

Published in final edited form as:

Neurosci Biobehav Rev. 2014 April ; 41: 78–84. doi:10.1016/j.neubiorev.2013.12.003.

A model of the temporal dynamics of multisensory enhancement

Benjamin A. Rowland* and Barry E. Stein

Wake Forest School of Medicine, Winston-Salem, NC 27157, United States

Abstract

The senses transduce different forms of environmental energy, and the brain synthesizes information across them to enhance responses to salient biological events. We hypothesize that the potency of multisensory integration is attributable to the convergence of independent and temporally aligned signals derived from cross-modal stimulus configurations onto multisensory neurons. The temporal profile of multisensory integration in neurons of the deep superior colliculus (SC) is consistent with this hypothesis. The responses of these neurons to visual, auditory, and combinations of visual–auditory stimuli reveal that multisensory integration takes place in real-time; that is, the input signals are integrated as soon as they arrive at the target neuron. Interactions between cross-modal signals may appear to reflect linear or nonlinear computations on a moment-by-moment basis, the aggregate of which determines the net product of multisensory integration. Modeling observations presented here suggest that the early nonlinear components of the temporal profile of multisensory integration can be explained with a simple spiking neuron model, and do not require more sophisticated assumptions about the underlying biology. A transition from nonlinear “super-additive” computation to linear, additive computation can be accomplished via scaled inhibition. The findings provide a set of design constraints for artificial implementations seeking to exploit the basic principles and potency of biological multisensory integration in contexts of sensory substitution or augmentation.

Keywords

Multisensory; Cross-modal; Modeling; Temporal dynamics; Enhancement

1. Introduction

The evolution of multiple sensory systems has enhanced the likelihood of survival for organisms living in a wide variety of environments. This is not only because the senses substitute for one another when necessary, but because they can interact synergistically, thereby providing far more information about external events than would otherwise be possible. This is because the different senses are not corrupted by the same sources of noise, and combining their conditionally independent estimates of the same event yields a better

© 2013 Published by Elsevier Ltd.

*Corresponding author: Tel.: +1 3365779513. browland@wakehealth.edu, benjaminarowland@gmail.com (B.A. Rowland), bestein@wakehealth.edu (B.E. Stein).

Uncited references

Alvarado et al. (2007), Koch (1998) and Koch and Segev (1999).

analysis of its features (Ernst and Banks, 2002). This advantage manifests physiologically as enhancements in the speed and robustness of reactions to concordant cross-modal stimuli (Rowland et al., 2007a; Rowland and Stein, 2008), which in turn lead to faster and more accurate behavioral responses to the originating event (Meredith and Stein, 1983; Gielen et al., 1983; Perrott et al., 1990; Hughes et al., 1994; Frens et al., 1995; Wilkinson et al., 1996; Goldring et al., 1996; Jiang et al., 2002). Such enhancements are particularly beneficial when the information provided by the inputs is otherwise impoverished and/or unreliable; that is, circumstances in which their individual utilities are minimized (Stein and Meredith, 1993).

The best studied system in which this occurs is the mammalian superior colliculus (SC), which mediates the detection, localization, and orientation toward environmental targets (Meredith et al., 1987; Stein and Meredith, 1993). Individual neurons within the SC are sensitive to cues derived from different sensory modalities (e.g., vision, audition, and somatosensation) within circumscribed and overlapping regions of space (Stein and Arigbede, 1972). When stimulated by cross-modal cues within their respective receptive fields (RFs), their net evoked response magnitude (i.e., total number of impulses) is elevated above the response magnitude evoked by only one of the cues individually (“multisensory enhancement”). For robust stimuli, this enhancement typically reflects the sum of the net unisensory response magnitudes, but can be greater than this sum when the unisensory responses are less robust.

However, recent analyses examining the temporal profile of multisensory enhancement suggest that this enhancement is not uniform over the duration of the response (i.e., the entire discharge train). As the multisensory response rises and falls, its instantaneous firing rate (IFR) rarely reflects a simple addition of the component unisensory firing rates, even when the overall enhancement in the net response magnitude is consistent with an additive model (Rowland et al., 2007a). Rather, response enhancements are proportionally largest at the beginning of the response, which leads to earlier-than-expected response onsets (Rowland et al., 2007a; Rowland and Stein, 2008). The timing and magnitude of these multisensory enhancements, especially when occurring early in the discharge train, have the potential to greatly influence downstream circuits responsible for overt behavioral responses, as well as other targets involved in more higher-order perceptual processes. The operational principles of these neurons are a subject of great interest to basic scientists and researchers in applied domains seeking to engineer devices for sensory augmentation and substitution. However, most computational approaches to understanding multisensory integration in the SC have been restricted to describing its net products (e.g., Anastasio et al., 2000; Rowland et al., 2007b; Cuppini et al., 2010), not its moment-to-moment operations.

The purpose of this paper is to describe how the nonlinearities evident at the beginning of the multisensory response can be explained by a simple spiking model of SC multisensory integration, and do not require more complex assumptions about the biological substrate. At a coarse temporal resolution, the behavior of this model is similar to those described previously. However, at the level of resolution addressed here, the timing and “shape” of the inputs are revealed as key determinants of the integrated multisensory response. It thereby

makes the neurobiological computations underlying the multisensory response more explicit.

