ORIGINAL ARTICLE

On quantifying multisensory interaction effects in reaction time and detection rate

Stefan Rach · Adele Diederich · Hans Colonius

Abstract Both mean reaction time (RT) and detection rate (DR) are important measures for assessing the amount of multisensory interaction occurring in crossmodal experiments, but they are often applied separately. Here we demonstrate that measuring multisensory performance using either RT or DR alone misses out on important information. We suggest an integration of RT and DR into a single measure of multisensory performance: the first index (MRE*) is based on an arithmetic combination of RT and DR, the second (MPE) is constructed from parameters derived from fitting a sequential sampling model to RT and DR data simultaneously. Our approach is illustrated by data from two audio–visual experiments. In the first, a redundant targets detection experiment using stimuli of different intensity, both measures yield similar pattern of results supporting the "principle of inverse effectiveness". The second experiment, introducing stimulus onset asynchrony and differing instructions (focused attention vs. redundant targets task) further supports the usefulness of both indices. Statistical properties of both measures are investigated via bootstrapping procedures.

Keywords Audio–visual interaction ·
Redundant target paradigm · Inverse effectiveness ·
Sequential sampling model · Inverse efficiency scores ·
Multisensory performance enhancement

Introduction

Reaction times (RT) and response frequencies are data commonly obtained in the behavioral sciences; mean RT and detection rate (DR), for instance, are accepted measures of performance in areas like sensation and perception. Often, these measures are studied in isolation, but sometimes it is more appropriate to consider them jointly, in particular when stimuli are very weak. Near to the detection threshold, both RT and DR are known to change with stimulus intensity (see, e.g., Luce, 1986; Woodworth & Schlosberg, 1956 for a summary), and it is important to know whether an improvement in one measure, say RT, indicates a change in overall performance rather than being compensated by a worsening in the other measure, DR. A general finding from studying such speed–accuracy trade-offs is that speeded responses tend to be less accurate; that is, a decrease in mean RT is often accompanied with a decreased DR (see, e.g., Luce, 1986 for a summary).

However, this phenomenon is not limited to stimuli near to threshold, as recently demonstrated by Arieh and Marks (2008) in a multisensory identification task. Multisensory tasks involve stimuli from two or more modalities, and a common finding, the so-called intersensory facilitation effect (IFE, Hershenson 1962), is that mean RT to crossmodal stimuli (e.g., light and tone) tends to be faster than to unimodal stimuli (e.g., light). A "true" IFE has been observed if the speeded reaction cannot be accounted for by other mechanism, such as statistical facilitation or response bias. Arieh and Marks (2008) evaluated the amount of multisensory interaction in a speeded

S. Rach (✉) · H. Colonius
Department of Psychology, University of Oldenburg,
26111 Oldenburg, Germany
e-mail: stefan.rach@uni-oldenburg.de

S. Rach · A. Diederich
Jacobs University Bremen, Bremen, Germany

identification of color with and without the presence of noise using speed–accuracy tradeoff functions (SATFs; Luce, 1986). Their results suggest that the facilitation of RT in audio–visual conditions is due to a change in the decision criterion induced by the auditory stimulus rather than to an increase of overall performance caused by the auditory stimulus.[1] Lowering the criterion means that the participant responds on the basis of less information, thereby speeding the response but reducing its accuracy.

These results demonstrate that studying multisensory interaction based on RT alone may lead to wrong conclusions. In this paper, we introduce two ways of quantifying overall performance by integrating RT and DR recorded in simple detection tasks. The first measure will make use of an arithmetic combination of RT and DR, namely, inverse efficiency scores (Townsend & Ashby, 1983). The second measure will utilize a sequential sampling model (see, e.g., Luce, 1986 for a summary) to quantify overall performance.

In the following, a brief overview of the quantification of multisensory effects is given, and both the descriptive and the model-based overall performance indices are briefly outlined. Then, a multisensory detection experiment is presented, and new overall performance indices are introduced in more detail before applying them to the results of the detection experiment and demonstrating the evaluation of differences in overall performance via non-parametric bootstrapping. After the presentation of the second experiment, which compares the influence of experimental instructions on overall performance in a detection task, we demonstrate how to adopt the new performance indices to experiments involving different instructions and stimulus onset asynchronies.

Quantification of multisensory interaction effects

Measures of multisensory speedup and detectability

Adaptive behavior in real world situations requires an organism to adequately combine cues from different sensory modalities. Especially in ambiguous or noisy situations (e.g., imagine a walker in a dark park), interpretation of vague information from one sensory modality can greatly be enhanced by further information delivered from other senses. The behavioral consequences of multisensory interaction have been the subject of a large body of research on both humans and animals. In the past hundred years, multisensory research has concentrated mainly on

two behavioral measures, RT and, although to a lesser degree, detectability indices.

RT to a visual stimulus tends to be faster when an auditory stimulus is presented in close temporal and spatial proximity (cf., IFE, Hershenson, 1962, or redundant targets effect, e.g., Miller, 1982). This effect proved robust in various behavioral replications when participants were instructed to respond to any stimulus they perceive (i.e., a redundant target paradigm, RTP, e.g., Diederich & Colonius, 1987; Gielen, Schmidt, & Van den Heuvel, 1983; Hershenson, 1962; Miller, 1982), as well as, when participant were instructed to respond only to stimuli of a pre-defined modality (target stimuli) and to ignore any other stimuli (i.e., a focused attention paradigm, FAP, e.g., Bernstein, Clark, & Edelstein, 1969b; Morrell, 1968; Rach & Diederich, 2006). The magnitude of IFE is modulated by the spatial and temporal alignment of the stimuli and it decreases with increasing temporal separation ("temporal rule", Bernstein, Clark, & Edelstein, 1969a; Bernstein, Rose, & Ashe,1970; Diederich & Colonius, 1987, 2004; Giray & Ulrich, 1993; Hershenson, 1962; Miller, 1986; Morrell, 1968), as well as with increasing spatial separation ("spatial rule", Amlôt, Walker, Driver, & Spence, 2003; Arndt & Colonius, 2003; Bernstein & Edelstein, 1971; Colonius & Diederich, 2004; Frens, Van Opstal, & Van der Willigen, 1995; Harrington & Peck, 1998; Walker, Deubel, Schneider, & Findlay, 1997). Moreover, the amount of IFE is larger when stimuli are less intense ("principle of inverse effectiveness", POIE, Corneil, Van Wanrooij, Munoz, & Van Opstal 2002; Diederich & Colonius, 2004; Rach & Diederich, 2006; see Holmes, 2009 for a critical view).

In addition to RT, multisensory interaction also shows up in detectability of stimuli. A task-irrelevant auditory stimulus can modulate visual perception (Bolognini, Frassinetti, Serino, & Làdavas, 2005; Frassinetti, Bolognini, & Làdavas, 2002), a task-irrelevant visual stimulus can enhance auditory perception (Lovelace, Stein, & Wallace, 2003), and task-irrelevant tactile stimuli can improve auditory detection (Gillmeister & Eimer, 2007). The amount of IFE, in terms of change in detectability, can also be modulated by the spatio-temporal alignment of stimuli (Bolognini et al., 2005; Frassinetti et al., 2002).

From a methodological point of view, all these findings rely on the ability to quantify an organism's performance in different experimental conditions. If we want to assess and compare the amount of multisensory interaction, it is necessary to compute measures that relate performance in unimodal conditions to performance in crossmodal ones. RT and DR measures can be computed under both conditions indicating the change in performance in crossmodal conditions relative to that in unimodal ones.

Note that, in the following section and in Experiment 1, we focus on the case of RTP, but all presented methods are

---

[1] Note that both a change of response bias and an change of overall performance can provide evidence for multisensory interaction if their occurrence can be attributed to the additional presentation of the auditory stimulus. In this study, however, our interest is directed towards the latter phenomenon.

adapted to the case of FAP in the second part of this manuscript (Experiment 2). For concreteness, we consider a bimodal (visual/auditory) simple RT experiment where participants have to react by a button press upon detecting a weak stimulus of either modality (redundant signals paradigm). Let RT_V, RT_A, and RT_{VA} denote mean RTs in the visual, auditory, and audio–visual condition; and let the detection rates (DRs) in the visual, auditory, and audio–visual condition be denoted by DR_V, DR_A, and DR_{VA}. To analyze the magnitude of IFE manifested in mean RT, Diederich and Colonius (2004) calculated multisensory response enhancement (MRE) by

$$MRE = \frac{\min(RT_V, RT_A) - RT_{VA}}{\min(RT_V, RT_A)} \times 100. \qquad (1)$$

MRE is a descriptive measure relating the fastest response on unimodal conditions (RT_V or RT_A) to that on bimodal audio–visual conditions (RT_{VA}).

When comparing detection rates (DR), the amount of IFE can be quantified analogously; we calculate multisensory detection enhancement (MDE) by relating the maximum unimodal performance to the bimodal performance:

$$MDE = \frac{DR_{VA} - \max(DR_V, DR_A)}{\max(DR_V, DR_A)} \times 100, \qquad (2)$$

where DR_{VA} indicates detection rate on audio–visual conditions and DR_V (DR_A) indicates detection rate on unimodal visual (auditory) conditions. Note that MRE and MDE share the following properties: (1) performance in the bimodal condition is compared to the best performance in the unimodal conditions; (2) increasing performance in the bimodal condition (e.g., shorter RT, or higher DR) leads to larger positive values; (3) decreasing performance in the bimodal condition (e.g., longer RT, or lower detection rate) leads to larger negative values; and (4) a value of zero indicates the absence of differences between the performance in the bimodal condition and the best performance in the unimodal conditions. Both MRE and MDE are unit free. The question is how to combine these two indices when both RT and DR are recorded simultaneously.

In the following, we consider two principally different ways of quantifying IFE from RT and DR. One measure is based on an arithmetical combination of RT and DR, whereas the other utilizes sequential sampling models to combine RT and DR and to provide model parameters from which a multisensory performance enhancement (MPE) measure is derived.

Inverse efficiency scores

In a simple RT experiment, a certain percentage of stimuli will be missed by the participant if the intensity level is weak enough. With increasing intensity, the percentage of missed stimuli will go down and, at the same time, mean RT will also decrease (e.g., Luce, 1986 for a review), resulting in an improved *overall performance*. However, a decrease in the percentage of misses could also be due to the participant being more careful, at the expense of taking more time to evaluate the stimuli (i.e., an increased mean RT) resulting in a speed–accuracy tradeoff (see, e.g., Luce, 1986 for a summary) without a change in overall performance. Furthermore, an increased carefulness of the participant could also result in an increased mean RT and a decreased detection rate due to an increased evidence threshold, resulting in a decreased overall performance. To differentiate between these possibilities, Townsend & Ashby (1983) introduced a measure combining accuracy and RT (in a choice task) by dividing mean RT by the percentage of correct responses. With this correction, which has later been termed inverse efficiency scores (IES), RTs are inflated in proportion to the error rate. Any difference in IESs between conditions is interpreted as a difference in overall performance; on the other hand, an IES invariant under differing mean RTs and choice frequencies is considered as evidence for a speed–accuracy tradeoff. In multisensory research, IES has been used to correct RT under low accuracy (e.g., Kitagawa & Spence, 2005; Röder, Kusmierek, Spence, & Schicke, 2007; Shore, Barnes, & Spence, 2006; Spence, McGlone, Kettenmann, & Kobal, 2001).

