Principles of Multisensory Behavior

Thomas U. Otto,^{1,2,3,4} Brice Dassy,^{1,2,5} and Pascal Mamassian^{1,2}

¹Paris Descartes University, Sorbonne Paris City, F-75006 Paris, France, ²CNRS UMR 8158, Laboratory of the Psychology of Perception, F-75006 Paris, France, ³Modeling of Cognitive Processes, Technical University of Berlin, D-10587 Berlin, Germany, ⁴Bernstein Center for Computational Neuroscience, D-10115 Berlin, Germany, and ⁵School of Psychology, Cardiff University, Cardiff, CF10 3XQ, United Kingdom

The combined use of multisensory signals is often beneficial. Based on neuronal recordings in the superior colliculus of cats, three basic rules were formulated to describe the effectiveness of multisensory signals: the enhancement of neuronal responses to multisensory compared with unisensory signals is largest when signals occur at the same location (“spatial rule”), when signals are presented at the same time (“temporal rule”), and when signals are rather weak (“principle of inverse effectiveness”). These rules are also considered with respect to multisensory benefits as observed with behavioral measures, but do they capture these benefits best? To uncover the principles that rule benefits in multisensory behavior, we here investigated the classical redundant signal effect (RSE; i.e., the speedup of response times in multisensory compared with unisensory conditions) in humans. Based on theoretical considerations using probability summation, we derived two alternative principles to explain the effect. First, the “principle of congruent effectiveness” states that the benefit in multisensory behavior (here the speedup of response times) is largest when behavioral performance in corresponding unisensory conditions is similar. Second, the “variability rule” states that the benefit is largest when performance in corresponding unisensory conditions is unreliable. We then tested these predictions in two experiments, in which we manipulated the relative onset and the physical strength of distinct audiovisual signals. Our results, which are based on a systematic analysis of response time distributions, show that the RSE follows these principles very well, thereby providing compelling evidence in favor of probability summation as the underlying combination rule.

Introduction

The availability of different senses is often beneficial. They not only increase the spectrum of perceivable signals but also provide potentially redundant signals that combined enable better estimates of external events and/or faster achievements of subjective goals. A central theme in multisensory research is consequently to better understand the processes and principles that guide an efficient combination of signals from the different senses (Ernst and Bühlhoff, 2004; Driver and Noesselt, 2008).

Based on the electrophysiology of neurons in the superior colliculus, three basic principles were formulated to describe the effectiveness of multisensory signals (Meredith and Stein, 1983; Stein and Meredith, 1993; Stein and Stanford, 2008). Enhancements of neuronal responses in multisensory compared with unisensory conditions are largest when physical signals occur approximately at the same location (“spatial rule”), when signals are presented at approximately the same time (“temporal rule”), and when signal strength is rather low, eliciting only weak uni-

sensory responses (“principle of inverse effectiveness”; PoIE). In view of this influential work, the same three principles are often also considered with respect to behavioral benefits that can be observed with multisensory signals (for recent examples, see Chandrasekaran et al., 2011; Senkowski et al., 2011; Buchholz et al., 2012; Cappe et al., 2012; Stevenson et al., 2012). However, it is remarkable that, for example, the speedup of responses in multisensory compared with unisensory conditions (known as the redundant signal effect or RSE; Todd, 1912; Hershenson, 1962; Miller, 1982) is actually not following these principles. The spatial rule fails because the RSE is not affected by stimulus alignment (Murray et al., 2005). The temporal rule fails because the RSE with auditory and visual signals is typically largest when auditory signals are delayed, but the optimal delay is highly variable for different signals (Hershenson, 1962; Miller, 1986). Finally, the PoIE fails at least when the physical signal strength is manipulated in some conditions (Chandrasekaran et al., 2011) but not in others (Senkowski et al., 2011). In summary, it is unclear how to best predict behavioral benefits with multisensory signals such as the speedup of responses and how to bridge the apparent gap between electrophysiology and behavior.

A step toward common principles to predict multisensory benefits may emanate from the observation that the effectiveness of multisensory signals actually depends not on their physical properties but roughly on the simultaneity of elicited responses, which seems to hold true for both neurons (Stein and Stanford, 2008) and behavior (Chandrasekaran et al., 2011). However, a general theoretical framework underlying these observations

Received Oct. 2, 2012; revised March 18, 2013; accepted March 20, 2013.

Author contributions: T.U.O., B.D., and P.M. designed research; T.U.O. and B.D. performed research; T.U.O. contributed unpublished reagents/analytic tools; T.U.O. analyzed data; T.U.O. and P.M. wrote the paper.

The research leading to these results has received funding from the European Community Seventh Framework Program FP7/2007–2013 under Grant 214728-2, as well as National Agency for Research Grant ANR-10-BLAN-1910. Parts of this paper have been presented previously at the 13th International Multisensory Research Forum, Oxford, UK, June 19, 2012.

Correspondence should be addressed to Thomas U. Otto, Modeling of Cognitive Processes, Technical University of Berlin, Marchstrasse 23, D-10623 Berlin, Germany. E-mail: tom.u.otto@gmail.com.

DOI:10.1523/JNEUROSCI.4678-12.2013

Copyright © 2013 the authors 0270-6474/13/337463-12\$15.00/0

is missing. Here, we show that probability summation (Raab, 1962), which has been discarded until recently (Otto and Mamassian, 2012), provides such a framework. Based on theoretical considerations, we first demonstrate that probability summation predicts two basic principles for multisensory benefits, which we will call the “principle of congruent effectiveness” and the “variability rule.” Then, in two experiments, we show that the RSE follows these newly derived principles very well when we varied the timing (experiment 1) and the physical strength (experiment 2) of audiovisual signals.

Materials and Methods

The redundant signal paradigm. To investigate behavioral benefits with multisensory signals, we used the classical redundant signal paradigm. We asked participants to detect the common onset of two distinct multisensory signals (with the term “distinct,” we refer to signals that do not refer to a common environmental property; Fig. 1*a*). This so-called redundant condition is randomly interleaved with two single conditions, in which only one or the other signal is presented (Fig. 1*b*). By design, signals are “redundant” in the sense that detection of either signal is sufficient for a correct response. In other words, signals are coupled by a logical OR operator by the definition of the task. Typically, response times (RTs) are faster in redundant compared with single signal conditions (which is the RSE). In our experiments, we used visual motion and auditory sound signals embedded in an audiovisual background, but in principle any combination of signals can be studied in both humans and nonhumans (Diederich and Colonius, 1987; Giray and Ulrich, 1993; Mordkoff and Yantis, 1993; Hughes et al., 1994; Schröger and Widmann, 1998; Gondan et al., 2004; Molholm et al., 2004; Murray et al., 2005; Katzner et al., 2006; Whitchurch and Takahashi, 2006; Martuzzi et al., 2007; Collignon et al., 2008; Hecht et al., 2008; Hirokawa et al., 2008; Tamietto and de Gelder, 2008; Poom, 2009; Suied et al., 2009; Zehetleitner et al., 2009; Cappe et al., 2010; Veldhuizen et al., 2010; Chandrasekaran et al., 2011; Senkowski et al., 2011; Brang et al., 2012; Buchholz et al., 2012). To uncover the conditions that maximize multisensory benefits in behavior, we varied the relative onset (experiment 1) and the strength (experiment 2) of the redundant motion and sound signals. Because of the similarity between the two experiments, we combine below the methods of both experiments.

Participants. We tested the RSE with human participants, who were 21–34 years old and had self-reported normal hearing and normal or corrected-to-normal vision (experiment 1: 6 females and 4 males; experiment 2: 5 females, 15 males). All but two participants (T.U.O. and B.D.) were naive to the purpose of the experiment. All participants gave informed consent before the experiment. The study was conducted in agreement with the Declaration of Helsinki.

Experimental setup. Auditory stimuli were presented to both ears simultaneously through Sennheiser HD-280 Pro headphones. Visual stimuli were presented on a Sony GDM-C520 CRT monitor (100 Hz refresh rate). Viewing distance was 60 cm supported by a chin rest. A computer running MATLAB (MathWorks) equipped with standard toolboxes (Brainard, 1997; Pelli, 1997) controlled stimulus presentation and the collection of responses via the keyboard.

