

Evaluating the Operations Underlying Multisensory Integration in the Cat Superior Colliculus

Terrence R. Stanford,¹ Stephan Quessy,² and Barry E. Stein¹

¹Department of Neurobiology and Anatomy, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157, and ²Department of Neurobiology and Anatomy, University of Rochester, Rochester, New York 14642

It is well established that superior colliculus (SC) multisensory neurons integrate cues from different senses; however, the mechanisms responsible for producing multisensory responses are poorly understood. Previous studies have shown that spatially congruent cues from different modalities (e.g., auditory and visual) yield enhanced responses and that the greatest relative enhancements occur for combinations of the least effective modality-specific stimuli. Although these phenomena are well documented, little is known about the mechanisms that underlie them, because no study has systematically examined the operation that multisensory neurons perform on their modality-specific inputs. The goal of this study was to evaluate the computations that multisensory neurons perform in combining the influences of stimuli from two modalities. The extracellular activities of single neurons in the SC of the cat were recorded in response to visual, auditory, and bimodal visual–auditory stimulation. Each neuron was tested across a range of stimulus intensities and multisensory responses evaluated against the null hypothesis of simple summation of unisensory influences. We found that the multisensory response could be superadditive, additive, or subadditive but that the computation was strongly dictated by the efficacies of the modality-specific stimulus components. Superadditivity was most common within a restricted range of near-threshold stimulus efficacies, whereas for the majority of stimuli, response magnitudes were consistent with the linear summation of modality-specific influences. In addition to providing a constraint for developing models of multisensory integration, the relationship between response mode and stimulus efficacy emphasizes the importance of considering stimulus parameters when inducing or interpreting multisensory phenomena.

Key words: superior colliculus; multisensory integration; single-unit electrophysiology; audition; vision; cat

Introduction

Single neurons in the mammalian superior colliculus (SC) integrate afferent visual, auditory, and somatosensory spatial information and generate efferent motor commands to control downstream targets that innervate the musculature of the eyes, neck, and body (for review, see Sparks, 1986; Stein et al., 2004). Despite having important implications for how the sensorimotor transformations necessary for orienting behavior are accomplished, the mechanisms that underlie multisensory integration are mostly unknown. Previous studies have firmly established two principles: (1) given sufficient overlap in time and space, stimuli from different modalities have synergistic effects on the response of an SC multisensory neuron and that (2) this synergy is most pronounced when the modality-specific stimuli are themselves least effective in driving the neuron (Stein and Meredith, 1993). The phenomena of multisensory enhancement and inverse effectiveness also have parallels in behavioral experiments showing facilitated stimulus detection/orientation for multisensory stimuli relative to their modality-specific components (Stein et al., 1988, 1989; Wilkinson et al., 1996; Jiang et al., 2002).

Although the benefits of integrating cross-modal sensory cues, both in terms of increased neuronal firing rate and improved orienting behavior, are amply documented at the phenomenological level, the mechanisms responsible for effecting multisensory integration at the single neuron level are poorly understood. Most previous studies have focused on multisensory response enhancement relative to the more effective of the two modality-specific component stimuli (Meredith and Stein, 1983) and, with the general goal of probing the limits of enhancement for a given neuron, have studied only a very limited range of weakly effective stimuli.

To date, no systematic study has evaluated SC neurons in terms of the computation(s) they perform in combining their unisensory inputs to yield their multisensory products. In the absence of quantitative information about the operation(s) that multisensory neurons implement, the mechanisms underlying multisensory integrations remain obscure, and there are few useful constraints for efforts to model established multisensory phenomena.

The current study evaluates the computations that produce multisensory integration and, more specifically, the multisensory enhancements that result from combining two excitatory modality-specific stimuli (i.e., visual and auditory). Toward this end, we have tested multisensory neurons at multiple levels of modality-specific stimulus effectiveness and at various temporal offsets between the two cross-modal stimuli for the purpose of

Received Dec. 14, 2004; revised May 29, 2005; accepted May 29, 2005.

This work was supported by National Institutes of Health Grants NS36916 and NS22543.

Correspondence should be addressed to Terrence R. Stanford, Department Neurobiology and Anatomy, Wake Forest University School of Medicine, Winston-Salem, NC 27157. E-mail: Stanford@wfuibmc.edu.

DOI:10.1523/JNEUROSCI.5095-04.2005

Copyright © 2005 Society for Neuroscience 0270-6474/05/256499-10\$15.00/0

evaluating the mode of multisensory integration across a large range of input values for each neuron. To our knowledge, this study represents the first systematic and quantitative assessment of the integrative modes of SC multisensory neurons. Some of these results have appeared in preliminary form (Quessy et al., 2000).

Materials and Methods

Surgical preparation. Each animal ($n = 2$) was implanted with a stainless-steel recording chamber using aseptic techniques in accordance with the *Guide for the Care and Use of Laboratory Animals* (National Institutes of Health publication 86-23) and an approved Institutional Animal Care and Use Committee protocol. The animal was rendered tractable with ketamine hydrochloride (20 mg/kg, i.m.) and acepromazine maleate (0.4 mg/kg, i.m.). Surgical anesthesia was maintained using halothane (1.5–4%). The animal was intubated and placed in a stereotaxic head-holder. Body temperature was measured with a rectal thermometer and maintained at 37–38°C with a circulating hot water heater. Appropriate hydration was maintained by intravenous injection (3–6 ml/h) of 5% dextrose Ringer's solution. A craniotomy was performed to expose the cortex overlying the SC. The recording chamber was then attached to the skull using surgical screws and orthopedic bone cement. The animal received postsurgical analgesic (butorphanol tartrate; 0.1–0.4 mg/kg for 6 h) as needed and antibiotics (cephazolin sodium; 25 mg/kg, twice daily) for 7 d.

Weekly recording sessions were begun after a postsurgical recovery period of at least 7 d. No wounds or pressure points were present, and for each recording session, the animal was anesthetized by injection of ketamine hydrochloride (20 mg/kg, i.m.) and acepromazine maleate (0.4 mg/kg, i.m.), intubated, and then paralyzed using an initial dose of pancuronium bromide (0.3 mg/kg). Anesthesia, paralysis, and proper hydration were maintained for the duration of the experiment by continuous injection of ketamine (10–15 mg · kg⁻¹ · h⁻¹, i.v.), pancuronium (0.1–0.2 mg · kg⁻¹ · h⁻¹, i.v.), and 5% dextrose Ringer's (1 mg · kg⁻¹ · h⁻¹, i.v.). The animal was artificially ventilated. Respiratory rate and volume were adjusted, and end-tidal CO₂ was monitored and maintained at ~4.0%. The recording chamber/head-holder was attached to a metal frame to allow unobstructed access to the animal. A calibrated X-Y slide was fitted to the chamber to guide the electrode to the SC. At the end of an experiment, injection of anesthetics and paralytics was terminated, and animal was allowed to recover normal respiration and locomotion before being returned to its home cage.

Conventional methods were used for electrophysiological recording. Single microelectrodes were advanced to the SC via hydraulic micromanipulator. Neuronal activity was amplified and monitored using an oscilloscope and an audiomonitor. Neurons were selected for study on the basis of responsiveness to both visual and auditory stimulation. Search stimuli consisted of a variety of flashing or moving spots of light along with tone bursts, clicks, or broadband noise bursts. When a visual/auditory neuron was identified and isolated, the visual receptive field was coarsely mapped by moving or flashing a bar or a spot along a translucent Plexiglas hemisphere. The auditory receptive field was similarly mapped by presenting brief (10 ms) noise bursts at various locations along the horizontal and vertical meridians. For data collection, stimulus positions were chosen on the basis of these preliminary receptive field maps. Because the modality-specific receptive fields of multisensory neurons overlap in space, we evaluated the integration of excitatory visual and auditory influences by presenting stimuli in nearby locations. Suppressive interactions, normally observed when one of the modality-specific stimuli is presented beyond the excitatory borders of the corresponding modality-specific receptive field (Kadunce et al., 1997), were not examined in this study.

