

Identifying and Quantifying Multisensory Integration: A Tutorial Review

Ryan A. Stevenson · Dipanwita Ghose · Juliane Krueger Fister ·
Diana K. Sarko · Nicholas A. Altieri · Aaron R. Nidiffer ·
LeAnne R. Kurela · Justin K. Siemann · Thomas W. James ·
Mark T. Wallace

Received: 10 July 2013/Accepted: 26 March 2014/Published online: 11 April 2014
© Springer Science+Business Media New York 2014

Abstract We process information from the world through multiple senses, and the brain must decide what information belongs together and what information should be segregated. One challenge in studying such multisensory integration is how to quantify the multisensory interactions, a challenge that is amplified by the host of methods that are now used to measure neural, behavioral, and perceptual responses. Many of the measures that have been developed to quantify multisensory integration (and which have been derived from single unit analyses), have been applied to these different measures without much consideration for the

nature of the process being studied. Here, we provide a review focused on the means with which experimenters quantify multisensory processes and integration across a range of commonly used experimental methodologies. We emphasize the most commonly employed measures, including single- and multiunit responses, local field potentials, functional magnetic resonance imaging, and electroencephalography, along with behavioral measures of detection, accuracy, and response times. In each section, we will discuss the different metrics commonly used to quantify multisensory interactions, including the rationale for their

R. A. Stevenson
Department of Psychology, University of Toronto, Toronto, ON, Canada

R. A. Stevenson · D. K. Sarko · A. R. Nidiffer · M. T. Wallace
Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, USA

R. A. Stevenson · J. K. Fister · L. R. Kurela ·
J. K. Siemann · M. T. Wallace
Vanderbilt Brain Institute, Nashville, TN, USA

R. A. Stevenson · J. K. Siemann · M. T. Wallace
Vanderbilt Kennedy Center, Nashville, TN, USA

R. A. Stevenson (✉)
Rm 523, Sidney Smith Bldg, 100 Saint George St., Toronto, ON, Canada
e-mail: ryan.andrew.stevenson@gmail.com

D. Ghose · M. T. Wallace
Department of Psychology, Vanderbilt University, Nashville, TN, USA

J. K. Fister · L. R. Kurela · J. K. Siemann
Neuroscience Graduate Program, Vanderbilt University Medical Center, Nashville, TN, USA

J. K. Fister · L. R. Kurela · J. K. Siemann
Clinical Neuroscience Scholars Program, Vanderbilt University Medical Center, Nashville, TN, USA

D. K. Sarko
Department of Anatomy, Cell Biology and Physiology, Edward Via College of Osteopathic Medicine, Spartanburg, SC, USA

N. A. Altieri
Department of Communication Sciences and Disorders, Idaho State University, Pocatello, ID, USA

J. K. Siemann
Silvio O. Conte Center for Neuroscience Research, Vanderbilt University Medical Center, Nashville, TN, USA

T. W. James
Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

M. T. Wallace
Department of Psychiatry, Vanderbilt University, Nashville, TN, USA

use, their advantages, and the drawbacks and caveats associated with them. Also discussed are possible alternatives to the most commonly used metrics.

Keywords Audiovisual · fMRI · EEG · Response time · Accuracy · Sensory processing

Historically, scientific research investigating the means with which humans transduce sensory information and subsequently filter, manipulate, and process that information has been structured to look at each sensory modality in isolation. While this singular focus has yielded an extraordinary amount of information about how our individual sensory systems function, in the natural environment it is rare to perceive an external event through only one sensory modality. Instead, information from the world converges on us via multiple sensory systems. As such, the brain must decide what information belongs together and what information should be segregated. Integrating information across the senses improves a wide range of behavioral outcomes, including detection (Lovelace et al. 2003; Stein and Wallace 1996), localization (Nelson et al. 1998; Wilkinson et al. 1996), and speed of response (Hershenson 1962; Diederich and Colonius 2004).

In the last few decades, our views concerning sensory processing have been revolutionized to now consider this from the perspective of a highly interactive, multisensory network of closely interrelated functional brain regions and mechanisms. Associated with the emergence of this field of inquiry have come a series of challenges, most notably the need to deliver stimuli that are carefully controlled in dimensions that are particularly relevant to multisensory integration (e.g., timing, spatial location, and effectiveness). An additional challenge to the field, and the one that will serve as the focus of this review, is how to quantify the multisensory interactions that occur upon the presentation of cues from multiple sensory modalities. This challenge is amplified by the host of methods that are now used to measure neural, behavioral, and perceptual responses. These range from those that index synaptic processes such as the local field potential (LFP) to those that measure the activity of large neuronal ensembles, such as the blood oxygenation level dependent (BOLD) signal that underpins most functional magnetic resonance imaging (fMRI) studies, to indices of behavioral and perceptual responses such as accuracy and response times (RTs). Perhaps most importantly, many of the measures that have been developed to quantify multisensory integration (and which have been derived from single unit analyses), have been applied to these different measures without much consideration for the nature of the process being studied. For example,

measurements from large neuronal populations (such as what is done with fMRI) must take into account that the metric reflects activity of a variety of types of neurons with very different functional properties, a feature that does not apply with single unit recordings.

Here, we attempt to provide a review focused on the means by which experimenters quantify multisensory processes and integration across a range of commonly used experimental methodologies. The emphasis will be on the most commonly employed measures, including single- and multiunit responses, LFP, fMRI, electroencephalography (EEG), along with behavioral measures of detection, accuracy, and RTs. In each of these sections, we will discuss the different metrics commonly used to quantify multisensory interactions, including the rationale for their use, their advantages, and the drawbacks and caveats associated with them. Also discussed will be possible alternatives to the most commonly used metrics. Table 1 presents a collection of the most common of these metrics, their uses, and their issues. For simplicity and illustrative purposes, we will focus our discussion on combinations of visual and auditory stimuli, but it should be noted that these measures and their associated advantages and disadvantages apply to the characterization and quantification of responses across any combination of sensory modalities.

Multisensory Integration in Single Cells

The earliest neuronal studies of multisensory integration in mammals focused on a midbrain structure, the superior colliculus (SC; Meredith and Stein 1983). The SC's well-established spatiotopic organization and responsiveness to auditory, visual, and somatosensory modalities, along with its high incidence of multisensory neurons, provided a functional framework to describe the basic response properties, integrative features, and mechanisms by which multisensory neurons integrate stimuli across sensory modalities (Meredith and Stein 1983, 1985, 1986a; Meredith et al. 1987). Rather than being specific to the SC, the characteristics of these multisensory neurons and their integrative features appear to generalize to the majority of the species and brain structures that have been examined to date (Stein and Wallace 1996; Wallace et al. 1992).

In these foundational studies of multisensory processing, multisensory integration has been defined according to the number of action potentials produced in response to a given stimulus. Spike counts that significantly differ between combined modality (e.g., visual-auditory) presentations and the most effective unisensory presentations have been used as the criterion to identify if the neuron is integrating

Table 1 Measures of Multisensory Integration

Methodology	Criteria	Equation	Uses	Issues
Single Unit Recordings	<i>ii</i>	$i = \frac{AV_{\max}(A,V)}{\max(A,V)} \times 100$	This metric assesses the gain attributable to having information from a second sensory modality available. Multisensory response is compared to strongest unisensory response	The weaker modality effects are not taken into account
Event-related potentials		$msc = \frac{\sum^n [AV_i - (A_i + V_i)]}{n}$	Incorporates both component unisensory responses in evaluating integration effects, thus basing multisensory gain (or loss) on the complete set of inputs effects of different modalities	Inensitive to possible subthreshold
BOLD fMRI	Additivity	$ERP_{AV} \neq ERP_A + ERP_V$	Based on the law of superposition, any significant deviation from the summed unisensory prediction is evidence of a multisensory interaction	Additivity cannot account for CA and may be limited to use in early responses (<200 ms)
	Additive Factors	$AV_H - AV_L \neq (A_H - A_L) + (V_H - V_L)$ $BOLD_{AV} > \max[BOLD_A, BOLD_V]$	Minimizes the impact of common activity, allowing for identification of later interactions Can be used to detect if information is being processed from both modalities. Best when used with tasks that are already known to induce multisensory integration	Additive factors multiplies the number of trials needed and decreases experimental effect sizes Cannot disambiguate if a region (a) processes both auditory and visual stimuli independently or (b) if it actively integrates them
	Maximum Unisensory		Can be used to identify active integration within a given voxel or region	Highly conservative and rarely observed, even in regions known to be involved in multisensory integration
	Superadditivity	$BOLD_{AV} > BOLD_A + BOLD_V$	Can be used to identify active integration within a given voxel or region without using the difficult to obtain superadditivity	Additive factors multiplies the number of trials needed and decreases experimental effect sizes
Accuracy and detection	Additive Factors	$AV_H - AV_L \neq (A_H - A_L) + (V_H - V_L)$	Can be used to detect if information is being used from both modalities. Best when used with tasks that are already known to induce multisensory integration	Implies that information is being used from both sensory modalities, but not necessarily that the information is being integrated, as it fails to account for statistical facilitation
	Maximum unisensory	$\hat{p}(AV) > \max[p(A), p(V)]$	Can identify active integration across sensory modalities, and accounts for statistical facilitation	This measure is conservative, assuming <i>optimal</i> statistical facilitation as the null hypothesis
	Probability Summation	$\hat{p}(AV) > p(A) + p(V) - p(A) \cdot p(V)$ $RT_{AV} > \max[RT_A, RT_V]$	Can be used to detect if information is being processed from both modalities. Best when used with tasks that are already known to induce multisensory integration	Mean response times collapse across the entire RT distribution, making this a very insensitive measure that often fails to detect interactions seen with the race model. Does not account for statistical facilitation
Response times	Fastest unisensory			Does not incorporate accuracy. Therefore, this method is subject to speed-accuracy tradeoffs.
	Race model predictions	$F_{AV}(t) = F_A(t) + F_V(t) - F_A(t) \cdot F_V(t) < F_A(t) + F_V(t)$	Uses the distribution of response times to identify multisensory interactions. Can identify where in the distribution and interaction occurred, and accounts for statistical facilitation	May be difficult to distinguish subtle differences from race model predictions (see capacity measure).
	Additive factors	$(A_H V_H - A_H V_L) - (A_L V_H - A_L V_L)$	The double difference can be used to help diagnose decision rule and architecture	Additive factors multiplies the number of trials needed and decreases experimental effect sizes
	Capacity	$C(t) = \frac{H_{AV}(t)}{H_A(t) + H_V(t)}$	Similar to Race model predictions. However, capacity uses integrated hazard functions, and compares processing to the benchmark of “1”, making it simpler to distinguish between unlimited, versus limited or super-capacity	Does not incorporate accuracy. May suffer from speed-accuracy tradeoffs similar to race model predictions

information across the two senses. Thus, the most effective unisensory response serves as the reference, and the metric assesses the gain attributable to having information from a second sensory modality available. In this manner, multisensory interactions have been quantified using the *Interactive Index* (*ii*) (Meredith and Stein 1983, 1986b).

$$ii = \frac{AV - \max(A, V)}{\max(A, V)} \times 100, \quad (1)$$

where the variables *AV*, *V*, and *A* represent spike count measures (e.g., mean spikes/trial, mean firing rate, etc.) for each stimulus condition after spontaneous activity has been removed (Fig. 1). While any significant non-zero value is indicative of a multisensory interaction, this measure can also be used to characterize the magnitude of integration or gain. A significant increase in firing rate compared to the most effective unisensory stimulus results in a positive *ii* and indicates response enhancement, whereas a significant decrease in firing rate results in a negative *ii*, and is indicative of response depression (Meredith and Stein 1983; Meredith et al. 1987; Meredith and Stein 1986b; Wallace et al. 1996, 1998; Jiang et al. 2002; Burnett et al. 2004). It must be emphasized that the interactive index is a proportionate measure that is scaled relative to the magnitude of the strongest unisensory response.

While a difference from the most effective unisensory response has been the traditional means used to identify and quantify a multisensory interaction, the *ii* is not the only way to characterize these interactions. Single unit spiking data can also be used to compare multisensory to unisensory responses using an additive model (i.e., the linear summation of both unisensory responses; Fig. 1). This differs from the *ii* in using the sum of each of the unisensory responses as the reference, and is referred to as the mean statistical contrast (*msc*):

$$msc = \frac{\sum_{i=1}^n [AV_i - (A_i + V_i)]}{n} \quad (2)$$

where *n* is the total number of stimulus presentations in each modality. This model assumes independence between inputs from each sensory modality and distinguishes between *subadditive* (*msc* < 0), *additive* (*msc* = 0), and *superadditive* (*msc* > 0) response modes (Perrault et al. 2003, 2005; Stanford et al. 2005). This characterization confers the advantage of incorporating both component unisensory responses in evaluating integration effects, thus basing multisensory gain (or loss) on the complete set of inputs. It should be noted here that the *ii* sets a threshold that defines the existence of multisensory interaction, and can be used to describe such an interaction, and provides an important view into the gain of information achieved by having information available from a second sensory modality. In contrast, the *msc* provides a more informative

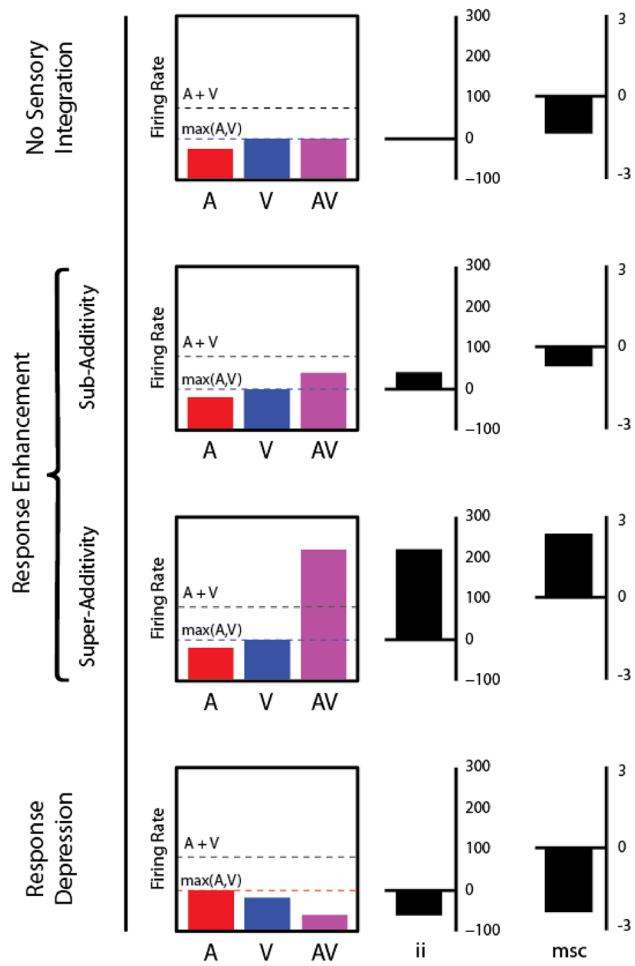


Fig. 1 Indexing multisensory integration in single neurons. Single neurons that respond to both auditory and visual stimuli can be characterized by two different quantifications, the *interactive index* or *mean statistical contrast*

view into the computation performed by the neuron under study. Thus, neurons exhibiting statistically significant positive *ii* values indicate response enhancement which can manifest as superadditive, additive or subadditive in nature. However, when the neurons exhibit statistically significant negative *ii* values, it is indicative of response depression which is always subadditive.