Adopting IES to a simple detection task, mean RT is divided by DR rather than by choice frequency:

$$RT^* = \frac{RT}{DR}. \qquad (3)$$

Substituting RT by RT* in Eq. 1 yields a measure reflecting relative multisensory enhancement with respect to IES:

$$MRE^* = \frac{\min(RT_V^*, RT_A^*) - RT_{VA}^*}{\min(RT_V^*, RT_A^*)} \times 100 \qquad (4)$$

where RT_V^*, RT_A^*, and RT_{VA}^* denote the transformed mean RTs in the visual, auditory, and audio–visual condition. Note that MRE* has the same properties as MRE stated above.

Multisensory performance enhancement

In a simple detection task with weak stimuli, participants are forced to make a decision in each trial. Given some evidence for the presence or absence of a stimulus, they have to decide whether or not to press the response button (cf., Luce, 1986, p. 140).

One approach to model such decision making are sequential sampling models (e.g., Diederich & Busemeyer, 2003; Luce, 1986; Ratcliff & Smith, 2004). Their basic

assumption is that the representation of a stimulus within the nervous system is noisy and that the organism accumulates samples of small quanta of information about this representation until a certain criterion of evidence is reached and a response is initiated. The time needed to reach the criterion is influenced by both the rate of the accumulation process (or drift rate) and the decision criterion (or boundary); see Fig. 1 for a schematic depiction of a single trial of the sequential sampling process. Formally, the rate of accumulation and the magnitude of the criterion are not separately identifiable as their influence on performance is compensatory. Nevertheless, their interpretation is different: the rate of accumulation is thought to be influenced by stimulus properties (e.g., brightness, duration, or salience), whereas the decision criterion is thought to be under the control of the participant (e.g., influenced by strategies). Postponing technical details for now (see "MPE calculated from drift rates" for a detailed mathematical description), the properties of sequential sampling models most relevant here are: (1) the rate of accumulation can be represented by a single model parameter, the drift rate δ, (2) more intense stimuli are represented by higher drift rates (and less intense stimuli are represented by lower drift rates). Note that sequential sampling models have been successfully used in the context of multisensory processes (e.g., Diederich, 1992, 1995; Jepma, Wagenmakers, Band, & Nieuwenhuis, 2008; Schwarz, 1994).

Returning to our audio–visual simple RT example experiment, let us assume that the drift rates for every unimodal and bimdodal condition have been estimated. We define multisensory performance enhancement (MPE) by

$$\text{MPE} = \frac{\delta_{\text{VA}} - \max(\delta_{\text{V}}, \delta_{\text{A}})}{\max(\delta_{\text{V}}, \delta_{\text{A}})} \times 100, \tag{5}$$
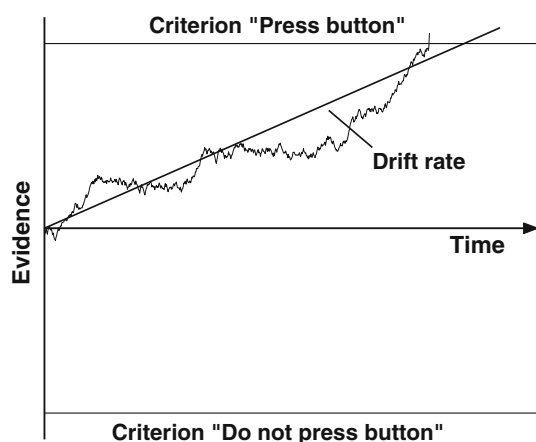

Fig. 1 Sampling path (trajectory) and drift rate (*solid line*) of a sequential sampling model. The sampling path (trajectory) represents sequential accumulation of evidence until a decision boundary (either "Press button" or "Do not press button") is reached. The average across multiple sampling paths is represented by the drift rate (*solid line*)

where δ_{V}, δ_{A}, and δ_{VA} indicate the drift rate for the visual, auditory, and bimodal condition, respectively. Like the previous indices before, MPE has the same properties as MRE noted above.

It is quite obvious that the presented multisensory performance measures, MRE* and MPE, differ in their theoretical foundation. Although being established empirically, the theoretical background of MRE* is rather weak: the calculation of IES is somewhat ad hoc without providing a specific justification for dividing RT by DR. In particular, the absolute magnitude of the "correction" applied to RT by this procedure will depend on the magnitude of both RT and DR. On the other hand, the sequential sampling models on which MPE is based are theoretically elaborate and there are analogies between the postulated mechanisms and the accumulation of neural activation found in organisms (cf., Diederich, 1995, 202). Sequential sampling models are frequently applied to RT and accuracy measures (e.g., Diederich, 1995; Diederich & Busemeyer, 2003, 2006).

To investigate the properties of the new measures, MRE* and MPE, we conducted two experiments with visual, auditory, and audio–visual stimuli of different intensities, followed by some cross-validation studies.

Experiment 1

A simple detection experiment with auditory and visual stimuli of several intensities near to the detection threshold was conducted. In concordance with the literature, we expect (1) RT to decrease with increasing stimulus intensity, (2) DR to increase with increasing stimulus intensity, and (3) performance in the bimodal condition to exceed performance in either unimodal condition.

Methods

Participants

Six students (ages 20–23 years, 3 female) served as participants and were paid for participation. All of them reported normal vision and hearing. Prior to their inclusion to this study, they were informed about the procedure and gave their informed consent. The experiment was conducted in accordance with the ethical standards described in the 1964 Declaration of Helsinki.

Apparatus

The study was conducted in a completely darkened and sound reflection attenuated room. Participants were seated in front of a black desk (180 × 130 × 75 cm), with their

head supported by a chin rest attached to the front edge of the desk.

Mounted on the desk, two red light-emitting diodes (LED, ø 5 mm) placed 20° to the left or right of a central fixation point marked by a third LED (fixation LED, red, ø 5 mm, 25 mA, 5.95 mCd) presented the visual stimuli. The three LEDs were arranged on a circle with a diameter of 35 cm centered on the base of the chin rest. Auditory stimuli were presented by two speakers (Canton Plus XS) placed horizontal to the participant's ear level at 20° to the left or right of the fixation LED. A PC multifunction card was used to control LEDs and speakers.

Responses were recorded using a button operated by the large toe. The toe rested on the button and was to be lifted in order to activate the button. This foot device was used because this experiment was part of a larger study that also employed tactile stimuli applied to the palms.

Stimuli

Visual stimuli were red lights of 500 ms duration. Intensity of visual stimuli was varied in 8 steps between 0.0045 and 0.0106 mcd, henceforth indicated by $V_1, \ldots, V_8$. For a complete list of intensities utilized in this experiment see Table 1. Auditory stimuli were bursts of white noise of 500 ms length. Intensity of auditory stimuli was varied in 8 steps between 14.4 and 24.7 dBA, henceforth indicated by $A_1, \ldots, A_8$. On bimodal trials, visual and auditory stimuli of corresponding intensity levels were paired (i.e., $\{V_1, A_1\}, \ldots, \{V_8, A_8\}$). Whenever it is referred to stimulus intensities without specifying the modality, the labels $I_1, \ldots, I_8$ will be used with I_1 representing the weakest and I_8 the strongest intensity.

Procedure

Participants were instructed to respond to every stimulus they detect regardless of its modality by lifting their toe as quickly as possible (redundant target paradigm). To keep participants from responding to the offset of the fixation LED, it was emphasized that the task aims at determining interindividual differences in perception, rather than perfect detection performance. In particular, it was underlined that some of the stimuli are very weak and unlikely to be detected at all.

The beginning of each trial was indicated by the onset of the fixation LED, which was turned off after a variable fixation time (800–1.500 ms). Stimulus presentation started simultaneously to the offset of the fixation LED. On unimodal trials either a visual or an auditory stimulus was presented; on bimodal trials both a visual and an auditory stimulus were presented simultaneously (no stimulus onset asynchrony). Both visual and auditory stimuli were presented either in the right or left hemifield. On audio–visual trials, both the visual and the auditory stimulus were always presented in the same hemifield. See Fig. 2 for a schematic depiction of the procedure.

A recording session of 45 min included two blocks separated by a 5-min break. Recording 32 trials for each of 24 conditions with each of 6 participants resulted in a total of 4.608 trials.

Data recording and preprocessing

A PC connected to the EyeLink was used for data storage and data preprocessing. Trials with RTs faster than 80 ms were classified as anticipation errors (0.9%) and considered as trial without response, as well as ones with RTs longer than 1,400 ms (misses, 0.6%). RTs from the right and the left hemifield showed no systematic difference and were therefore combined across hemifields of stimulus presentation (left or right).
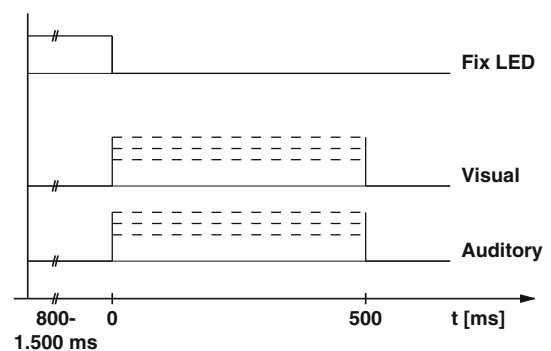
Table 1 Experiment 1: intensities utilized given as luminance (mcd) of visual and loudness (dBA) of auditory stimuli

Visual		Auditory	
Index	mcd	Index	dBA
V_1	0.0045	A_1	14.3
V_2	0.0052	A_2	16.3
V_3	0.0059	A_3	17.3
V_4	0.0065	A_4	18.3
V_5	0.0069	A_5	19.2
V_6	0.0086	A_6	21.1
V_7	0.0096	A_7	23.0
V_8	0.0106	A_8	24.7

Fig. 2 Experiment 1: time course of a trial. Onset of the fixation LED defined the start of a trial. Simultaneously to the offset of the fixation LED, either the visual, the auditory, or both were presented. Stimuli from both sensory modalities were presented at 8 different intensity levels

Results

The manipulation of stimulus intensity showed to have the hypothesized effect on both RTs and DRs. With increasing intensity of stimuli, mean RT became faster for the visual, auditory, and the bimodal condition. The difference between the slowest and the fastest mean RT was about 110 ms for the visual condition, 120 ms for the auditory, and 130 ms for the bimodal condition. At the same time, DR increased with increasing stimulus intensity for the visual, auditory, and the bimodal condition. In the visual and the auditory condition, the lowest and the highest DR differed by 0.4, whereas this difference was about 0.2 for the bimodal condition. Mean RT and DR observed in this experiment are summarized in Table 2.