Stimulus generation and stimulus conditions. Visual stimuli consisted of circular spots of light produced using high-brightness red or white light-emitting diodes (LEDs). The visual stimuli, projected through a semi-transparent screen, were presented in a region of the visual receptive field that yielded a robust response. Auditory stimuli were bursts of pink noise presented through small speakers positioned in close proximity to each

LED. Stimulus levels and durations were controlled by a Pentium personal computer via a Spike II (Cambridge Electronics Design, Cambridge, UK) analog-to-digital converter.

Each neuron was tested with modality-specific (visual alone, auditory alone) and multisensory (combined visual and auditory) stimuli. The goal of the study was to evaluate the multisensory computation for stimulus combinations ranging widely in terms of the efficacies of the modality-specific stimulus components. To this end, three stimulus levels and a single stimulus duration for each modality-specific stimulus component were determined empirically for each neuron. The procedure for selecting the parameter ranges was similar for both visual and auditory stimulation. First, an on-line estimate of approximate response threshold was determined to find the lowest stimulus level for which a response could be reliably evoked. Three stimulus intensities were then chosen in an attempt to span the dynamic range of the neuron. Although a few cases of a nonmonotonic relationship between stimulus intensity and stimulus efficacy (i.e., response magnitude) were observed (see Fig. 7), more typically, the low-, medium-, and high-intensity stimulus choices yielded responses that were near or slightly below threshold (low), slightly to moderately above threshold (medium), and well above threshold (high), respectively (see Fig. 1, left column and top row). Visual stimuli ranged from 0.65 to 13.0 candelas/m² and auditory stimuli from 0.7 to 70 dB sound pressure level. On average, the number of impulses (i.e., stimulus effectiveness) resulting from the lowest- and highest-intensity unisensory stimuli differed by just less than twofold, from an average of 3.4 ± 5.2 to 5.8 ± 7.2 impulses per trial (highest/lowest = 1.7) for visual stimuli and from 4.0 ± 5.9 to 7.4 ± 6.9 impulses per trial (highest/lowest = 1.85) for the auditory modality. For each modality, stimulus duration was made to be as brief as possible while still producing a clear response when presented at relatively high intensity.

The parameters (stimulus levels and durations) of the modality-specific components of the combined stimuli were the same as those selected for the modality-specific conditions, and multisensory pairings consisted of all nine possible intensity combinations (3 × 3). For each multisensory pairing, the relative timing of the modality-specific stimulus components was also varied. To choose the range of stimulus onset asynchrony (SOA), we first made an online estimate of the response latency to each modality-specific stimulus component (in all recorded neurons, responses latencies were longer for the visual stimulus). Typically, four timing configurations were tested in an attempt to bracket the value of interstimulus timing estimated to produce maximal temporal coincidence (estimated as the difference in visual and auditory response latencies). Generally, the smallest and largest SOAs differed by 100 ms (in increments of 25 ms). Thus, most neurons (35 of 41) were tested with at least 36 (3 × 3 × 4), multisensory stimulus combinations. A few exceptions include five neurons tested with three SOAs and one neuron tested with two visual stimulus intensities. Additionally, one neuron was tested with five visual stimulus intensities, and five neurons were tested with more than four SOAs. During testing, unisensory and multisensory stimulus configurations were presented randomly (with replacement). Neurons were tested with an average of 20 trials per stimulus configuration (minimum, 5; maximum, 30), with most neurons tested (38 of 41) having a minimum of 10 trials per stimulus combination.

Results

The response matrix of Figure 1 illustrates multisensory integration in an SC neuron and demonstrates the integrative products that can result from combining weakly effective versus more strongly effective modality-specific stimuli. The matrix consists of rasters and peristimulus time histograms (PSTHs) for three intensities of a visual stimulus (left column; open histograms), three intensities of an auditory stimulus (top row; filled histograms), and each of their nine possible cross-modal combinations (filled histograms).

The unisensory responses of this neuron to both the visual and auditory stimuli increased monotonically with stimulus level. They ranged from very weak to absent for low-intensity stimulation to relatively robust for high-intensity stimulation. Similarly,

