A New Measure of Multisensory Integration in a Single Neuron Based on Coupling

Hans Colonius∗
Department of Psychology
Oldenburg University
26111 Oldenburg, Germany
hans.colonius@uni-oldenburg.de

Adele Diederich†
Psychology
Jacobs University Bremen
28759 Bremen, Germany
a.diederich@jacobs-university.de

Abstract

A single neuron is categorized as “multisensory” if there is a statistically significant difference between the response evoked by a cross-modal stimulus combination and that evoked by the most effective of its components individually. The most widely applied quantitative index expresses multisensory enhancement (or inhibition) as a proportion of the strongest unisensory response. However, it has no theoretical foundation in terms of the possible operations a neuron may perform in combining unisensory inputs to yield the multisensory response. In particular, being responsive to multiple sensory modalities does not guarantee that a neuron has actually engaged in integrating its multiple sensory inputs rather than simply responding to the most salient stimulus. Here, a new index is proposed based on a dependent probability summation mechanism. It compares the mean observed cross-modal response of a neuron with the largest cross-modal mean achievable by stochastically coupling its unisensory responses. Because this new measure is, in general, more restrictive than the traditional one, many neurons previously categorized as “multisensory” may possibly lose that property. The approach is illustrated by data from a single cat superior colliculus neuron. Computation of the new index is straightforward and it is as amenable to statistical testing as the traditional one. Moreover, it is completely analogous to a widely-used procedure in behavioral studies of multisensory integration (“race model inequality”).

1 Defining and measuring multisensory integration

Single neurons in the deep layers of the mammalian superior colliculus (SC) have been shown to integrate afferent visual, auditory, and somatosensory cues and to generate efferent motor commands to structures innervating the musculature of, e.g., the eyes and hands [1, 2]. From a wealth of neurophysiological studies, some general “rules” for multisensory integration (MI) have emerged. Cross-modal cues (e.g., visual-auditory) that are in close spatial and temporal register enhance the response of multisensory neurons implementing the “spatial and temporal rules”, whereas those that are spatially or temporally disparate often elicit response depression or fail to be integrated [3] (see also [4]). A third principle refers to the observation that proportionately greater effects of cross-modal cues are obtained when those individual cues are weakly effective. Thus, the magnitude of multisensory integration is inversely related to the efficacy of the stimuli being integrated (“inverse effectiveness rule”) [5]. Moreover, those rules are closely mirrored by findings of behavioral studies measuring (saccadic or manual) reaction time [6]. Whereas empirical results are, at least in general, unequivocal regarding these rules, the issue of exactly how to measure “multisensory integration”...
has been under debate for some time [7]. The purpose of this paper is to suggest a novel, theoretically founded index to measure MI in a single neuron.

At the level of a single (SC) neuron, response strength has traditionally been defined by the absolute number of impulses (spikes) registered within a fixed time interval after stimulus presentation, or by the firing rate within this interval. Operationally, a neuron is categorized as “multisensory” if there is a statistically significant difference between the response evoked by a cross-modal stimulus combination and that evoked by the most effective of its components individually [2]. Moreover, if a neuron responds, for example, to visual but not to auditory stimulation and if the response to a visual-auditory combination differs significantly from the response to the visual stimulus, it is also considered “multisensory”. The most widely accepted descriptive measure of the magnitude of multisensory integration is the multisensory integration index (also termed interaction index) (MSI) defined as

$$\text{MSI} = \frac{\text{CM} - \text{SM}_{\text{max}}}{\text{SM}_{\text{max}}} \times 100$$

where CM is the mean number of impulses in response to the cross-modal stimulus and SM_{max} is the mean number of impulses to the most effective (modality-specific) component stimulus [8]. Thus, MSI expresses multisensory enhancement as a proportion of the strongest unisensory response.

Some minor modifications of MSI have been discussed as well (e.g., [9]). One alternative measure is the “additive model” suggesting to replace SM_{max} in equation 1 by the sum of the unisensory responses [10]. This measure has raised some controversy; in particular, it has been argued that any cross-modal response larger than the largest unisensory response but smaller than the sum might then be misinterpreted as response depression [7].

A main weakness of MSI as measure of multisensory integration, however, is that it has no theoretical foundation in terms of the possible operations SC neurons may perform in combining their unisensory inputs to yield the multisensory response [11]. Being responsive to multiple sensory modalities does not guarantee that a neuron has actually engaged in integrating its multiple sensory inputs rather than simply responding to the most salient stimulus; or, as Stein and colleagues [7] (p. 114) have put it, “At the time of the early physiology studies in the 1980s, it was considered possible that these neurons only represented a common route by which independent inputs from a variety of senses could gain access to the same motor apparatus in generating behavior (e.g., possibly employing a ‘winner-take-all’ algorithm).”

