Measuring Multisensory Integration in Reaction Time: Relative Entropy Approach

Hans Colonius

Department of Psychology

Carl von Ossietzky Universität, Oldenburg, Germany

and

Department of Psychological Sciences

Purdue University

Adele Diederich

Department of Medical Physics

Carl von Ossietzky Universität, Oldenburg, Germany

and

Department of Psychological Sciences

Purdue University

Abstract

A classic definition of *multisensory integration* (MI) has been proposed as "the presence of a (statistically) significant change in the response to a crossmodal stimulus complex compared to unimodal stimuli". However, this general definition did not result in a broad consensus on how to quantify the amount of MI in the context of reaction time (RT). In this brief note, we argue that numeric measures of reaction times that only involve mean or median RTs do not uncover the information required to fully assess the effect of multisensory integration. We suggest instead novel measures that include the entire RT distributions functions. The central role is played by *relative entropy* (aka *Kullback-Leibler divergence*), a statistical concept in information theory, statistics, and machine learning to measure the (non-symmetric) distance between probability distributions. We provide a number of theoretical examples, but empirical applications and statistical testing are postponed to later study.

Measuring Multisensory Integration in Reaction Time:
Relative Entropy Approach

1. Introduction

The study of how information from different sensory modalities is merged to produce a unified percept is an important topic in many research fields including the behavioral sciences. A pragmatic definition of *multisensory integration* (MI) as "*the presence of a (statistically) significant change in the response to a crossmodal stimulus complex compared to unimodal stimuli*" has been proposed in Stein et al. (2009). In the realm of reaction time (RT) measures for MI, this amounts to comparing the average time, e.g., to detect a visual-auditory stimulus to the average time to detect a unisensory, visual or auditory, stimulus. The study of crossmodal interaction effects in RTs goes back more than 100 years (Todd, 1912), and has generated a huge number of studies (see reviews Hershenson (1962); Welch and Warren (1986); Rach et al. (2011)).

In this theoretical note, we argue that current MI measures for reaction time based solely on parameters of central tendency, like means or medians, do not fully reveal the information available to assess effects of integration across the senses. We suggest a novel approach to quantifying MI that involves the entire RT distributions but without underlying parametric assumptions. The basic idea is to measure, in a sense to be specified below, how "far away" the crossmodal RT distribution is from the unimodal RT distributions. A central role is played by the concept of *relative entropy* (aka *Kullback-Leibler divergence*), a statistical concept in information theory, statistics, and machine learning to measure the (oriented, i.e. non-symmetric) distance between probability distributions (Cover and Thomas, 1991). The new measure is illustrated by some theoretical examples. Empirical applications including simulation and testing are postponed to a future study.

We introduce the traditional measure of multisensory integration for RTs and point out its shortcomings due to being based on means or medians only. After presenting the notion of relative entropy and some of its properties, we give a general definition of crossmodal response enhancement (CRE) based on relative entropy. Subsequently, we discuss two different areas of application. In the first, measures of CRE are derived for different statistical distributions (exponential, normal, and lognormal) yielding an alternative to classic ones based on means (expected values) only. In the second, we use relative entropy to quantify by how much the prediction of an (arbitrary) model deviates from observed data, followed by the analysis of the new measure to two specific MI models, the race model and a mixture model. We conclude with some remarks concerning potential application of our approach to neurosensory data.

1.1 Response Enhancement in Redundant Signals Paradigm: Traditional Measure

In the redundant signals paradigm, also known as divided attention paradigm, a participant is instructed to respond as soon as a uni- or crossmodal signal occurs. A traditional measure of crossmodal response enhancement in RTs is defined as (e.g., Rach et al., 2011)

$$\mathrm{CRE}_{RT} = \frac{\min\{\mathrm{E}RT_V, \mathrm{E}RT_A\} - \mathrm{E}RT_{VA}}{\min\{\mathrm{E}RT_V, \mathrm{E}RT_A\}} \times 100. \tag{1}$$

Here, E stands for expected (mean) value, but the median is often used instead as well. The numerator compares the faster of the unisensory RTs (here, visual or auditory) to the crossmodal (visual-auditory) RT, and the denominator and multiplication factor simply serve to standardize the measure. Thus, CRE_{RT} expresses multisensory enhancement or inhibition as a proportion of the faster unisensory response. For example, $\mathrm{CRE}_{RT} = 10$ means that mean RT to the visual-auditory stimulus is 10% faster than the faster of the expected RTs to the unimodal visual or auditory stimuli. For simplicity, we neglect occurrence of erroneous responses, like failure to detect a stimulus.