AUDITORY

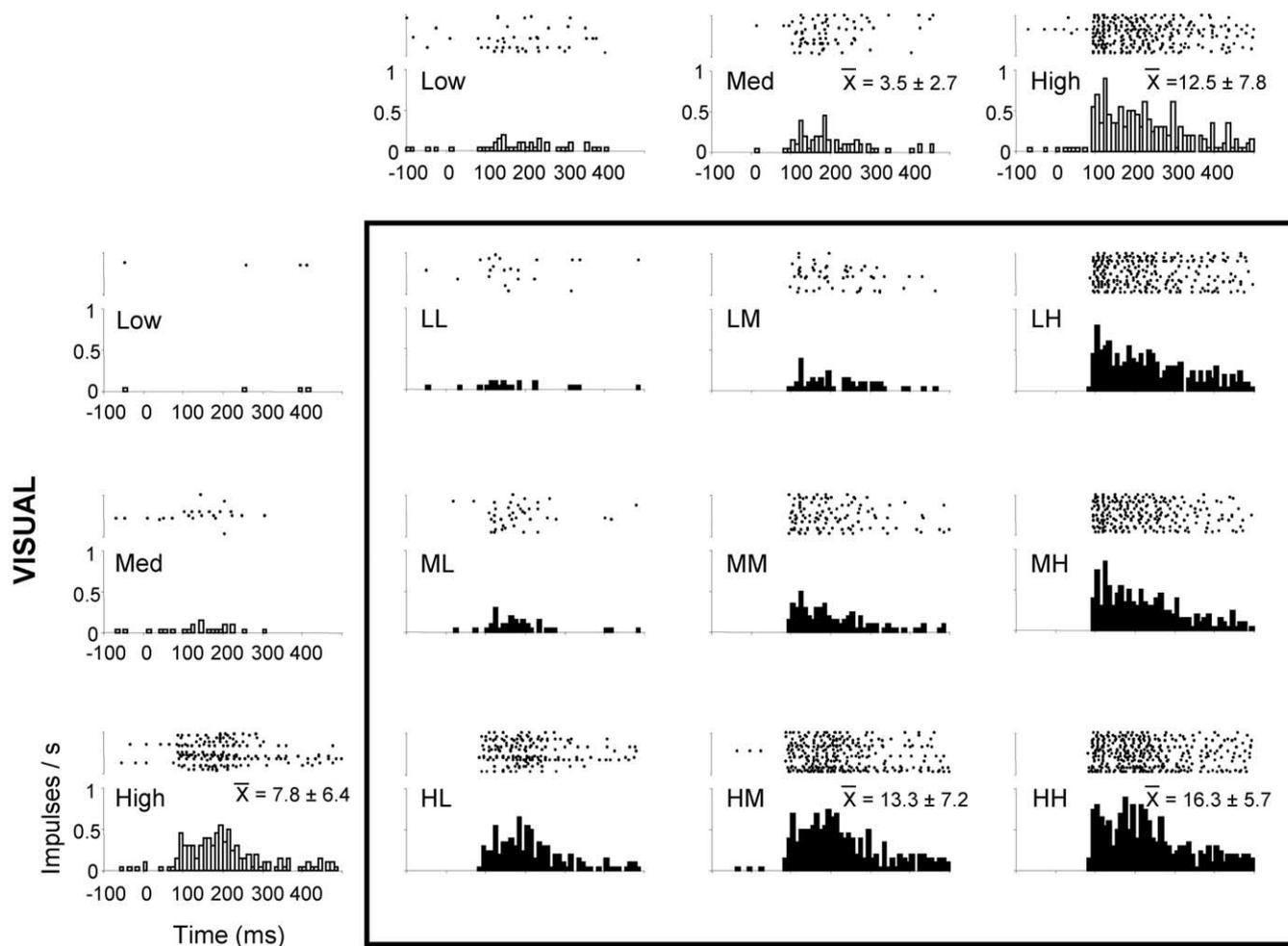


Figure 1. The responses of a single SC neuron to visual, auditory, and multisensory stimuli. Rasters and PSTHs compose a matrix with responses to three intensities (low, medium, high) of modality-specific visual (left column) and modality-specific auditory (top row) stimuli (open histograms). Responses to the nine possible multisensory combinations of these modality-specific stimuli fill in the cells of the matrix (filled histograms within box outline) according to the usual convention (e.g., histogram LL depicts the response to combination of the low-intensity visual and low-intensity auditory stimulus). All PSTHs are binned at a resolution of 10 ms and plotted on the same scales. Stimulus onset is at $t = 0$. Values corresponding to the mean number of impulses per trial and the SDs are given for selected histograms (see text for details). Impulse rates (y -axis) are computed as number impulses per bin width (10 ms) \times 100. Med, Medium.

when these stimuli were combined, the multisensory responses of the neuron ranged from relatively weak (e.g., histogram LL) to comparatively strong (histogram HH).

Multisensory integration can be assessed here qualitatively by comparing the multisensory responses (filled histograms) to those for the corresponding level-matched modality-specific stimuli. In some cases, integration was evident because the multisensory stimulus evoked activity greater than that evoked by either of the corresponding modality-specific stimulus components. For example, pairing of the high-intensity visual stimulus with the medium-intensity auditory stimulus (histogram HM) yielded an average of 13.3 ± 7.2 impulses per trial, which was greater than the response to either the corresponding visual (7.8 ± 6.4 impulses per trial) or auditory (3.5 ± 2.7 impulses per trial) stimulus presented alone.

As noted above, multisensory integration of excitatory influences has traditionally been quantified in the form of an enhancement index (E.I.) (Stein and Meredith, 1993) that relates the magnitude of the response to the multisensory stimulus to that evoked by the more effective of the two modality-specific stimu-

lus components. Indeed, according to the operational definition of multisensory integration (Stein and Meredith, 1993), a neuron is said to integrate excitatory influences only if the multisensory response exceeds the best unisensory response. Thus, in the above example, the mean multisensory response (13.3 impulses per trial) exceeded the best unisensory response (7.8 impulses per trial) by 5.5 impulses per trial, a net enhancement of 71% ($5.5/7.8 \times 100 = 70.5\%$). Although the measure of relative enhancement effectively gauges the benefit of adding a stimulus from a second modality in terms of increased activity, it does not provide any insight into the operation (in the mathematical sense) that a multisensory neuron performs on its unisensory inputs to generate that integrated product. Evaluating this operation was the main objective of the present study.

The approach taken can be summarized by, once again, referring to the example in Figure 1. Nominally, the multisensory response (13.3 spikes per trial) (Fig. 1, histogram HM) slightly exceeded a simple sum of the responses to the individually presented visual and auditory stimulus components (sum, 11.3 spikes per trial). On the other hand, a nominally subadditive

integration is shown in histogram HH (Fig. 1), in which the multisensory response of 16.3 spikes per trial exceeded the best unisensory response of 12.5 spikes per trial (visual, high) but fell short of the sum of the responses to the visual and auditory stimulus components ($12.5 + 7.8 = 20.3$ spikes per trial). Although both responses appear to have satisfied the criterion for multisensory integration (i.e., enhancement), the response to one stimulus pairing suggested superadditivity and that for another pairing suggested subadditivity, although it is unclear whether either of these responses is significantly distinct (in a statistical sense) from that predicted by simply summing the influences of the modality-specific stimulus components. Accordingly, for each neuron and for each stimulus combination, we evaluated the multisensory response against the null hypothesis of summation.

Generating the predicted multisensory response by summation

Figure 2 illustrates the procedure for generating a predicted multisensory response given simple summation of the unisensory inputs. The method is shown for a single stimulus pair for each of three multisensory neurons (Fig. 2A–C). These particular neurons, which differed considerably with regard to their responsiveness to both unisensory and multisensory stimulation, were chosen to be representative of the range of outcomes that was observed.

During data collection, multisensory trials were interleaved randomly, with an approximately equal number of level-matched unisensory visual and unisensory auditory trials for each cross-modal stimulus pairing (see Materials and Methods for details). Response rasters for each trial of visual, auditory, and multisensory stimulation are shown in order starting from the left of Figure 2 for each of the three example neurons (Fig. 2A–C).

To generate the predicted multisensory response for a given condition, a distribution of possible impulse counts was generated by computing the sum of counts for all possible combinations of the trials in which a visual or auditory stimulus was presented alone. For the neuron shown in Figure 2A, for example, this distribution contained the set of the 900 impulse counts derived from all possible combinations of the 30 visual and 30 auditory trials shown (note that the number of counts varied for different neurons according to the number of trials per condition; see Materials and Methods for details). Thus, each bin of the distribution (data not shown) represents the relative probability (given summation) of observing a particular number of impulses on a single multisensory trial. To generate a distribution of possible impulse