2. Results

2.1. Empirical observations

In multisensory SC neurons, concordant cross-modal signals typically evoke responses containing more impulses (i.e., enhanced net response magnitude), higher firing rates, longer durations, and shorter latencies than do their individual component stimuli (Stein and Meredith, 1993). The magnitude of the total multisensory response is generally related to the efficacy of the component stimuli: typically greater than the sum of these constituent unisensory response magnitudes when they are individually weak, and equal to their sum when they are more robust (Meredith and Stein, 1986). Fig. 1A and B provide typical examples from the cat SC. In Fig. 1A, both unisensory responses are very weak, engaging a net superadditive computation. In Fig. 1B, the unisensory responses are more robust, revealing a net additive integrative computation. These examples are consistent with the “principle of inverse effectiveness”, which specifies an inverse relationship between the proportional multisensory enhancement and the magnitude of the response to the most effective modality-specific component stimulus. However, despite the difference in the net products of these two examples, an examination of their temporal profiles (captured by the instantaneous firing rate and cumulative impulse count trace comparisons) reveals similarities in the underlying computational schematic. Importantly, both responses have aspects in which the multisensory response is more robust, and more dynamic, than an additive model would predict.

2.2. Enhancements in magnitude

On a moment-by-moment basis, most enhanced multisensory responses evidence “instantaneously” superadditive computations in portions of the response when the signals are weak or modest (typically at the beginning or end of one of the component responses), additive computations when they are modest, and sub-additive computations if and when they are very potent. The net computation evident in the overall response magnitude or firing rate reflects the sum of these linear and nonlinear instantaneous computations; thus, the difference between a net superadditive (Fig. 1A) and a net additive (Fig. 1B) product is related to not just the potency, but the relative incidence of superadditive, additive, and subadditive computations that took place during each.

2.3. Enhancements in timing

Most enhanced multisensory responses are more dynamic than their component unisensory responses. In both examples in Fig. 1, the multisensory response rises to its maximum firing rate more rapidly than predicted by the summed unisensory responses. However, it also descends from its peak earlier than expected, prior to the time at which the summed unisensory responses are expected to peak. Later, in its declining phase, the multisensory response can have a slower dynamic than expected, leading to a longer-than-expected response duration, although this result is variable across samples.

These examples illustrate the computational schematic that underlies the multisensory responses of most neurons in the SC, one in which the relationship between the unisensory and multisensory responses is neither linear nor strictly fixed in time. Clearly, the temporal alignment of the unisensory inputs (reflected in the unisensory responses) is a critical determinant in the instantaneous and overall products. Aligning weak or modest portions of the unisensory responses leads to large, typically superadditive enhancements, while aligning robust portions yields additive or subadditive enhancements. Multisensory responses change quickly when they are robust, and slowly when less robust. These underlying dynamics, only visible at a fine temporal resolution, can have profound consequences for the overall multisensory product achieved in any particular circumstance.

Fig. 2 illustrates the generality of these observations across a population of 324 samples of multisensory and unisensory responses recorded from the cat SC (a subset of the data originally published in Stanford et al., 2005). This analysis is restricted to multisensory neurons that are overtly responsive to brief presentations (50 ms) of visual and auditory stimuli individually and exhibit an enhanced response when the auditory stimulus is presented between 30 and 100 ms after the visual stimulus onset. This window of time typically yields the greatest likelihood for cross-modal interactions (Meredith et al., 1987). The restriction of the analysis to neurons overtly responsive to both modalities reduces the incidence of net superadditive observations relative to the incidence of the general population (see Perrault et al., 2005; Stanford et al., 2005).

Fig. 2A compares the net multisensory response magnitude to that of the largest component unisensory response (top) and the sum of the component responses (bottom). Due to the selection criteria, the net multisensory response is always enhanced relative to the largest unisensory component response, and is approximated by their sum throughout much of the range of tested efficacies. However, despite the appearance of a linear system suggested in the net response magnitudes, individual samples evidence moment-by-moment computations that are superadditive, additive, and subadditive. This is illustrated in Fig. 2B and C, which shows the average (mean and median) differences between the multisensory and summed unisensory IFRs as a function of the time (Fig. 2B) and percent of total multisensory response duration (Fig. 2C) (positive values indicate superadditivity). To generate these plots, all of the data samples were time-shifted to the expected onset of the multisensory response. The maximum enhancement occurs at or near this time (often due to the shorter multisensory response latency), and is typically superadditive. The computation is reduced to additivity within 30–50 ms (or 50% total response duration). The small increase in superadditivity evident at the end of Fig. 2C reflects the increased multisensory response duration relative to the unisensory. This trailing enhancement is averaged out in Fig. 2B due to the variability in response duration across the population.

Fig. 2D compares the times (from stimulus onset) at which the multisensory and summed unisensory response traces first cross one of two threshold criteria: the maximum IFR of the summed unisensory response (or multisensory if it is less, which is rare), and the rate that is 1/2 that maximum value. Negative values indicate that the multisensory response reaches the criterion at an earlier time; that is, the multisensory response trace is speeded relative to the summed unisensory trace. The cumulative frequency distributions of the time differences for

both crossing criteria are significantly biased to negative values, indicating the substantial speed enhancements evident in the multisensory response. Equivalent findings are observed using similar analytic methods, e.g., comparisons after normalizing by respective response magnitudes and comparisons of the mean cumulative impulse count functions, as the multisensory response begins earlier than expected based on the timing of the second response (Rowland et al., 2007a,b).