Figure 1. The redundant signal paradigm and the probability summation framework. **a**, Participants indicated the onset of two target signals (motion or sound; see **b**) that were embedded in a continuous audiovisual background. Except for similar stimulus onset, these distinct signals did not refer to a common environmental property. **b**, We presented signals in three basic conditions according to the classical redundant signal paradigm. In two single signal conditions, either the motion (M) or the sound (S) signal was presented. In a redundant signal condition (R), both signals were presented. Signals are “redundant” in the sense that detection of either signal was sufficient for a correct response. Hence, the two signals are coupled by a logical OR operator by the definition of the task. **c**, According to the LATER model of perceptual decision making, on presentation of M, evidence for the motion signal is accumulated from the start level L_0 until the threshold L_T is reached and a response is triggered. The drift rate is subject to Gaussian noise (see D_M ; the drift rate of an exemplary trial is indicated by the thick diagonal line). The resulting RT distribution follows the reciprocal of the Gaussian, which is a reci-normal distribution that is skewed to the right like empirical distributions. **d**, Evidence for the sound signal is accumulated analogously. **e**, On presentation of R, according to race models, a response on a given trial can be triggered by the faster of the two parallel, stochastic decision processes (in the illustrated exemplary trial, when evidence for sound reaches L_T). Consequently, compared with the single signal conditions, a speedup of RTs is expected on average across trials. Within the probability summation framework (for more details, see Materials and Methods), the resulting RT distribution can be computed by the minimum function of the RT distributions shown in **c** and **d** (which corresponds to the maximum function of the corresponding drift rate distributions).

Auditory stimulation. Auditory background noise was continuously presented in all conditions. The noise was Gaussian (i.e., a sequence of normally distributed random numbers with zero mean at a sample rate of 44.1 kHz), which was filtered so that most of the power was between 262 and 330 Hz. As measured using a Brüel and Kjær sound level meter equipped with an artificial ear adaptor, the presentation level of the background noise was 53 dB(A) and 52 dB(A) in experiments 1 and 2, respectively. As sound signals, we presented 294 Hz tones (the note D) that were embedded in the noise. Tones were presented for 500 ms, including sinusoidal ramp onset (10 ms in experiment 1; 5 ms in experiment 2) and offset (100 ms in experiments 1; 5 ms in experiments 2). In experiment 1, the presentation level of tones was 53 dB(A). Presentation level of tones in experiment 2 were 49.5, 50.5 and 53 dB(A) in conditions called weak, medium, and strong, respectively.

Visual stimulation. Visual background noise was continuously presented in all conditions. The noise was composed of 200 white dots (2×2 pixels) on a dark gray background. Dots moved linearly in random directions with a speed of $1^\circ/\text{s}$. Dots were uniformly distributed within the area of a notional annulus with an inner/outer diameter of $0.5/5.0^\circ$ around central fixation. On each refresh, dots falling outside the area and some of the remaining dots (with a probability of 1%) were randomly relocated within the area of the notional annulus. As visual signals, some of the dots changed for 500 ms from random motion to coherent rotation around fixation (0.88 and 0.75 rad/s in experiments 1 and 2, respectively). In experiment 1, 50% of the dots moved coherently. In experi-

Table 1. Summary of conditions in experiment 1

Condition	SOA (ms) ^a	Performance summary ^b			LATER model fit ^c		
		Misses (%)	Median (ms)	MAD (ms)	μ (s ⁻¹)	σ (s ⁻¹)	RMSE
Sound	(30)	0.5 ± 0.2	383 ± 8	72 ± 4	2.651	0.537	0.004
	(60)				(2.451)	(0.461)	(0.005)
Motion	(-30)	1.8 ± 0.4	416 ± 10	74 ± 5	2.440	0.422	0.012
	(-60)				(2.268)	(0.365)	(0.014)
Redundant	-60	0	373 ± 5	48 ± 2	2.724	0.422	0.013
	-30	0	361 ± 6	44 ± 2	2.813	0.407	0.009
	0	0	347 ± 6	43 ± 2	2.903	0.393	0.007
	30	0	362 ± 6	42 ± 2	2.788	0.360	0.010
	60	0	373 ± 6	42 ± 2	2.702	0.339	0.012

^aNegative (positive) values indicate that motion (sound) signals were delayed.

^bMean ± SEM across 40 blocks of 40 valid RTs each.

^cVincentized group distributions (averaged on the 1/RT scale) based on 1600 RTs per condition. Values in parentheses are estimates of the delayed distributions.

ment 2, 20, 30, and 50% of the dots moved coherently in the conditions weak, medium, and strong, respectively.

Task and procedures. Participants were instructed to detect target signals by pressing a button. We asked participants to respond as fast as possible but to avoid false alarms and missed targets. We used a partially self-paced, continuous stimulation paradigm. Auditory and visual background noise was presented throughout a block. After a random interval of 1500–3500 ms (uniformly distributed), a target was presented. At the time of a response, there was a new random interval until the next target was generated.

We considered responses with RTs within 100–1500 ms after signal onset as valid. Responses falling outside this range were false alarms (~2% across all conditions) and missed targets (see Tables 1, 3) if too early or too late, respectively. After an error, a feedback screen, which indicated the type of the error, interrupted the continuous stimulation for 1500 ms before a new random interval preceding the next target was triggered. If a target was missed, a corresponding trial was repeated at the end of the experiment.

Conditions were randomly interleaved within a block. In experiment 1, this included two single signal conditions and five redundant signal conditions, in which both signals were presented with one of five signal onset asynchronies (SOAs) of -60, -30, 0, 30, and 60 ms (positive values indicate that the sound signal was delayed relative to the motion signal). For each condition, we collected 45 valid responses per block. A block lasted ~30 min, including occasions for participants to take a short rest. Each participant of the first experiment was tested on 2 d with two blocks per day (i.e., each participant performed four blocks in total). In experiment 2, we presented six single conditions (only the auditory or only the visual signal was presented at one of three levels of signal strength, i.e., weak, medium, and strong) and nine redundant conditions (both signals were presented simultaneously, and we tested all 3 × 3 possible combinations of signal strength). For each condition, we collected 55 valid responses per block. A block lasted ~1 h, including several occasions (i.e., every 75 trials) for participants to take a short rest. To reduce fatigue effects, participants of the second experiment were tested on 2 d with one block per day. Participants in both experiments were familiarized with all conditions in a short practice session before the experiment proper.

RT analyses. We analyzed RTs of valid responses only. We used a reciprocal scale (1/RT) for the analysis of group RT distributions and for the modeling approach. For the presentation of our data, we back-transformed these data to the easier to understand RT scale. All analysis and modeling was performed using MATLAB (MathWorks).

Regarding the computation of group RT distributions, we noted that median RTs of one condition measured for one participant but in different blocks/sessions could differ by several tens of milliseconds, which is possibly related to learning, fatigue, or other general effects that might occur across blocks/sessions. To avoid that such differences affected the shape (and particularly the variance) of the estimated group RT distributions, we considered each block as an independent sample instead of pooling RTs across sessions.

To reduce the impact of anticipatory responses and lapses of attention, we performed an outlier correction on the basis of the valid responses in a given block. For each condition, we rejected trials with RTs deviating by >3 SDs from the mean on the reciprocal scale (<1% of the trials in all conditions). Then, to obtain equal-sized data blocks, we selected the latest 40 trials of each condition in experiment 1 and the latest 50 trials of each condition in experiment 2 (the first few trials were considered as training and were not analyzed further). In total, we collected 40 equal-sized data blocks per condition, summing up to 1600 and 2000 RTs per distribution in experiments 1 and 2, respectively. Hence, for the whole study, we analyzed 41,200 RTs.

To obtain cumulative group RT distributions, we rank ordered the RTs of each block and averaged RTs of each rank on the reciprocal scale (Vincent averaging; Ratcliff, 1979; e.g., the RT of the fastest group quantile is computed by selecting the fastest RT from each block and by averaging these fastest RTs, and so forth for the remaining ranks). To obtain continuous distribution functions, we used the linear approach to threshold with ergodic rate (LATER) model (Carpenter and Williams, 1995; Reddi et al., 2003; Fig. 1*c,d*). According to this model, empirical RT distributions can be described by reci-normal distributions (Reddi et al., 2003; the reci-normal distribution is defined as the distribution of a random variable X whose reciprocal $Y = 1/X$ is normally distributed with mean μ and SD σ). Despite its simple nature, the LATER model captures the main principles of the accumulation framework of perceptual decision making (Gold and Shadlen, 2007) and yields reasonable fits to empirical RT distributions in detection tasks with single signals (at least when signal strength is suprathreshold). We fitted the LATER model to the group quantiles of each condition by minimizing the root mean squared error (RMSE; using the “fminsearch” routine of MATLAB). Estimates of the mean and the SD of the Vincentized group RT distribution (which were averaged on the reciprocal scale) were virtually identical to the averaged estimates of the distributions in the individual blocks (which has motivated our choice to perform the Vincent averaging on the reciprocal scale). To estimate the RT distributions for delayed motion or sound signals in experiment 1, we added to all 1600 RTs of these conditions the corresponding physical delay of either 30 or 60 ms and then fitted the LATER model as described above (best-fitting estimates for these conditions are specified in parentheses in Table 1).

