotemporally consistent multisensory stimulation is determined by whether or not the particular stimulus has properties that can be detected by more than one sensory system. Properties of sensory stimuli that may be detected in more than one sensory system are referred to as ***amodal stimulus properties***. For instance, shape and texture are perceived in both visual and haptic modalities, just as intensity and spatiotemporal information may be perceived in both visual and auditory modalities. Properties of sensory stimuli that may only be detected in one sensory system are referred to as modality-specific properties. As examples, color can only be perceived in the visual modality, whereas tone can only be perceived in the auditory modality. Throughout ontogeny, the interaction between the neural structure of the organism and its environment result in a neural system that processes certain properties best in multisensory contexts and others best in unisensory contexts.

The importance of the distinction between multisensory and unisensory perceptual input is most evident in consideration of the ontogeny of ***multisensory neurons*** in the evolutionarily early subcortical structure of the ***superior colliculus*** (SC). Although multisensory neurons have been found in several brain regions and in many species, the anatomical nature of this region has been most extensively studied in cats [Stein et al. 2014]. The SC is unique, in that decidedly ***unisensory neurons*** from the retina, ear, and/or skin may all converge onto a single SC neuron [Wallace et al. 1992, 1993, Stein and Meredith 1993, Fuentes-Santamaria et al. 2009]. Note that ***convergence*** is not the same as ***integration***, which is something that has recently been shown to be ***experience-dependent***, at least in the sequentially early sensory convergence region of the SC [Wallace et al. 1992, Stein and Rowland 2011, Stein et al. 2014]. In other words, neurons from several sensory systems may converge at one point, but whether or not these signals are integrated into a single output depends upon the history that a give neuron had with receiving sensory signals. SC neurons in kittens are largely unisensory and only become multisensory if they receive stimulation from multiple sensory systems at the same time. Perhaps even more interesting from a developmental perspective: the emergence of multisensory neurons and their ***receptive fields*** are also experience-dependent [Stein and Rowland 2011]. Receptive fields not only change their size through experience, they also change their sensitivity to physical locations. Repeated exposure to a visual stimulus and an auditory stimulus that occur at the same time and at the same physical location will result in overlapping receptive fields. Overlapping receptive fields from more than one sensory system results in faster, stronger, and more reliable neural responses for that combination of stimuli at that location in the future [Stein et al. 2014]. The implications of these phenomena are that events that violate the learned correspondences among sensory modalities are readily detected and, because they converge at such an early stage in sensory processing, are difficult for the perceiver to overcome. Furthermore, projections from SC are widely distributed throughout the cortex and are one of the major pathways by which sensory information reaches the cortex where, presumably, higher-level cognitive functions are carried out, such as object and symbol recognition.

Neural plasticity, the ability of neuronal structures, connections, and processes in the brain to undergo experience-dependent changes, in the adult is well documented with respect to the initial learning of multimodal-multisensory mappings. However, the *re*learning of these mappings is a decidedly more laborious process, though initial and relearning stages both follow a few basic principles. Although several principles could be mentioned here, we will focus on two: ***Multisensory enhancement*** and ***multisensory depression***. First, information about objects and events in the environment can be gleaned by simply relying upon spatiotemporal coincidence to indicate the presence of an object or event; this type of coordinated environmental stimuli contribute to multisensory enhancement. This phenomenon results in an increase in the system's ability to detect the same object or event based on multisensory, or, to use another word, amodal, input in the future. Second, information about objects and events in the environment can be gleaned by simply relying upon spatiotemporal disparities to indicate the separability of objects or events; this type of uncoordinated environmental stimuli contribute to multisensory depression. This phenomenon results in a decrease in the system's ability to detect either of those objects or events based on multisensory input in the future. In the case of multisensory enhancement, amodal cues are emphasized and in the case of multisensory depression, modality-specific cues are more readily emphasized. Thus, the functioning of the SC and corresponding connections appear to be foundational attention mechanisms associated with orienting to salient environmental stimulation based on the system's history of sensory experiences [Johnson et al. 2015].

These principles reflect our knowledge regarding the cellular mechanisms of neural plasticity. The principle of ***Hebbian learning*** is foundational to theories of experience-dependent brain changes as it proposes that axonal connections between neurons undergo activity-dependent changes. It has two basic tenants: (1) when two neurons repeatedly fire in a coordinated manner, the connections between them are strengthened, effectively increasing the likelihood of firing together in the future; and (2) when two neurons repeatedly fire in an uncoordinated manner, the connections between them weaken, effectively reducing the likelihood of firing together in the future [Hebb 1949]. The relevance of this theory to experience-dependent brain changes is most readily understood when considering the differences in brain systems supporting the recognition of objects that would result from *active*, as opposed to *passive*, behaviors. For this discussion, the crucial difference between active and passive behaviors is that the perceiver in active behaviors performs an action on stimuli in the environment. The crucial implication of that difference is that active behaviors are inherently multimodal, involving both action systems and perceptual systems, and multisensory, involving haptic and visual inputs at a minimum. Passive behaviors, on the other hand, often involve the stimulation of only one sensory modality, rendering them unisensory (see Chapter 3).

Therefore, active interactions with the environment, as opposed to passive interactions, are inherently multisensory and multimodal. Not only do they entail input from various effectors (e.g., hands, eyes) based on the action (e.g., reaching to grasp, saccading), but they also simultaneously produce input to numerous sensory modalities (e.g., somatosensation, vision). Therefore, active interactions produce multisensory information that allow for coactivation of signals, resulting in multisensory enhancement. Beyond multisensory enhancement, however, active interactions have been thought to be necessary for any type of percept, be it unisensory or multisensory. That is, without physical movement, sensory information quickly fades. This phenomenon occurs because sensory receptor organs and neurons stop responding to repeated stimulation. One well-known example is that of the ***stabilized image phenomenon***: if eye movements are inhibited, resulting in a stable retinal image, the resultant percept will dissipate within seconds (for review see [Heckenmueller 1965]). Although it is beyond the scope of this chapter to outline sensory and *neural adaptation*, it is important to note that for the brain to code change, and thus detect similarities and differences in incoming percepts, action is necessary.

What effect does this knowledge of the mechanisms that underlie experience-dependent brain plasticity have on our understanding of learning in general? One argument for the usefulness of these findings is that they have led to numerous experiments that have shown that active interactions with the environment facilitate learning more than passive "interactions". In their pivotal work, Held and Hein [1963] showed that kittens that were initially reared without visual experience were unable to learn to understand their world, even after their vision was restored, unless they were given the freedom to move around on their own volition. For comparison with the active learning experience given to this cohort of kittens, an experimental apparatus moved a separate cohort of kittens, who, therefore, experienced "passive" movements. The crucial difference is that the first cohort of kittens received visual stimulation that was contingent with their own self-generated movements, whereas the second cohort experienced visual stimulation that was not contingent with their own movements. This study spurred a wealth of research on the role of active vs. passive experiences on learning. These lines of research have largely confirmed the value of self-generated action for visual perception in many domains.

In the context of the intersensory redundancy discussed earlier, active learning is the avenue by which multimodal-multisensory signals in the brain arise from the body and environment. In what follows, our goal is to provide a brief overview of some of the empirical work on how these multimodal-multisensory signals affect learning. We briefly outline empirical work that has shown that active interactions facilitate learning in surprising ways at times. We discuss findings from behavioral work and from neuroimaging work. The behavioral data show the usefulness of learning through action in numerous domains and the neuroimaging work suggests the mechanisms, at a brain systems level, that support active learning.

2.3 Facilitatory Effects on Learning as a Function of Active Interactions

Actively interacting with one's environment involves multisensory, multimodal, processing. It is therefore, not surprising then, that these types of interactions facilitate learning in many domains. Here we briefly review findings from behavioral experiments that demonstrate the far-reaching beneficial effects that active experience has on learning.

2.3.1 Behavioral Research in Adults

Visual Perception Although somewhat controversial, the majority of research in psychology separates sensation from perception. A primary separation results from findings that perception relies on learning, in that information from prior experience changes how we perceive. When these experiences involve active interaction, we can assume that perception is changed by action. Our first example comes from the phenomenon of *size constancy*. Size constancy refers to our ability to infer that an object maintains its familiar size despite large changes in retinal size from visual sensation, due to distance. Distance cues that lead to size constancy can be from object movement, observer movement, or a combination of the two. Research has shown that size constancy is dependent on observer movement, not object movement. If size is manipulated through object movement, constancy is worse than if size is manipulated through observer movement [Combe and Wexler 2010].

2.3.2 Spatial Localization

It seems intuitive that locomotion is important for understanding the three-dimensional space around us. However, several lines of research support the claim that *self-generated* movement is the key to understanding spatial location of objects. For instance, if we encode the location of an object and then are blindfolded and walk to a new location, we can point to its original location with great accuracy. If we simply imagine moving to the new location [Rieser et al. 1986] or see the visual information that would occur if we were to move, but without actually moving we cannot accurately locate the target object [Klatzky et al. 1998]. Even if an individual is moved in a wheelchair by another person [Simons and Wang 1998] or in a virtual-reality environment, object localization is worse than if the movement were self-generated [Christou and Bülthoff 1999].

2.3.3 Three-dimensional Object Structure

Learning object structure is facilitated by active interactions in numerous ways. For instance, slant perception [Ernst and Banks 2002] and shape from shading cues for object structure [Adams et al. 2004] are both cues for depth perception. Both of these perceptual competencies are facilitated by manual active interaction with the stimuli. Remembering object structure is also facilitated by active manipulation of objects with one's hands. In a series of studies, novel objects were studied through either an active interaction, where participants rotated the 3D images on a computer screen, or through a passive interaction, where participants observed object rotations that had been generated from another participant. By pairing active and passive learning between subjects, one could compare the learning of a given object that resulted from either visual and haptic sensation with the involvement of the motor system (multimodal) or learning through visual sensation alone (unisensory). Subsequently, participants were tested on object recognition and object matching (see Figure 2.2). Results were straightforward: when objects were learned through active interactions, recognition was enhanced relative to learning through passive interactions [Harman et al. 1999, James et al. 2001, 2002]. These results demonstrate that understanding object structure is facilitated by multimodal exploration relative to unimodal exploration. We hypothesized from these results that active interactions allow the observer to control what parts of the object they saw and in what sequence. This control allows an observer to "hypothesis test" regarding object structure and may be guided by preferences for certain view-points that the observer's perceptual system has learned to be informative.

Understanding object structure requires that we have some notion that an object looks different from different viewing perspectives and that does not change its identity (as in Figure 2.2 left, upper-right quadrant presents the same object from different viewpoints, lower-right presents different objects from different view-

Figure 2.2 **Examples of novel object decision: same or different object?** (From James et al. [2001])

points). This form of object constancy is referred to as ***viewpoint-independent object recognition***. Many researchers and theorists agree that one way that we achieve viewpoint-independent recognition is through ***mental rotation***—the rotation of an incoming image to match a stored representation. A considerable amount of research has shown that the ability to mentally rotate an object image is related to manual rotation in important ways. Manual rotation is a specific form of active interaction that involves the volitional rotation of objects with the hands. That is, mental rotation is facilitated more by manual rotation practice compared to mental rotation practice, and further, that mental and manual rotations are reliant on common mechanisms [Wohlschläger and Wohlschläger 1998, Wexler 1997, Adams et al. 2011]. Thus, multimodal experience contributes to the development of mental rotation ability, a basic process in spatial thinking, which leads to an increase in the ability of an observer to understand its environment.