To evaluate the effects of multisensory stimulation, the relative amount of multisensory facilitation was quantified by calculating MRE from RTs and MDE from DRs. Mean RT and DR, both with standard errors (black vertical lines) as a function of stimulus intensity and stimulus modality, are given in Fig. 3 (panels a, c). MRE and MDE as function of stimulus intensity are given in Fig. 3 (panels b, d).

For intensity levels I_4–I_7, MRE was greater than zero, indicating that responses on bimodal presentation outperformed the best unimodal response in terms of speed. For the remaining intensity levels I_1 to I_3 and I_8 the bimodal RT did not undercut the fastest unimodal RT, and therefore MRE was about zero or even slightly negative, i.e., bimodal stimulation did not speed up reactions compared to the fastest unimodal condition. MDEs exhibited a pattern almost opposite. Positive values of MDE were found for intensity levels I_1–I_5, indicating enhanced detectability on bimodal conditions. MDE was about zero for intensity levels I_6–I_8, which, however, is no surprise since DR was already almost close to 1 for unimodal conditions, that is, a ceiling effect was observed here.
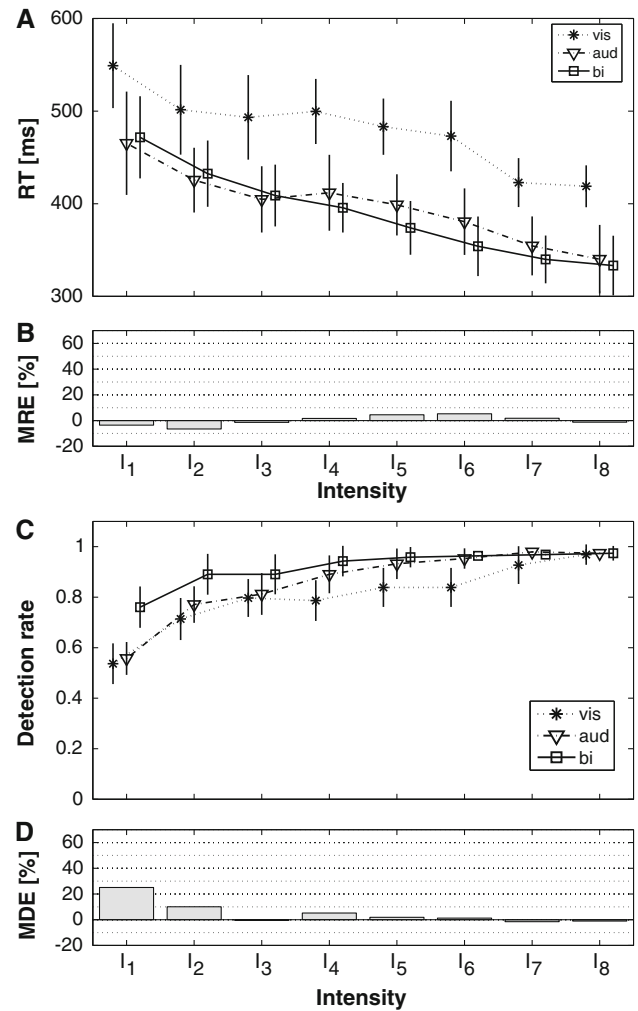

Fig. 3 Experiment 1: performance in a detection experiment quantified from reaction time and detection rate. **a** Mean foot reaction time (ms) with standard errors (*black vertical lines*) as a function of stimulus intensity and modality. **b** Multisensory response enhancement (%) as a function of stimulus intensity. **c** Detection rate with standard errors (*black vertical lines*) as a function of stimulus intensity. **d** Multisensory detection enhancement (%) as a function of stimulus intensity. Stimulus intensity increases from I_1 to I_8

To summarize, quantifying the magnitude of IFE with respect to either only RT (i.e., in terms of MRE) or only DR (i.e., in terms of MDE) led to contrary results: MRE indicates multisensory facilitation in conditions where MDE indicated none and vice versa. To integrate these opposite findings inferred from RT and DR, we evaluated overall performance in terms of MRE* and MPE.

MRE* calculated from inverse efficiency scores

RT* and MRE* calculated from RT and DR according to Eqs. 3 and 4 are given in Fig. 4.

For all three modalities RT* decreased with increasing stimulus intensity, indicating that the overall performance

Table 2 Experiment 1: mean reaction time (with standard error) and detection rate as a function of stimulus intensity and sensory modality

Intensity	Response time in ms (SE)			Detection rate		
	V	A	VA	V	A	VA
I_1	549 (45.8)	465 (55.8)	472 (44.3)	0.54	0.56	0.76
I_2	501 (48.5)	425 (34.9)	432 (35.8)	0.71	0.77	0.89
I_3	493 (45.7)	405 (35.7)	409 (33.3)	0.80	0.81	0.89
I_4	500 (35.1)	412 (40.9)	396 (26.8)	0.79	0.89	0.94
I_5	483 (30.4)	399 (33.0)	374 (29.0)	0.84	0.93	0.96
I_6	473 (38.0)	381 (35.9)	354 (32.1)	0.84	0.95	0.96
I_7	423 (26.4)	354 (31.8)	340 (25.8)	0.93	0.98	0.97
I_8	419 (22.6)	340 (37.0)	333 (32.2)	0.97	0.97	0.97

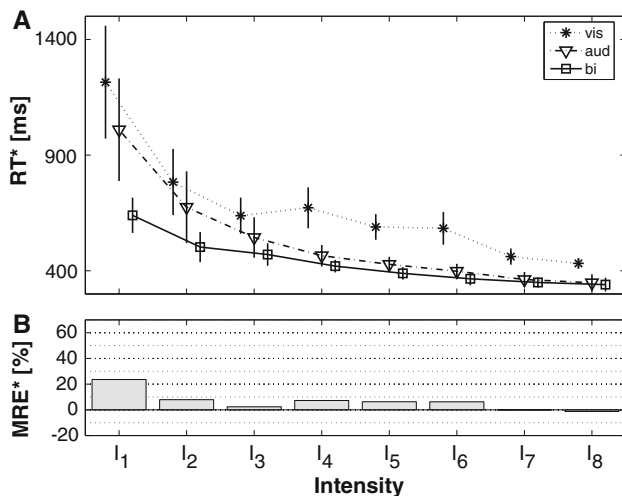Stimulus intensity is ordered ascending from I_1 to I_8

Fig. 4 Experiment 1: overall performance in a detection experiment quantified from inverse efficiency scores. **a** Mean inverse efficiency scores, RT*, as a function of stimulus intensity and modality. **b** Relative response enhancement (MRE*) calculated from inverse efficiency scores, RT*, as a function of stimulus intensity. Stimulus intensity increases from I_1 to I_8

increased with stimulus intensity. Compared to RT, this decrease was much steeper for RT* (note the differently scaled y-axes in panel a of Figs. 3, 4). For intensity conditions I_1– I_6, bimodal RT* was lower than either of the unimodal ones; for I_7 and I_8 bimodal and auditory RT* were about the same. The relative enhancement calculated from mean RT*, MRE* is larger than zero for all conditions, indicating multisensory facilitation across all intensity conditions. Moreover, with increasing stimulus intensity MRE* decreased (see Fig. 4, panel b). For the lowest intensity level I_1, MRE* was 26.6%, for the highest intensity level I_8 it was 1.7%. Note that this pattern is in accordance with the POIE.

MPE calculated from drift rates

To calculate MPE, we fitted a sequential sampling model with two absorbing boundaries to the data. Such models are often applied to two-alternative choice tasks to account for RTs and choice probabilities (Diederich, 1997, 2008; Diederich & Busemeyer, 2003, 2006; Ratcliff & Smith, 2004). However, it is important to note one difference between the typical two-alternatives sequential sampling model and the version utilized here. From an alternative choice experiment with two alternatives, say A and B, three independent measures can be recorded: the choice probabilities for both alternatives (p_A and $p_B = 1 - p_A$), as well as the corresponding RTs (RT_A and RT_B). In a detection task, only two independent measures can be observed: the detection probabilities ($p_{detected}$ and $p_{not-detected} = 1 - p_{detected}$, and mean RT for trials where participants responded because a stimulus was detected. The RT on trials where participants decided not to respond is not observable.

Different stochastic processes can be used to define a sequential sampling model; the Wiener process, $X(t)$ with drift and bias is considered here for simplicity of demonstration (Diederich & Busemeyer, 2003, 2006; Ratcliff & Smith, 2004). The Wiener process is determined by two parameters: the drift rate δ and the decision criterion θ. The decision criterion θ determines how much activation has to be accumulated in favor of one alternative until an absorbing boundary is reached. We assume θ to be the same for both alternatives. As soon as $X(t) \geq \theta$ or $X(t) \leq -\theta$ a response is initiated with t being called the first passage time (FPT). For constant drift rates, low values of θ result in faster mean FPTs, while large values result slower mean FPTs.

The drift parameter δ represents the effectiveness of a stimulus. For a given boundary θ, a higher value of δ leads to a higher choice probability and a faster mean FPT for one alternative, say A, and at the same time, leads to a lower choice probability and a slower first passage time for the opposite alternative, say B. A decreased magnitude of δ has an opposite effect: choice probability for alternative A decreases and its mean FPT increases, while choice probability for alternative B increases and the respective mean FPT decreases. For detection tasks, A would be interpreted as decision to respond, while B would be interpreted as decision not to respond (cf. Gomez, Ratcliff, & Perea, 2007). Consequently, only RT for alternative A would be examined, ignoring RT for alternative B, because it is unobservable in this case.

Furthermore, to determine the initial state of evidence, $X(0) = \beta \times \theta$, the parameter β was estimated. The initial state of evidence accounts for differences in the amount of activation accumulated necessary between the two alternatives: $X(0) = 0$ indicates an process not favoring any of the alternatives; $X(0) \geq 0$ indicating evidence in favor of one alternative, and $X(0) \leq 0$ indicates evidence in favor of the opposite. For the application presented in this paper, the main purpose of introducing β is to prevent negative drift rates which occur when probabilities are lower than 0.5.