Fig. 2E summarizes findings pertaining to the timing of the peaks of the multisensory and summed unisensory firing rates (left) and a schematic of the magnitude and timing enhancements in the general population (right). All values reflect the population averages for each variable, but the traces are only for illustration (each value varies neuron by neuron). In the schematic, the multisensory response peaks 8 ms earlier than the summed unisensory trace. At the time of the multisensory peak, the multisensory response is actually 22 Hz larger than the summed unisensory trace, while by the time of the summed unisensory peak, it has decreased to be approximately equal. In other words, the computation has transitioned to approximately linearity.

These dynamics provide a set of general design constraints for systems seeking to replicate the biological principles of multisensory integration. The key dynamic common to almost all multisensory responses is the appreciable nonlinearity at the beginning of the response onset, which is our focus here. A priori, this observation provides evidence for a sort of multiplicative rule; e.g., one implemented by special synaptic configurations or computation within dendritic branches. However, as demonstrated below, a simpler alternative exists to explain it.

2.4. Model architecture

The nonlinearity at the beginning of the response can be described by an artificial neural network model that incorporates a minimal set of assumptions. This model provides a link between the observed physiology and what is believed to be its biological underpinnings. There are three key components to the model: (1) It receives and processes transient input signals that rise and fall, (2) the model unit (mimicking the multisensory SC neuron) contains a simple spiking dynamic, and (3) the unit is assumed to receive input from a number of spontaneously-activating afferents that produce random positive and negative fluctuations in its input current. This is a realistic assumption for SC neurons, which receive input from a large number of cortically-derived, subcortically-derived, and local excitatory and inhibitory sources (Baleydier et al., 1983; Behan and Appell, 1992; Berson, 1985; Berson and McIlwain, 1982; Blomqvist et al., 1978; Clemo and Stein, 1984; Edwards et al., 1979; Meredith and Clemo, 1989; Mize, 1983; Tortelly et al., 1980). For reasons of clarity, we explicitly model each sensory channel as providing a single input. As we have suggested elsewhere (cf. Cuppini et al., 2010), this input is best conceptualized as being largely sourced from unisensory areas of association cortex.

Thus, the input signal I for a modality-specific sensory source J is represented by a piecewise continuous function that begins rising from zero at time D according to a slope parameter λ_1 , peaks with value I_{\max} at time T , and subsequently decreases back to zero according to another slope parameter λ_2 .

$$\begin{aligned}
I_J(t) &= 0 & \text{for } t < D \\
I_J(t) &= I_{\max} \times \frac{1 - \exp(-(t-D)/\lambda_1)}{1 - \exp(-(T-D)/\lambda_1)} & \text{for } D < t < T \\
I_J(t) &= I_{\max} \times \exp\left(-\frac{t-T}{\lambda_2}\right) & \text{for } t > T
\end{aligned} \quad (1)$$

Fig. 3A illustrates an example input signal. The net input to the model unit at a given time t ($I(t)$) reflects the sum of the contributions from each sensory input channels and a noise current representing contributions from other afferents. Its value is selected from a Gaussian distribution with mean μ and standard deviation σ on each time step (represented by $N_{\mu,\sigma}(t)$).

The net input from a particular channel is presumed to incorporate both excitatory and inhibitory components. However, real inhibitory inputs are derived from local sources that receive input from integrating multisensory neurons, which will be engaging a nonlinear computation. This feature is incorporated in the model by an extra inhibitory term applied to multisensory conditions that is delayed with respect to the principal input. This does not affect the initial nonlinear computation (i.e., our principal focus), but mediates the output levels in the later portions of the response. For the sake of convenience, we introduce this component as term H (scaled by parameter h) applied at the peak time of the input. This component is only applied to multisensory conditions.

$$I(t) = \sum_J I_J(t) + N_{\mu,\sigma}(t) - H(t) \quad (2)$$

$$H(t) = \frac{h \times \prod_j I_j(t)}{I_{\max}} \quad \text{for } t > D, \text{ multisensory conditions only} \quad (3)$$

The multisensory SC neuron is modeled as a leaky integrate-and-fire unit with a membrane voltage differential (V) that responds to the net input according to time constant τ . In this model, an impulse is generated when the voltage exceeds the arbitrarily-selected value of 1 (0 is considered “rest”), after which V is clamped to 0 for R time steps to model a brief refractory period (Gabbiani and Koch, 1999).

$$\frac{\tau \Delta V}{\Delta t} = -V + I(t); \quad V(t) = 0 \quad \text{if } V(x) \geq 1 \text{ for any } x \text{ in } [t-R, t] \quad (4)$$

Numerical simulation of Eqs. (1)–(4) with reasonable parameter selections generates realistic-looking voltage traces, response rasters, and instantaneous firing rates (Fig. 3B and C). The links between the model architecture and the empirical results are explained below.