The probability summation framework. A straightforward explanation of the RSE is probability summation (or statistical facilitation; Raab, 1962). Using this framework, predictions for the RT distribution with redundant signals can be computed based on the RT distributions with single signals. Let P_X and P_Y be the cumulative distribution functions (CDFs) of RTs in single conditions with arbitrary signals X and Y , respectively. Then, the CDF of RTs $P_{X \cup Y}$ with redundant signals $X \cup Y$ can be computed from the following:

$$P_{X \cup Y}(t) = P_X(t) + P_Y(t) - P_{X \cap Y}(t), \quad (1)$$

where the joint probability $P_{X \cap Y}$ is given by the product of P_X and P_Y if RTs to single signals are statistically independent:

$$P_{X \cup Y}(t) = P_X(t) + P_Y(t) - P_X(t) \times P_Y(t). \quad (2)$$

To illustrate benefits attributable to probability summation and to uncover the conditions under which the largest speedup is expected, we simulated the RSE for arbitrary signals X and Y using Equation 2. We selected for signal X a fixed RT distribution chosen from a set of reciprocal normal distributions with a given median RT and median absolute deviation (MAD). Then, as an optimization problem, we searched for signal Y the reciprocal normal distribution from a set of distributions with reasonable median RTs and MADs that yielded the maximal speedup. As a measure of benefit, we considered the so-called multisensory index, which is defined as “the proportionate difference between a multisensory response to a cross-modal stimulus and the unisensory response to the most effective modality-specific component stimulus” (Stein et al., 2010). Accordingly, as a measure of response speedup in redundant compared with single conditions, we defined the benefit B with redundant signals as the area between the CDFs in the redundant and the faster of the two corresponding single signal conditions:

$$B_{X,Y} = \int_T P_{X \cup Y}(t) - \max[P_X(t), P_Y(t)] dt, \quad (3)$$

where the integral is the definite integral taken for response time t ranging over the set of possible response times T (for an illustration, see Fig. 5c). Note that this measure of benefit is based on entire RT distributions, which is a more accurate measure than for example differences in median RT (a related approach also using an integral measure was introduced by Colonius and Diederich, 2006). Note also that the term max in Equation 3, which is used as reference to compute benefits with redundant signals, corresponds to the so-called Grice bound (Grice et al., 1984), which is a lower bound for predictions based on probability summation (Townsend and Wenger, 2004). The Grice bound corresponds to the prediction assuming not statistical independence (as in Eq. 2) but a maximal positive correlation (Colonius, 1990; the effect of potential correlations on predictions is illustrated by Otto and Mamassian, 2012, their Fig. S1). Equation 3 can be rewritten as

$$B_{X,Y} = \int_T \min[P_{X \cup Y}(t) - P_X(t), P_{X \cup Y}(t) - P_Y(t)] dt, \quad (4)$$

and by substituting $P_{X \cup Y}$ using Equation 2, we obtain the benefit predicted by probability summation under the assumption of statistical independence:

$$B_{X,Y} = \int_T \min[P_Y(t)(1 - P_X(t)), P_X(t)(1 - P_Y(t))] dt. \quad (5)$$

Model fitting. To explain the exact RT distribution in redundant conditions, we used the probability summation model implemented by Otto and Mamassian (2012). This model is constrained by the reciprocal normal distributions fitted in the single signal conditions and has two degrees of freedom. First, it takes into account a correlation coefficient ρ , which is critically needed because the assumption of statistical independence can be violated, for instance, when RTs show strong history dependence (for a more detailed discussion of the history effect and its relationship to the correlation parameter, see Results, Additional observations). Second, it allows the parallel decision processes to interact by the extra noise η , which is added to the σ of the best-fitting reciprocal normal distributions in both single signal conditions. The noise interaction violates the so-called context invariance assumption (Ashby and Townsend, 1986; Luce, 1986; Townsend and Wenger, 2004), which implies that the frequently used race model test (Miller, 1982) has no additional implication here. We fitted the model to the empirical distributions by minimizing the RMSE.

Results

Simulations

The probability summation framework provides a straightforward explanation of the RSE (Raab, 1962; Otto and Mamassian, 2012). Within this framework, the speedup of RTs with redundant signals depends on a “race” between two stochastic decision

Figure 2. Numerical analysis of the principle of congruent effectiveness. We assume that RTs to a given signal X follow a fixed cumulative distribution P_X . To maximize the benefit, signal Y can be manipulated freely so that P_Y can take the form of any cumulative distribution. **a**, According to Equation 5 (see Materials and Methods), the benefit for a fixed value of P_X , here for example equal to 0.3, is given by the minimum function of the two linear functions $f_1 = P_Y(1 - P_X)$ and $f_2 = P_X(1 - P_Y)$. This minimum function is strictly increasing with P_Y when $f_1 < f_2$ and is strictly decreasing with P_Y when $f_1 > f_2$. It follows that the minimum function has an absolute maximum when $f_1 = f_2$, that is, when P_Y is equal to 0.3 as well. **b**, The same is true for any value of P_X between 0 and 1, that is, the benefit is always largest when P_Y is equal to P_X (the dotted line indicates the cross-section shown in **a**). It follows that the expected benefit using a fixed signal X is largest if signal Y is chosen so that RTs in the corresponding single conditions are equal in distribution. Moreover, this analysis shows that the principle of congruent effectiveness does not depend on distributional assumptions (as made, for example, by the model presented in Fig. 1).

processes, one for each signal. On a given trial, a response can be triggered by the faster of the two processes (therefore, also the name race model). If RT distributions in the single signal conditions overlap, such a race predicts on average a statistical benefit in redundant compared with single signal conditions (Fig. 1e).

To illustrate potential benefits that can be expected from such a race mechanism, we first considered the probability summation rule, for simplicity, under the assumption of statistical independence (Eq. 2; the assumption of statistical independence can be violated because of the history dependence of empirical RTs; Otto and Mamassian, 2012). Under this assumption, expected benefits can be directly computed based on the cumulative RT distributions in the single signal conditions (Eq. 5), that is, based on the cumulative probabilities P_X and P_Y to observe a response to signals X and Y at time t . We can now consider the optimization problem, in which we have a prearranged signal X (yielding a fixed RT distribution) and in which we are free to change signal Y (and consequently also the resulting RT distribution), to maximize the benefit in RTs that can be expected from probability summation with these two signals. To solve this problem, we first assumed that P_X is, for instance, equal to 0.3 for an arbitrary time t . We then asked how signal Y should be chosen (i.e., what would be the optimal value of P_Y) to maximize the benefit at this time t . For this question, the numerical analysis in Figure 2a shows that the maximal benefit occurs when P_Y is equal to 0.3 as well. Moreover, the same holds true for any value of P_X (between 0 and 1), that is, the benefit at any time t is always largest when P_Y is identical to P_X (Fig. 2b). It follows that the overall benefit attributable to probability summation is largest when RTs with single signals X and Y are equal in distribution. We refer to this property of the probability summation framework as the principle of congruent effectiveness, which states that benefits with redundant signals are expected to be largest when corresponding single signals yield similar performance levels.