2.3.4 Symbol Understanding

Thus far, we have reviewed how physical interactions with the world affect our learning and understanding of existing objects. Symbols are objects that have several qualities that make learning and understanding different from 3D objects. When trying to understand symbols and symbolic relations, we must take something arbitrary and relate it to real world entities or relationships. Evidence suggests that symbol understanding is facilitated when symbols are actively produced. Producing symbols by hand, using multiple modalities during encoding (seeing them and producing them manually) facilitates learning symbol structure and meaning more than visual inspection alone. Lakoff and Nunez [2000], p. 49 argue that for symbols to be understood they must be associated with "something meaningful in human cognition that is ultimately grounded in experience and created via neural mechanisms." Here, grounding in experience is in reference to theories of ***grounded cognition*** and refers to the self-generation of meaningful actions, whose purpose is to control the state of the organism within its physical environment by providing it with interpretable sensory stimulation [Cisek 1999]. An important point, then, is that this grounding not only affects our learning and understanding of existing objects by providing expected action-sensory contingencies, it also guides the active production of objects.

One type of object production that has been shown to facilitate symbol learning, is writing by hand. Because the production of symbols has the defined purpose of communicating through the written word, humans have created writing tools that allow a greater degree of accuracy in the written form and that allow the transfer of permanent marks through ink or similar media. The manipulation of writing implements introduces yet another source of sensory information: haptic and kinesthetic cues that can augment and change motor production and visual perception. Nonetheless, visual-motor guidance of the writing device is still required, and this coupling of multimodal-multisensory information facilitates learning in much the same manner as other forms of active interactions. Handwriting symbols has been repeatedly correlated with increased competencies in symbol understanding (e.g., [Berninger et al. 1998, Richards et al. 2011, Li and James 2016). Although experimental studies are somewhat limited in adults, there is nonetheless a growing amount of research implicating the importance of self-generated action through writing on symbol understanding. In a series of pivotal and still highly influential studies, Freyd and colleagues showed that symbol recognition is significantly affected by our experiences with writing. For example, we are better able to identify static letters of the alphabet whose stroke directions and trajectories conform to how an individual creates those forms through handwriting, compared to similar letters that do not conform to the observer's own stroke directions and trajectories [Freyd 1983, Babcock and Freyd 1988, Orliaguet et al. 1997].

The interplay between the motor production of letters and the visual perception of letters has been demonstrated in individuals who have letter perception deficits. For example, individuals with *pure alexia* have difficulty recognizing letters. However, in some cases, their letter recognition can be facilitated by tracing the letter in the air with their finger or making hand movements that mimic the letter shape [Bartolomeo et al. 2002, Seki et al. 1995]. The interpretation of these case studies rests upon the fact that pure alexia results from damage to a specific brain region. The fact that damage to one location can result in a deficit in letter recognition that can be recovered to some extent by active interaction with the symbol's form suggests that both visual and motor brain systems are involved in the perception of a symbol. In these cases, the patients' actions facilitated their visual perceptions, evidence that the neural mechanisms subserving letter perception span both visual and motor brain regions as a result of prior multimodal experience.

Moreover, the production of symbols may rely upon the same neural mechanisms as the recognition of symbols. When adults are asked to simultaneously identify letters and shapes presented in varying degrees of *visual noise*, their thresholds for detection were increased (worsened) if they were writing a perceptually similar letter or shape during symbol identification, compared to when writing a perceptually dissimilar shape [James and Gauthier 2006]. The interpretation of this study rests upon theories of neural interference, which suggest that if one behavioral task is significantly affected by another concurrent task, then the two share common mechanisms. The fact that an interference effect was observed, rather than a facilitation effect, suggests that the production of a form and the perception of the form overlapped in terms of their underlying mechanisms. Again, we are provided with evidence that active interactions mold neural systems in such a way that the system will seek expected action-sensory contingencies, either by producing them through manual rotation or, in this case, producing them through the manipulation of writing implements.

2.4 Behavioral Results in Children

By exploring their surroundings, infants and children discover object properties and uncover possibilities for actions afforded by many objects. Importantly, they also learn about the functional changes that those novel possibilities for action imply for the action capabilities of the limbs [Lockman 2000]. The profound effects that multimodal-multisensory learning has on cognitive development was originally outlined in an extensive psychological theory proposed by Jean Piaget [1952, 1954]. Piaget believed that because information gained by watching others in early life was extremely limited, most of what children learn prior to two years of age is gained through multimodal experience. Importantly, Piaget argued that manual actions served to link information from multiple sensory systems together. For example, vision, somatosensation, and audition could all be linked through action on objects, because an action would result in the simultaneous production of visual, somatosensory, and auditory sensory input.

Early in life, each sensory modality is still relatively immature. However, some sensory systems are more mature than others. For example, the visual system of infants is very poor (e.g., Banks and Salapatek 1983), but their sense of touch is remarkably mature [Sann and Streri 2008]. This observation highlights one of the major benefits of active interactions in infancy: each action is accompanied by, at least, tactile stimulation. As infants reach and hold objects of interest, they often bring them into sight, and in doing so, produce simultaneous visual, motor, and tactile information. Importantly, the somatosensory modality provides the infant with rich information about the object (e.g., texture, shape, weight, temperature, size) that is simultaneously paired with relatively immature visual percepts (e.g., color, global shape). The tactile information is only gained through actions, and due to the somatosensory system's relative maturity, has the ability to aid in the development of the visual percept. Therefore, in infancy, although visual and motor systems are immature, actions still provide a wealth of information to the developing brain, because they are inherently multimodal.

In what follows, we will discuss a sample of empirical work that underlines the importance of multimodal-multisensory learning during development.

2.4.1 Surface Perception

Visual perception is affected by early locomotion abilities in very young children. Locomotion is an active interaction that allows children to explore their surroundings and that presents a variety of new multimodal experiences. For infants, the ability to move themselves represents unprecedented opportunity for self-generated actions on a variety of objects that were previously unreachable, but first, on surfaces. Visual competencies can develop from merely experiencing different surfaces. One well-known demonstration of visual development as a result of multimodal experience is the ***visual cliff*** paradigm [Gibson and Walk 1960]. These experiments require an apparatus constructed of two surface levels, the lower surface being a large drop from the higher surface. However, glass covers the lower surface at the same height as the higher surface such that one could locomote from the high surface over the low surface by moving across the glass (Figure 2.3). The point of the apparatus is to provide conflicting visual and somatosensory information to an infant. If one relies on vision, one will perceive a large drop off (danger). If one relies on somatosensation, the feel of the glass would reassure the actor that the surface was safe to cross. Most infants that can crawl will *not* cross the visual cliff to get to their caregiver—relying on visual cues rather than haptic ones

Figure 2.3 **Visual cliff apparatus.** From The Richard D. Walk papers, courtesy Drs. Nicholas and Dorothy Cummings Center for the History of Psychology, The University of Akron.

[Gibson and Walk 1960, Bertenthal et al. 1984]. These original studies documented this reliance on visual depth cues in numerous diurnal species [Gibson and Walk 1960].

However, visual development is prolonged compared to somatosensory development, which suggests that there should exist a time point in development when infants should not be so affected by visual cues. As a multimodal experience that binds visual and somatosensation, one would expect that experience with locomotion is important for infants to learn that "clear" (i.e., non-existent) surfaces do not afford locomotion. The early studies did not compare crawlers to non-crawlers to test experimentally whether experience with locomotion was *necessary* for this behavior to develop. Campos et al. [1992] showed, however, that if one lowers a non-crawling infant over the low side of the visual cliff, they will not demonstrate a fear response (increase in heart rate), but a crawler will exhibit considerable distress when put in the same situation [Campos et al. 1992]. In other words, non-crawling infants were not afraid of being atop a clear surface, but crawlers were. This suggests that non-crawling infants are not as affected by visual cues of surfaces as are crawling infants and that the ability of locomotion to bind visual and somatosensory cues has provided the infant with the knowledge that "clear" surfaces are not accompanied by the expected somatosensory cues and, therefore, do not afford locomotion.

In addition, the perception of depth is affected differentially by different modes of locomotion. For example, crawlers will crawl down slopes or across surfaces that are safe for crawling but that are not safe for walking, and early walkers will walk down slopes and across surfaces that are only safe for walking [Gibson and Walk 1960, Adolph et al. 2008, 2010, Adolph 1997]. Further, crawlers and walkers detect safe drop-offs vs. unsafe drop-offs that are specific to their particular mode of locomotion [Kretch and Adolph 2013]. A safe drop for a crawler is quite different from one for an early walker, as the risk in falling farther is greater. This is detected and used in this form of multimodal experience. Interestingly, infants use multimodal-multisensory information when approaching and understanding slope and drop-off: they reach over the edges and feel with their hands, before descending.

Why would locomotion have such a profound effect on depth perception? The assumption from these and other studies is that the motor experience of locomotion allows the individual to experience depth through their own visual-motor experience. The multimodal-multisensory information that results from locomotion (motor action) including proprioception, somatosensation, and vision, are all simultaneous, or ***time-locked***, allowing for integration of the multiple inputs (sensory and motor) into a representation that combines the inputs. New perception-action correspondences are dynamically coupled as a response to the environmental stimulation that accompanies new behaviors.

2.4.2 Three Dimensional Object Structure

Multimodal exploration of objects emerges in the first year of life. Very young infants will bring an object into their mouths for tactile exploration, whereas older infants (7 months) will begin to bring an object in front of their eyes for visual inspection [Rochet 1989, Ruff et al. 1992]. Early multimodal object exploration leads to significant gains in object knowledge. For example, infants who explore objects through self-generated actions are better able to apprehend two objects as separate entities compared to infants who do not explore, even when the objects explored and the objects seen are completely different [Needham 2000]. That is, the exploration served to help the infants learn about object properties that generalized beyond their direct experience.

Generally, infants are only able to learn about objects and their properties after they have the ability to reach and hold objects in the environment (multimodal input). However, recent work has allowed pre-grasping infants to be able to hold objects in the hopes of seeing increases in learning about object structure before intentional reaching and holding can occur. This body of work employs "sticky mittens"—mittens that are put onto infants that have Velcro on the palms such that objects (that also have Velcro on them) can be held by the infants at an earlier age than their manual dexterity would allow them to do so [Needham et al. 2002]. The infants are able to "grab" the objects through swiping at them by 3 months of age. With the sticky mittens experience, infants showed significantly more reaches toward objects than infants that wore non-sticky mittens [Libertus and Needham 2010]. But the more interesting aspect of these studies was how learning was changed when infants manually interacted with the objects.

Very brief experience with sticky mittens led to 3-month infants' understanding that the actions of others are goal directed, a behavior that is not typically observed until much later [Sommerville et al. 2005]. The implication is that even a very brief experience (2–3 min in this case) seeing one's own actions as goal directed led to understanding that same experience in others.

Other research with older children (18–36-month-olds) suggests that multimodal interactions with objects through visually guided action leads to enhanced understanding of object structure. When 18-month-olds are allowed to manually rotate a novel object during initial exposure they do so in a somewhat random manner (see Figure 2.4 top two graphs). But by 24 months, toddlers will manually rotate objects in the same way as older children and adults (see Figure 2.4 bottom two graphs) [Pereira et al. 2010]. Adult's rotations of similar three-dimensional objects are not statistically different from the 2½ year-olds depicted in Figure 2.4 (see Harman et al. 1999 for adult data). Furthermore, the rotations are not determined by haptic cues to object structure as 24-month-olds will rotate uniquely shaped objects in plexiglass cubes and spheres in the same way [James et al. 2014]. Showing oneself specific views of objects through multimodal interaction (in this case ***planar views***) was also correlated with object recognition scores in 24-month-old children: the more planar views were focused upon, the higher the recognition of objects [James et al. 2013].

