



JACOBS
UNIVERSITY

On quantifying multisensory interaction effects in the vicinity of detection thresholds

by

Stefan Rach

A thesis submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy in Psychology

Approved, Thesis committee:

Prof. Dr. Adele Diederich

Prof. Dr. Claus Hilgetag

Prof. Dr. Hans Colonius

Prof. Dr. Thomas Münte

Prof. Dr. Benjamin Godde

Date of Defence: 26.11.2009

School of Humanities and Social Science

Abstract

Responses to multiple stimuli from different modalities tend to be faster compared to responses to each of these stimuli alone. In two experiments, we investigated the relation between stimulus duration and the relative amount of response enhancement. Visual, tactile, and visuo-tactile stimuli of different durations were presented. Participants were required to gaze only at visual stimuli and to ignore tactile stimuli (focused attention paradigm). Saccadic reaction times (RT) were recorded. Results suggest that the amount of response enhancement was largest for the shortest stimulus duration and decreases with increasing stimulus duration, i.e., inverse effectiveness of stimulus duration.

The assessment of multisensory interaction effects critically depends on the investigator's choice of measuring a participant's performance. For very weak stimuli, performance quantification from either RT or detection rates (DR) alone possibly misses out important information. We consider two principally different ways of integrating RT and DR to quantify overall performance. One measure, inverse efficiency scores (IES), is based on an arithmetical combination of RT and DR, and the other utilizes sequential sampling models. Statistical properties of both measures are investigated via bootstrapping procedures.

In an audio-visual detection task, we investigated the influence of different experimental instructions ("respond to any stimulus" vs. "respond only to visual stimuli") and the presented modalities (visual and audio-visual stimuli vs. visual, audio-visual, and auditory stimuli) on overall performance and the amount of multisensory interaction. RT and DR were recorded, overall performance was quantified in terms of IES, and the amount of multisensory interaction was quantified by calculating multisensory response enhancement from IES.

A probability inequality introduced in Miller (1982) puts an upper limit on the amount of RT facilitation within the redundant target paradigm that is consistent with a race model. Here, it is shown that inferences from this inequality test may become invalid when the experimenter misses a proportion of the responses by limiting the recording interval (right-censoring) or excludes outliers from analysis (left- and/or right-censoring). Moreover, a correction of the inequality test for right-censored RT distributions is proposed.

Contents

List of Figures	IV
List of Tables	V
List of Abbreviations	VI
1 General introduction	1
2 Visual-tactile integration: Does stimulus duration influence the relative amount of response enhancement?	11
2.1 Introduction	11
2.2 Experiment 1	14
2.2.1 Methods and Apparatus	14
2.2.2 Results	17
2.3 Experiment 2	19
2.3.1 Methods and Apparatus	20
2.3.2 Results	21
2.4 Discussion	23
3 On quantifying multisensory interaction effects in reaction time and detection rate	26
3.1 Introduction	26
3.2 Quantification of multisensory interaction effects	28
3.2.1 Measures of Multisensory Speed-up and Detectability	28
3.2.2 Inverse efficiency scores	30
3.2.3 Multisensory Performance Enhancement	31
3.3 Methods	33
3.3.1 Participants	33
3.3.2 Apparatus	34
3.3.3 Stimuli	34
3.3.4 Procedure	35
3.4 Results	35
3.4.1 MRE* calculated from Inverse Efficiency Scores	38
3.4.2 MPE calculated from drift rates	39
3.4.3 Evaluation of differences in overall performance utilizing bootstrapping procedures	42
3.5 Summary and Discussion	44
3.6 Appendix	48
3.6.1 Parameter estimation	48

3.6.2	Bootstrap procedures	48
3.6.3	MatLab script for the simple Wiener process	49
4	Comparing redundant target and focused attention paradigm:	
	A descriptive analysis	52
4.1	Introduction	52
4.2	Methods & Apparatus	56
4.3	Results	59
4.3.1	Detection rates	59
4.3.2	Reaction times	61
4.3.3	Overall performance in terms of inverse efficiency scores . . .	63
4.3.4	Relative amount of response enhancement in terms of MRE*	65
4.4	Summary & Discussion	68
4.5	Appendix	70
5	The race model inequality for censored reaction time distributions	74
5.1	Introduction	74
5.2	Race model inequality for right-censored RT distributions	77
5.3	Race model inequality for left-censored RT distributions	79
5.4	Correcting for right-censoring	82
5.5	Amount of inferential errors caused by censoring: Simulation results	84
5.6	Concluding Remarks	88
5.7	Appendix	90
6	General summary and discussion	92
	Bibliography	97
	Acknowledgements	111
	Declaration by Word of Honor	112

List of Figures

2.1	Experiment 1: Multisensory integration as a function of stimulus duration and laterality.	18
2.2	Experiment 2: Multisensory integration as a function of stimulus duration and laterality	22
3.1	Sampling path and drift rate of a sequential sampling model	32
3.2	Time course of a trial	36
3.3	Performance in a detection experiment quantified from reaction time and detection rate	37
3.4	Overall performance in a detection experiment quantified from inverse efficiency scores	39
3.5	Overall performance in a detection task quantified from parameters of a sequential sampling model	41
3.6	Magnitude and variation of multisensory indices as a function of stimulus intensity	44
4.1	Detection rate as a function of stimulus onset asynchrony, stimulus intensity, and experimental paradigm.	60
4.2	Reaction time as a function of stimulus onset asynchrony, stimulus intensity, and experimental paradigm.	62
4.3	Inverse efficiency scores as a function of stimulus onset asynchrony, stimulus intensity, and experimental paradigm.	64
4.4	Multisensory response enhancement calculated from inverse efficiency scores as a function of stimulus onset asynchrony, stimulus intensity, and experimental paradigm.	66
5.1	Violations of the race model inequality for right-censored distributions	80
5.2	Violations of the race model inequality for left-censored distributions	83
5.3	Influence of censoring on the diagnosticity of the race model inequality as a function of coactivation gain, censoring method, and percentage of excluded RTs	86

List of Tables

3.1	Intensities utilized given as luminance of visual and loudness of auditory stimuli	35
3.2	Mean reaction time and detection rate as a function of stimulus intensity and sensory modality	36
3.3	Confidence intervals calculated by the bootstrap percentile method as a function of stimulus intensity	43
4.1	Four variations of two experimental paradigm	53
4.2	Participant specific luminance of visual stimuli, loudness of auditory stimuli and stimulus onset asynchronies utilized in the experiment. .	58
4.3	Detection rate as a function of intensity, stimulus onset asynchrony, and experimental paradigm.	71
4.4	Reaction time as a function of intensity, stimulus onset asynchrony, and experimental paradigm.	72
4.5	Multisensory response enhancement calculated from inverse efficiency scores as a function of intensity, stimulus onset asynchrony, and experimental paradigm.	73

List of Abbreviations

CI	confidence interval
DLSC	deep and intermediate layers of the colliculus superior
DR	detection rate
FPT	first passage time
FAP	focused attention paradigm
IES	inverse efficiency score
IFE	intersensory facilitation effect
LED	red light emitting diode
MDE	multisensory detection enhancement
MPE	multisensory performance enhancement
MRE	multisensory response enhancement
MSE	modality shift effect
POIE	principle of inverse effectiveness
RMI	race model inequality
RSE	redundant signals effect
RT	reaction time
RTP	redundant target paradigm
RTTM	regression towards the mean
SATF	speed–accuracy tradeoff function
SRT	saccadic reaction time

1 General introduction

*Dark night, that from the eye his function takes,
The ear more quick of apprehension makes;
Wherein it doth impair the seeing sense,
It pays the hearing double recompense.*

Shakespeare,
Midsummer Night's Dream, III, 2, 1217

In experimental psychology, sensory modalities are typically investigated separately from each other. When examining a particular sensory system, the experimenter usually tries to eliminate possible influences from other modalities for not having the outcome of the experiments contaminated by unintentional effects. However, for most animals including humans interaction between the senses is crucial for adequate behavior in the environment, and intersensory interaction is the rule rather than an exception. Over the last two decades or so the research on intersensory interaction has caught the interest of both the neurosciences, in particular neurophysiology, and the behavioral sciences, in particular psychology. The studies presented in this thesis are intended to contribute to the huge jigsaw revealed so far.

Most organisms perceive the surrounding world through different sensory systems which are – at least in the most peripheral parts of the nervous system – separated from each other. To provide a coherent impression of the environment, the information delivered by these distinct sensory systems has to be combined at some point within the nervous system to create a unitary mental representation of objects in the physical world, based on temporal and spatial characteristics of the unimodal percepts. Normally these processes are automated, and we only rarely get notice of their existence. For instance, it is unpleasant and confusing to watch a movie with a synchronization that does not match the pace of the lip movements of the actors. Apart from such rather rare exceptions, *multisensory interaction* happens without us being aware of it: we might benefit from faster or more accurate responses when congruent information from more than one sensory modality

is available, and our responses can be slower or less accurate when the incoming information is contradictory. In the following, some of the main findings of behavioral studies on multisensory interactions are briefly reviewed, a glance at the neural basis underlying these processes is presented, and then the course of this thesis is outlined.

Starting from the pioneering study of Todd (1912), a large body of research has focused on the behavioral consequences of multisensory interaction. From an empirical or experimental perspective, multisensory interaction has occurred when the response¹ to an event from one sensory modality is changed by the presence of a stimulus from a second sensory modality (cf., Welch & Warren, 1986). Consequently, a typical experimental design compares the performance in one (or more) unimodal conditions to the performance in crossmodal conditions. If the performance in crossmodal conditions is higher than that in unimodal ones, we speak of *multisensory enhancement*; if the crossmodal performance is lower we speak of *multisensory depression*. Response enhancement and depression do not constitute distinct phenomena, but rather they are two instances along the same continuum of multisensory interaction (Meredith & Stein, 1986). In the behavioral experiments reported in this thesis, response enhancement relates to an increase in response speed or an increase in the accuracy of the response, whereas response depression relates to the opposite. Many behavioral experiments utilize one out of two very prominent experimental paradigms. In the redundant target paradigm (RTP), participants are advised to respond to any stimulus they detect regardless of its sensory modality. In the focused attention paradigm (FAP), participants are instructed to respond only to stimuli from a specific modality and to ignore stimuli from other modalities. The findings reviewed in the following are mostly comparable for both paradigms unless noted otherwise.

For the first time, Todd (1912) reported that reaction time (RT) in crossmodal conditions is shorter than RT in unimodal conditions. This *intersensory facilitation*

¹Note that Welch & Warren (1986) refer to the “perception” changing due to the presence of a second stimulus. Because we cannot observe changes in perceptions, we are left with observing changes in overt behavior.

effect (IFE, cf., Hershenson, 1962) has been replicated for visual–auditory stimulation (*e.g.*, Bernstein et al., 1969; Colonius & Arndt, 2001; Diederich & Colonius, 2008; Gielen et al., 1983; Steenken et al., 2008), for visual–tactile stimulation (*e.g.*, Amlôt et al., 2003; Diederich & Colonius, 2007; Gielen et al., 1983; Rach & Diederich, 2006), as well as for trimodal stimulation (*e.g.*, Diederich, 1992b, 1995; Diederich & Colonius, 2004).

The amount of RT facilitation is modulated by the spatial and temporal alignment of the stimuli. It decreases with increasing temporal separation (*temporal rule*, Bernstein et al., 1969, 1970; Diederich & Colonius, 1987, 2004; Giray & Ulrich, 1993; Hershenson, 1962; Miller, 1986; Morrell, 1968), as well as with increasing spatial separation (*spatial rule*, Amlôt et al., 2003; Bernstein & Edelman, 1971; Colonius & Diederich, 2004; Diederich & Colonius, 2007; Frens et al., 1995; Harrington & Peck, 1998; Walker et al., 1997). Moreover, the amount of IFE is larger when stimuli are less intense (*principle of inverse effectiveness*, POIE, Corneil et al., 2002; Diederich & Colonius, 2004; Gillmeister & Eimer, 2007; Ma et al., 2009; Rach & Diederich, 2006; Serino et al., 2007; but see Holmes, 2007, 2009, for a critical view).

These findings on the behavioral level are nicely complemented by evidence from neurophysiological studies (see Stein & Meredith, 1993, for a review). Single cell recordings from neurons in the deep and intermediate layers of the colliculus superior (DLSC) of higher mammals demonstrated that large populations of DLSC neurons are multisensory, *i.e.*, receive input from more than one sensory modality (*e.g.*, Meredith & Stein, 1986). These neurons exhibit enhanced responses on crossmodal stimulation, *i.e.*, the firing rates elicited by crossmodal stimulation often exceed the firing rates elicited by unimodal stimulation (*e.g.*, Bell et al., 2005; Meredith & Stein, 1986; Perrault et al., 2005; Stanford & Stein, 2007; Stein & Stanford, 2008) and adhere to similar spatio–temporal rules (King & Palmer, 1985; Populin & Yin, 2002; Wallace et al., 1996), as well as to the POIE (Meredith & Stein, 1986; Perrault et al., 2005; Stein & Meredith, 1993). Moreover, there is increasing evidence that many cortical structures receive inputs from more than one sensory modality, and even the visual and the auditory primary sensory cortices cannot be considered as

entirely unimodal anymore (Ghazanfar & Schroeder, 2006; Stein & Stanford, 2008).