To illustrate this principle, we simulated the RSE for arbitrary signals X and Y using the model depicted in Figure 1. For example, in the middle plot of Figure 3, RTs to signal X followed a

Figure 3. Benefits expected from probability summation. We simulated the RSE for arbitrary signals *X* and *Y* using the model illustrated in Figure 1*c–e* (under the assumption of statistical independence). In each plot, RTs to signal *X* follow a reci-normal distribution with a fixed median and MAD (as indicated by vertical and horizontal lines, respectively). RTs to signal *Y* could follow any reci-normal distribution within the plotted range of median RTs and MADs. We then computed the expected benefit for all combinations of signals *X* and *Y* (Eq. 5). These computations reveal two basic properties to predict benefits within the probability summation framework. First, for each plot, the maximal benefit is observed when RTs to signals *X* and *Y* had identical medians as well as identical MADs (or, more precisely, when RTs to signal *X* and *Y* are equal in distribution). We refer to this property as the principle of congruent effectiveness. Second, when we varied independently the median and the MAD of RTs to signal *X* across plots, benefits are mostly enhanced with increasing MADs (middle column). In contrast, manipulations of the median had basically no effect (middle row). We refer to this property as the variability rule.

prearranged reci-normal distribution with a median RT of 0.5 s and an MAD of 0.10 s. We then asked how signal *Y* should be chosen to maximize the expected benefit. To solve the optimization problem, we assumed in our simulations that signal *Y* could be chosen freely so that RTs could follow any reci-normal distribution within a reasonable range of median RTs and MADs. Next, for all potential combinations, we computed the benefit expected from probability summation (Eq. 5), which is color coded in Figure 3. As shown in the middle plot, benefits occurred over a wide range of potential RT distributions for signal *Y*. Importantly, the maximal benefit was observed when signal *Y* was chosen so that both the median and the MAD of RTs were identical to signal *X*. In contrast, almost no benefits were observed when the overlap of the RT distributions for signals *X* and *Y* was very small. To illustrate the generality of this observation, we simulated the RSE with another four prearranged RT distributions for signal *X*. Notably, the maximal benefit occurred in all simulations (i.e., in each plot of Fig. 3) when RTs to signals *X* and *Y* were identical in distribution (i.e., when both the median and the MAD were identical). Hence, each of these simulations is, as expected from our numerical analysis (Fig. 2), in accordance with the principle of congruent effectiveness.

The principle of congruent effectiveness explains that the maximal benefit in each plot of Figure 3 occurred when RTs to signals *X* and *Y* were equal in distribution. However, the maximal benefit varied substantially across plots and particularly in the middle column of plots. For these plots, the MAD of RTs to signal *X* increased from the bottom to the top. Concurrently, benefits were increased toward the top plot. In contrast, varying the median RT of signal *X* had basically no effect on the expected benefit (Fig. 3, middle row; small differences as well as the asymmetries in benefits within plots are explained by the changing skewness of the different reci-normal distributions). These observations illustrate that it is fundamentally the variability of RTs with single signals that determines the overall magnitude of benefits attributable to probability summation. If there is no variability in the responses, no benefit would occur. We refer to this second property of the probability summation framework as the variability rule, which states that benefits with redundant signals are expected to be larger when the variability of performance with single signals is increased.

Next, we tested the two principles derived from probability summation in experiments using the redundant signal paradigm with motion and sound signals (Fig. 1*a,b*). In a first experiment, we varied the effectiveness of the redundant signals by manipulating their relative onset. In a second experiment, we manipulated both the effectiveness of signals and the variability of responses by manipulating independently the physical strength of both signals.

Experiment 1: SOA

To manipulate the effectiveness of the motion and sound signals, we added small physical delays to either the sound or the motion signal. Adding such delays should shift the entire RT distribution compared with the condition without delay, without changing the variability of RTs. Hence, if probability summation is in play, the expected benefit should be predicted by the principle of congruent effectiveness only. Precisely, benefits are expected to be largest when RTs in the single conditions are identical in median (see simulations in Fig. 3). A straightforward prediction is consequently that the benefit should be largest when the added delay matches and counteracts the difference in median RTs as measured in single signal conditions (without physical delays).

Our results show that RTs in redundant conditions were fastest when none of the single signals was delayed (SOA of 0 ms; Fig. 4*a*; for an overview of conditions, see Table 1). When either the sound or the motion signal was delayed, expectedly, median RTs in the redundant conditions increased. Importantly, compared with the faster of the two single signal conditions, the speedup in median RTs was largest for the SOA of 30 ms, which closely

counteracted the median difference of 33 ms in single motion and sound conditions (Fig. 4a). This effect can be best illustrated by considering the benefits, as computed on the basis of the RT distributions (Eq. 3), as a function of the median difference taking into account the added delays (Fig. 4b). Given that the maximal benefit occurred for a median difference close to 0 ms, for which the maximal benefit was expected, this analysis showed that the RSE followed the principle of congruent effectiveness. We highlight that this result is of course perfectly in agreement with previous findings showing that the optimal temporal offset between signals, which maximizes the RSE, is highly variable and that largest benefits occur typically when median RTs in the single signal conditions including additional physical delays are similar (Hershenson, 1962; Miller, 1986).

To explain the exact distribution of RTs in the redundant conditions, we fitted the probability summation model as implemented by Otto and Mamassian (2012) to the empirical distributions. This model is constrained by the RT distributions in single conditions and has two degrees of freedom (i.e., the correlation ρ and the additional noise η). It should be noted that the correlation parameter ρ has an additional effect on the expected benefit, which is expected to be largest for maximally negatively correlated RTs and zero for maximally positively correlated RTs (Colonius, 1990; Otto and Mamassian, 2012). Critically, when signals are presented in random order as in the present experiment, Otto and Mamassian (2012) showed that RTs are negatively correlated because of trial history effects. Hence, benefits can be larger than predicted by probability summation under the assumption of statistical independence (Eq. 2). We fitted the model to the RT distribution in the redundant condition with an SOA of 0 ms (Fig. 4c; for best-fitting parameters, see Table 2). The model not only explained the empirical distribution well but also, by keeping the best-fitting parameters, predicted the benefit in the remaining conditions reasonably well (Fig. 4d). In summary, our results show that the probability summation framework and the newly formulated principle of congruent effectiveness account very well for the RSE when the relative onset times of the redundant signals are manipulated.

Experiment 2: signal strength

In the second experiment, we manipulated the signal strength of both signals independently in a 3×3 design. Manipulations of signal strength typically have an effect on both the location and the variability of RT distributions (Wagenmakers and Brown, 2007). Consequently, this experiment provides a critical test of both the principle of congruent effectiveness and the variability rule as predicted by probability summation.

Our manipulations of signal strength were very effective. For example, with single sound signals, changing from weak to strong

Figure 4. The RSE for different values of SOA (experiment 1). *a*, Median RTs as a function of SOA (mean across 40 blocks of 40 trials each; error bars indicate the SEM). RTs were fastest for redundant signals presented with an SOA of 0 ms. Critically, the speedup compared with the corresponding single signal conditions was largest when the SOA was 30 ms (broken lines indicate projected median RTs for single signals for different values of SOA). *b*, Benefits as a function of the difference in median RT in single conditions (negative values indicate that median RTs to sound compared with motion signals were faster taking into account the delays resulting from the SOA manipulation; labeling of SOA conditions as in *d*). As expected from probability summation, benefits were largest when the difference in median RTs approached 0 ms. Estimated benefits were based on cumulative group RT distributions with 1600 trials per distribution. Vertical lines show 95% confidence intervals, calculated by 1000 repetitions of a bootstrap procedure. *c*, Cumulative group RT distributions in the condition with an SOA of 0 ms. For each distribution, symbols indicate the empirical quantiles (each quantile is based on 40 trials). In the single conditions, the dotted lines indicate best fitting reci-normal distributions (see Table 1). Based on these distributions, we fitted the probability summation model with the correlation ρ and the additional noise η as free parameters to the empirical distribution with redundant signals (Fit). The model fitted the empirical distribution very well (for best-fitting model parameters, see Table 2). M, Motion; S, sound; R, redundant. *d*, Benefits (see *b*) as a function of predictions based on the model fitted in the redundant condition with an SOA of 0 ms (see *c*). Interestingly, the model predicted the benefits for the remaining SOA conditions reasonably well.

Table 2. Best-fitting race model parameters (experiment 1)

SOA (ms) ^a	ρ	η (s ⁻¹)	η (%) ^b	RMSE
-60	-0.39	0.017	~4%	0.011
-30	-0.49	0.059	~13%	0.011
0	-0.48	0.060	~13%	0.016
30	-0.49	0.076	~17%	0.019
60	-0.57	0.086	~21%	0.024

^aNegative (positive) values indicate that motion (sound) signals were delayed.

