

Viewpoints

Multisensory Integration and the Society for Neuroscience: Then and Now

 Barry E. Stein,  Terrence R. Stanford, and  Benjamin A. Rowland

Department of Neurobiology and Anatomy, Wake Forest School of Medicine, Winston-Salem, North Carolina 27157

The operation of our multiple and distinct sensory systems has long captured the interest of researchers from multiple disciplines. When the Society was founded 50 years ago to bring neuroscience research under a common banner, sensory research was largely divided along modality-specific lines. At the time, there were only a few physiological and anatomical observations of the multisensory interactions that powerfully influence our everyday perception. Since then, the neuroscientific study of multisensory integration has increased exponentially in both volume and diversity. From initial studies identifying the overlapping receptive fields of multisensory neurons, to subsequent studies of the spatial and temporal principles that govern the integration of multiple sensory cues, our understanding of this phenomenon at the single-neuron level has expanded to include a variety of dimensions. We now can appreciate how multisensory integration can alter patterns of neural activity in time, and even coordinate activity among populations of neurons across different brain areas. There is now a growing battery of sophisticated empirical and computational techniques that are being used to study this process in a number of models. These advancements have not only enhanced our understanding of this remarkable process in the normal adult brain, but also its underlying circuitry, requirements for development, susceptibility to malfunction, and how its principles may be used to mitigate malfunction.

Introduction

When the Society for Neuroscience was founded 50 years ago, efforts to understand the functional properties of neural systems were growing rapidly, but were scattered among a variety of disciplines (e.g., physiology, biology, psychology, linguistics, philosophy, computer science). One of the overarching ambitions of the founding committee was to bring us together into a single Society that would encompass the enormous diversity of this burgeoning field. Investigators using very different conceptual frameworks, scientific approaches, model species, and concerned with different neurological issues, would be gathered under a single umbrella. Even choosing the appropriate name of such a society was an issue, as the term “neuroscience” was not yet in vogue.

The founders did an excellent job, and one clear indicator of the Society’s success is its rapid and continued increase in membership, from several hundred at its inception to >37,000 today. If you want to be involved in the neuroscientific community, the first step is to become a member of SfN.

One of the larger research contingents at the first official meeting of the Society in 1971 was devoted to understanding the operation of vertebrate sensory systems. This was still evident at the most recent meetings and is, perhaps, not surprising. As noted in the 18th century by Immanuel Kant, all our knowledge of the world begins with the senses. And yet, as René Descartes articulated a century prior, the senses are fallible, and our perception is necessarily based on inference. How could one not be interested in how these systems operate?

There was enormous excitement at that initial meeting of the Society, an excitement that has only increased, about using an increasingly sophisticated and diverse set of experimental techniques that have made it possible to better understand how sensory organs transduce signals, how the brain segregates and distributes that information for processing, how sensation relates to perception, how perception relates to behavior, how all of these sensory abilities arose and changed during the evolution of extant species, and how they change during the maturation of the individual.

The standard neurological approach at the time the Society was founded, and for some time thereafter, was sensory-specific. There was no field of neuroscience representing “multisensory” research. Existing concepts emphasized the segregation of the senses, including the 19th century “Law of Specific Nerve Energies” and various versions of the “Labeled Line” theory, in which each sense was believed to have dedicated receptors, fibers, and target regions. Textbooks were organized around the “5 senses” proposed by Aristotle (many more are now recognized), with little discussion of their possible interaction. It was also not unusual for a researcher to be identified by the sensory modality he or she studied (e.g., a “visual scientist” or an “auditory scientist”).

Experimental design reflected this view, and common experimental controls involved minimizing cues from senses not under study to eliminate their possible “confounding” influences. The expressed concern was that they would change the measures of interest due to their general influence on arousal. But it is important to recognize that researchers were not unaware that other senses could also have more specific effects on one another. There was already a long history of perceptual research demonstrating the potent effects of intersensory interactions (especially visual-auditory) on perception and reaction time (for discussion, see Marks, 1978; Walk and Pick, 1981; Stein and Meredith, 1993), a field that has expanded and continues to thrive (see Bruno and

Received April 1, 2019; revised Sept. 10, 2019; accepted Sept. 21, 2019.

This work was supported in part by National Institutes of Health Grants EY026916, EY024458, and EY016916.

The authors declare no competing financial interests.

Correspondence should be addressed to Barry E. Stein at bestein@wakehealth.edu.

<https://doi.org/10.1523/JNEUROSCI.0737-19.2019>

Copyright © 2019 the authors

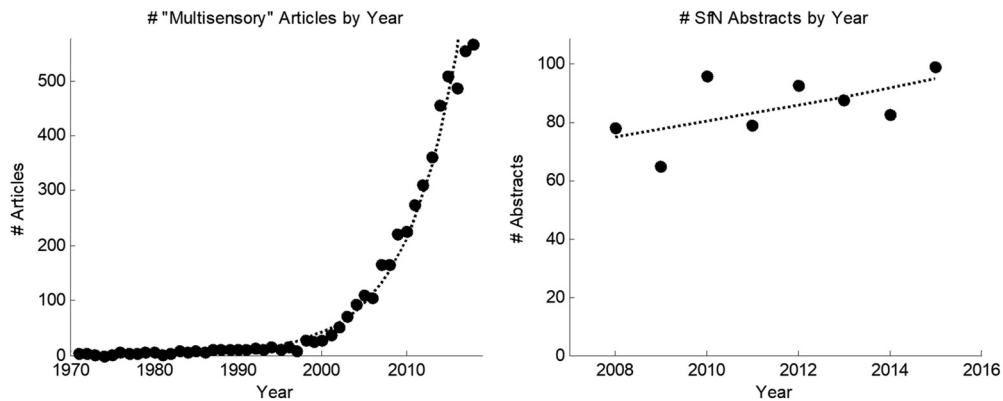


Figure 1. The rapid growth of interest in multisensory integration. Left, Number of research articles indexed by the key word “multisensory” (on PubMed) published each year since the inception of the annual meeting in 1971. Right, Number of multisensory-related abstracts at the annual meeting of the Society for Neuroscience (years 2008–2015), including key words multisensory, polysensory, intersensory, cross-modal, heteromodal, multimodal, polymodal, supramodal, and amodal.

Pavani, 2018). But at the time, these interactions were thought to take place somewhere “out there” in higher-order association cortex. There were no systematic efforts to understand how and where the response properties of neurons were altered to produce these multisensory perceptual effects. The dominant focus in sensory physiology was on sensory-specific questions along the primary sensory pathways, with only isolated reports identifying the presence of neurons responding to sensory inputs that should not be available (e.g., auditory-responsive neurons in cat visual cortex), and anatomical studies identifying cross-projections between ostensibly “unisensory” cortical areas.