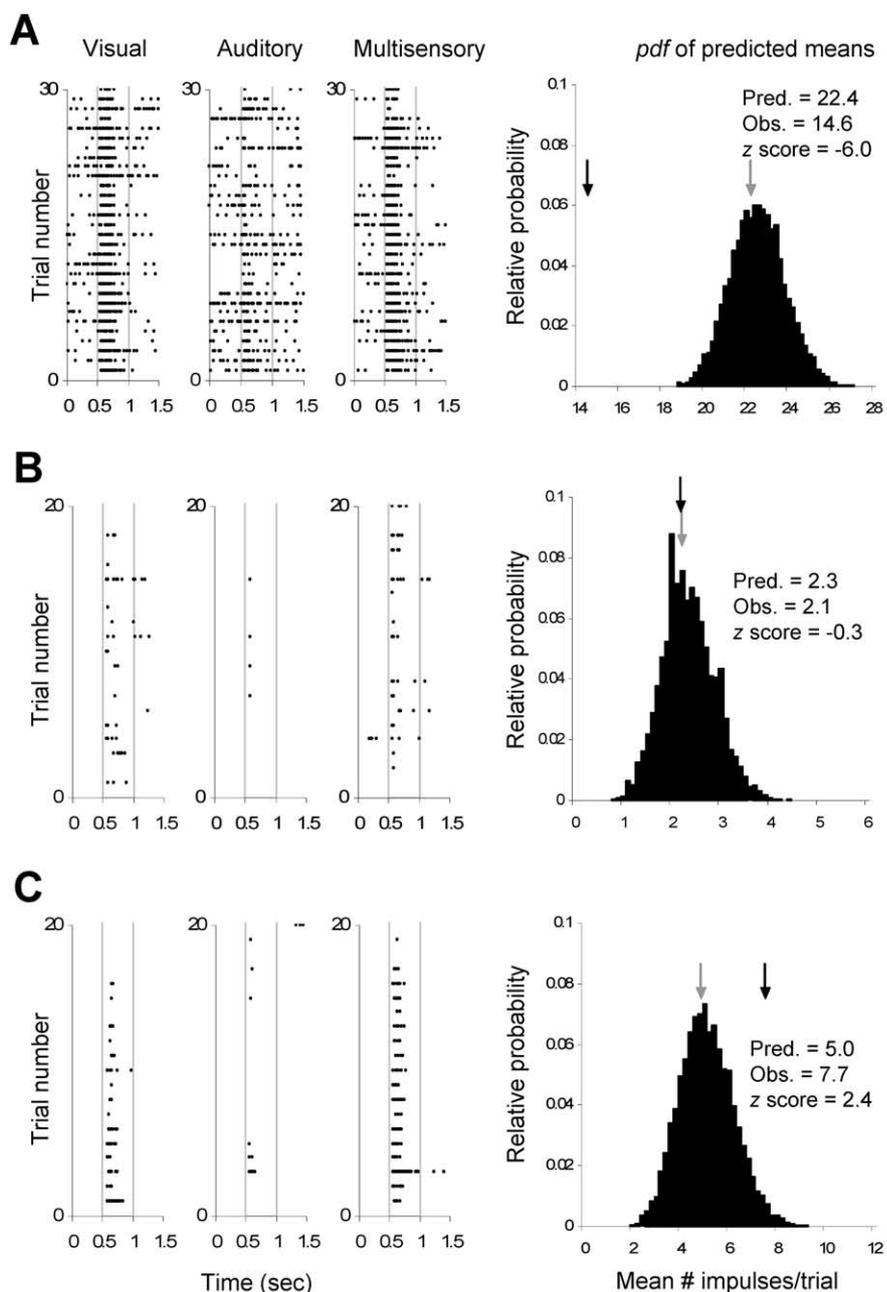


Figure 2. Generating the predicted multisensory sum. Visual, auditory, and multisensory responses and the distributions of predicted sums are shown for three different SC neurons (A–C). Left column, Rasters of impulses are shown for each trial of visual (left), auditory (middle), and visual–auditory multisensory (right) stimulation. In all cases, stimulus onset occurs at $t = 0.5$ s (first vertical gray line). Impulse counts were based on the 500 ms period from stimulus onset to $t = 1$ s (second vertical gray line). Right column, Probability density functions (pdf) of the predicted mean multisensory responses (mean number of impulses per trial) based on summing the modality-specific visual and auditory responses (see text for details). Gray arrow, Mean of the predicted distribution; black arrow, Mean of the actual multisensory response shown in the left column. The predicted mean, actual mean observed, and their difference expressed in SDs (z score) are shown for each example neuron. **A**, Superadditive interaction. **B**, Additive interaction. **C**, Superadditive interaction. Pred., Predicted mean; Obs., actual mean observed.

count means, this distribution was then randomly sampled (with replacement), each time drawing the same number of trials (n) used to compute the mean of the actual multisensory response (e.g., $n = 30$ trials for the neuron in Fig. 2A). With each draw of n trials, mean impulse count was computed and the process repeated until 10,000 samples were drawn to create the approximately normal distributions shown in the right column of Figure 2.

Having generated the expected values of the mean multisensory

sory response, one could then evaluate the probability that the mean of the observed multisensory response deviated from the prediction. For example, summation for the neuron shown in Figure 2A predicted a mean of 22.4 impulses per trial (gray arrow), whereas the actual multisensory response that was obtained was 14.6 impulses per trial (black arrow). The z score of this comparison was -6.0 , indicating that observing a mean this low when the underlying operation was actually summation was highly improbable ($p < 0.001$). Accordingly, this interaction was considered to be significantly subadditive.

The two relatively less responsive neurons shown in Figure 2, B and C, illustrate cases of additivity and superadditivity, respectively. For the neuron shown in Figure 2B, the predicted and actual means were nearly identical ($z = -0.3$), and the null hypothesis of additivity could not be rejected. The interaction was categorized as additive. In contrast, the multisensory response shown in Figure 2C exceeded the predicted mean by more than two SDs ($z = 2.36$) and was, therefore, categorized as significantly ($p < 0.01$) superadditive.

Returning to the examples shown in Figure 1 (histograms HM and HH), this method confirms that the nominally subadditive interaction shown in histogram HH is, in fact, significantly less than summation ($z = -2.3$), whereas the nominally superadditive response in histogram HM fails to reject the null hypothesis ($z = 1.2$) and is, therefore, additive.

Correction for spontaneous activity

Neurons in the SC vary substantially in their levels of spontaneous activity. For those without spontaneous activity, the estimates described above are valid. However, for spontaneously active neurons (Fig. 2A), a corrective subtraction was required to yield the predicted sum, because simply summing the unisensory impulse counts inappropriately doubles the contribution of spontaneous activity, thereby inflating the predicted multisensory response. The correction is of little consequence for neurons such as those depicted in Figure 2, B and C, but has a significant impact on the calculation of the predicted mean for a neuron such as that shown in Figure 2A, which has a high rate of spontaneous activity. Subtracting the erroneous spontaneous contribution adjusts the predicted mean multisensory response downward, in this case, from 22.4 to 19.5 impulses per trial. Although still significantly subadditive, the predicted response is now somewhat closer to the observed response (14.6 spikes per trial), a change that is reflected in the z score (uncorrected, -5.96 ; corrected, -3.79).

The principal effect of scaling down the predicted multisensory response by correcting for spontaneous activity was to eliminate spurious cases of subadditivity (decreased from 24.6 to 6.8%), with most of these cases now falling within the additive range. This is shown in Figure 3, A and B, which compares the uncorrected (open bars) and corrected (filled bars) z score distributions for all neurons and all stimulus combinations. Corrected or not, most multisensory response means fell within two SDs ($z \pm 1.96$; dotted lines) of the predicted sum. Accordingly, the mean z scores for the uncorrected and corrected distributions fell near zero at -0.4 ± 2.9 and 0.6 ± 2.4 , respectively. Adjusting for spontaneous activity did, however, lead to a substantial increase in the proportion of additive interactions (56.4–69.4%) and a slight increase (19.0–23.8%) in the proportion of superadditive interactions. The disproportionate effect on the subadditive tail of the distribution is clearly evident in the cumulative density functions (CDFs) of Figure 3B (corrected, filled squares; uncorrected, open circles).