^bCompared with the average σ in corresponding single signal conditions (see Table 1).

signals decreased the median RT from 483 to 401 ms and the MAD from 123 to 67 ms (Table 3). A stronger effect was found with single motion signals: changing from weak to strong signals decreased the median RT from 573 to 417 and the MAD from 196 to 73 ms. It should be noted that our manipulations also had an effect on accuracy. Although performance was close to perfect in most conditions, this was particularly not the case for weak motion signals with an average miss rate of 21.2%. We recall that, with our experimental procedures, missed trials were repeated at the end of an experimental block until a sufficient number of valid responses was collected (see Materials and Methods, Task and procedures). Hence, regarding the estimated RT distribution, missed trials were replaced by valid trials, which led to an

Table 3. Summary of conditions in experiment 2

Condition	Signal strength		Performance summary ^a			LATER model fit ^b		
	Sound	Motion	Misses (%)	Median (ms)	MAD (ms)	μ (s ^{−1})	σ (s ^{−1})	RMSE
Single	Strong		0.6 ± 1.8	401 ± 9	67 ± 5	2.547	0.468	0.008
	Medium		1.2 ± 3.3	444 ± 9	88 ± 5	2.301	0.481	0.008
	Weak		5.5 ± 6.2	483 ± 10	123 ± 5	2.106	0.552	0.007
		Strong	1.2 ± 2.1	417 ± 9	73 ± 4	2.438	0.436	0.006
		Medium	8.0 ± 7.6	495 ± 11	119 ± 6	2.079	0.494	0.006
		Weak	21.2 ± 12.6	573 ± 10	196 ± 8	1.777	0.621	0.013
Redundant	Strong	Strong	0.1 ± 0.6	362 ± 7	45 ± 3	2.811	0.384	0.006
	Strong	Medium	0.5 ± 2.6	379 ± 8	48 ± 2	2.700	0.396	0.006
	Strong	Weak	0.6 ± 1.7	386 ± 7	56 ± 3	2.630	0.414	0.005
	Medium	Strong	0.3 ± 1.6	373 ± 7	53 ± 3	2.719	0.433	0.007
	Medium	Medium	0.4 ± 1.3	400 ± 8	62 ± 4	2.533	0.416	0.008
	Medium	Weak	1.0 ± 2.0	417 ± 8	69 ± 3	2.450	0.440	0.006
	Weak	Strong	0.4 ± 1.4	388 ± 8	58 ± 4	2.635	0.439	0.007
	Weak	Medium	0.7 ± 2.3	416 ± 8	71 ± 4	2.444	0.446	0.007
	Weak	Weak	2.5 ± 4.4	448 ± 9	94 ± 4	2.274	0.504	0.004

^aMean ± SEM across 40 blocks of 50 valid RTs each.
^bVincitized group distributions (averaged on the 1/RT scale) with 2000 RTs per condition.

Figure 5. RT distributions (experiment 2). *a–i*, Each plot shows the cumulative group RT distributions for the combination of motion and sound signals at different levels of signal strength [bottom to top rows: weak (W), medium (M), and strong (S) sound signals; left to right columns: weak, medium, and strong motion signals]. For single signal conditions (Motion and Sound), broken lines indicate the best-fitting reci-normal distributions (for best-fitting parameters, see Table 3; empirical quantiles are not shown for clarity). For redundant signal conditions, dots indicate the empirical quantiles of the cumulative group RT distribution (each of the 50 quantiles is based on 40 trials, i.e., each distribution is based on 2000 trials). We fitted the probability summation model, which is constrained by the single signal conditions and has two degrees of freedom (i.e., the correlation ρ and the additional noise η), to the empirical quantiles in the redundant condition with medium signals (see *e*). The model fitted the data virtually perfectly (as indicated by the solid dark gray line; Fit). Interestingly, by keeping identical parameters, the model predicted the distributions of the remaining conditions, except for the condition with a pair of weak signals (see *g*), surprisingly well (as indicated by the solid light gray lines; Prediction). For the subsequent analysis shown in Figure 6, we computed the benefit for each pair of signals based on the area between the CDF with redundant signals and the faster of the CDFs with single signals (as illustrated in *c*; see also Eq. 3).

Figure 6. Benefits for different combinations of signal strength (experiment 2). *a*, Benefits as a function of signal strength of motion and sound signals (w, weak; m, medium; s, strong signals). The arrows highlight the conditions with pairs of weak, medium, and strong signals in both modalities. *b*, Benefits as a function of the parameter free predictions based on probability summation under the assumption of statistical independence (Eq. 5). *c*, Benefits as a function of predictions based on the model fitted in the redundant condition with medium signals (see Fig. 5*e*). Benefits in each condition were estimated based on the cumulative group RT distributions with 2000 trials per distribution (as illustrated in Fig. 5*c*). Vertical lines show 95% confidence intervals, calculated by 1000 repetitions of a bootstrap procedure.

underestimation of the RT distribution (toward faster and less variable responses). Finally, with redundant signals, median RTs in the nine conditions varied between 362 and 448 ms and the MAD between 45 and 94 ms (a summary of all conditions is shown in Table 3; cumulative group RT distributions are shown in Fig. 5).

As in experiment 1, our main interest was in the benefit in redundant compared with single signal conditions. To obtain precise estimates of benefits, we considered again the area between the CDF in the redundant and the faster CDF in the single signal conditions (Eq. 3; for an illustration, see Fig. 5*c*). As shown in Figure 6*a*, our manipulation of signal strength was very effective regarding benefits, which were highly variable across conditions. The smallest benefit of only 17 ms occurred for the combination of strong sound and weak motion signals. The largest benefit of 84 ms occurred for the combination of weak sound and medium motion signals. However, the exact benefit was hardly predicted by the physical signal strength and appeared at first sight random in Figure 6*a*. Notably, when considering only the combinations of two weak, two medium, and two strong signals, the PoIE, when interpreted in terms of physical signal strength, predicts a large benefit for the pair of weak signals and gradually less benefit as signals get stronger. However, empirical benefits for these three pairs were rather constant at ~50 ms in these conditions (Fig. 6*a*, arrows). Hence, our results agree with a recent study showing that the RSE does not follow the PoIE when interpreted in terms of physical signal strength (Chandrasekaran et al., 2011; but see Senkowski et al., 2011).

Strikingly, the pattern of benefits that appeared random when considered as a function of signal strength got immediately structured when we considered the probability summation framework with the two newly formulated principles. For example, regarding the principle of congruent effectiveness, we should note that RTs with strong signals were rather similar for motion and sound signals (median difference, 22 ms) but highly different for weak signals (median difference, 118 ms). Hence, following this principle, we would expect that the benefit with our signals should decrease when switching from strong to weak signals. Conversely, regarding the variability rule, we should note that the MAD of RTs was much larger for weak (123 and 196 ms, respectively) than for strong (63 and 73 ms, respectively) signals. Hence, following this rule, we would expect that the benefit with our signals should

Table 4. Best-fitting race model parameters (experiment 2)

Sound	Motion	ρ	η (s^{-1})	η (%) ^a	RMSE
Strong	Strong	−0.31	0.057	~13%	0.008
	Medium	−0.43	0.043	~9%	0.009
	Weak	−0.26	0.001	~0%	0.010
Medium	Strong	−0.24	0.108	~24%	0.007
	Medium	−0.25	0.055	~11%	0.010
	Weak	−0.22	0.029	~5%	0.005
Weak	Strong	−0.17	0.075	~15%	0.005
	Medium	−0.24	0.057	~11%	0.007
	Weak	+0.15	0.010	~2%	0.004

^aCompared with the average σ in corresponding single signal conditions (see Table 3).

increase when switching from strong to weak signals. Hence, together, the two effects cancelled out each other, and it is consequently not surprising that the benefit with our signals stayed rather constant when changing from pairs of strong to pairs of weak signals (Fig. 6*a*, arrows).

How well the probability summation framework with its two principles accounts for the observed benefits can be demonstrated using the probability summation rule assuming statistical independence (Eq. 2). For each redundant condition, we computed the parameter free prediction of the benefit based on the RT distributions in the corresponding single signal conditions (Eq. 5). Strikingly, the empirical benefits followed the expected benefit remarkably well (Fig. 6*b*). One outlier occurred for the combination of weak sound and motion signals, in which the empirical benefit was slightly smaller than predicted (which is a finding opposite to the other conditions). This effect might be explained by the fact that we have underestimated the RT distribution in the weak motion condition (see above) and, consequently, that the prediction based on probability summation is underestimated as well.

Finally, to explain the exact distribution of RTs in the redundant conditions, we used as in experiment 1 the probability summation model as implemented by Otto and Mamassian (2012). As illustrated in Figure 5*e*, the model fitted the RT distribution, for example, in the redundant condition with a pair of medium signals virtually perfectly (for best-fitting parameters, see Table 4). Strikingly, keeping the best-fitting parameters of this condition, the model predicted the RT distributions in the remaining eight conditions remarkably well (Fig. 5). This strong explanation

tory power is best illustrated by plotting the empirical benefits of all nine redundant conditions as a function of the predicted benefit based on the model fitted to the condition with medium signals (Fig. 6c). The outstanding agreement of predicted and empirical benefits is mainly achieved by assuming a small negative correlation in RTs to single signals, which appeared to be rather consistent across conditions (Table 4; see also the next section below). The noise parameter, which appeared to be more variable across conditions, has only a minor effect on predicted benefits. In summary, our results show that the probability summation framework, with the newly formulated principle of congruent effectiveness and the variability rule, accounts very well for the RSE when the physical strength of the redundant signals is manipulated.