Since this time, research in the field of multisensory processing and integration has increased exponentially (Fig. 1). The enormous amount of interest and energy now devoted to studying interactions between the senses makes it difficult in this brief overview to give proper credit to the individual contributions of the many researchers involved, so we have elected to steer the reader to published books and compendiums that more fully explore issues relating to neural computation, sensory development, perceptual psychophysics, and clinical relevance from a multisensory perspective. The number of species in which this has been studied now spans a wide range, from insect to human, and the functional impact of a host of modality-convergence patterns has been studied, or is currently being studied. How did we get here? Below we provide a brief review of evolution in thinking about sensory organization and representations, the principles of multisensory integration, the multisensory transform and its computational bases, its development and plasticity, and new translational applications of this knowledge.

Sensory organization

Sensory function was actively being studied long before the Society was begun; and by the 1950s, the cat had become the model of choice for many neurological studies. Vernon Mountcastle had used this model to examine the microstructure of somatosensory cortex. He found that it had a modular composition wherein receptive fields were organized into interconnected vertical columns of 300–600 μm in diameter. The properties of neurons in individual columns were similar to one another but differed systematically from those in adjacent columns. He then found a similar organization in the cortex of the monkey. This elemental organizational feature of the mammalian brain has guided research to this day.

Mountcastle’s findings prompted similar studies, with similar findings, in the cat visual cortex by David Hubel and Torsten

Wiesel. The obvious organizational constancies across the different sensory representations led to a host of studies to determine how local features of cortical representations are created. Although Hubel and Wiesel’s groundbreaking studies of the visual system in the 1960s and 1970s were directed at understanding how neurons in visual cortex responded to the features of a stimulus and, ultimately, how the brain uses their individual contributions to recreate a scene, they were also deeply interested in how the mammalian visual system developed. Of special interest was the impact of experience on its functional organization. Their work on the developing kitten visual cortex was transformative. It had direct implications for dealing with developmental problems in human vision, inspired generations of visual scientists, and led to an upsurge in inquiries into activity-dependent neural plasticity, often using the visual system as a model. They were awarded the Nobel Prize for Physiology or Medicine in 1981.

Overlapping sensory representations

At the first two SfN meetings, one of us (B.E.S.) presented data regarding the response properties of visual neurons and the development of visual and nonvisual (auditory and somatosensory) neurons in the cat superior colliculus (SC). At the time, studies of the SC were being conducted in a variety of animals, and many of the findings supported a fundamental organizational pattern that superseded species. The SC was known primarily as a visual structure with the role of initiating orientation responses (e.g., shifts of gaze). It was initially believed that its sensory and motor functions were separated in a laminar fashion. The superficial layers were purely sensory (visual) with a map-like representation of visual space. The deeper layers had the motor representation (map), in which adjacent regions directed gaze shifts to adjacent locations in space (a similar organization for the control of other body parts was discovered later). However, later findings revealed that the deeper layers, which are also responsive to visual, auditory, and somatosensory stimuli, are the sites at which sensory inputs are converted to premotor outputs (i.e., sensorimotor transformation) and underlie the role of the SC in detection, localization, and orientation behavior. The SC soon became known as a multisensory structure that could integrate its multiple sensory inputs to facilitate these behaviors (Fig. 2). It also became a useful model for the study of this phenomenon and comprises much of the discussion here. Although the SC does not contain all possible modality convergence patterns, the manner in which it integrates its multiple sensory inputs is instructive and helpful in examining the consequences of other modality conver-

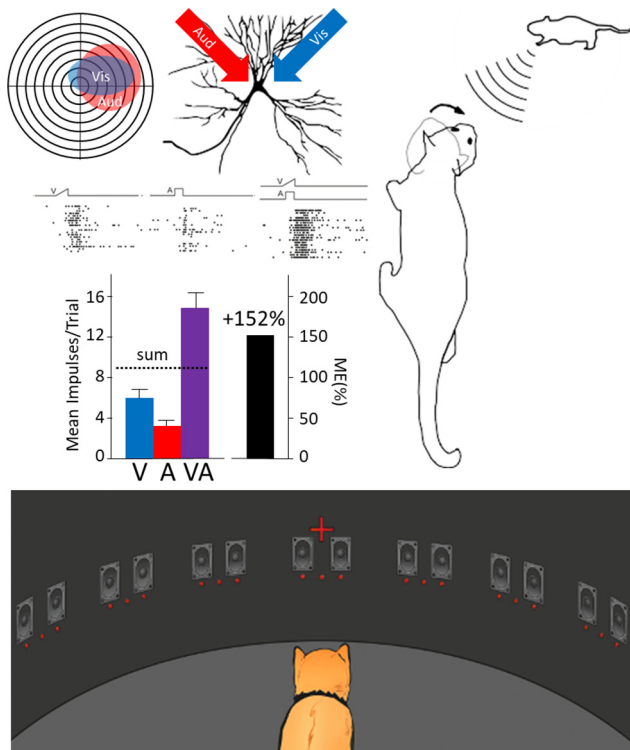


Figure 2. The cat model used to study the multisensory principles of SC neurons. Top left, The overlapping visual and auditory receptive fields of a multisensory SC neuron on a polar map of visual-auditory space. Below it are impulse rasters showing the neuron's unisensory (V, Visual; A, auditory) and multisensory (VA) responses. Bar graph below them represents the magnitude of the multisensory enhancement evoked by their combination (ME, the proportionate amplification relative to the best unisensory response). Error bars indicate s.e.m., dashed line indicates the sum of the unisensory response magnitudes. Top right, This physiological enhancement facilitates the detection, localization, and orientation roles of the structure. Bottom, A perimetry device with LEDs and speakers used to probe multisensory behavioral enhancement.

gence patterns elsewhere in the nervous system, but more about this later.

The first description of the topographic representation of the body in the cat SC was at the 1974 SfN meeting. As noted above, this deep layer somatotopic map followed the same organizational pattern as did the map of visual space. In both cases, there was an expanded representation of central or forward space (e.g., macular vision and the face and head). Contemporaneous and later research revealed that there was also a similar spatiotopic auditory representation and that all these sensory maps were in spatial register with the deep layer motor maps. This proved to be a general mammalian plan. Similar research on the optic tectum, the nonmammalian homolog of the SC, strongly suggested that this scheme of overlapping maps was retained during the evolutionary transition from pre-mammalian vertebrates. Yet, despite the similarities in the sensory representation across species, there are notable species-specific specializations. These make intuitive sense, and a few obvious examples include the expansion of the whisker representation in the rodent, the expansive representation of the exotic star nose organ in the blind mole, the specialized auditory representation in echolocating bats, and the representation of the infrared organ in the pit viper. But whenever examined, the topographic nature of the representations and the overlap between sensory and motor maps were shown to be conserved organizational principles.