The occurrence of MI has traditionally been defined as “inputs from two or more senses are combined to form a product that is distinct from, and thus cannot be easily deconstructed to reconstitute the components from which it is created” [2]. As of now, those multisensory computations performed by SC neurons are still not fully understood [12]. However, in view of the above definition, and even in the absence of detailed knowledge about those processes, here we show that it is feasible to define a multisensory index that is theoretically founded. Simply put, this new index, which is based on a dependent probability summation mechanism, compares the mean observed cross-modal response of a neuron with the largest cross-modal mean achievable by stochastically coupling its unisensory responses. Because this new measure is, in general, more restrictive than the traditional MSI, many neurons previously categorized as “multisensory” may possibly lose that property.

2 Deriving the new index of multisensory integration

To fix ideas, let NV, NA, and NV,A denote the random number of impulses emitted in a given time interval by a neuron, following unisensory (visual and auditory) and cross-modal stimulation, respectively, without assuming any specific parametric distributions. Their expected values (more precisely, sample averages) would be inserted in the traditional MSI value (Equation 1)

$$\text{MSI} = \frac{\text{E}N_{V,A} - \max\{\text{E}N_V, \text{E}N_A\}}{\max\{\text{E}N_V, \text{E}N_A\}} \times 100$$

where E stands for expected value.

Realizations of random variables NV and NA, with distribution functions FV and FA, respectively, are collected in different experimental trials, that is, under different experimental stimulus conditions. Therefore, they refer to distinct probability spaces and, a-priori, there is no natural way to
combine the results from visual and auditory trials. In particular, any assumption about stochastic (in-)dependence between \( N_V \) and \( N_A \) is void. Nevertheless, it is always possible to define a stochastic coupling of the two random variables (see [13]). Coupling of \( N_V \) and \( N_A \) means defining a distribution \( H_{V,A} \) for a bivariate random vector \((N_V, N_A)\) in such a way that its marginal distributions are identical to \( F_V \) and \( F_A \).

Let \( H_{V,A}(m, n) = \Pr(\hat{N}_V \leq m, \hat{N}_A \leq n) \), \( m, n = 0, 1, \ldots \), be some coupling of \( N_V \) and \( N_A \). As a bivariate (discrete) distribution, it obeys the Fréchet inequality (14):

\[
\max\{0, F_V(m) + F_A(n) - 1\} \leq H_{V,A}(m, n) \leq \min\{F_V(m), F_A(n)\},
\]

for all \( m, n = 0, 1, \ldots \). Setting \( m = n \), we get

\[
H_{V,A}(m, m) = \Pr(\max\{N_V, N_A\} \leq m),
\]

and from [3],

\[
\max\{0, F_V(m) + F_A(m) - 1\} \leq H_{V,A}(m, m) \leq \min\{F_V(m), F_A(m)\},
\]

for \( m = 0, 1, \ldots \). In [4], both the upper bound \( H^+(m) = \min\{F_V(m), F_A(m)\} \) and the lower bound \( H^-(m) = \max\{0, F_V(m) + F_A(m) - 1\} \) are univariate distribution functions of random variable \( \max\{N_V, N_A\} \). Moreover, it is well known (see [15]) that \( H^+ \) and \( H^- \) represent distributions with maximal positive, respectively negative, dependence between \( N_V \) and \( N_A \) (assuming non-degenerate marginal distributions \( F_V \) and \( F_A \)). For example, explicit values for the correlation, within the class of bivariate Poisson distributions, have been given in [16].

In order to derive the new measure of MI, one chooses the coupling of \( N_V \) and \( N_A \) corresponding to the lower bound \( H^- \), with maximal negative dependence between the variables. It turns out that this coupling predicts the maximum expected value from combining the unisensory responses:

**Proposition 1** \(^1\) For any coupling of the univariate response random variables \( N_V \) and \( N_A \) with expected value \( \mathbb{E}\max\{N_V, N_A\} \),

\[
\max\{\mathbb{E}N_V, \mathbb{E}N_A\} \leq \mathbb{E}\max\{N_V, N_A\} \leq \mathbb{E}^- \max\{N_V, N_A\},
\]

where \( \mathbb{E}^- \max\{N_V, N_A\} \) is the expected value under maximal negative dependence between the univariate response random variables.