Findings from both the behavioral and the neurophysiological research suggest that the relative magnitude of multisensory interaction effects might be largest for weak stimuli. However, controversy surrounds current discussions of the POIE (cf. Holmes, 2007, 2008, 2009; Ross et al., 2007). The main argument of this ongoing debate claims that the POIE might be a statistical artefact rather than an empirical principle. In this thesis, behavioral experiments are reported that provide evidence in line with the predictions based on the POIE, but cannot be explained by these alternative statistical explanations. Furthermore, the analysis of the results of these experiments point to some methodological problems that, independently from the validity of the inverse effectiveness hypothesis, arise if very weak stimuli are utilized in behavioral experiments. Some of these problems will be addressed in this thesis.

The outline of the thesis is as follows.

In Chapter 2, two experiments providing evidence supporting the POIE are reported.^{2,3} From neurophysiological studies on higher mammals it is known that multisensory integration effects are especially large when stimuli are weak (see Stein & Meredith, 1993, for a review). This observation led to the postulation of the POIE: stimuli of low intensity are relatively more effective in eliciting multisensory enhancement than more intense stimuli. Such a mechanism makes sense from an evolutionary perspective, as it enables the organism to benefit from multisensory interaction especially when the stimuli are weak and ambiguous. In such situations, the combination of information from different senses can reduce ambiguity and therefore enhance the adaptiveness of the organism. When, on the other hand, the unimodal percepts are unobscured the benefits of including information from other senses are lower as less additional gain of knowledge can be expected. The

²This chapter has been published previously as Rach, S. & Diederich, A. (2006). Visual–Tactile Integration: Does Stimulus Duration Influence the Relative Amount of Response Enhancement? *Experimental Brain Research* 173, 514–520. (DOI:10.1007/s00221-006-0452-4)

³The first experiment reported in this chapter was part of a thesis submitted in partial fulfillment of the requirements for the degree of Diplom–Psychologe: Rach S., (2004). Visual–tactile spatial interaction: Inverse effectiveness and variation of stimulus duration. Diplomarbeit (unpublished), Universität Oldenburg.

POIE has been often replicated on the neurophysiological level, but only to a lesser degree in behavioral studies (Bell et al., 2005; Corneil et al., 2002; Diederich & Colonius, 2004).

In this study, the duration of visual and tactile stimuli was varied to investigate whether duration, as a further determinant of stimulus effectiveness, is also inversely related to the relative amount of response enhancement. In the first experiment, visual and tactile stimuli with durations ranging from 30 to 500 ms were presented, and even shorter durations ranging from 10 to 100 ms were utilized in the second experiment. In each trial, either a visual (red light emitting diode, LED), a tactile (vibration applied to the palm) or a visuo-tactile stimulus was presented. Participants were instructed to gaze only at visual stimuli and to ignore tactile stimuli, *i.e.* an FAP was conducted. Saccadic reaction times were recorded, and the amount of multisensory interaction was calculated in terms of *multisensory response enhancement* (MRE; Diederich & Colonius, 2004), a descriptive measure that relates mean RT in the unimodal condition (LED-only; *i.e.*, RT_V), to that in the bimodal condition, (*i.e.*, RT_{VT}):

$$\text{MRE} = \frac{RT_V - RT_{VT}}{RT_V} \times 100. \quad (1.1)$$

In the first experiment, the relative amount of response enhancement was largest for the shortest stimulus durations, indeed providing evidence for inverse effectiveness. However, this effect was only observable for the condition with the shortest stimulus durations, and the results might have lead to a completely different interpretation if, for some reasons, this condition would not have been presented in this experiment. Therefore a second experiment focusing on a range of shorter durations was conducted. The results of the second experiment replicated the main findings from the first experiment, as the shortest stimulus durations again elicited the largest MRE. Moreover, MRE on ipsilateral conditions decreased with increasing stimulus duration, providing further evidence for inverse effectiveness of stimulus durations.

Although the stimuli utilized in these two experiments were weak, they were still

reliably detected by our participants, as indicated by the percentage of erroneous trials which was below 5 % in the first and below 15 % in the second experiment. But how would the outcomes change if even weaker stimuli are used? It is known that multisensory interaction also affects the detectability of stimuli. For instance, a task-irrelevant auditory stimulus can modulate visual perception (Bolognini et al., 2005; Frassinetti et al., 2002), a task-irrelevant visual stimulus can enhance auditory perception (Lovelace et al., 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). Therefore the approach of utilizing even weaker stimulus intensities raised a methodological question: if multisensory interaction effects come into effect not only for RT, but also for detection rates (DRs), how can these two measures be combined?

In Chapter 3 a study is reported that was specifically directed towards this question.⁴ Near to the detection threshold, both RT and DR change when stimulus intensity is varied (see, *e.g.*, Luce, 1986; Woodworth & Schlosberg, 1956, for a summary). To quantify multisensory interaction effects, it is important to know whether an improvement in response speed indicates a change in overall performance rather than being compensated by a worsening in accuracy. A general finding from studying such speed-accuracy tradeoffs is that speeded responses tend to be less accurate; for instance, in simple detection tasks, an increase in mean RT is often accompanied with a decreased DR and vice versa (see, *e.g.*, Luce, 1986, for a summary). The reported study presented and compared two principally different approaches to quantify overall performance in a crossmodal task utilizing weak stimuli. The first approach utilized inverse efficiency scores, a descriptive combination of RT and DR known from the speed-accuracy tradeoff literature (Townsend & Ashby, 1983) to combine RT and DR and compute a measure reflecting relative enhancement in crossmodal conditions. The second approach made use of a sequential sampling

⁴A major revision of this chapter has been invited for possible publication in Psychological Research as: Rach, S., Diederich, A., & Colonius, H. On quantifying multisensory interaction effects in reaction time and detection rate.

model which was fitted to RTs and DRs. Here we used specific parameters of this model to compute a measure reflecting relative enhancement in crossmodal conditions.

To investigate multisensory interaction effects for very weak stimuli we conducted an audio–visual simple detection experiment. In each trial either a visual stimulus (red LED), an auditory stimulus (burst of white noise), or both were presented. The intensity of both the visual and the auditory stimulus was varied in 8 steps. Participants were advised to activate a button with their large toe whenever they detected any stimulus (RTP). RT and DR were recorded as dependent measures.

We were able to demonstrate, that the assessment of multisensory interaction effects can lead to contradictory conclusions when either only RTs or only DRs were used for quantification: on some intensity conditions, multisensory enhancement was evident in DR, but it was absent in RT and vice versa. Both of the integrative measures developed in this study were able to pick up the characteristics of the results from both DR and RT, yielding an estimate of *overall performance*. Furthermore, for both measures, bootstrap confidence intervals indicated that the multisensory enhancement was largest for the stimuli with the lowest intensity. Note that this replicates the main findings of the first chapter and therefore adds evidence supporting the POIE hypothesis.

In Chapter 4, we used inverse efficiency scores to assess the influence of experimental instructions on overall performance in an audio–visual detection task. Many studies on multisensory interaction utilize either RTP tasks, where participants are instructed to respond to any stimulus they perceive, or FAP tasks, where participants are instructed to respond only to a target stimulus and to ignore all other stimuli. FAP tasks can be further distinguished by whether unimodal non-targets, so-called catch trials, are presented or not. The combination of two experimental instructions (RTP vs. FAP) and two stimulus presentations schemes (visual and audio–visual vs. visual, auditory, and audio–visual) resulted in four different experimental conditions, for which the following predictions were derived. RTP and FAP differ in task complexity, because RTP involves only the detection of stimuli,

whereas it is necessary to discriminate between different modalities in FAP tasks. Since performance should be higher in the easier task, faster RTs should be observed in the RTP task. Responses to a visual stimulus are expected to be slower, on average, if the preceding stimulus was of a different modality (modality shift effect, MSE, Gondan et al., 2004; Spence & Driver, 1997; Spence et al., 2001). Thus, the presentation of unimodal auditory trials in RTP tasks should prolong RTs to unimodal visual stimuli. In FAP tasks, it has been shown that the presentation of catch trials can prolong RT (Gielen et al., 1983). Presenting unimodal auditory trials in FAP tasks should therefore result in slower RTs.

RTs and DRs were recorded and overall performance was quantified in terms of inverse efficiency scores. For both RTs and inverse efficiency scores, the above mentioned predictions were confirmed. Furthermore, for three out of four experimental conditions, presenting two low intense stimuli resulted in higher amounts of response enhancement in terms of MRE*, compared to the presentation of two stimuli of high intensity, which adds further evidence in favor of the POIE.

In Chapter 5, we demonstrate that the utilization of very weak stimuli can also influence the diagnosticity of the race model inequality (RMI), a very common tool for the evaluation of multisensory interaction effects.⁵ The RMI bases on a model proposed by Raab (1962), which postulates that, for simple RT, each individual stimulus elicits a detection process performed in parallel to the others, and the winner's time determines the observable RT. According to this race model, the observed reduction in RT might be due to statistical effects, rather than multisensory processes: if detection latencies are interpreted as random variables, the time to detect the first of several redundant signals is faster, on average, than the detection time for any single signal. The RMI puts an upper limit on the amount of RT reduction that can be explained by statistical facilitation, *i.e.*, a violation against the RMI

⁵This chapter was accepted for publication in *Attention, Perception, & Psychophysics* on November, 14th, 2009 as: Rach, S., Diederich, A., Steenken, R., & Colonius, H. The race model inequality for censored reaction time distributions. (Scheduled to be published in volume 72, issue 3, April 2010.) The copyright is held by Psychonomic Society Publications. This document may not exactly correspond to the final published version. Psychonomic Society Publications disclaims any responsibility or liability for errors in this manuscript.

suggests that an interaction between the signals must have occurred, contributing to the observed RT reduction.

In this chapter, two proofs illustrate that the exclusion of data below or above a certain criterion (*i.e.*, censoring the RT distributions) can influence the results of RMI tests. Such data exclusions are often reported in the literature when researchers exclude data from further analysis by considering RTs below a certain criterion as anticipations and/or RTs above a certain criterion as misses (*e.g.* Giray & Ulrich, 1993; Leo et al., 2008; Miller, 2007b; Savazzi & Marzi, 2008; Schwarz, 2006). We show that the exclusion of RTs faster than a certain criterion (left-censoring) can lead to a more liberal RMI test, *i.e.*, more violations are detected than actually present in the uncensored data; and that excluding RTs slower than a criterion can lead to a more conservative test, *i.e.*, violations of the RMI that are present in the uncensored data are not detected in the censored data.

The results presented in this chapter clearly indicate that censoring should be avoided whenever possible. Nevertheless, there are situations when censoring is caused by the experimental setup. When very weak stimuli are utilized in an experiment, a participant's responses to a very weak stimulus can occur quite late. It is principally possible that an experimenter misses a proportion of the responses by not registering reactions that occur after some arbitrary upper bound. These trials are often considered as misses, *i.e.*, it is assumed that the participant failed to respond to the stimulus (*e.g.*, Gondan et al., 2007; Miller, 2007a,b,c). In such cases, the experimental setup imposes a right-censoring on the data, and we suggest a correction procedure that can restore the uncensored distributions if the proportion of missing trials is a known quantity. Furthermore, the appropriateness of this correction is demonstrated by simulation results. Similar correction can also be obtained for left-censoring or simultaneous censoring from the left and the right, but simulation results suggest that they provide only inaccurate restorations of the uncensored distributions.

In the final chapter, the studies presented in this thesis are put into a coherent context, and their relevance is discussed in the light of the ongoing debate regarding

the validity of the POIE (Holmes, 2007, 2008, 2009).

2 Visual-tactile integration: Does stimulus duration influence the relative amount of response enhancement?^{1,2}

Abstract

Responses to multiple stimuli from different modalities tend to be faster compared to responses to each of these stimuli alone. Neurophysiological studies on higher mammals and behavioral studies on humans suggest that the relative amount of enhancement is inversely related to stimulus intensity.

In two experiments, the duration of visual and tactile stimuli was varied to investigate whether duration, as a further determinant of stimulus effectiveness, is also inversely related to the relative amount of response enhancement. Visual and tactile stimuli were presented left or right of fixation either in the same or different hemifields. Participants were required to gaze only at visual stimuli and to ignore tactile (focused attention paradigm). Saccadic reaction times were recorded.

Results from both experiments show that the relative amount of response enhancement was largest for the shortest stimulus duration and decreases with increasing stimulus duration, *i.e.*, inverse effectiveness of stimulus duration.

Keywords Multisensory integration · Visual-tactile interaction · Stimulus duration · Inverse effectiveness · Saccadic reaction time

2.1 Introduction

Complex organisms usually perceive the surrounding world through more than one sensory modality. The information delivered through separate sensory channels has to be arranged according to temporal and spatial features in order to provide a coherent picture of the environment. This process of multisensory integration may hold further benefits for the organism, as reaction time (RT) studies have revealed. For instance, it has been demonstrated that responses to the onset of a visual stimulus are faster if the visual stimulus is accompanied by a spatially

¹This chapter has been published previously as Rach, S. & Diederich, A. (2006). Visual-Tactile Integration: Does Stimulus Duration Influence the Relative Amount of Response Enhancement? *Experimental Brain Research* 173, 514–520. (DOI:10.1007/s00221-006-0452-4) The original publication is available at www.springerlink.com.