The sensory maps are not formed in complete independence of one another. Using the optic tectum of the owl and the SC of

the ferret as developmental models, researchers showed that the visual representation guides the formation of the auditory map. Shifts of the visual axis induced by prisms or surgery on the extraocular muscles produced corresponding shifts in the auditory map. Intermapping can also be quite rapid during overt behavior and has been shown in the adult monkey. When the animal voluntarily shifts its eyes to look at some environmental event, the physical change in the visual axis reorients an SC neuron's eye-centered visual receptive field so that it now samples information from a different area of space. However, there is also a corresponding shift in its auditory and/or somatosensory receptive field to minimize any misalignment among the maps and keep them in approximate register (see Stein and Meredith, 1993).

This makes good sense. Maintaining the alignment among maps ensures that different sensory inputs that derive from the same event will produce activation in the same SC locus and be referred to the same point in its motor maps. Because many of the neurons forming these maps are multisensory (e.g., visual-auditory, visual-somatosensory, trisensory), the same neurons participate in multiple sensory maps. Thus, it may not be entirely appropriate to think of these representations as independent unisensory entities that communicate with one another during their formation in early life and/or during their function in adulthood, but as part of an overarching multisensory representation. Because these SC "sensory" neurons can also have "motor" properties, even the distinction between sensory and motor representations is fuzzy. Similar interpretive issues have arisen in the study of other sensorimotor and decision-making areas of the brain, including those in "higher-order" cortex (e.g., frontal eye fields).

Principles of multisensory integration

When first described at the SfN meeting, the utility of these overlapping sensory topographies in the SC was believed only to reflect a principle of biological conservatism. There was obvious efficiency in having different sensory inputs initiate orientation movements through a common motor map in a structure like the SC, avoiding the necessity of creating a sensory-motor interface for each.

It was not until later in the 1970s and early 1980s that a different and complementary function of this scheme was demonstrated. SC neurons integrate the inputs they receive from different sensory modalities to produce new neural products. The type of product elicited is dependent on stimulus configuration according to a logic that is consistent with the SC's functional role. Cross-modal stimulus configurations that are likely to be derived from the same event (i.e., are in spatial and temporal concordance) enhance the discharges of SC neurons and, thus, the physiological salience of the initiating event. The signals transduced from such stimuli are mutually reinforcing indicators of the presence of the event obtained from independent sources. Cross-modal stimuli that are discordant in either or both of these dimensions either fail to induce an interaction or degrade the neuron's response. In this case, the transduced signals are likely to refer to different events that will compete for an orientation response.

The first neurophysiological observations of this sensitivity were followed by behavioral studies showing that reasonable intuitions about the behavioral consequences of these physiological observations were correct. The same stimulus configurations that increase the physiological salience of the event also enhance its behavioral impact (i.e., they increase the likelihood of detecting, localizing, and orienting toward the originating stimuli). On the other hand, stimulus configurations that reduce physiological salience reduce their behavioral impact, suppressing or eliminat-

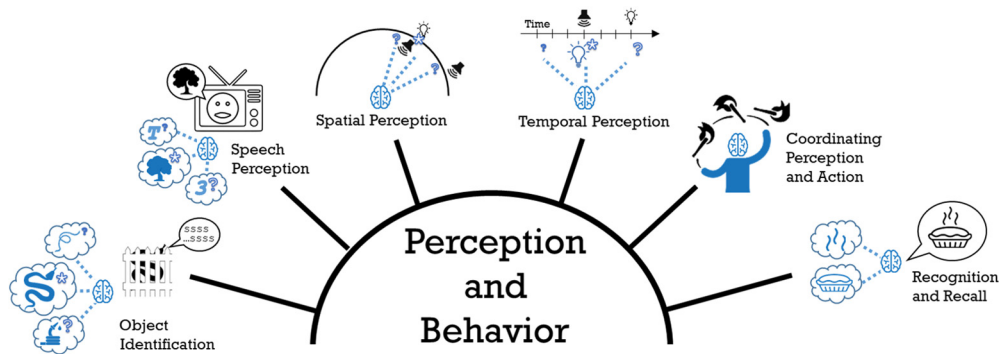


Figure 3. Multisensory integration enhances performance in a number of perceptual and behavioral domains.

ing behavioral responses to the stimuli. Elegant in its simplicity, this organizational scheme could use any combination of cross-modal inputs to put the animal in the best position to deal with initiating events (Stein and Meredith, 1993).

The spatial and temporal principles of multisensory integration proved to be independent of species and have been observed not only in single neuron recordings in the cat, but in the monkey, ferret, owl, guinea pig, rat, snake, and even the frog tadpole. Enhancement with spatiotemporally concordant stimuli was also observed using fMRI in the human SC. That these very basic principles supersede species and phyletic level is likely because the physical constancies of space and time supersede most ecological niches, thereby rendering these same principles broadly adaptive. Undoubtedly, species-specific adaptations are overlaid on these core principles to ensure that the multisensory system is sensitive to the particular needs of a given animal, and may even be modified in different circumstances (e.g., in dealing with events in near or far space), events that are transient or sustained, predictable or not, with strong emotional content or not. These are all active areas of research (see Stein, 2012).

The SfN meetings served as a major opportunity for discussions about these findings, and these discussions were a major impetus to examine questions about whether these principles of space and time reflected sensitivities specific to the SC (and its functional role) or more general neural strategies. And indeed, it now appears as if there are different guiding principles that govern the integration of stimuli for the many different perceptual and behavioral domains in which multisensory integration has been documented (see also Stein and Meredith, 1993; Calvert et al., 2004; Spence and Driver, 2004; Naumer and Kaiser, 2010; Stein, 2012), some of which are illustrated in Fig. 3. The emerging unified principle from this work is that, in all of these domains, multisensory integration obeys an intuitive logic that depends on congruency of information: when the cross-modal signals offer congruent information, they are integrated to produce enhanced neural and behavioral products. When they offer incongruent information, stimuli are either not integrated (i.e., are segregated), or lead to degraded products.