Measure CRE_{RT} is a simple way of quantifying MI that is amenable to standard statistical testing. However, it does not take into account that integrating information from different modalities may also affect other, more fine-grained aspects of the associated RT distributions. For example, one possible result of integrating information might be that short RTs become more frequent while long RTs tend to be even longer, leaving the difference between uni- and crossmodal mean RTs more or less invariant. Because stimulus detection is conceived of as a stochastic event generating some random variability in information accumulation, the way this variability is modified under crossmodal stimulation may yield important insights into the integration process itself (Otto et al., 2013).

2. Multisensory Integration Measures Based on Relative Entropy

All available information about the MI process is contained in how the multisensory RT distribution differs from the unimodal RT distributions. Thus, an MI measure should be some function of this difference. We identify two issues: first, how should this difference be formally defined? Second, how should the two unisensory RT distributions be combined to enter into that expression?

Recall that a metric d on a set S is defined as a function $d\colon S \times S \to \mathbb{R}$ ($\mathbb{R}$ the set of real numbers) such that, for all $x, y, z \in S$, (i) $d(x, y) \geq 0$ (*non-negativity*); (ii) $d(x, y) = 0$ if and only if $x = y$; (iii) $d(x, y) = d(y, x)$ (*symmetry*); and (iv) $d(x, y) \leq d(x, z) + d(z, y)$ (*triangle inequality*).

There is a huge number ways to define a metric on a set of probabilities (Deza and Deza, 2009). However, it turns out that not all properties of a metric are actually needed for our approach. We want a measure that quantifies how the unimodal distributions have to be "modified" in order to attain the crossmodal distribution; thus, neither symmetry nor the triangle inequality are required. This suggests using the following concept:

Definition 1 *The relative entropy (RE) between two probability mass functions $p(x)$ and*

$q(x)$ is defined as

$$D(p||q) = \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)} \tag{2}$$

$$= E_p \log \frac{p(X)}{q(X)}. \tag{3}$$

Here, X is a discrete real-valued random variable and Equation (3) means that $D(p||q)$ equals the expected value of random variable $\log \frac{p(X)}{q(X)}$ with respect to probability mass function $p(x)$. We use the convention that $0 \log \frac{0}{q} = 0$ and $p \log \frac{p}{0} = 0$. Relative entropy is also known as Kullback-Leibler Divergence (KLD). Relative entropy for continuous random variables with probability density functions (pdf) f and g is defined as

$$D(f||g) = \int f \log \frac{f}{g}.$$

$D(p||q)$ can be interpreted as measuring how well the "target" distribution $p(x)$ is approximated by $q(x)$; it plays an important role in several fields including information theory, statistical physics, neural networks, and Bayesian statistics (Kullback and Leibler, 1951; Cover and Thomas, 1991; MacKay, 2003). Relevant properties for our purposes are:

1. $D(p||q) = 0$ if and only if $p = q$ (self-identification)

2. $D(p||q) \geq 0$ for all p, q (non-negativity)

Non-negativity, also known as *Gibb's inequality* or *information inequality*, follows from Jensen's inequality (for proofs, see the above references).

2.1 Defining a Measure of MI Based on Relative Entropy

In order to define a measure of crossmodal response enhancement based on relative entropy, we take the crossmodal pdf, $f_{VA}(t)$, as "target" function $p(x)$, or f, and unimodal pdfs f_V, f_A as $q(x)$, or g (indexes VA, V, and A here stand again for visual-auditory crossmodal and unimodal conditions). Without adding any modeling assumption, we take the smaller of the Kullback-Leibler divergences (note that we are using the shorthand KLD from now on) with respect to the unisensory distributions to define :

Definition 2

$$\text{CRE}_{KLD} = \min\{D(f_{VA}||f_V), D(f_{VA}||f_A)\}. \tag{4}$$

This is analogous to the traditional measure CRE_{RT} of Equation (1). It equals zero if and only if $f_{VA} = f_V$ or $f_{VA} = f_A$. Note that KLD values can take large values going towards infinity. In order to make CRE_{KLD} values from different data sets comparable, a standardization like in Equation (1) would be desirable, but it seems not obvious how to do this.