In analogy to the traditional index MSI defined in [2], the new index based on the above is defined as

\[
\text{MSI}^* = \frac{\mathbb{E}N_{V,A} - \mathbb{E}^- \max\{N_V, N_A\}}{\mathbb{E}^- \max\{N_V, N_A\}} \times 100,
\]

In an empirical context, the expected value \( \mathbb{E}N_{V,A} \) is replaced by the sample mean of crossmodal responses and \( \mathbb{E}^- \max\{N_V, N_A\} \) is estimated using the method of antithetic variates as demonstrated in the next section (see also [17]). The test for multisensory enhancement then amounts to comparing the observed mean number of impulses to crossmodal stimulation with the estimate for \( \mathbb{E}^- \max\{N_V, N_A\} \).

### 3 Empirical Example

We illustrate the approach by a sample of recordings from a single cat SC neuron.\(^2\)

\(^1\)For proof, see section [6].

\(^2\)Information obtained from members of lab Z. Standard PSTHs were computed. Spontaneous activity was computed from the 500 ms preceding each stimulus onset (allowing at least 1500 ms between each trial). A threshold of mean S.A. rate per 10 ms bin plus 2 standard deviations was computed, only used to determine onset and offset. Response onset was defined when the first spike occurred within the bin that rises above this threshold and remained above for at least 3 bins. Offset was counted as the last spike in the bin just before the response fell back below this threshold and remained below for 3 bins. The response window (duration) is the time between onset and offset. Total number of spikes (left columns in the table) include all spikes within the response window, which will inevitably include some S.A. The right columns include responses with S.A. removed. The expected number of S.A. spikes within the given window (i.e., S.A. times window size in seconds) was removed. This is never an integer and can sometimes cause negative values on some trials. This number represents “change from baseline firing”.

3
3.1 Data

The data set consists of the total number of spikes occurring within a response window, including spontaneous activity (A.S.) in the left-hand columns and with A.S. removed in the right-hand columns of Table 1. The number of trials was \( N = 20 \) each for visual, auditory, and visual-auditory stimulation. The antithetic variates method involves pairing the unisensory responses, sorted by increasing order (\( V \)) and by decreasing order (\( A \)), and computing \( \max(V, A) \) for each pair. The mean of the \( \max(V, A) \) values constitutes an estimate of \( E^{-\max\{N_V, N_A\}} \), that is, of the maximum expected value from combining the unisensory responses achievable via negatively dependent probability summation.

Table 1: **Sample of recordings from a single cat SC:**
Columns 2 and 6 are arranged by increasing order, 3 and 7 by decreasing order. S.A. stands for “spontaneous activity” (4.26 spikes/s in this sample).

| trial | V | A | max(V, A) | VA | V | A | max(V, A) | VA |
|-------|---|---|-----------|----|---|---|-----------|----|
| 1     | 3 | 8 | 8         | 11 | 1.113 | 7.493 | 7.493 | 18.933 |
| 2     | 4 | 8 | 8         | 22 | 2.113 | 7.493 | 7.493 | 13.933 |
| 3     | 5 | 7 | 7         | 17 | 3.113 | 6.493 | 6.493 | 15.933 |
| 4     | 5 | 7 | 7         | 19 | 3.113 | 6.493 | 6.493 | 14.933 |
| 5     | 5 | 7 | 7         | 18 | 3.113 | 6.493 | 6.493 | 9.933  |
| 6     | 6 | 7 | 7         | 13 | 4.113 | 6.493 | 6.493 | 14.933 |
| 7     | 6 | 6 | 6         | 18 | 4.113 | 5.493 | 5.493 | 7.933  |
| 8     | 7 | 6 | 7         | 11 | 5.113 | 5.493 | 5.493 | 22.933 |
| 9     | 7 | 6 | 7         | 26 | 5.113 | 5.493 | 5.493 | 16.933 |
| 10    | 8 | 6 | 8         | 20 | 6.113 | 5.493 | 6.113 | 24.933 |
| 11    | 8 | 6 | 8         | 28 | 6.113 | 5.493 | 6.113 | 15.933 |
| 12    | 9 | 6 | 9         | 19 | 7.113 | 5.493 | 7.113 | 21.933 |
| 13    | 9 | 5 | 9         | 25 | 7.113 | 4.493 | 7.113 | 11.933 |
| 14    | 10| 5 | 10        | 15 | 8.113 | 4.493 | 8.113 | 13.933 |
| 15    | 10| 5 | 10        | 17 | 8.113 | 4.493 | 8.113 | 15.933 |
| 16    | 10| 4 | 10        | 19 | 8.113 | 3.493 | 8.113 | 15.933 |
| 17    | 11| 4 | 11        | 19 | 9.113 | 3.493 | 9.113 | 14.933 |
| 18    | 11| 4 | 11        | 18 | 9.113 | 3.493 | 9.113 | 27.933 |
| 19    | 13| 4 | 13        | 31 | 11.113| 3.493| 11.113| 13.933 |
| 20    | 14| 4 | 14        | 17 | 12.113| 3.493| 12.113| 7.933  |
| mean  | 8.05 | 5.75 | 8.85 | 19.15 | 6.163 | 5.243 | 7.484 | 16.083 |
| SD    | 2.999 | 1.333 | 2.159 | 5.204 | 2.999 | 1.333 | 1.791 | 5.204 |