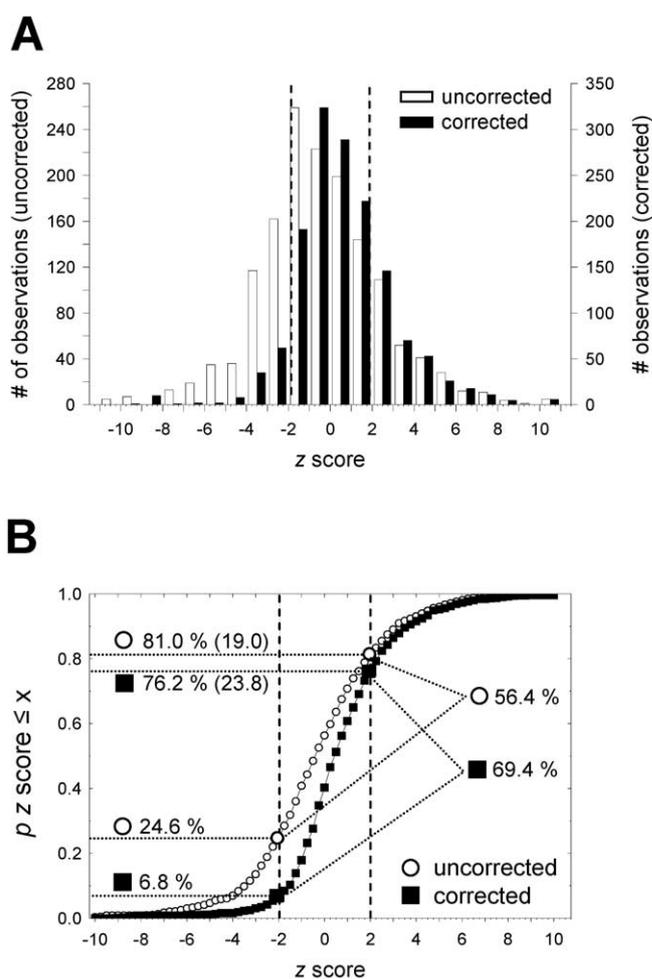


Figure 3. The majority of multisensory interactions are additive. **A**, Frequency distributions of z scores, including all stimulus combinations ($n = 1482$), are shown for all neurons ($n = 41$). Open bars plot z scores relating the actual multisensory responses to predicted sums that did not incorporate a corrective subtraction of spontaneous activity. Filled bars show the z scores relating the same multisensory responses to predicted sums that are corrected for spontaneous activity. **B**, Cumulative density functions of the same data shown in **A** illustrate the relative frequency of subadditivity, additivity, and superadditivity with (filled squares) and without (open circles) the correction for spontaneous activity. Vertical dotted lines at z score values of -1.96 and 1.96 indicate the points of transition from subadditivity to additivity (-1.96) and from additivity to superadditivity (1.96). For each neuron, a single estimate of mean spontaneous rate was obtained from the prestimulus epochs of all unisensory and multisensory trials. The rate was then converted to an impulse count by multiplying by the time interval over which impulses were counted (e.g., 500 ms) (see Fig. 2). The duplicated spontaneous term (see Results for details) was then eliminated by subtracting one-half of the number of spontaneous impulses from the visual and auditory responses before creating the distribution of all possible sums (see Fig. 2). Uncorrected for spontaneous activity, the relative percentages of subadditivity, additivity, and superadditivity were 24.6, 56.4, and 19%, respectively. Corrected for spontaneous activity, the relative percentages of subadditivity, additivity, and superadditivity were 6.8, 69.4, and 24.8%, respectively.

Qualifications for simple summation

Tested with a wide range of stimulus configurations, the majority (69.4%) of the multisensory stimulus combinations yielded responses consistent with simple summation of the modality-specific influences (Fig. 3A, B). We note, however, that a substantial number of these cases could have been the trivial consequence of there being no multisensory interaction at all. If, for example, one or both of the modality-specific component stimuli were ineffective in driving the inputs to the neuron, failure to reject the null hypothesis is not particularly meaningful. Indeed, Figure 1

provides several examples of a stimulus combination in which at least one of the modality-specific stimuli apparently failed to evoke postsynaptic activity.

By itself, failure to evoke postsynaptic activity is not sufficient evidence to conclude that a modality-specific input channel is inactive; it is well known that, when combined, subthreshold unisensory stimuli can produce a postsynaptic response. Thus, we also considered whether or not an apparently ineffective stimulus enhanced the response to a stimulus from the second modality. For example, in Figure 1, the low-intensity visual stimulus failed to meet either of these criteria; it did not evoke activity when presented alone and, when presented in combination with an auditory stimulus, (histograms LL, LM, LH), failed to enhance the response relative to that evoked by the auditory stimulus alone. Thus, in this case, the failure to activate one of the modality-specific channels precluded the possibility of a multisensory interaction.

To eliminate potentially trivial instances of additivity, we filtered the data set to include only multisensory stimulus conditions for which there was clear evidence of activation on both modality-specific input channels. As detailed above, inclusion was warranted if both of the modality-specific stimulus components, when presented during unisensory trials, evoked a response that exceeded spontaneous activity (one-tailed *t* test; $p < 0.05$). Failing that, inclusion was also justified whenever the multisensory response exceeded (one-tailed *t* test; $p < 0.05$) the response evoked by the more effective of the two modality-specific stimulus components (i.e., multisensory enhancement).

Consideration only of cases that met these more stringent inclusion criteria ($n = 870$) shifted the *z* score distribution toward more positive values, thereby yielding a greater proportion of superadditive interactions. The distribution, shown in Figure 4, distinguishes between cases in which multisensory enhancement was evident (filled bars) and those in which, despite activity on both channels, the multisensory response did not significantly exceed the activity evoked by the more effective modality-specific stimulus component (open bars). On the whole, the mean *z* score was 1.17 ± 2.6 (compare with 0.6 ± 2.4 in Fig. 3). Although the majority of interactions (59.4%) still failed to exceed summation, the incidence of superadditive interactions was substantially increased to 33.2% (compare with 23.8% in Fig. 3). The proportion of subadditive cases remained low at 7.4% (compare with 6.8% in Fig. 3).

Multisensory enhancement and the multisensory computation

As noted in the Introduction, previous studies have focused on multisensory enhancement and the benefit of combining stimulus modalities. Although the nature of the computation has not been evaluated quantitatively, many published examples of multisensory enhancement exceed 100% and thus are nominally superadditive interactions. In fact, the *z* score distribution shown in Figure 4 indicates that superadditivity is quite common among cases of multisensory enhancement. For this subset of cases (filled bars), the mean multisensory response is on average >2 SDs greater than the summation prediction (mean *z* score, 2.5 ± 2.3), with slightly more than one-half of all cases (52.7%) significantly superadditive. The tendency toward superadditivity for cases in which multisensory enhancement was observed is further illustrated in the scatterplot of Figure 5, which directly relates the observed multisensory response to the summation prediction for this subset. Accordingly, most points fall above the line of equality, indicating that the observed response typically exceeds the

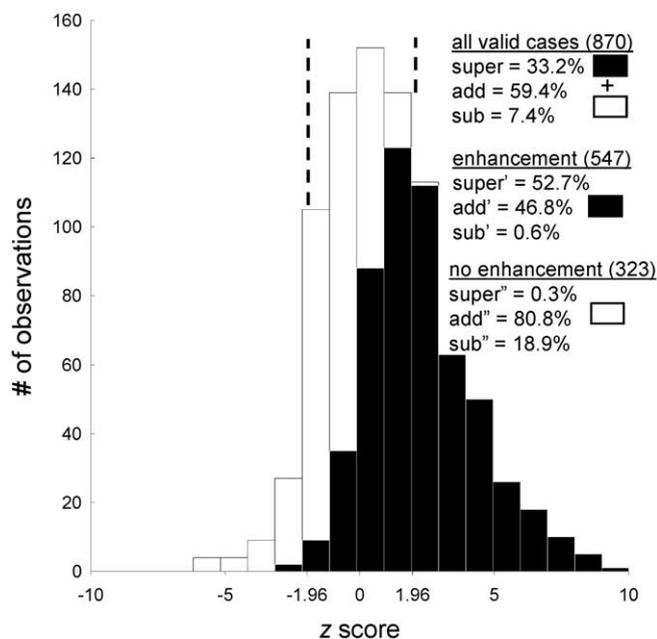


Figure 4. Imposing criteria for efficacy yields populations enriched for superadditive interactions. Shown are frequency distributions of *z* scores including only conditions that met either of two criteria for the efficacy of the modality-specific channels (valid cases; see Results for details). Filled bars are cases in which multisensory enhancement was observed (i.e., the multisensory response was greater than the greatest unisensory response). Unfilled bars are cases in which statistically significant activity was observed in response to each of the unisensory stimuli presented alone, but enhancement was not observed when they were presented in combination. Vertical dotted lines at *z* score values of -1.96 and 1.96 indicate the points of transition from subadditivity to additivity (-1.96) and from additivity to superadditivity (1.96). super, Superadditivity; add, additivity; sub, subadditivity.

summation prediction, with statistically significant instances of superadditivity indicated by filled circles.

Although multisensory enhancement was observed over a fairly wide range of predicted sums (i.e., 0–26 impulses per trial), the overwhelming majority of these cases were produced when weakly effective modality-specific stimulus components were combined. As shown in the inset cumulative density function (filled squares), fully 50% of the instances of multisensory enhancement corresponded to cases in which the predicted sums of the responses to the two modality-specific stimuli were three impulses per trial or fewer, with $\sim 90\%$ (485 of 547) contained in the range of predicted sums below 14 impulses per trial. Overall, the mean predicted response for this subset was 5.7 ± 6.1 impulses per trial and less than one-half the predicted sum for cases in which enhancement was not observed (12.4 ± 8.7 impulses per trial).