2.2 Measures of MI for Some Specific Distributions

The first two examples (exponential and normal distributions) are only for illustration of the measure defined in Equation (4). The third (lognormal distributions) is a plausible RT distribution. In particular, an example demonstrates how a measure based on relative entropy is more informative than the traditional measure: the latter does not detect any MI effect whereas the former clearly does.

Example 1 (Exponential) *We assume that f_{VA}, f_V, f_A are exponential distributions with parameters $\lambda_{VA}, \lambda_V, \lambda_A$, respectively, and let $\lambda_{VA} > \lambda_A > \lambda_V > 0$. Then*

$$\begin{aligned} D(f_{VA}||f_V) &= \int_0^\infty \lambda_{VA} \exp(-\lambda_{VA}\, t) \log \left[\frac{\lambda_{VA} \exp(-\lambda_{VA}\, t)}{\lambda_V \exp(-\lambda_V\, t)} \right] dt \\ &= \log \frac{\lambda_{VA}}{\lambda_V} + \frac{\lambda_V}{\lambda_{VA}} - 1, \end{aligned}$$

so that we obtain the same as above, due to assuming $\lambda_A > \lambda_V$:

$$\begin{aligned} \text{CRE}_{KLD} &= \min\{D(f_{VA}||f_V), D(f_{VA}||f_A)\} \\ &= \min\{\log \frac{\lambda_{VA}}{\lambda_V} + \frac{\lambda_V}{\lambda_{VA}} - 1, \log \frac{\lambda_{VA}}{\lambda_A} + \frac{\lambda_A}{\lambda_{VA}} - 1\} \\ &= \log \frac{\lambda_{VA}}{\lambda_V} + \frac{\lambda_V}{\lambda_{VA}} - 1. \end{aligned}$$

With $(\lambda_{VA}/\lambda_V) \to \infty$, that is, if the effect of MI increases without bound, then $\text{CRE}_{KLD} \to \infty$ *as well. Note that the exponential is not a plausible RT distribution and is presented here for illustration only.*

Example 2 (Normal) *The density of the normal distribution of a real-valued random variable X is*

$$f(x;\mu_x,\sigma_x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right)$$

with $\mu_x \in (-\infty,+\infty)$ and $\sigma_x>0$, abbreviated as $X\sim\mathcal{N}(\mu_x,\sigma_x^2)$. The KLD for two random variables X and Y with $Y\sim\mathcal{N}(\mu_y,\sigma_y^2)$ with densities f_x, f_y, respectively, is known to equal

$$D(f_x||f_y)=\frac{1}{2}\left[\frac{(\mu_x-\mu_y)^2}{\sigma_y^2}+\frac{\sigma_x^2}{\sigma_y^2}-\log\frac{\sigma_x^2}{\sigma_y^2}-1\right]. \tag{5}$$

We have a unisensory (visual) distribution $\mathcal{N}(\mu_V,\sigma_V^2)$, a unisensory (auditory) distribution $\mathcal{N}(\mu_A,\sigma_A^2)$, and a bisensory distribution $\mathcal{N}(\mu_{VA},\sigma_{VA}^2)$; then

$$D(f_{VA}||f_V)=\frac{1}{2}\left[\frac{(\mu_{VA}-\mu_V)^2}{\sigma_V^2}+\frac{\sigma_{VA}^2}{\sigma_V^2}-\log\frac{\sigma_{VA}^2}{\sigma_V^2}-1\right].$$

Because $EX=\mu$ and $VarX=\sigma^2$ are functionally independent, that is, their values can vary separately, the KLD depends on both the means and variances of the distributions, and so does $\mathrm{CRE}_{KLD}=\min\{D(f_{VA}||f_V),D(f_{VA}||f_A)\}$. Importantly, even under equality of the means, CRE remains non-zero. In particular, if the bisensory distribution has a larger or a smaller variance compared to the unisensory distributions, this is taken into account in the KLD-based measure of MI.

Because of the symmetry of the normal distribution, this example is again not a realistic one for empirical RT data. The following example, however, is often considered to be of a plausible shape for RTs.