3.2 Results

Computing the traditional MSI value by inserting the estimates from Table 1 in Equation 2, i.e., replacing the expected values by the means, yields

\[
\text{MSI} = \frac{E_{NV,A} - \max\{E_{NV}, E_{NA}\}}{\max\{E_{NV}, E_{NA}\}} \times 100 \approx 19.15 - \frac{\max\{8.05, 5.75\}}{\max\{8.05, 5.75\}} \times 100 = 137.89\%.
\]

for responses containing S.A (left-hand columns). The corresponding value for the new index is estimated by inserting the estimates from Table 1 in Equation 5

\[
\text{MSI}^* = \frac{E_{NV,A} - E^{-\max\{N_V, N_A\}}}{E^{-\max\{N_V, N_A\}}} \times 100 \approx 19.15 - \frac{8.85}{8.85} \times 100 = 116.64\%.
\]

The corresponding values for responses with S.A. removed (right-hand columns) amount to

\[
\text{MSI} \approx 16.083 - \frac{\max\{6.163, 5.243\}}{\max\{6.163, 5.243\}} \times 100 = 160.96\%.
\]
and

\[ \text{MSI}^* \approx \frac{16.083 - 7.484}{7.484} \times 100 = 114.90\% \].

The results are quite clearcut. For this neuron, all MSI and MSI* are statistically significant \((p < .01, t\text{-test, one-tailed})\). So it will be labeled “multisensory” under the traditional and the new index. Moreover, there is no statistically significant difference between MSI and MSI* when responses include spontaneous activity \((p > .17, t\text{-test, one-tailed})\). However, with S.A. removed, the difference between MSI and MSI* is highly significant \((p < .001)\). Thus, for this neuron there is a clear drop in the estimated level of multisensory activity by using the new index.

4 Discussion and Conclusion

We have introduced a new way of gauging a neuron’s multisensory integration capacity. In contrast to the traditional multisensory index (MSI), the new index (MSI*) introduced here has a clear theoretical foundation, without the need to specify the multisensory computational mechanism proper: it measures the degree by which a neuron’s observed multisensory response surpasses the level obtainable by optimally combining the unisensory responses under a probability summation rule, that is, assuming that the neuron simply reacts to the more salient modality in any given cross-modal trial. The new index is easy to compute and is as amenable to statistical testing as the traditional one\(^4\). For the example neuron studied here, both measures indicated it as “multisensory”. However, given that, by its definition, index MSI* always yields lower values than MSI, it is a clear possibility that in many cases neurons previously labeled as “multisensory” could no longer be considered as such.

The extent to which this holds can only be determined by a large-scale investigation of a multitude of neurons from empirical studies. Obviously, at the level of a (sub-)population of neurons, such a relabeling may lead to a reassessment of the distribution of multisensory neurons and different types of unisensory neurons for that region. Moreover, studies probing the entire scope of the behavior of multisensory neurons, e.g. by looking at intrinsic differences in the dynamic range of these neurons (see \[9\]), may come to different conclusion when using the new index.

It is worth mentioning that the new approach can also be applied to an alternative measure, comparing cross-modal responses to the sum of the unisensory responses (“additive model”) (see also \[11\]) . From \[18\] (and more recent papers in actuarial statistics), it is possible to compute the maximally achievable sum of two random variables and, using the same logic as for computing MSI*, cross-modal responses can be compared with the response level obtainable by adding the unisensory responses in an optimal way.

Finally, it should be noted that the present approach is analogous to a well-established procedure in the behavioral domain. In comparing cross-modal and unimodal reaction time (RT) to test for multisensory facilitation, the so-called “race model inequality” (\[19\]) has become a standard tool to probe whether cross-modal RT is faster than predictable by a dependent probability summation mechanism (see \[20\]). An inequality analogous to the one in Proposition 1, replacing the max operator by \(\min\) has been presented in \[21\] .