Additional observations

We point out some additional observations. First, predictions based on probability summation depend critically on the quality of the extracted RT distributions in single signal conditions. This dependency is evident in experiment 2 in the condition with a pair of weak signals but points to a more general issue. For example, we used the appealingly simple (cf. Gold and Shadlen, 2007) LATER model and, consequently, reci-normal distributions to estimate continuous distribution functions. However, reci-normal distributions (despite their advantage to compute predictions including the correlation ρ , which we use to account for the trial history effect) are not necessarily the best distributions to describe empirical RT distributions. Moreover, the LATER model appears too simple to be consistent with physiological recordings of neurons involved in decision making (Churchland et al., 2011). Regarding our modeling approach, we can say that reci-normal distributions explained our data reasonably well. For example, to explain the RT distributions with redundant signals in experiment 2, the probability summation model with two degrees of freedom, which are constrained by the single signal RT distributions, provided fits (average RMSE, 0.007; Table 4) that are only marginally worse than the LATER model, which has also two degrees of freedom but which are not constrained (average RMSE, 0.006; Table 3). Most importantly, we highlight that the principles derived from the probability summation framework (see Eqs. 1–5 and Fig. 2) do not depend on the distributional assumptions made by our modeling approach using the LATER model and, hence, are rather general.

Second, as mentioned previously, potential correlations can have a huge impact on expected benefits. In this context, it is important to consider trial history effects, that is, the slowdown of RTs when signals (or tasks) are switched (Spence et al., 2001; Monsell, 2003; Waszak et al., 2003). If RTs are simply pooled across trials to estimate cumulative RT distributions (which is a standard procedure), such history effects systematically contribute to the variability of the RT distributions because rather fast RTs after a signal/modality repetition are mixed with rather slow RTs after a signal/modality switch. If redundant signals are presented, for example, after a single auditory signal, there is a repetition of the auditory signal but a switch regarding the visual signal (and vice versa for redundant signals presented after a single visual signal). Hence, the trial history effect generates an (artificial) negative correlation in the RT distributions, which needs to be considered for predictions based on probability summation. Although Otto and Mamassian (2012) assumed a correlation of -0.59 to account for the RSE in their experiment with randomly presented signals, we here found correlations of only -0.48 ± 0.03 and -0.27 ± 0.03 in the conditions of experiments

1 and 2, respectively (Tables 2, 4). A simple explanation might be that the number of redundant compared with single conditions was larger in the present experiments so that more trials were preceded by trials with signals in both modalities. Hence, the frequency of modality switches was reduced. In addition, the probability of repeating exactly identical signals was reduced in experiment 2 because of the different levels of signal strength. Compared with our previous study, these effects should have reduced the systematic influence of the trial history in the distributions (because most RTs were determined after a switch) and, consequently, yielded smaller negative correlations as revealed by the model fitting. Critically, within each of the two experiments here, the history effect (i.e., the extracted correlation) was rather constant across conditions. However, it is important to note that this is not necessarily always the case. If the history effect is changing across conditions for some reason, the resulting correlations have an additional effect on expected benefits, which has to be considered in addition to the two main principles. In particular, this could be the case if benefits are compared across different subject populations such as in patient studies (for example, when a patient population shows a stronger history effect compared with a control group).

Finally, Otto and Mamassian (2012) proposed a noise interaction when two distinct signals are processed in parallel (for a potentially related finding of reduced efficiency in parallel evidence accumulation using the psychological refractory period paradigm, see Zylberberg et al., 2012). Consequently, we expected that noise, as estimated by our modeling approach, should not be constant but increase the more likely it is that two decision processes occur simultaneously. This is likely the case when single signal conditions show equal median RTs, which is the condition for which we expect the noise interaction to be largest. In contrast, when one signal is more effective than the other in triggering a response (i.e., when single signal conditions differ in median RTs), the noise interaction should decrease. For example, by adding a physical delay to one signal, the other signal should be processed without additional noise during the delay period. When the delay is increased, the noise interaction is expected to decrease further. A similar rationale holds when the strength of signals is manipulated. Interestingly, taking into account the best-fitting noise estimates of the 14 redundant conditions tested in this study (Tables 2, 4), we found a corresponding relationship. The noise interaction was largest when the median difference between corresponding single signal conditions was small (Fig. 7). Furthermore, it may appear that the maximal noise interaction occurred when our motion signals slightly preceded the sound signals, which may point to an asymmetry in the noise interaction. However, we cannot completely describe this relationship here because we miss conditions in which median RTs to motion signals are much faster than to sound signals.

Discussion

Our theoretical considerations within the probability summation framework (Raab, 1962; Otto and Mamassian, 2012) revealed that the RSE should follow two main principles. First, the principle of congruent effectiveness states that benefits with redundant signals are expected to be largest when the corresponding single signals yield similar performance levels. Second, the variability rule states that benefits with redundant signals are expected to be largest when responses in corresponding unisensory conditions are variable. Our systematic analysis of RT distributions shows that the RSE with distinct multisensory signals follows these principles very well, which provides compelling evidence in favor of

Figure 7. Additional noise. For each of the 14 tested redundant signal conditions, we show the best-fitting estimates of the additional noise parameter η as a function of the median difference in the corresponding single signal conditions. Estimates of the noise were based on the model fitting with cumulative group RT distributions with 1600 and 2000 trials per distribution in experiments 1 and 2, respectively (see Tables 2, 4). Negative values of median difference indicate that median RTs to sound signals were faster compared with motion signals.

probability summation as the underlying combination rule. This finding implies further that distinct multisensory signals are actually not integrated within a single decision process but processed in parallel and that distinct multisensory signals are flexibly coupled by basic logical decision gates only after unisensory decisions have been made (Otto and Mamassian, 2012). Whether such parallel processing will also apply to situations in which stimuli are not distinct (as tested here and in most experiments on the RSE) but refer to a unique environmental property remains to be seen.

It is worth noting that similar principles to predict benefits have been extracted previously in experiments with the precision of responses as behavioral measure and using signals that refer to a unique environmental property, such as the size of an object or its position in space. For example, many reports show that humans and monkeys combine information from different sensory cues in a very efficient way (van Beers et al., 1996; Jacobs, 1999; Ernst and Banks, 2002; Alais and Burr, 2004; Fetsch et al., 2012). Interestingly, maximum likelihood estimation, as the underlying combination rule, predicts that benefits from optimal cue combination are largest when the single cues, tested in isolation, yield congruent performance levels, which corresponds to the proposed principle of congruent effectiveness. Moreover, maximum likelihood estimation predicts that benefits from optimal cue combination are proportional to the uncertainty associated with the single cues, which corresponds well to the proposed variability rule. Given these concordant principles despite the fundamental differences in behavioral measures (speed vs precision), physical signals (distinct signals that are redundant only by the definition of the task vs signals that refer to a unique environmental property), and assumed combination rules (probability summation vs maximum likelihood integration), these principles may be considered as general principles of multisensory behavior with redundant signals. However, although research on perceptual decision making with one sensory signal has converged on a compelling framework to explain both the speed and the accuracy of responses and tradeoffs between the two (Gold and Shadlen, 2007; Bogacz et al., 2010), such a unique framework is critically missing for multisensory responses (for a recent attempt to implement optimal cue combination within the accumulation framework of perceptual decision making, see Drugowitsch et al., 2010).

Importantly, the proposed principles resolve discrepancies reported in previous studies. For example, Senkowski et al. (2011) reported recently that the RSE follows the PoIE. In this study, mean RTs to auditory and visual signals were more similar in conditions with weak compared with strong signals (Senkowski et al., 2011, their Fig. 3). Thus, following the principle of congruent effectiveness, the RSE should be larger for weak signals (the variability of RTs was not reported, but we presume that the variability rule would point in the same direction). Consequently, the RSE appears here to be in accordance with the PoIE. In contrast, Chandrasekaran et al. (2011) reported that the RSE does not follow the PoIE. In this study, manipulations were mostly effective for auditory signals, and mean RTs to auditory and visual signals were more similar for strong compared with weak auditory signals (Chandrasekaran et al., 2011, their Fig. 3). Consequently, the RSE appears here to violate the PoIE. However, it should be obvious that both studies are actually in agreement with the principles of multisensory behavior proposed here. Moreover, in contrast to the opposite claims regarding the PoIE, it is worth mentioning that both studies, as many follow-up studies of Miller (1982), agreed on the rejection of the probability summation framework, a view that we challenged recently (Otto and Mamassian, 2012). To the best of our knowledge, the probability summation framework together with the proposed noise interaction is to date the only model that, based on the RT distributions for single signals, directly explains the exact shape of the RT distribution with redundant signals.