Finally, a residual time T_r was estimated. It can be interpreted as base time, that is, the time taken by those non-decisional cognitive and motor processes that are not influenced by experimental manipulations under interest, and therefore remains constant across all experimental conditions (i.e., the component of the measured RTs that is independently and identically distributed across trials and across experimental conditions, cf., Townsend & Honey, 2007, pp. 259–260).[2]

---

[2] We are aware of the ongoing debate on whether coactivation effects might be located in the motor component (eg., Diederich & Colonius, 1987; Giray & Ulrich, 1993) or not (eg., Miller, Ulrich & Lamarre, 2001; Mordkoff, Miller, & Roch, 1996). However, the purpose of this paper is not to take a stand in this debate, but rather to provide methods to combine different response measures into one single index of overall performance. Hence, base time as defined here does not include components that are influenced by coactivation effects.
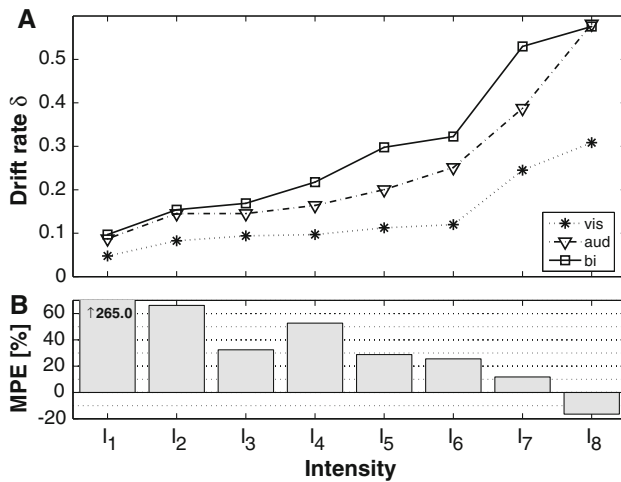
Fig. 5 Experiment 1: overall performance in a detection task quantified from parameters of a sequential sampling model. **a** Drift rates δ as a function of stimulus intensity and modality. **b** Multisensory performance enhancement (MPE) as a function of stimulus intensity. Stimulus intensity is ordered ascending from I_1 to I_8

For each participant, drift rates were estimated for every experimental condition, i.e., 8 visual, 8 auditory, and 8 bimodal conditions, whereas θ, β, and T_r were estimated once for the whole data set by minimizing Pearson's χ^2 statistic

$$\chi^2 = \sum_i \left(\left(\frac{\overline{RT}_i - \widehat{RT}_i}{\sigma_{\overline{RT}_i}} \right)^2 + \left(\frac{\overline{DR}_i - \widehat{DR}_i}{\sigma_{\overline{DR}_i}} \right)^2 \right) \qquad (6)$$

where $\overline{RT}_i$ and $\widehat{RT}_i$ indicate observed and predicted RT; $\overline{DR}_i$ and $\widehat{DR}_i$ indicate observed and predicted DR; and $\sigma_{\overline{RT}_i}$ and $\sigma_{\overline{DR}_i}$ indicate standard errors of RT and DR.[3] Estimates for drift rates averaged across participants are given in Fig. 5, panel a. The effect of manipulating intensity displayed clearly in the drift rates, as they increased with increasing stimulus intensity for the visual, auditory and bimodal condition. Since higher drift rates correspond to shorter RTs and higher DRs, this means that overall performance increases with increasing stimulus intensity. Moreover, there was a clear effect of bimodal stimulation, since the drift rates for bimodal stimulation exceeded that for the best unimodal for all intensity conditions. The remaining parameters (averaged across participants) were

---

[3] A statistical test for the goodness of fit is given by $\chi^2(21) = 29.6$, $p = 0.10$ (27 parameters were fitted to 48 data points). Excellent fits were indicated for 5 out of 6 participants (observed χ^2 values of 22.2, 11.9, 21.0, 10.9, and 18.6) and a very poor fit for the sixth participant (observed χ^2 values of 54.6) Note, however, that we did not intend to test the diffusion model with this fit. Instead we utilized the drift rates in a descriptive way to quantify overall performance as indicated by RT and DR.

$\theta = 23$, $\beta = -0.23$ and $T_r = 249$ ms. MPE calculated according to Eq. 5 is displayed in Fig. 5, panel b.

Note that MPE for the lowest intensity level I_1, where MPE indicated a response enhancement of 265.0 %, the bar did not fit in the plot, which was kept at the same scale as those for MRE, MDE, and MRE* to allow for better comparison. MPE clearly decreased with increasing intensity levels, thus, evidence for the POIE was observed here.

MRE* and MPE exhibited very similar patterns. Both indices picked up characteristics from MRE and MDE. Like MDE, both measures indicated the largest enhancement for condition I_1 and decreasing enhancement from I_1 to I_3, and like MRE, both indices indicated a comparable amount of enhancement for conditions I_4–I_6. Nevertheless, this pattern is much more pronounced in the case of MPE. Since for both measures it was not clear whether the observed differences between intensity conditions are statistically meaningful, both were subjected to bootstrapping procedures.

Evaluation of differences in overall performance utilizing bootstrapping procedures

To allow for the examination of variability and the calculation of confidence intervals, the obtained performance indices were subjected to a non-parametric bootstrap procedure. The bootstrap is a Monte Carlo technique that generates simulated data sets by resampling from empirical data observed in the original experiment (Efron & Tibshirani, 1986; Wichmann & Hill, 2001). The non-parametric bootstrap samples simulated data sets by drawing with replacement from the original data and provides a distribution of simulated RTs and a distribution of simulated DRs. These distributions can be used to calculate confidence intervals. We will report 68% confidence intervals, CI_{68}, calculated by the bootstrap percentile method, because they are comparable to common standard error bars. CI_{68} spans from the 16th to the 84th percentile of the bootstrap distribution, which approximately compares to the original estimate $\pm$ 1 standard deviation of a Gaussian (Wichmann & Hill, 2001).

We generated 1,000 non-parametric bootstrap samples from the original data set (see Appendix for details); for each of them we calculated MRE* from RT*. From the resulting distributions of MRE* we calculated CI_{68}. To evaluate the variation of MPE, a sequential sampling model (Wiener processes) was fitted to each of the 1,000 bootstrap samples and 68% confidence intervals, CI_{68}, were calculated from the resulting parameter distributions.

Bootstrapped median MPE (asterisks), and median MRE* (open triangles) with CI_{68} (black vertical lines) are given in Fig. 6. CI_{68} are also listed in Table 3.
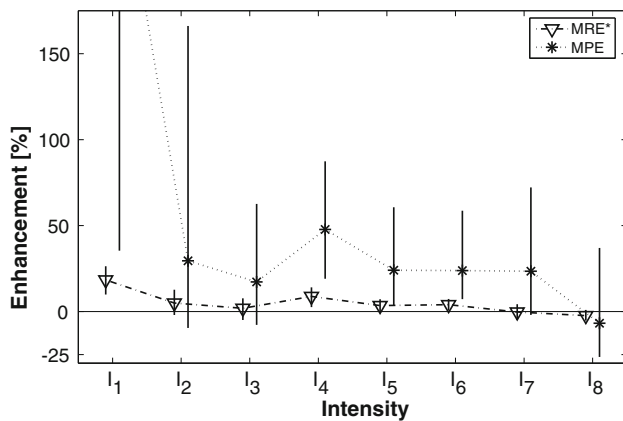
Fig. 6 Experiment 1: magnitude and variation of multisensory indices as a function of stimulus intensity. *Open triangles* display multisensory response enhancement calculated from inverse efficiency scores (MRE*); *asterisks* display multisensory performance enhancement (MPE). *Black vertical lines* indicate confidence intervals, CI_{68}, obtained in a non-parametric bootstrap procedures. Stimulus intensity is ordered ascending from I_1 to I_8. Median MPE for the lowest intensity level, I_1, was 265.4 %, with the corresponding CI_{68} spanning from 35.4 to 910.1

Table 3 Experiment 1: confidence intervals, CI_{68}, calculated by the bootstrap percentile method as a function of stimulus intensity

Intensity	Percentiles			
	MRE*		MPE	
	16th	84th	16th	84th
I_1	9.90	26.37	35.42	910.09
I_2	−1.96	12.69	−9.54	166.12
I_3	−4.90	7.72	−7.74	62.60
I_4	2.59	14.07	19.11	87.37
I_5	−0.36	7.19	3.38	60.69
I_6	0.63	7.33	7.26	58.68
I_7	−4.75	4.28	−1.94	72.22
I_8	−5.26	0.96	−26.39	37.00

For MRE*, and MPE the 16th and the 84th percentile are given as lower and upper limit of the confidence interval. Stimulus intensity is ordered ascending from I_1 to I_8

Obviously the variation was considerably larger for MPE, than for to MRE*. Nevertheless, both indices displayed similar patterns. MRE* was significantly larger than zero for intensity conditions I_1, I_4, and I_6, indicating an enhanced performance due to bimodal stimulation for the respective stimulus intensities. For MPE intensity conditions I_1, I_4, I_5, and I_6 were significantly larger than zero, i.e., the conclusions suggested by both indices differed only in one out of eight conditions.

MRE* for intensity condition I_1 was significantly larger than MRE* for condition I_3 and for conditions I_5 to I_8, as the confidence intervals did not overlap for those

conditions. No significant differences between intensity conditions were found for MPE.

Discussion

An experiment was conducted to investigate the influence of intensity levels close to detection threshold on multisensory interaction and to test two proposed indices of overall performance, MRE* and MPE. Visual, auditory, and audio–visual stimuli of eight different intensities were presented to participants, and MRE was calculated from RT, as well as MDE from DR. Furthermore, MRE* was calculated from IES, and MPE from drift rates estimated from the experimental data. MRE* and MPE were subjected to bootstrapping procedures to evaluate the variability of both indices.

From earlier reports (e.g., Diederich & Colonius, 1987, 2004; Gielen et al., 1983; Hershenson, 1962; Miller, 1982), one would expect bimodal RTs to be faster than unimodal RTs. However, as can be seen from the MRE bar plots (see Fig. 3, panel b), this was only the case for intensity levels I_4–I_7. For intensity levels I_1–I_3 and I_8, the fastest unimodal RT was equal to or even faster than the bimodal one. For the highest intensity level I_8, the absence of multisensory enhancement might be due to a performance limit: the "irreducible minimum" of RT puts a lower boundary on RT facilitation (Woodworth & Schlosberg, 1956). For the lower intensity levels, an explanation is not as easy since the POIE (Meredith & Stein, 1986; Perrault, Vaughan, Stein, & Wallace, 2005; Stein & Meredith, 1993) would predict multisensory response enhancement to be most prominent when stimuli are weak. However, if we broaden our analysis from RT and also take DRs into account, it becomes obvious from the MDE bar plots (see Fig. 3, panel d) that, for intensity levels I_1–I_5, multisensory enhancement actually did manifest in DR. The fact that no enhancement was observed for DR on intensity levels I_6–I_8 could be explained by ceiling effects: DR on the best unimodal conditions was already nearly perfect, therefore, further improvement through bimodal stimulation was not possible. Multisensory enhancement was evident in DR when it was absent in RT and vice versa; thus, isolated inspection of either MRE or MDE ould have missed out essential information.