5 Acknowledgment

We are indebted to X and Y from the Z lab in providing the data from an SC neuron.

6 Proof of Proposition 1

From Equation \[4\]

\[ 1 - H^+(m) \leq 1 - H_{VA}(m, m) = P(\max\{N_V, N_A\} > m) \leq 1 - H^-(m) \],

for \(m = 0, 1, \ldots\).

Summing over all \(m\) yields

\[ \sum_{m=0}^{\infty} [1 - H_{VA}(m, m)] = \mathbb{E}\max\{N_V, N_A\} \leq \mathbb{E}^\min\{N_V, N_A\} \].

\(^4\)Bootstrap procedures for further refining the statistical testing procedure are being developed.
The left-hand side, \( \max\{E_N V, E_N A\} \leq E \max\{N_V, N_A\} \) follows from *Jensen's inequality* ([17], p. 51).
References

[1] B.E. Stein, editor. The New Handbook of multisensory processing. MIT Press, 2012.
[2] B.E. Stein and M.A. Meredith. The merging of the senses. MIT Press, 1993.
[3] M.A. Meredith, J.W. Nemitz, and B.E. Stein. Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. Journal of Neuroscience, 7:3215–3229, 1987.
[4] B.E. Stein, T.R. Stanford, and B.A. Rowland. Development of multisensory integration from the perspective of the individual neuron. Nature Reviews Neuroscience, 15:520–535, 2014.
[5] M.A. Meredith and B.E. Stein. Spatial factors determine the activity of multisensory neurons in cat superior colliculus. Brain Research, 365:350–354, 1986.
[6] A. Diederich and H. Colonius. Modeling the time course of multisensory perception in manual and saccadic responses. In G. Calvert, C. Spence, and E. Stein, B., editors, Handbook of Multisensory Processes, pages 395–408. Cambridge: MIT Press, 2004.
[7] B.E. Stein, T.R. Stanford, R. Ramachandran, T.J. Perrault Jr, and B.A. Rowland. Challenges in quantifying multisensory integration: alternative criteria, models, and inverse effectiveness. Experimental Brain Research, 198:113–126, 2009.
[8] M.A. Meredith and B.E. Stein. Interactions among converging sensory inputs in the superior colliculus. Science, 221:389–391, 1983.
[9] T.J. Perrault Jr, J.W. Vaughan, B.E. Stein, and M.T. Wallace. Superior colliculus neurons use distinct operational modes in the integration of multisensory stimuli. Journal of Neurophysiology, 93:2575–2586, 2005.
[10] L.C. Populin and T.C. Yin. Bimodal interactions in the superior colliculus of the behaving cat. Journal of Neuroscience, 22:2826–2834, 2002.
[11] T.R. Stanford, S. Quessy, and B.E. Stein. Evaluating the operations underlying multisensory integration in the cat superior colliculus. Journal of Neuroscience, 25(28):6499–6508, 2005.
[12] R.L. Miller, S.R. Pluta, B.E. Stein, and B.A. Rowland. Relative unisensory strength and timing predict their multisensory product. Journal of Neuroscience, 35(13):5213–5220, 2015.
[13] H. Thorisson. Coupling, stationarity, and regeneration. Springer Verlag, New York, NY, 2000.
[14] M. Fréchet. Sur les tableaux de corrélation dont les marges sont donnés. Annales de l’Université de Lyon, Section A, Séries 3(14):53–77, 1951.
[15] H. Joe. Multivariate models and dependence concepts. Number 73 in Monographs on Statistics and Applied Probability. Chapman & Hall, London, UK, 1997.
[16] R.C. Griffiths, R.K. Milne, and R. Wood. Aspects of correlation in bivariate Poisson distributions and processes. Australian Journal of Statistics, 21(3):238–255, 1979.
[17] S.M. Ross. Stochastic processes. John Wiley & Sons, New York, NY, Second edition, 1996.
[18] L. Rüschendorf. Random variables with maximum sums. Advances in Applied Probability, 14:623–632, 1982.
[19] J. O. Miller. Divided attention: Evidence for coactivation with redundant signals. Cognitive Psychology, 14:247—279, 1982.
[20] A. Diederich and H. Colonius. Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. Perception and Psychophysics, 66:1388–1404, 2004.
[21] H. Colonius and A. Diederich. Race model inequality: Interpreting a geometric measure of the amount of violation. Psychological Review, 113(1):148–154, 2006.