²The first experiment reported in this chapter was part of a thesis submitted in partial fulfillment of the requirements for the degree of Diplom-Psychologe: Rach S., (2004). Visual-tactile spatial interaction: Inverse effectiveness and variation of stimulus duration. Diplomarbeit (unpublished), Universität Oldenburg.

adjacent stimulus from a different modality (*e.g.*, Diederich et al. 2003). Moreover, it was suggested that the amount of RT facilitation is determined also by their intensity (*e.g.*, Bernstein et al. 1970). If the amount of RT reduction is sensitive to the intensity of the stimuli, can it also be influenced by the duration of the stimuli?

Almost 100 years ago, Todd (1912) reported that manual responses to stimulus onsets are faster when the target stimulus is accompanied by a stimulus of another modality compared to the presentation of the target alone, even when the participants are advised to ignore the second stimulus (focused attention paradigm, FAP). This shortening of RTs due to the presentation of additional stimuli of different modalities has been termed intersensory facilitation effect (IFE; Hershenson 1962) and was replicated for visual-auditory stimulation (Bernstein et al. 1969; Colonius & Arndt 2001; Frens et al. 1995; Gielen et al. 1983), as well as for visual-tactile stimulation (Amlôt et al. 2003; Diederich et al. 2003; Gielen et al. 1983). It has been shown that the intensities of both target and non-target have an effect on the amount of RT facilitation (Bernstein et al. 1970, 1973, Colonius & Arndt 2001).

Moreover, utilizing a redundant target paradigm (RTP; *i.e.*, the participant is instructed to respond to any of the stimuli upon detection) Diederich & Colonius (2004) varied the intensity of tactile and auditory stimuli that were presented in addition to a visual stimulus and found larger IFE for lower intensities. In a visual-auditory RTP task, Corneil et al. (2002) varied the intensity of an auditory stimulus presented in a visual-auditory background noise and found larger response enhancement for lower signal-to-noise ratios.

Interestingly, the inverse relation between stimulus intensity and the amount of IFE reported when utilizing RTP is supported and put into a broader context by neurophysiological observations (Stein & Meredith 1993, Perrault et al. 2005). Cell recordings from the deep layers of the superior colliculus (DLSC) of anesthetized cats revealed that the majority of multisensory neurons show enhanced responses when they get input from more than one sensory modality, compared with the input from any of these modalities alone. Note that this paradigm is formally equivalent to an RTP task in the behavioral setting, since responses are evoked by

any stimulus without focusing on a certain target modality and ignoring the other. The amount of response enhancement was shown to depend on spatial and temporal disparity and stimulus effectiveness as well. The effectiveness of stimuli was varied by manipulating their size, velocity, directional preference, and intensity. In particular, it was found that the enhancement is larger when the stimuli are less effective. This principle, termed *inverse effectiveness*, culminates in the observation that two very weak stimuli of different modalities can elicit a response from a multisensory neuron, even if this neuron responds to none of these stimuli appearing alone.

If transferred to a behavioral setting, inverse effectiveness relates to the observation that combinations of weak stimuli elicit larger IFE than do combinations of more intense ones (Colonius & Diederich 2002) which is in line with the results reported by Corneil et al. (2002) and Diederich & Colonius (2004) from RTP studies.

Bernstein et al. (1970, 1973) reported that the amount of IFE (measured as the difference between RT on the target stimulus presented alone, RT_V , and RT on a bimodal condition, RT_{VA} , increased with the intensity of an auditory non-target and decreased with the intensity of the visual target. These observations make sense, because a stronger non-target can provide more response activation in order to elicit a response, whereas a stronger target would diminish the influence of the non-target because it already provides enough response activation itself (cf. Colonius & Diederich 2004).

However, Bernstein and colleagues observed inverse effectiveness only when the intensity of the visual target was decreased but not when the intensity of the auditory non-target was decreased. Note that the intensities of the target and the non-target were not varied concurrently, but each one separately. In general, increasing the intensity of one stimulus while leaving the intensity of the other stimulus constant decreases the relative effectiveness of the latter. In the FAP, the non-target enhances response activation and the more intense the stimulus the larger the possible enhancement. Therefore, the IFE is larger due to a more intense non-target, *i.e.*, direct effectiveness. On the other hand, a less intense target increases the relative effectiveness of the non-target. Therefore, the IFE is larger due to the less intense

target, *i.e.*, inverse effectiveness.

Furthermore, stimulus effectiveness has been varied only in terms of stimulus intensity. The duration of a stimulus could be yet another determinant of the amount of IFE, because single stimulus RT studies have revealed that the perceived intensity of a stimulus can be changed by varying its duration and an increase in stimulus duration results in a decrease in manual RT, although this holds only for very short durations (for a summary see Luce 1986). Analogous to inverse effectiveness obtained from stimulus combination with weak intensities, an inverse relation between stimulus duration and the relative amount of IFE is plausible if we assume that it is beneficial for the organism to detect minimal stimuli.

The purpose of the present study is twofold. First, to explore whether inverse effectiveness can be observed in a visual-tactile FAP task when the effectiveness of the visual target and the tactile target is varied concurrently, *i.e.*, when both stimuli are either of shorter or longer duration. Second, to probe the influence of different stimulus durations on the amount of multisensory integration in a FAP task. We hypothesize that the presentation of shorter stimuli results in a larger IFE compared with the presentation of longer stimuli, *i.e.*, inverse effectiveness of stimulus duration.

2.2 Experiment 1

2.2.1 Methods and Apparatus

Participants

Four students (ages 26–27 years) served as paid voluntary participants. All of them had normal vision and right-eye dominance. All participants were informed about the procedure and gave their informed consent prior to their inclusion in the study. The experiment was conducted in accordance with the ethical standards described in the 1964 Declaration of Helsinki.

Apparatus

The experiment took place in a completely darkened room. Stimuli were presented on a black desk (180×130×75 cm) with a recess for a chair (referred to as *vertex*). A chin rest was attached to the front edge of the vertex.

Two red light-emitting diodes (LEDs, \varnothing 5 mm, 2.4 V, 0.4 cd/m²) served as visual targets. They were placed 10° to the left or right of a central fixation point marked by a third LED (fixation LED). All LEDs were arranged on a circle with a diameter of 35 cm centered on the base of the chin rest.

Tactile stimuli were vibrations (0.6 V, 50 Hz, 1–2 mm amplitude) generated by two silenced oscillation exciters (Brüel & Kjær Mini-Shaker type 4810, referred to as *shaker*) positioned 10° to the left or right of the fixation LED, 5 cm away from the visual target in the respective hemifield. They were applied to the center of the palms through wooden balls that covered the upper ends of the shakers.

The durations of both the visual target and the tactile non-target varied. The presented stimulus durations (ms) were (30/50), (60/100), (90/150), (120/200) and (500/500), where the first entry refers to the duration of the visual stimulus.³ For convenience, the different duration combinations are labeled D1, D2, D3, D4, and D5, respectively.

Both LEDs and shakers were controlled by a PC multifunction card.

Eye movements were recorded by the infra-red light video system EyeLink (Sensomotoric Instruments, referred to as *EyeLink*) with a spatial resolution of 0.01° and a 250 Hz sampling rate.

Procedure

The participants were advised to respond only to the onset of the visual target and to pay no attention to the tactile non-target (*i.e.*, FAP). They were instructed to

³Stimulus durations were determined in unimodal pilot studies such that: (a) RT decreased monotonically with stimulus duration, (b) difference in RT was as large as possible between the shortest and longest durations, and (c) steps between the lower 4 durations were equal in size for each modality. The longest condition (500 ms) was added to allow comparisons to earlier experiments.

gaze as quickly and as accurately as possible to the position of the visual stimulus.

Every experimental session contained two blocks of 250 trials each, separated by a break of 10 min. At the beginning of each session the participants were customized with the experimental setup (adjusting the headband of EyeLink and the position and height of the chair). Before each block, EyeLink was calibrated. For each participant the first two sessions (1,000 trials) were omitted from the data analysis (training). Afterwards a total of 1,500 trials was recorded in three regular sessions on three separate days.

Each trial started with the onset of the fixation LED. After a variable fixation time (800-1,500 ms) the fixation LED was turned off and stimulus presentation started immediately (no gap). On *unimodal* trials the visual target appeared alone, on *bimodal* trials it was accompanied by a tactile non-target that was presented simultaneously. On *bimodal-ipsilateral* trials both visual and tactile stimuli were presented in the same hemifield; on *bimodal-contralateral* trials they were presented in different hemifields. After the offset of the stimuli a pause (2-2.45 s, depending on the duration of the stimuli presented before) in complete darkness followed, before the beginning of the next trial was indicated by the onset of the fixation LED.

No stimulus onset asynchronies were introduced, *i.e.*, the presentation of both LED and shaker started simultaneously.

For each of the 30 different conditions (five durations \times three lateralities (ipsi, contra, LED-only) \times two hemifields) 50 trials were recorded. All conditions appeared randomized within a block.

Data recording and preprocessing

The EyeLink was connected to a PC used for data storage and preprocessing. Saccades were detected trial by trial using the criteria of velocity ($22^\circ s^{-1}$) and acceleration ($8,000^\circ s^{-2}$).

Because no systematic differences between responses from the right and the left hemifield were observed, saccadic RT (SRT) were combined across hemifields of stimulus presentation (left or right).

For data analysis two factors were defined: *laterality* with levels LED-only, bimodal-ipsilateral, and bimodal-contralateral and *duration* with levels D1–D5.

2.2.2 Results

Trials with SRT smaller than 80 ms were excluded from evaluation (anticipation errors; cf. Becker 1991), as well as ones with SRTs longer than 500 ms (misses). Also trials with saccadic amplitudes smaller than 5° or greater than 20° (50° for participant SL⁴) and/or direction errors (responses to the hemifield opposite to the visual target) were excluded from further evaluation (a total of less than 4.6 % errors). The rate of direction errors was the same for both lateralities.⁵

Mean SRTs with standard errors (black vertical lines) are plotted as a function of stimulus duration and laterality in Fig. 2.1 (upper plot). If presentation of the tactile non-target had not affected SRT, the three curves would be identical. Furthermore, if it made no difference whether the non-target was presented ipsilateral or contralateral to the target the dotted and the slash-dotted curves would be the same. If duration of the stimulus had no effect on SRT, all curves would be aligned horizontally.

A two-way ANOVA on SRT with factors *laterality* and *duration* revealed significant results for *laterality* ($F_{(2,6)} = 21.85, P \leq 0.001$), indicating that the spatial alignment of stimuli had an effect on the amount of IFE. Post hoc tests (Tukey's) showed that a tactile non-target presented ipsilateral to the visual target resulted in significantly shorter SRTs than presenting it contralateral ($P \leq 0.001$) or presenting the visual stimulus alone ($P \leq 0.001$). However, contralateral presentation of the tactile non-target did not lead to significant differences compared with presentation of the visual stimulus alone.

Moreover, the factor *duration* was statistically significant ($F_{(4,12)} = 4.42, P \leq 0.001$), suggesting that different durations of stimuli led to different amounts of IFE.

⁴Participant SL performed saccades that systematically overshoot beyond the position of the target. This resulted in a higher mean amplitude and a larger scatter.

⁵This strongly suggests that participants responded according to the FAP, *i.e.*, ignored the non-target. Otherwise, many more direction errors should have been observed on contralateral trials than on ipsilateral.

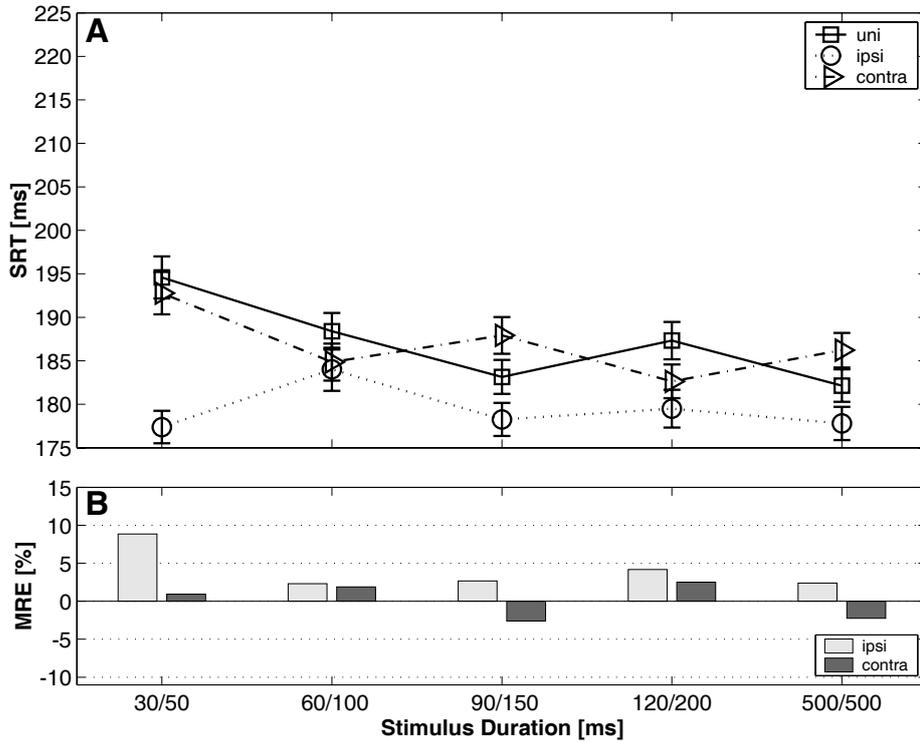


Figure 2.1: Experiment 1: Multisensory integration in a focused attention paradigm as a function of stimulus duration (ms) and laterality (across all participants). In each pair of durations the first entry refers to the duration of the visual stimulus. (A) Mean saccadic reaction time [SRT, (ms), vertical black lines indicate standard error] decreases with increasing duration for unimodal and bimodal-contralateral presentation, but stays about the same for bimodal-ipsilateral. Unimodal mean SRTs are marked with squares, bimodal-ipsilateral SRTs with circles, and bimodal-contralateral SRTs with triangles. (B) Multisensory response enhancement (MRE, [%]) is largest for ipsilateral presentation of the stimuli with the shortest durations.