Example 3 (Log-Normal) *The density of the log-normal distribution of a non-negative random variable X is*

$$f(x;\mu,\sigma)=\frac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\frac{(\log(x)-\mu)^2}{2\sigma^2}\right),$$

with $\mu\in(-\infty,+\infty)$ and $\sigma>0$, abbreviated as $X\sim\mathcal{LN}(\mu_x,\sigma_x^2)$. Moreover,

$$EX=\exp\left(\mu+\frac{\sigma^2}{2}\right)\ \textit{and}\ VarX=[\exp(\sigma^2)-1]\exp(2\mu+\sigma^2).$$

For random variables $X \sim \mathcal{LN}(\mu_x, \sigma_x^2)$ and $Y \sim \mathcal{LN}(\mu_y, \sigma_y^2)$ with pdfs f_x, f_y, respectively, the KLD can be shown Dittrich (2013) to be

$$D(f_x||f_y) = \log \frac{\sigma_y}{\sigma_x} + \frac{1}{2\sigma_y^2}[(\mu_x - \mu_y)^2 + \sigma_x^2 - \sigma_y^2.] \tag{6}$$

Assume f_{VA}, f_V, f_A are all log-normal distributions, thus $\mathcal{LN}(\mu_{VA}, \sigma_{VA}^2)$, $\mathcal{LN}(\mu_V, \sigma_V^2)$, and $\mathcal{LN}(\mu_A, \sigma_A^2)$, respectively. Then,

$$D(f_{VA}||f_V) = \log \frac{\sigma_V}{\sigma_{VA}} + \frac{1}{2\sigma_V^2}[(\mu_{VA} - \mu_V)^2 + \sigma_{VA}^2 - \sigma_V^2]$$

and

$$D(f_{VA}||f_A) = \log \frac{\sigma_A}{\sigma_{VA}} + \frac{1}{2\sigma_A^2}[(\mu_{VA} - \mu_A)^2 + \sigma_{VA}^2 - \sigma_A^2].$$

Both mean and variance of log-normal random variables are functions of both parameters (μ and σ), so they cannot vary independently. Still,

$$\text{CRE}_{KLD} = \min\{D(f_{VA}||f_V), D(f_{VA}||f_A)\}$$

depends on both moments.

Two special cases are of interest as well:

1. $\sigma_{VA} = \sigma_V = \sigma_A = \sigma$:

$$\text{CRE}_{KLD} = \frac{\min\{(\mu_{VA} - \mu_V)^2), (\mu_{VA} - \mu_A)^2\}}{\sigma^2};$$

2. $\sigma_{VA} = \sigma_V = \sigma_A = \sigma$ *and* $\mu_V = \mu_A = \mu$:

$$\text{CRE}_{KLD} = \frac{(\mu_{VA} - \mu)^2}{\sigma^2}$$

Even in these restricted cases, the MI measure depends on parameter σ^2 modulating both mean and variance.

The last example serves to illustrates that measure CRE_{KLD} is more informative than the traditional one. Indeed, assume that $\text{E}RT_{VA} = \text{E}RT_A$; for convenience, we also assume $\mu_A = \mu_V$ and $\sigma_A = \sigma_V$ implying $f_V = f_A$ and $\mathrm{E}RT_V = \mathrm{E}RT_A$. Then $\mathrm{CRE}_{RT} = 0$ indicates a null effect of multisensory integration. On the other hand, it also means that

$$\exp\left[\mu_{VA} + \frac{\sigma_{VA}^2}{2}\right] = \exp\left[\mu_A + \frac{\sigma_A^2}{2}\right] \quad \text{or,}$$
$$\mu_{VA} + \frac{\sigma_{VA}^2}{2} = \mu_A + \frac{\sigma_A^2}{2}, \tag{7}$$

which does not imply $f_{VA} = f_A$ except if $\mu_{VA} = \mu_A$ and $\sigma_{VA} = \sigma_A$. For example, let $\mu_{VA} = 3, \sigma_{VA} = 2$ and $\mu_A^{(1)} = -7.5$ and $\sigma_A^{(1)} = 5$. While both pdfs, f_{VA} and $f_A^{(1)}$, have the same mean (148), their shape is very different: $f_A^{(1)}$ has much more probability mass on short values than f_{VA} (see Figure1).

(FIGURE 1 ABOUT HERE)

Moreover, by simple calculation

$$\mathrm{CRE}_{KLD} = D(f_{VA}||f_A^{(1)}) = 2.70129, \tag{8}$$

For another pdf, $f_A^{(2)}$, with $\mu_A^{(2)} = -13$ and $\sigma_A^{(2)} = 6$, again with the same mean, we get

$$\mathrm{CRE}_{KLD} = D(f_{VA}||f_A^{(2)}) = 4.20972, \tag{9}$$

representing an even larger effect of multisensory integration.