2.5. Model results

The dynamics of interest evident in the empirical data are that multisensory responses are initially more robust, rise faster, and peak earlier than the sum of their unisensory component responses. Fig. 4A illustrates these dynamics in the unisensory and multisensory responses produced by the model (10,000 simulations with fixed parameters), which are similar to the “mean schematic” presented in Fig. 2E. Plotted is the sum of the model’s IFR

for two identical unisensory responses (dashed line) versus the instantaneous firing rate of the multisensory response achieved by summing the input signals (solid line). The model's multisensory response rises earlier, reaches a higher maximum IFR earlier, and later transitions to an additive computation.

Fig. 4B illustrates the difference between the model's multisensory response and the sum of its unisensory responses as a function of time from response onset. This figure is analogous to that of Fig. 2B, with a parameter range appropriate for modeling the neuron in Fig. 1B (similar plots for the neuron in Fig. 1A are obtainable through different parameter selection). As observed empirically, the greatest superadditive nonlinearity is observed at the beginning of the response. Here, the difference trace begins at 0 owing to the simulation having a perfectly-controlled onset time. The rapid decrease in the nonlinearity and subsequent linear computation initiated at time = 50 ms indicates the contribution of the extra inhibitory component ($H(t)$). Note that this transition is quite a bit more rapid, and more severe, than one would expect in a biological circuit due to the simple and discrete method with which we have implemented the extra inhibition.

We now give intuitions for how the simple spiking model produces the early nonlinearity in the multisensory response (i.e., in the exclusion of the extra inhibition). The simple spiking model exhibits a nonlinear relationship between input magnitude and firing rate which can be appreciated via the relationship between input magnitude and the inter-spike interval (ISI) of the evoked response. The ISI is the reciprocal of the firing rate. For a simple leaky integrate-and-fire unit with time constant τ , refractory period R , threshold 1, and resting state 0, the ISI in response to constant input I is determined by solving Eq. (4) for t with $V(t) = 1$ and $V(0) = 0$.

$$ISI_I = R - \tau \ln \left[\frac{I-1}{I} \right] \quad (5)$$

More generally, the ISI in response to a constant input I to which another input of magnitude $a \times I$ has been added is:

$$ISI_{(1+a)I} = R - \tau \ln \left[\frac{(1+a)I-1}{(1+a)I} \right] \quad (6)$$

Given two observed responses with firing rates FR_1 and FR_2 corresponding to ISIs ISI_1 and ISI_2 , the predicted summed firing rate $FR_{1+2} = FR_1 + FR_2$ has an ISI determined by Eq. (7):

$$ISI_{1+2} = \frac{ISI_1 ISI_2}{ISI_1 + ISI_2} \quad (7)$$

Eq. (4) can be perceived as representing the unisensory ISI, Eq. (5) as representing the multisensory ISI (for $0 < a < 1$), and Eq. (6) as representing the ISI predicted by the summation of two unisensory response traces. Even though input currents are linearly summed in the model, this integration yields a nonlinear product owing to the refractory

period (which is a constant addition to ISI) and logarithmic transformation of the input current. Fig. 4C plots this relationship between the multisensory and summed unisensory firing rates for different levels of constant input, values of $a = 0.5$ and $a = 1$, while omitting the extra multisensory inhibition component H . The inset in the figure shows the relationship between the input magnitudes and the resulting firing rates for the multisensory ($a = 1$), unisensory, and summed unisensory conditions. The multisensory response is superadditive when the predicted sum is relatively low and transitions to additivity and subadditivity when it is higher. The potency of the multisensory enhancement is greater for unisensory inputs whose magnitudes are matched. The responses of the model to transient stimuli can also be appreciated from this relationship. For transient inputs, $I(t)$ rises from zero, peaks, and returns to zero. Thus, during the course of the response, the system effectively traverses across the x -axis of Fig. 4C (and inset), always passing through a phase of superadditivity and possibly reaching additivity and subadditivity, if the inputs are sufficiently robust. In this way, the model produces output patterns consistent with the rising phases of real neurons (e.g., Fig. 1).

The noise current in the model alters the normal relationship between the input and output firing frequency for the model unit (Fig. 4D, inset), and consequently alters the relationship of the multisensory and summed unisensory responses (Fig. 4D). Normally this relationship shows a sharp increase in firing rate for input values near threshold (i.e., slightly larger than 1). By contaminating the input current with random noise, this function becomes effectively smoothed around the threshold value according to the parameter σ . This causes some reduction in enhancement by smoothing the nonlinear relationship between firing frequency and input magnitude around the threshold. However, with higher levels of noise, rising input does not need to reach very high levels before evoking impulses from the model unit. Thus, the unit becomes responsive to transient signals while they are still in the most dynamic portions of their rising phases, which contributes to the speeding of the multisensory responses relative to unisensory responses. Thus, a given neuron subject to low noise currents will be likely to show large and sudden multisensory enhancements, while a neuron subject to high noise will show smaller enhancements that occur earlier in time.

3. Discussion

Below we summarize the relationship between the properties of the model, its underlying assumptions, and the key results.

A critical feature of the model is the amplification of the neuron's responsiveness to inputs that would not otherwise evoke impulses; that is, "stochastic resonance" (Benzi et al., 1981). This is provided by the noise current source as well as any input modalities whose values are (instantaneously) insufficient to generate impulses on their own. Because each input signal rises from an initial value of 0, for each multisensory sample there are times at which one (or more) of the signals is providing such an input, particularly when they are in their rising phases. Thus, at these times, the combined cross-modal signals are capable of eliciting overt activity earlier and at higher firing rates, leading to accelerated multisensory responses. An important perspective advanced here is that the nonlinearities observed in the magnitude and timing of the multisensory interactions do not explicitly require complicated assumptions

about the internal components of the unit model; rather, they are a natural product of dynamic signal processing in a model with simple biological constraints.