Importantly, the multisensory performance indices MRE* and MPE both indicate enhancement for conditions where either MRE or MDE, or both, did as well; i.e., they integrated the information revealed in RT and DR. Despite differences in absolute magnitude, the patterns exhibited by MRE* and MPE were very similar, i.e., the ordinal relations across index values of different intensity levels are almost invariant. Nevertheless, the differences in absolute magnitude were substantial, especially for the lowest intensity condition and, based on the current data set, it

cannot be decided whether one of the indices displays a numerical value more appropriate than the other, or whether both are somewhat out of range. This issue may not be settled with behavioral experiments alone but, rather, by simulation studies because the latter allow to manipulate the magnitude of enhancement directly.

To evaluate the variability of MRE* and MPE, we performed bootstrapping studies and, interestingly, the results where similar for both indices. Although the width of the resulting confidence intervals differed substantially (small for MRE* vs. large for MPE), both measures indicated significantly enhanced performance due to bimodal stimulation for almost the same conditions (concordance in seven out of eight conditions). Nevertheless, differences between MRE* and MPE were also observed. MRE* indicated significantly more enhancement for the lowest intensity condition, compared to five out of the seven remaining intensity conditions. This finding presents strong evidence in favor of the POIE (cf., Corneil et al., 2002; Diederich & Colonius, 2004; Rach & Diederich, 2006). Although a similar trend was present, no significant differences between intensity conditions were observed for MPE, probably due to the very large bootstrap confidence intervals. Unfortunately, it is not clear whether the magnitude of these confidence intervals represents true variability of the data, or rather a lack of reliability in the parameter estimation of the bootstrap samples. Nevertheless it is important to note that the model-based approach underlying MPE also allows for a parametric bootstrap (cf., Wichmann & Hill, 2001) which often results in smaller confidence intervals.

To summarize the characteristics of both MRE* and MPE: both measures pick up and integrate characteristics exhibited by RT and DR resulting in very similar patterns, which, however, may differ in magnitude. Both measures lead to converging evidence in line with earlier research, although only MRE* allowed for statistical conclusions derived from bootstrap confidence intervals.

One of the reviewers raised the question whether the offset of the fixation LED simultaneously to the onset of the stimuli might have led the participants to respond only to the offset of the fixation LED and to ignore the visual and auditory stimuli (i.e., performing a focused attention task with the offset of the fixation LED acting as target and the visual and auditory stimuli acting as non-targets). Although we do agree that a detection task without catch trials may lead participants to respond to the offset of the fixation LED, we judge this possibility as very unlikely for this particular set of data, for the following two reasons: (1) the intensity of the fixation LED exceeded the intensity of the strongest visual stimulus by orders of magnitude; since the fixation LED was clearly detectable in every trial, one would not expect that its offset results in a pattern of detection rates as observed in Experiment 1 (see Fig. 3, panel c). In particular,

one would not expect DRs much lower than 1 as observed for the lower intensity conditions; (2) participants were informed that some of the stimuli would be very weak and unlikely to be detected at all. Furthermore, it was emphasized that the task aims at determining interindividual differences in perception, rather than perfect detection performance.

Still it cannot be ruled out that participants used the offset of the fixation LED as a warning cue since it reliably announced the presentation of a stimulus. Such a cue would reduce or even diminish the effect of the random durations of the fixation LED, resulting in faster RTs (Luce, 1986) and perhaps higher DRs. Since faster RTs typically result in lower response time gains on bimodal stimulation, the observed results would be more conservative compared to results with catch trials (cf., Gielen et al., 1983).

Nevertheless, this issue poses an interesting problem addressed in a subsequent experiment: Given that different task instructions lead to different patterns of results under identical stimulus conditions, in which conditions would these differences show up?

Experiment 2

Two experimental paradigms have been proposed to investigate multisensory interaction. In the FAP task participants are instructed to respond only to stimuli of one particular modality (target stimuli) and to ignore all stimuli from other modalities (non-targets). In the RTP task participants are instructed to respond to any stimulus regardless of its modality. In previous studies comparing RTs recorded from FAP and RTP tasks, it was reported that both unimodal and bimodal responses are faster in RTP (Giray & Ulrich, 1993; Morrell, 1968). Both studies, however, compared data from separate experiments recorded with different participants. Here both experimental tasks are compared in one and the same experiment (i.e., identical stimuli and participants) using a blocked design, allowing to attribute any difference in results to differences between the tasks. Note, in particular, that the influence of stimulus intensity on DR and RT in bimodal conditions is different for RTP and FAP.

For instance, let us assume that both stimuli of a bimodal stimulus complex are perfectly detectable on unimodal presentation, and, hence, on bimodal presentation. Decreasing the intensity of one of the stimuli while leaving the intensity of the other stimulus constant should have different implications in RTP and FAP tasks. For RTP, decreasing the intensity of either of the stimuli should not worsen the detectability of the bimodal stimulus complex, because the unaltered stimulus alone should still be intense enough to be always detected. For FAP, however, it should make a difference whether the intensity of the target

or the non-target is decreased. Decreasing the intensity of the non-target should not decrease the detectability of the bimodal stimulus, because the unaltered target stimulus would still be intense enough to be always detected. On the other hand, gradually decreasing the intensity of the target should lead to a decreased detectability of the bimodal stimulus sooner or later, because even if participants still perfectly detect the non-target, they are only allowed to respond if they detect the target too, which should become rarer with decreasing target intensity.

For mean RT, the situation is similar. In an RTP task, decreasing the intensity of either of the stimuli while leaving the other unaltered should increase bimodal RT, but only up to a certain limit where it equals the RT to the unaltered stimulus, and any further decrease of stimulus intensity should not influence bimodal RT anymore. For FAP, again it should make a difference whether the intensity of the target or the non-target is manipulated. Decreasing the intensity of the non-target should increase bimodal RT up to the point where it equals the RT to the target stimulus and further decreasing the intensity of the non-target should have no effect on bimodal RT. On the contrary, gradually reducing the target's intensity should increase bimodal RT up to the point where no responses occur at all because the target stimulus is too weak to be detected.

In general, decreasing the intensity of one stimulus while leaving the intensity of the other stimulus constant increases the relative effectiveness of the latter. In the FAP, pairing a high-intensity non-target with a low-intensity target maximizes the possible influence of the non-target, because a stronger non-target can provide more response activation in order to elicit a response, whereas the opposite situation (i.e., a high-intensity target presented with a low-intensity non-target) would diminish the influence of the non-target because the high-intensity target already provides enough response activation itself. In the RTP, the influence of stimulus intensity is symmetrical, because no matter which stimulus' intensity is decreased, the other high-intensity stimulus always provides enough response activation itself, reducing the influence of the low-intensity stimulus. Hence, the difference in IFE between the conditions with a high-intensity and a low-intensity stimulus should be larger in the FAP than in the RTP.

For all intensity conditions of Experiment 1, the RTs to auditory stimuli were about 80 ms faster than the RTs to visual stimuli. Although this difference is well below the 200 ms temporal window of integration that has been reported previously (e.g., Eimer, 2001; Meredith, Nemitz, & Stein, 1987), a smaller difference in unimodal RTs might have led to larger integration effects, because it is assumed that having equal unimodal RTs maximizes the probability of multisensory interaction (physiological simultaneity, cf.,

Hershenson, 1962). Nickerson (1973) argued that differences in RT between sensory modalities might be caused by differences in "internal arrival times", that is, the time "that is required for the nervous system to do whatever it does before a response is evoked" (p. 501), assuming that response execution does not differ between sensory modalities. He concluded that stimuli that are matched for internal arrival times have a greater chance to interact than those that are not. Thus, experimenters often try either to match stimuli in terms of RT or, if that is not feasible, delay the faster of the stimuli by a stimulus onset asynchrony (SOA) in bimodal conditions to increase the probability of multisensory interaction and the expected magnitude of IFE (Hershenson, 1962; Hilgard, 1933; Miller, 1986; Morrell, 1968). The latter approach was utilized in this experiment to study how the different measures of overall performance can be adopted to experimental designs with SOA.

Regardless of the experimental task, we expect (1) bimodal RT to be faster than unimodal RTs for both RTP and FAP; (2) IFE to be larger for pairs of low-intensity stimuli compared to pairs of high-intensity stimuli (i.e., the POIE). Furthermore, we expect (3) RTs to be faster in the RTP task than in the FAP task, and (4) the difference in IFE between the condition where one stimulus is of high intensity and the other is of low intensity to be larger in the FAP than in the RTP.

Methods and apparatus

Participants

Four students (ages 20–23 years) served as paid voluntary participants. All of them had participated in Experiment 1. Prior to their inclusion to this study, all participants were informed about the procedure and gave their informed consent. The experiment was conducted in accordance with the ethical standards described in the 1964 Declaration of Helsinki.

Apparatus

The study was conducted in exactly the same setup as Experiment 1.

Stimuli

Stimulus intensities and SOA were determined from the data recorded in Experiment 1 for each participant individually. For each modality, two intensities where selected according to three criteria: (1) the higher intensity (in the following represented by uppercase letters: V, A) was detected more than 90% of the times; (2) the lower

intensity (in the following indicated by lowercase letters v, a) was detected less than 80% of the times; (3) the difference in RT between low-intensity and high-intensity stimuli (i.e., RT(v) − RT(V), respectively RT(a) − RT(A) for the auditory modality) was between 70 and 80 ms (see Table 4 for the resulting intensities). Stimulus duration was 500 ms for both visual and auditory stimuli.

For optimal interaction between the modalities (e.g., Hershenson, 1962), an SOAs was determined for each participant individually by choosing

$$SOA = RT(V) − RT(A)$$

(see Table 4 for the resulting SOAs). Note that, due to the criteria for selecting stimulus intensities, the difference between the low intensity stimuli, RT(v) − RT(a), was about the same, i.e., presenting stimuli with these SOA should result in physiological simultaneity in both conditions VA and va.