The specific stimulus features and configurations determining congruency vary across functional domains. Thus, for example, when integrating visual and vestibular signals to estimate an animal's heading direction, stimuli with aligned direction yield enhancement, and misaligned directions yield depression. Similarly, the congruency of visual shapes and felt textures significantly impacts judgments of the physical shapes of objects; the similarity of the temporal patterning of visual and auditory stimuli directly impacts the perception of their rhythmicity; and the ability to interpret social communication cues, such as human

speech and primate vocalizations, is enhanced when sound is paired with congruent facial expression. Multisensory neurons in PFC have also been shown to be sensitive to both the synchrony and semantic context (i.e., facial expression/sound) of vocalizations. These examples are also illustrative of the range, from simple to complex, of perceptual problems for which there is benefit in combining independent sources of information. Space and time are important determinants in many, but not all, of these cases. The guiding principle of congruency appears dependent on whether or not, in a specific functional domain, the stimulus features indicate that the stimuli are derived from a common cause. But when stimulus features are congruent and are able to elicit an enhanced response, there is still an issue of how this response should be scaled to the inputs. In short, how should the multisensory transform operate?

The multisensory transform

The issue of the multisensory transform was first addressed at SfN meetings in the mid-1980s. It was intuitive that space and time should be guiding principles for integrating cross-modal stimuli in the SC, but it was less obvious how the products of integration should be scaled. For example, if visual and auditory stimuli each elicited two impulses from an SC neuron when presented individually, what should be expected from their congruent copresentation: three impulses, six impulses, 12 impulses? Exploring this issue led to identification of the principle of inverse effectiveness: greater proportionate enhancement levels were found to be associated with combinations of less effective stimuli. This too makes intuitive sense, as the weaker the sensory evidence about the initiating event, the greater the "benefit" the brain derives from augmenting it with information from another, independent sensory source. The most impressive amplifications were seen when individual cues were below threshold and failed to activate the neuron, but when combined produced reliable responses, multisensory integration produced "something" from "nothing." The products of this integration are very different from those that occur when multiple stimuli are registered within the same modality-specific sensory channel, logically reflecting the difference in informational gains when integrating cross-modal information (which are derived from independent sources) versus within-modal information.

SC neurons are rendered capable of multisensory integration by virtue of being embedded in a specific circuit. They receive their converging unisensory inputs from many sources, including a variety of cortical regions. One region in particular within association cortex (the anterior ectosylvian sulcus [AES] in cat) proved to be of special importance. It contains regions that are primarily visual, auditory, or somatosensory, and unisensory

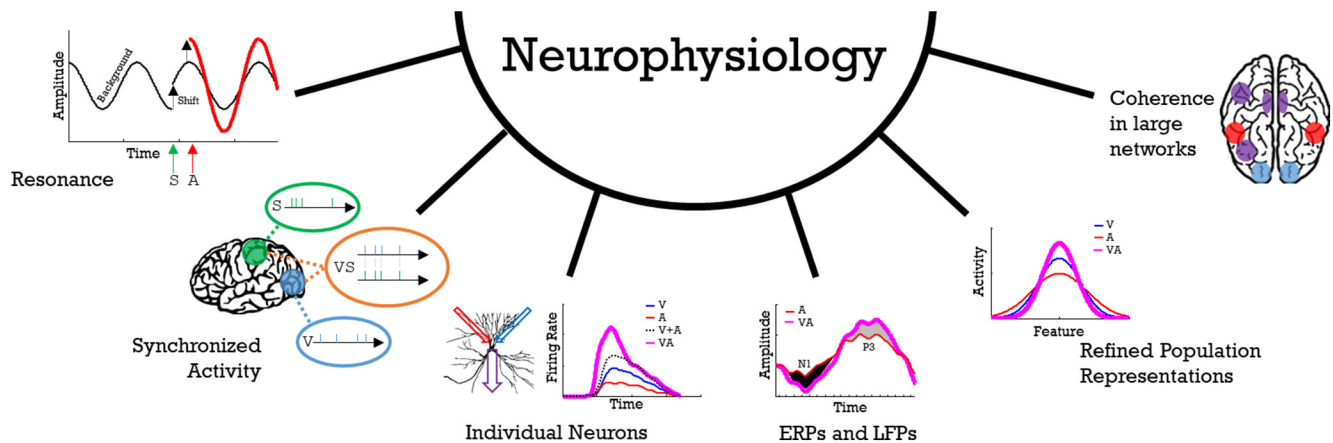


Figure 4. The physiological bases of multisensory integration can be measured in different ways, and can manifest differently in different brain regions. Left to right, Inputs from another modality (S) can shift the phase of ongoing background oscillations to resonate with an incoming signal (A). “Unisensory” neurons in different brain regions can synchronize their activity to amplify the impact of their signals on target structures. Individual multisensory neurons can integrate incoming signals as soon as they arrive to amplify responses. Event-related potential and local field potential methods are used to detect gross changes in responses at the ensemble level. Multisensory integration can improve discrimination by producing more reliable distributions of activity in a feature map, and can lead to more coherent activation patterns within large-scale networks, reflecting more efficient information processing.

neurons from these regions converge directly onto multisensory SC neurons. Experiments involving deactivating or ablating the visual and/or auditory inputs from AES failed to disrupt the modality profile of their common SC target neurons (they were still visual-auditory). But it did render those SC neurons unable to integrate visual and auditory inputs to enhance their responses. At best, the neurons now responded as if only one (the most effective) of the component stimuli was present and often responded with diminished responses that approximated the average of the two inputs. These physiological effects were reflected behaviorally. Deactivation of AES did not alter an animal’s unisensory performance but prevented it from benefiting from combinations of congruent cross-modal cues. That a cortical region oversees how an SC neuron will deal with its converging sensory inputs was surprising. Determining whether this is a general mammalian plan and, if so, what homologs of AES exist in other species are obvious targets of future studies.

In the SC, the products of multisensory integration are fully characterized at the single-neuron level in terms of the response magnitude, latency, and firing rate, and they are readily related to behavioral impact. Supralinear products of the transform are especially common when congruent unisensory signals first arrive at their target neurons (a phenomenon referred to as the initial response enhancement). But its products change, becoming additive or subadditive, as the response evolves and winds down, potentially affecting different components of a behavioral response. Similar patterns of multisensory integration have been noted in cortical regions that engage similar computations. It is important to note that, while a supralinear product based on impulse counts or firing rate is logical for the immediate detection of multiple ambiguous cues in a noisy environment, as detailed below, it is not the ideal product in other sensory discriminative contexts, such as calculating a singular direction of self-motion from visual and vestibular sources (see below). And this schematic contrasts with other brain mechanisms, such as those that use response timing rather than magnitude, to integrate information (e.g., in forming resonant assemblies of oscillating circuits).