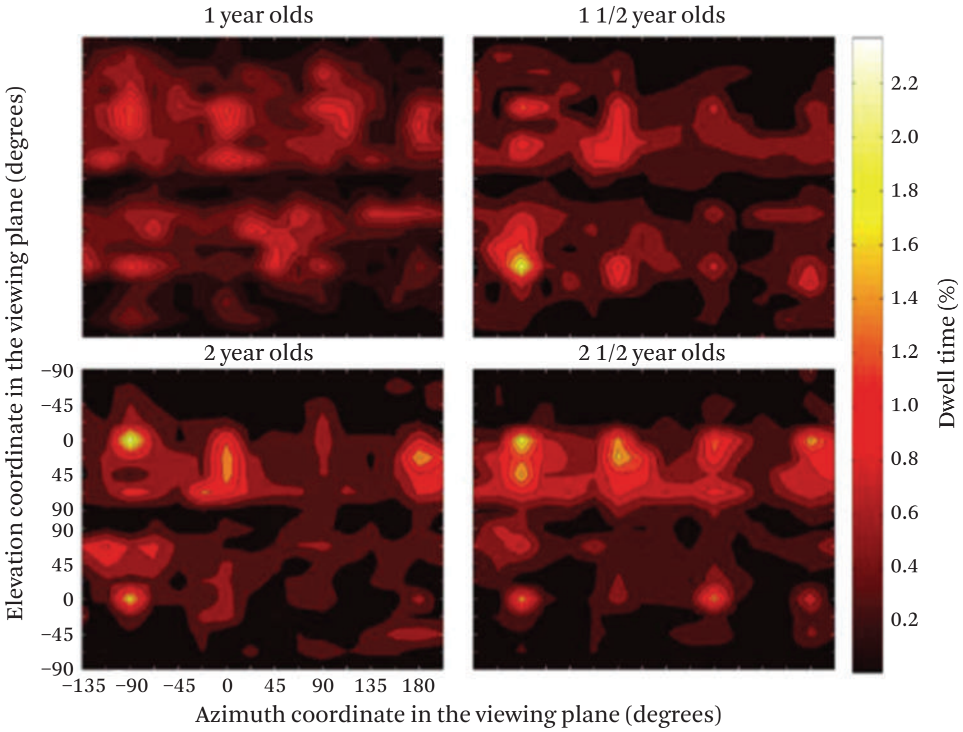

Figure 2.4 **Flattened viewing space of objects rotated manually at various ages.** The "hot spots" of increased dwell time in older children reflect planar views of objects. The focus on these planar views in adults manually exploring novel objects is well documented [Harman et al. 1999, James et al. 2001]. (From Pereira et al. [2010])

Another study investigated how moving an object along a trajectory influences the perception of elongation in object structure. When young children actively moved an object horizontally, they were better able to generalize the object's structure to new objects that were similarly elongated horizontally [Smith 2005]. However, when children simply watched an experimenter perform the same movement, there was no benefit in elongation perception.

The studies outlined above show that once children can manually interact with objects, their perception of the world changes significantly. They use their hands to manipulate objects in a way that impacts their learning and reflects their understanding of themselves and their environment.

2.4.3 Symbol Understanding

As discussed in Section 2.3.4, the extant literature has shown that handwriting symbols is especially effective for early symbol learning. Indeed, handwriting represents an important advance in motor skills and tool use during preschool years. Children progress from being barely able to hold a pencil to the production of mostly random scribbles before being able to produce specific, meaningful forms. In the following, we consider a special case of symbol learning: learning letters of the alphabet.

When children are first learning letters, they must map a novel, 2D shape onto the letter name and the letter sound. Eventually, they must put combinations of the symbols together to create words, which is another level of symbolic meaning. This is by no means a trivial task. The first step alone, learning to perceive a set of lines and curves as a meaningful unit, has a protracted development in the child and requires explicit teaching. One of the problems is that letters do not conform to what the child has already learned about 3D objects. Specifically, if a letter is rotated from upright, its identity can change. For instance, rotating a “p” 180° results in a different letter identity, a “d”. Change of identity after a 180°degree rotation does not occur with other types of objects: an upright cup and an upside-down cup are both cups. This quality of symbols alone makes things difficult for the early learner and manifests in the numerous reversal errors children make when perceiving and producing letters. Like other objects, however, the symbols must be distinguished from one another by detecting similarities and dissimilarities. In terms of letters, the similarities and differences may be very slight changes in the visual input, for example, the difference between an uppercase C and an uppercase G. Things get even more difficult when one introduces the many examples of letters that are handwritten that one must decipher. However, the variability present in handwritten letters may be important in understanding why the visual recognition of letters is facilitated by handwriting experience.

The ability to recognize written symbols, such as letters, is made easier by producing them by hand [Molfese et al. 2011, Longcamp et al. 2005, Hall et al. 2015]. For instance, the National Research Council and the National Early Literacy Panel both found that letter writing in preschool had a significant impact on future literacy skills [Snow et al. 1998]. Why handwriting facilitates letter recognition above and beyond other types of practice can be understood from the multimodal-multisensory learning perspective. Although it is generally accepted that the multisensory learning of letters (e.g., hearing and seeing with no motor action) facilities letter learning beyond unisensory learning, incorporating multimodal production of letters contributes even more to the learning experience. The act of producing a letterform by hand is a complicated task, requiring efficient coordination between multiple systems. We have hypothesized that handwriting accomplishes this through the production of variable forms. Each letter production is accompanied by a unique combination of visual and somatosensory stimulation. Manual dexterity in children is somewhat poor, resulting in a variety of possible visual and tactile combinations every time a child attempts to write a letter. The variability of this experience is exacerbated by the use of tools (writing implements), which requires fine motor skill, an ability that matures at a slower rate than gross motor skill. The perceptual result is the production of letterforms that are often quite variable and "messy" (see Figure 2.5). We have recently found that children who produce variable forms while handwriting or through tracing handwritten symbols are better able to recognize a novel set of symbols than their peers who trace the same typed symbols [Li and James 2016]. It is well known that learning a category through variable exemplars facilitates learning of that category compared to studying more similar exemplars [Namy and Gentner 2002]. The more variability that is perceived and integrated into a named category (such as the letter "A"), the more novel instances that can then be matched to this information. Put simply, once children understand the many instances of the letter "p" they can begin to recognize new, unique instances of that letter. Thus, the multimodal production of a letterform has the benefit of creating perceptually variable instances that facilitate category learning, that is in addition to the development of a visual and somatosensory history with that category.

2.4.4 Neural Changes as a Result of Active Experience

The fact that the brain produces all behavior would not receive the attention it deserves, if we were to try to understand human behavior without some understanding of the neural circuitry that underlies the behavior in question. The claim here is that to truly understand behavior, we must also understand how the brain produces that behavior. Because our understanding of the brain is in its infancy, however, this is difficult and controversial. Nonetheless, we can still interpret and predict behavior based on what we know so far regarding neural functioning from comparative work with other species and human neuroimaging studies. This is especially true when we consider the effects that multimodal-multisensory experiences have on learning. As outlined in the introduction, learning in the brain occurs through association of inputs. We argue here that human action serves to combine multisensory inputs, and as such, is a crucial component of learning. This claim is based on the assumption that there are bidirectional, reciprocal relations between

Figure 2.5 **Examples of handwritten letters by 4- year-old children.** Top row are traces, bottom two rows are handwritten free-hand.

perception and action (e.g., [Dewey 1896, Gibson 1979]). From this perspective, action and perception are intimately linked: the ultimate purpose of perception is to guide action (see, e.g., [Craighero et al. 1996]), and actions (e.g., movements of the eyes, heads, and hands) are necessary in order to perceive (e.g., [Campos et al. 2000, O'Regan and Noë 2001]). When humans perceive objects, they automatically generate actions appropriate for manipulating or interacting with those objects if they have had active interactions with them previously [Ellis and Tucker 2000, Tucker and Ellis 1998]. Therefore, we know that actions and perceptions form associated networks in the brain under some circumstances. Thus, perception and action become linked through our multimodal experiences.

In what follows, we provide some evidence of this linking in the brain and the experiences that are required to form these multimodal-multisensory networks. We will focus on ***functional magnetic resonance imaging (fMRI)*** as a method of human neuroimaging, given its high ***spatial resolution*** of patterns of neural activation, safety, widespread use in human research, and applicability to research involving the neural pathways created through learning. We will focus on a multimodal network that includes: (1) the ***fusiform gyrus***, a structure in the ventral temporal-occipital cortex that has long been known to underlie visual object processing and that becomes tuned, with experience, for processing faces in the right hemisphere (e.g., [Kanwisher et al. 1997]) and letters/words in the left hemisphere (e.g., [Cohen and Dehaene 2004]); (2) the dorsal precentral gyrus, which is in the top portion of the primary motor cortex, a region that has long been known to produce actions [Penfield and Boldfrey 1937]; (3) the middle frontal gyrus, a region in the premotor cortex, involved with motor programming and traditionally thought to underlie fine-motor skills (e.g., [Exner 1881, Roux et al. 2009]); and (4) the ventral primary motor/premotor cortex, that overlaps with Broca's area, a region thought to underlie speech production (e.g., [Broca 1861]]) and that has, more recently, become associated with numerous types of fine motor production skills (for review see [Petrides 2015]). As outlined below, this network only becomes linked after individuals experience the world through multimodal-multisensory learning.

2.5 Neuroimaging Studies in Adults

Functional neuroimaging methods provide us with unique information about neural patterns that occur with an overt behavioral response. As such, the resultant data is correlational, but highly suggestive of the neural patterns that underlie human behavior.

In additional to providing the neural correlate of overt behavior, neuroimaging studies can also generate hypotheses and theories about human cognition. In what follows, we briefly outline the use of this method in the service of understanding the experiences that are required to link sensory and motor systems.

2.5.1 The Effects of Action on Sensory Processing of Objects

According to embodied cognition models, a distributed representation of an object concept is created by brain-body-environment interactions (see [Barsalou et al. 2003]). Perception of a stimulus via a single sensory modality (e.g., vision) can therefore engage the entire distributed representation. In neuroimaging work, this is demonstrated when motor systems in the brain are activated when participants simply look at objects that they are accustomed to manipulating (e.g., [Greezes and Decety 2002]), even without acting upon the objects at that time. Motor system activation is more pronounced when participants need to make judgments about the manipulability of objects rather than about the function of objects [Boronat et al. 2005, Buxbaum and Saffran 2002, Simmons and Barsalou 2003]. The motor system is also recruited in simple visual perception tasks. For example, in a recent study, we asked adult participants to study actual novel objects that were constructed to produce a sound upon specific types of interactions (e.g., pressing a top made a novel, rattling sound) (see Figure 2.6). Participants learned these novel sound-action-object associations in one of two ways: either by actively producing the sound themselves (active interaction) or through watching an experimenter produce the

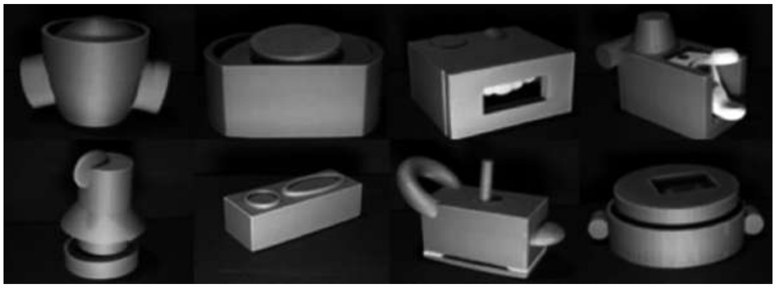

Figure 2.6 **Examples of novel, sound producing objects.** (From Butler and James [2013])

sounds (passive interaction). Note that in both cases, there was equal exposure to the visual and auditory information. The only difference was that in one case they produced the sound themselves instead of watching another produce the sound.