The nature of the multisensory interactions demonstrable in single neurons, as defined using either *ii* or *msc*, are not fixed values, but rather are dependent upon a wide array of factors. Changes in these metrics can, in addition to identifying and quantifying integration, be applied to measure the manner in which sensory integration varies with changes in any given factor. Such applications have shown that receptive field characteristics of multisensory neurons strongly influence the nature and magnitude of their multisensory integration (Meredith and Stein 1986b, a; Meredith et al. 1987). The most important and best

characterized of these appear to be the spatiotemporal structure of neuronal receptive fields, the spatial and temporal relationship between the presented multisensory stimuli, and the effectiveness of the responses within the large and heterogeneous receptive fields. More specifically, if a multisensory (e.g., visual-auditory) stimulus combination is presented such that the visual stimulus is within its visual receptive field and the auditory stimulus is within its auditory receptive field, it will most likely produce response enhancement, even when the stimuli are not found at the exact same spatial location. However, if the visual stimulus is within its receptive field while the auditory stimulus is outside its receptive field, the likely result is a response depression. Similarly, the temporal processing of multisensory information was revealed in a study that parametrically manipulated the temporal relationship between paired multisensory stimuli and examined the consequent impact on the firing profiles of multisensory neurons. In these experiments, it was found that the largest response gains were seen upon multisensory combinations when the peak discharge periods of the individual sensory responses overlapped. As the stimulus onset asynchronies (SOAs) were shifted to larger asynchronies and the peak discharge periods became less overlapping, the magnitude of the enhancement was generally found to decline. In fact, if the temporal disparity between the stimuli was sufficiently large, response enhancement could transition to response depression. (Meredith et al. 1987; Meredith and Stein 1986a). That both spatial and temporal factors play a deterministic role in multisensory integration is not surprising, given that multisensory stimuli related to the same object or event in a real world setting will be spatially and temporally proximate. Hence, the spatiotemporal statistics of a multisensory stimulus complex and the relationship of these stimuli to the receptive field organization of multisensory neurons play key roles in their integration or “binding.” The third major factor affecting multisensory integration is the efficacy of the component stimuli within the neuronal receptive fields. In this context, the weaker the component unisensory stimuli in eliciting a response, the larger the response to multisensory presentations relative to the response to unisensory presentations, reflecting a larger magnitude of multisensory integration (Meredith and Stein 1986b). Conversely, as the effectiveness of the individual stimuli increased, the amount of multisensory integration decreased. Hence, minimally effective visual and auditory stimulus combinations produced the largest response enhancements, whereas strongly effective stimulus combinations produced weaker (or even non-existent) enhancements. Collectively, this third principle has been referred to as inverse effectiveness. Though these factors affecting the nature and magnitude of multisensory integration were initially described in multisensory neurons of

the cat SC, later studies have found similar results in multisensory neurons in a variety of brain structures and species (Kayser et al. 2005, 2007; Bizley and King 2008; Lakatos et al. 2007; Ghazanfar et al. 2005).

The consequences of these factors on the integrated product should play an important role in the design of multisensory experiments. For example, when designing an experimental paradigm aimed at studying multisensory enhancement, experimenters can maximize the chances of achieving such enhancement by using stimuli that are spatially proximate, temporally synchronous, and weakly effective. Conversely, if the aim is to study multisensory depression, stimulus pairs that are spatially disparate, temporally asynchronous, and highly effective should be used.

Other alternative measures to *ii* and *msc* are available to characterize and quantify multisensory interactions in the spiking domain. For example, one can measure response duration, response latency, and peak discharge frequency (or peak firing rate, measured from the time bin in which the greatest number of spikes occurred as the average number of spikes per trial within that time bin; Meredith et al. 1987; Royal et al. 2009). Many of these measures are extremely valuable in providing important insights into temporal response dynamics, and may reveal encoding strategies beyond simple spike rate changes for multisensory stimuli.

Finally, it should be noted that the majority of single-unit studies focus on characterizing the responses of overt multisensory neurons (i.e., neurons that respond to both auditory and visual stimuli presented in isolation). More recent studies have begun to analyze modulatory multisensory neurons, or neurons that only respond to stimulus presentations in a single modality, but that exhibit response changes with the added presence of a second stimulus modality (Allman et al. 2008, 2009; Allman and Meredith 2007; Carriere et al. 2007; Kayser et al. 2008; Lakatos et al. 2007). While these neurons qualitatively differ from overt multisensory neurons, the operational principles outlined above appear to be very similar.

Multisensory Integration in Neural Populations

Local Field Potentials, Current Source Density and Cross-Correlational Analyses

Though activity at the level of the single neuron has been one of the most commonly implemented approaches to the study of sensory and multisensory processing, spiking activity of single neurons fails to capture subthreshold processing at the synaptic level. This has led to an upsurge in the study of LFPs, which provide a measure of synaptic

processing and local input architecture to a specific brain region. The LFP provides a measure of pooled voltage changes surrounding the electrode tip (Mitzdorf 1987; Buzsaki and Draguhn 2004; Kajikawa and Schroeder 2011), relates well to the BOLD signal recorded in fMRI (Logothetis et al. 2001; Goense and Logothetis 2008), and has gained popularity as a candidate signal for neural prostheses because it is easier to collect than spikes and is more tolerant to small changes in electrode position (Andersen et al. 2004a; Pesaran et al. 2006). The LFP signal has two major components: a high frequency component (0.5–3 kHz) believed to largely reflect local spiking activity within a spatial scale of 150–300 μm (Logothetis 2008; Gray et al. 1995; Henze et al. 2000), and a low frequency component (<200 Hz) believed to be derived largely from postsynaptic potential changes surrounding the electrode tip (Logothetis 2008, 2003; Berens et al. 2008a). The LFP signal can be further decomposed into its component frequency bands similar to EEG waves known as *delta* (1–4 Hz), *theta* (4–8 Hz), *alpha* (8–12 Hz), *beta* (12–30 Hz), and *gamma* (>30 Hz). The power within these different frequency bands has been related to different aspects of sensory processing including stimulus selectivity (Berens et al. 2008a, b), attentional allocation (Fries et al. 2001; Taylor et al. 2005), and object category selectivity (Kreiman et al. 2006).

Because of its ability to provide complementary information to that seen via neuronal spiking, the LFP is becoming an increasingly important tool for the study of multisensory processes. This has stemmed from recent observations of sub-threshold influences of stimuli from the non-driving or non-dominant modality on sensory processing in brain areas traditionally considered sensory specific—including auditory and visual cortices (Allman et al. 2008; Lakatos et al. 2007). Kayser and colleagues reported visual modulation of activity in auditory cortex of rhesus monkeys, as demonstrated by changes in amplitude (enhancement and suppression) under cross modal conditions (Kayser et al. 2008). Multisensory integration of faces and voices using LFP and single unit activity in the primate auditory cortex has been demonstrated as well (Ghazanfar et al. 2005). Study of LFP signals in conjunction with information theoretic approaches have demonstrated that, in visual cortex, the timing of spikes relative to the phase of LFP oscillations in the delta band carries information about the contents of naturalistic images (Montemurro et al. 2008). Similarly in auditory cortex, the timing of spikes with respect to LFP phase in the theta frequency range carries information about the type of sound presented to awake and passively listening monkeys (Kayser et al. 2009). These studies provide evidence that the information encoded by the “phase of firing” (revealed by LFP signals) is complementary to spike rate code, thus enhancing the

sensory acuity of spike code through global network activation (Kayser et al. 2009, 2008; King and Walker 2012; Ghazanfar et al. 2005; Montemurro et al. 2008).

Though LFPs can convey additional information relevant for subthreshold processing, the inherent nature of LFP signals raises significant concerns about the extent of spatial spread due to volume conduction. The extent of this spatial spread is still controversial and affected by numerous factors but the upper range has been estimated to be 500–800 μm (Liu and Newsome 2006; Xing et al. 2009; Katzner et al. 2009; Berens et al. 2013). Thus, this complicates the spatial interpretation of LFP recordings since it can comprise of local as well as distant electrophysiological events. This caveat can be easily resolved especially in cortical recordings by using multicontact electrodes that span the entire cortical thickness and can be used to provide a laminar analysis of LFP signals. The spatial derivative of this LFP signal can be used to reduce the spatial spread considerably such that only local events are indexed. The second order spatial derivative of the LFP signal is termed the current source density (CSD) (Mitzdorf 1985; Pettersen et al. 2008). This provides a measure of localized current flow across the cell membrane, which is measured as sources and sinks and helps to index sub-threshold synaptic processing in a layer specific manner in various cortical areas and also has much less spatial spread than LFP signals (Kajikawa and Schroeder 2011; Lakatos et al. 2007; Einevoll et al. 2013).

One example of the application and power of CSD analyses is to the auditory cortex of non-human primates. In this study, Schroeder and colleagues showed that somatosensory stimulation could modulate ongoing oscillations in primary auditory cortex by resetting their phase (Lakatos et al. 2007) (Fig. 2). Such phase resetting could represent a powerful mechanism by which somatosensory inputs could prime the auditory system to amplify the incoming auditory signal. Indeed, the concept of phase resetting is likely not unique to somatosensory-auditory interactions. Much interest of late is focused on the concept that visual inputs may also have the capacity to cause phase resets in auditory cortex, a possible mechanistic basis for the gains in speech intelligibility seen when the mouth of a speaker is viewed (Zion Golumbic et al. 2012, 2013b, 2013a).

In addition to the use of LFP and CSD analyses to evaluate multisensory-mediated activity changes at the population level, other population-based measures have focused on coincident spiking activity and the role that it may play in neural coding. One of the most common tools for measuring such correlations is the cross-correlogram (CCG), which correlates activity between neuronal pairs and depicts changes in the probability of a target neuron’s discharge relative to the discharge timing of a reference

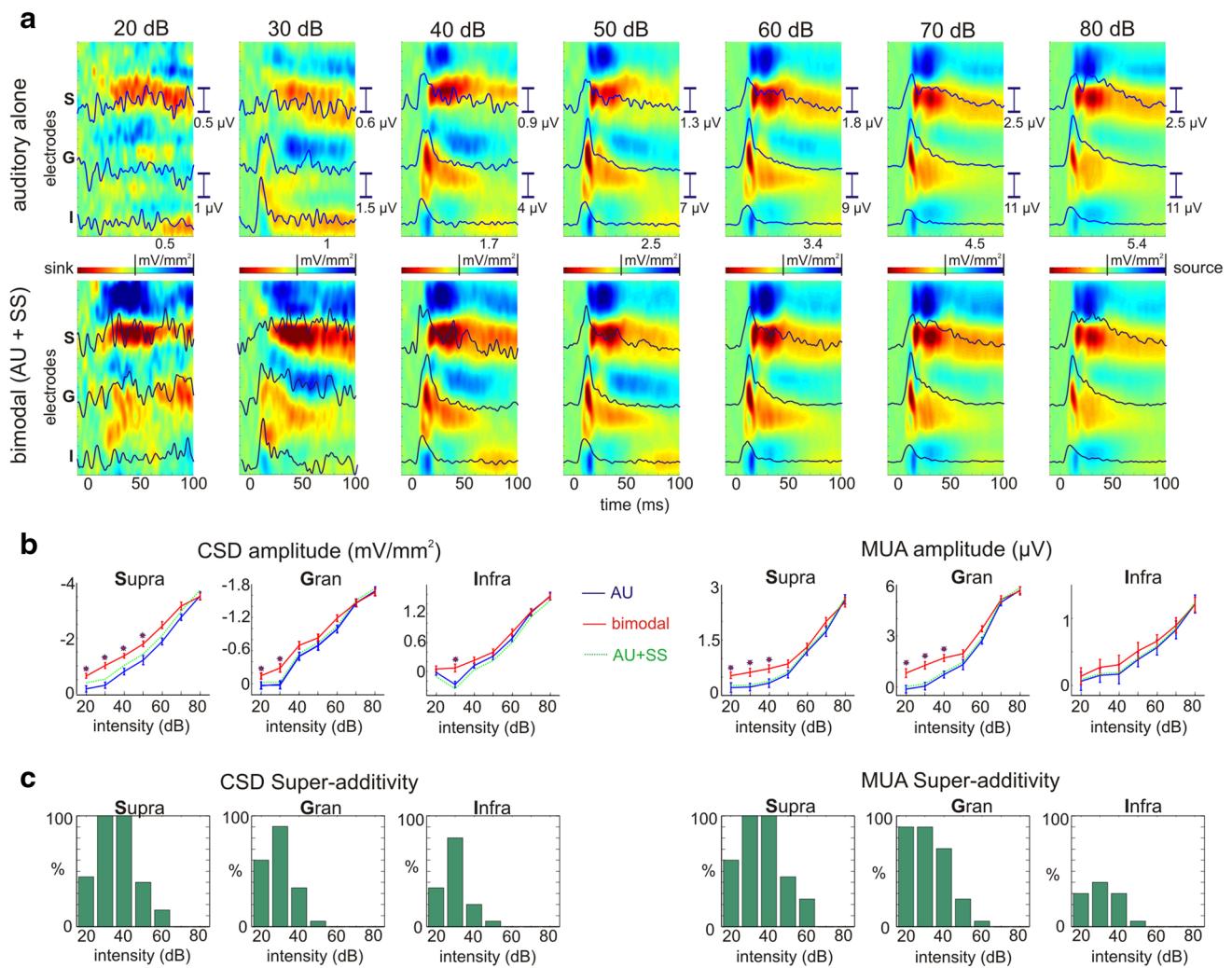


Fig. 2 Current source density. Color plots depicting laminar profile of CSD responses from Lakatos et al., 2007, under auditory (upper) and bimodal (lower) conditions at different auditory stimulus intensities are shown in Panel A. The overlaid traces show MUA in the selected supragranular (S), granular (G) and infragranular (I) channels. Panel B shows single trial CSD and MUA amplitudes on selected S, G and I channels averaged across the time interval of

15–60 ms. Standard errors represented as error bars and asterisks denote bimodal response amplitudes that are significantly higher than the sum of unimodal responses (one sample t tests, $p < 0.01$). Panel C shows the percentage of experiments ($n = 20$) at each auditory intensity where single trial CSD and MUA bimodal responses were significantly larger than sum of unimodal responses in each layer

neuron. When a pair of neurons fire synchronously, it appears as peaks or valleys in the CCG and indicate excitatory or inhibitory interactions, respectively (Perkel et al. 1967; Vroomen and Baart 2009). More recently, correlated activity has been studied through the use of the joint peristimulus time histogram, due to its enhanced temporal resolution and facilitated observation of spike coincidence over time after a stimulus or behavioral event (relative to the CCG; Gerstein et al. 1989). Though spike synchrony has been most widely studied in the visual system (Singer 1993; Usrey and Reid 1999; Jermakowicz and Casagrande 2007; Brecht et al. 1999), it has also been used within the auditory (Ahissar et al. 1992) and somatosensory systems (Merzenich et al. 1987; Jenkins

et al. 1990). Somewhat surprisingly, it has yet to be applied to multisensory systems, where it will likely reveal how coordinated activity across multisensory neurons (and between multisensory and unisensory neurons) contributes to the encoding of a multisensory stimulus complex.

The advantages of these population-based measures, when contrasted with spiking data, are that they provide greater insight into how larger neural ensembles may encode multisensory objects and events, and reveal facets of the distributed code that is not readily evident from single (or multi) units. However, the disadvantages of these methods are closely tied to these advantages, in that population-based measures have a host of interpretational caveats associated with them due to the heterogeneous

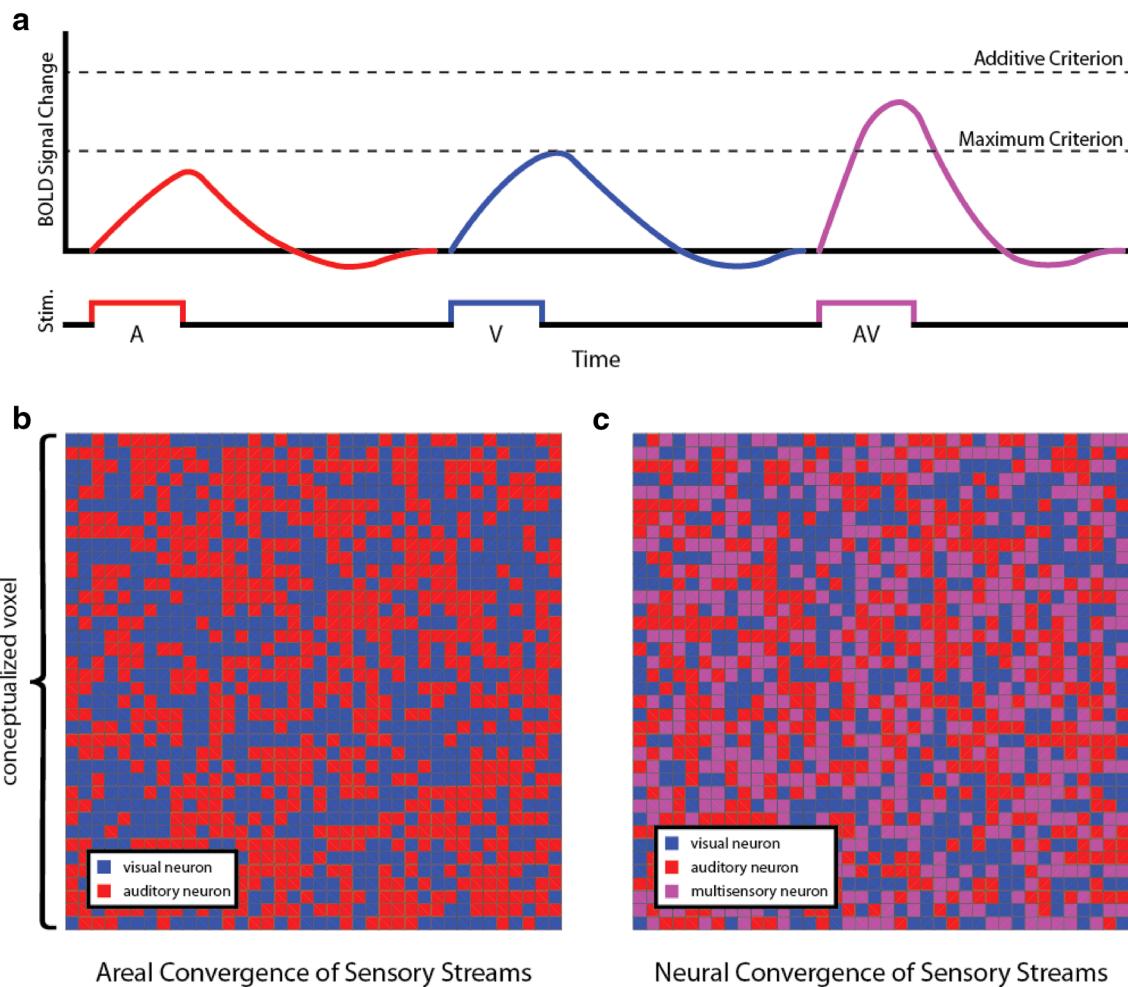


Fig. 3 Multisensory enhancement in BOLD fMRI. BOLD responses are the product of a population of neurons which may include only unisensory neurons (**b**; areal convergence) or may include

multisensory neurons (**c**; neuronal convergence). These heterogeneous populations must be taken into account when considering how to index multisensory integration in the BOLD signal (**a**)

nature of the elements contributing to these signals. Rather than describe these limitations here, they are equally germane for other methods that have been used to assess multisensory function in human subjects, which are discussed in the following sections.