Over the years, it has come to seem more likely that the multisensory transform varies with the computational problem that a given perceptual evaluation poses. Whereas the supralinear computation often seen in the SC is ideal for detection and localiza-

tion functions, a strictly linear computation is best for weighted cue integration in which a single estimate of some composite feature is needed. A good example of this is the strategy used for estimating self-motion (see Angelaki et al., 2009). Within the dorsal medial superior temporal (MSTd) area of extrastriate cortex, so-called “congruent” neurons show similar tuning preferences for visual and vestibular motion direction. It is the activity of these particular neurons that appears to underlie the benefit that multisensory integration provides to estimate the direction of self-motion. MSTd neurons combine their inputs in a weighted linear fashion to generate steeper direction-tuning functions, precisely the transformation needed to optimize the discrimination of heading direction and a direct neural correlate of the increased behavioral sensitivity that is observed when the perception of heading is based on both visual and vestibular cues.

In contrast to these examples, in which the actions of well-defined multisensory circuits yield mechanistically intuitive outcomes in changing response magnitude or tuning functions of neurons, in other brain circuits the principal effect of multisensory integration may relate to changes in the phase of ongoing neuronal oscillations, or in the relative timing of impulses or oscillations, spectral coherence, or even the relative changes in these measures in different neural populations. How these relate to perception and behavior is less well understood, as is how they relate to the issues of congruency described above. In the case of the SC, the temporal alignment of the incoming cross-modal signals is a crucial factor in producing the characteristic multisensory enhancement. But for oscillation coherence, exact stimulus onsets and offsets are less crucial. There is a possibility that this variance in the sensitivity of different physiological products will be predictive of variance in the different types of behavioral improvements that are observed in different contexts (e.g., enhanced perception, or the accuracy of behavioral decisions, or in the speed of reacting to sudden events). But this is only beginning to be examined using a host of physiological techniques, some of which are illustrated in Figure 4 (Stein, 2012; see also Murray and Wallace, 2012). Nevertheless, it is safe to assume that products of integration within each circuit are tailored to its functional role(s), and how this is accomplished will have to be examined in each multisensory region.

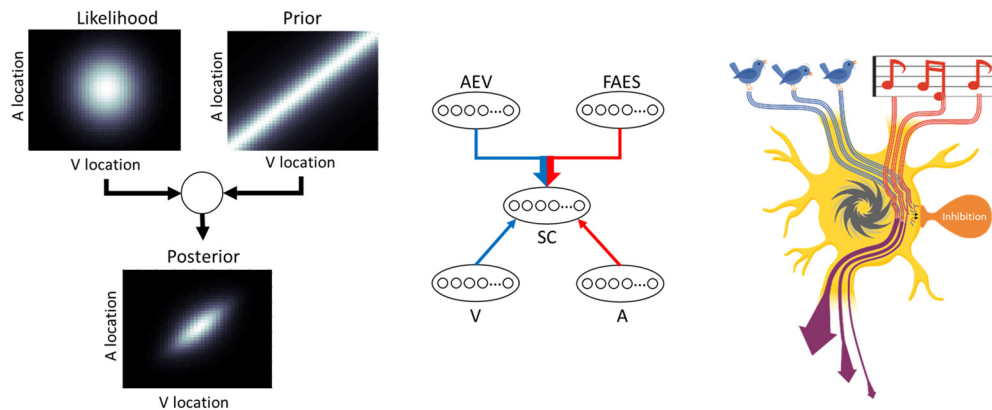


Figure 5. Computational modeling of multisensory integration. Left, Bayesian frameworks describe how ambiguous sensory signals can be combined with prior expectations that the information being offered refers to the same event to form optimal multisensory estimates. Depicted are variable estimates of a sensory feature from each modality that combine to form a joint distribution, which is then combined with a prior distribution peaked along the diagonal, representing an assumption that these signals have a common cause. The product is a distribution of multisensory estimates that represents an optimal combination of the unisensory inputs and is more reliable than either alone. Middle, Network models of multisensory integration have increased in sophistication from simple, abstract architectures involving three areas (e.g., two unisensory areas and one multisensory area) to models that include multiple biologically realistic inputs. In this diagram, ovals represent processing areas containing multiple units (circles). There are a total of four modeled input areas: two derived from cortical regions AES (AEV, visual; FAES, auditory) and two derived from non-AES sources (V, visual; A, auditory). These areas extend projections to integrating neurons in the SC. Right, Models of single units performing multisensory integration no longer seek to describe the responses of a “canonical” or “average” multisensory product calculated over a wide window of time but can successfully predict the responses of individual neurons at a millisecond-by-millisecond resolution. Illustrated is one such model in which excitatory visual and auditory streams (depicted at three time points) are integrated in real-time by a model neuron, which also receives input from inhibitory sources.

It is interesting to note in this context that the number of identified multisensory circuits is constantly increasing. It is becoming hard to find an area beyond the first synapse or two in an ascending pathway that does not have at least some multisensory inputs. As noted earlier, classically defined unisensory areas of cortex have been shown to have at least some multisensory neurons, and the border regions between the different “unisensory” areas are well populated with multisensory neurons (although their functional roles are not always well defined). This reveals how aggressively the brain combines information from its different senses, and is in alignment with a dominant modern perspective that the brain develops mechanisms that allow it to use all of the information available in a given context. It has also led some researchers to wonder whether the entire cortex (brain?) is multisensory. How many of these regions will prove to be primary sites of convergence (as in SC and MSTd), and how many will prove to be relays for multisensory computations that have occurred elsewhere, remains to be determined. This issue is likely to be clarified by current efforts to map the anatomical and functional connections between brain regions (the “connectome”).

These efforts should also provide the impetus to determine how different multisensory computations are linked to the perceptual and/or behavioral roles of different brain regions. This may prove to be a difficult task in regions in which neurons appear to yield multisensory products that are not always predictable (e.g., in PFC). In these cases, it is often unclear what role of the circuit can be facilitated by multisensory integration and/or whether a diverse set of integrated products is created to serve different output streams. These are areas of active exploration.

Computational modeling of multisensory integration

These conceptual frameworks for understanding the principles and circuit dynamics engaged in multisensory integration have benefited greatly from the introduction of theoretical and computational perspectives that began being discussed at SfN meetings in the early 2000s (Fig. 5).