The noise current effectively smoothes the relationship between firing rate and input at low input levels, allowing the model unit to overtly respond to weaker inputs. One consequence is that neurons can begin responding to inputs at lower levels, and depending on the shape of the input function, this can further accelerate the multisensory response profile relative to that of the unisensory. However, another consequence is that it limits the sub- and peri-threshold range of interactions between two signals, reducing the potency of enhancement. Biologically, noise current distributions are expected to show significant interneuronal variability, a variability which is likely to be reflected by different spontaneous rates. The predicted consequence is that individual neurons can be stereo-typed according to whether they will produce robust or marginal multisensory enhancements relative to the larger population on the basis of their noise current distributions. Spontaneous firing rates provide a coarse measure of these values, and there is empirical evidence that neurons with high spontaneous firing rates evidence poorer multisensory integration capabilities (Perrault et al., 2005).

The extra inhibition provided in the multisensory conditions here is principally responsible for transitioning the neuron from a state of superadditivity to additivity and/or subadditivity depending on the scaling. This inhibition is intended to model the stronger competitive dynamics within the SC associated with cross-modal stimuli (Pluta et al., 2011; see also Fetsch et al., 2012). Parametric manipulations of this component are beyond the scope of the present study; however, they can greatly influence the temporal profile of the multisensory response and consequently the net product of integration. These dynamics vary across neurons, and are likely not to be strictly fixed even for a given neuron (for example, varying with stimulus features). Thus, they offer an intriguing possibility for theoretical and empirical future study.

The model also does not seek to incorporate slower-acting dynamics (e.g., associated with repeated stimulus presentations, see Yu et al., 2013), and the links between the abstract computations of the model and the neurobiological substrate are not rigorously supported by empirical findings, only suggested by them. Although the present model design is modular and self-contained, these factors should be considered when inserting it into larger, more powerful network architectures. Fortunately, the components specified here are simple and presumably inexpensive.

From a practical perspective, the model captures findings for the early temporal dynamics of multisensory enhancement that are critical to applications based on reverse-engineering the biological principles. One issue that commonly surfaces in these endeavors is that of calibration; that is, given a particular set of (artificially-generated) “cross-modal” signals with measured fidelity, what is the appropriate fused signal that should be produced? The analyses presented here suggest that the seemingly intricate nonlinearities in the relationship between multisensory and unisensory responses can be explained by relatively simple biologically-realistic mechanisms. It does not strictly require broader circuit dynamics other than those required to appropriately shape the input signals. The model introduces two

additional component, the noise currents to which the neuron is subject and extra inhibitory components associated with multisensory integration, as potential mechanisms to modulate the timing and efficacy of the integrative process. These components may vary with changes in the state of the neuron, circuit, or animal, and thereby produce changes in the potency of the products of integration that are contingent upon experience or circumstance.