Accordingly, the efficacy (in terms of number of impulses evoked) of the modality-specific stimulus components of a cross-modal stimulus pair was also a strong predictor of the relative probability of observing superadditivity, additivity, and subadditivity. Figure 6A plots the relative incidence of superadditive (filled circles), additive (open circles), and subadditive (filled squares) multisensory interactions as a function of the magnitude of the predicted sum. Lighter, dotted line traces include all cases that met either of the two inclusion criteria described above ($n = 870$) (Fig. 4), whereas darker, solid line traces correspond to the subset of cases in which multisensory enhancement was verified ($n = 547$). In either case, as the modality-specific components of the multisensory stimulus become more effective (predicting a larger sum), the probability of superadditivity declines systemat-

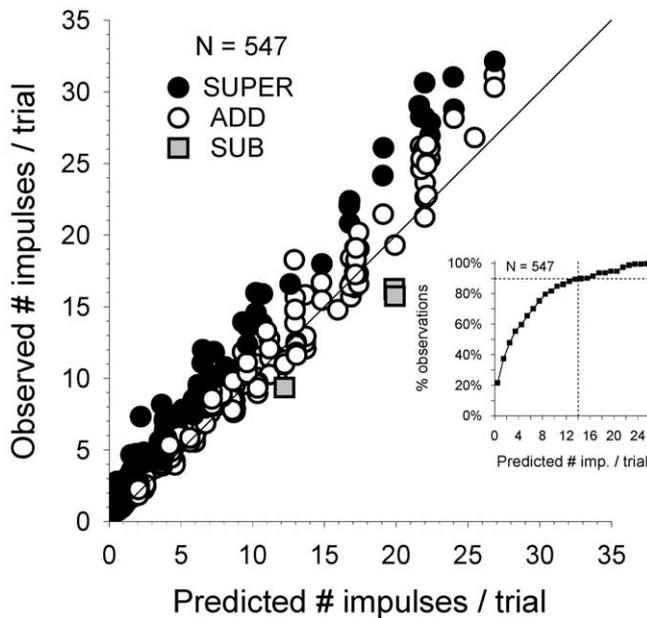


Figure 5. The observed versus the predicted multisensory response. A scatter plot relates the observed multisensory responses (mean number of impulses per trial) to the predicted sums (mean number of impulses per trial) for all cases in which multisensory enhancement was observed. Black circles and gray squares indicate statistically significant superadditivity (SUPER; $z > 1.96$) and subadditivity (SUB; $z < -1.96$), respectively. The open circles indicate cases of additivity (ADD). The solid diagonal line represents equality. Inset, A cumulative function of the same data illustrates that 90% of all observations correspond to a predicted sum of 14 impulses (imp.) per trial (dotted vertical lines on main panel and inset) or less.

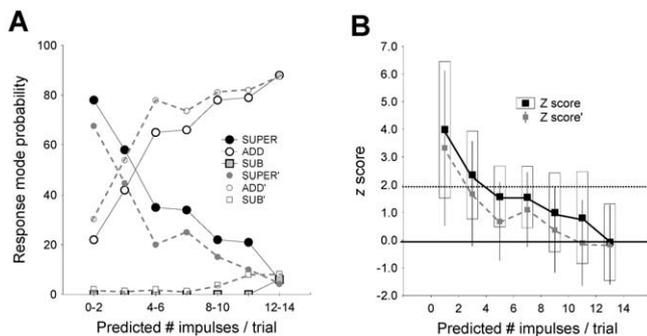


Figure 6. The mode of multisensory integration varies as an inverse function of stimulus efficacy. **A**, The plot shows the relative probability of observing subadditivity (SUB), additivity (ADD), and superadditivity (SUPER) as a function of the predicted sum magnitude (mean number of impulses per trial). The larger symbols connected by solid lines correspond to cases in which multisensory enhancement was observed ($n = 547$). The smaller symbols connected by dotted lines correspond to all valid cases (for details, see text and Fig. 4 legend). **B**, A plot of z score versus predicted sum (mean number of impulses per trial) for the same data shown in **A**. The horizontal dotted line at $z = 1.96$ marks the transition from additivity to superadditivity, and the solid line marks $z = 0$. SDs are shown as bars (enhancement) and open rectangles (valid cases).

ically and gives way to a steadily increasing incidence of additive interactions. Beyond 14 impulses per trial, there were too few data points to be binned at this resolution; however, when considered as an aggregate (impulses per trial, >14 ; $n = 203$) of the multisensory enhancement ($n = 58$) and nonenhancement groups ($n = 145$), there is evidence of a continuing trend: the incidence of superadditivity remains low (7.9%) and that for additivity remains high (69.5%) but gives way somewhat to increasing subadditivity (23.7%). In keeping with these trends, Figure 6B illustrates the corresponding declines in average z score

for both the population as a whole (gray symbols; dotted traces) and the multisensory enhancement subset (black symbols; solid traces).

The inverse trend observed in the population as a whole was also apparent in the integration of individual neurons. For 35 of the 41 neurons tested, it was possible (minimum of three valid cases) to perform a linear regression on plots of z score versus the predicted multisensory sum. For 29 of 35 (83%) of these neurons, the slope of the line of best fit was negative, with 20 of 29 (70%) of these fits revealing a significant negative correlation ($p < 0.05$). As a result, many neurons illustrated different multisensory response modes (e.g., superadditivity, additivity, subadditivity) across different stimulus combinations. Of the 35 neurons with multiple valid stimulus conditions (see inclusion criteria for Fig. 4), 31 (89%) showed more than one statistically defined response mode. Of these, seven (22.6%) showed all three modes of interaction, 18 (51.4%) demonstrated superadditivity and additivity, and six (19.4%) showed both additivity and subadditivity.

It is important to note, however, that individual neurons varied considerably in the degree to which this inverse relationship was expressed, and exceptions to this general trend were noted. For example, the response matrix in Figure 7 depicts a neuron for which the opposite trend was observed; the greatest superadditivity occurred for stimulus pairs that included either one or both of the most effective visual (medium intensity; middle row) and auditory (high intensity; right column) stimuli. This neuron was also unusual in that, for the visual modality, the relationship between stimulus intensity and stimulus efficacy was nonmonotonic such that the intermediate intensity (middle row; “med”) was the most effective in driving the neuron. Despite these anomalies, considered in the context of the population, this and similar neurons still contributed to the trends depicted in Figure 6, A and B. In relation to the population as a whole, the response of this neuron to the modality-specific stimulus components (auditory, top row; visual, left column) were quite weak, and the z scores, accordingly, were quite high (significantly superadditive for seven of the nine stimulus pairs). On average, the predicted sum was 1.8 ± 1.1 impulses per trial, and the z score was 3.0 ± 2.4 , values that fall in line with the population based plots of Figure 6B.