Procedure

Two different experimental tasks were utilized in this experiment. Participants were either instructed to respond to any stimulus they perceive (RTP) or to respond only to the visual stimulus (FAP). The tasks were presented balanced over experimental blocks, ensuring that each session consisted of blocks of either task in randomized order. Prior to each block, participants were informed about the particular task of this block by written instructions, which were additionally read out loud to them once more before the recording started. Within a block visual, auditory, and audio–visual stimuli were presented. Visual and auditory stimuli were either of low or high intensity, resulting in four unimodal conditions (v, V, a, A) and four audio–visual conditions (va, vA, Va, VA). Note that unimodal auditory conditions served as catch trials in the FAP condition. Visual and auditory stimuli were presented either in the right or left hemifield. On audio–visual trials, both the visual and the auditory stimulus were always

Table 4 Experiment 2: participant specific luminance (mcd) of visual stimuli, loudness (dBA) of auditory stimuli and stimulus onset asynchronies (ms) utilized in this experiment

Participant	Visual (mcd)		Auditory (dBA)		SOA [ms]
	v	V	a	A	
as	0.0045	0.0076	18.3	21.1	−10
da	0.0087	0.0112	14.3	24.8	90
ig	0.0076	0.0106	15.3	20.2	135
sr	0.0059	0.0106	15.3	19.2	70
Mean	0.0067	0.0100	15.8	21.3	71

Lower-case letters indicate low-intensity stimuli; upper-case letters indicate high-intensity stimuli

presented in the same hemifield. The beginning of each trial was indicated by the onset of the fixation LED, which was turned off after a variable fixation time (800–1,500 ms). On unimodal trials, stimulus presentation started simultaneously to the offset of the fixation LED. On audio–visual trials, presentation of visual stimuli always started simultaneously to the offset of the fixation LED, whereas presentation of the auditory stimuli was shifted by an SOA[4]. Except for participant AS, the presentation of the auditory stimuli started simultaneously to or after the onset of the visual stimuli.

Prior to the main study, each participant completed 2 h of training. In a recording session of 1 h, a participant completed two blocks separated by a break of 10 min. Recording 80 trials for each of 16 conditions with each of 4 participants resulted in a total of 5.120 trials.

Data recording and preprocessing

A PC connected to the EyeLink was used for data storage and data preprocessing. Trials with RTs faster than 80 ms were classified as anticipation errors (0.1%) and considered as trial without response, as well as ones with RTs longer than 1,000 ms (misses, 0.9%). RTs from the right and the left hemifield showed no systematic difference and were therefore combined across hemifields of stimulus presentation (left or right).

Results

For both experimental paradigms, unimodal stimuli of high intensity were detected more often and elicited faster RTs than stimuli of low intensity. The presentation of audio–visual stimuli always led to higher DRs and faster RT, compared to presenting either of the unimodal stimuli alone. For both paradigms, the fastest RTs were recorded in audio–visual conditions when both stimuli were of high intensity (VA). Overall, RTs were slower in the FAP task than in the RTP task, which is in line with previous reports (Giray & Ulrich, 1993; Morrell, 1968). The influence of stimulus intensity on DR in audio–visual conditions was different for FAP and RTP. For FAP tasks, the highest DRs were observed when the visual stimulus was of high intensity regardless of the intensity of the auditory stimulus (Va and VA). For RTP tasks, the highest DRs were recorded when either of the stimuli was of high intensity (Va, vA, VA). Since the presented stimuli and the presentation scheme were the same for both tasks, this difference strongly suggests that participants correctly performed two different experimental tasks, rather than, for

---

[4] SOAs where determined individually for each participant in a pilot study. See section "Stimuli" for details.

instance, simply responded to the offset of the fixation LED.

To evaluate the effects of multisensory stimulation, the relative amount of multisensory enhancement was quantified by calculating MRE from RTs and MDE from DRs. For the results of the FAP tasks, only responses to unimodal visual stimuli were available since participants were instructed not to respond to auditory stimuli. Therefore, the minimum (respectively the maximum) of the unimodal responses was replaced by the visual response in Eqs. 1 and 2 for the evaluation of the FAP condition.

Since the presentation of auditory stimuli was shifted by an SOA, RT_A was corrected accordingly for the calculation of MRE for the RTP condition

$$MRE = \frac{\min(RT_V, RT_A + SOA) - RT_{VA}}{\min(RT_V, RT_A + SOA)} \times 100. \qquad (7)$$

Mean RT and DR with standard errors (black vertical lines) as a function of stimulus intensity and stimulus modality are summarized in Fig. 7. Panels on the left depict results from the FAP task, panels on the right depict results from the RTP task. Panels in the upper two rows display RTs and MRE, panels in the lower two rows display DRs and MDE. Note that RTs to auditory stimuli were corrected for SOA.

MRE and MDE displayed similar patterns for the FAP task, but differed in magnitude. For both measures, the largest magnitude of IFE was indicated for the presentation of a weak visual stimulus together with a high-intensity auditory stimulus (vA) and the smallest amount of IFE was observed when high-intensity visual stimuli were paired with low-intensity auditory stimuli (Va). MRE was slightly smaller for the presentation of two weak stimuli (va) compared to the presentation of two high-intensity stimuli (VA), whereas MDE was much larger for two weak stimuli (va) compared to two high-intensity stimuli (VA).

For the RTP task, the patterns of MRE and MDE were different. MRE was about the same magnitude for conditions with high-intensity auditory stimuli (VA and vA) and smaller for conditions with low-intensity auditory stimuli (Va and va). MDE was about 0 for condition VA and increased stepwise for the remaining conditions Va, vA, and va.

Evaluation of differences in overall performance utilizing bootstrapping procedures

For the quantification of overall performance, 1,000 bootstrap samples were generated from the original data for both FAP and RTP. Median IES and median model parameter for the bootstrapped data sets are given in Table 5. Both IES and drift rates exhibited similar patterns of overall performance. Regardless of the experimental instruction, performance for auditory and
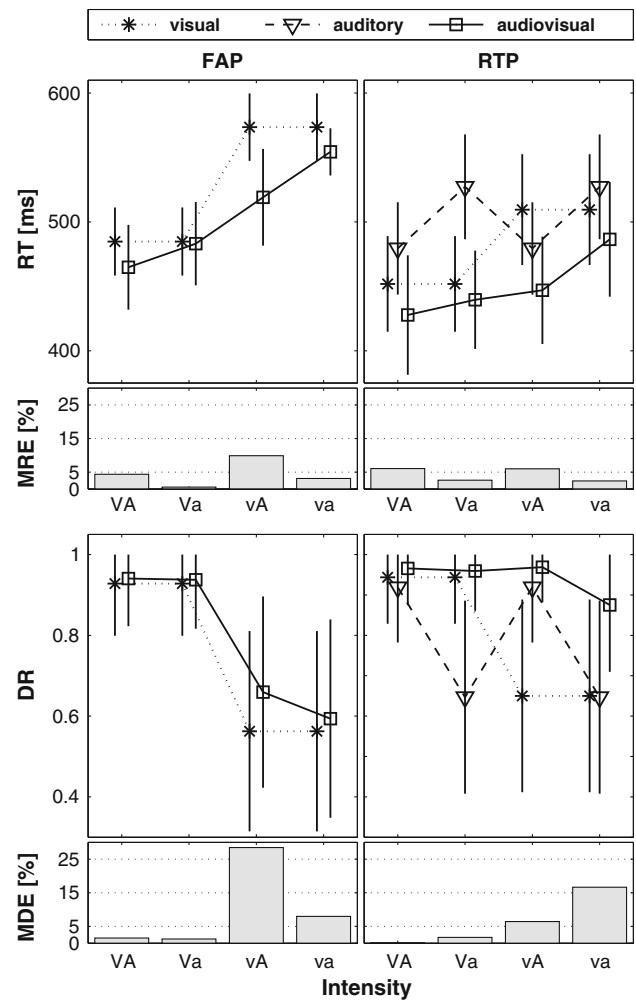

Fig. 7 Experiment 2: performance in a detection experiment quantified from reaction time (*RT*) and detection rate (*DR*). *Left column* data recorded in a focused attention paradigm (*FAP*); *right column* data recorded in a redundant target paradigm (*RTP*). *Upper row* Mean foot RT (ms) with standard errors (*black vertical lines*) as a function of stimulus intensity and modality. The *bar plot* presents multisensory response enhancement (%) as a function of stimulus intensity. *Lower row* DR with standard errors (*black vertical lines*) as a function of stimulus intensity. The *bar plot* presents multisensory detection enhancement (%) as a function of stimulus intensity

visual unimodal conditions was higher for high-intensity stimuli than for low-intensity stimuli (i.e., lower IES or higher drift rates). For the bimodal conditions, the pattern of performance differed between RTP and FAP. In the RTP, the highest performance was observed for condition VA, in conditions vA and Va it was lower, and it much lower for va. Thus, in the RTP, high performance was observed whenever at least one of the stimuli was of high intensity. In the FAP, again the highest performance was observed for VA, in condition Va it was lower, and it was much lower for conditions vA and va. Thus, in the FAP, high performance was observed whenever the intensity of the visual stimulus was high.

Table 5 Experiment 2: median diffusion model parameters and median inverse efficiency scores (IES) from 1,000 non-parametric bootstrap samples

Condition	Diffusion model			IES	
	Parameter	RTP	FAP	RTP	FAP
V	δ_V	0.2440	0.3461	466.0	533.5
v	δ_v	0.0389	0.0658	820.9	1,183.6
A	δ_A	0.1657	–	500.8	–
a	δ_a	0.0343	–	857.2	–
VA	δ_{VA}	0.5623	0.4117	425.8	497.1
Va	δ_{Va}	0.3297	0.3201	439.3	521.7
vA	δ_{vA}	0.3131	0.1043	440.9	811.8
va	δ_{va}	0.1198	0.0853	552.5	1,032.1
	θ	11	13		
	β	0	−0.32		
	T_r	382	384		

Table 6 Experiment 2: confidence intervals, CI_{68}, calculated by the bootstrap percentile method as a function of stimulus intensity

Task	Intensity	Percentiles			
		MRE*		MPE	
		16th	84th	16th	84th
FAP	VA	3.18	10.14	3.49	64.61
	Va	−1.79	5.65	−13.64	14.23
	vA	13.87	34.95	36.97	693.61
	va	1.53	17.71	3.72	238.33
RTP	VA	4.77	11.36	50.53	193.59
	Va	2.36	8.00	10.83	81.53
	vA	9.27	14.82	55.55	170.67
	va	12.67	21.41	63.26	218.56

For MRE*, and MPE the 16th and the 84th percentile are given as lower and upper limit of the confidence interval. Visual and auditory stimuli were presented with two different intensities. Lower-case letters indicate low-intensity stimuli; upper-case letters indicate high-intensity stimuli
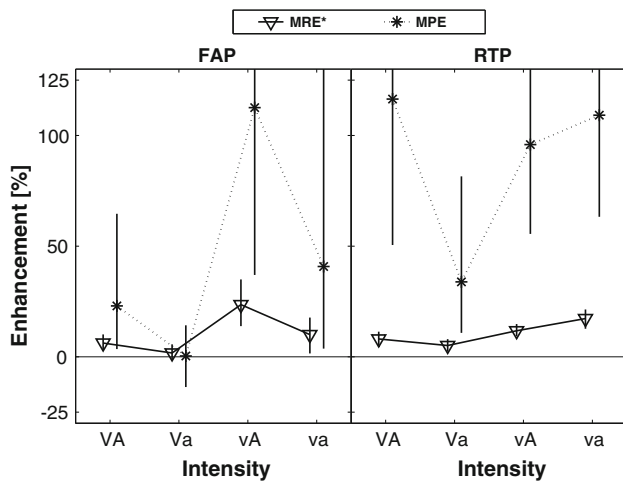