Electroencephalography (EEG)

The methods to study multisensory integration in neural populations discussed above are invasive in nature and thus are almost invariably performed in animal models (but see the ECoG discussion below). In humans, the need for non-invasive methods has driven the application of event-related potentials (ERP) as measured by EEG to the study of multisensory questions (Lakatos et al. 2008; Buzsaki and Draguhn 2004; Fries 2005; Lakatos et al. 2005; Magri et al. 2012; Logothetis et al. 2001; Viswanathan and Freeman 2007; Rauch et al. 2008). While EEG measures neural

activity as a series of voltage signals derived from an array of scalp electrodes, the spatial resolution of these techniques is quite coarse, creating a number of issues that must be considered when interpreting data gathered via these methods. Perhaps most importantly, the measured EEG signal at the scalp reflects the concatenated activity of a host of different neural generators, the localization of which is exceedingly difficult. State-of-the-art EEG source modeling approaches were developed to solve what has been referred to as the “inverse problem”, but still have a spatial resolution on the order of $6 \times 6 \times 6$ mm (de Peralta Menendez et al. 2001, 2004). As a result of this spatial limitation, millions of neurons reside within this large volume of tissue, and these populations of neurons are extremely heterogeneous (Fig. 3b, c). As such, the criterion for assessing multisensory integration using EEG/ERP approaches must differ from that used to evaluate single neurons.

In discussing single-unit recordings, we described how referencing the maximum unisensory response has been well-established and validated as a criterion for identifying multisensory integration. However, in population measures such as EEG/ERP, exceeding the maximum unisensory response may be achieved by recordings from a population including unisensory neurons responding to two (or more) different sensory modalities (areal convergence), and does not necessarily imply multisensory integration (neuronal convergence; Fig 3) (Beauchamp et al. 2004a; Laurienti et al. 2005; Stevenson et al. 2009; Besle et al. 2009; James and Stevenson 2012).

In order to account for the possibility of areal convergence, one must consider how to quantify the appropriate null hypothesis. ERP recordings directly measure the electrical fields generated by neuronal activity, and electrical fields sum linearly. Because of this, if there are two populations of synchronously firing unisensory neurons, the predicted ERP response would be the linear sum of the responses recorded with the presentation of the two respective unisensory stimulus components (Besle et al. 2004; Giard and Besle 2010; Besle et al. 2009). The electrical activities that originate from the brain region of interest travel equally in all directions and thus impact the electrical recording at each node across the entire scalp. As such, this ability to account for independently firing pools of unisensory neurons (i.e., those that are not integrating) cannot be indexed through the application of the maximum criterion (Eq. 1). Due to this constraint, multisensory research using ERPs have generally used the additive criterion (Berman 1961; Barth et al. 1995; Besle et al. 2004):

$$ERP_{AV} \neq ERP_A + ERP_V. \quad (3)$$

The use of this criterion has been challenged in a number of forums (for an in-depth discussion, see Besle et al. 2004; Giard and Besle 2010; Gondan and Röder 2006; Teder-Salejarvi et al. 2002). The most pressing of issues associated with this metric is referred to as common activation (CA). That is, neural activity not directly related to sensory processing, such as motor activity, is also summed across unisensory conditions, while it is only represented once in the multisensory response (Besle et al. 2004; Giard and Besle 2010). Due to CA, the typical multisensory responses as seen using EEG are subadditive (Puce et al. 2007; Vroomen and Stekelenburg 2010; Stekelenburg and Vroomen 2007). Given this concern, a number of methods have attempted to account for CA in EEG. The first assumes that processes that are not strictly related to sensory processing usually occur later than 200 ms after stimulus onset. Hence, one strategy is to simply exclude the components of the ERP later than this point in time (Hillyard et al. 1998). While this lowers the chance of CA as a confound, it also precludes the ability to

find any multisensory interaction after this time. A second strategy to deal with CA is the inclusion of null trials in which participant's perform the same task as they would with the auditory, visual, and AV trials, but in the absence of sensory information (Talsma and Woldorff 2005). Null trials, which in theory elicit the same CA seen in the other conditions, are then subtracted from each condition, eliminating their effect. Although task-related CAs may differ in the presence of a stimulus, this method can reduce the impact of CA. Thus, while the additive criteria is often used to measure multisensory integration in EEG, one must be cautious of using subadditivity as a marker for active sensory integration when there are issues that relate to common activity.

Finally, one can use the additive factors method, adapted from studies of RTs (Sternberg 1969a, b, 1998). The additive factors approach consists of parametrically modulating some component of the auditory, visual, and AV stimuli and measuring a change in the relative responses to unisensory and multisensory presentations across levels of that modulation. For example, one study varied the signal-to-noise ratio of auditory and visual speech stimuli, showing that ERPs previously thought to be exclusively visual were in fact modulated by auditory stimulation (Stevenson et al. 2012a). If the change in responses across these variations in unisensory stimuli is different from the respective change with multisensory stimuli,

$$AV_H - AV_L \neq (A_H - A_L) + (V_H - V_L), \quad (4)$$

then effectiveness is not having a selective influence (i.e. changing the effectiveness of the stimulus in one modality impacts the processing of the second modality), and thus there is evidence for multisensory interaction. Through the calculation of these differences, the CA is also subtracted out. While there may be differences in CA across the added factor, this method will reliably reduce the impact of CA (Stevenson et al. 2012b). Therefore, use of the additive factors method provides a more conservative metric for identifying active integration across sensory modalities. For a more in-depth discussion of additive factors and its application to multisensory behavioral research, see the section *Behavioral measures: response times* below.

While multisensory interactions have historically been characterized using peak-based analyses of the ERP voltage waveforms, other analysis tools are being increasingly used to assess EEG data from a more dynamical network perspective. A particularly powerful set of analyses is referred to as electrical neuroimaging (Michel and Murray 2012). This technique is based on reference-independent measures of the global electric field at the scalp to differentiate between modulations in response strength, topography, and latency (Murray et al. 2008; Michel et al. 2004, 2009). These analyses characterize ERPs according to the

sequence and amplitude of topographical maps that in turn can disambiguate (a) effects that are the result of changes in response amplitudes originating from a stable set of generators from (b) effects that are the result of changes in the configuration of generators. Electrical neuroimaging analyses can be directly applied to multisensory integration, where a change in the neural generators underlying multisensory processing relative to unisensory processing, or changes in the timing that those generators are active, can represent distinct integrative processes (Altieri et al. 2013; Hauthal et al. 2013; Thelen et al. 2012; Murray et al. 2004, 2005). Indeed, this technique has now provided direct links between neural measures of multisensory integration and metrics of behavioral performance, such as RTs (Cappe et al. 2012; Stevenson et al. 2012b; Sperdin et al. 2010).

In addition to EEG recordings on the scalp, under certain exceptional circumstances one can make these measurements intracranially in human subjects using a method referred to as electrocorticography (ECoG). In ECoG recordings an array of electrodes is generally placed on the surface of the brain, allowing better spatial resolution when compared with EEG [down to approximately 1 mm (Asano et al. 2005)], and offering the additional advantage of being able to detect much weaker signals, as the current does not have to travel through the skull. Analysis of ECoG data for multisensory interactions generally follows the same principles as that for scalp recordings, and has been used to anatomically identify specific regions such as the temporo-parietal junction as multisensory in nature (Matsuhashi et al. 2004), as well as to demonstrate that audiovisual interactions occur as early as secondary sensory cortices (Besle et al. 2008; Mercier et al. 2013; Gomez-Ramirez et al. 2011; Pockett et al. 2013).

Functional Magnetic Resonance Imaging

In contrast to EEG, which has excellent temporal resolution and relatively poor spatial resolution, fMRI has excellent spatial resolution and relatively poor temporal resolution. However, even at the upper boundaries of fMRI's spatial resolution (e.g., in 7T magnets used in human research), tens- to hundreds-of-thousands of neurons are contained within a single voxel. Furthermore, BOLD fMRI's measure of neural activity is indirect being derived from blood flow and oxygenation. Thus, as with EEG, one must always be aware that the sample is derived from a population of heterogeneous neurons (and in this case from an indirect measure of activity in these neurons). In general, the spatial-temporal tradeoff between EEG and fMRI has led to the use of EEG as a primary tool for studying *when* an integrative process occurs, and fMRI for studying *where* in the brain multisensory interactions are taking place.

In the case of fMRI, it has been shown that BOLD activations from a given voxel sum in a roughly linear, time-invariant fashion, assuming ceiling and floor effects are avoided (Boynton et al. 1996; Dale and Buckner 1997; Glover 1999; Heeger and Ress 2002). As such, a conservative null hypothesis would be equal to the sum of the BOLD responses seen with the unisensory stimulus presentations (i.e., the additive criterion; Fig. 3a). Although superadditivity was seen in an early study examining multisensory cortical processing (Calvert et al. 2001), it has rarely been seen in subsequent studies (Stevenson et al. 2007; Stevenson and James 2009; Werner and Noppeney 2010; Foxe et al. 2002). The inability of fMRI to reveal superadditivity on a consistent basis is likely due to several features. First, even in the most "multisensory" of structures, multisensory neurons are interdigitated among substantial populations of unisensory neurons. Second, the response properties of multisensory neurons are highly heterogeneous, and include neurons that show response enhancements to a certain stimulus complex, others that show response depressions, and still others that show no interactions (Fig. 1, right) (James and Stevenson 2012; Laurienti et al. 2005). Given these factors, the likelihood of even a single voxel showing a superadditive response to a multisensory stimulus combination is quite low.

With these practical constraints, the majority of fMRI studies opt to use the maximum criterion (see Eq. 5 below). Nonetheless, the limitations described above mean that the maximum criterion is still unable to conservatively disambiguate areal and neuronal convergence, resulting in a conceptual dilemma. Since the BOLD response is derived from a heterogeneous population of neurons, one must use the additive criterion to eliminate the possibility of mere areal convergence, yet the heterogeneity makes this practically unfeasible. To address this limitation, a recent study examined auditory, visual and audiovisual processing using both criteria (Beauchamp 2005). This study successfully identified established multisensory regions using the maximum criterion, but failed to do so using the additive criteria. This finding reflects the bulk of the multisensory fMRI literature, in which a failure to achieve superadditivity while still exceeding the maximum criterion has been the typical result (Beauchamp et al. 2004a, b; Stevenson and James 2009). Nonetheless, studies that have taken advantage of the principle of inverse effectiveness and presented very weakly effective stimuli have had more success (Stevenson et al. 2007; Stevenson and James 2009). Thus, when superadditivity is found, the results strongly point to active integration across sensory modalities, as this pattern of activation cannot be accounted for through independent unisensory processing (James et al. 2009, 2012). In comparison, when only the maximum criteria is surpassed, this provides weaker evidence that an

active integrative process has occurred, and cannot definitely rule out concurrent, independent unisensory processing that has resulted from simple areal convergence.

While the maximum and additive criteria are the most used and discussed metrics for quantifying multisensory integration using fMRI, they are not the only ones. Three other approaches that have been used are fMR-adaptation (fMR-A), additive factors, and BOLD dynamics. While the previously-described criteria are hindered by the problem of heterogeneity, fMR-A is specifically designed to isolate a sub-population of neurons (Grill-Spector and Malach 2001). With fMR-A, the participant is repeatedly presented with the same stimulus, resulting in adaptation/fatigue of neurons responsive to that stimulus. When used in a multisensory context, the adapting stimulus is presented in one modality (e.g. auditory), while the next stimulus presentation is in a different modality (e.g. vision). If the adapted neurons are multisensory, then the response to the presentation of the same stimulus in a different modality will result in a decreased BOLD response relative to a novel presentation of that stimulus. In this manner, fMR-A has been successfully used to identify multisensory convergence (Tal and Amedi 2009). Additionally, the additive factors paradigm and approach has been used successfully with fMRI to show multisensory integration (Kim and James 2010; Stevenson et al. 2009, 2010; James and Stevenson 2012; Kim et al. 2012). Finally, another means to address issues related to using measures of BOLD response amplitude is to evaluate the dynamics of the BOLD response, namely the slopes and peak latencies of the hemodynamic signal (Martuzzi et al. 2007). Peak latencies of responses to multisensory stimuli that occur sooner than predicted by BOLD responses using unisensory stimuli are thus indicative of a multisensory interaction.

As should be clear, none of the metrics used to identify multisensory integration in neural populations is flawless and without conceptual and/or practical caveats. Nonetheless, much has been learned through their application to questions specific to multisensory function. However, recognition of the limitations associated with each is vitally important, and experimental conclusions should be tempered with respect to these limitations. Incorporating new methods of analysis as well as converging evidence from these established methods will provide important insights into multisensory convergence and integration in large-scale neuronal networks.

Multisensory Integration in Behavior and Perception

Both animal and human studies have served to highlight the fundamental importance of multisensory interactions

from a behavioral and perceptual perspective. Animal studies have demonstrated that multisensory stimuli that increase activity at the neuronal level also enhance an animal's orientation abilities, yielding heightened detection and localization of external events (Stein et al. 1988; Kinsella et al. 1989; Wilkinson et al. 1996; Jiang et al. 2002; Burnett et al. 2004; Calvert et al. 2004; Stanford and Stein 2007; Driver and Noesselt 2008; Stein and Stanford 2008; Sarko et al. 2012; Van Wanrooij et al. 2009; Corneil et al. 2002; Bell et al. 2005; Frens and Van Opstal 1998). Within human psychophysical studies [and non-human primate investigations as well (Cappé et al. 2010; Miller et al. 2001)] there are a number of behavioral measurements used to quantify differences in responses mediated by multisensory integration, including, but not limited to, detection, accuracy, and reaction/RTs. As highlighted above, with each distinct type of measurement, it is essential to understand the strengths and weaknesses of the metrics that are employed.

Behavioral Measures: Detection and Accuracy

Two of the simplest and most commonly employed measures used to identify multisensory interactions in behavior are stimulus detection and accuracy. In a typical psychophysical paradigm, a multisensory (e.g., audiovisual) stimulus combination is repeatedly presented, often interleaved with the presentation of the component unisensory stimuli. The participant is asked to detect, localize, identify, or make a judgment about the stimulus(i) on each trial. A multisensory interaction is identified when the response to the multisensory stimulus significantly deviates from response predicted by the unisensory stimulus presentations.