Bayesian frameworks have been used to explore how multisensory integration can combine individual sensory estimates to

provide overall estimates that make best use of all available information. In a classic Bayesian model of sensory function, sensory estimates of environmental features are combined with prior knowledge to produce a “posterior” probability distribution from which the most likely value of the external feature can be inferred. In a multisensory model, signals obtained from different sensory sources are combined as independent indicators of the same feature. When the signals are congruent, this combination improves the accuracy of the inferred estimates (Fig. 5). These models have been applied to multiple functional domains, including visual-auditory, visual-haptic, and visual-vestibular, and have lent two important computational concepts to the discussion. The first is a standard for optimal information synthesis, which has added a robust quantitative standard for the evaluation of empirical data. Given a set of unisensory responses and assumptions about the functional goal of the structure, the Bayesian framework provides a prediction for the multisensory response under the assumption of optimal signal combination. Second, they introduced the concept of the prior distribution as a way of codifying the way in which multiple sensory inputs are bound and integrated together. This provides a framework for examining assumptions of common cause described above, and of examining how the weights/biases of different modalities are adjusted to represent their relative reliabilities within a particular context.

At the same time, neural network models have been developed at multiple levels of abstraction to examine the validity of posited verbal theories of circuit function. These have extended from early work in which single modality-specific units sent direct projections onto multisensory targets, to increasingly more elaborate architectures. In each case, efforts have been made to link the operation of these models to Bayesian frameworks and to focus on incorporating more extensive biological constraints to explain broader empirical findings. This has led to modern models that provide a moment-by-moment accounting of the multisensory transform, circuit models to explain normal and abnormal function and development, whole-brain models linking multisensory circuit computations to behavior, and sophisticated abstract models that distill essential biological computation

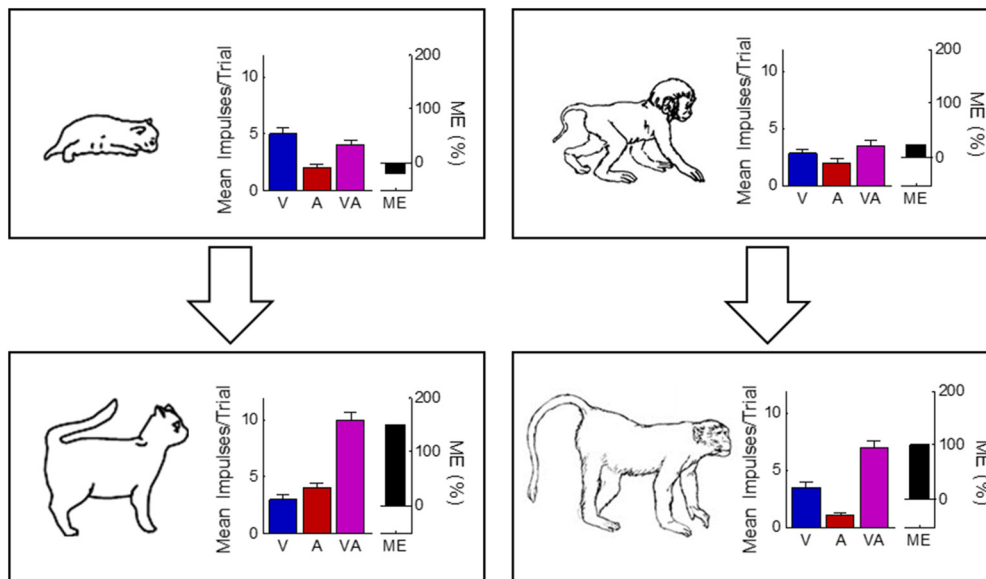


Figure 6. SC multisensory integration develops gradually in postnatal life. Top, Multisensory enhancement in SC neurons of cat and monkey are not present in early neonatal life. The visual-auditory (VA) response in the exemplar neurons is not significantly better than the V response. Bottom, Normal adult exemplars. Bar graph conventions are the same as Fig. 2.

and make them available for implementation in artificial sensor-fusion devices (for more discussion, see Trommershauser et al., 2011).

Throughout their emergence, these computational approaches have raised questions about how multisensory structures develop their ability to implement these computations. To answer these questions, they look to the data acquired in empirical studies.

Development of multisensory integration

Early empirical studies of how individual neurons develop their multisensory integration capabilities followed a trend common in studies of unisensory systems. Once the adult condition of feature selectivity was revealed, there was an immediate attempt to understand how it came to be. Was it an inherent capability of the brain that was either present at birth or elaborated soon after birth and independent of experience? Or, did its maturation depend on postnatal experience? These were general developmental themes in sensory neuroscience and have been the source of many of the presentations at SfN meetings through the years. They continue to be of great interest to this day.

The cat SC model was well suited to facilitate our understanding of visual, auditory, and somatosensory development. Much had already been learned about SC organization in the adult cat; and because this species is altricial, born with its eyes and ear canals closed, its visual, auditory, and somatosensory systems are functionally inactive or rudimentary at birth. Its protracted postnatal period of sensory development is ideal for assessing the maturational changes in these sensory representations. There are similar advantages in the increasingly popular rodent models that have also always been in heavy use for studies of sensory development, especially of the chemical senses. The advent of transgenic techniques in rodents offers a number of additional possibilities that remain to be fully used. However, it would also be helpful to expand the number of available biological models with which to examine issues of multisensory circuitry, development, and adaptation.

Nevertheless, the cat SC continues to be an effective model for understanding multisensory development. Such studies began in the 1970s and found that just before and after birth its neurons

were largely “silent”: few were spontaneously active, and even fewer were responsive to sensory stimulation. Those neurons that were responsive to external events were activated only by tactile stimuli, primarily by stimulation of the face, and appeared to be unisensory. Neurons responsive to auditory stimuli, most of which also seemed to be unisensory, appeared a number of days later (some somatosensory-auditory neurons were also noted at this time). It was not until after a considerable lag that neurons responsive to visual stimuli appeared and multisensory neurons became more numerous. This sensory chronology parallels the animal’s behavioral capabilities. But neonatal multisensory neurons are unable to integrate their multiple sensory inputs to produce enhanced responses. Instead, they appear to act as a common conduit for different senses to reach the same motor output systems. The lack of coordination between them can even produce competitive interactions.