Given that the redundant signal paradigm is frequently used in studies on sensory processing within and across senses and in both humans and nonhumans (Diederich and Colonius, 1987; Giray and Ulrich, 1993; Mordkoff and Yantis, 1993; Hughes et al., 1994; Schröger and Widmann, 1998; Gondon et al., 2004; Molholm et al., 2004; Murray et al., 2005; Katzner et al., 2006; Whitchurch and Takahashi, 2006; Martuzzi et al., 2007; Collignon et al., 2008; Hecht et al., 2008; Hirokawa et al., 2008; Tamietto and de Gelder, 2008; Poom, 2009; Suied et al., 2009; Zehetleitner et al., 2009; Cappe et al., 2010; Veldhuizen et al., 2010; Chandrasekaran et al., 2011; Senkowski et al., 2011; Brang et al., 2012; Buchholz et al., 2012), we think it will be very important for a broad audience to consider the implications of our theoretical framework. For example, studies on the RSE have contributed to the view that multisensory interactions take place in sensory cortical areas that have classically been considered to be unisensory (Foxe and Schroeder, 2005; Ghazanfar and Schroeder, 2006; Kayser, 2010). Following the rejection of the probability summation framework as an explanation of the RSE and findings indicating that multisensory interactions can occur as early as 40 ms after stimulus onset (Giard and Peronnet, 1999; Molholm et al., 2002, 2004; Murray et al., 2005), these early interactions suggested that multisensory signals converge in a feedforward manner within a single decision process. However, our analysis of RT distributions shows that such a feedforward convergence or pooling of sensory evidence is not required to explain the RSE. Instead, our framework rather suggests that sensory evidence for distinct multisensory signals is accumulated separately in parallel decision processes that are coupled by flexible logical decision gates (Otto and Mamassian, 2012). Based on our behavioral data, we still observe two major interactions in the processing of distinct signals that may correspond to the observed early multisensory interactions. On the one hand, the processing of distinct signals that are presented in a random order shows strong history dependences (Spence et al., 2001; Monsell, 2003; Waszak et al., 2003), which corresponds to the negative

correlation that is required to explain the RSE in these conditions (Otto and Mamassian, 2012). If early sensory interactions relate to these history effects, this would imply that the early interactions are feedback rather than feedforward because they depend on previously processed signals and the actual state of the observer. On the other hand, we found that the parallel processing of redundant signals yielded RTs that are more variable than predicted by probability summation. To account for this increased variability, we proposed that parallel perceptual decision processes mutually interact by increased noise (for a review on noise in the nervous system, see Faisal et al., 2008; for a potentially related finding, see Zylberberg et al., 2012). Interestingly, we showed here that the proposed noise interaction is largest when the RT distributions of the corresponding single conditions overlap and tends toward zero when, presumably, the decision processes are temporally non-overlapping (Fig. 7). Although we can only speculate about exact sources, increased noise may be related to a recent electrophysiological result on early interactions with distinct auditory and tactile signals (Lemus et al., 2010), which shows that the activity of some neurons in early sensory cortices is affected by signals in the nonpreferred modality. Critically, this activity is unspecific regarding the features of the signals and may thus relate to the proposed noise interaction. In any case, we are confident that the probability summation framework with the revealed interactions in the processing of distinct multisensory signals will contribute to solve the puzzle of the early multisensory interactions (Kayser, 2010).

Finally, we should come back to the principles of multisensory integration as formulated based on the electrophysiology of the superior colliculus (Meredith and Stein, 1983; Stein and Meredith, 1993; Stein and Stanford, 2008). It is well known that this structure is primarily involved in the control of eye movements. Therefore, it is worth noting that the RSE can be observed in the latencies of eye movements in a very similar manner compared with the RTs of manual responses (Hughes et al., 1994). Hence, we are rather confident that the speedup of saccadic latencies can be equally explained by the probability summation framework and, thus, should follow the principles described here as well. Future research may address the question whether the speedup of neuronal responses (Rowland et al., 2007) could be explained by probability summation as well. Such an analysis would bridge the apparent gap between the principles described for behavior and the principles described for electrophysiology.

References

- Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 14:257–262. [CrossRef Medline](#)
- Ashby FG, Townsend JT (1986) Varieties of perceptual independence. *Psychol Rev* 93:154–179. [CrossRef Medline](#)
- Bogacz R, Wagenmakers EJ, Forstmann BU, Nieuwenhuis S (2010) The neural basis of the speed-accuracy tradeoff. *Trends Neurosci* 33:10–16. [CrossRef Medline](#)
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436. [CrossRef Medline](#)
- Brang D, Williams LE, Ramachandran VS (2012) Grapheme-color synesthesia shows enhanced crossmodal processing between auditory and visual modalities. *Cortex* 48:630–637. [CrossRef Medline](#)
- Buchholz VN, Goonetilleke SC, Medendorp WP, Corneil BD (2012) Greater benefits of multisensory integration during complex sensorimotor transformations. *J Neurophysiol* 107:3135–3143. [CrossRef Medline](#)
- Cappe C, Murray MM, Barone P, Rouiller EM (2010) Multisensory facilitation of behavior in monkeys: effects of stimulus intensity. *J Cogn Neurosci* 22:2850–2863. [CrossRef Medline](#)
- Cappe C, Thelen A, Romei V, Thut G, Murray MM (2012) Looming signals reveal synergistic principles of multisensory integration. *J Neurosci* 32:1171–1182. [CrossRef Medline](#)
- Carpenter RH, Williams ML (1995) Neural computation of log likelihood in control of saccadic eye-movements. *Nature* 377:59–62. [CrossRef Medline](#)
- Chandrasekaran C, Lemus L, Trubanova A, Gondon M, Ghazanfar AA (2011) Monkeys and humans share a common computation for face/voice integration. *PLoS Comput Biol* 7:e1002165. [CrossRef Medline](#)
- Churchland AK, Kiani R, Chaudhuri R, Wang XJ, Pouget A, Shadlen MN (2011) Variance as a signature of neural computations during decision making. *Neuron* 69:818–831. [CrossRef Medline](#)
- Collignon O, Girard S, Gosselin F, Roy S, Saint-Amour D, Lassonde M, Lepore F (2008) Audio-visual integration of emotion expression. *Brain Res* 1242:126–135. [CrossRef Medline](#)
- Colonius H (1990) Possibly dependent probability summation of reaction-time. *J Math Psychol* 34:253–275. [CrossRef](#)
- Colonius H, Diederich A (2006) The race model inequality: interpreting a geometric measure of the amount of violation. *Psychol Rev* 113:148–154. [CrossRef Medline](#)
- Diederich A, Colonius H (1987) Intersensory facilitation in the motor component: a reaction-time analysis. *Psychol Res* 49:23–29. [CrossRef](#)
- Driver J, Noesselt T (2008) Multisensory interplay reveals crossmodal influences on “sensory-specific” brain regions, neural responses, and judgments. *Neuron* 57:11–23. [CrossRef Medline](#)
- Drugowitsch J, Pouget A, DeAngelis GC, Angelaki DE (2010) Optimal decision-making in multisensory integration. Presented at the Computational and Systems Neuroscience Conference, Salt Lake City, UT, February 25 to March 2, 2010.
- Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429–433. [CrossRef Medline](#)
- Ernst MO, Bühlhoff HH (2004) Merging the senses into a robust percept. *Trends Cogn Sci* 8:162–169. [CrossRef Medline](#)
- Faisal AA, Selen LP, Wolpert DM (2008) Noise in the nervous system. *Nat Rev Neurosci* 9:292–303. [CrossRef Medline](#)
- Fetsch CR, Pouget A, DeAngelis GC, Angelaki DE (2012) Neural correlates of reliability-based cue weighting during multisensory integration. *Nat Neurosci* 15:146–154. [CrossRef Medline](#)
- Foxe JJ, Schroeder CE (2005) The case for feedforward multisensory convergence during early cortical processing. *Neuroreport* 16:419–423. [CrossRef Medline](#)
- Ghazanfar AA, Schroeder CE (2006) Is neocortex essentially multisensory? *Trends Cogn Sci* 10:278–285. [CrossRef Medline](#)
- Giard MH, Peronnet F (1999) Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci* 11:473–490. [CrossRef Medline](#)
- Giray M, Ulrich R (1993) Motor coactivation revealed by response force in divided and focused attention. *J Exp Psychol Hum Percept Perform* 19:1278–1291. [CrossRef Medline](#)
- Gold JL, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574. [CrossRef Medline](#)
- Gondan M, Lange K, Rösler F, Röder B (2004) The redundant target effect is affected by modality switch costs. *Psychonom Bull Rev* 11:307–313. [CrossRef Medline](#)
- Grice GR, Canham L, Gwynne JW (1984) Absence of a redundant-signals effect in a reaction-time task with divided attention. *Percept Psychophys* 36:565–570. [CrossRef Medline](#)
- Hecht D, Reiner M, Karni A (2008) Enhancement of response times to bi- and tri-modal sensory stimuli during active movements. *Exp Brain Res* 185:655–665. [CrossRef Medline](#)
- Hershenson M (1962) Reaction time as a measure of intersensory facilitation. *J Exp Psychol* 63:289–293. [CrossRef](#)
- Hirokawa J, Bosch M, Sakata S, Sakurai Y, Yamamori T (2008) Functional role of the secondary visual cortex in multisensory facilitation in rats. *Neuroscience* 153:1402–1417. [CrossRef Medline](#)
- Hughes HC, Reuter-Lorenz PA, Nozawa G, Fendrich R (1994) Visual-auditory interactions in sensorimotor processing: saccades versus manual responses. *J Exp Psychol Hum Percept Perform* 20:131–153. [CrossRef Medline](#)
- Jacobs RA (1999) Optimal integration of texture and motion cues to depth. *Vision Res* 39:3621–3629. [CrossRef Medline](#)
- Katzner S, Busse L, Treue S (2006) Feature-based attentional integration of color and visual motion. *J Vis* 6:269–284. [CrossRef Medline](#)