At first glance, one may assume that it would be quite simple to calculate the predicted multisensory response rate if the detection/accuracy rate exceeds that seen to both unisensory stimuli. Thus, if

$$\hat{p}(AV) > \max[p(A), p(V)], \quad (5)$$

then one would infer a multisensory interaction (Fig. 4). However, there is a caveat with this interpretation. Although a significant increase relative to the maximum unisensory response rate does indeed illustrate that the participant is using information from both senses, it does not necessarily imply that the participant is *integrating* this information, as it does not account for response gains associated with mere statistical facilitation. To account for statistical facilitation, one should instead use the equation:

$$\hat{p}(AV) > p(V) + [1 - p(V)] \times p(A), \quad (6)$$

which can be rewritten the following way if independence is assumed,

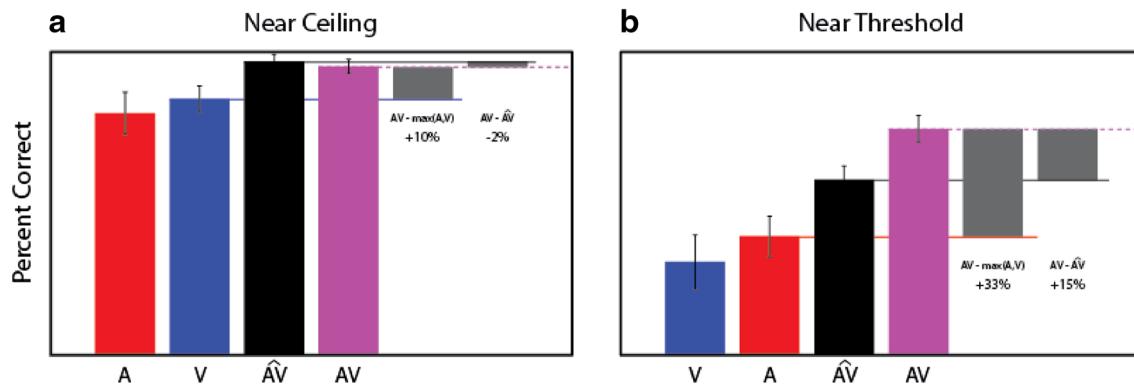


Fig. 4 Multisensory enhancement in accuracy. The manner with which multisensory accuracies are predicted from responses to unisensory stimuli vary. When near ceiling, the maximum criterion

$$\hat{p}(AV) > p(A) + p(V) - p(A) \cdot p(V) \quad (7)$$

In these two equations, the term on the right represents the response rate that would be expected when both auditory and visual stimuli were presented together and were processed independently. In short, the predicted response rate for a paired AV stimulus is the sum of the detection rates for each of the unisensory presentations, less the probability that both the unisensory auditory and the unisensory visual stimulus were detected on the same trial (Fig. 4). Any detection rate that exceeds this predicted rate implies that information is being integrated.

It is important to note here that the maximum criterion, despite the limitations elaborated on above, is often useful. Certain unavoidable factors can preclude the use of a more conservative metric. Many paradigms can only use correct trials for additional analyses, making it impractical to induce large error rates, whereas others have practical constraints imposed by the special populations being studied (e.g., children or clinical populations) or by expense (e.g., in fMRI experiments). As such, it is often necessary to design a behavioral experiment such that the participants are at or near ceiling performance, which precludes the ability to observe multisensory enhancements using more conservative criteria (Fig. 4a). In these cases, important information can still be gleaned from using more liberal criteria, particularly when the effect being studied is well-established. However, one must be cautious regarding the conclusions drawn from such results, particularly if the results are not supported by previous research or converging evidence.

Finally, alternative methods to those described above may be applied to assess the extent to which multisensory integration occurs and the associated magnitude of the effects. To use one example from speech perception, multisensory integration has been assessed using the fuzzy logical model of perception (FLMP; Massaro 1987, 2004). FLMP utilizes a formula similar to Luce's choice rule (Luce 1959)

is often a more practical metric (a), whereas responses not approaching ceiling can take advantage of more conservative predictions that allow for stronger conclusions to be drawn (b)

to describe how auditory and visual cues are combined. In this model, auditory and visual information is combined in a weighted fashion based on how reliably the information within the sensory signal predicts a given stimulus percept relative to alternative stimuli in the response set. As an example, in FLMP the probability of correctly identifying the speech stimulus/ba/, instead of a similar speech token, given the available auditory and visual information, can be described using the following rule:

$$\hat{p}(/ba|A_i, V_j) = \frac{A_i V_j}{A_i V_j + (1 - A_i)(1 - V_j)}. \quad (8)$$

The values A_i and V_j denote the level of auditory and visual reliability for a particular stimulus. A crucial feature of FLMP is that the prediction for AV identification probabilities (and thus the presence of integration) involves a non-linear function of how the multisensory systems weight information from each individual modality.

Behavioral Measures: Response Times

Measuring multisensory interactions via reaction or RTs employs a similar principle to that used with detection and accuracy, but with a slightly more nuanced logic. It is tempting to assume that a faster mean RT under multisensory conditions when compared with the corresponding unisensory conditions is reflective of multisensory integration, but this is not necessarily the case. In fact, simple statistical facilitation predicts that even under independent circumstances, multisensory RTs should be faster than the unisensory correlates. This statistical facilitation, when applied to multisensory systems, has been referred to as the redundant target or race-model effect (Miller 1982; Raab 1962). To conceptualize two processes as independent in terms of RTs, one can think of two channels racing against each other (Raab 1962) with a choice made as soon as

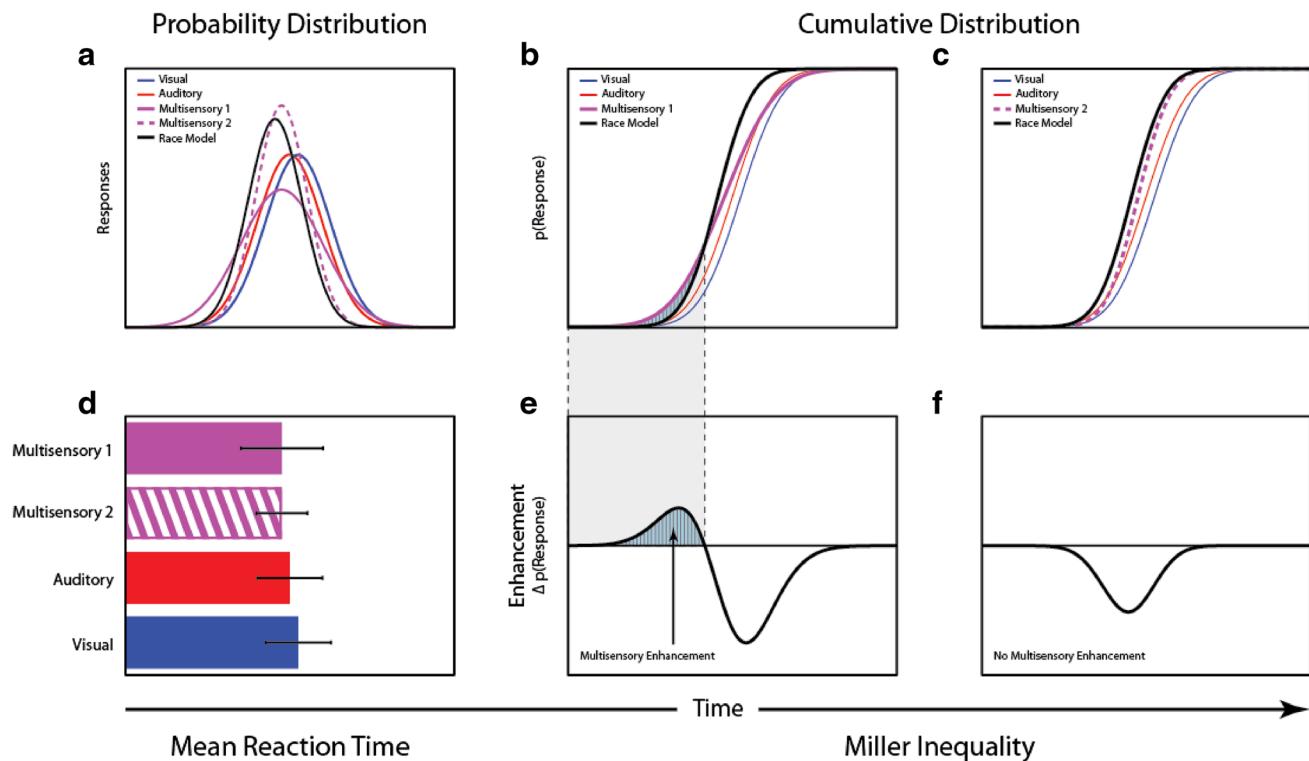


Fig. 5 Multisensory enhancement in response times. Response times are best analyzed as a distribution (a). Faster mean RTs with multisensory relative to the fastest unisensory presentations can occur without any interaction between the senses (d), and should instead be

compared as a distribution against a race model (b–c). Race model violations measured as the difference between the multisensory CDF and the race model CDF, known as Miller's Inequality (e–f), can provide a rigorous test of multisensory interactions

either of the processes is complete. When this is the case, the predicted mean RT with an AV presentation will be faster than either of the two unisensory RTs *without any interaction between the two processes*.

Redundant target facilitation may be modeled by creating two independent normal distributions, one representing the RT distribution with auditory-only and one with visual-only presentations (Fig. 5a, red and blue). To calculate a distribution predicted from two racing, non-interactive processes, we first randomly selected a RT from each unisensory distribution. The faster of these two RTs (i.e., the first unisensory process to finish) is then recorded as the predicted AV response from that trial (a minimum stopping-time rule). The resulting \hat{AV} distribution is plotted in Fig. 5a (in black).

The race model thus provides a useful benchmark against which to test for multisensory interactions. Using the race model, such interactions are most commonly measured in terms of cumulative distribution functions (CDFs; Fig. 5b, c). CDFs (denoted by $F(t)$) represent the cumulative probability that a response has been made by a given point in time following a stimulus presentation. An upper bound on parallel independent race model predictions can be formulated in terms of CDFs (Miller 1982). Violations of the upper bound at any time imply that

independent race model predictions cannot account for the facilitation observed in the audiovisual condition, and hence that a multisensory interaction is taking place (Fig. 5c, f). This upper boundary is defined as:

$$F_{AV}(t) = F_A(t) + F_V(t) - F_A(t) \cdot F_V(t) < F_A(t) + F_V(t) \quad (9)$$

where $F_V(t)$ is the probability that a response has been made to a visual stimulus by time t and $F_A(t)$ is the probability that the response has been made to the auditory stimulus by time point t . Race-model violations occur when the CDF associated with multisensory RTs is to the left of the upper bound of the non-interactive race-model prediction (Fig. 5b and 5e; statistically measured using a Kolmogorov–Smirnov test). This leftward shift in the CDF implies the following non-mutually exclusive processes: (a) a change in capacity or cognitive resources, (b) interactions between processes, or (c) a pooling of resources, otherwise known as *coactivation* (Miller 1982; Townsend and Nozawa 1995). These differences between the race-model and the actual multisensory (i.e., AV) CDF can be visualized as the difference of these two functions, known as the Miller inequality (Fig. 5e, f). In the Miller inequality, positive values are taken as evidence for multisensory

integration. It should be noted here that negative values do *not* imply any type of interaction as these could indicate, but not differentiate, sub-optimal integration, independent processing, or interference.

The race model approach to the analysis of multisensory RTs has been recently extended by employing a measure known as *capacity* to assess audiovisual integration efficiency under easy and difficult conditions (O’Riordan 2004). Similar to the race model test, *capacity* compares RTs from trials where both auditory *and* visual information are available to RTs obtained from trials where only auditory *or* visual information is presented. The capacity metric represents an expansion of the traditional race model test for at least a few reasons. First, a transformed hazard function ratio (described below) is used to establish a benchmark of efficiency (capacity = 1), to which processing on audiovisual trials may be compared. A related advantage of capacity is that the measure is computed using integrated hazard functions, which yield an instantaneous assessment of the work completed (see Townsend and Nozawa, 1995). The hazard function approach captures the notion of “capacity” and “efficient audiovisual processing” more closely than mean RTs or mean accuracy and also has many statistical advantages over the use of mean RTs and CDFs. The capacity coefficient thus uses the entire distribution of RTs, at the level of the integrated hazard function. It is measured by dividing the probability density function by the survivor function ($S(t)$)—the probability that recognition has not occurred by time t) and then obtaining the cumulative sum (see below)(Gauthier et al. 2003; Townsend and Ashby 1978).

$$H(t^*) = \int_0^{t^*} h(t)dt. \quad (10)$$

We may interpret the cumulative hazard function as denoting the total amount of work completed until a specified unit of time. As such, Townsend and Nozawa (Townsend and Nozawa 1995) derived the benchmark capacity coefficient in tasks where observers are presented with 0, 1, or 2 sources of stimuli. For our present purposes, let $H_{AV}(t)$ denote the integrated hazard function obtained from trials in which audiovisual information is available, and let $H_A(t)$ and $H_V(t)$ denote the integrated hazard functions obtained from the auditory and visual-only trials. We may define the capacity coefficient, $C(t)$, as:

$$C(t) = \frac{H_{AV}(t)}{H_A(t) + H_V(t)} \quad (11)$$

The sum in the denominator corresponds to the independent race model prediction. Capacity provides a non-parametric measure of integration with three possible outcomes. First, $C(t)$ can be greater than 1 at time t , indicating

more work completed in the audiovisual condition compared to the auditory and visual-only conditions, and hence, the presence of integration. Second, capacity can be less than 1 at time t , indicating limited resources or the presence of cross-modal inhibition. Finally, capacity may equal 1 at time t . This would suggest that audiovisual processing is neither faster nor slower than parallel predictions, and hence, evidence for multisensory integration is lacking. For example, recent reports show evidence for high capacity, or efficient integration, in a word recognition task when the auditory S/N ratio was low, but evidence for limited capacity when the auditory signal was clear (Altieri and Townsend 2011). Finally, recent work has sought to combine both RT and accuracy in a single capacity measure that tests redundant target responses against standard parallel predictions (O’Riordan and Plaisted 2001).

Yet another way to identify multisensory integration using RTs, or in a broader sense, to detect an interaction between any two processes, is through the use of the previously described additive factors (Sternberg 1969a, b, 1975, 1998). Additive factors were originally used to address problems with Donder’s subtraction method (1868), a method used to measure the length of cognitive processes. The subtraction method assumes that a given process can be timed by comparing RTs between a task with and a task without the measured process, which we will call y :

$$RT_{xyz} - RT_{xz} = RT_y. \quad (12)$$

While this subtraction method has provided substantial insights into cognitive processing, it relies upon a number of assumptions, including the assumption that each process is independent. In the study of multisensory integration, however, what is being studied is a potential interaction between cognitive processes. To address this assumption, Sternberg (1969a) created an additive-factors paradigm in which the process in question is manipulated by some added factor as the experimental conditions:

$$RT_{xy} - RT_{xy'} = RT_y - RT_{y'} \quad (13)$$

where some change to the process y is denoted as y' . When results fit this equation, the experimental factor selectively influences process y , suggesting that processes x and y are independent. An inequality, on the other hand, indicates a lack of selective influence of the added factor. As such, this implies that the cognitive processes x and y interact. In the current framework, the two cognitive processes can be considered to be the auditory and visual sensory processes (for extensions and more rigorous generalizations of selective influence and additive factors, see Ashby 1982; Ashby and Townsend 1986; Townsend 1984; Townsend and Ashby 1980; Townsend and Thomas 1994; Wenger

and Townsend 2000; Sternberg 2001; Taylor 1976; Schweickert 1978; Pieters 1983).

One example of such an additive-factors paradigm being applied to multisensory research is varying the effectiveness of auditory and visual stimuli through variations in stimulus intensity or strength. The null hypothesis for such additive factors modeling of multisensory redundant targets is displayed as:

$$(A_H V_H - A_H V_L) = (A_L V_H - A_L V_L), \quad (14)$$

where H and L refer to low and high stimulus effectiveness, respectively. Non-zero results provide evidence that the relationship between processing the auditory and visual stimuli differs as a factor of stimulus efficacy, and thus are indicative of a multisensory interaction (Altieri and Townsend 2011).

Perceptual Measures of Multisensory Integration

Although multisensory interactions can be quantitatively measured using approaches that index changes in detection, accuracy, and RTs, there are also a number of more qualitative experimental manipulations that provide evidence supporting multisensory integration. These manipulations provide evidence that multisensory cues are merged during the formation of the perception of an external event, commonly referred to as perceptual fusion or binding. Importantly, findings such as these are not reliant upon quantifications of multisensory gain as detailed above and can thus be used as supporting evidence for the presence of multisensory integration.

One of the most well-known examples illustrating the perceptual fusion of multisensory cues is the McGurk effect (McGurk and MacDonald 1976), an illusion in which the combination of incongruent visual and auditory speech signals produces a novel percept that represents a synthesis of the two sensory channels. For example, the presentation of an auditory/ba/ and a visual/ga/ often yields the fused percept/da/or/tha/.