Fig. 8 Experiment 2: magnitude and variation of multisensory indices as a function of experimental paradigm and stimulus intensity. *Left panel* data recorded in a focused attention paradigm (*FAP*); *right panel* data recorded in a redundant target paradigm (*RTP*). *Open triangles* display multisensory response enhancement calculated from inverse efficiency scores (MRE*); *asterisks* display multisensory performance enhancement (MPE). *Black vertical lines* indicate confidence intervals (CI), CI_{68}, obtained in a non-parametric bootstrap procedures (see also Table 6)

To quantify multisensory interaction effects, MRE* and MPE were calculated for FAP and RTP tasks. For the FAP results, the minimum (respectively the maximum) of the unimodal responses was replaced by the visual response in Eqs. 4 and 5. For the RTP results, $RT_A^* + SOA$ instead of RT_A^* was entered into Eq. 4 and $RT_A + SOA$ was entered into the model for the calculation of RDC. Bootstrapped MRE* (asterisks), and MPE (open triangles) with CI_{68} (black vertical lines) are presented in Fig. 8. Results for the FAP task are given in the left panel, results for the RTP

task are displayed in the right panel. CI_{68} are also listed in Table 6.

For the FAP task, MRE* and MPE yielded similar patterns. The highest amount of IFE was observed for condition vA and both measures were about zero for condition Va. Furthermore, the magnitude of IFE was larger for condition va compared to condition VA, which is in line with the POIE. Significant differences in terms of non-overlapping bootstrap confidence intervals, CI_{68}, were observed as follows. Significant increases of MRE* were observed between conditions Va and VA, as well as between Va and vA. For MPE, only the difference between vA and Va was significant.

The patterns of MRE* and MPE were also similar in the RTP task, as both measures increased stepwise from condition Va, to vA, and va. However, MRE* was significantly smaller for VA than for va, supporting the POIE, whereas MPE yielded an insignificant opposite trend. The difference between MRE* in conditions Va and vA was also significant. No significant differences were observed for MPE.

Discussion

We conducted an experiment comparing the influence of different experimental instructions on overall performance in an audio–visual detection task. Participants were either instructed to respond to any stimulus they perceive (RTP) or to respond only to visual stimuli and to ignore auditory ones (FAP).

For all experimental tasks, performance on bimodal conditions was higher than performance on the unimodal

visual condition, no matter whether performance was quantified in terms of RT, DR, or IES. This replicates earlier reports of RT reductions and DR improvements due to crossmodal stimulation. In the RTP task, the amount of multisensory interaction measured in terms of MRE* and MPE was larger when two stimuli of low intensity were presented together (va), compared to conditions where two stimuli of high intensity were presented (VA). This finding is in accordance with other studies reporting evidence in favor of the POIE.

In the FAP task, MRE* and MPE indicated no multisensory enhancement for the conditions Va and significantly more for condition vA. This observation makes sense in an FAP task, because the stronger auditory non-target in condition vA provides more response activation, whereas the high-intensity visual target in condition Va reduces the influence of the auditory non-target because it already provides enough response activation itself (cf., Bernstein, Chu, Briggs & Schurman, 1973; Bernstein et al., 1970; Colonius & Diederich, 2004). Furthermore, multisensory enhancement observed in condition vA was even larger than in condition va, what may appear as evidence against the POIE at a first glance, because the overall intensity was smaller for the latter stimulus combination. However, in the FAP the POIE only accounts for the intensity of the target stimulus and not for the non-target, because participants are advised to respond only to the target. In the FAP, multisensory interaction occurs if a non-target is able to provide additional response activation, facilitating the response to the target. The relative effectiveness of the non-target, and therefore the amount of additional response activation, can be increased by either rasing its intensity (i.e., direct effectiveness), or by decreasing the intensity of the target (i.e., inverse effectiveness). Note, however, that this principle only works within a limited range: if the intensity of the non-target is increased beyond a certain point, the unimodal RT difference between target and non-target may become too large to result in multisensory interaction. On the other hand, participants will stop responding if the intensity of the target is decreased below the detection threshold.

In concordance with our hypotheses, the difference between the relative response enhancement observed in conditions vA and Va was smaller in the RTP than in the FAP. The assumed symmetry between the stimuli in the RTP was nicely reflected in both IES and drift rates, as both did not differ between vA and Va. For FAP, both IES and drift rates did differ between these conditions: both did indicate a lower overall performance in condition vA. This suggests that the intensity of the visual target stimulus had a larger influence on overall performance than the intensity of the non-target, which, as explained above, had a larger

influence on the amount of response enhancement than the intensity of the target.

The diffusion model framework utilized for the calculation of MPE invites speculation about the nature of the observed differences between RTP and FAP. Despite the overall mean RT difference between RTP and FAP, the estimated base time T_r was the same in both paradigms (see Table 5). The slower mean RT in the FAP was reflected by a higher decision criterion θ and a negative β, that is, an initial state of evidence in favor of the alternative "Do not respond". These differences suggest that in the FAP it was necessary to accumulate more information to elicit a response. This interpretation parallels the actual difference in task, as the RTP demands to decide whether a stimulus is present or not, whereas for the FAP it is necessary to identify the modality of the stimuli.

Summary and general discussion

In many experimental studies, multisensory interaction effects have been observed for both RT and DR. Here we suggested two quantitative indices for these effects that combine RT and DR into a single measure. Both measures relate the unimodal condition with higher performance to the performance under bimodal stimulation. The first index, MRE*, is a descriptive measure based on an arithmetical combination of RT and DR. It is calculated from inverse efficiency scores (IES), that is, RT divided by DR. The second index, MPE, is a model-based measure founded on a sequential sampling model. It is calculated from drift rates of a sequential sampling process (Wiener process) fitted to RT and DR.

Applied to the results from two detection experiments, the two approaches provided converging evidence. Both indices showed patterns that were similar to each other and in line with previous research. Nevertheless, statistical conclusions in terms of bootstrap confidence intervals derived for both indices did not always coincide. The non-parametric bootstrap was utilized for both MRE* and MPE for the sake of better comparability, although the model-based approach of MPE would have allowed for a parametric bootstrap, which often results in smaller confidence intervals.

Apart from the results obtained in the experiments reported here, both measures can be compared in terms of their validity, reliability, and practicability. Validity describes to which extent a measure really assesses the underlying theoretical construct. As mentioned before, a theoretical foundation for calculating IES (by dividing RT by DR) is lacking, it appears a bit ad hoc, and this questions the validity of MRE*. On the other hand, for MPE the theoretical background of sequential sampling models is

quite elaborate, these models are successful in describing results of simple reaction tasks as well as choice reaction tasks, and neural underpinning of the model mechanisms have been identified (Diederich, 1995, 1997, 2008; Diederich & Busemeyer, 2003; Luce, 1986; Ratcliff & Smith, 2004).

In order to be reliable, a measure is to remain nearly invariant when calculated repeatedly for the same data. For MRE* this is clearly the case as only elementary arithmetic is involved. This is not necessarily the case for MPE as it involves parameter estimation procedures that may lead to different results depending on particular starting values, objective functions, and minimization routines. Furthermore, achieving a fit does not imply the global minimum to be found, i.e., there might be better solutions. Reliability can be crucial in bootstrap studies where thousands of model fits are necessary, since very frequent poor model fits can add to the width of observed confidence intervals. Therefore, the reliability of MPE critically depends on adequate choices for optimization (relevant expertise can be found in the literature, cf., Wichmann & Hill, 2001).

When it comes to practicability, MRE* is quite easy to handle because IES is computed right away, and so is MRE*. On recent computers, non-parametric bootstrap takes only a couple of minutes, even for large data sets. On the contrary, MPE requires model fits with large numbers of parameters and thereby relies on the quality of the estimation routine. Although we utilized a rather "brute force" approach (we passed estimated parameters multiple times through the estimation routine as start values, see Appendix for details), the computational effort was not substantial. However, for bootstrap studies computation time multiplies and becomes an issue indeed because a model needs to be fitted to each of the generated data sets.

We are aware of approaches that combine a good theoretical foundation with high reliability and practicability. For instance, it is possible to define multisensory indices based on average information rates (AIR, e.g., Fitts, 1966; Hick, 1952) calculated for unimodal and bimodal conditions. AIR is one of the information theoretical approaches that were adopted in psychology in the course of the so-called *cognitive revolution* more than 50 years ago (see Proctor & Vu, 2006, for a review, but also Baird, 1984, and Luce, 2003, for critical views). It can be used to measure the capacity of sensory channels in absolute identification tasks adopting the concepts of entropy and transmitted information (e.g., Baird, 1984; Fitts, 1966; Hick, 1952). AIR is calculated from RTs and accuracy data and yields channel capacity in terms of transmitted information units per time unit. However, since at least two signals (i.e., at least stimulus vs. no stimulus) are necessary to compute AIR, this approach cannot be applied to data obtained in simple detection experiments without catch trials like those

presented in this study. Nevertheless, it appears as a promising approach for multisensory interaction assessment, in particular because it allows for the evaluation of choice tasks with more than two alternatives.

To conclude, with two different approaches we have demonstrated that speed and accuracy can be integrated into a single measure quantifying multisensory enhancement, and how the calculation of confidence intervals can be attempted with bootstrap procedures. Although both approaches led to good results, it appears indicated to further examine the properties of the presented measures in simulation studies which, however, are not within in the scope of this paper. Whenever computational power and time are available, we advocate the use of MPE together with the parametric bootstrap, as it outperforms MRE* in terms of validity. When computational power is an issue or a quick first impression is intended, MRE* can be beneficial. Ideally, both measures are used together in a multi-method approach to cross-validate conclusions.

Appendix

Bootstrap procedure

For each of 6 participants, the original data set consisted of responses to 32 trials for each of 24 conditions. RTs for trials without a valid response were set to 0. For the non-parametric bootstrap, 1,000 random samples were generated as follows.

Repeat 1,000 times:

1. Draw a set of six participants with replacement from the original set of participants.
2. For each of these participants:

   (a) for each of 24 conditions, draw a random sample with replacement from the 32 responses and calculate mean RT (mean of non-zero elements) and DR (number of non-zero elements divided by 32), resulting in a simulated data set (24 pairs of mean RT and DR).