Mean SRT on the shortest duration D1 was significantly longer than mean SRT on durations of D3, D4, and D5, as revealed by post hoc tests ($P \leq 0.05$, $P \leq 0.05$, and $P \leq 0.01$, respectively).

A significant interaction of *laterality* and *duration* ($F_{(8,24)} = 3.06$, $P \leq 0.002$) indicates that the influence of different durations varies at different conditions of laterality. Further exploration of this interaction using post hoc tests (Tukey's) revealed that significant differences were obtained only for level LED-only between durations D1 and D3, and durations D1 and D5 (both $P \leq 0.01$). Durations did not differ significantly for levels bimodal ipsilateral and bimodal contralateral.

In order to check whether or not inverse effectiveness was obtained, *multisensory*

response enhancement (MRE) was calculated from observed mean SRT. Diederich & Colonius (2004) measured IFE in terms of MRE, a descriptive measure that relates mean RT in the unimodal condition (LED-only; *i.e.*, RT_V), to that in the bimodal condition, (*i.e.*, RT_{VT}):

$$\text{MRE} = \frac{RT_V - RT_{VT}}{RT_V} \times 100. \quad (2.1)$$

Note that this measure is applied to means tested for significant differences in an ANOVA (see above). Therefore, no further tests on this measure are necessary.

Bar plots in Fig. 2.1 (bottom row) show that IFE in terms of MRE was largest for ipsilateral presentation of the shortest stimuli D1.

Results of Experiment 1 show that the relative amount of IFE varied with different stimulus durations and that the shortest duration led to the largest MRE, supporting our hypothesis. However, for durations above condition D2 the amount of IFE does not appear to be influenced by stimulus duration.

Since neurophysiological studies have shown that response enhancement is greatest when stimulus intensity is near threshold (Perrault et al. 2005; Stein & Meredith 1993), evidence for inverse effectiveness might still be more convincing if stimulus duration is even shorter. This idea led us to our second experiment.

2.3 Experiment 2

The results of Experiment 1 revealed inverse effectiveness of stimulus duration only for durations below 60 ms (respectively, 100 ms for tactile stimuli). To probe whether inverse effectiveness holds for even shorter stimulus durations, a second experiment utilizing shorter stimulus durations was conducted.

2.3.1 Methods and Apparatus

Participants

Five students (ages 20–23 years) served as paid voluntary participants. All of them had normal vision; three had right-eye and two had left-eye dominance. All participants were informed about the procedure and gave their informed consent prior to their inclusion in the study. The experiment was conducted in accordance with the ethical standards described in the 1964 Declaration of Helsinki.

Apparatus

The experimental setup was the same as in Experiment 1, except for the following differences.

Eye movements were recorded by the infra-red light video system EyeLink II (SR Research, referred to as *EyeLink II*) with a spatial resolution of 0.01° and a 500 Hz sampling rate.

To allow for the presentation of vibrations shorter than 50 ms, the shakers were operated with a higher frequency of 250 Hz. Because the shaker frequently produced noise when operated at this frequency, a background noise (white noise, 54 dB) was presented during the whole experiment, except for 10 ms of silence at the beginning of each trial for technical reasons.

Again, durations of both the visual target and the tactile non-target varied simultaneously. For both modalities the presented stimulus durations (ms) were (10, 20, 30, 40, 50, and 100 ms). Visual-tactile stimuli were presented simultaneously.

Procedure

A recording session took approximately 50 min and consisted of 2 blocks with 216 trials each, separated by a break of 10 min. After a training session (432 trials, excluded from analysis) four regular sessions were recorded on four separate days (1,728 trials for each participant).

Procedure and instructions to participants were the same as in Experiment 1.

For each of the 36 different conditions (six durations \times three lateralities (ipsilateral, contralateral, LED-only) \times two hemifields) 48 trials were recorded. All conditions appeared randomized within a block.

Data recording and preprocessing

SRTs were collapsed across hemifields because no significant differences were observed between responses from the right and left hemifield.

For data analysis two factors were defined, *laterality* with levels LED-only, bimodal-ipsilateral and bimodal-contralateral, and *duration* with levels 10, 20, 30, 40, 50, and 100 ms.

2.3.2 Results

Data were tested for RT (min 80 ms; max 500 ms), amplitude [min: 5° ; max: 20° (50° for participant IM⁶)], and/or direction errors as in Experiment 1.

Data from participant SI were excluded from further analysis because the proportion of correct responses was below 0.4 for all lateralities. The fact that a large majority of incorrect responses failed to meet the minimum amplitude of 5° , with about half of them also being indicated as direction errors, suggests that SI had problems detecting very short stimuli.

For the four remaining participants a total of 15% of all trials were indicated as incorrect and therefore excluded from further analyses.

Mean SRTs with standard errors (black vertical lines) are drawn as a function of stimulus duration and laterality in the upper plot of Fig. 2.2. Again, the three curves would be identical if the presentation of the tactile non-target had no influence on SRT. If the laterality of the tactile non-target had not affected SRT, the dotted and the slash-dotted curves would be the same. If there had been no effect of stimulus duration on SRT, all curves would be aligned horizontally.

A two-way ANOVA on SRT with factors *laterality* and *duration* revealed sig-

⁶Participant IM performed saccades that systematically overshoot beyond the position of the target. This resulted in a higher mean amplitude and a larger scatter.

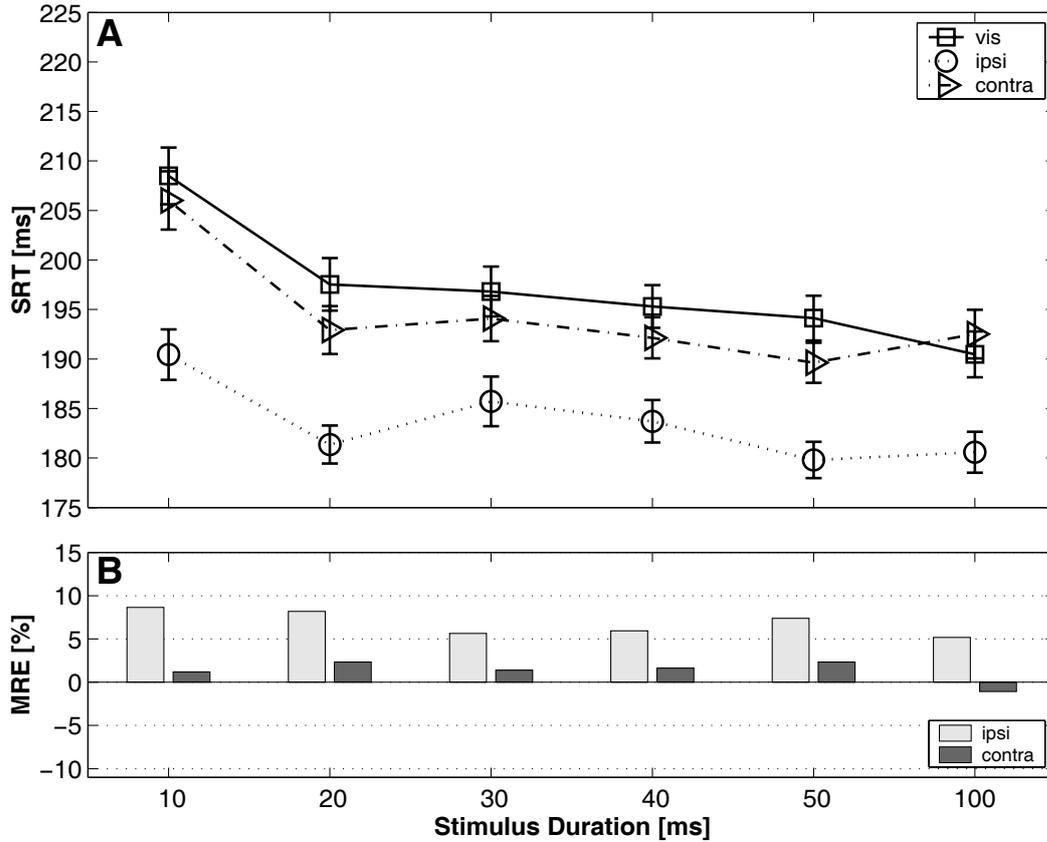


Figure 2.2: Experiment 2: Multisensory integration as a function of stimulus duration [ms] and laterality (across all participants). (A) Mean saccadic reaction time [SRT, (ms), vertical black lines indicate standard error] decreases with increasing duration for unimodal, bimodal-ipsilateral, and bimodal-contralateral presentation. Unimodal visual mean SRTs are marked with squares, bimodal-ipsilateral SRTs with circles, and bimodal-contralateral SRTs with triangles. (B) Multisensory response enhancement [MRE, (%)] is largest for ipsilateral presentation of the stimuli with the shortest durations and slightly decreases with increasing stimulus duration. MRE on contralateral presentation does not appear to be influenced by stimulus duration.

nificant results for *laterality* ($F_{(2,6)} = 56.2, P \leq 0.001$), indicating that the spatial alignment of stimuli had an effect on the amount of IFE. The presentation of a tactile non-target ipsilateral to the visual target led to a significant decrease in SRT compared to SRT on presentation of the visual target alone (Tukey's, $P \leq 0.001$) or the contralateral presentation of the tactile non-target ($P \leq 0.001$). SRT on contralateral presentation of the tactile non-target did not differ significantly from the SRT on presentation of the visual target alone.

Significant results were also obtained for the factor *duration* ($F_{(5,15)} = 14.1, P \leq 0.001$). Post hoc tests (Tukey's) showed that SRT on the shortest stimuli (10 ms) was

significantly longer than SRT on all remaining stimulus durations (all $P \leq 0.001$).

The interaction between *laterality* and *duration* was not significant (Tukey's, $F_{(10,30)} = 0.71, P = 0.714, \text{n.s.}$).

As shown in Fig. 2.2 (bottom row), ipsilateral MRE is largest for the shortest stimulus duration (10 ms) and decreases slightly with increasing stimulus duration. For contralateral MRE no such trend is present.

Very short durations yielded reliable response enhancement, confirming the findings of Experiment 1, and showing that inverse effectiveness of stimulus duration can also be demonstrated for durations below 60 ms and that ipsilateral MRE does decrease while stimulus duration increases. Ipsilateral MRE observed in Experiment 2 is about the same magnitude as in the shortest condition D1 of Experiment 1.

2.4 Discussion

The influence of stimulus intensity on the amount of RT facilitation in behavioral studies has been shown in studies utilizing both RTP and FAP. For RTP, larger relative amounts of IFE were observed when weaker stimuli were presented (Cornel et al. 2002; Diederich & Colonius 2004), a finding that is supported by neurophysiological studies on higher mammals (Perrault et al. 2005; Stein & Meredith 1993). For FAP, the pattern of results is more complicated (Bernstein et al. 1970, 1973).

The present study investigated whether duration might have a similar effect on IFE. In two experiments, utilizing an FAP with visual target and tactile non-target, we showed that this is indeed the case. In both experiments, the relative amount of IFE on ipsilateral conditions measured in terms of MRE was largest for the shortest stimulus durations and moreover, in Experiment 2, MRE on ipsilateral conditions decreased with increasing stimulus duration.

For Experiment 1, a significant interaction between laterality and duration was observed. A full exploration of the interaction revealed that stimulus duration had no effect on bimodal SRT, neither ipsilateral nor contralateral. Inverse effectiveness of stimulus duration, however, can still be postulated for the current set of data

because inverse effectiveness, as defined by Stein & Meredith (1993), refers to an inverse relation between stimulus effectiveness and the relative amount of intersensory facilitation rather than between stimulus effectiveness and bimodal RT. Nevertheless, the relation between IFE and stimulus intensity, *i.e.*, inverse effectiveness, could be explained by a ceiling effect. However, since SRT on bimodal-ipsilateral stimulation was always shorter than unimodal SRT, this possibility is ruled out for this set of data. For Experiment 2 no such interaction could be observed, ruling out the ceiling effect for our results and therefore strongly supporting the hypothesis of inverse effectiveness for very short durations. Thus, the results of Experiment 2 are consistent with the notion that inverse effectiveness can be observed in particular when stimuli are very weak, as suggested by Perrault et al. (2005).