- Kayser C (2010) The multisensory nature of unisensory cortices: a puzzle continued. *Neuron* 67:178–180. [CrossRef Medline](#)
- Lemus L, Hernández A, Luna R, Zainos A, Romo R (2010) Do sensory cortices process more than one sensory modality during perceptual judgments? *Neuron* 67:335–348. [CrossRef Medline](#)
- Luce RD (1986) Response times: their role in inferring elementary mental organization. New York: Oxford UP.
- Martuzzi R, Murray MM, Michel CM, Thiran JP, Maeder PP, Clarke S, Meuli RA (2007) Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cereb Cortex* 17:1672–1679. [CrossRef Medline](#)
- Meredith MA, Stein BE (1983) Interactions among converging sensory inputs in the superior colliculus. *Science* 221:389–391. [CrossRef Medline](#)
- Miller J (1982) Divided attention: evidence for coactivation with redundant signals. *Cognit Psychol* 14:247–279. [CrossRef Medline](#)
- Miller J (1986) Time course of coactivation in bimodal divided attention. *Percept Psychophys* 40:331–343. [CrossRef Medline](#)
- Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, Foxe JJ (2002) Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res Cogn Brain Res* 14:115–128. [CrossRef Medline](#)
- Molholm S, Ritter W, Javitt DC, Foxe JJ (2004) Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. *Cereb Cortex* 14:452–465. [CrossRef Medline](#)
- Monsell S (2003) Task switching. *Trends Cogn Sci* 7:134–140. [CrossRef Medline](#)
- Mordkoff JT, Yantis S (1993) Dividing attention between color and shape: evidence of coactivation. *Percept Psychophys* 53:357–366. [CrossRef Medline](#)
- Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, Schroeder CE, Foxe JJ (2005) Grabbing your ear: rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb Cortex* 15:963–974. [CrossRef Medline](#)
- Otto TU, Mamassian P (2012) Noise and correlations in parallel perceptual decision making. *Curr Biol* 22:1391–1396. [CrossRef Medline](#)
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442. [CrossRef Medline](#)
- Poom L (2009) Integration of colour, motion, orientation, and spatial frequency in visual search. *Perception* 38:708–718. [CrossRef Medline](#)
- Raab DH (1962) Statistical facilitation of simple reaction times. *Trans N Y Acad Sci* 24:574–590. [CrossRef Medline](#)
- Ratcliff R (1979) Group reaction-time distributions and an analysis of distribution statistics. *Psychol Bull* 86:446–461. [CrossRef Medline](#)
- Reddi BA, Asrress KN, Carpenter RH (2003) Accuracy, information, and response time in a saccadic decision task. *J Neurophysiol* 90:3538–3546. [CrossRef Medline](#)
- Rowland BA, Quessy S, Stanford TR, Stein BE (2007) Multisensory integration shortens physiological response latencies. *J Neurosci* 27:5879–5884. [CrossRef Medline](#)
- Schröger E, Widmann A (1998) Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology* 35:755–759. [CrossRef Medline](#)
- Senkowski D, Saint-Amour D, Höfle M, Foxe JJ (2011) Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *Neuroimage* 56:2200–2208. [CrossRef Medline](#)
- Spence C, Nicholls ME, Driver J (2001) The cost of expecting events in the wrong sensory modality. *Percept Psychophys* 63:330–336. [CrossRef Medline](#)
- Stein BE, Meredith MA (1993) Merging of the senses. Cambridge, MA: Massachusetts Institute of Technology.
- Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. *Nat Rev Neurosci* 9:255–266. [CrossRef Medline](#)
- Stein BE, Burr D, Constantinidis C, Laurienti PJ, Alex Meredith M, Perrault TJ Jr, Ramachandran R, Röder B, Rowland BA, Sathian K, Schroeder CE, Shams L, Stanford TR, Wallace MT, Yu L, Lewkowicz DJ (2010) Semantic confusion regarding the development of multisensory integration: a practical solution. *Eur J Neurosci* 31:1713–1720. [CrossRef Medline](#)
- Stevenson RA, Bushmakim M, Kim S, Wallace MT, Puce A, James TW (2012) Inverse effectiveness and multisensory interactions in visual event-related potentials with audiovisual speech. *Brain Topogr* 25:308–326. [CrossRef Medline](#)
- Suied C, Bonneel N, Viaud-Delmon I (2009) Integration of auditory and visual information in the recognition of realistic objects. *Exp Brain Res* 194:91–102. [CrossRef Medline](#)
- Tamietto M, de Gelder B (2008) Affective blindsight in the intact brain: neural interhemispheric summation for unseen fearful expressions. *Neuropsychologia* 46:820–828. [CrossRef Medline](#)
- Todd J (1912) Reaction to multiple stimuli. *Arch Psychol* 3:1–65.
- Townsend JT, Wenger MJ (2004) A theory of interactive parallel processing: new capacity measures and predictions for a response time inequality series. *Psychol Rev* 111:1003–1035. [CrossRef Medline](#)
- van Beers RJ, Sittig AC, Denier van der Gon JJ (1996) How humans combine simultaneous proprioceptive and visual position information. *Exp Brain Res* 111:253–261. [CrossRef Medline](#)
- Veldhuizen MG, Shepard TG, Wang MF, Marks LE (2010) Coactivation of gustatory and olfactory signals in flavor perception. *Chem Senses* 35:121–133. [CrossRef Medline](#)
- Wagenmakers EJ, Brown S (2007) On the linear relation between the mean and the standard deviation of a response time distribution. *Psychol Rev* 114:830–841. [CrossRef Medline](#)
- Waszak F, Hommel B, Allport A (2003) Task-switching and long-term priming: role of episodic stimulus-task bindings in task-shift costs. *Cogn Psychol* 46:361–413. [CrossRef Medline](#)
- Whitchurch EA, Takahashi TT (2006) Combined auditory and visual stimuli facilitate head saccades in the barn owl (*Tyto alba*). *J Neurophysiol* 96:730–745. [CrossRef Medline](#)
- Zehetleitner M, Krummenacher J, Müller HJ (2009) The detection of feature singletons defined in two dimensions is based on salience summation, rather than on serial exhaustive or interactive race architectures. *Atten Percept Psychophys* 71:1739–1759. [CrossRef Medline](#)
- Zylberberg A, Ouellette B, Sigman M, Roelfsema PR (2012) Decision making during the psychological refractory period. *Curr Biol* 22:1795–1799. [CrossRef Medline](#)