After participants studied these objects and learned the pairings, they underwent fMRI scanning while they were shown static photos of the objects with sound and without sound, as well as hearing the sounds alone. We then probed brain responses though ***blood-oxygen-level-dependent (BOLD)*** activation in regions of the multimodal network (see Section 2.5 and Figure 2.7). We observed that there was significantly greater activation after the active learning experience compared to the passive learning experience [Butler and James 2013]. Only active learning served to recruit the extended multimodal network including frontal (motor) and sensory brain regions. Thus, simply perceiving the objects or hearing the sound that was learned automatically recruited the sensory and motor brain regions used during the learning episode. Furthermore, we were interested in whether or not the visual regions were functionally connected to the motor regions. Assessing ***functional connectivity*** allows one to investigate whether the active regions are recruited together because of the task, or whether the recruitment is due to other factors (such as increased physiological responses). Indeed, only after active learning were visual seed regions (in blue in Figure 2.7) functionally connected to motor regions (in orange Figure 2.7).

2.5.2 Neural Systems Supporting Symbol Processing in Adults

Symbols represent an interesting category with which to study questions of interactions among sensory and motor brain systems. Just as multisensory experiences help shape the functioning of the subcortical region of the superior colliculus, active interactions and their inherent multisensory nature have been shown to shape

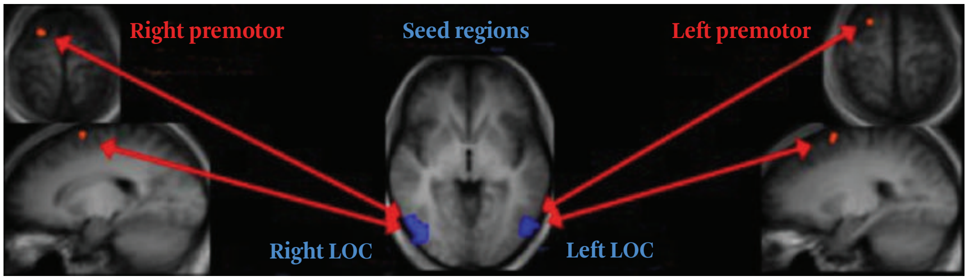

Figure 2.7 **Functional connectivity between the visual *Lateral Occipital Complex (LOC)* regions and motor regions in the brain after active learning.** Note that left side of the brain in figure is right hemisphere due to radiological coordinate system. (From Butler and James [2013])

the functioning of cortical brain regions. During handwriting, the motor experience is spatiotemporally coordinated with the visual and somatosensory input. Although we focus largely on the visual modality, it is important to note that, to the brain, an action such as handwriting is marked by a particular set of multisensory inputs along with motor to sensory mappings, or multimodal mappings. For example, the creation of a capital “A” results in a slightly different multimodal pairing than the creation of a capital “G”. Similarly, writing a capital “A” requires a different motor command than writing a capital “G” and each action looks and feels different. If we were to extrapolate single neuron data and apply it to human systems neuroscience, we could speculate that the visual stimulation of a capital “A” invokes subcortical multisensory neurons tuned through multisensory experiences (or, in the case of visual only learning, unimodal neurons) that pass information to cortical regions associated with the multimodal experience of handwriting.

Indeed, the perception of individual letters has been shown to be supported by a neural system that encompasses both sensory and motor brain regions, often including ventral-temporal, frontal, and parietal cortices [Longcamp et al. 2003, 2008, 2011, 2014, James and Gauthier 2006, James and Atwood 2009]. We investigated the neural overlap between writing letters and perceiving letters and found that even when participants did not look at the letters they were writing, significant overlap in brain systems emerged for writing and perception tasks with letters (see Figure 2.8). This network included not only the usual visual regions observed when one perceives letters, but also an extended frontal network that included dorsal and ventral regions that are known to code actions and motor programs. As depicted in Figure 2.8, the putative visual region was active during letter writing, and the

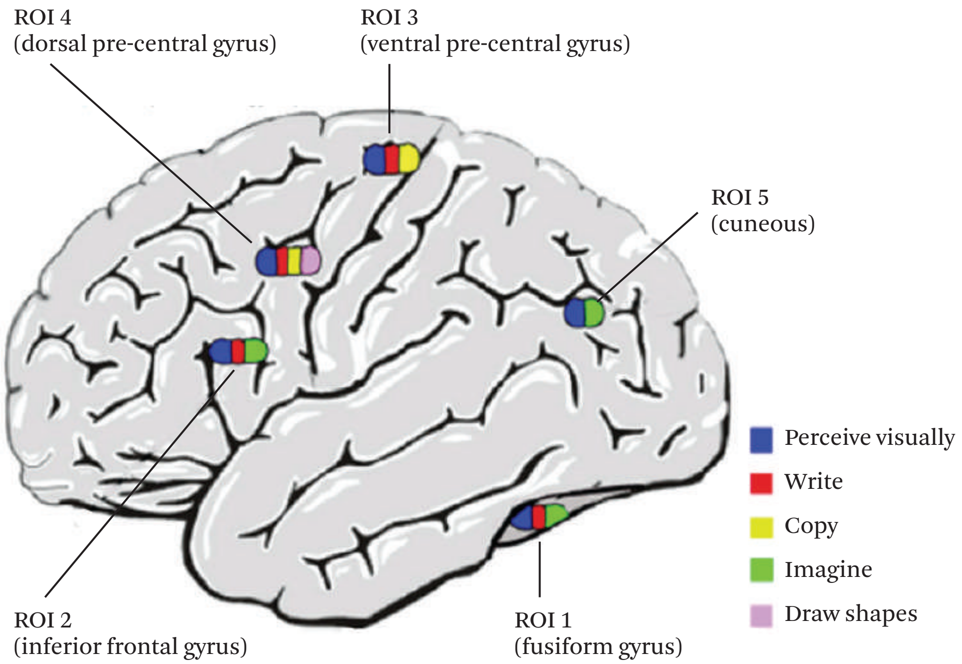

Figure 2.8 **A schematic of results from James and Gauthier [2006], showing the overlap in brain activation patterns as a result of handwriting, perceiving and imagining letters.** (From James and Gauthier [2006])

traditional motor areas were active during letter perception. These results suggest that action and perception, even in the case of symbol processing, automatically recruit an integrated multimodal network.

This overlap in activation led us to the next obvious question: What are the experiences that create such a system? It is possible that any motor act with symbols would recruit such a system, but it could also be that the creation of the symbols by hand, feature-by-feature, may serve to pair the visual input and motor output during this behavior. To test this idea, we had adult participants learn a novel symbol system, ***pseudoletters***, through two types of multimodal-multisensory learning (writing + vision and typing + vision) and one type of multisensory learning (vision + audition) [James and Atwood 2009]. Results demonstrated that only after learning the pseudoletters through handwriting the left fusiform gyrus active for these novel forms (Figure 2.9). This finding was the first to show that this "visual" region was affected by how a stimulus was learned. More specifically, it was not responding only to the visual presentation of letters, but was responding to the visual presentation

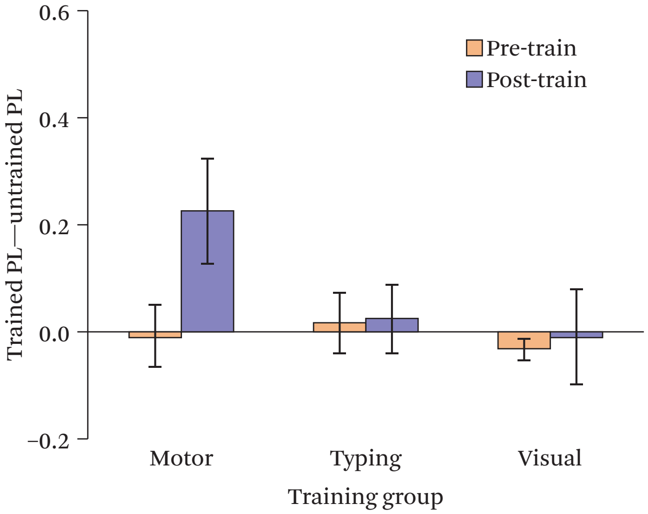

Figure 2.9 **The difference between trained and untrained pseudoletters in the left fusiform as a function of training experience.** Note that the left fusiform gyrus is a visual processing region, in which only handwriting (labeled as motor in this figure) experience resulted in greater activation after learning. (From James and Atwood [2009])

of a letter with which the observer had a specific type of motor experience: handwriting experience. Furthermore, the dorsal precentral gyrus seen in the above study for letter perception and writing was also active only after training that involved multimodal learning (writing + vision and typing + vision) but not through training that involved only multisensory (vision + audition) learning. Thus, the network of activation seen for symbol processing is formed by multimodal experience of handwriting.

2.6 Neuroimaging Studies in Developing Populations

Methodological limitations precludes the use of fMRI in infants and toddlers that are fully awake, because it requires that individuals stay still for a minimum of 5 min at a time, usually for 30 min in total. Nonetheless, because we are very interested in how multimodal-multisensory learning serves to create neural systems, we routinely scan 4–6-year-old children in the hopes that we can observe the development of these neural systems. In the summary that follows, we outline studies that have investigated how multimodal networks emerge. In other words, what experiences are required for this functional network to be automatically activated?

2.6.1 The Multimodal Network Underlying Object and Verb Processing

There is now ample evidence that reading verbs recruits the motor systems that are used for performing the action that a verb describes [Hauk et al. 2003, Pulvermüller 2005, 2012, 2013]. These findings suggest that the perception of verbs re-activates motor systems that were used when the word was encountered during action, and, by extension, also suggest that only actions that we have performed will recruit the motor system during perception. However, not all verbs describe actions with which we have personal experience (e.g., skydiving), which begs the question: Do our action systems then become linked with perception if we simply watch an action? The research in this realm is controversial. Some studies have shown that indeed, action systems are recruited during action observation (e.g., [Gallese et al. 1996, Rizzolatti and Craighero 2004]), while others have shown that we have to have had personal experience performing the action for these systems to be automatically recruited [Lee et al. 2001].

We wished to address this question with children, given their relatively limited personal experience with actions and their associated verbs. We asked whether multimodal recruitment would occur after children watched another person perform an action during verb learning, or if they had to perform the action themselves in order for these multimodal systems to be recruited. To test this idea, we asked 5-year-old children to learn new verbs (such as “yocking”) that were associated with an action performed on a novel object. Children learned the action/verb/object associations either by watching an experimenter perform the action or through performing the action themselves. We then measured their BOLD activation patterns when they saw the novel objects or heard the novel words in a subsequent fMRI scanning session [James and Swain 2011]. The results were clear: only when children acted on the objects themselves while hearing/seeing the novel stimuli did a multimodal network emerge (see Figure 2.10).

A follow-up study directly tested the effects of multimodal-multisensory learning by having 5-year-old children learn the noises that objects made as they interacted with them, or watched an experimenter [James and Bose 2011]. The procedure was the same as above except that instead of learning a verb, the participants heard a sound that was associated with object manipulation. Again, only after self-produced action during learning was the extended multimodal network of activation recruited in the brain.

The action that is coded during learning is, therefore, part of the neural representation of the object or word. In essence, the multimodal act of object manipulation serves to link sensory and motor systems in the brain.