Another audiovisual illusion, the ventriloquist effect, provides evidence of the influences of the different sensory modalities on the localization of a stimulus in space (Bertelson and Radeau 1981). In this illusion, named for the famous act, the ventriloquist has the ability to induce a shift in the perceived location of the sound source by: (1) minimizing the articulatory cues of his own mouth and lips, and (2) moving the “dummy’s” mouth in concert with the audible signal. This perceived spatial shift of the auditory signal to the location of the visual signal provides evidence that the auditory and visual systems interact. Although in its most extreme form the ventriloquist is an example of perceptual “capture” (i.e., the localization judgment is made at the location of the biasing stimulus), other

paradigms have also been used to highlight the spatial biases induced by the presentation of spatially incongruent multisensory cues. For example, the spatial localization of a sound source in a darkened room can be strongly “pulled” toward an extraneous visual cue, even when subjects are told to actively ignore the light (Bertelson 1998; Spence and Driver 2000; Radeau 1994; Hairston et al. 2003).

Another common multisensory illusion used to assess the magnitude of integration or binding is the sound-induced flash illusion (Shams et al. 2000). In this illusion, the participant is presented with a single visual flash paired with multiple auditory beeps and is instructed to count the number of flashes while ignoring the beeps. The multiple beeps induce the perception of multiple flashes even when only a single flash is presented. Also, the pairing of a single beep with multiple flashes can induce a fusion effect, where only a single flash is perceived (Andersen et al. 2004b; Shams et al. 2005; Mishra et al. 2008). In addition to their utility in assessing the strength or susceptibility to perceptual integration (Calvert and Thesen 2004), these various illusions can be used to examine factors such as how the temporal relationship of the paired stimuli impacts the illusory percepts. Indeed, studies have used such tasks to examine the multisensory temporal binding window (Foss-Feig et al. 2010; Colonius et al. 2009; Stevenson et al. 2014; Stevenson and Wallace 2013; Stevenson et al. 2012c; Hairston et al. 2005; Powers et al. 2009; 2012; Stevenson et al. 2013; Dixon and Spitz 1980; Vroomen and Keetels 2010; Conrey and Pisoni 2006; van Wassenhove et al. 2007)—the span of SOAs within which the illusion is highly likely to take place.

Linking Neural Activity with Perception and Behavior

While each of the previously described methods used to study multisensory integration are useful and provide pertinent information as to how sensory stimuli are synthesized to drive behavior and shape perception, the use of these methods in isolation provides only a partial picture concerning multisensory processing. To understand these phenomena more fully, it is necessary to complement these behavioral and perceptual analyses with neural studies designed to reveal the mechanistic underpinnings of multisensory interactions and how they relate to behavior and perception.

Neurophysiological recordings in animal models conducted during the performance of a behavioral task are becoming increasingly commonplace within the individual sensory systems [vision; (Britten et al. 1992, 1996; Yang and Maunsell 2004; Lakatos et al. 2008; David et al. 2008; Leopold and Logothetis 1996; Churan and Ilg 2001; Thiele

et al. 1999; Maier et al. 2007; Li and Basso 2005; Snyder et al. 2006), audition; (Recanzone et al. 2000b; Niwa et al. 2012; Recanzone et al. 2000a), and touch; (Romo et al. 1998, 2002; Steinmetz et al. 2000)], however they have been slow to be adopted within multisensory systems (but see Bell et al. 2004; Frens and Van Opstal 1998; Wang et al. 2008; Burton et al. 1997). Such studies allow strong inferences to be drawn between the activity of individual neurons (and groups of neurons) and the behavioral processes under study. The most common use to date of these neurometric/psychometric analyses has been within the visual system. Thus, in the middle temporal area (MT) of monkeys, a higher order visual region implicated in visual motion processing single- and multi-unit spikes and LFPs have been shown to be informative regarding task-related behavior (i.e., motion judgments; Britten et al. 1992, 1996; Price and Born 2010; Uka et al. 2012; Dodd et al. 2001; Liu and Newsome 2005; Bradley et al. 1998; Liu and Newsome 2006).

Albeit less well developed than work in the visual system, several studies have attempted to link neurophysiology to behavior during the performance of a multisensory task (Wang et al. 2008; Bell et al. 2003; Iriki et al. 1996; Van Opstal and Munoz 2004; Frens and Van Opstal 1998). For instance, several of these studies have correlated the activity of neurons in the superior colliculus (SC) during an audiovisual task with the timing of saccadic eye movements to the target (Van Opstal and Munoz 2004; Frens and Van Opstal 1998).

One area in which the connections from neurons to behavior is beginning to be addressed is in the realm of cue integration. When cues from multiple modalities are available simultaneously, behavioral precision is improved in a manner predicted by statistically optimal (Bayesian) cue integration models (Fetsch et al. 2013; Angelaki et al. 2011). Such examples of multisensory cue integration and use of optimal cue integration models are important in understanding how multisensory neurons in different brain regions combine multisensory information to improve behavior, and are very helpful in bridging the gap between theories of sensory cue integration and physiology of multisensory neuronal populations (for review, see Fetsch et al. 2013). Briefly, as suggested by Fetsch et al. (2013), in a highly simplified system, we may assume the existence of two populations of primary sensory neurons, each receiving unisensory information from different modalities. Information from these unisensory neurons is then relayed to a multisensory neural population, the activity of which generates a particular behavior or perceptual choice. In the scenario mentioned above, the output of two primary sensory neurons converges onto a multisensory neuron with varying synaptic weights that reflects the number and/or efficacy of synaptic connections associated with each modality. Such synaptic activity cannot be directly

measured by extracellular spike recordings but LFP recordings (discussed above) provide a more direct measure of such subthreshold synaptic changes. Spiking activity, however, provides the final output of network computation with firing rates that vary when different unisensory stimuli are presented alone or when multisensory stimuli are presented in combination. We can then ask how multisensory firing rate is best predicted from firing rates under unisensory conditions (for example by a weighted sum with different neural weights for distinct unisensory conditions). Finally, the population activity of the multisensory neurons is read out by downstream circuits to generate a behavioral response or perceptual choice.

Another set of powerful tools that can brought to bear on multisensory questions and to establish better links between neurons and behavior are those derived from signal detection theory (SDT) and receiver operating characteristics (ROC) analysis (Green and Swets 1966). Though widely used in behavioral analysis, SDT and ROC analysis has proved useful in linking neurons and behavior in unisensory research (de Lafuente and Romo 2005; Romo et al. 2004). In the SDT framework, a subject is asked to detect the presence of a signal in a yes–no response choice. Responses can be divided into four categories: hits (correct detects), misses (signal present but no response), false alarms (response with signal absent), and correct rejects. SDT posits that two probability distributions are formed along some continuum of internal representation of stimulus state (i.e., a noise distribution and a signal + noise distribution). A typical observer will set a criterion level along this continuum; every time the internal representation exceeds this criterion, the observer will respond “yes.” More salient stimuli will typically increase the separation between these two distributions. Two valuable measures are then derived from the response matrix: detectability (d') and bias (β). Whereas d' reflects the sensory discriminability of the stimuli, β represents the internal criterion set by the participant on the task. Usually, d' is given as the difference between noise and signal + noise distribution means divided by the standard deviation of the distributions:

$$d' = \frac{\bar{m}_{s+n} - \bar{m}_n}{\delta}, \quad (15)$$

However, for quantifying multisensory gain, we would take the difference between the means of a multisensory distribution and a prediction distribution based on the unisensory components (for example, a joint probability distribution):

$$d' = \frac{\bar{m}_{AV} - \bar{m}_{\hat{AV}}}{\delta}, \quad (16)$$

In a different form, d' is given as the difference between the z-transformed hit rate and the z-transformed false alarm

rate (Macmillan and Creelman 2004). In a multisensory context we would substitute a predictive model for false alarm rate:

$$d' = z(HR_{AV}) - z(HR_{A\bar{V}}) \quad (17)$$

In ROC analysis, the probability distributions are compared using a sliding criterion threshold to assess the degree of separation between the distributions. As the criterion slides along the distributions, area under the noise distribution is plotted against area under the signal + noise distribution, yielding the ROC curve. The power of this ROC analysis is that it can also be applied to neural datasets, where now the distributions are the spontaneous activity (i.e., noise) and the evoked activity (i.e., signal + noise). By comparing the separation of the neurometric distributions with the separation of the psychometric distributions, ROC analysis can thus begin to link neurons with behavior. In order to adapt this method for quantifying multisensory neural gain and compare it with behavioral gain, a multisensory predictive model distribution derived from the unisensory response distributions (for example, a simple joint probability) can be compared to the observed multisensory response distribution.

Although SDT and ROC analyses represent valuable tools in relating the activity of neurons to behavioral and perceptual processes, they are still tools that assess correlations between neural and behavioral datasets. The next frontier in multisensory research is that which extends these approaches to more causally-based methods. Indeed, assessing such causal links has been done in multisensory systems using methods such as cortical deactivation (Koval et al. 2011; Lomber and Malhotra 2008; Malhotra et al. 2004, 2008; Jiang et al. 2001), transcranial magnetic stimulation (TMS; Takarae et al. 2008; Beauchamp et al. 2010; Romei et al. 2007, 2009; Ramos-Estebanez et al. 2007; Romei et al. 2013; Spierer et al. 2013), pharmacological injections (Lee et al. 1988), and lesions (Burnett et al. 2007; Baum et al. 2012; for review, see Bolognini et al. 2013). One powerful new method that will undoubtedly shed important light on multisensory neurons and their contributions to behavior and perception is optogenetics, because of its ability to excite and inhibit specific neuron populations (Deisseroth 2011; Boyden 2011; Mattis et al. 2012; Yizhar et al. 2011a, b; Fenn et al. 2011). By comparing measures of behavior with and without optogenetic stimulation, it may be possible to causally link neuronal populations with certain behaviors.

Conclusions

The increased number of studies investigating interactions across sensory modalities has led to a marked improvement

in our understanding of what these interactions are, where and how they take place, and their behavioral and perceptual implications. With the rapid growth in this field comes an increase in the breadth of methodologies used to measure multisensory interactions. While the overarching goals of studies of multisensory integration may be similar across these methodologies, there are also significant differences between the assumptions that must be accounted for when analyzing the dependent variables observed with each of these measures. From neuronal response measures to behavioral studies of detection, accuracy, and RTs, and ultimately to applying analysis tools that link neurometric properties to behavioral outcomes (Britten et al. 1992; Guido et al. 1995; Thompson et al. 1996; Dahmen et al. 2010), each dependent variable measured with these techniques must be analyzed and assessed within its own framework. Each method has unique benefits, drawbacks, and assumptions that must be considered not only when choosing how to measure a given multisensory effect, but also when identifying and quantifying multisensory interactions.

Here, we have detailed the manner with which these underlying factors influence how multisensory interactions are evaluated using a number of common experimental methodologies. Additionally, we have provided an overview of the more canonical metrics used to quantify multisensory interactions with each of these methodologies, as well as the conclusive claims that can and cannot be made with each criterion. While by no means is this an exhaustive review of these topics, it is our hope that the discussion of method-specific quantifications will spur future research to take into account these differences between methodologies and the manner in which multisensory integration is quantified in each.

Acknowledgments Funding for this work was provided by a Banting Postdoctoral Fellowship, *It's only a matter of time: Neural networks underlying multisensory perceptual binding*, National Institutes of Health grant F32 DC011993, *Multisensory Integration and Temporal Processing in ASD*, a grant from the Vanderbilt Institute for Clinical and Translational Research, VICTR VR5807.1, *Development and Modulation of Multisensory Integration*, National Institutes of Health grant R34 DC010927, *Evaluation of Sensory Integration Treatment in ASD*, National Institutes of Health grant 5T32 MH018921-24, *From Brain and Behavioral Science to Intervention*, a Vanderbilt Kennedy Center MARI/Hobbs Award, the Vanderbilt Brain Institute, the Vanderbilt University Kennedy Center, the Idaho IDEA Network of Biomedical Research Excellence, and National Institutes of Health grants P20 RR016454 and P20 GM103408.

References

- Ahissar M, Ahissar E, Bergman H, Vaadia E (1992) Encoding of sound-source location and movement: activity of single neurons and interactions between adjacent neurons in the monkey auditory cortex. *J Neurophysiol* 67(1):203–215

- Allman BL, Meredith MA (2007) Multisensory processing in “unimodal” neurons: cross-modal subthreshold auditory effects in cat extrastriate visual cortex. *J Neurophysiol* 98(1):545–549
- Allman BL, Keniston LP, Meredith MA (2008) Subthreshold auditory inputs to extrastriate visual neurons are responsive to parametric changes in stimulus quality: sensory-specific versus non-specific coding. *Brain Res* 1242:95–101. doi:[10.1016/j.brainres.2008.03.086](https://doi.org/10.1016/j.brainres.2008.03.086)
- Allman BL, Keniston LP, Meredith MA (2009) Not just for bimodal neurons anymore: the contribution of unimodal neurons to cortical multisensory processing. *Brain Topogr* 21(3–4):157–167. doi:[10.1007/s10548-009-0088-3](https://doi.org/10.1007/s10548-009-0088-3)
- Altieri N, Townsend JT (2011) An assessment of behavioral dynamic information processing measures in audiovisual speech perception. *Front Psychol* 2:238. doi:[10.3389/fpsyg.2011.00238](https://doi.org/10.3389/fpsyg.2011.00238)
- Altieri N, Stevenson RA, Wallace MT, Wenger MJ (2013) Learning to associate auditory and visual stimuli: behavioral and neural mechanisms. *Brain Topogr*. doi:[10.1007/s10548-013-0333-7](https://doi.org/10.1007/s10548-013-0333-7)
- Andersen RA, Musallam S, Pesaran B (2004a) Selecting the signals for a brain-machine interface. *Curr Opin Neurobiol* 14(6):720–726. doi:[10.1016/j.conb.2004.10.005](https://doi.org/10.1016/j.conb.2004.10.005)
- Andersen TS, Tiippainen K, Sams M (2004b) Factors influencing audiovisual fission and fusion illusions. *Brain Res Cogn Brain Res* 21(3):301–308. doi:[10.1016/j.cogbrainres.2004.06.004](https://doi.org/10.1016/j.cogbrainres.2004.06.004)
- Angelaki DE, Gu Y, DeAngelis GC (2011) Visual and vestibular cue integration for heading perception in extrastriate visual cortex. *J Physiol* 589(4):825–833
- Asano E, Juhasz C, Shah A, Muzik O, Chugani DC, Shah J, Sood S, Chugani HT (2005) Origin and propagation of epileptic spasms delineated on electrocorticography. *Epilepsia* 46(7):1086–1097. doi:[10.1111/j.1528-1167.2005.05205.x](https://doi.org/10.1111/j.1528-1167.2005.05205.x)
- Ashby FG (1982) Testing the assumptions of exponential, additive reaction time models. *Mem Cogn* 10:125–134
- Ashby FG, Townsend JT (1986) Varieties of perceptual independence. *Psychol Rev* 93(2):154–179
- Barth DS, Goldberg N, Brett B, Di S (1995) The spatiotemporal organization of auditory, visual, and auditory-visual evoked potentials in rat cortex. *Brain Res* 678(1–2):177–190
- Baum SH, Martin RC, Hamilton AC, Beauchamp MS (2012) Multisensory speech perception without the left superior temporal sulcus. *NeuroImage*. doi:[10.1016/j.neuroimage.2012.05.034](https://doi.org/10.1016/j.neuroimage.2012.05.034)
- Beauchamp MS (2005) Statistical criteria in fMRI studies of multisensory integration. *Neuroinformatics* 3(2):93–113
- Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A (2004a) Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nat Neurosci* 7(11):1190–1192
- Beauchamp MS, Lee KE, Argall BD, Martin A (2004b) Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron* 41(5):809–823
- Beauchamp MS, Nath AR, Pasalar S (2010) fMRI-guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. *J Neurosci* 30(7):2414–2417. doi:[10.1523/JNEUROSCI.4865-09.2010](https://doi.org/10.1523/JNEUROSCI.4865-09.2010)
- Bell AH, Corneil BD, Munoz DP, Meredith MA (2003) Engagement of visual fixation suppresses sensory responsiveness and multisensory integration in the primate superior colliculus. *Eur J Neurosci* 18(10):2867–2873
- Bell AH, Fecteau JH, Munoz DP (2004) Using auditory and visual stimuli to investigate the behavioral and neuronal consequences of reflexive covert orienting. *J Neurophysiol* 91(5):2172–2184
- Bell AH, Meredith MA, Van Opstal AJ, Munoz DP (2005) Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements. *J Neurophysiol* 93(6):3659–3673
- Berens P, Keliris GA, Ecker AS, Logothetis NK, Tolias AS (2008a) Comparing the feature selectivity of the gamma-band of the local field potential and the underlying spiking activity in primate visual cortex. *Front Syst Neurosci* 2:2. doi:[10.3389/neuro.06.002.2008](https://doi.org/10.3389/neuro.06.002.2008)
- Berens P, Keliris GA, Ecker AS, Logothetis NK, Tolias AS (2008b) Feature selectivity of the gamma-band of the local field potential in primate primary visual cortex. *Front Neurosci* 2(2):199–207. doi:[10.3389/neuro.01.037.2008](https://doi.org/10.3389/neuro.01.037.2008)
- Berens P, Logothetis NK, Tolias AS (2013) Local field potentials, BOLD and spiking activity–relationships and physiological mechanisms. *Nature* 2:67
- Berman AL (1961) Interaction of cortical responses to somatic and auditory stimuli in anterior ectosylvian gyrus of cat. *J Neurophysiol* 24:608–620
- Bertelson P (1998) Starting from the ventriloquist: The perception of multimodal events. In: Sabourin M, Craik FIM, Robert M (eds) *Advances in psychological science*, vol 2., Biological and cognitive aspectsPsychological Press, Hove, pp 419–439
- Bertelson P, Radeau M (1981) Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Percept Psychophys* 29(6):578–584
- Besle J, Fort A, Giard M (2004) Interest and Validity of the additive model in electrophysiological studies of multisensory interactions. *Cogn Process* 5:189–192
- Besle J, Fischer C, Bidet-Caulet A, Lecaignard F, Bertrand O, Giard MH (2008) Visual activation and audiovisual interactions in the auditory cortex during speech perception: intracranial recordings in humans. *J Neurosci* 28(52):14301–14310. doi:[10.1523/JNEUROSCI.2875-08.2008](https://doi.org/10.1523/JNEUROSCI.2875-08.2008)
- Besle J, Bertrand O, Giard MH (2009) Electrophysiological (EEG, sEEG, MEG) evidence for multiple audiovisual interactions in the human auditory cortex. *Hear Res* 258:143–151. doi:[10.1016/j.heares.2009.06.016](https://doi.org/10.1016/j.heares.2009.06.016)
- Bizley JK, King AJ (2008) Visual-auditory spatial processing in auditory cortical neurons. *Brain Res* 1242:24–36. doi:[10.1016/j.brainres.2008.02.087](https://doi.org/10.1016/j.brainres.2008.02.087)
- Bolognini N, Convento S, Rossetti A, Merabet LB (2013) Multisensory processing after a brain damage: clues on post-injury crossmodal plasticity from neuropsychology. *Neurosci Biobehav Rev* 37(3):269–278
- Boyden ES (2011) Optogenetics: using light to control the brain. *Cerebrum* 2011:16
- Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16(13):4207–4221
- Bradley DC, Chang GC, Andersen RA (1998) Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature* 392(6677):714–717. doi:[10.1038/33688](https://doi.org/10.1038/33688)
- Brecht M, Singer W, Engel AK (1999) Patterns of synchronization in the superior colliculus of anesthetized cats. *J Neurosci* 19(9):3567–3579
- Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12(12):4745–4765
- Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neurosci* 13(1):87–100
- Burnett LR, Stein BE, Chaponis D, Wallace MT (2004) Superior colliculus lesions preferentially disrupt multisensory orientation. *Neuroscience* 124(3):535–547. doi:[10.1016/j.neuroscience.2003.12.026](https://doi.org/10.1016/j.neuroscience.2003.12.026)
- Burnett LR, Stein BE, Perrault TJ Jr, Wallace MT (2007) Excitotoxic lesions of the superior colliculus preferentially impact multisensory neurons and multisensory integration. *Exp Brain Res Experimentelle Hirnforschung* 179(2):325–338. doi:[10.1007/s00221-006-0789-8](https://doi.org/10.1007/s00221-006-0789-8)