   (b) calculate RT* and MRE* from the simulated data set.

   (c) fit a sequential sampling model to the simulated data set to obtain drift rates and calculate RDC.

3. Average the obtained measures across the set of six participants yields RT*, MRE*, the drift rates, and RDC for one random sample.

Parameter estimation

For the original data set and each of 1,000 resampled non-parametric bootstrap data sets, 24 drift rates $\delta_i \in \Delta$, one boundary separation θ, one β, and one residual time T_r were estimated from 24 RT and 24 DRs (8 intensity levels $\times$ 3 modalities). The estimation routine utilized the following constraints on parameters: each $\delta_i \in \Delta$ was bound to $0.001 < \delta_i < 0.999$, β was bound to $-1 < \delta_i < 1$, and the residual time was bound to $T_r \geq 0$. Parameter estimation was attempted in a three-step procedure:

> Step 1: Estimate parameters θ, β, T_r, and Δ by minimizing χ^2. Store χ^2.
>
> Step 2: Hold constant θ, β, and T_r, while estimating each $\delta_i \in \Delta$ individually by minimizing χ^2.[5]
>
> Step 3: Estimate parameters θ, β, T_r, and Δ by minimizing χ^2. If the new χ^2 is smaller than the stored one, jump to Step 2, or break, otherwise.

MatLab's (The Mathworks, Natick, MA) routine FMINSEARCH was used for parameter estimation.

References

Amlôt, R., Walker, R., Driver, J., & Spence, C. (2003). Multimodal visual–somatosensory integration in saccade generation. *Neuropsychologica, 41*, 1–15.

Arieh, Y., & Marks, L. E. (2008). Cross-modal interaction between vision and hearing: A speed–accuracy analysis. *Perception & Psychophysics, 70*, 412–421.

Arndt, P. A., & Colonius, H. (2003). Two stages in crossmodal saccadic integration: Evidence from a visual–auditory focused attention task. *Experimental Brain Research, 150*, 417–426.

Baird, J. C. (1984). Information theory and information processing. *Information Processing and Management, 20*, 373–381.

Bernstein, I. H., Chu, P. K., Briggs, P., & Schurman, D. L. (1973). Stimulus intensity and foreperiod effects in intersensory facilitation. *Journal of Experimental Psychology, 25*, 171–181.

Bernstein, I. H., Clark, M. E., & Edelstein, B. A. (1969a). Effects of an auditory signal on visual reaction time. *Journal of Experimental Psychology, 80*(3), 567–569.

Bernstein, I. H., Clark, M. H., & Edelstein, B. A. (1969b). Intermodal effects in choice reaction time. *Journal of Experimental Psychology, 81*(2), 405–407.

Bernstein, I. H., & Edelstein, B. A. (1971). Effects of some variation in auditory input upon visual choice reaction time. *Journal of Experimental Psychology, 87*(2), 241–247.

Bernstein, I. H., Rose, R., & Ashe, V. M. (1970). Energy integration in intersensory facilitation. *Journal of Experimental Psychology, 86*(2), 196–203.

Bolognini, N., Frassinetti, F., Serino, A., & Làdavas, E. (2005). "acoustical vision" of below threshold stimuli: interaction among spatially converging audiovisual inputs. *Experimental Brain Research, 160*, 273–282.

Colonius, H., & Diederich, A. (2004). Multisensory interaction in saccadic reaction time: A time-window-of-integration model. *Journal of Cognitive Neuroscience, 16*(6), 1–10.

Corneil, B. D., Van Wanrooij, M., Munoz, D. P., & Van Opstal, A. J. (2002). Auditory–visual interactions subserving goal-directed saccades in a complex scene. *Journal of Neurophysiology, 88*(1), 438–454.

Diederich, A. (1992). *Intersensory facilitation: Race, superposition, and diffusion models for reaction time to multiple stimuli (Dissertation)*. Frankfurt: Verlag Peter Lang.

Diederich, A. (1995). Intersensory facilitation of reaction time: Evaluation of counter and diffusion coactivation models. *Journal of Mathematical Psychology, 39*, 197–215.

Diederich, A. (1997). Dynamic stochastic models for decision making with time constraints. *Journal of Mathematical Psychology, 41*, 260–274.

Diederich, A. (2008). A further test of sequential-sampling models that account for payoff effects on response bias in perceptual decision tasks. *Perception & Psychophysics, 70*, 229–256.

Diederich, A., & Busemeyer, J. R. (2003). Simple matrix methods for analyzing diffusion models of choice probability, choice response time and simple response time. *Journal of Mathematical Psychology, 47*, 304–322.

Diederich, A., & Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: Threshold-bound, drift-rate-change, or two-stage-processing hypothesis. *Perception & Psychophysics, 68*, 194–207.

Diederich, A., & Colonius, H. (1987). Intersensory facilitation in the motor component? A reaction time analysis. *Psychological Research, 49*, 23–29.

Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics, 66*(8), 1388–1404.

Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science, 1*, 54–75.

Eimer, M. (2001). Crossmodal links in spatial attention between vision, audition, and touch: evidence from event-related brain potentials. *Neuropsychologia, 39*, 1292–1303.

Fitts, P. M. (1966). Cognitive aspects of information processing. III. Set for speed versus accuracy. *Journal of Experimental Psychology, 71*, 849–857.

Frassinetti, F., Bolognini, N., & Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research, 147*, 332–343.

Frens, M. A., Van Opstal, A. J., & Van der Willigen, R. F. (1995). Spatial and temporal factors determine auditory–visual interaction in human saccadic eye movements. *Perception & Psychophysics, 57*(6), 802–816.

Gielen, S. C. A. M., Schmidt, R. A., & Van den Heuvel, P. J. M. (1983). On the nature of intersensory facilitation of reaction time. *Perception & Psychophysics, 34*(2), 161–168.

Gillmeister, H., & Eimer, M. (2007). Tactile enhancement of auditory detection and perceived loudness. *Brain Research, 1160*, 58–68.

Giray, M., & Ulrich, R. (1993). Motor coactivation revealed by response force in divided and focused attention. *Journal of Experimental Psychology: Human Perception and Performance, 19*(6), 1278–1291.

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General, 136*, 389–413.

Harrington, L.K., & Peck, C.K. (1998). Spatial disparity affects visual–auditory interactions in human sensorimotor processing. *Experimental Brain Research, 122*, 247–252.

---

[5] Note that, for constant θ, β, and T_r, all $\delta_i \in \Delta$ are independent from each other, thus can be estimated separately.

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

Hick, W. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*, 11–26.

Hilgard, E. R. (1933). Reinforcement and inhibition of eyelid reflexes. *Journal of General Psychology*, *8*, 85–113.

Holmes, N. P. (2009). The principle of inverse effectiveness in multisensory integration: Some statistical considerations. *Brain Topography*, *21*, 168–176.

Jepma, M., Wagenmakers, E.-J., Band, G. P. H., & Nieuwenhuis, S. (2008). The effects of accessory stimuli on information processing: Evidence from electrophysiology and a diffusion model analysis. *Journal of Cognitive Neuroscience*, *21*, 847–864.

Kitagawa, N., & Spence, C. (2005). Investigating the effect of a transparent barrier on the crossmodal congruency effect. *Experimental Brain Research*, *161*, 62–71.

Lovelace, C. T., Stein, B. E., & Wallace, M. T. (2003). An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Cognitive Brain Research*, *17*, 447–453.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organisation*. Oxford Psychology Series; no. 8. New York: Oxford University Press.

Luce, R. D. (2003). Whatever happened to information theory in psychology? *Review of General Psychology*, *7*, 183–188.

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience*, *7*(10), 3215–3229.

Meredith, M. A. & Stein, B. E. (1986). Visual, auditory and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, *56*(3), 640–662.

Miller, J. O. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*, 247–279.

Miller, J. O. (1986). Time course of coactivation in bimodal divided attention. *Perception & Psychophysics*, *40*(5), 331–343.

Miller, J. O., Ulrich, R., & Lamarre, Y. (2001). Locus of the redundant-signals effect in bimodal divided attention: a neurophysiological analysis. *Perception & Psychophysics*, *63*(3), 555–562.

Mordkoff, J. T., Miller, J., & Roch, A.-C. (1996). Absence of coactivation in the motor component: evidence from psychophysiological measures of target detection. *JJournal of Experimental Psychology: Human Perception and Performance*, *22*(1), 25–41.

Morrell, E. K. (1968). Temporal characteristics of sensory interaction in choice reaction times. *Journal of Experimental Psychology*, *77*(1), 14–18.

Nickerson, R. S. (1973). Intersensory facilitation of reaction time: Energy summation or preparation enhancement. *Psychological Review*, *80*, 489–509.

Perrault, T. J., Vaughan, J. W., Stein, B. E., & Wallace, M. T. (2005). Superior colliculus neurons use distinct operational modes in the integration of multisensory stimuli. *J Neurophysiology*, *93*, 2575–2586.

Proctor, R. W., & Vu, K.-P.L. (2006). The cognitive revolution at age 50: Has the promise of the human information-processing approach been fulfilled? *International Journal of Human–Computer Interaction*, *21*, 253–284.

Rach, S., & Diederich, A. (2006). Visual–tactile integration: does stimulus duration influence the relative amount of response enhancement? *Experimental Brain Research*, *173*, 514–520.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*(2), 333–367.

Röder, B., Kusmierek, A., Spence, C., & Schicke, T. (2007). Developmental vision determines the reference frame for the multisensory control of action. *PNAS*, *104*, 4753–4758.

Schwarz, W. (1994). Diffusion, superposition, and the redundant-targets effect. *Journal of Mathematical Psychology*, *38*, 504–520.

Shore, D. I., Barnes, M. E., & Spence, C. (2006). Temporal aspects of the visuotactile congruency effect. *Neuroscience Letters*, *392*, 96–100.

Spence, C., McGlone, F. P., Kettenmann, B., & Kobal, G. (2001). Attention to olfaction: A psychophysical investigation. *Experimental Brain Research*, *138*, 432–437.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. London, Cambridge: The MIT Press.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modelling of elementary psychological processes*. New York: Cambridge Univ Press.

Townsend, J. T., & Honey, C. J. (2007). Consequences of base time for redundant signals experiments. *Journal of Mathematical Psychology*, *51*, 242–265.

Walker, R., Deubel, H., Schneider, W. X., & Findlay, J. M. (1997). Effect of remote distractors on saccade programming: Evidence for an extended fixation zone. *Journal of Neurophysiology*, *78*, 1108–1119.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, *63*, 1314–1329.

Woodworth, R. S., & Schlosberg, H. (1956). *Experimental psychology* (rev. ed.). New York: Holt.