References

- Alvarado JC, Vaughan JW, Stanford TR, Stein BE. Multisensory versus unisensory integration: contrasting modes in the superior colliculus. *J Neurophysiol.* 2007; 97(5):3193–3205. [PubMed: 17329632]
- Anastasio TJ, Patton PE, Belkacem-Boussaid K. Using Bayes' rule to model multisensory enhancement in the superior colliculus. *Neural Comput.* 2000; 12:1165–1187. [PubMed: 10905812]
- Baleydier C, Kahungu M, Mauguiere F. A crossed corticotectal projection from the lateral suprasylvian area in the cat. *J Comp Neurol.* 1983; 214(3):344–351. [PubMed: 6853761]
- Behan M, Appell PP. Intrinsic circuitry in the cat superior colliculus: projections from the superficial layers. *J Comp Neurol.* 1992; 315(2):230–243. [PubMed: 1372013]
- Benzi R, Suter A, Vulpiani A. The mechanism of stochastic resonance. *J Phys A.* 1981; 14:453–457.
- Berson DM. Cat lateral suprasylvian cortex: Y-cell inputs and corticotectal projection. *J Neurophysiol.* 1985; 53(2):544–556. [PubMed: 3981229]
- Berson DM, McIlwain JT. Retinal Y-cell activation of deep-layer cells in superior colliculus of the cat. *J Neurophysiol.* 1982; 47(4):700–714. [PubMed: 6279793]
- Blomqvist A, Flink R, Bowsher D, Griph S, Westman J. Tectal and thalamic projections of dorsal column and lateral cervical nuclei: a quantitative study in the cat. *Brain Res.* 1978; 141(2):335–341. [PubMed: 626905]
- Clemo HR, Stein BE. Topographic organization of somatosensory corticotectal influences in cat. *J Neurophysiol.* 1984; 51(5):843–858. [PubMed: 6726314]
- Cuppini C, Ursino M, Magosso E, Rowland BA, Stein BE. An emergent model of multisensory integration in superior colliculus neurons. *Front Integr Neurosci.* 2010; 4:6. [PubMed: 20431725]
- Edwards SB, Ginsburgh CL, Henkel CK, Stein BE. Sources of subcortical projections to the superior colliculus in the cat. *J Comp Neurol.* 1979; 184(2):309–329. [PubMed: 762286]
- Ernst MO, Banks MS. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature.* 2002; 415:429–433. [PubMed: 11807554]
- Fetsch CR, Pouget A, DeAngelis GC, Angelaki DE. Neural correlates of reliability-based cue weighting during multisensory integration. *Nat Neurosci.* 2012; 15:146–154. [PubMed: 22101645]
- Frens MA, Van Opstal AJ, Van der Willigen RF. Spatial and temporal factors determine auditory–visual interactions in human saccadic eye movements. *Percept Psychophys.* 1995; 57:802–816. [PubMed: 7651805]
- Gabbiani, F.; Koch, C. Principles of spike train analysis. In: Koch, C.; Segev, I., editors. *Methods in Neuronal Modeling*. MIT Press; MA: 1999. p. 313–360.
- Gielen SC, Schmidt RA, van den Heuvel PJ. On the nature of intersensory facilitation of reaction time. *Percept Psychophys.* 1983; 34:161–168. [PubMed: 6634374]
- Goldring JE, Dorris MC, Corneil BD, Ballantyne PA, Munoz DP. Combined eye–head gaze shifts to visual and auditory targets in humans. *Exp Brain Res.* 1996; 111:68–78. [PubMed: 8891638]
- Hughes HC, Reuter-Lorenz PA, Nozawa G, Fendrich R. Visual–auditory interactions in sensorimotor processing: saccades versus manual responses. *J Exp Psychol Hum Percept Perform.* 1994; 20:131–153. [PubMed: 8133219]
- Jiang W, Jiang H, Stein BE. Two corticotectal areas facilitate multisensory orientation behavior. *J Cogn Neurosci.* 2002; 14:1240–1255. [PubMed: 12495529]
- Koch, C. *Biophysics of Computation: Information Processing in Single Neurons*. Oxford University Press; New York: 1998.
- Koch, C.; Segev, I. Oxford University Press; New York: 1999.

- Meredith MA, Clemo HR. 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.* 1989; 289(4):687–707. [PubMed: 2592605]
- Meredith MA, Stein BE. Interactions among converging sensory inputs in the superior colliculus. *Science.* 1983; 221:389–391. [PubMed: 6867718]
- Meredith MA, Stein BE. Spatial determinants of multisensory integration in cat superior colliculus neurons. *J Neurophysiol.* 1986; 75:1843–1857. [PubMed: 8734584]
- Meredith MA, Nemitz JW, Stein BE. Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J Neurosci.* 1987; 7:3215–3229. [PubMed: 3668625]
- Mize RR. Patterns of convergence and divergence of retinal and cortical synaptic terminals in the cat superior colliculus. *Exp Brain Res.* 1983; 51(1):88–96. [PubMed: 6884465]
- Perrault TJ Jr, Vaughan JW, Stein BE, Wallace MT. Superior colliculus neurons use distinct operational modes in the integration of multisensory stimuli. *J Neurophysiol.* 2005; 93(5):2575–2586. [PubMed: 15634709]
- Perrott DR, Saberi K, Brown K, Strybel TZ. Auditory psychomotor coordination and visual search performance. *Percept Psychophys.* 1990; 48:214–226. [PubMed: 2216648]
- Pluta SR, Rowland BA, Stanford TR, Stein BE. Alterations to multisensory and unisensory integration by stimulus competition. *J Neurophysiol.* 2011; 106:3091–3101. [PubMed: 21957224]
- Rowland BA, Stein BE. Temporal profiles of response enhancement in multisensory integration. *Front Neurosci.* 2008; 2(2):218–224. [PubMed: 19225595]
- Rowland BA, Quessy S, Stanford TR, Stein BE. Multisensory integration shortens physiological response latencies. *J Neurosci.* 2007a; 27(22):5879–5884. [PubMed: 17537958]
- Rowland BA, Stanford TR, Stein BE. A model of the neural mechanisms underlying multisensory integration in the superior colliculus. *Perception.* 2007b; 36 (10):1431–1443. [PubMed: 18265826]
- Stanford TR, Quessy S, Stein BE. Evaluating the operations underlying multisensory integration in the cat superior colliculus. *J Neurosci.* 2005; 25:6499–6508. [PubMed: 16014711]
- Stein BE, Arigbede MO. Unimodal and multimodal response properties of neurons in the cat's superior colliculus. *Exp Neurol.* 1972; 36(1):179–196. [PubMed: 4558413]
- Stein, BE.; Meredith, MA. *The Merging of the Senses.* MIT Press; Cambridge, MA: 1993.
- Tortelly A, Reinoso-Suarez F, Llamas A. Projections from non-visual cortical areas to the superior colliculus demonstrated by retrograde transport of HRP in the cat. *Brain Res.* 1980; 188(2):543–549. [PubMed: 7370774]
- Wilkinson LK, Meredith MA, Stein BE. The role of anterior ectosylvian cortex in cross-modality orientation and approach behavior. *Exp Brain Res.* 1996; 112:1–10. [PubMed: 8951401]
- Yu L, Rowland BA, Xu J, Stein BE. Multisensory plasticity in adulthood: cross-modal experience enhances neuronal excitability and exposes silent inputs. *J Neurophys.* 2013; 109(2):464–474.