- Burton H, Sinclair RJ, Hong SY, Pruitt JR Jr, Whang KC (1997) Tactile-spatial and cross-modal attention effects in the second somatosensory and 7b cortical areas of rhesus monkeys. *Somatosens Mot Res* 14(4):237–267
- Buzsaki G, Draguhn A (2004) Neuronal oscillations in cortical networks. *Science* 304(5679):1926–1929. doi:[10.1126/science.109745](https://doi.org/10.1126/science.109745)
- Calvert GA, Thesen T (2004) Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of physiology, Paris* 98(1–3):191–205
- Calvert GA, Hansen PC, Iversen SD, Brammer MJ (2001) Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *NeuroImage* 14(2):427–438
- Calvert GA, Spence C, Stein BE (2004) The handbook of multisensory processes. MIT, Cambridge
- Cappe C, Murray MM, Barone P, Rouiller EM (2010) Multisensory facilitation of behavior in monkeys: effects of stimulus intensity. *J Cogn Neurosci* 22(12):2850–2863. doi:[10.1162/jocn.2010.21423](https://doi.org/10.1162/jocn.2010.21423)
- Cappe C, Thelen T, Romei V, Thut G, Murray MM (2012) Looming signals reveal synergistic principles of multisensory integration. *J Neurosci* 32(4):1171–1182. doi:[10.1523/JNEUROSCI.5517-11.2012](https://doi.org/10.1523/JNEUROSCI.5517-11.2012)
- Carriere BN, Royal DW, Perrault TJ, Morrison SP, Vaughan JW, Stein BE, Wallace MT (2007) Visual deprivation alters the development of cortical multisensory integration. *J Neurophysiol* 98(5):2858–2867. doi:[10.1152/jn.00587.2007](https://doi.org/10.1152/jn.00587.2007)
- Churan J, Ilg UJ (2001) Processing of second-order motion stimuli in primate middle temporal area and medial superior temporal area. *J Opt Soc Am A Opt Image Sci Vis* 18(9):2297–2306
- Colonius H, Diederich A, Steenken R (2009) Time-window-of-integration (TWIN) model for saccadic reaction time: effect of auditory masker level on visual-auditory spatial interaction in elevation. *Brain Topogr* 21(3–4):177–184. doi:[10.1007/s10548-009-0091-8](https://doi.org/10.1007/s10548-009-0091-8)
- Conrey B, Pisoni DB (2006) Auditory-visual speech perception and synchrony detection for speech and nonspeech signals. *The Journal of the Acoustical Society of America* 119(6):4065–4073
- Corneil BD, Van Wanrooij M, Munoz DP, Van Opstal AJ (2002) Auditory-visual interactions subserving goal-directed saccades in a complex scene. *J Neurophysiol* 88(1):438–454
- Dahmen JC, Keating P, Nodal FR, Schulz AL, King AJ (2010) Adaptation to stimulus statistics in the perception and neural representation of auditory space. *Neuron* 66(6):937–948. doi:[10.1016/j.neuron.2010.05.018](https://doi.org/10.1016/j.neuron.2010.05.018)
- Dale AM, Buckner RL (1997) Selective averaging of rapidly presented individual trials using fMRI. *Hum Brain Mapp* 5:329–340
- David SV, Hayden BY, Mazer JA, Gallant JL (2008) Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. *Neuron* 59(3):509–521. doi:[10.1016/j.neuron.2008.07.001](https://doi.org/10.1016/j.neuron.2008.07.001)
- de Lafuente V, Romo R (2005) Neuronal correlates of subjective sensory experience. *Nat Neurosci* 8(12):1698–1703. doi:[10.1038/nn1587](https://doi.org/10.1038/nn1587)
- de Peralta Menendez RG, Andino SG, Lantz G, Michel CM, Landis T (2001) Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. *Brain Topogr* 14(2):131–137
- de Peralta Grave, Menendez R, Murray MM, Michel CM, Martuzzi R, Gonzalez Andino SL (2004) Electrical neuroimaging based on biophysical constraints. *NeuroImage* 21(2):527–539. doi:[10.1016/j.neuroimage.2003.09.051](https://doi.org/10.1016/j.neuroimage.2003.09.051)
- Deisseroth K (2011) Optogenetics. *Nat Methods* 8(1):26–29. doi:[10.1038/nmeth.f.324](https://doi.org/10.1038/nmeth.f.324)
- Diederich A, Colonius H (2004) Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Percept Psychophys* 66(8):1388–1404
- Dixon NF, Spitz L (1980) The detection of auditory visual desynchrony. *Perception* 9(6):719–721
- Dodd JV, Krug K, Cumming BG, Parker AJ (2001) Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *J Neurosci* 21(13):4809–4821
- Donders FC (1868) Over de Snelheid van Psychische Processen. Onderzoeken Gedaan in het Psychologisch Laboratorium der Utrechtse Hoogeschool
- Driver J, Noesselt T (2008) Multisensory interplay reveals crossmodal influences on ‘sensory-specific’ brain regions, neural responses, and judgments. *Neuron* 57(1):11–23
- Einevoll GT, Kayser C, Logothetis NK, Panzeri S (2013) Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nat Rev Neurosci* 14(11):770–785
- Fenno L, Yizhar O, Deisseroth K (2011) The development and application of optogenetics. *Annu Rev Neurosci* 34:389–412. doi:[10.1146/annurev-neuro-061010-113817](https://doi.org/10.1146/annurev-neuro-061010-113817)
- Fetsch CR, DeAngelis GC, Angelaki DE (2013) Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nat Rev Neurosci* 14(6):429–442
- Foss-Feig JH, Kwakye LD, Cascio CJ, Burnette CP, Kadivar H, Stone WL, Wallace MT (2010) An extended multisensory temporal binding window in autism spectrum disorders. *Exp Brain Res Experimentelle Hirnforschung* 203(2):381–389. doi:[10.1007/s00221-010-2240-4](https://doi.org/10.1007/s00221-010-2240-4)
- Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, Ritter W, Murray MM (2002) Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J Neurophysiol* 88(1):540–543
- Frens MA, Van Opstal AJ (1998) Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. *Brain Res Bull* 46(3):211–224
- Fries P (2005) A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn Sci* 9(10):474–480. doi:[10.1016/j.tics.2005.08.011](https://doi.org/10.1016/j.tics.2005.08.011)
- Fries P, Reynolds JH, Rorie AE, Desimone R (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291(5508):1560–1563. doi:[10.1126/science.291.5508.1560](https://doi.org/10.1126/science.291.5508.1560)
- Gauthier J, Joober R, Mottron L, Laurent S, Fuchs M, De Kimpe V, Rouleau GA (2003) Mutation screening of FOXP2 in individuals diagnosed with autistic disorder. *Am J Med Genet A* 118A(2):172–175. doi:[10.1002/ajmg.a.10105](https://doi.org/10.1002/ajmg.a.10105)
- Gerstein G, Bedenbaugh P, Aertsen AM (1989) Neuronal assemblies. *IEEE Trans Biomed Eng* 36(1):4–14
- GHazanfar AA, Maier JX, Hoffman KL, Logothetis NK (2005) Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J Neurosci* 25(20):5004–5012. doi:[10.1523/JNEUROSCI.0799-05.2005](https://doi.org/10.1523/JNEUROSCI.0799-05.2005)
- Giard M, Besle J (2010) Methodological considerations: Electrophysiology of multisensory interactions in humans. In: Kaiser J, Naumer MJ (eds) Multisensory object perception in the primate brain. Springer, New York
- Glover GH (1999) Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage* 9(4):416–429
- Goense JB, Logothetis NK (2008) Neurophysiology of the BOLD fMRI signal in awake monkeys. *Curr Biol* 18(9):631–640. doi:[10.1016/j.cub.2008.03.054](https://doi.org/10.1016/j.cub.2008.03.054)
- Gomez-Ramirez M, Kelly SP, Molholm S, Sehatpour P, Schwartz TH, Foxe JJ (2011) Oscillatory sensory selection mechanisms during intersensory attention to rhythmic auditory and visual inputs: a human electrocorticographic investigation. *J Neurosci* 31(50):18556–18567

- Gondan M, Röder B (2006) A new method for detecting interactions between the senses in event-related potentials. *Brain Res* 1073–1074:389–397. doi:[10.1016/j.brainres.2005.12.050](https://doi.org/10.1016/j.brainres.2005.12.050)
- Gray CM, Maldonado PE, Wilson M, McNaughton B (1995) Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. *J Neurosci Methods* 63(1–2):43–54
- Green DM, Swets JA (1966) Signal detection theory and psychophysics, vol 1974. Wiley, New York
- Grill-Spector K, Malach R (2001) fMRI-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol* 107(1–3):293–321
- Guido W, Lu SM, Vaughan JW, Godwin DW, Sherman SM (1995) Receiver operating characteristic (ROC) analysis of neurons in the cat's lateral geniculate nucleus during tonic and burst response mode. *Vis Neurosci* 12(4):723–741
- Hairston WD, Wallace MT, Vaughan JW, Stein BE, Norris JL, Schirillo JA (2003) Visual localization ability influences cross-modal bias. *J Cogn Neurosci* 15(1):20–29
- Hairston WD, Burdette JH, Flowers DL, Wood FB, Wallace MT (2005) Altered temporal profile of visual-auditory multisensory interactions in dyslexia. *Exp Brain Res Experimentelle Hirnforschung* 166(3–4):474–480. doi:[10.1007/s00221-005-2387-6](https://doi.org/10.1007/s00221-005-2387-6)
- Hauthal N, Thorne JD, Debener S, Sandmann P (2013) Source localisation of visual evoked potentials in congenitally deaf individuals. *Brain Topogr*. doi:[10.1007/s10548-013-0341-7](https://doi.org/10.1007/s10548-013-0341-7)
- Heeger DJ, Ress D (2002) What does fMRI tell us about neuronal activity? *Nat Rev* 3(2):142–151. doi:[10.1038/nrn730](https://doi.org/10.1038/nrn730)
- Henze DA, Borhegyi Z, Csicsvari J, Mamiya A, Harris KD, Buzsaki G (2000) Intracellular features predicted by extracellular recordings in the hippocampus in vivo. *J Neurophysiol* 84(1):390–400
- Hershenson M (1962) Reaction time as a measure of intersensory facilitation. *J Exp Psychol* 63:289–293
- Hillyard SA, Teder-Salejarvi WA, Munte TF (1998) Temporal dynamics of early perceptual processing. *Curr Opin Neurobiol* 8(2):202–210
- Iriki A, Tanaka M, Iwamura Y (1996) Coding of modified body schema during tool use by macaque postcentral neurones. *NeuroReport* 7(14):2325–2330
- James TW, Stevenson RA (2012) The use of fMRI to assess multisensory integration. In: Wallace MH, Murray MM (eds) *Frontiers in the neural basis of multisensory processes*. Taylor & Francis, London
- James TW, Kim S, Stevenson RA (2009) Assessing multisensory interaction with additive factors and functional MRI. In: The International Society for Psychophysics, Galway, Ireland, 2009
- James TW, Stevenson RA, Kim S (2012) Inverse effectiveness in multisensory processing. In: Stein BE (ed) *The new handbook of multisensory processes*. MIT Press, Cambridge
- Jenkins WM, Merzenich MM, Ochs MT, Allard T, Guic-Robles E (1990) Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. *J Neurophysiol* 63(1):82–104
- Jermakowicz WJ, Casagrande VA (2007) Neural networks a century after Cajal. *Brain Res Rev* 55(2):264–284. doi:[10.1016/j.brainresrev.2007.06.003](https://doi.org/10.1016/j.brainresrev.2007.06.003)
- Jiang W, Wallace MT, Jiang H, Vaughan JW, Stein BE (2001) Two cortical areas mediate multisensory integration in superior colliculus neurons. *J Neurophysiol* 85(2):506–522
- Jiang W, Jiang H, Stein BE (2002) Two corticotectal areas facilitate multisensory orientation behavior. *J Cogn Neurosci* 14(8):1240–1255
- Kajikawa Y, Schroeder CE (2011) How local is the local field potential? *Neuron* 72(5):847–858. doi:[10.1016/j.neuron.2011.09.029](https://doi.org/10.1016/j.neuron.2011.09.029)
- Katzner S, Nauhaus I, Benucci A, Bonin V, Ringach DL, Carandini M (2009) Local origin of field potentials in visual cortex. *Neuron* 61(1):35–41. doi:[10.1016/j.neuron.2008.11.016](https://doi.org/10.1016/j.neuron.2008.11.016)
- Kayser C, Petkov CI, Augath M, Logothetis NK (2005) Integration of touch and sound in auditory cortex. *Neuron* 48(2):373–384
- Kayser C, Petkov CI, Augath M, Logothetis NK (2007) Functional imaging reveals visual modulation of specific fields in auditory cortex. *J Neurosci* 27(8):1824–1835. doi:[10.1523/JNEUROSCI.4737-06.2007](https://doi.org/10.1523/JNEUROSCI.4737-06.2007)
- Kayser C, Petkov CI, Logothetis NK (2008) Visual modulation of neurons in auditory cortex. *Cereb Cortex* 18(7):1560–1574. doi:[10.1093/cercor/bhm187](https://doi.org/10.1093/cercor/bhm187)
- Kayser C, Petkov CI, Logothetis NK (2009) Multisensory interactions in primate auditory cortex: fMRI and electrophysiology. *Hear Res* 258(1–2):80–88. doi:[10.1016/j.heares.2009.02.011](https://doi.org/10.1016/j.heares.2009.02.011)
- Kim S, James TW (2010) Enhanced effectiveness in visuo-haptic object-selective brain regions with increasing stimulus salience. *Hum Brain Mapp* 31(5):678–693. doi:[10.1002/hbm.20897](https://doi.org/10.1002/hbm.20897)
- Kim S, Stevenson RA, James TW (2012) Visuo-haptic neuronal convergence demonstrated with an inversely effective pattern of BOLD activation. *J Cogn Neurosci* 24(4):830–842. doi:[10.1162/jocn_a_00176](https://doi.org/10.1162/jocn_a_00176)
- King AJ, Walker KM (2012) Integrating information from different senses in the auditory cortex. *Biol Cybern* 106(11–12):617–625. doi:[10.1007/s00422-012-0502-x](https://doi.org/10.1007/s00422-012-0502-x)
- Kinsella TJ, Trivette G, Rowland J, Sorace R, Miller R, Fraass B, Steinberg SM, Glatstein E, Sherins RJ (1989) Long-term follow-up of testicular function following radiation therapy for early-stage Hodgkin's disease. *J Clin Oncol* 7(6):718–724
- Koval MJ, Lomber SG, Everling S (2011) Prefrontal cortex deactivation in macaques alters activity in the superior colliculus and impairs voluntary control of saccades. *J Neurosci* 31(23):8659–8668. doi:[10.1523/JNEUROSCI.1258-11.2011](https://doi.org/10.1523/JNEUROSCI.1258-11.2011)
- Kreiman G, Hung CP, Kraskov A, Quiroga RQ, Poggio T, DiCarlo JJ (2006) Object selectivity of local field potentials and spikes in the macaque inferior temporal cortex. *Neuron* 49(3):433–445. doi:[10.1016/j.neuron.2005.12.019](https://doi.org/10.1016/j.neuron.2005.12.019)
- Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE (2005) An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J Neurophysiol* 94(3):1904–1911. doi:[10.1152/jn.00263.2005](https://doi.org/10.1152/jn.00263.2005)
- Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE (2007) Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53(2):279–292. doi:[10.1016/j.neuron.2006.12.011](https://doi.org/10.1016/j.neuron.2006.12.011)
- Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320(5872):110–113. doi:[10.1126/science.1154735](https://doi.org/10.1126/science.1154735)
- Laurienti PJ, Perrault TJ, Stanford TR, Wallace MT, Stein BE (2005) On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. *Exp Brain Res Experimentelle Hirnforschung* 166(3–4):289–297
- Lee C, Rohrer WH, Sparks DL (1988) Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature* 332(6162):357–360. doi:[10.1038/332357a0](https://doi.org/10.1038/332357a0)
- Leopold DA, Logothetis NK (1996) Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* 379(6565):549–553
- Li X, Basso MA (2005) Competitive stimulus interactions within single response fields of superior colliculus neurons. *J Neurosci* 25(49):11357–11373
- Liu J, Newsome WT (2005) Correlation between speed perception and neural activity in the middle temporal visual area. *J Neurosci* 25(3):711–722. doi:[10.1523/JNEUROSCI.4034-04.2005](https://doi.org/10.1523/JNEUROSCI.4034-04.2005)