For Experiment 1, condition (500/500) was similar to a condition reported in Diederich et al. (2003) (the 10° conditions in all three experiments). Whereas the magnitude of MRE on ipsilateral conditions observed in our experiment was comparable to the one reported there, the magnitude of MRE on contralateral conditions was not. While Diederich et al. observed MREs between 1.8 and 4.3% for the contralateral 10° conditions, an MRE of -2.3 % for the contralateral D5 condition in Experiment 1, *i.e.*, response depression, was found. We do not have an explanation for these differing results at the moment. However, it should be noted that response enhancement and depression do not constitute distinct phenomena but rather they are two instances along the same continuum of multisensory interaction (Meredith & Stein 1986).

For both experiments, MRE on contralateral conditions seems not to be affected by stimulus duration. This suggests that inverse effectiveness of stimuli with short durations only applies if target and non-target are in close proximity. Since this effect is relatively small it might disappear with larger distances between target and non-target, as some behavioral visual-auditory studies suggest that the amount of IFE also gets smaller with increasing disparity between stimuli (*e.g.*, Arndt & Colonius 2003; Frens et al. 1995). However, there are also studies that did not find a systematic relationship between spatial disparity and MRE (Colonius & Diederich

2004; Diederich & Colonius 2007; Diederich et al. 2003). Furthermore, neurophysiological studies on cat's DLSC neurons did not reveal a systematic relationship between stimulus disparity and the magnitude of MRE on the level of individual neurons (cf. Kadunce et al. 2003).

The evidence from DLSC neurons is especially relevant to the investigation of SRTs, because this region has been reported not only to be an important site of multisensory integration, but also to be involved in saccade generation (Munoz & Wurtz 1995). Recently, two mechanisms facilitating shorter SRT due to multisensory integration have been revealed: the reduction of neuronal response onset latencies and an increase of premotor activity of DLSC neurons (Bell et al. 2005).

3 On quantifying multisensory interaction effects in reaction time and detection rate¹

Abstract

The assessment of multisensory interaction effects critically depends on the investigator's choice of measuring a participant's performance. For an audio–visual simple detection task, we demonstrate that performance quantification from either reaction times (RTs) or detection rates (DRs) alone possibly misses out important information. We consider two principally different ways of integrating RT and DR to quantify overall performance. One measure is based on an arithmetical combination of RT and DR, and the other utilizes sequential sampling models. 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

3.1 Introduction

Response times (RT) and response frequencies are commonly obtained data in the behavioral sciences, and mean RT and detection rates (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 tradeoffs is that

¹A major revision of this chapter has been invited for possible publication in Psychological Research as: Rach, S., Diederich, A., & Colonius, H. On quantifying multisensory interaction effects in reaction time and detection rate.

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 & Marks (2008) in a multisensory identification task. Multisensory tasks involve stimuli from two or more modalities, and a common finding is that mean RT to crossmodal stimuli (*e.g.*, light and tone) tends to be faster than to unimodal stimuli (*e.g.*, light). “True” multisensory interaction has occurred if the speeded reaction cannot be accounted for by other mechanism, such as statistical facilitation or response bias. Arieh & Marks (2008) evaluated the amount of multisensory interaction in a speeded identification of color with and without the presence of noise using speed-accuracy tradeoff function (SATF; 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 multisensory interaction. 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 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. Finally, we demonstrate the evaluation of differences in overall performance via non–parametrical and parametrical bootstrapping.

3.2 Quantification of multisensory interaction effects

3.2.1 Measures of Multisensory Speed-up 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, reaction time (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. This so-called intersensory facilitation effect (IFE, Hershenson 1962) proved robust in various behavioral replications (*e.g.*, Diederich & Colonius, 1987, 2004; Gielen et al., 1983; Hershenson, 1962; Miller, 1982). The amount of IFE is modulated by the spatial and temporal alignment of the stimuli and it decreases with increasing temporal separation (“temporal rule”, Bernstein et al., 1969, 1970; Diederich & Colonius, 1987, 2004; Giray & Ulrich, 1993; Hershenson, 1962; Miller, 1986; Morrell, 1968), as well as with increasing spatial separation (“spatial rule”, Amlôt et al., 2003; Bernstein & Edelstein, 1971; Colonius & Diederich, 2004; Frens et al., 1995; Harrington & Peck, 1998; Walker et al., 1997). Moreover, the amount of IFE is larger when stimuli are less intense (“principle of inverse effectiveness”, Corneil et al., 2002; Diederich & Colonius, 2004; Rach & Diederich, 2006; see Holmes, 2007, 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 et al., 2005; Frassinetti et al., 2002), a task-irrelevant visual stimulus can enhance auditory perception (Lovelace et al., 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.

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 & Colonius (2004) calculated multisensory response enhancement (MRE) by

$$\text{MRE} = \frac{\min(RT_V, RT_A) - RT_{VA}}{\min(RT_V, RT_A)} \times 100. \quad (3.1)$$

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:

$$\text{MDE} = \frac{DR_{VA} - \max(DR_V, DR_A)}{\max(DR_V, DR_A)} \times 100, \quad (3.2)$$

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) is 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.

3.2.2 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). 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 et al., 2007; Shore et al., 2006; Spence et al., 2001).

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

$$RT^* = \frac{RT}{DR} \quad (3.3)$$

Substituting RT by RT^* in Eq. 3.1 yields a measure reflecting relative multisensory enhancement with respect to IES:

$$MRE^* = \frac{\min(RT_V^*, RT_A^*) - RT_{VA}^*}{\min(RT_V^*, RT_A^*)} \times 100 \quad (3.4)$$

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.

3.2.3 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. 3.1 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 Section 3.4.2 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, 1992a, 1995; Schwarz, 1994). Returning to our audio–visual simple RT example experiment, let us assume that

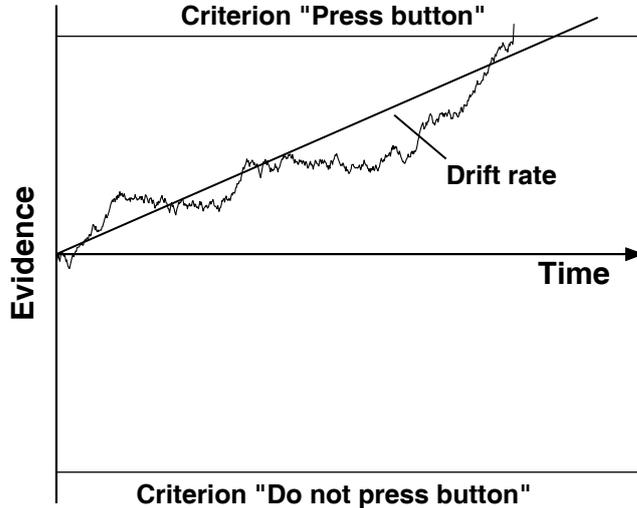


Figure 3.1: Sample path (trajectory) and drift rate (solid line) of a sequential sampling model. The sampling path (trajectory) represents sequential accumulation of evidence until a decision boundary (either “Press button” or “Do not press button”) is reached. The average across multiple sampling paths is represented by the drift rate (solid line).

the drift rates for every unimodal and bimodal condition have been estimated. We define multisensory performance enhancement (MPE) by

$$\text{MPE} = \frac{\delta_{VA} - \max(\delta_V, \delta_A)}{\max(\delta_V, \delta_A)} \times 100, \quad (3.5)$$

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 prop-

erties 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 both measures, MRE* and MPE, we conducted a detection experiment with visual, auditory, and audio–visual stimuli of different intensities, followed by some cross-validation studies. 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.

3.3 Methods

3.3.1 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.

3.3.2 Apparatus

The study was conducted in a completely darkened and sound reflection attenuated room. Participants were seated in front of a black desk ($180 \times 130 \times 75$ 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, \varnothing 5 mm) placed 20° to the left or right of a central fixation point marked by a third LED (fixation LED) 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.

3.3.3 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, \dots, V_8 . For a complete list of intensities utilized in this experiment see Table 3.1. 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 dB A, henceforth indicated by A_1, \dots, A_8 . On bimodal trials, visual and auditory stimuli of corresponding intensity levels were paired (*i.e.*, $\{V_1, A_1\}, \dots, \{V_8, A_8\}$). Whenever it is referred to stimulus intensities without specifying the modality, the labels I_1, \dots, I_8 will be used with I_1 representing the weakest and I_8 the strongest intensity.

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

Table 3.1: Intensities utilized given as luminance [mcd] of visual and loudness [dBA] of auditory stimuli.

3.3.4 Procedure

Participants were instructed to respond to every stimulus regardless of its modality by lifting their toe as quickly as possible (redundant target paradigm).

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). See Fig. 3.2 for a schematic depiction of the procedure.

A recording session of 45 minutes included two blocks of 192 trials each separated by a 5 minutes. Recording 32 trials for each of 24 conditions with each of 6 participants resulted in a total of 4,608 trials.

3.4 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

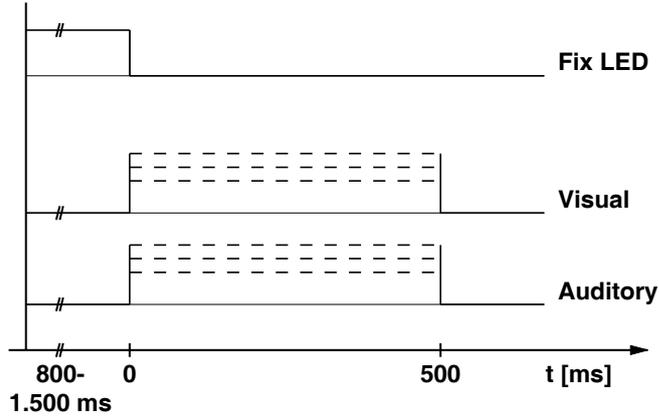


Figure 3.2: Time course of a trial. Onset of the fixation LED defined the start of a trial. Simultaneously to the offset of the fixation LED, either the visual, the auditory, or both were presented. Stimuli from both sensory modalities were presented at 8 different intensity levels.

Intensity	Reaction Time [ms] (Standard error)						Detection Rate		
	V		A		VA		V	A	VA
I_1	530	(23.5)	465	(22.1)	466	(17.1)	0.54	0.56	0.76
I_2	489	(17.0)	428	(15.6)	430	(11.4)	0.71	0.77	0.89
I_3	490	(14.6)	402	(13.4)	411	(10.1)	0.80	0.81	0.89
I_4	496	(13.7)	415	(12.6)	397	(10.7)	0.79	0.89	0.94
I_5	480	(12.1)	400	(11.4)	376	(8.5)	0.84	0.93	0.96
I_6	470	(12.9)	384	(11.3)	357	(8.7)	0.84	0.95	0.96
I_7	424	(9.7)	354	(9.5)	342	(8.5)	0.93	0.98	0.97
I_8	420	(8.1)	341	(9.0)	335	(9.1)	0.97	0.97	0.97

Table 3.2: Mean reaction time (with standard error) and detection rate as a function of stimulus intensity and sensory modality. Stimulus intensity is ordered ascending from I_1 to I_8 .

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 Tab. 3.2.

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. 3.3 (panels A and C). MRE and MDE as function of stimulus intensity are given in Fig. 3.3 (panels B and D).

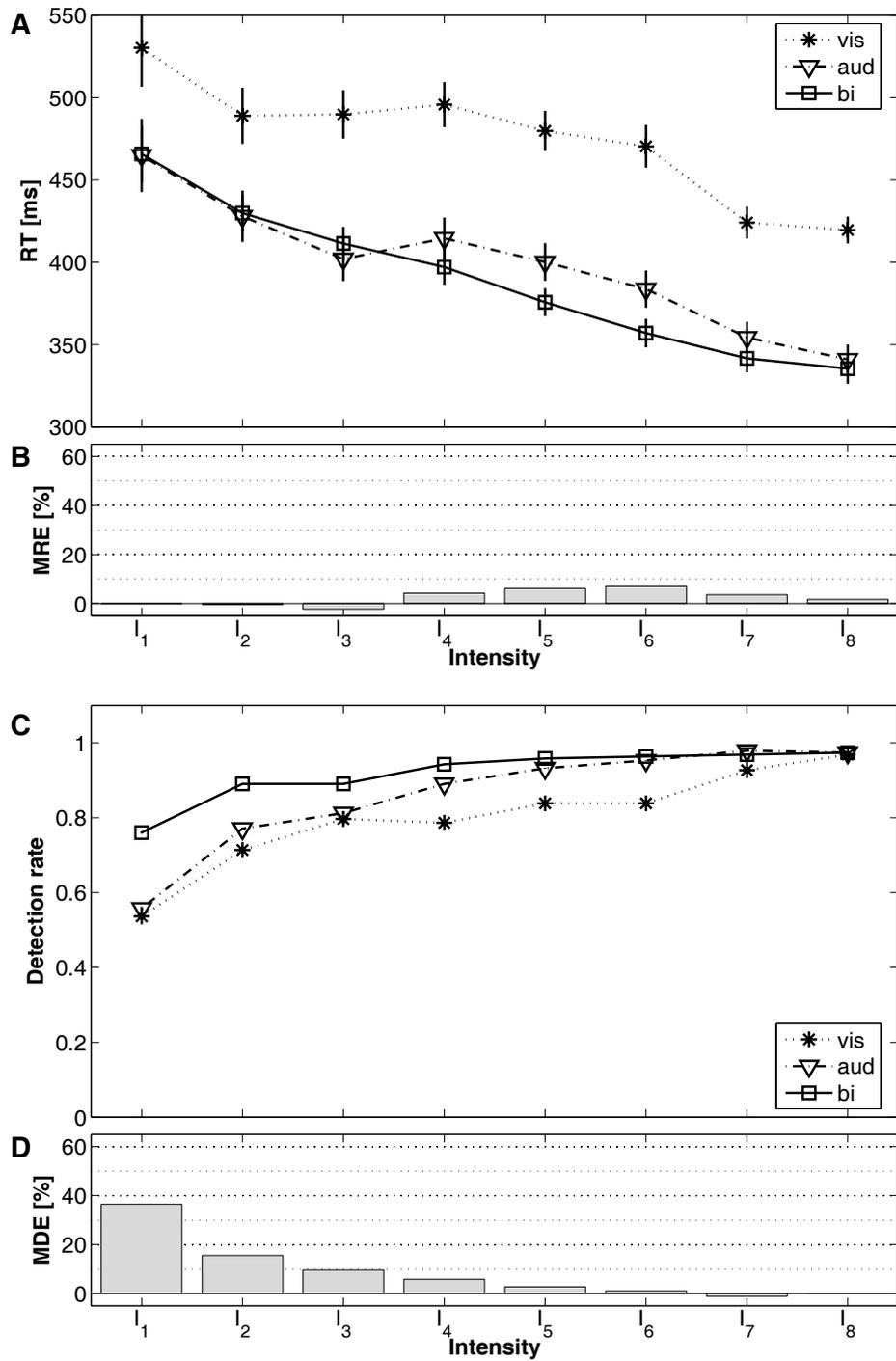


Figure 3.3: Performance in a detection experiment quantified from reaction time and detection rate. (a) Mean foot reaction time [ms] with standard errors (black vertical lines) as a function of stimulus intensity and modality. (b) Multisensory response enhancement [%] as a function of stimulus intensity. (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 is ordered ascending from I_1 to I_8 .