Testing cats of different ages revealed that it took months of development for these multisensory neurons to begin showing multisensory integrative capabilities, a trend also observed in multisensory neurons in association cortex. This delay has also been observed in monkey SC (and later, in frog optic tectum). But unlike the altricial cat, the precocial monkey can already see and hear quite well at birth, and its SC already has multisensory neurons; yet, just like the neonatal cat, those neurons cannot integrate their converging inputs to enhance their responses (Fig. 6). They too lack experience with combinations of visual, auditory, and/or tactile cues. A similar development lag has been found in human subjects using behavioral/perceptual techniques. Yet, these studies also showed that human infants could judge whether cross-modal cues were related to the same object, and whether they were synchronous. It appears that at least some cross-modal comparisons are possible in the absence of substantial cross-modal experience, even if integration is not. Why this should be is not entirely clear (see Lewkowicz and Lickliter, 1994; Bremner et al., 2012).

The importance of experience with cross-modal events for normal multisensory maturation was made clear by sensory restriction experiments (for review, see Stein et al., 2014). When animals were deprived of visual-auditory experience by rearing

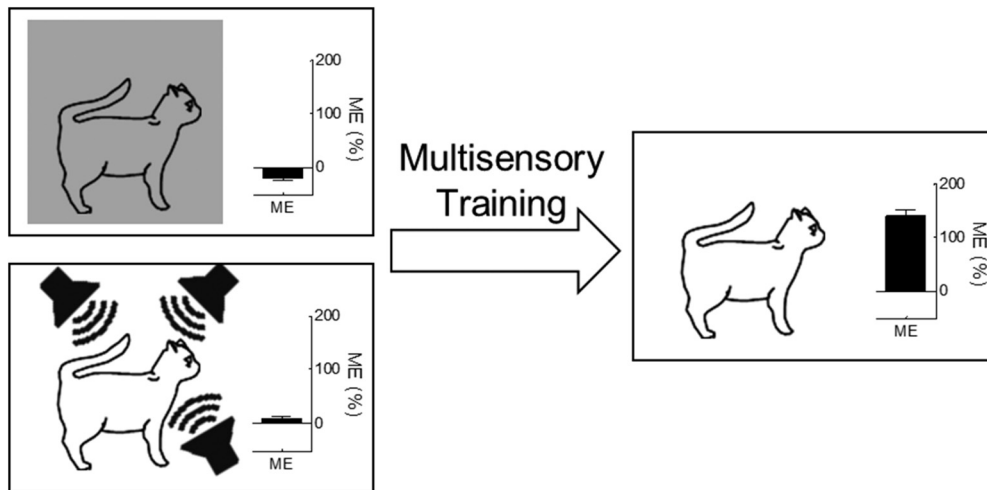


Figure 7. Lack of multisensory experience compromises multisensory development, but explicit training can compensate. Left, Experimental manipulations to preclude visual-auditory experience include dark-rearing (top) and rearing with masking noise (bottom). Both rearing conditions disrupt the development of the ability to integrate those cross-modal stimuli. Multisensory enhancement (ME) is not significantly above zero. Right, Explicit training with spatiotemporally concordant visual-auditory cues can mitigate these deficits, even in adulthood.

them in the dark, or with masking noise, or with randomly appearing visual and auditory cues, their SC neurons (and dark-reared multisensory neurons in AES) did not develop the ability to integrate visual-auditory cues. This was not because they had developed some general disability that rendered them incapable of integrating information from different senses. They were able to do so quite well when presented with cross-modal combinations that they had experienced (i.e., neurons in dark-reared animals were able to integrate auditory-somatosensory stimuli), and those in noise-reared animals were able to integrate visual-somatosensory stimuli. In both cases, the magnitude of the integrative products appeared to be much like that seen in normally reared animals.

There are a number of human conditions in which multisensory processing appears to be disrupted or is anomalous, such as autism, dyslexia, sensory processing disorder, schizophrenia, post-traumatic stress disorder, and traumatic brain injury. In some cases, for example in autism, these processing disruptions may represent a delay in normal development that, in some circumstances, will become ameliorated with age (Beker et al., 2018). Whether and how such defects may relate to the acquisition of multisensory experience remain to be determined.

Multisensory plasticity

Neural plasticity has been a major theme at SfN meetings throughout their 50 year history. One of the major tenets of sensory plasticity is that it degrades with age. This proved to be true for the ability of SC neurons to acquire multisensory integration capabilities, but it also appeared as if this degradation could be minimized in certain circumstances.

For example, SC neurons in cats reared without visual-auditory experience were initially unable to integrate visual-auditory cues, but were later able to develop this integrative capacity with little difficulty when given training sessions. These sessions involved repeated exposure to identical (i.e., invariant) pairs of visual-auditory stimuli (Fig. 7). The resultant integrative characteristics of these neurons were linked more closely to the spatial and temporal characteristics of these stimuli than expected based on studies of their normal counterparts. Of course, the latter had experienced combinations of cross-modal cues that varied in spatiotemporal concordance. This would likely pre-

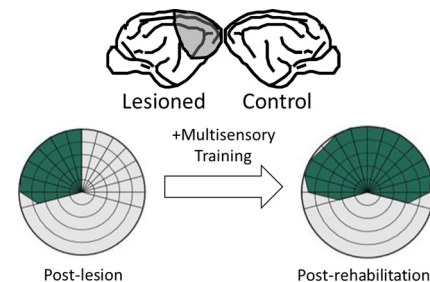


Figure 8. Multisensory training can restore visual function in cortically blinded animals. Lesions of all contiguous regions of visual cortex on one side (e.g., left) of the cat brain (shaded area on the schematic) result in complete blindness in contralateral (right) space. Visual responses are now restricted to left visual space (the proportion of correct responses at each location is shown in green on the polar plots). But after repeated exposure to congruent visual-auditory stimuli in the blinded visual field, vision is restored there.

clude them from developing the narrower spatial or temporal focus of animals whose only visual-auditory experience was with the invariant training stimulus configuration.

The focused training experience also greatly facilitated the speed of acquiring these capabilities, producing results within weeks. In contrast, when adult animals lacking integration capabilities were simply placed into a normal housing environment where cross-modal events abound, they took years to develop their multisensory integration capacity. Presumably, the spatial and temporal variation in natural cross-modal events also impedes the older brain's ability to extract the spatiotemporal relationships it needs to develop multisensory integration capabilities.