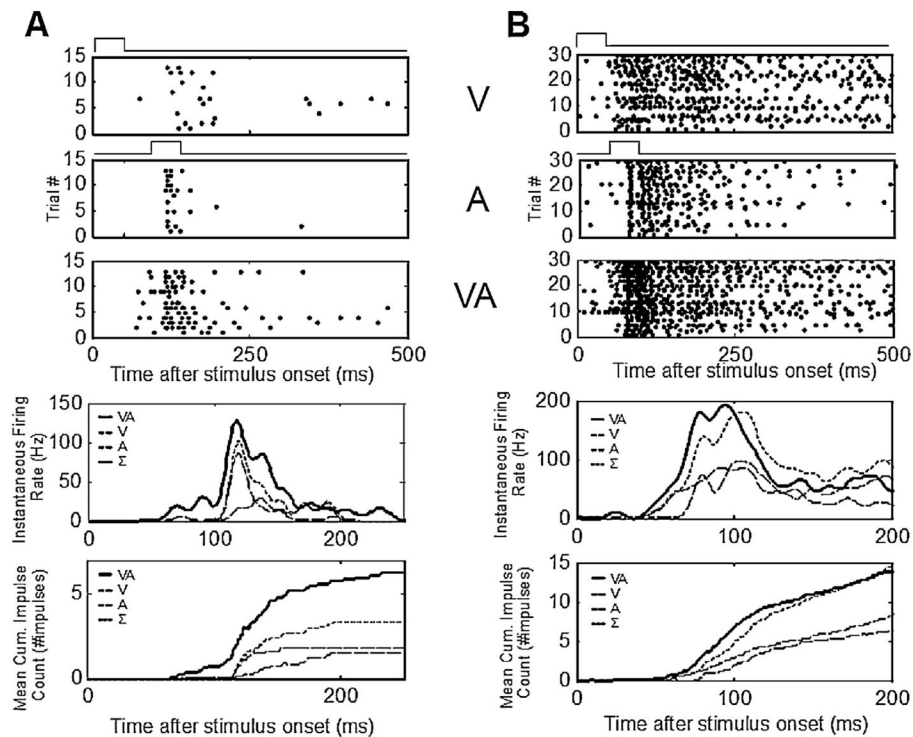
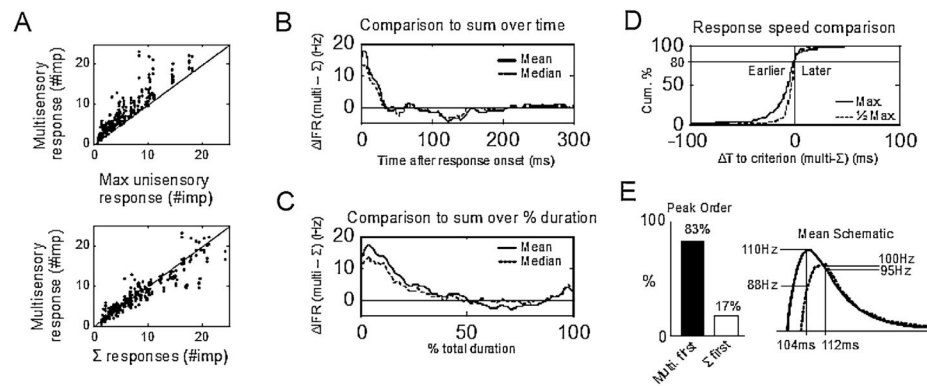


Fig. 1.

Unisensory and multisensory responses from two exemplar neurons in the cat SC. Both neurons overtly respond to visual (V) and auditory (A) stimuli presented in isolation and in combination (VA). Top plots indicate the traces and responses corresponding to each tested stimulus condition (dots indicate impulses). The middle plots illustrate the instantaneous firing rate (calculated by convolving the impulse train with a Gaussian kernel, 5 ms standard deviation) traces for each stimulus condition as well as the sum of the unisensory traces (dotted line). The scale has been changed to provide better visualization of the behavior near response onset. Bottom plots illustrate the mean cumulative impulse count, a running tally of the number of impulses elicited by each stimulus on or before each moment in time (i.e., without any smoothing or transformation). In the exemplar in (A), the total multisensory response magnitude is statistically greater ($p < 0.05$) than the sum of the two unisensory response magnitudes. In the exemplar in (B), it is not statistically different than the predicted sum. However, both show similar nonlinear computations at the beginning of the response.

**Fig. 2.**

Empirical findings of nonlinear enhancements in magnitude and timing in a population of multisensory SC neurons. (A) Comparison of the multisensory response magnitude (mean # stimulus-elicited impulses) to the largest component unisensory response magnitude (top) and to the sum (Σ) of the unisensory response magnitudes (bottom). (B) and (C) Average difference between the multisensory and summed unisensory instantaneous firing rates for all traces in the population, synchronized to the onset of the multisensory response (solid line = mean, dashed line = median) and plotted as a function of time (B) and total multisensory response duration (C). (D) Speed comparisons between the multisensory and summed unisensory responses. Plotted is the cumulative frequency distribution for the time difference between when the multisensory and summed unisensory responses first crossed each of two criterion levels: the peak firing rate of the weaker of the two (max) and a firing rate 1/2 in magnitude (1/2 max). Negative sample values indicate a faster multisensory response. (E) Left: bar graph comparing the relative incidence of multisensory responses peaking before the peak of the summed unisensory responses (multi. first) versus the reverse (Σ first). Right: schematic of the mean values for peak firing times and rates at those times for the population. Time values are relative to visual stimulus onset.