For intensity levels I_4 to 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 to I_5 , indicating enhanced detectability on bimodal conditions. MDE was about zero for intensity levels I_6 to 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.

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.

3.4.1 MRE* calculated from Inverse Efficiency Scores

RT* and MRE* calculated from RT and DR according to Eq. 3.3 and 3.4 are given in Fig. 3.4. For all three modalities RT* decreased with increasing stimulus intensity, indicating that the overall performance increased with stimulus intensity. Compared to RT, this decrease was much steeper for RT* (note the differently scaled y-axes in panel (a) of Fig. 3.3 and 3.4). For intensity conditions I_1 to 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. 3.4, 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 principle of inverse effectiveness.

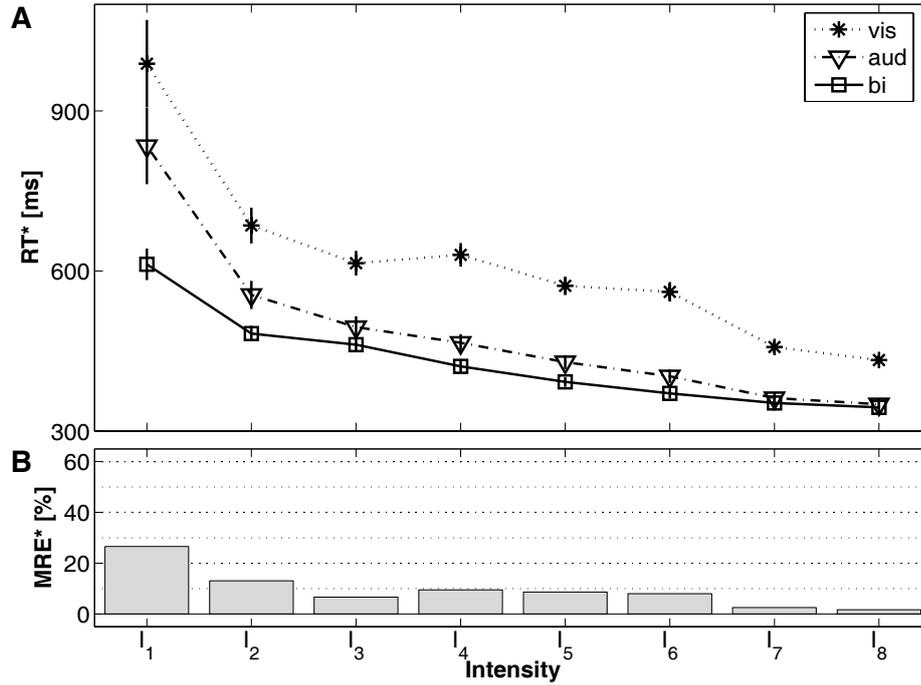


Figure 3.4: Overall performance in a detection experiment quantified from inverse efficiency scores. (a) Mean inverse efficiency scores, RT^* , as a function of stimulus intensity and modality. (b) Relative response enhancement (MRE^*) calculated from inverse efficiency scores, RT^* , as a function of stimulus intensity. Stimulus intensity is ordered ascending from I_1 to I_8 .

3.4.2 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 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. Consequently, only RT for alternative A would be examined, ignoring RT for alternative B, because it is unobservable in this case.

Finally, a residual time T_r was estimated. It can be interpreted as base time, that is, the time taken by non-decisional cognitive and motor processes that are not influenced by experimental manipulations under interest, and therefore remains constant across all experimental conditions (Townsend & Honey, 2007).

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

$$\chi^2 = \sum_i \left(\left(\frac{\overline{RT}_i - \widehat{RT}_i}{\sigma_{\overline{RT}_i}} \right)^2 + \left(\frac{\overline{DR}_i - \widehat{DR}_i}{\sigma_{\overline{DR}_i}} \right)^2 \right) \quad (3.6)$$

where \overline{RT}_i and \widehat{RT}_i signify observed and predicted RT; \overline{DR}_i and \widehat{DR}_i signify ob-

served and predicted DR; and $\sigma_{\overline{RT}_i}$ and $\sigma_{\overline{DR}_i}$ signify standard deviation of RT and DR. ²Estimates for drift rates are given in Fig. 3.5, 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 were $\theta = 18$ and $T_r = 187.8$ ms. MPE calculated according to Eq. 3.5 is displayed in Fig. 3.5, panel (b). Note that MPE for the lowest intensity level, I_1 where MPE

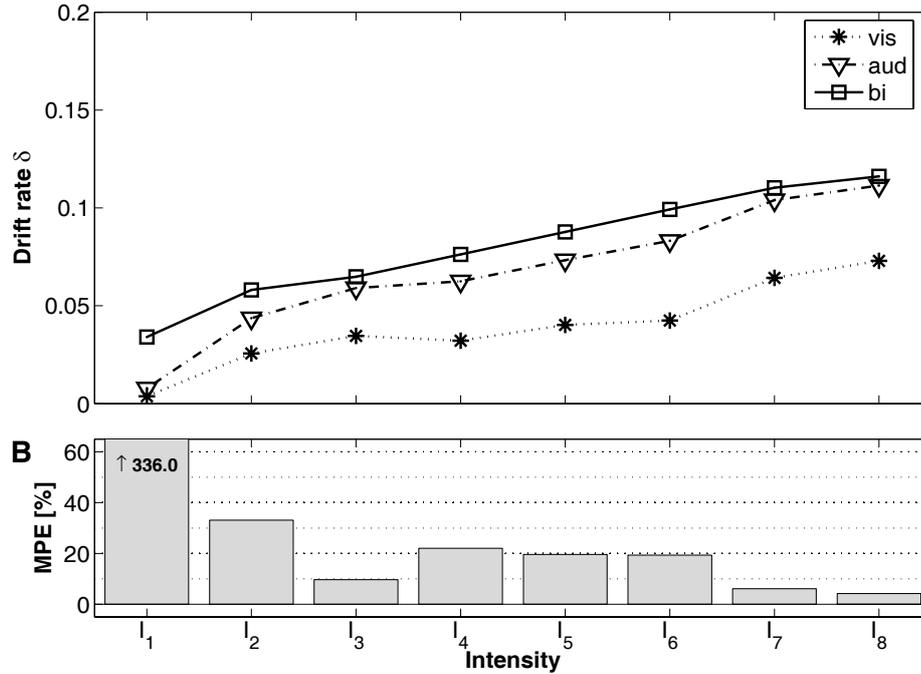


Figure 3.5: Overall performance in a detection task quantified from parameters of a sequential sampling model. (a) Drift rates δ as a function of stimulus intensity and modality. (b) Multisensory performance enhancement (MPE) as a function of stimulus intensity. Stimulus intensity is ordered ascending from I_1 to I_8 .

indicated a response enhancement of 336.0 %, the bar did not fit in the plot, which

²A statistical test for the goodness of fit is given by $\chi^2(22) = 30.8, p = 0.10$ (26 parameters were fitted to 48 data points). The observed χ^2 value was 0.31 indicating an excellent fit. 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.

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 principle of inverse effectiveness 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 to 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.

3.4.3 Evaluation of differences in overall performance utilizing bootstrapping procedures

To allow for the examination of variability and the calculation of confidence intervals, the estimated parameters were subjected to a 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. In the parametric bootstrap, the resampling method utilizes a generating model, that is in our case, the sequential sampling model that was fitted to the original data set. Thus, we used the parameters estimated to generate large numbers of simulated data sets, to which, in turn, the sequential sampling model was fitted again. Both procedures resulted in distributions: the non-parametric bootstrap results in a distribution of simulated RTs and a distribution of simulated DRs; and the the parametric bootstrap results in a distribution for each parameter estimated from generated data sets. 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 ± 1 standard deviation of a Gaussian (Wichmann & Hill, 2001). As it was pointed out by Wichmann & Hill (2001), the parametric bootstrap relies heavily on one’s belief in the model used to generate the simulated data sets, but, however, the same belief already led us to fitting this model to the original data set and therefore no additional assumptions are to be made here.

For the examination of MRE^* , we generated 1000 non-parametric bootstrap samples from the original data set; 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, 1000 parametric bootstrap samples were generated from the parameters estimated from the original data (see Appendix for technical details). Sequential sampling models (Wiener processes) were fitted to the 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. 3.6.

Intensity	Percentiles			
	MRE^*		MPE	
	16th	84th	16th	84th
I_1	19.12	33.08	185.55	577.48
I_2	7.62	18.82	18.45	58.30
I_3	0.64	11.85	-2.05	23.61
I_4	5.69	14.64	11.82	40.85
I_5	4.60	12.68	11.09	39.09
I_6	4.01	12.11	11.38	41.00
I_7	-1.13	6.46	-2.90	21.40
I_8	-2.47	5.25	-4.10	18.75

Table 3.3: Confidence intervals, CI_{68} , calculated by the bootstrap percentile method as a function of stimulus intensity. For both 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 MRE^* . Nevertheless, both indices indicated that multisensory response enhancement for intensity condition I_1 significantly exceeded that for any other condition, as in both

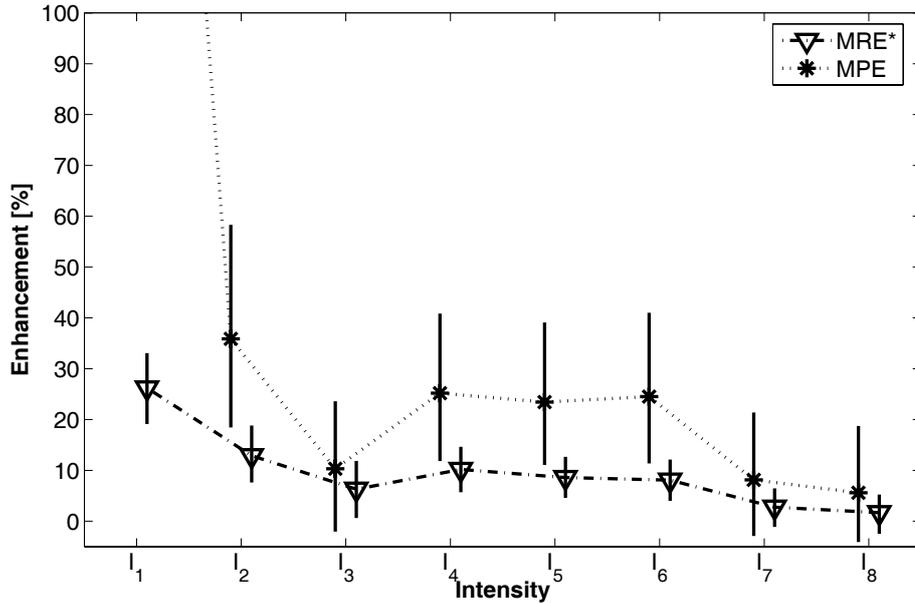


Figure 3.6: Magnitude and variation of multisensory indices as a function of stimulus intensity. Open triangles display multisensory response enhancement, MRE^* , calculated from inverse efficiency scores with confidence intervals, CI_{68} , obtained in a non-parametric bootstrap procedure (black vertical lines). Asterisks display multisensory performance enhancement (MPE) with confidence intervals, CI_{68} , obtained in a parametric bootstrap procedure (black vertical lines). Stimulus intensity is ordered ascending from I_1 to I_8 . MPE for the lowest intensity level, I_1 , was 336.0 %, with the corresponding CI_{68} spanning from 185.6 to 577.5.

cases the confidence intervals for both conditions did not overlap (see Tab. 3.3). MRE^* was significantly larger than zero for intensity conditions I_1 to I_6 , indicating an enhanced performance due to bimodal stimulation for the respective stimulus intensities. For MPE all intensity conditions were significantly larger than zero.