Very similar treatments can be performed with other 2-parameter distributions, e.g., the gamma. Obviously, measure CRE_{KLD} can also be defined when f_{VA}, f_V, f_A all belong to different distributional families, e.g., log-normal unisensory distributions together with a bisensory Weibull.

3. MI measures based on model predictions

Besides calculating empirical measures of MI like CRE_{RT}, measures based on models of the integration process are in use as well. Specifically, given a model predicting performance in the crossmodal condition from the unimodal conditions, a KLD-based measure quantifies how "far away" the prediction is from the observed data. This is analogous to the role of relative entropy in statistical testing, namely to quantify how much an empirical data set deviates from a hypothesized distribution or model.

Let $\tilde{f}_{VA,\theta}(t)$ denote the bisensory density predicted by some model with parameter space $\theta \subset \mathbb{R}^d$. The less the observed MI distribution is predictable from the model, the larger the CRE measure should be. The KLD-based MI measure then is defined by

$$\text{CRE}_{KLD} = \min_{\theta \subset \mathbb{R}^d} D(f_{VA} || \tilde{f}_{VA,\theta}). \tag{10}$$

Of course, minimization over the parameter space will be void when a model is parameter-free.

3.1 The race model: traditional vs. KLD-based MI measures. One of the earliest multisensory models is the (horse) race model, that is, a visual-auditory stimulus complex is supposed to trigger random visual and auditory processes such that the observed RT equals the minimum time of the two, i.e., the 'winner of the race' Raab (1962). Thus, combination of the unisensory distributions is here simply defined by probability summation. Under stochastic independence, the bisensory distribution function of the race model is obtained as

$$\tilde{F}_{VA}(t) = F_V(t) + F_A(t) - F_V(t)\, F_A(t),$$

with corresponding density

$$\tilde{f}_{VA}(t) = f_V(t)(1 - F_A(t)) + f_A(t)(1 - F_V(t)), \;\; t \geq 0. \tag{11}$$

A violation of the race model occurs if the observed distribution $F_{VA}(t)$ is larger than $\tilde{F}_{VA}(t)$ for some t. The most traditional MI measure quantifies the amount of violation by defining

$$R_{VA}^{IND} = \int_0^{\infty} [F_{VA}(t) - (F_V(t) + F_A(t) - F_V(t)\, F_A(t))]^+ \;\; dt. \tag{12}$$

Thus, it simply takes the area between the observed bisensory distribution function and the one predicted via the race model (Colonius and Diederich, 2020). Without assuming stochastic independence, the measure R_{VA}^{IND} can be replaced by the, generally smaller, measure

$$R_{VA}^{MND} = \int_0^\infty [F_{VA}(t) - \min\{F_V(t) + F_A(t), 1\}]^+ \ dt,$$

corresponding to maximally negative dependence between the 'racers' (Colonius and Diederich, 2006). It has been shown that areas R_{VA}^{IND} and R_{VA}^{MND} are simply equal to the difference between the observed mean (expected value) of the bisensory distribution and the mean predicted by a race model under stochastic independence and maximal negative dependence, respectively (Colonius and Diederich, 2006).

3.2 KLD-based MI measures for race models. Under the independent race model (IND), inserting the bisensory density into the KLD measure yields

$$\begin{aligned} &D(f_{VA}||\tilde{f}_{VA}) \\ &= \int_0^\infty f_{VA}(t) \log \frac{f_{VA}(t)}{\tilde{f}_{VA}(t)} \, dt \\ &= \int_0^\infty f_{VA}(t) \log \frac{f_{VA}(t)}{f_V(t)(1 - F_A(t)) + f_A(t)(1 - F_V(t))} \, dt \end{aligned} \tag{13}$$

In this parameter-free form, the integral (13) can be taken as CRE$_{KLD}$. If some specific distributions for the IND model are assumed, minimization of $D(f_{VA}||\tilde{f}_{VA})$ over the parameter space would be required.

Comparing the KLD-based measure with traditional one, R_{VA}^{IND}, suggests that the former one should be more sensitive with respect to the distributional shapes. The reason is that in the traditional measures, integration is over distributions functions, whereas integration is over densities in KLD measures (13). Moreover, instead of race models, any other model type (e.g., diffusion co-activation models) predicting $\tilde{f}_{VA}(t)$ can be inserted in CRE$_{KLD}$.