Nevertheless, the observed training effects suggest that neurons can adapt their responses to the feature-specific characteristics of the cross-modal stimuli they experience. Further evidence for such multisensory plasticity has been documented in the rodent gustatory cortex. The degree to which multisensory plasticity can be engaged in different animals, different systems, and at different ages has yet to be determined, and it is already apparent that there are constraints on this plasticity, even during early development. Rearing animals with only spatially disparate visual-auditory cues impacted only a minority of SC neurons. These neurons developed poor visual-auditory receptive field overlap and showed response enhancement to spatially disparate

visual-auditory cues as might be expected by their receptive field disparities. But the vast majority of neurons appeared to be unaffected by the rearing condition. They looked very much like those found in dark-reared animals. Apparently, there is an inherent bias for spatially aligned cues in the SC that constrains the effectiveness of different multisensory experiences.

Trends observed in animals deprived of visual-auditory experience have also been seen in human patients who had similar restrictions due to congenital cataracts or deafness. Following removal of the congenital cataracts, or the introduction of cochlear implants (especially when done early), these individuals developed the ability to integrate visual-auditory cues; albeit the time frame in which they first showed such capabilities is not entirely clear, and they may always have deficits, especially with complex events (e.g., speech) (Stevenson et al., 2017). Whether more rapid and more pronounced enhancements in performance would be induced with a multisensory training paradigm remains to be determined. But it seems like a reasonable expectation.

The effectiveness and benefits associated with the plasticity induced by multisensory training have also been demonstrated on memory-related tasks. For example, combinations of cross-modal stimuli are more reliably recalled after multisensory training than are their modality-specific components after unisensory training. And even the unisensory components of the cross-modal training stimulus are more readily recalled after multisensory training than if they were originally encoded in a modality-specific training context.

This multisensory-unisensory transfer effect is also evident in results from studies of how the perception of cross-modal synchrony can change with training. Repeated presentation of asynchronous cross-modal cues makes them seem more synchronous, and training with synchronous cross-modal cues sharpens one's ability to detect even minor variations from their synchrony. Training with cross-modal cues also produces a number of robust perceptual illusions, one of the best known of which is the "ventriloquism aftereffect." In this case, repeated exposure to spatially disparate visual-auditory cues produces a comparatively long-lasting bias in localizing auditory cues; and indeed, the changes supporting this bias can take place very rapidly.

The ability of multisensory processing to enhance or otherwise modulate the salience of events through noninvasive means has opened the door to interesting translational strategies. One of the more dramatic of these is the restoration of sight after blindness created by extensive damage to visual cortex. Cats, like humans, are rendered blind by these cortical lesions, and the animals have been shown to lose visual responses in their multisensory SC neurons. Repeatedly presenting visual-auditory cues in the blinded hemifield of both species restores their vision (Fig. 8); and, in cat, this has been shown to be accompanied by the restoration of visual responses in SC multisensory neurons. Whether these training techniques can ameliorate the multisensory integration deficits of other developmental anomalies, such as autism spectrum disorder, sensory processing disorder, dyslexia, etc., remains to be determined. But there is reason to be hopeful.

Where are we now?

The neuroscientific study of multisensory processes now uses the latest investigative techniques and a rich variety of model species. It uses the entire panoply of behavioral, physiological, anatomical, and computational approaches; and enjoys an active collaboration among researchers with very different expertise. It has begun to look very much like other sensory research endeavors that predate it. This is a good thing.

SfN continues to play a major role in disseminating multisensory research findings. Despite the formation of a specialty society (the International Multisensory Research Forum), which also has annual meetings, and the advent of multisensory sessions at other specialty meetings (e.g., the Visual Sciences Society), SfN remains a primary meeting for multisensory research. This is because its annual meeting attracts investigators who deal with all the different individual senses, thereby providing a convenient opportunity for them to find common areas of interest with those studying how those senses interact. It is also the best place to learn about technical innovations that can dramatically alter approaches to sensory research, be it unisensory or multisensory. It is clear that over the years SfN has played a critical role in creating the current appreciation among neuroscientists of the impact of multisensory integration on normal perception.

So, it is interesting to note that this knowledge is often underrepresented in graduate training and in the training of medical professionals. Although the number of recognized senses has increased, most textbooks continue to divide sensory function by modality without the counterbalance of how they interact and change as a result of that interaction. From this perspective, the field has not yet fully matured, and there exists an odd disconnect between what is generally recognized and what is generally communicated to students. Change is often slow, but the future of multisensory research is bright. It promises to provide answers to fundamental questions of sensory function, perception, cognition, and decision making, and to provide new strategies to ameliorate sensory dysfunction. It will be interesting to see the changes in this field that will be summarized in the Journal 50 years from now.

References

- Angelaki DE, Gu Y, DeAngelis GC (2009) Multisensory integration: psychophysics, neurophysiology, and computation. *Curr Opin Neurobiol* 19:452–458.
- Beker S, Foxe JJ, Molholm S (2018) Ripe for solution: delayed development of multisensory processing in autism and its remediation. *Neurosci Biobehav Rev* 84:182–192.
- Bremner AJ, Lewkowicz DJ, Spence C (eds) (2012) *Multisensory development*, Ed 1. Oxford: Oxford UP.
- Bruno N, Pavani F (2018) *Perception: a multisensory perspective*. Oxford: Oxford UP.
- Calvert G, Spence C, Stein BE (eds) (2004) *Handbook of multisensory processes*. Cambridge, MA: Massachusetts Institute of Technology.
- Lewkowicz DJ, Lickliter R (eds) (1994) *The development of intersensory perception: comparative perspectives*. Hillsdale, NJ: LEA.
- Marks I (1978) *The unity of the senses: interrelationships among the modalities*. San Diego: Academic.
- Murray MM, Wallace MT (eds) (2012) *The neural bases of multisensory processes*. Boca Raton, FL: CRC.
- Naumer MJ, Kaiser J (eds) (2010) *Multisensory object perception in the primate brain*. New York: Springer.
- Spence C, Driver J (eds) (2004) *Crossmodal space and crossmodal attention*. Oxford: Oxford UP.
- Stein BE (2012) *The new handbook of multisensory processing*. Cambridge, MA: Massachusetts Institute of Technology.
- Stein BE, Stanford TR, Rowland BA (2014) Development of multisensory integration from the perspective of the individual neuron. *Nat Rev Neurosci* 15:520–535.
- Stein BE, Meredith MA (1993) *The merging of the senses*. Cambridge, MA: Massachusetts Institute of Technology.
- Stevenson RA, Sheffield SW, Butera IM, Gifford RH, Wallace MT (2017) Multisensory integration in cochlear implant recipients. *Ear Hear* 38:521–538.
- Trommershauser J, Kording K, Landy MS (eds) (2011) *Sensory cue integration*. Oxford: Oxford UP.
- Walk RD, Pick HL Jr (eds) (1981) *Intersensory perception and sensory integration*. New York: Plenum.