3.5 Summary and Discussion

Multisensory interaction is known to show up in both RT and DR. We developed two indices for the quantification of IFE combining RT and DR in 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.

An experiment was conducted to investigating the influence of stimuli, close to detection threshold, on multisensory interaction and to test the proposed indices. Visual, auditory and audio–visual stimuli of 8 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. 3.3, panel b), this was only the case for intensity levels I_4 to I_7 . For intensity levels I_1 to 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: for RT an “irreducible minimum” is known to exist which puts a lower boundary on RT facilitation (Woodworth & Schlosberg, 1956). For the lowest intensity levels an explanation might not be as easy since the inverse effectiveness rule (Meredith & Stein, 1986; Perrault et al., 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. 3.3, panel d), that, for intensity levels I_1 to I_5 multisensory enhancement actually did manifest in DR. The fact that no enhancement was observed for DR on intensity levels I_6 to 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 missed out important 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 were similar for both indices. Although the bootstrap procedures were different (non-parametric MRE* vs. parametric for MPE) and the width of the resulting confidence intervals differed substantially (small for MRE* vs. large for MPE), both indices suggest almost identical conclusions. (1) Bimodal stimulation led to enhanced performance in most intensity conditions. This observation is in line with a large body of research (see Section 3.2, for an overview). (2) There was significantly more enhancement for the lowest intensity condition, compared to all remaining intensities. This finding presents strong evidence for the rule of inverse effectiveness (*cf.*, Corneil et al., 2002; Diederich & Colonius, 2004; Rach & Diederich, 2006).

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. Still both measures would lead to the same statistical conclusions when evaluated in terms of bootstrap confidence intervals and, in particular, these conclusions are in line with earlier research.

Independently from the results obtained in this particular experiment, both measures can be compared in terms of validity, reliability, and practicability. Validity describes to which extent a measure really assesses the underlying theoretical con-

struct. As mentioned before, a theoretical foundation for calculating IES (by dividing RT by DR) is lacking and 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 common in the examination of simple reaction tasks as well as choice reaction tasks, and parallels between model mechanisms and properties of nervous processes have been identified (Diederich, 1995, 1997, 2008; Diederich & Busemeyer, 2003; Luce, 1986; Ratcliff & Smith, 2004).

To be reliable, a measure is to remain nearly invariant when calculated repeatedly on the same data. For MRE* this is clearly the case as only elementary arithmetics are involved. However, this is not necessarily the case for MPE as it involves parameter estimation 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. 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, when it comes to parametrical bootstrap, computation time multiplies and becomes an issue indeed because a model needs to be fitted to each of the generated data sets.

To conclude, with two different approaches we demonstrated that speed and accuracy can be integrated in a single measure quantifying multisensory enhancement, and how the calculation of confidence intervals can be attempted with bootstrap pro-

cedures. Whenever computational power and time are available, we advocate the use of MPE, as it outperforms MRE* in terms of validity and reliability. Nevertheless, when computational power is an issue or a quick first impression is intended, MRE* can be beneficial. Although both approaches led to good results, it appears indicated to further examine the properties of the presented indices in simulation studies, which, however, are not within in the scope of this paper.

3.6 Appendix

3.6.1 Parameter estimation

For the original data set and each of 1000 resampled parametric bootstrap data sets, 24 drift rates $\delta_i \in \Delta$, one boundary separation θ , 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 $0 < \delta_i < 1$, 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 , T_r , and Δ by minimizing χ^2 . Store χ^2 .

Step 2: Hold constant θ and T_r , while estimating each $\delta_i \in \Delta$ individually by minimizing χ^2 .³

Step 3: Estimate parameters θ , 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.

3.6.2 Bootstrap procedures

Two different bootstrap procedures were applied; a non-parametric bootstrap to examine MRE*, and a parametric bootstrap to examine MPE.

³Note that, for constant θ and T_r , all $\delta_i \in \Delta$ are independent from each other, thus can be estimated separately.

For the non-parametric bootstrap, 1000 random samples were generated as follows. For each of 24 experimental conditions, the original data set consisted of RTs from 192 trials, with zeros in any trial without a valid response. From this 1 by 192 vector, 192 values were drawn randomly with replacement, resulting in a simulated 1 by 192 data vector, from which mean RT (mean of non-zero elements) and DR (number of non-zero elements divided by 192) were calculated. This procedure resulted in $24 \times 1,000$ pairs of mean RT and DR.

For the parametric bootstrap procedure, the parameter θ , T_r , and Δ estimated from data recorded in Experiment 1 were utilized. The diffusion rate σ^2 was not estimated, but set to 1 (Diederich, 1995; Diederich & Busemeyer, 2003, 2006). A transition matrix was calculated for each $\delta_i \in \Delta$. Then, for each of 1000 simulations, 192 trajectories were calculated. Trajectories reaching the upper boundary θ were considered as “stimulus detected” and detection probabilities and first passage times for this alternative were calculated accordingly. Finally, RTs were calculated by adding the residual time T_r to the first passage time. This procedure also resulted in 24×1000 pairs of mean RT and DR, to which sequential sampling models were fitted afterwards. Finally, for each condition a distribution of 1000 drift rates was obtained.

3.6.3 MatLab script for the simple Wiener process

```
function [prob_s, prob_d, et_s, et_d] = wiener(para)
% Simple Wiener process with two boundaries; no bias
% Diederich & Busemeyer (2005)

drift = para(4:end);
alpha = 1;
tau = 1;
sigma = 1;
delta = alpha*sqrt(sigma)*sqrt(tau);
m = 2*round(100*para(1)/delta)+1;
x = -(m-1)/2:(m-1)/2;

g1 = 0; % no gamma
```

```

nDrift = length(drift); % number of drift parameters

% pre-allocating memory
prob_s(np,1) = zeros;
prob_d(np,1) = zeros;
et_a(np,1) = zeros;
et_b(np,1) = zeros;
mux1(1,35) = zeros;
tm1(m,m) = zeros;

n = m-2;
id = eye(n,n);

e1 = -0.999;
e2 = 0.999;

sqrttausigma=sqrt(tau)/sigma;

for k = 1:nDrift % for every drift rate
    mux1 = drift(k)-g1*x;

    mux1(mux1 >= e2) = e2;
    mux1(mux1 <= e1) = e1;

    px1 = 1/(2*alpha)*(1-mux1*sqrttausigma);
    qx1 = 1/(2*alpha)*(1+mux1*sqrttausigma);
    pq = 1-(1/alpha)-0*x;

    % transition matrix
    tm1(1,1) = 1; tm1(m,m)=1;
    i = (2:m-1)';
    tm1((i-2)*m+i) = px1(i);
    tm1((i-1)*m+i) = pq(i);
    tm1((i)*m+i) = qx1(i);

    % R-vector and Q-matrix
    rb = tm1(2:(m-1),1); % two boundaries
    ra = tm1(2:(m-1),m);
    q = tm1(2:(m-1),2:(m-1));
    invq = inv(id-q);

    % start vector
    z = (zeros(n,1));

    % probabilities
    zinvq = z'*invq;
    probb = zinvq*rb;

```

```
proba = zinvq*ra;

prob_d(k) = sum(proba);
prob_s(k) = sum(probb);

%expected value conditioned on the alternative
eta = zinvq*invq;
etb = eta*rb;
eta = eta*ra;
et_a(k) = sum(eta);
et_b(k) = sum(etb);
end

et_d = et_a./prob_d;
et_s = et_b./prob_s;
```

4 Comparing redundant target and focused attention paradigm: A descriptive analysis

Abstract

Presenting an auditory stimulus in addition to a visual stimulus often results in decreased reaction times and increased detection rates compared to presenting the visual stimulus alone, no matter whether or not participants are instructed to attend to the auditory stimulus. Here we investigated the influence of different experimental instructions (“respond to any stimulus” vs. “respond only to visual stimuli”) and the presented modalities (visual and audio–visual stimuli vs. visual, audio–visual, and auditory stimuli) on overall performance and the amount of multisensory interaction. Reaction time (RT) and detection rate (DR) were recorded as dependent measures. Overall performance was quantified in terms of inverse efficiency scores (IES), a descriptive measure calculated from RT and DR, and the amount of multisensory interaction was quantified by calculating multisensory response enhancement from IES.

Keywords redundant target paradigm · focused attention paradigm · multisensory interaction · reaction time · detection rate · inverse efficiency scores · multisensory response enhancement

4.1 Introduction

Reaction time (RT) to a visual stimuli tends to be faster if an auditory stimulus is presented in temporal and spatial proximity. This observation has been termed intersensory facilitation effect (IFE, *e.g.*, Hershenson, 1962) or redundant signals effect (*e.g.*, Miller, 1982) and has often been replicated (*e.g.*, Bernstein et al., 1969; Colonius & Arndt, 2001; Diederich & Colonius, 2008; Diederich et al., 2003; Gielen et al., 1983; Steenken et al., 2008). The magnitude of this effect is dependent on the spatial and temporal alignment of the stimuli. It decreases with increasing temporal separation (*temporal rule*, Bernstein et al., 1969, 1970; Diederich & Colonius, 1987, 2004; Giray & Ulrich, 1993; Hershenson, 1962; Miller, 1986; Morrell, 1968), as well as with increasing spatial separation (*spatial rule*, Amlôt et al., 2003; Bernstein & Edelstein, 1971; Colonius & Diederich, 2004; Diederich & Colonius, 2007; Frens

Instruction	Stimuli presented	
	V, A, VA	V, VA
respond to all stimuli	RTP	RTP _{¬A}
respond to visual stimulus	FAP	FAP _{¬A}

Table 4.1: Four variations of the experimental paradigm as presented in the experiment. The upper case letters represent the sensory modality of the stimuli: V - visual; A - auditory; VA - audio-visual

et al., 1995; Harrington & Peck, 1998; Walker et al., 1997). Moreover, the amount of IFE is larger when stimuli are less intense (*principle of inverse effectiveness*, Corneil et al., 2002; Diederich & Colonius, 2004; Rach & Diederich, 2006; but see Holmes, 2007, 2008, for a critical view).

A further behavioral measure that can be modulated by multisensory interaction is the detection rate (DR) of stimuli. For instance, a task-irrelevant auditory stimulus can modulate visual perception (Bolognini et al., 2005; Frassinetti et al., 2002), a task-irrelevant visual stimulus can enhance auditory perception (Lovelace et al., 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).

RT and DR have often been dealt with separately in studies on multisensory interaction, *i.e.*, either RT or DR has been studied. Here we present an analysis approach that intends to integrate both measures. Near to the detection threshold both RT and DR are known to change simultaneously with stimulus intensity (cf. Luce, 1986) and therefore it appears reasonably to assume that multisensory interaction effects can also be observed in both measures (cf. Chapters 3 and 5). To integrate both measures into one measure of overall performance, inverse efficiency scores (IES, Townsend & Ashby, 1983), a descriptive measure computed from RT and DR, can be used to quantify performance and to assess multisensory interaction effects (cf. Chapter 3).

Two experimental paradigms have been proposed to investigate multisensory interaction. In the *focused attention paradigm* (FAP), 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 *redundant target paradigm* (RTP) participants are instructed to respond to any stimulus regardless of its modality. In the following, we refer to the example of a bimodal (visual/auditory) experiment where stimuli are presented in either three conditions (randomized across trials): unimodal visual (V), unimodal auditory (A), and audio–visual (VA), or in only two conditions (randomized across trials): unimodal visual (V) and audio–visual (VA). If we present both versions either in an RTP task (“respond to any stimulus”) or an FAP task (“respond only to visual stimuli”), four experimental tasks can be distinguished: RTP, FAP, RTP_{¬A}, and FAP_{¬A}, where the subscript $\neg A$ indicates that unimodal auditory trials were not presented (see Table 4.1). A number of differences between these four tasks have been reported previously.

Studies comparing performance in RTP and FAP tasks revealed that RTs recorded in RTP tasks are faster than those in FAP tasks (Giray & Ulrich, 1993; Morrell, 1968). Morrell (1968) recorded RTs in an audio–visual FAP experiment (visual target, auditory non–target; auditory–only catch-trials included) and compared them to RTs sampled in an audio–visual RTP study (Morrell, 1967). RTs recorded in the RTP task were about 50 ms faster than RTs in the FAP task. Giray & Ulrich (1993) compared RTs in an audio–visual RTP task to RTs in an audio–visual FAP task (visual target, auditory non–target; auditory–only catch trials included) and found that RTs in the RTP task were faster than those in the FAP task (visual stimuli: 47 ms difference, audio–visual stimuli: 79 ms difference). However, both studies compared results from different participants recorded in separated experiments.

The difference in RT as reported by Morrell (1968) and Giray & Ulrich (1993) can be attributed to a relevant difference in the task: RTP tasks involve only the detection of stimuli, *i.e.*, the decision whether a stimulus is present or not. For FAP tasks it is not only necessary to decide whether or not a stimulus is present, but in addition, whether or not it belongs to the target modality (*i.e.*, a discrimination task) and it appears quite reasonable to assume, that this additional process increases RT. Another important contributor to the difference in RT between FAP

and RTP tasks is statistical facilitation. Raab (1962) suggested that the RT reduction observed in crossmodal conditions can be explained by a statistical effect which does not involve interaction between the senses: if detection latencies are interpreted as random variables, the time to detect the first of several redundant signals is faster, on average, than the detection time for any single signal. Statistical facilitation can occur in redundant target tasks, where participants respond to any stimulus they perceive, but it cannot occur in focused attention tasks, since participants are instructed to respond only to the target modality.