- Liu J, Newsome WT (2006) Local field potential in cortical area MT: stimulus tuning and behavioral correlations. *J Neurosci* 26(30): 7779–7790
- Logothetis NK (2003) The underpinnings of the BOLD functional magnetic resonance imaging signal. *J Neurosci* 23(10): 3963–3971
- Logothetis NK (2008) What we can do and what we cannot do with fMRI. *Nature* 453(7197):869–878. doi:[10.1038/nature06976](https://doi.org/10.1038/nature06976)
- Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001) Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412(6843):150–157
- Lomber SG, Malhotra S (2008) Double dissociation of ‘what’ and ‘where’ processing in auditory cortex. *Nat Neurosci* 11(5): 609–616. doi:[10.1038/nn.2108](https://doi.org/10.1038/nn.2108)
- Lovelace CT, Stein BE, Wallace MT (2003) An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Brain Res Cogn Brain Res* 17(2):447–453
- Luce RD (1959) Individual choice behavior: a theoretical analysis. Courier Dover Publications, New York
- Macmillan NA, Creelman CD (2004) Detection theory: a user’s guide, 2nd edn. Psychology Press, USA
- Magri C, Schridde U, Murayama Y, Panzeri S, Logothetis NK (2012) The amplitude and timing of the BOLD signal reflects the relationship between local field potential power at different frequencies. *J Neurosci* 32(4):1395–1407. doi:[10.1523/JNEUROSCI.3985-11.2012](https://doi.org/10.1523/JNEUROSCI.3985-11.2012)
- Maier A, Logothetis NK, Leopold DA (2007) Context-dependent perceptual modulation of single neurons in primate visual cortex. *Proc Natl Acad Sci USA* 104(13):5620–5625. doi:[10.1073/pnas.0608489104](https://doi.org/10.1073/pnas.0608489104)
- Malhotra S, Hall AJ, Lomber SG (2004) Cortical control of sound localization in the cat: unilateral cooling deactivation of 19 cerebral areas. *J Neurophysiol* 92(3):1625–1643. doi:[10.1152/jn.01205.2003](https://doi.org/10.1152/jn.01205.2003)
- Malhotra S, Stecker GC, Middlebrooks JC, Lomber SG (2008) Sound localization deficits during reversible deactivation of primary auditory cortex and/or the dorsal zone. *J Neurophysiol* 99(4): 1628–1642. doi:[10.1152/jn.01228.2007](https://doi.org/10.1152/jn.01228.2007)
- Martuzzi R, Murray MM, Michel CM, Thiran J-P, Maeder PP, Clarke S, Meuli RA (2007) Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cereb Cortex* 17(7):1672–1679
- Massaro DW (1987) Speech Perception by Ear and Eye. In: Dodd B, Campbell BA (eds) Hearing by eye: the psychology of lip reading. Lawrence Erlbaum Associates Ltd., London, pp 53–84
- Massaro DW (2004) From Multisensory Integration to Talking Heads and Language Learning. In: Calvert G, Spence C, Stein BE (eds) The handbook of multisensory processes. MIT Press, Cambridge, pp 153–176
- Matsuhashi M, Ikeda A, Ohara S, Matsumoto R, Yamamoto J, Takayama M, Satow T, Begum T, Usui K, Nagamine T, Mikuni N, Takahashi J, Miyamoto S, Fukuyama H, Shibasaki H (2004) Multisensory convergence at human temporo-parietal junction: epicortical recording of evoked responses. *Clin Neurophysiol* 115(5):1145–1160. doi:[10.1016/j.clinph.2003.12.009](https://doi.org/10.1016/j.clinph.2003.12.009)
- Mattis J, Tye KM, Ferenczi EA, Ramakrishnan C, O’Shea DJ, Prakash R, Gunaydin LA, Hyun M, Fenno LE, Gradinari V, Yizhar O, Deisseroth K (2012) Principles for applying optogenetic tools derived from direct comparative analysis of microbial opsins. *Nat Methods* 9(2):159–172. doi:[10.1038/nmeth.1808](https://doi.org/10.1038/nmeth.1808)
- McGurk H, MacDonald J (1976) Hearing lips and seeing voices. *Nature* 264(5588):746–748
- Mercier MR, Foxe JJ, Fiebelkorn IC, Butler JS, Schwartz TH, Molholm S (2013) Auditory-driven phase reset in visual cortex: human electrocorticography reveals mechanisms of early multisensory integration. *NeuroImage* 79:19–29
- Meredith MA, Stein BE (1983) Interactions among converging sensory inputs in the superior colliculus. *Science* 221(4608):389–391
- Meredith MA, Stein BE (1985) Descending efferents from the superior colliculus relay integrated multisensory information. *Science* 227(4687):657–659
- Meredith MA, Stein BE (1986a) Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Res* 365(2):350–354
- Meredith MA, Stein BE (1986b) Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J Neurophysiol* 56(3):640–662
- Meredith MA, Nemitz JW, Stein BE (1987) Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J Neurosci* 7(10):3215–3229
- Merzenich MM, Nelson RJ, Kaas JH, Stryker MP, Jenkins WM, Zook JM, Cynader MS, Schoppmann A (1987) Variability in hand surface representations in areas 3b and 1 in adult owl and squirrel monkeys. *J Comp Neurol* 258(2):281–296. doi:[10.1002/cne.902580208](https://doi.org/10.1002/cne.902580208)
- Michel CM, Murray MM (2012) Towards the utilization of EEG as a brain imaging tool. *NeuroImage* 61(2):371–385. doi:[10.1016/j.neuroimage.2011.12.039](https://doi.org/10.1016/j.neuroimage.2011.12.039)
- Michel CM, Murray MM, Lantz G, Gonzalez S, Spinelli L, Grave de Peralta R (2004) EEG source imaging. *Clin Neurophysiol* 115(10):2195–2222. doi:[10.1016/j.clinph.2004.06.001](https://doi.org/10.1016/j.clinph.2004.06.001)
- Michel CM, Koenig T, Brandeis D, Gianotti LR, Wackermann J (2009) Electrical neuroimaging. Cambridge University Press, Cambridge
- Miller J (1982) Divided attention: evidence for coactivation with redundant signals. *Cognit Psychol* 14(2):247–279
- Miller J, Ulrich R, Lamarre Y (2001) Locus of the redundant-signals effect in bimodal divided attention: a neurophysiological analysis. *Percept Psychophys* 63(3):555–562
- Mishra J, Martinez A, Hillyard SA (2008) Cortical processes underlying sound-induced flash fusion. *Brain Res* 1242:102–115. doi:[10.1016/j.brainres.2008.05.023](https://doi.org/10.1016/j.brainres.2008.05.023)
- Mitzdorf U (1985) Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. *Physiol Rev* 65(1):37–100
- Mitzdorf U (1987) Properties of the evoked potential generators: current source-density analysis of visually evoked potentials in the cat cortex. *Int J Neurosci* 33(1–2):33–59
- Montemurro MA, Rasch MJ, Murayama Y, Logothetis NK, Panzeri S (2008) Phase-of-firing coding of natural visual stimuli in primary visual cortex. *Curr Biol* 18(5):375–380. doi:[10.1016/j.cub.2008.02.023](https://doi.org/10.1016/j.cub.2008.02.023)
- Murray MM, Michel CM, Grave de Peralta R, Ortigue S, Brunet D, Andino SG, Schnider A (2004) Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *NeuroImage* 21(1):125–135
- Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, Schroeder CE, Foxe JJ (2005) Grabbing your ear: rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb Cortex* 15(7):963–974. doi:[10.1093/cercor/bhh197](https://doi.org/10.1093/cercor/bhh197)
- Murray MM, Brunet D, Michel CM (2008) Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr* 20(4): 249–264. doi:[10.1007/s10548-008-0054-5](https://doi.org/10.1007/s10548-008-0054-5)
- Nelson WT, Hettinger LJ, Cunningham JA, Brickman BJ, Haas MW, McKinley RL (1998) Effects of localized auditory information on visual target detection performance using a helmet-mounted display. *Hum Factors* 40(3):452–460
- Niwa M, Johnson JS, O’Connor KN, Sutter ML (2012) Active engagement improves primary auditory cortical neurons’ ability to discriminate temporal modulation. *J Neurosci* 32(27): 9323–9334. doi:[10.1523/JNEUROSCI.5832-11.2012](https://doi.org/10.1523/JNEUROSCI.5832-11.2012)

- O'Riordan MA (2004) Superior visual search in adults with autism. *Autism* 8(3):229–248. doi:[10.1177/1362361304045219](https://doi.org/10.1177/1362361304045219)
- O'Riordan M, Plaisted K (2001) Enhanced discrimination in autism. *Q J Exp Psychol* 54(4):961–979
- Perkel DH, Gerstein GL, Moore GP (1967) Neuronal spike trains and stochastic point processes II. Simultaneous spike trains. *Biophys J* 7(4):419–440. doi:[10.1016/S0006-3495\(67\)86597-4](https://doi.org/10.1016/S0006-3495(67)86597-4)
- Perrault TJ Jr, Vaughan JW, Stein BE, Wallace MT (2003) Neuron-specific response characteristics predict the magnitude of multisensory integration. *J Neurophysiol* 90(6):4022–4026. doi:[10.1152/jn.00494.2003](https://doi.org/10.1152/jn.00494.2003)
- Perrault TJ Jr, Vaughan JW, Stein BE, Wallace MT (2005) Superior colliculus neurons use distinct operational modes in the integration of multisensory stimuli. *J Neurophysiol* 93(5):2575–2586
- Pesaran B, Musallam S, Andersen RA (2006) Cognitive neural prosthetics. *Curr Biol* 16(3):R77–80. doi:[10.1016/j.cub.2006.01.043](https://doi.org/10.1016/j.cub.2006.01.043)
- Pettersen KH, Hagen E, Einevoll GT (2008) Estimation of population firing rates and current source densities from laminar electrode recordings. *J Comput Neurosci* 24(3):291–313
- Pieters JPM (1983) Sternberg's additive factor method and underlying psychological processes: some theoretical consideration. *Psychol Bull* 93:411–426
- Pockett S, Purdy SC, Brennan BJ, Holmes MD (2013) Auditory click stimuli evoke event-related potentials in the visual cortex. *NeuroReport* 24(15):837–840
- Powers AR 3rd, Hillock AR, Wallace MT (2009) Perceptual training narrows the temporal window of multisensory binding. *J Neurosci* 29(39):12265–12274. doi:[10.1523/JNEUROSCI.3501-09.2009](https://doi.org/10.1523/JNEUROSCI.3501-09.2009)
- Powers AR 3rd, Hevey MA, Wallace MT (2012) Neural correlates of multisensory perceptual learning. *J Neurosci* 32(18):6263–6274. doi:[10.1523/JNEUROSCI.6138-11.2012](https://doi.org/10.1523/JNEUROSCI.6138-11.2012)
- Price NS, Born RT (2010) Timescales of sensory- and decision-related activity in the middle temporal and medial superior temporal areas. *J Neurosci* 30(42):14036–14045. doi:[10.1523/JNEUROSCI.2336-10.2010](https://doi.org/10.1523/JNEUROSCI.2336-10.2010)
- Puce A, Epling JA, Thompson JC, Carrick OK (2007) Neural responses elicited to face motion and vocalization pairings. *Neuropsychologia* 45(1):93–106. doi:[10.1016/j.neuropsychologia.2006.04.017](https://doi.org/10.1016/j.neuropsychologia.2006.04.017)
- Raab DH (1962) Statistical facilitation of simple reaction times. *Trans N Y Acad Sci* 24:574–590
- Radeau M (1994) Auditory-visual spatial interaction and modularity. *Curr Psychol Cogn* 13(1):3–51
- Ramos-Estebanez C, Merabet LB, Machii K, Fregni F, Thut G, Wagner TA, Romei V, Amedi A, Pascual-Leone A (2007) Visual phosphene perception modulated by subthreshold cross-modal sensory stimulation. *J Neurosci* 27(15):4178–4181. doi:[10.1523/JNEUROSCI.5468-06.2007](https://doi.org/10.1523/JNEUROSCI.5468-06.2007)
- Rauch A, Rainer G, Logothetis NK (2008) The effect of a serotonin-induced dissociation between spiking and perisynaptic activity on BOLD functional MRI. *Proc Natl Acad Sci USA* 105(18):6759–6764. doi:[10.1073/pnas.0800312105](https://doi.org/10.1073/pnas.0800312105)
- Recanzone GH, Guard DC, Phan ML (2000a) Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *J Neurophysiol* 83(4):2315–2331
- Recanzone GH, Guard DC, Phan ML, Su TK (2000b) Correlation between the activity of single auditory cortical neurons and sound-localization behavior in the macaque monkey. *J Neurophysiol* 83(5):2723–2739
- Romei V, Murray MM, Merabet LB, Thut G (2007) Occipital transcranial magnetic stimulation has opposing effects on visual and auditory stimulus detection: implications for multisensory interactions. *J Neurosci* 27(43):11465–11472. doi:[10.1523/JNEUROSCI.2827-07.2007](https://doi.org/10.1523/JNEUROSCI.2827-07.2007)
- Romei V, Murray MM, Cappe C, Thut G (2009) Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Curr Biol* 19(21):1799–1805. doi:[10.1016/j.cub.2009.09.027](https://doi.org/10.1016/j.cub.2009.09.027)
- Romei V, Murray MM, Cappe C, Thut G (2013) The contributions of sensory dominance and attentional bias to cross-modal enhancement of visual cortex excitability. *J Cogn Neurosci* 25(7):1122–1135. doi:[10.1162/jocn_a_00367](https://doi.org/10.1162/jocn_a_00367)
- Romo R, Hernandez A, Zainos A, Salinas E (1998) Somatosensory discrimination based on cortical microstimulation. *Nature* 392(6674):387–390
- Romo R, Hernandez A, Zainos A, Lemus L, Brody CD (2002) Neuronal correlates of decision-making in secondary somatosensory cortex. *Nat Neurosci* 5(11):1217–1225
- Romo R, Hernandez A, Zainos A (2004) Neuronal correlates of a perceptual decision in ventral premotor cortex. *Neuron* 41(1):165–173
- Royal DW, Carriere BN, Wallace MT (2009) Spatiotemporal architecture of cortical receptive fields and its impact on multisensory interactions. *Exp Brain Res* 198(2–3):127–136. doi:[10.1007/s00221-009-1772-y](https://doi.org/10.1007/s00221-009-1772-y)
- Sarko DK, Nidiffer AR, Powers AR, Ghose D, Fister MC, Hillock-Dunn A, Krueger J, Wallace MT (2012) Spatial and temporal features of multisensory processes: Bridging animal and human studies. In: Murray MM, Wallace MT (eds) *Frontiers in the neural bases of multisensory processes*. CRC Press, Boca Raton, pp 191–215
- Schweickert R (1978) A critical path generalization of the additive factor method: analysis of a Stroop task. *J Math Psychol* 18:105–139
- Shams L, Kamitani Y, Shimojo S (2000) Illusions. What you see is what you hear. *Nature* 408(6814):788
- Shams L, Ma WJ, Beierholm U (2005) Sound-induced flash illusion as an optimal percept. *NeuroReport* 16(17):1923–1927
- Singer W (1993) Neuronal representations, assemblies and temporal coherence. *Prog Brain Res* 95:461–474
- Snyder LH, Dickinson AR, Calton JL (2006) Preparatory delay activity in the monkey parietal reach region predicts reach reaction times. *J Neurosci* 26(40):10091–10099. doi:[10.1523/JNEUROSCI.0513-06.2006](https://doi.org/10.1523/JNEUROSCI.0513-06.2006)
- Spence C, Driver J (2000) Attracting attention to the illusory location of a sound: reflexive crossmodal orienting and ventriloquism. *NeuroReport* 11(9):2057–2061
- Sperdin HF, Cappe C, Murray MM (2010) The behavioral relevance of multisensory neural response interactions. *Front Neurosci* 3:9
- Spierer L, Manuel AL, Bueti D, Murray MM (2013) Contributions of pitch and bandwidth to sound-induced enhancement of visual cortex excitability in humans. *Cortex; a journal devoted to the study of the nervous system and behavior* 49(10):2728–2734. doi:[10.1016/j.cortex.2013.01.001](https://doi.org/10.1016/j.cortex.2013.01.001)
- Stanford TR, Stein BE (2007) Superadditivity in multisensory integration: putting the computation in context. *NeuroReport* 18(8):787–792. doi:[10.1097/WNR.0b013e3280c1e315](https://doi.org/10.1097/WNR.0b013e3280c1e315)
- Stanford TR, Quesey S, Stein BE (2005) Evaluating the operations underlying multisensory integration in the cat superior colliculus. *J Neurosci* 25(28):6499–6508
- Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. *Nat Rev Neurosci* 9(4):255–266. doi:[10.1038/nrn2331](https://doi.org/10.1038/nrn2331)
- Stein BE, Wallace MT (1996) Comparisons of cross-modality integration in midbrain and cortex. *Prog Brain Res* 112:289–299
- Stein BE, Huneycutt WS, Meredith MA (1988) Neurons and behavior: the same rules of multisensory integration apply. *Brain Res* 448(2):355–358
- Steinmetz PN, Roy A, Fitzgerald PJ, Hsiao SS, Johnson KO, Niebur E (2000) Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature* 404(6774):187–190