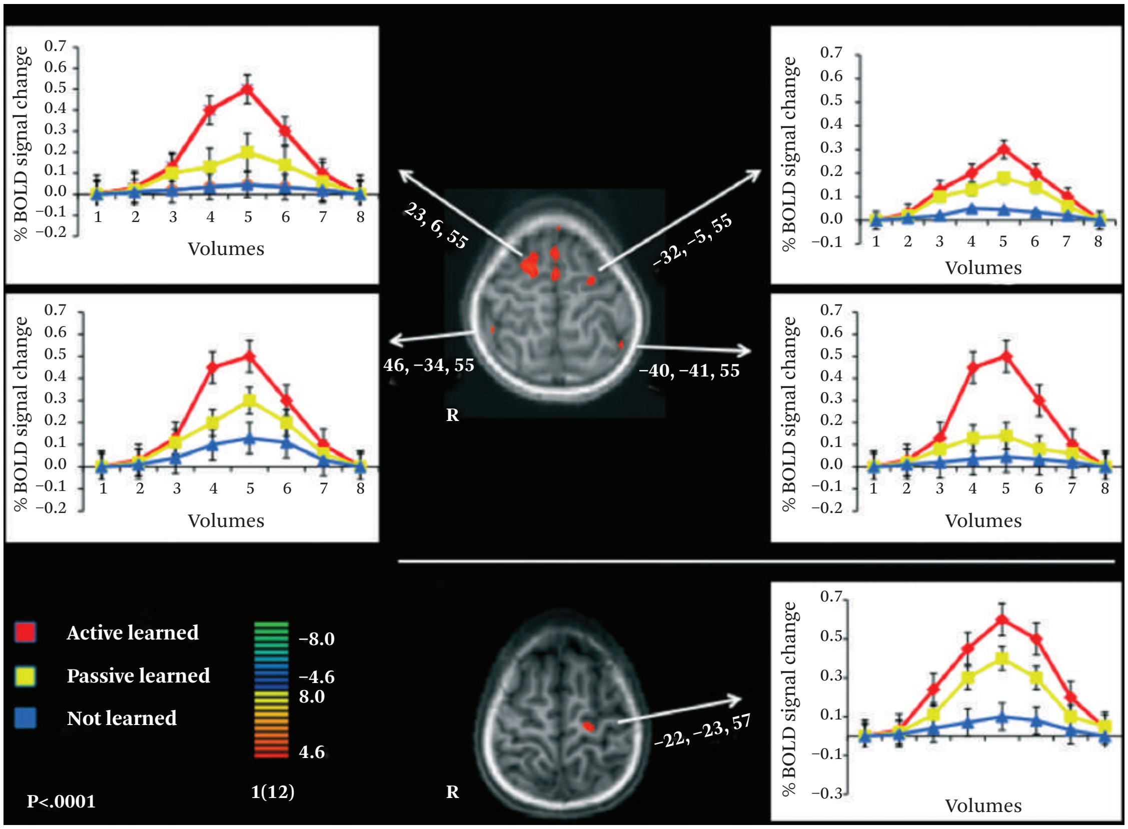

Figure 2.10 **fMRI results after children learn verbs through their own actions or by passively watching the experimenter compared with unlearned verbs.** The upper panel shows the response to hearing the new verbs, while the lower panel depicts activation when children saw the novel objects. Only learning through action resulted in high motor system activity. Top left and right graphs: Middle frontal gyri; middle graphs: Inferior parietal sulci; bottom graph: left primary motor cortex. Note: left is right hemisphere. (From James and Swain [2011])

2.6.2 Neural Systems Supporting Symbol Processing in Children

The developmental trajectory of the neural system supporting letter perception clearly displays the importance of multimodal-multisensory experiences through the act of handwriting on brain development. In preliterate children, the perception of individual letters recruits bilateral fusiform gyri in visual association areas, bilateral intra-parietal sulcus in the visual-motor association brain regions, and left

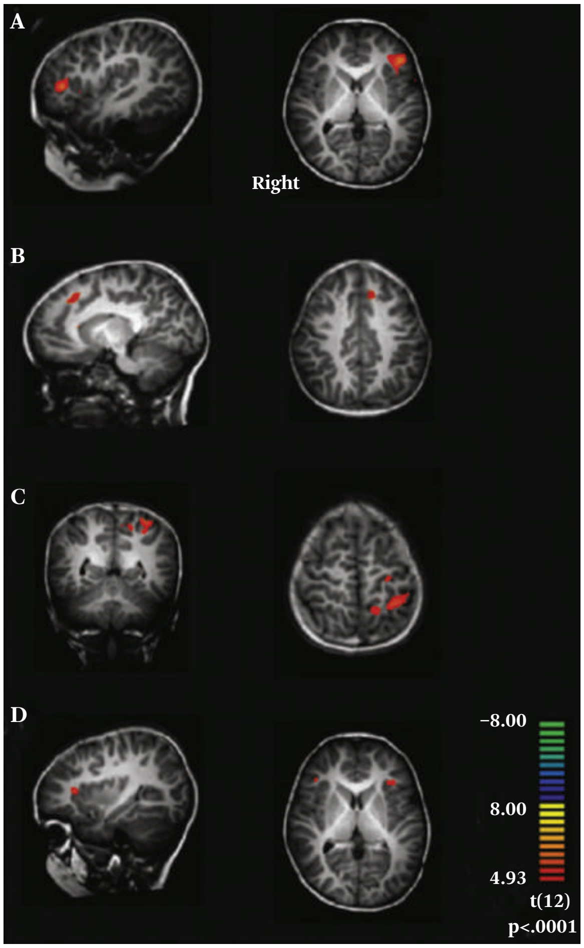

Figure 2.11 **(A) and (B) the difference in BOLD signal between handwriting > typing in the frontal premotor cortices; (C) difference between handwriting > tracing in precentral gyrus and parietal cortex; and (D) traced > typed activation in frontal cortex.** (From James and Engelhardt [2012])

dorsal precentral gyrus in motor cortex, but *only* for letters with which the child has had handwriting experience (Figure 2.11) [James and Engelhardt 2012]. Indeed, the intraparietal sulcus responds more strongly for form-feature associations with letterforms (i.e., stronger for letters learned through handwriting than shapes learned through handwriting) whereas the left dorsal precentral gyrus responds more strongly for motor associations (i.e., stronger for letters learned through handwriting than tracing or typing). The functional connections between these three regions and the left fusiform gyrus show a similar pattern in preliterate children after handwriting practice [James and Engelhardt 2012, Vinci-Booher et al. 2016].

The action of handwriting a letter feature-by-feature allows for multimodal input of information that can also be encoded though a single sense. It therefore transforms an inherently unisensory behavior (visual letter recognition before handwriting experience) into a multimodal behavior (visual letter recognition after handwriting experience). Multimodal behaviors promote the emergence of multisensory integration in evolutionarily early subcortical brain regions. They effectively structure the input to cortical regions and engender multimodal integration in the cortex, as outlined previously. The ability to transform inherently unimodal behaviors, such as visual perception of written forms or stationary objects, into inherently multimodal behaviors is the major utility provided by interacting with tools, such as writing implements.

Interestingly, the pattern of activation that is seen after children learn to write letters by hand is only observed if the writing is self-produced. That is, if a child watches an experimenter produce the same forms, the multimodal network will not be evident during subsequent letter perception [Kersey and James]. This latter result suggests that it is the multimodal *production*, not the multisensory perception that results in the emergence of the distributed brain network observed during letter perception.

The extant literature on the neural substrates underlying multimodal-multisensory learning in young children clearly demonstrate that the visual perception of objects learned actively is not purely visual. Learning through action creates multimodal brain networks that reflect the multimodal associations learned through active interactions.

This brief review of empirical studies suggests that the brain is highly adaptive to modes of learning. The input from the environment, through the body, is processed by the brain in a manner that requires high plasticity both in childhood and adulthood. The way in which we learn changes brain systems that, in turn, change our behaviors. As such, we argue that human behavior cannot be fully understood without a consideration of environmental input, bodily constraints, and brain functioning. Valuable insights can be gained by a thoughtful consideration of the rich data sets produced by brain imaging techniques. Understanding brain mechanisms that underlie behavior change our understanding of the *hows and whys* of the efficacy of technologies to assist learning.

2.7 Theoretical Implications–Embodied Cognition

The embodied cognition perspective encompasses a diverse set of theories that are based on the idea that human cognitive and linguistic processes are rooted in perceptual and physical interactions of the human body with the world [Barsalou 2008, Wilson 2002]. According to this perspective, cognitive structures and processes—including ways of thinking, representations of knowledge, and methods of organizing and expressing information—are influenced and constrained by the specifics of human perceptual systems and human bodies. Put simply, cognition is shaped through actions by the possibilities and limitations afforded by the human body.

The research outlined in this chapter clearly supports this embodied perspective. Learning is facilitated though bodily interactions with the environment. Such necessary competencies, such as visual perception, object knowledge, and symbol understanding, are determined by physical action. Often we consider these basic human abilities to be reliant on unimodal (usually visual) processing. However, more research is emerging that supports the notion that multimodal processing is key to acquiring these abilities, and further, that multisensory processing is created through action, an inherently multimodal behavior. Furthermore, the mechanisms that support multimodal-multisensory learning are becoming better understood. Multimodal-multisensory learning results in the recruitment of widespread neural networks that serve to link information that creates highly adaptive systems for supporting human behavior. Because of the importance of action for learning, and because action is reliant on the body, the work reviewed here outlines the importance of the embodied cognition standpoint for understanding human behavior.

2.8 Implications for Multimodal-Multisensor Interface Design

When we think about modern society and its reliance on human-computer interactions, it is wise to remember that our brains have adapted to an environment that existed hundreds of thousands of years prior to our use of screens and multimodal interfaces. Track balls, keyboards, and pens are not a part of the environment for which our brains have adapted. However, our early hominid ancestors did use tools (e.g., [Ambrose 2001]). Considering that many of our modern interfaces are tools helps to understand how humans have become so adept at using computer interfaces. Our brains are highly plastic in terms of integrating these bodily extensions into a pre-existing system, although they evolved to capitalize on relatively primitive forms of these instruments.

Glossary

Active behaviors: Overt movements performed by the observer. In this chapter, we restrict this definition to include intentional, or goal-directed actions (as opposed to reflexive actions).

Amodal stimulus properties: Properties of an object that can be conveyed by more than one sensory system [Lickliter and Bahrick 2004]. By some definitions, amodal also refers to information that is no longer connected to the sensory modality by which it was encoded.

Binding problem: Theoretical problem stemming from the required ability of the brain to bind sensory signals that trigger neuronal firing to the environmental stimulus from which those signals originate. This also entails the problem of how the brain combines signals from multiple senses into a unified percept, given that they detect different features of the stimuli.

Blood-oxygen-level-dependent (BOLD): The primary measure of interest in fMRI. BOLD computes the ratio of deoxygenated to oxygenated hemoglobin in the brain. The more oxygen an area is consuming, the more active its neurons.

Constructivism: A broad theoretical approach that considers the organism to be constructing meaning based on interactions with objects and people in the world.

Convergence: The unique property of neural connections by which more than one type of unisensory neuron connects to the same neuron and may independently evoke neural activity from that neuron. This can be contrasted with integration, whereby the sensory signals are combined to produce a single response based on more than one input. Convergence results in multiple separable signals in a given neuron, whereas integration refers to combining multiple inputs into a single response (see Figure 2.12).

Experience-dependent: Changes in the brain that are caused by experience, which may involve active or passive interaction with the environment.

Feedforward-feedback loop: Information travels in at least a bi-directional manner in the cortex and subcortex. Information affects, and is affected by, information included in these loops.

Functional connectivity: Statistical dependencies in time between spatially distinct neural regions. Used on a BOLD signal from fMRI, this method can result in revealing brain regions whose activities are correlated in time, leading to the inference that they are working together for that given behavior.

Functional magnetic resonance imaging (fMRI): A neuroimaging method broadly used in medical and psychological sciences to non-invasively observe extremely small changes in tissue properties, particularly changes in ferromagnetism associated with changes in the ratio of deoxygenated to oxygenated hemoglobin, that can be used to make inferences concerning the functioning of cortical structures.

Glossary *(continued)*

Fusiform gyrus: A gyrus in the brain that lies on the ventral aspect of the temporal lobe, spans from the occipital lobe to midway along the temporal lobe and, is located approximately along the midline of the temporal lobe. In functional terms, it is associated with object perception and recognition.