Effect of stimulus timing

It is self-evident that activity carried on modality-specific inputs must cooccur within a finite temporal window for a multisensory integration to occur (Meredith et al., 1987; Stein and Meredith, 1993). To provide the best opportunity for obtaining a representative sample of multisensory interactions for each stimulus configuration and for each neuron, the modality-specific components of the multisensory stimuli were presented at multiple SOAs. (The details of the timing effects will be presented in a follow-up report.) It is relevant to note here that variations in timing did not alter that general pattern of results (i.e., additivity always most prevalent) but did have an impact on the relative proportions of the three interaction types. To examine this, SOA was ranked from best to worst for each stimulus combination according to multisensory response magnitude. Because all stimulus combinations were presented with at least three SOAs, z scores were compared for the most effective and third most effective SOA. The effect of timing was manifest as increased likelihoods of superadditivity (first, 39%; third, 16%) and additivity (first, 57%; third, 75%) and a lower likelihood of subadditivity (first, 4%; third, 16%) for the optimal timing. Accordingly, the mean z score for the optimal timing configuration was signifi-

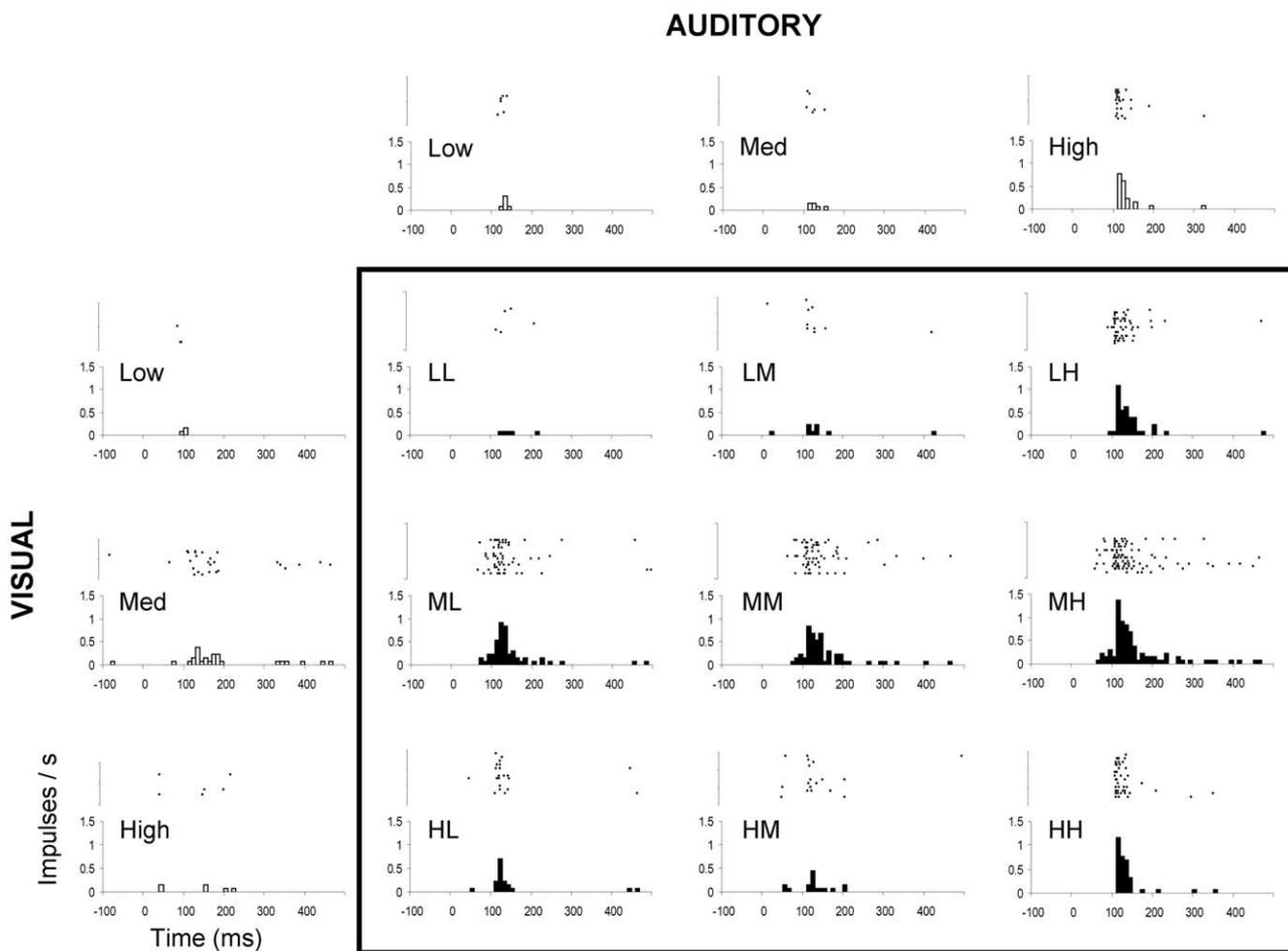


Figure 7. A response matrix illustrating the predominantly superadditive interactions for a weakly responsive SC neuron. All conventions are the same as for Figure 1.

cantly higher (first, 1.68 ± 2.64 ; third, 0.26 ± 2.14 ; $t = 25.5$; $df = 371$; $p \ll 0.01$).

Discussion

Despite the wealth of data documenting multisensory principles such as response enhancement and inverse effectiveness (for review, see Stein et al., 2004), there are little data bearing more directly on the specific mechanisms that give rise to such multisensory phenomena. This gap is primarily attributed to previous emphasis on the benefits of multisensory integration. The enhancement index is one manifestation of this focus in that it quantifies the difference between the multisensory response and the most vigorous modality-specific response. In previous studies, efforts to probe the limits of multisensory enhancement revealed that, proportionately, the largest increases from adding a stimulus of a second modality occurred when the two modality-specific stimulus components were themselves weakly effective. As such, the majority of extant data assess multisensory integration over a restricted range of input values at the presumed low end of the response range of the neuron.

Unlike previous studies, the current dataset evaluates the operation performed by multisensory neurons on their unisensory inputs and does so for a wide range of modality-specific input values. Our main finding is that the efficacy of the modality-specific inputs is the primary factor dictating whether the multisensory computation is superadditive, additive, or subadditive.

Superadditivity, the apparently nonlinear combination of the modality-specific inputs, is common within a narrow range of stimulus efficacies, being most prevalent when the modality-specific influences are very weak. With more effective stimuli, and over the majority of the responsiveness range displayed by the sample population, simple summation seemed a good approximation to the way in which multisensory neurons combine their inputs. Subadditivity was rare, occurring only for combinations of the most effective modality-specific stimuli.

Implications for underlying mechanism

That most multisensory interactions were consistent with simple summation of modality-specific influences suggests that a very basic linear model of the SC might account for many of the current observations. For example, the fact that superadditive interactions were common only in response to combinations of very weak modality-specific stimuli suggests that such superadditivity reflects temporal summation of the EPSPs consequent to near-threshold activity on the auditory and visual input channels. Subadditive interactions were relatively uncommon; however, their correspondence with combinations of the most effective stimuli could likewise be the simple consequence of approaching an intrinsic firing frequency limit of the SC neuron. In principle, a qualitatively similar relationship between computational mode and stimulus efficacy could be produced by an integrate-and-fire model that includes threshold and saturating nonlinearities. Us-

ing passive membrane properties alone, it should be possible to produce a transition from superadditivity through additivity and to subadditivity as the efficacy of the modality-specific stimuli driving the inputs to the multisensory neuron transition from minimally effective (subthreshold to near-threshold) to maximally effective (saturating).

Although appealing in its simplicity, ultimately, such a straightforward model may not capture the complexity and variety in the responses that have been observed among multisensory neurons in the SC. Although most individual neurons displayed evidence of the transition from one computational mode to another, and this is readily apparent in the population as a whole (Fig. 6), few displayed the transition in its entirety from superadditivity to subadditivity. The considerable variability in response type that we observed is clearly evident from comparing the neuron depicted in Figure 1 to that shown in Figure 7. Whereas the former demonstrates smooth and monotonic auditory and visual rate-intensity functions reminiscent of neurons in more primary sensory areas, the responses of the latter were modulated over a much more restricted range.