- Stekelenburg JJ, Vroomen J (2007) Neural correlates of multisensory integration of ecologically valid audiovisual events. *J Cogn Neurosci* 19(12):1964–1973. doi:[10.1162/jocn.2007.19.12.1964](https://doi.org/10.1162/jocn.2007.19.12.1964)
- Sternberg S (1969a) The discovery of processing stages: extensions of Donders' method. *Acta Psychol* 30:276–315
- Sternberg S (1969b) Memory-scanning: mental processes revealed by reaction-time experiments. *Am Sci* 57(4):421–457
- Sternberg S (1975) Memory scanning: new findings and current controversies. *Exp Psychol* 27:1–32
- Sternberg S (1998) Discovering mental processing stages: The method of additive factors. In: Scarborough D, Sternberg S (eds) An invitation to cognitive science, vol 4., Methods, models, and conceptual issues MIT Press, Cambridge, pp 739–811
- Sternberg S (2001) Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. *Acta Psychol* 106:147–246
- Stevenson RA, James TW (2009) Audiovisual integration in human superior temporal sulcus: inverse effectiveness and the neural processing of speech and object recognition. *NeuroImage* 44(3):1210–1223. doi:[10.1016/j.neuroimage.2008.09.034](https://doi.org/10.1016/j.neuroimage.2008.09.034)
- Stevenson RA, Wallace MT (2013) Multisensory temporal integration: task and stimulus dependencies. *Exp Brain Res Experimentelle Hirnforschung Experimentation cerebrale* 227(2):249–261. doi:[10.1007/s00221-013-3507-3](https://doi.org/10.1007/s00221-013-3507-3)
- Stevenson RA, Geoghegan ML, James TW (2007) Superadditive BOLD activation in superior temporal sulcus with threshold non-speech objects. *Exp Brain Res* 179(1):85–95
- Stevenson RA, Kim S, James TW (2009) An additive-factors design to disambiguate neuronal and areal convergence: measuring multisensory interactions between audio, visual, and haptic sensory streams using fMRI. *Exp Brain Res Experimentelle Hirnforschung* 198(2–3):183–194. doi:[10.1007/s00221-009-1783-8](https://doi.org/10.1007/s00221-009-1783-8)
- Stevenson RA, Altieri NA, Kim S, Pisoni DB, James TW (2010) Neural processing of asynchronous audiovisual speech perception. *NeuroImage* 49(4):3308–3318. doi:[10.1016/j.neuroimage.2009.12.001](https://doi.org/10.1016/j.neuroimage.2009.12.001)
- Stevenson R, Bushmakin M, Kim S, Wallace M, Puce A, James T (2012a) Inverse effectiveness and multisensory interactions in visual event-related potentials with audiovisual speech. *Brain Topogr* 1:19. doi:[10.1007/s10548-012-0220-7](https://doi.org/10.1007/s10548-012-0220-7)
- Stevenson RA, Bushmakin M, Kim S, Wallace MT, Puce A, James TW (2012b) Inverse effectiveness and multisensory interactions in visual event-related potentials with audiovisual speech. *Brain Topogr* 25(3):308–326. doi:[10.1007/s10548-012-0220-7](https://doi.org/10.1007/s10548-012-0220-7)
- Stevenson RA, Zemtsov RK, Wallace MT (2012c) Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *J Exp Psychol Hum Percept Perform*. doi:[10.1037/a0027339](https://doi.org/10.1037/a0027339)
- Stevenson RA, Wilson MM, Powers AR, Wallace MT (2013) The effects of visual training on multisensory temporal processing. *Exp Brain Res Experimentelle Hirnforschung Experimentation cerebrale* 225(4):479–489. doi:[10.1007/s00221-012-3387-y](https://doi.org/10.1007/s00221-012-3387-y)
- Stevenson RA, Siemann JK, Schneider BC, Eberly HE, Woynaroski TG, Camarata SM, Wallace MT (2014) Multisensory temporal integration in autism spectrum disorders. *J Neurosci* 34(3): 691–697. doi:[10.1523/JNEUROSCI.3615-13.2014](https://doi.org/10.1523/JNEUROSCI.3615-13.2014)
- Takarae Y, Luna B, Minshew NJ, Sweeney JA (2008) Patterns of visual sensory and sensorimotor abnormalities in autism vary in relation to history of early language delay. *J Int Neuropsychol Soc* 14(6):980–989. doi:[10.1017/S1355617708081277](https://doi.org/10.1017/S1355617708081277)
- Tal N, Amedi A (2009) Multisensory visual-tactile object related network in humans: insights gained using a novel crossmodal adaptation approach. *Exp Brain Res Experimentelle Hirnforschung Experimentation cerebrale* 198(2–3):165–182. doi:[10.1007/s00221-009-1949-4](https://doi.org/10.1007/s00221-009-1949-4)
- Talsma D, Woldorff MG (2005) Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *J Cogn Neurosci* 17(7):1098–1114. doi:[10.1162/0898929054475172](https://doi.org/10.1162/0898929054475172)
- Taylor DA (1976) Stage analysis of reaction time. *Psychol Bull* 83:161–191
- Taylor K, Mandon S, Freiwald WA, Kreiter AK (2005) Coherent oscillatory activity in monkey area v4 predicts successful allocation of attention. *Cereb Cortex* 15(9):1424–1437. doi:[10.1093/cercor/bhi023](https://doi.org/10.1093/cercor/bhi023)
- Teder-Salejarvi WA, McDonald JJ, Di Russo F, Hillyard SA (2002) An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Brain Res Cogn Brain Res* 14(1):106–114
- Thelen A, Cappe C, Murray MM (2012) Electrical neuroimaging of memory discrimination based on single-trial multisensory learning. *NeuroImage* 62(3):1478–1488. doi:[10.1016/j.neuroimage.2012.05.027](https://doi.org/10.1016/j.neuroimage.2012.05.027)
- Thiele A, Distler C, Hoffmann KP (1999) Decision-related activity in the macaque dorsal visual pathway. *Eur J Neurosci* 11(6): 2044–2058
- Thompson KG, Hanes DP, Bichot NP, Schall JD (1996) Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. *J Neurophysiol* 76(6):4040–4055
- Townsend JT (1984) Uncovering mental processes with factorial experiments. *J Math Psychol* 28(4):363–400
- Townsend JT, Ashby FG (1978) Methods of modeling capacity in simple processing systems. *Cogn Theory* 3:200–239
- Townsend JT, Ashby FG (1980) Decomposing the reaction time distribution: pure insertion and selective influence revisited. *J Math Psychol* 21:93–123
- Townsend JT, Nozawa G (1995) Spatio-temporal properties of elementary perception: an investigation of parallel, serial, and coactive theories. *J Math Psychol* 39:321–359
- Townsend JT, Thomas RD (1994) Stochastic dependencies in parallel and serial models: effects on systems factorial interactions. *J Math Psychol* 38(1):1–34
- Uka T, Sasaki R, Kumano H (2012) Change in choice-related response modulation in area MT during learning of a depth-discrimination task is consistent with task learning. *J Neurosci* 32(40):13689–13700. doi:[10.1523/JNEUROSCI.4406-10.2012](https://doi.org/10.1523/JNEUROSCI.4406-10.2012)
- Usrey WM, Reid RC (1999) Synchronous activity in the visual system. *Annu Rev Physiol* 61:435–456. doi:[10.1146/annurev.physiol.61.1.435](https://doi.org/10.1146/annurev.physiol.61.1.435)
- Van Opstal AJ, Munoz DP (2004) Auditory-visual interactions subserving primate gaze orienting. In: Calvert GA, Spence C, Stein BE (eds) The handbook of multisensory processes. MIT Press, Cambridge, pp 373–394
- Van Wanrooij MM, Bell AH, Munoz DP, Van Opstal AJ (2009) The effect of spatial-temporal audiovisual disparities on saccades in a complex scene. *Exp Brain Res.* doi:[10.1007/s00221-009-1815-4](https://doi.org/10.1007/s00221-009-1815-4)
- van Wassenhove V, Grant KW, Poeppel D (2007) Temporal window of integration in auditory-visual speech perception. *Neuropsychologia* 45(3):598–607. doi:[10.1016/j.neuropsychologia.2006.01.001](https://doi.org/10.1016/j.neuropsychologia.2006.01.001)
- Viswanathan A, Freeman RD (2007) Neurometabolic coupling in cerebral cortex reflects synaptic more than spiking activity. *Nat Neurosci* 10(10):1308–1312. doi:[10.1038/nn1977](https://doi.org/10.1038/nn1977)
- Vroomen J, Baart M (2009) Recalibration of phonetic categories by lipread speech: measuring aftereffects after a 24-hour delay. *Lang Speech* 52(Pt 2–3):341–350
- Vroomen J, Keetels M (2010) Perception of intersensory synchrony: a tutorial review. *Atten Percept Psychophys* 72(4):871–884. doi:[10.3758/APP.72.4.871](https://doi.org/10.3758/APP.72.4.871)

- Vroomen J, Stekelenburg JJ (2010) Visual anticipatory information modulates multisensory interactions of artificial audiovisual stimuli. *J Cogn Neurosci* 22(7):1583–1596. doi:[10.1162/jocn.2009.21308](https://doi.org/10.1162/jocn.2009.21308)
- Wallace MT, Meredith MA, Stein BE (1992) Integration of multiple sensory modalities in cat cortex. *Exp Brain Res Experimentelle Hirnforschung Experimentation cerebrale* 91(3):484–488
- Wallace MT, Wilkinson LK, Stein BE (1996) Representation and integration of multiple sensory inputs in primate superior colliculus. *J Neurophysiol* 76(2):1246–1266
- Wallace MT, Meredith MA, Stein BE (1998) Multisensory integration in the superior colliculus of the alert cat. *J Neurophysiol* 80(2): 1006–1010
- Wang Y, Celebreni S, Trotter Y, Barone P (2008) Visuo-auditory interactions in the primary visual cortex of the behaving monkey: electrophysiological evidence. *BMC Neurosci* 9:79. doi:[10.1186/1471-2202-9-79](https://doi.org/10.1186/1471-2202-9-79)
- Wenger MJ, Townsend JT (2000) Basic response time tools for studying general processing capacity in attention, perception, and cognition. *J Gen Psychol* 127(1):67–99
- Werner S, Noppeney U (2010) Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. *Cereb Cortex* 20(8):1829–1842. doi:[10.1093/cercor/bhp248](https://doi.org/10.1093/cercor/bhp248)
- Wilkinson LK, Meredith MA, Stein BE (1996) The role of anterior ectosylvian cortex in cross-modality orientation and approach behavior. *Exp Brain Res* 112(1):1–10
- Xing D, Yeh CI, Shapley RM (2009) Spatial spread of the local field potential and its laminar variation in visual cortex. *J Neurosci* 29(37):11540–11549. doi:[10.1523/JNEUROSCI.2573-09.2009](https://doi.org/10.1523/JNEUROSCI.2573-09.2009)
- Yang T, Maunsell JH (2004) The effect of perceptual learning on neuronal responses in monkey visual area V4. *J Neurosci* 24(7):1617–1626
- Yizhar O, Fenno L, Zhang F, Hegemann P, Deisseroth K (2011a) Microbial opsins: a family of single-component tools for optical control of neural activity. *Cold Spring Harb Protoc* 2011(3): top102
- Yizhar O, Fenno LE, Davidson TJ, Mogri M, Deisseroth K (2011b) Optogenetics in neural systems. *Neuron* 71(1):9–34. doi:[10.1016/j.neuron.2011.06.004](https://doi.org/10.1016/j.neuron.2011.06.004)
- Zion Golumbic EM, Poeppel D, Schroeder CE (2012) Temporal context in speech processing and attentional stream selection: a behavioral and neural perspective. *Brain Lang* 122(3):151–161. doi:[10.1016/j.bandl.2011.12.010](https://doi.org/10.1016/j.bandl.2011.12.010)
- Zion Golumbic E, Cogan GB, Schroeder CE, Poeppel D (2013a) Visual input enhances selective speech envelope tracking in auditory cortex at a “cocktail party”. *J Neurosci* 33(4):1417–1426. doi:[10.1523/JNEUROSCI.3675-12.2013](https://doi.org/10.1523/JNEUROSCI.3675-12.2013)
- Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, Poeppel D, Schroeder CE (2013b) Mechanisms underlying selective neuronal tracking of attended speech at a “cocktail party”. *Neuron* 77(5):980–991. doi:[10.1016/j.neuron.2012.12.037](https://doi.org/10.1016/j.neuron.2012.12.037)