Grounded cognition: A central concept embedded in a range of cognitive theories, including simulation theory, situated cognition, and embodied cognition that opposes the existence of abstract knowledge systems. Instead, they consider knowledge systems to be represented implicitly by sensory and motor mechanisms during interaction with the world.

Hebbian learning: A well-documented neural mechanism of learning in which axonal connections between neurons undergo activity-dependent changes. There are two basic tenants: (1) when two neurons repeatedly fire in a coordinated manner, the connections between them are strengthened, effectively increasing the likelihood of firing together in the future; and (2) when two neurons repeatedly fire in an uncoordinated manner, the connections between them weaken, effectively reducing the likelihood of firing together in the future.

Integration: The unique property of some multisensory neurons by which they combine converging input signals from more than one sensory modality and output an *integrated* signal. Input may be from primary sensory organs, secondary regions, or association areas. The integrated output combines the signals into a single output. See also definition for convergence for a comparison (see Figure 2.12).

Intersensory redundancy: Recent theory of perceptual development that evades the binding problem by portraying the human perceptual system as a multimodal system. It does so by responding to unisensory and multisensory inputs with differential weightings: greater weight on multisensory input and less weight on unisensory input.

Lateraloccipital complex (LOC): A broad functional zone encompassing the anatomical regions of lateral occipital cortex and the ventral and lateral surfaces of temporal cortex. It is traditionally defined as cortex that is more active when presented with pictures of intact objects than pictures of textures or of scrambled objects.

Mental rotation: A form of mental imagery in which an observer rotates imaginary visual representations of two- or three-dimensional objects in their mind.

Modality-specific stimulus properties: Properties of objects that may only be detected by one sensory system. An example is wavelengths of light, which are only detectable by the visual system.

Multimodal (cognitive neuroscience meaning): Combined sensory and motor system input to the brain, for example, the interaction of vision and motor systems during encoding of an event. On the other hand, passive viewing of visually presented stimuli is not generally considered multimodal, because it can be accomplished without motor input or input from other sensory systems.

Glossary *(continued)*

Multimodal-multisensory: See *multimodal* and *multisensory* definitions.

Multisensory: The involvement of at least two sensory modalities in a given process. For example, vision and haptics are combined and provide multisensory input to the brain when we see and touch objects. Multisensory processes are most often multimodal. However, multisensory processes can occur without motor input. For example, temporally coincident visual and auditory cues need not be accompanied by head or eye movements.

Multisensory depression: Information about objects and events in the environment gleaned by relying upon spatiotemporal disparities across more than one sense (e.g., the visual experience of color is not always coincident with tactile cues), which leads to a depressed (decreased) neural response. That is, the neural response to the same multisensory input in the future will be reduced. This process effectively increases the response to modality-specific stimulus properties, such as color, because it decreases the physiological response to multisensory signal combinations, such as simultaneous visual-auditory stimuli.

Multisensory enhancement: Information about objects and events in the environment gleaned by relying upon spatiotemporal coincidence of more than one sense (e.g., vision and haptics that coincide during visually guided touch), which leads to an enhanced (increased) neural response. That is, the neural response to the same multisensory input in the future will be increased. This process effectively increases the response to amodal stimulus properties, such as visual-haptic cues, because it increases the physiological response to amodal stimulus properties. Multisensory enhancement is distinct from Super-additivity (see Chapter 1 of this volume), which occurs when a multisensory response is greater than the sum of the unimodal responses. Multisensory enhancement provides a developmental mechanism for multisensory perception whereas supper-additivity describes the neural response of a population of neurons responding to multisensory stimuli.

Multisensory neurons: A neuron that responds to a stimulus from more than one sensory modality.

Neural Adaptation: A change over time of the responses of neurons as a result of ongoing stimulation. As a stimulus is repeated, neurons sensitive to that stimulus will decrease their firing rates.

Neural plasticity: The ability of neuronal structures, connections, and processes in the brain to undergo experience-dependent changes.

Passive behaviors: The perceiver receives sensory input, although no action has been performed.

Planar views: View of an object, often corresponding to a non-canonical view, “in which the major axis of the object is perpendicular or parallel to the line of sight” [James et al. 2013]. Planar views are often referred to as “side views” or “front views” of 3D objects (see Figure 2.13).

Glossary *(continued)*

Pseudoletters: Written forms that are designed to match features of a particular written script (e.g., Roman letter). Such constructions often are used to control for certain dimensions of the written form of interest, such as stroke thickness, orientation, or size (see Figure 2.14).

Pure alexia: A selective deficit in letter perception associated with lesions in left ventral-temporal cortex, that results in poor reading ability.

Receptive field: The spatial or temporal range of sensory stimulation that affects the activity pattern of a neuron, or group of neurons. For example, neurons in the visual system respond to specific areas in space that are coded on the retina. Neurons in the auditory system respond to a specific range of sound frequencies.

Size constancy: A form of perceptual constancy that refers to the tendency of an observer to infer a constant object size from different viewing distances, despite the fact that the different viewing distances project different image sizes onto the retina (see Figure 2.15).

Spatial resolution: The number of units (e.g., pixels) per squared area used to construct an image, which is related to the amount of detail that an image can convey. Higher resolution images have more units per squared area than lower resolution images.

Stabilized image phenomenon: Phenomenon whereby a percept provided by a visual stimulus will fade away if the visual stimulus is continually projected onto the same retinal cells. This phenomenon only occurs when the eyes are artificially stabilized to prevent movement.

Superior colliculus: A layered subcortical structure containing multiple classes of neurons that interconnect with an array of sensory and motor neurons, including early sensory receptors and later cortical areas.

Time-locked: Signal processing concept that describes the response behavior of signals to a particular stimulus onset as being inextricably bound in time. For example, if a ganglion cell in the retina is alternately exposed to light and dark, the cell will respond during light and rest during dark. That is, the response of the ganglion cell is time-locked with the onset of changes in the light stimulus.

Unisensory neuron: A neuron that receives input from a single sense. For example, neurons in the primary visual cortex receive input from other neurons that transduce light waves into neural signal, but not from neurons that transduce sound waves.

Viewpoint-independent object recognition: The ability of an observer to know that an object remains the same from different viewpoints, despite potential variations in shape contour, size, or other features projected onto the retina.

Visual cliff: A psychological assessment tool, often used as a measure of depth perception development, in which an infant is placed atop an apparatus and allowed to explore the visually displayed depth difference by crawling (Figure 2.3).

Visual noise: Any image manipulation that degrades, distorts, or hides aspects of an underlying visual image. One common perturbation is applying Gaussian noise to an image (see Figure 2.16).

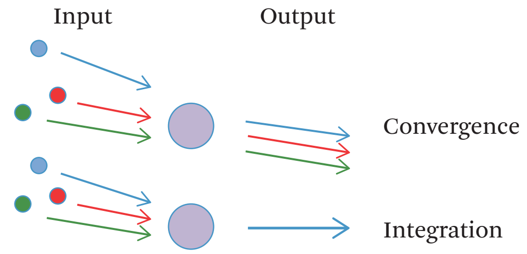

Figure 2.12 Graphical depiction of convergence and integration of neural signals.

Figure 2.13 In this image, the left crib is a planar, side view, and the right crib is a non-planar view. (From Pereira et al. [2010])

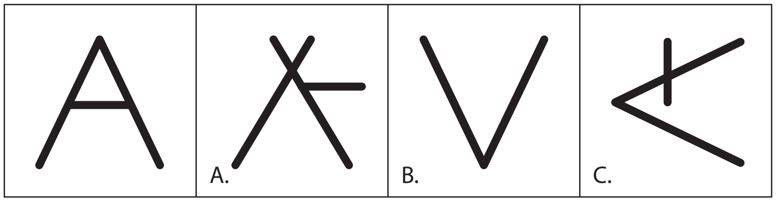

Figure 2.14 The (C) pseudoletter would control for stroke features, thickness, orientation, and size of the Roman letter A, whereas (B) controls for thickness and size, and (A) controls for features, thickness, and size.

Figure 2.15 **Classic demonstration of size constancy.** Both persons appear to be nearly the same size, although if the image of the person further away is moved closer it becomes obvious that the image of the second person is substantially smaller. (From Boring [1964])

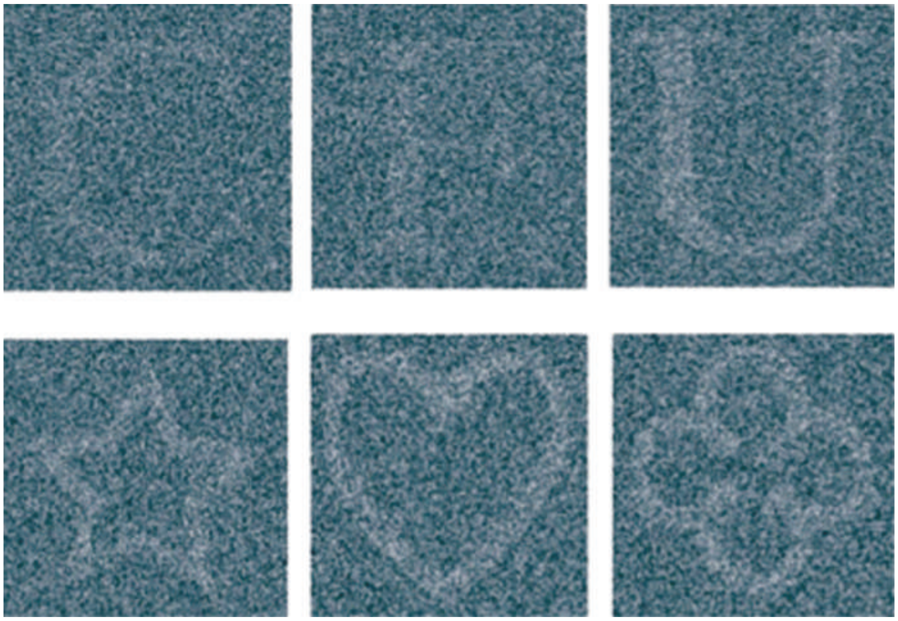

Figure 2.16 **Example of stimuli with different levels of visual noise added.** The top row has more noise added, and is slightly above visual detection threshold. The bottom row has less noise added, and is well above visual detection threshold. (From James and Gauthier [2009])

How can we capitalize on our knowledge of embodied learning to better design human-computer interfaces? First, we must remember that self-generated action is important for learning. Therefore, interfaces that capitalize on self-produced actions, such as touchscreens and styluses, will facilitate learning, because they allow the visual and motor systems to interact and form important links for learning (see Chapter 3 this volume). Further, the types of actions may also be important. For example, recall that multimodal neural systems are not created through typing to the same extent as writing. We would be well advised therefore, to continue the use of tools that require construction of the 2D form rather than actions that are farther removed from the produced symbol, such as typing. In addition, multisensory coding of information aids learning. Although we discussed one way in which this occurs (through action), one can imagine that adding multisensory information in general through interface design may be beneficial for student learning, potentially through haptic interfaces that make use of somatosensation (see Chapter 3 this volume). Immersive environments in which learners can self-locomote may also benefit learning, especially when considering understanding of spatial relations.

In short, self-guided actions facilitate learning. To increase the efficacy of human-computer interfaces, we would do well to support such actions in interface design. As such, the goal would be to capitalize on our knowledge of how the brain supports human behavior in the service of increasing learning through interface design.

Focus Questions

2.1. In what ways are actions multimodal? Compare and contrast multisensory and multimodal as described in this chapter. Do you think these concepts are completely separable?

2.2. Sensory signals are combined in many brain areas, the first being the superior colliculus. How are sensory signals combined in this structure to overcome the binding problem? How is this related to intersensory redundancy?

2.3. What types of behaviors are enhanced by multimodal interactions with the environment across development? In what way are these behaviors changed by this experience?