It is also important to note that a simple linear model cannot account for the apparently unique role of cortical inputs for affecting multisensory enhancement. It is believed that cat SC neurons are a primary site of interaction for both ascending and descending modality-specific inputs but that it is the cortical inputs deriving from the anterior ectosylvian sulcus and the rostral lateral suprasylvian cortex (Stein et al., 1983; Meredith and Clemo, 1989; Wallace et al., 1993) that are critical for multisensory enhancement to occur (Wallace and Stein, 1994; Jiang et al., 2001; Stein et al., 2002). Deactivation of one or both of these cortical areas appears to leave modality-specific responsiveness intact but eliminates multisensory enhancement and thus, by definition, rules out linear summation of the modality-specific inputs (Jiang et al., 2001). Additional evaluation of cortical deactivation data using methods comparable with those currently used will be necessary to fully consider such results in the present context.

Relationship to previous studies

As noted, most previous studies have quantified multisensory integration in terms of enhancement. Because computation of the enhancement index does not consider the contribution of the less effective of the two modality-specific stimuli, there is no direct relationship between the value it returns and the *z* score of the current classification scheme based on summation as the referent. For example, a multisensory response of 20 impulses and a best modality-specific response of 15 impulses would give an enhancement index value of 33% [$E.I. = (20 - 15)/15 \times 100 = 33\%$], whether the interaction that gave rise to the 20 impulses was superadditive (i.e., if the response to the other modality-specific stimulus was $\ll 5$ impulses), additive (other modality response, ≈ 5 impulses), or subadditive (other modality response, $\gg 5$ and ≤ 15 impulses). The only absolute relationship between the two measures is that enhancement values exceeding 100% are nominally superadditive. This is particularly relevant, because the results of this study provide an intuitive explanation for the preponderance of the apparently supralinear interactions (enhancements exceeding 100%) that have been observed in previous studies, an explanation that fits well with the related concept of inverse effectiveness.

With regard to previous studies, in which near-threshold stimuli were chosen to test the upper limits of multisensory enhancement, casual inspection suggests that the majority of pub-

lished examples of superadditivity fall in line with the functions shown in Figure 6, with most estimated to fall near the origins of the stimulus efficacy (i.e., predicted number of impulses) axes (Stein and Meredith, 1993). As originally defined by Stein and Meredith (1993), inverse effectiveness refers to the fact that, proportionately (i.e., expressed as a percentage of the best modality-specific response), the greatest multisensory response enhancements are observed for combinations of the weakest modality-specific stimuli. Again, although there is no unique relationship between enhancement index and the *z* score, it is intuitively obvious and analytically verifiable that inverse effectiveness is a manifestation of the observed trend from superadditivity toward subadditivity reported here (Fig. 6) and emphasizes that the benefits of multisensory integration for overt behavior are greatest when stimuli are weakest (Stein et al., 1989; Wilkinson et al., 1996; Jiang et al., 2002).

To our knowledge, there are two previous studies that have used summation as the referent for evaluating multisensory integration in SC neurons. However, direct comparison to our findings is not possible because the effect of stimulus efficacy was not systematically explored in either study. In a relatively early study, King and Palmer (1985) reported superadditive, additive, subadditive, and suppressive interactions for guinea pig SC neurons in response to combinations of visual and auditory stimuli. Although generally consistent with our findings, only a few examples are shown, and comparable analyses were not performed. More recently, Populin and Yin (2002) examined multisensory integration for neurons in the SC of the awake behaving cat and, using the sum of the responses to the modality-specific stimuli as the referent, reported that a large proportion of the interactions were sublinear when tested with stimuli that were fixed at specific suprathreshold values. Because neither the modality-specific nor multisensory response magnitudes was reported, it is not possible to know whether this sample would fall in line with the data shown in Figure 6. Furthermore, it is not clear whether, when computing the sum of the modality-specific responses, a correction for spontaneous activity was implemented (Fig. 3). Failure to do so overestimates the sum, thereby favoring subadditivity. Without estimates of spontaneous activity, it is not possible to determine if and how this factor influenced these previous findings.

Summary

These results represent the first systematic exploration of the operations underlying multisensory integration and provide a more mechanistic framework for considering the long-standing multisensory phenomena of enhancement and inverse effectiveness. With regard to the latter, it is known that neural (number of impulses) and behavioral (detection) responses to the weakest stimuli benefit the most from the addition of a stimulus from a second modality. It is clear from the current dataset that inverse effectiveness is a manifestation of the observed trend from superadditivity to subadditivity. From a practical perspective, the relationship between integration and stimulus efficacy illustrates that the nature of the multisensory interaction can be predicted reliably from the efficacy of the modality-specific stimulus components. Whether other multisensory phenomena (e.g., cortical dependence of multisensory integration) will prove to be consistent with equally simple mechanistic underpinnings remains to be determined.

References

- Jiang W, Wallace MT, Jiang H, Vaughan JW, Stein BE (2001) Two cortical areas mediate multisensory integration in superior colliculus neurons. *J Neurophysiol* 85:506–522.
- Jiang W, Jiang H, Stein BE (2002) Two corticotectal areas facilitate multisensory orientation behavior. *J Cogn Neurosci* 14:1240–1255.
- Kadunce DC, Vaughan JW, Wallace MT, Benedek G, Stein BE (1997) Mechanisms of within- and cross-modality suppression in the superior colliculus. *J Neurophysiol* 78:2834–2847.
- King AJ, Palmer AR (1985) Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp Brain Res* 60:492–500.
- Meredith MA, Clemo HR (1989) Auditory cortical projection from the anterior ectosylvian sulcus (field AES) to the superior colliculus in the cat: an anatomical and electrophysiological study. *J Comp Neurol* 289:687–707.
- Meredith MA, Stein BE (1983) Interactions among converging sensory inputs in the superior colliculus. *Science* 221:389–391.
- Meredith MA, Nemitz JW, Stein BE (1987) Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J Neurosci* 7:3215–3229.
- Populin LC, Yin TC (2002) Bimodal interactions in the superior colliculus of the behaving cat. *J Neurosci* 22:2826–2834.
- Quesy S, Sweatt A, Stein BE, Stanford TR (2000) The influence of stimulus intensity and timing on the responses of multisensory neurons in the superior colliculus: comparison to a model's prediction. *Soc Neurosci Abstr* 26:455.7.
- Sparks DL (1986) Translation of sensory signals into commands for control of saccadic eye movements: role of primate superior colliculus. *Physiol Rev* 66:118–171.
- Stein BE, Meredith MA (1993) *The merging of the senses*. Cambridge, MA: MIT.
- Stein BE, Spencer RF, Edwards SB (1983) Corticotectal and corticothalamic efferent projections of SIV somatosensory cortex in cat. *J Neurophysiol* 50:896–909.
- Stein BE, Huneycutt WS, Meredith MA (1988) Neurons and behavior: the same rules of multisensory integration apply. *Brain Res* 448:355–358.
- Stein BE, Meredith MA, Huneycutt WS, McDade L (1989) Behavioral indices of multisensory integration: orientation to visual cues is affected by auditory stimulation. *J Cogn Neurosci* 1:12–24.
- Stein BE, Wallace MW, Stanford TR, Jiang W (2002) Cortex governs multisensory integration in the midbrain. *Neuroscientist* 8:306–314.
- Stein BE, Jiang W, Stanford TR (2004) Multisensory integration in single neurons of the midbrain. In: *Handbook of multisensory integration* (Calvert G, Spence C, Stein BE, eds), pp 243–264. Cambridge, MA: MIT.
- Wallace MT, Stein BE (1994) Cross-modal synthesis in the midbrain depends on input from cortex. *J Neurophysiol* 71:429–432.
- Wallace MT, Meredith MA, Stein BE (1993) Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. *J Neurophysiol* 69:1797–1809.
- Wilkinson LK, Meredith MA, Stein BE (1996) The role of anterior ectosylvian cortex in cross-modality orientation and approach behavior. *Exp Brain Res* 112:1–10.