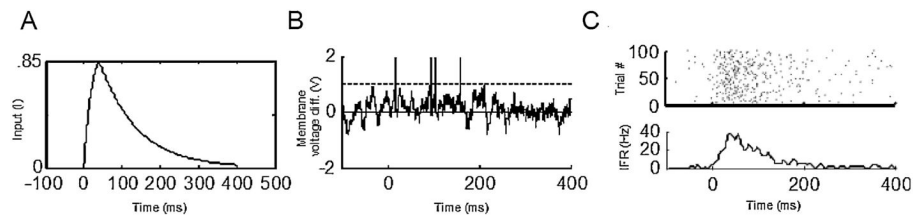
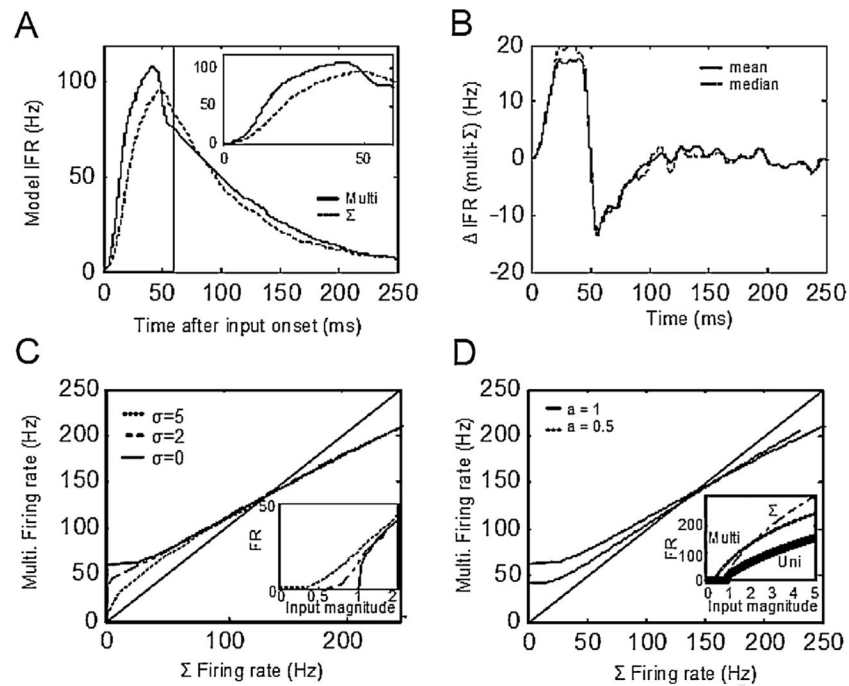


Fig. 3.

Model architecture. (A) An example of an input trace that begins rising at time 0 and peaks at $t = 40$ ms ($D = 0$, $\lambda_1 = 20$ ms, $I_{\max} = 1$, $T = 40$ ms, $\lambda_2 = 100$ ms). (B) An example of the membrane potential over time in a single trial of a neuron responding to the stimulus in (A), contaminated with a noise current having parameters ($\mu = 0$, $\sigma = 5$). Vertical lines are drawn at times where action potentials are generated. The dotted line indicates the unit's threshold. (C) Top: an impulse raster of 100 simulations having the same parameters as (B). Bottom: instantaneous firing rate of the above raster using a Gaussian kernel with 5 ms standard deviation.

**Fig. 4.**

Model results. (A) Example multisensory (solid line) and summed unisensory (dotted line) response traces from a simulated multisensory neuron ($D = 0$, $\lambda_1 = 20$ ms, $I_{\max} = 1$, $T = 40$ ms, $\lambda_2 = 100$ ms, $\mu = 0$, $\sigma = 5$, $\tau = 10$ ms, $R = 2$ ms, $h = 0.5$). The two unisensory signals contributing to the multisensory response were assumed to be exactly equal. The inset expands the activity in the window 0–50 ms. The multisensory response begins earlier, rises faster, and peaks earlier at a higher magnitude than predicted by the summed unisensory traces. (B) The difference between the multisensory and summed unisensory inputs for multiple simulations (solid = mean, dashed = median) with the parameters of (A). (C) Relationship between the multisensory and summed unisensory firing rates for model units ($\mu = 0$, $\sigma = 5$, $\tau = 10$ ms, $R = 2$ ms, $h = 0$) responding to constant input ($0 < I < 5$). The line of unity indicates an additive multisensory response. The two curves illustrate the relationship for circumstances in which the two unisensory inputs are equal in magnitude ($a = 1$, solid line) one input is 50% as strong as the other ($a = 0.5$). The inset shows the relationship between the firing rates for the multisensory (multi, with $a = 1$), summed unisensory (Σ), and unisensory (Uni) responses and the input magnitude. (D) The relationship between firing rate and input magnitude for three different values of the noise current standard deviation (σ).