2.4. How does learning through action change the brain? Include examples of learning about objects and symbols. What would be a practical implication of learning about three-dimensional objects through action vs. through observation?

2.5. Why would handwriting facilitate learning and recruit brain systems used for reading more than typing?

2.6. How can we use our knowledge of brain systems to aid in the construction of multimodal interfaces?

2.7. Given the information provided in this chapter, what are some other multimodal interfaces that would facilitate learning?

2.8. Would the understanding of word meaning be facilitated by multimodal learning? Would some words benefit more from this type of learning than others?

2.9. Are there concepts that we learn independently of our multimodal interactions with the world? What are some examples? How would such concepts be represented in brain systems?

References

W. J. Adams, E. W. Graf, and M. O. Ernst. 2004. Experience can change the "light-from-above" prior. *Nature Neuroscience,* 7(10):1057–1058. DOI: 10.1038/nn1312. 58

D. M. Adams, A. T. Stull, I. Roewer, and M. Hegarty. 2011. Effects of using mental and manual rotation training on mental and manual rotation performance. In L. A. Carlson, C. Hölscher, and T. F. Shipley, editors. *Expanding the Space of Cognitive Science: Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*. pp. 1268–1273. Cognitive Science Society, Austin, TX. 59

K. Adolph. 1997. Learning in the development of infant locomotion. *Monographs of the Society for Research in Child Development,* 62(3):1–162. 64

K. Adolph, L. B. Karasik, and C. Tamis-LeMonda. 2010. Motor skill. In M. H. Bornstein, editor. *Handbook of Cultural Developmental Science*. pp. 61–88. Psychology Press, New York. 64

K. E. Adolph, C. S. Tamis-LeMonda, S. Ishak, L. B. Karasik, and S. A. Lobo. 2008. Locomotor experience and use of social information are posture specific. *Developmental Psychology*, 44(6):1705–1714. DOI: 10.1037/a0013852. 64

S. H. Ambrose. 2001. Paleolithic technology and human evolution. *Science*, 291(5509):1748–1753. 79

M. K. Babcock and J. J. Freyd. 1988. Perception of dynamic information in static handwritten forms. *American Journal Psychology,* 101(1):111–130. 60

L. E. Bahrick and R. Lickliter. 2002. Intersensory redundancy guides early perceptual and cognitive development. In R. Kail, editor. *Advances in Child Development and Behavior*. vol. 30, pp. 153–187. Academic Press, New York. 52

M. Banks and P. Salapatek. 1983. Infant visual perception. In M. M. Haith and J. J. Campos, editors. P. H. Mussen (Series Ed.). *Handbook of Child Ppsychology*: *Vol. 2. Infancy and Developmental Psychobiology*. pp. 43–572. Wiley, New York. 62

L. W. Barsalou. 2008. Grounded cognition. *Annual Review of Psychology,* 59:617–645. DOI: 10.1146/annurev.psych.59.103006.093639. 79

L. W. Barsalou, W. K. Simmons, A. Barbey, and C. Wilson. 2003. Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Science,* 7(2):84–91. 70

P. Bartolomeo, A. Bachoud-Lévi, S. Chokron, and J. Degos. 2002. Visually- and motor-based knowledge of letters: evidence from a pure alexic patient. *Neuropsychologia,* 40:1363–1371. 61

V. W. Berninger, K. Vaughan, R. D. Abbott, A. Brooks, S. P. Abbott, L. Rogan, E. Reed, and S. Graham. 1998. Early intervention for spelling problems: Teaching functional spelling units of varying size with a multiple-connections framework. *Journal of Educational Psychologu,* 90(4):587–605. 60

B. I. Bertenthal, D. R. Proffitt, and J. E. Cutting. 1984. Infant sensitivity to figural coherence in biomechanical motions. *Journal of Experimental Child Psychology*, 37:213–230. 63

E. G. Boring. 1964. Size-constancy in a picture. *American Journal of Psychology,* 77(3):494–498. 85

C. B. Boronat, L. J. Buxbaum, H. B. Coslett, K. Tang, E. M. Saffran, D. Y. Kimberg, and J. A. Detre. 2005. Distinctions between manipulation and function knowledge of objects: evidence from functional magnetic resonance imaging. *Cognitive Brain Research,* 23(2–3):361–373. DOI: 10.1016/j.cogbrainres.2004.11.001. 70

S. Broaders, S. W, Cook, Z. Mitchell, and S. Goldin-Meadow. 2007. Making children gesture reveals implicit knowledge and leads to learning. *Journal of Experimental Psychology General,* 136(4):539–550. DOI: 10.1037/0096-3445.136.4.539.

P. Broca. 1861. Perte de la parole, ramollissement chronique et destruction partielle du lobe antérieur gauche du cerveau. *Bull Soc Anthropol*, 2(1), pp. 235–238. 70

A. J. Butler and K. H. James. 2013. Active learning of novel sound-producing objects: motor reactivism and enhancement of visuo-motor connectivity. *Journal of Cognitive Neuroscience,* 25(2):203–218. DOI: 10.1162/jocn_a_00284. 71, 72

L. J. Buxbaum and E. M. Saffran. 2002. Knowledge of object manipulation and object function: dissociations in apraxic and nonapraxic subjects. *Brain Language,* 82:179–199. 70

C. G. Christou and H. H. Bülthoff. 1999. View dependence in scene recognition after active learning. *Memory & Cognition,* 27(6):996–1007. 58

J. Campos, D. Anderson, M. Barbu-Roth, E. Hubbard, M. Hertenstein, and D. Witherington. 2000. Travel broadens the mind. *Infancy,* 1(2):149–219. 69

J. Campos, B. I. Bertenthal, and R. Kermoian. 1992. Early experience and emotional development: The emergence of wariness of height. *Psychological Science,* 3(1):61–64. 64

P. Cisek. 1999. Beyond the computer metaphor: behaviour as interaction. *Journal of Consciousness Studies,* 6(11–12):125–142. 60

L. Cohen and S. Dehaene. 2004. Specialization within the ventral stream: the case for the visual word form area. *Neuroimage,* 22(1):466–476. DOI: 10.1016/j.neuroimage.2003.12.049. 69

E. Combe and M. Wexler. 2010. Observer movement and size constancy. *Psychological Science*, 21(5):667–675. DOI: 10.1177/0956797610367753. 57

L. Craighero, L. Fadiga, C. Umilta, and G. Rizzolatti. 1996. Evidence for visuomotor priming effect. *Neuroreport,* 8:347–349. 69

J. Dewey. 1896. The reflex arc concept in psychology. *Psychological Review* 3:357–370. 69

S. B. Ehrlich, S. C. Levine, and S. Goldin-Meadow. 2006. The importance of gesture in children's spatial reasoning. *Developmental Psychology,* 42(6):1259–1268. DOI: 10.1037/0012-1649.42.6.1259.

R. Ellis and M. Tucker. 2000. Micro-affordance: The potentiation of components of action by seen objects. *British Journal of Psychology,* 91:451–471. 69

M. O. Ernst and M. S. Banks. 2002. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature,* 415:430–433. 58

S. Exner. 1881. *Untersuchungen über die Lokalisation der Functionen in der Grosshirnrinde des Menschen*. Wilhelm Braunmüller, Vienna, Austria. 70

L. M. Flevares and M. Perry. 2001. How many do you see? The use of nonspoken representations in first-grade mathematics lessons. *Journal of Educational Psychology,* 93(2):330–345. DOI: 10.1037/0022-0663.93.2.330

J. J. Freyd. 1983. Representing the dynamics of a static form. *Memory & Cognition,* 11:342–346. 60

V. Fuentes-Santamaria, J. C. Alvarado, J. G. McHaffie, and B. E. Stein. 2009. Axon morphologies and convergence patterns of projections from different sensory-specific cortices of the anterior ectosylvian sulcus onto multisensory neurons in the cat superior colliculus. *Cerebral Cortex,* 19(12):2902–2915. DOI: 10.1093/cercor/bhp060. 54

V. Gallese, L. Fadiga, L. Fogassi, and G. Rizzolatti. 1996. Action recognition in the premotor cortex. *Brain,* 119:593–609. 75

E. J. Gibson. 1988. Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology,* 39,:1–41.

E. J. Gibson and R. D. Walk. 1960. The "visual cliff." *Scientific America*, 202:67–71. 62, 63, 64

J. J. Gibson. 1979. *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston. 69

J. Grezes and J. Decety. 2002. Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia,* 40:212–222. 70

A. H. Hall., A. Simpson, Y. Guo, and S. Wang. 2015. *Examining the effects of Preschool Writing Instruction on Emergent Literacy Skills: A Systematic Review of the Literature*. Clemson University Press, Clemson, SC. 67

K. L. Harman, G. K. Humphrey, and M. A. Goodale. 1999. Active manual control of object views facilitates visual recognition. *Current Biology,* 9:1315–1318. 58, 65, 66

O. Hauk, I. Johnsrude, and F. Pulvermüller. 2004. Somatotopic representation of action words in human motor and premotor cortex. *Neuron,* 41:301–307. 75

D. O. Hebb. 1949. *The Organization of Behavior—A Neuropsychological Theory*. John Wiley, New York. 55

E. G. Heckenmueller. 1965. Stabilization of the retinal image: A review of method, effects, and theory. *Psychological Bulletin,* 63(3):157–169. DOI: 10.1037/h0021743. 56

R. Held and A. Hein. 1963. Movement-produced stimulation in the development of visually guided behaviour. *Journal of Comparative and Physiological Psychology*, 56:872–876. 56

K. H. James and T. Atwood. 2009. The role of sensorimotor learning in the perception of letter-like forms: Tracking the causes of neural specialization for letters. *Cognntive Neuropsychology*, 26(1):91–110. DOI: 10.1080/02643290802425914. 72, 73, 74

K. H. James and P. Bose. 2011. Self-generated actions during learning objects and sounds create sensori-motor systems in the developing brain. *Cognitive Brain Behavior*, 15(4):485–503. 75

K. H. James and L. Engelhardt. 2012. The effects of handwriting experience on functional brain development in per-literate children. *Trends in Neuroscience and Education*, 1:32–42. DOI: 10.1016/j.tine.2012.08.001. 77, 78

K. H. James and I. Gauthier. 2006. Letter processing automatically recruits a sensory-motor brain network. *Neuropsychologia,* 44(14):2937–2949. DOI: 10.1016/j.neuropsychologia.2006.06.026. 61, 72, 73

K. H. James and I. Gauthier. 2009. When writing impairs reading: Letter perception's susceptibility to motor interference. *Journal Experimental Psychology General*, 138(3):416. 85

K. H. James, G. K. Humphrey, and M. A. Goodale. 2001. Manipulating and recognizing virtual objects: where the action Is. *Canadian Journal of Experimental Psychology,* 55(2):111–120. 58, 59, 66

K. H. James, G. K. Humphrey, T. Vilis, B. Corrie, R. Baddour, and M. A. Goodale. 2002. "Active" and "passive" learning of three-dimensional object structure within an immersive virtual reality environment. *Behavioral Research Methods, Instruments, and Compututers*, 34(3):383–390. 58

K. H. James, S. S. Jones, S. Swain, A. Pereira, and L. B. Smith. 2014. Some views are better than others: evidence for a visual bias in object views self-generated by toddlers. *Developmental Science,* 17(3):338–351. DOI: 10.1111/desc.12124. 65

K. H. James and S. N. Swain. 2011. Only self-generated actions create sensori-motor systems in the developing brain. *Developmental Science,* 14(4):673–687. DOI: 10.1111/j.1467-7687.2010.01011.x. 75, 76

K. H. James, S. N. Swain, S. S. Jones, and L. B. Smith. 2013. Young children's self-generated object views and object recognition. *Journal of Cognitive Development*, 15(3):393–401. DOI: 10.1080/15248372.2012.749481. 65, 82, 626