3.3 KLD measures based on mixtures. Consider a mixture of the unisensory distributions,

$$\tilde{f}_{VA,\alpha}(t) = \alpha f_V(t) + (1-\alpha) f_A(t),$$

with $0 \leq \alpha \leq 1$. This model holds that the bisensory response is determined by just one modality while ignoring the other, with probability α by the visual and $1-\alpha$ the auditory modality. For a given set of RT distributions, a value of α can be determined that gives the smallest KLD value of $D(f_{VA}||\alpha f_V + (1-\alpha) f_A)$, that is:

$$\alpha^* = \arg \min_{\alpha \in [0,1]} D(f_{VA}||\alpha f_V + (1-\alpha) f_A). \tag{14}$$

We then have

$$\text{CRE}_{KLD} = D(f_{VA}||\tilde{f}_{VA,\alpha^*}). \tag{15}$$

The cases of $\alpha = 1$ or 0 would yield the components of CRE_{KLD} again. This example is special because the model does not predict response enhancement but inhibition of the bisensory RT (for examples, see Welch and Warren (1986)). The value of α^* may be of interest when interpreted as the relative weight given to the visual component in approximating the bisensory distribution.

4. Dicussion

Relative entropy (*aka* Kullback-Leibler divergence, KLD) is an oriented (i.e., non-symmetric) measure of "distance" between probability distributions. Here we demonstrate that it can be used to define measures of crossmodal response enhancement that more fully uncover the information about the integration process than classic measures based solely on means or medians of RT data. These novel measures are defined by the relative entropy between some combination of the unimodal RT distributions and the (observed) crossmodal RT distribution, thus gauging the "distance" between the former and the latter. We present examples where the classic measures are (close to) zero, because the difference between uni-and crossmodal means is (close to) zero whereas the KLD based measure is sensitive to changes in the shape of the RT distributions. We limit our presentation to theoretical examples, but an extension of the approach to empirical data is straightforward drawing upon the ubiquitous applications of relative entropy in various areas of statistics and machine learning.

While our focus here was on measuring multisensory effects on reaction time, the relative entropy approach could easily be extended to the realm of neuronal data. Note that the most widely used descriptive measure of the magnitude of multisensory integration, measured by absolute spike frequency, is defined as

$$\mathrm{CRE} = \frac{\mathrm{CM} - \mathrm{SM}_{\max}}{\mathrm{SM}_{\max}} \times 100, \tag{16}$$

where, at the sample level, CM is the mean absolute number of spikes in response to the crossmodal stimulus and $\mathrm{SM}_{\max}$ is the mean absolute number of spikes to the most effective modality-specific component stimulus (Meredith and Stein, 1983; Stein and Meredith, 1993). We have previously suggested to replace (16) by a measure taking into account possible stochastic dependency between the sensory channels under crossmodal stimulation (Colonius and Diederich, 2017). Using KLD measures on the (theoretical or empirical) spike count frequency distributions, in analogy to the one suggested for RTs here, would go one step further in maximizing the amount of information uncovered by numerical measures of MI.

It had been acknowledged early on that measures based solely on mean numbers of impulses have limitations (Rowland and Stein, 2007). In particular, it is not sensitive to the temporal variation of the rate of information across the stimulus processing ("initial response enhancement"). The suggested replacement of the mean spike counts in (16) by the KLD measure would not resolve that problem, however. On the other hand, Miller et al. (2017) developed a continuous-time multisensory model that postulates a moment-by-moment operation transforming unisensory inputs into a multisensory output ("multisensory transform"). It would be interesting to investigate whether it is possible to define KLD measures on the unisensory-multisensory transforms, after appropriate standardization into probability distributions, to extract additional information from the data.

References

Colonius, H. and Diederich, A. (2006). The race model inequality: interpreting a geometric measure of the amount if violation. Psychological Review, 113(1):148–154.

Colonius, H. and Diederich, A. (2017). Measuring multisensory integration: from reaction times to spike counts. Scientific Reports, 7(1):3023. http://dx.doi.org/10.1038/s41598-017-03219-5.

Colonius, H. and Diederich, A. (2020). Formal models and quantitative measures of multisensory integration: a selective overview. European Journal of Neuroscience, 51:1161–1178.

Cover, T. and Thomas, J. (1991). Elements of information theory. John Wiley & Sons, Inc., New York, NY.

Deza, M. M. and Deza, E. (2009). Encyclopedia of distances. Springer-Verlag, Berlin Heidelberg.