M. H. Johnson, A. Senju, and P. Tomalski. 2015. The two-process theory of face processing: modifications based on two decades of data from infants and adults. *Neuroscience & Biobehavioral Review*, 50:169–179. DOI: 10.1016/j.neubiorev.2014.10.009. 55

N. Kanwisher, J. McDermott, and M. Chun. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17(11):4302–4311. 69

A. J. Kersey and K. H. James. 2013. Brain activation patterns resulting from learning letter forms through active self-production and passive observation in young children. *Frontiers in Psychology*, 4:1–15. DOI: 10.3389/fpsyg.2013.00567. 78

R. L. Klatzky, J. M. Loomis, A. C. Beall, S. S. Chance, and R. G. Golledge. 1998. Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science* 9(4):293–298 58

K. S. Kretch and K. E. Adolph. 2013. Cliff or step? Posture-specific learning at the edge of a drop-off. *Child Development*, 84:226–240. 64

G. Lakoff and R. Nunez. 2000. *Where Mathematics Comes From.* Basic Books, New York. 60

K. M. Lee, C. M. Kim, and S. H. Woo. 2001. fMRI comparison between expert and novice perception of dance. *NeuroImage*, 13(6):907. 75

R. Lickliter and L. E. Bahrick. 2004. Perceptual development and the origins of multisensory responsiveness. In G. Calvert, C. Spence & B. Stein, editors. *Handbook of Multisensory Processes*. MIT Press, Cambridge, MA. 80

J. X. Li and K. H. James. 2016. Handwriting generates variable visual output to facilitate symbol learning. *Journal of Experimental Psychology: General*, 145(3):298–313. DOI: 10.1037/xge0000134. 60, 68

K. Libertus and A. Needham. 2010. Teach to reach: the effects of active versus passive reaching experiences on action and perception. *Vision Research*, 50(24):2750–2757. 65

J. J. Lockman. (2000). A perception-action perspective on tool use development. *Child Development*, 71:137–144. 61

M. Longcamp J. L. Anton, M. Roth, and J. L. Velay. 2003. Visual presentation of single letters activates a premotor area involved in writing. *Neuroimage*, 19(4), pp. 1492–1500. 72

M. Longcamp, M. T. Zerbato-Poudou, and J. L. Velay. 2005. The influence of writing practice on letter recognition in preschool children: A comparison between handwriting and typing. *Acta Psychologica*, 119:67–79. DOI: 10.1016/j.actpsy.2004.10.019. 67

M. Longcamp, T. Tanskanen, and R. Hari. 2006. The imprint of action: motor cortex involvement in visual perception of handwritten letters. *Neuroimage*, 33:681–688. DOI: 10.1016/j.neuroimage.2006.06.042.

M. Longcamp, C. Boucard, J. C. Gilhodes, J. L. Anton, M. Roth, B. Nazarian, and J. L. Velay. 2008. Learning through hand-or typewriting influences visual recognition of new graphic shapes: Behavioral and functional imaging evidence. *Journal of Cognitive Neuroscience*, 20(5), pp. 802–815. 72

M. Longcamp, Y. Hlushchuk, and R. Hari. 2011. What differs in visual recognition of handwritten vs. printed letters? An fMRI study. *Human brain mapping*, 32(8), pp. 1250–1259. 72

M. Longcamp, A. Lagarrigue, B. Nazarian, M. Roth, J. L. Anton, F. X. Alario, and J. L. Velay. 2014. Functional specificity in the motor system: evidence from coupled fMRI and kinematic recordings during letter and digit writing. *Human Brain Mapping*, 35:6077–6087. 72

V. J. Molfese, J. L. Beswick, J. L. Jacobi-Vessels, N. E. Armstrong, B. L. Culver, J. M. White, M. C. Ferguson, K. M. Rudasill, and D. L. Molfese. 2011. Evidence of alphabetic knowledge in writing: Connections to letter and word identification skills in preschool and kindergarten. *Reading and Writing*, 24(2), 133–150. DOI: 10.1007/s11145-010-9265-8. 67

L. L. Namy and D. Gentner. 2002. Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General*, 131(1):5–15. DOI: 10.1037/0096-3445.131.1.5.. 68

M. J. Nathan. 2008. An embodied cognition perspective on symbols, gesture and grounding instruction. In M. DeVega, A. M. Glenberg, and A. C. Graesser, editors. *Symbols, Embodiment and Meaning: A Debate*. Oxford University Press, Cambridge, UK.

A, Needham. 2000. Improvements in object exploration skills may facilitate the development of object segregation in early infancy. *Journal of Cognitive Development*, 1:131–156. 65

A. Needham, T. Barrett, and K. Peterman. 2002. A pick-me-up for infants' exploratory skills: Early simulated experiences reaching for objects using "sticky mittens" enhances young infants' object exploration skills. *Infant Behavior and Development*, 25:279–295. 65

J. O'Regan and A. Noë. 2001. A sensorimotor account of vision and visual consciousness. *Behavior and Brain Sciences*, 24:939–1031. 69

J. Orliaguet, S. Kandel, and L. Boe. 1997. Visual perception of motor anticipation in cursive handwriting: Influence of spatial and movement information on the prediction of forthcoming letters. *Perception*, 26:905–912. 60

S. Oviatt. 2017. Please proivde Biblio Info.

W. Penfield and E. Boldfrey. 1937. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain*, 60(4):389–443. 70

A. F. Pereira, K. H. James, S. S. Jones, and L. B. Smith. 2010. Early biases and developmental changes in self-generated object views. *Journal of Vision*, 10(11):22:1–13. DOI: 10.1167/10.11.22. 65, 66, 84

M. Petrides. 2015. The ventrolateral frontal region. *Neurobiology of Language*, 25. 70

J. Piaget. 1952. *The Origins of Intelligence in Children. (M. Cook, Translation).* International University Press, New York. 61

J. Piaget. 1954. *The Construction of Reality in the Child*. Basic Books, New York 61

K. J. Pine, N, Lufkin, and D. Messe 2004. More gestures than answers: Children learning about balance. *Developmental Psychology*, 40:1059–106.

F. Pulvermüller. 2005. Brain mechanisms linking language and action. *Nature Revviews Neuroscience*, 6:576–583. 75

F. Pulvermüller. 2012. Meaning and the brain: The neurosemantics of referential, interactive, and combinatorial knowledge. *Journal of Neurolinguistics*, 25(5):423–459. DOI: 10.1016/j.jneuroling.2011.03.004. 75

F. Pulvermüller. 2013. How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends Cognitive Science*, 17(9):458–470. DOI: 10.1016/j.tics.2013.06.004. 75

F. Pulvermüller, O. Hauk, V. V. Nikulin, and R. J. Ilmoniemi. 2005. Functional links between motor and language systems. *European Journal of Neuroscience*, 21(3):793–797. DOI: 10.1111/j.1460-9568.2005.03900.x.

T. L. Richards, V. W. Berninger, P. Stock, L. Altemeier, P. Trivedi, and K. R. Maravilla. 2011. Differences between good and poor child writers on fMRI contrasts for writing newly taught and highly practiced letter forms. *Reading and Writing*, 24*(5)*:493–516. DOI: 10.1007/s11145-009-9217-3. 60

L. E. Richland. 2015. Linking gestures: cross-cultural variation during instructional analogies. *Cognition and Instruction*, 33(4):295–321.

L. E. Richland, O, Zur, and K. Holyoak. 2007. Cognitive supports for analogies in the mathematics classroom. *Science,* 316(5828):1128–1129. DOI: 10.1126.science.1142103.

J. J. Rieser, D. A. Guth, and E. W. Hill. 1986. Sensitivity to perspective structure while walking without vision. *Perception,* 15:173–188. 58

G. Rizzolatti and L. Craighero. 2004. The mirror-neuron system. *Annual Review of Neuroscience*, 27:169–192. DOI: 10.1146/annurev.neuro.27.070203.144230. 75

P. Rochet. 1989. Object manipulation and exploration in 2- to 5-month-old infants. *Developmental Psychology*, 25:871–884. 64

F. E. Roux, O. Dufor, C. Giussani, Y. Wamain, L. Draper, M. Longcamp, and J. F. Démonet. 2009. The graphemic/motor frontal area Exner's area revisited. *Annals of neurology*, 66(4), pp. 537–545. 70

F. E. Roux, L. Draper, B. Köpke, and J. F. Démonet. 2010. Who actually read Exner? Returning to the source of the frontal "writing centre" hypothesis. *Cortex*, 46(9):1204–1210.

H. A. Ruff, L. M. Saltarelli, M. Capozzoli, and K. Dubiner. 1992. The differentiation of activity in infants' exploration of objects. *Developmental Psychology*, 28:851–861 64

C. Sann and A. Streri. 2008. The limits of newborn's grasping to detect texture in a cross-modal transfer task. *Infant Behavior and Development*, 31(3):523–531. DOI: 10.1016/j.infbeh.2008.03.001. 62

K. Seki, M.Yajima, and M. Sugishita. 1995. The efficacy of kinesthetic reading treatment for pure alexia. *Neuropsychologia,* 33:595–609 61

W. K. Simmons and L. W. Barsalou. 2003. The similarity-in-topography principle: reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, 20(3):451–486. DOI: 10.1080/02643290342000032. 70

D. J. Simons and R. F. Wang. 1998. Perceiving real-world viewpoint changes. *Psychological Science*, 9(4):315–320. 58

L. B. Smith. 2005. Action alters shape categories. *Cognitive Science*, 29:665–679. 66

C. E. Snow, M. S. Burns, and P. Griffin, editors. 1998. *Preventing Reading Difficulties in Young Children*. National Academy Press, Washington, DC. 67

J. A. Sommerville, A. L. Woodward, and A. Needham. 2005. Action experience alters 3-month-old infants' perception of others' actions. *Cognition,* 96:B1–B11. 65

B. E. Stein and M. A. Meredith. 1993. *The Merging of the Senses*. MIT, Cambridge, MA. 54

B. E. Stein and B. A. Rowland. 2011. Organization and plasticity in multisensory integration: early and late experience affects its governing principles. *Progress in Brain Research*, 191:145–163. DOI: 10.1016/B978-0-444-53752-2.00007-2. 52, 54

B. E. Stein, T. R. Stanford, and B. A. Rowland. 2014. Development of multisensory integration from the perspective of the individual neuron. *Nature Reviews Neuroscience*, 15(8):520–535. DOI: 10.1038/nrn3742. 52, 53, 54

M. Tucker and R. Ellis. 1998. On the relations between seen objects and components of potential actions. *Journal of Experimental Pyschology: Human Perception and Performance*, 24(3):830–846. 69

S. Vinci-Booher, T. W. James, and K. H. James, 2016. Visual-motor functional connectivity in preschool children emerges after handwriting experience. *Trends in Neuroscience and Education*, 5(3).107-120. DOI: 10.1016/j.tine.2016.07.0. 78

M. T. Wallace, M. A. Meredith, and B. E. Stiem. 1992. Integration of multiple sensory modalities in cat cortex. *Experimental Brain Research*, 91:484–488. 54

M. T. Wallace, M. A. Meredith, and B. E. Stiem. 1993. Converging influences from visual, auditory and somatosensory cortices onto output neurons of the superior colliculus. *Journal of Neuropsychology*, 69(6):1797–1809. 54

M. Wexler. 1997. Is rotation of visual mental images a motor act? In K. Donner, editor. *Proceedings of the 20th European Conference on Visual Perception, ECUP'97*. Pion Ltd., London. 59

M. Wilson. 2002. Six views of embodied cognition. *Psychonomic Bulletin Review*, 9(4):625–636. 79

A. Wohlschläger and A. Wohlschläger. 1998. Mental and manual rotation. *Journal of Experimental Psychology: Human Perception and Performance*, 24(2):397–412.