Dittrich, D. (2013). Statistical methods for estimating loss distributions. Master's thesis, Dept. Mathematics, TU München.

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. Journal of Experimental Psychology, 63(3):289–293.

Kullback, S. and Leibler, R. (1951). On information and sufficiency. The Annals of Mathematical Statistics, 22(1):79–86.

MacKay, D. (2003). Information theory, inference, and learning algorithms. Cambridge University Press, Cambridge, UK.

Meredith, M. and Stein, B. (1983). Interactions among converging sensory inputs in the superior colliculus. Science, 221:389–391.

Miller, R. L., Stein, B., and Rowland, B. (2017). Multisensory integration uses a real-time unisensory-multisensory transform. The Journal of Neuroscience, 37(20):5183–5194.

Otto, T., Dassy, B., and Mamassian, P. (2013). Principles of multisensory behavior. The Journal of Neuroscience, 33:7463–7474.

Raab, D. (1962). Statistical facilitation of simple reaction time. Transactions of the New York Academy of Sciences, 24:574–590.

Rach, S., Diederich, A., and Colonius, H. (2011). On quantifying multisensory interaction effects in reaction time and detection rate. Psychol Res, 75(2):77–94.

Rowland, B. and Stein, B. (2007). Multisensory integration produces an initial response enhancement. Frontiers in Integrative Neuroscience, 1:1–8.

Stein, B. and Meredith, M. (1993). The merging of the senses. MIT Press.

Stein, B., Stanford, T., Ramachandran, R., Perrault Jr, T., and Rowland, B. (2009). Challenges in quantifying multisensory integration: alternative criteria, models, and inverse effectiveness. Experimental Brain Research, 198(2):113–126.

Todd, J. (1912). Reaction to multiple stimuli. Archives of Psychology No. 25. Columbia Contributions to Philosophy and Psychology, XXI(8). New York: The Science Press.

Welch, R. and Warren, D. (1986). Intersensory interactions. In Boff, K., Kaufman, L., and Thomas, J., editors, Handbook of perception and human performance, volume 1, chapter 25, pages 1–36. John Wiley & Sons, New York, NY.

Acknowledgments

This research was supported by grants from German Science Foundation (DFG): HC by grant CO 94/8-1, AD by grant DI 506/19-1. Computations were performed using Mathematica©(13.3).

Author contributions statement

HC and AD wrote the main manuscript text and prepared the figure. Both authors reviewed the manuscript.

Additional information

Competing financial interests The authors declare no competing financial interest.

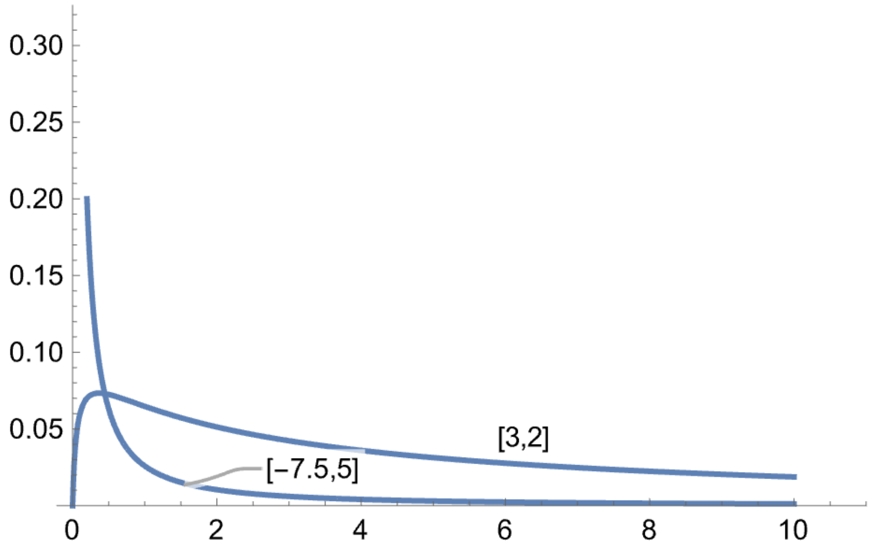

Figure 1. Two lognormal probability distribution functions with the same mean (E = 148) but different parameters: $(\mu_1 = 3, \sigma_1 = 2)$ vs. $(\mu_2 = -7.5, \sigma_2 = 5)$.