For focused attention tasks, it has been shown that the presentation of catch trials can prolong RT (Gielen et al., 1983). Catch trials are conditions where a non-target is presented alone and therefore participants should withhold their response. In terms of our example, responses in an FAP task (*i.e.*, a focused attention task with catch trials) are expected to be slower, on average, than responses in an FAP_{-A}. One possible reason for this slowdown might be the effort of the participant to avoid responses on catch trials, *i.e.*, to avoid wrong responses. Similar effects are reported for RTP tasks: responses to unimodal visual stimuli in RTP_{-A} tasks (where only visual and audio-visual trials are presented) are expected to be faster, on average, than RTP tasks (utilizing visual, auditory, and audio-visual trials), because of the *modality shift effect* (MSE, Gondan et al., 2004; Spence & Driver, 1997; Spence et al., 2001). The MSE describes the observation that RT to a stimulus is increased, on average, when the modality of the preceding stimulus was different. When only visual and audio-visual stimuli are presented, no switch in modality occurs, because a visual stimulus is presented in every trial. If visual, auditory, and audio-visual trials are presented, a modality switch occurs whenever a unimodal visual trial is followed by a unimodal auditory trial and vice versa. RTs to unimodal stimuli can be increased up to 15 ms due to the MSE (Gondan et al., 2004). Note that the effect of catch trials and the MSE differ in one important aspect: MSE is expected to prolong only RTs in unimodal conditions, whereas catch trials are expected to prolong RTs in both the unimodal and the crossmodal conditions.

The aim of this study is to investigate these predictions by presenting different

RTP and FAP variations blocked within one single experiment. An experiment was conducted to compare the effects of different instructions and stimulus configurations on RT, DR, and the relative amount of response enhancement.

4.2 Methods & Apparatus

Participants

Four students (ages 20–23 years) served as paid voluntary participants. All of them reported normal vision and hearing. 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 a completely darkened and sound reflection attenuated room. Participants were seated in front of a black desk ($180 \times 130 \times 75$ cm), with their heads fixed by a chin rest attached to the front edge of the desk.

Mounted on the desk, two red light emitting diodes (LED, \varnothing 5 mm) placed 20° to the left or right of a central fixation point marked by a third LED (fixation LED) were used to present 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 (white noise) were presented by two speakers (Canton Plus XS) placed horizontal to the participant's ear level 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 utilized tactile stimuli applied to the palms, ruling out manual responses.

Stimuli

Stimulus intensities and stimulus onset asynchronies were determined for each participant individually. In a simple response task (*i.e.*, RTP), foot RT (fRT) to unimodal visual and auditory stimuli of varying intensities were recorded. Stimulus duration was 500 ms for both the visual and the auditory stimuli. For each modality, two intensities were 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 between RT on the lower and on the higher intensity, (*i.e.*, $RT(v) - RT(V)$, respectively $RT(a) - RT(A)$ for the auditory modality) was between 70 and 80 ms (see Table 4.2 for the resulting intensities).

It has been reported that a stimulus onset asynchrony (SOA) equal to the difference in RT between two stimuli results in physiological simultaneity and therefore optimal interaction between the modalities (Hershenson, 1962; Hilgard, 1933). Thus, SOAs were determined for each participant individually by choosing

$$\tau = RT(V) - RT(A)$$

and two additional SOAs at $\tau - 35$ ms and $\tau + 35$ ms (see Table 4.2 for the resulting SOAs). Note that, due to the criteria for selecting stimulus intensities, this difference was about the same for the low intensity stimuli, *i.e.*, presenting stimuli with an SOA of τ should result in physiological simultaneity in the conditions VA and va.

Procedure

Four variations of two different experimental paradigms were utilized in this experiment. Participants were either instructed to respond to any stimulus they perceive (redundant target paradigm) or to respond only to the visual stimulus (focused attention paradigm). Within a block, either visual, auditory, and audio–visual stimuli were presented, or only visual and audio–visual stimuli were presented ($\neg A$). This

Participant	Visual [mCd]		Auditory [dB A]		SOA [ms]		
	v	V	a	A	$\tau - 35$	τ	$\tau + 35$
as	0.0045	0.0076	18.3	21.1	-45	-10	25
da	0.0087	0.0112	14.3	24.8	55	90	125
ig	0.0076	0.0106	15.3	20.2	100	135	170
sr	0.0059	0.0106	15.3	19.2	35	70	105
Mean	0.0067	0.0100	15.8	21.3	36	71	106

Table 4.2: Participant specific luminance [mCD] of visual stimuli, loudness [dBA] of auditory stimuli and stimulus onset asynchronies [ms] utilized in this experiment. Lower-case letters represent low-intense stimuli; upper-case letters represent high-intense stimuli.

resulted in four different paradigms (FAP, FAP_{-A}, RTP, and RTP_{-A}), which were presented randomized across experimental blocks (see Table 4.1 for a summary). 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.

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 the presentation of the auditory stimuli was shifted by an SOA¹. Except for two bimodal conditions for participant AS, the presentation of the auditory stimuli started simultaneously to or after the onset of the visual stimuli. 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.

In a recording session of one hour, a participant completed two blocks separated by a break of 10 minutes. Each participant recorded a total of 20 blocks on 10 different days. Prior to the main study, each participant completed 2 hours of training. Recording 40 trials for each of 120 conditions with each of 4 participants resulted in a total of 19,200 trials.

¹SOAs were determined individually for each participant in a pilot study. See section “Stimuli” for details.

Data recording and preprocessing

A PC connected to the EyeLink was used for data storage and data preprocessing.

Trials with fRTs faster than 80 ms were classified as anticipation errors (0.1 %) and therefore excluded from further analyses, as well as ones with fRTs longer than 1000 ms (misses, 1.3 %).

RTs from the right and the left hemifield showed no systematic difference and were therefore combined across hemifields of stimulus presentation (left or right).

4.3 Results

4.3.1 Detection rates

DRs recorded in this experiment are depicted in Fig. 4.1 (see also Table 4.3 in the appendix). The figure consist of four panels representing four different intensity conditions. The panels are arranged as follows: panels in the upper row depict conditions with visual stimuli of high intensity, and the panels in the bottom row depict conditions with low intense visual stimuli. Conditions with auditory stimuli of high intensity are depicted in the left column and conditions with low intense auditory stimuli are presented in the right column. For instance, the intensity condition vA is depicted in the lower left panel. Different experimental conditions are coded as follows: RTP_{-A} is represented by red squares, RTP by blue asterisks, FAP_{-A} by green triangles, and FAP by black circles. For all conditions, horizontal dotted lines represent responses to unimodal visual stimuli; dashed lines represent responses to audio-visual stimuli for three different SOA ($\tau - 35$, τ , $\tau + 35$).

When unimodal visual stimuli of high intensity were presented, DR was about the same for all paradigms (see panels in the upper row). DR on unimodal visual stimuli with low intensity was smaller for FAP and FAP_{-A} compared to the two redundant target variants (see panels in the lower row). For most of the conditions, audio-visual DR was higher than the DR in the unimodal visual condition, which replicates earlier reports of improved detection performance due to crossmodal stim-

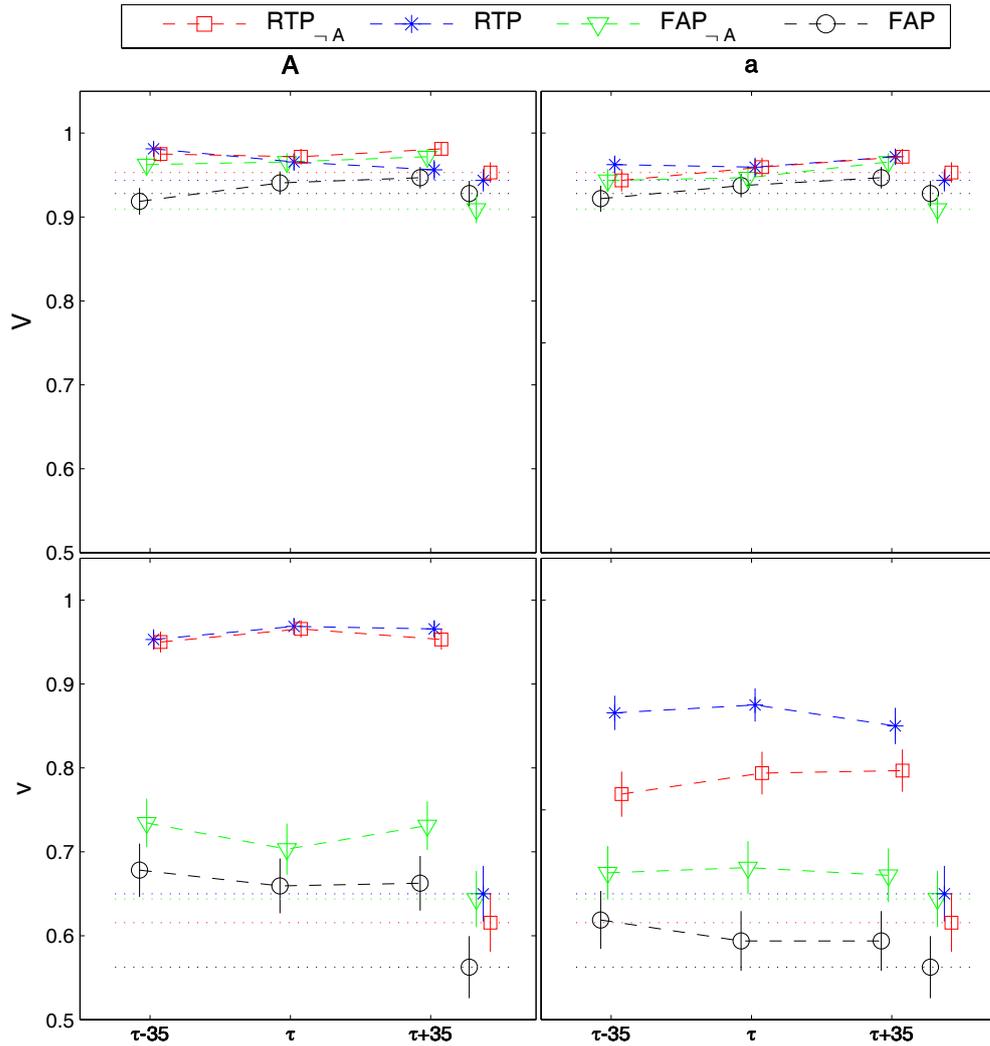


Figure 4.1: Detection rate (DR) with standard errors (vertical lines) as a function of stimulus onset asynchrony, stimulus intensity, and experimental paradigm. Panels representing four different intensity conditions are arranged as follows: intensity of the visual stimulus was high in the left column and low in the right column; intensity of the auditory stimulus was high in the upper row and low in the bottom row. Horizontal dotted lines represent responses to unimodal visual stimuli; dashed lines represent responses to audio-visual stimuli. Please see text for further details.

ulations (see Introduction). For the two redundant target tasks, RTP and RTP_{-A}, crossmodal DR was about the same for the conditions VA, Va, and vA and was considerably lower for the condition va. The first three intensity conditions all contain at least one high intensity stimulus, whereas the latter does contain only low intensity stimuli. This suggest that DR in the crossmodal conditions was mainly determined by the stimulus with the higher intensity, because DR was considerably

higher if at least one stimulus was of high intensity. For the focused attention tasks, FAP and FAP_{-A}, DR was about the same only for the intensity conditions VA and Va, and was considerably lower for vA and va. The first two conditions have in common that the intensity of the visual stimulus was high, whereas the intensity of the visual stimulus was low in the latter two conditions. This finding suggests that the response was mainly determined by the intensity of the visual stimulus, that is, the target stimulus.

4.3.2 Reaction times

RTs recorded in this experiment are depicted in Fig. 4.2 (see also Table 4.4 in the appendix), which is exactly arranged like Fig. 4.1 (see above for a detailed explanation). Different experimental conditions are coded as follows: RTP_{-A} is represented by red squares, RTP by blue asterisks, FAP_{-A} by green triangles, and FAP by black circles. For all conditions, horizontal dotted lines represent responses to unimodal visual stimuli; dashed lines represent responses to audio-visual stimuli for three different SOA ($\tau - 35$, τ , $\tau + 35$).

For most of the conditions, RT in the audio-visual condition was shorter than RT in the unimodal visual condition, which replicates earlier reports of RT facilitation in crossmodal conditions (see Introduction). For all intensity conditions, the fastest RTs were observed in condition RTP, followed by RTP_{-A}, FAP, and FAP_{-A} with the slowest RTs. Note that this resembles our reasoning concerning the differences between those experimental paradigms:

1. RTs from the redundant target type tasks, RTP_{-A}, and RTP, (paradigms where only stimulus detection is involved and statistical facilitation can occur) are faster than those from the focused attention type tasks, FAP_{-A} and FAP_{-A} (where stimulus discrimination is necessary and statistical facilitation can not occur);
2. RTs from RTP (a paradigm where MSE can occur) are slower than those from RTP_{-A} (where MSE should not occur);

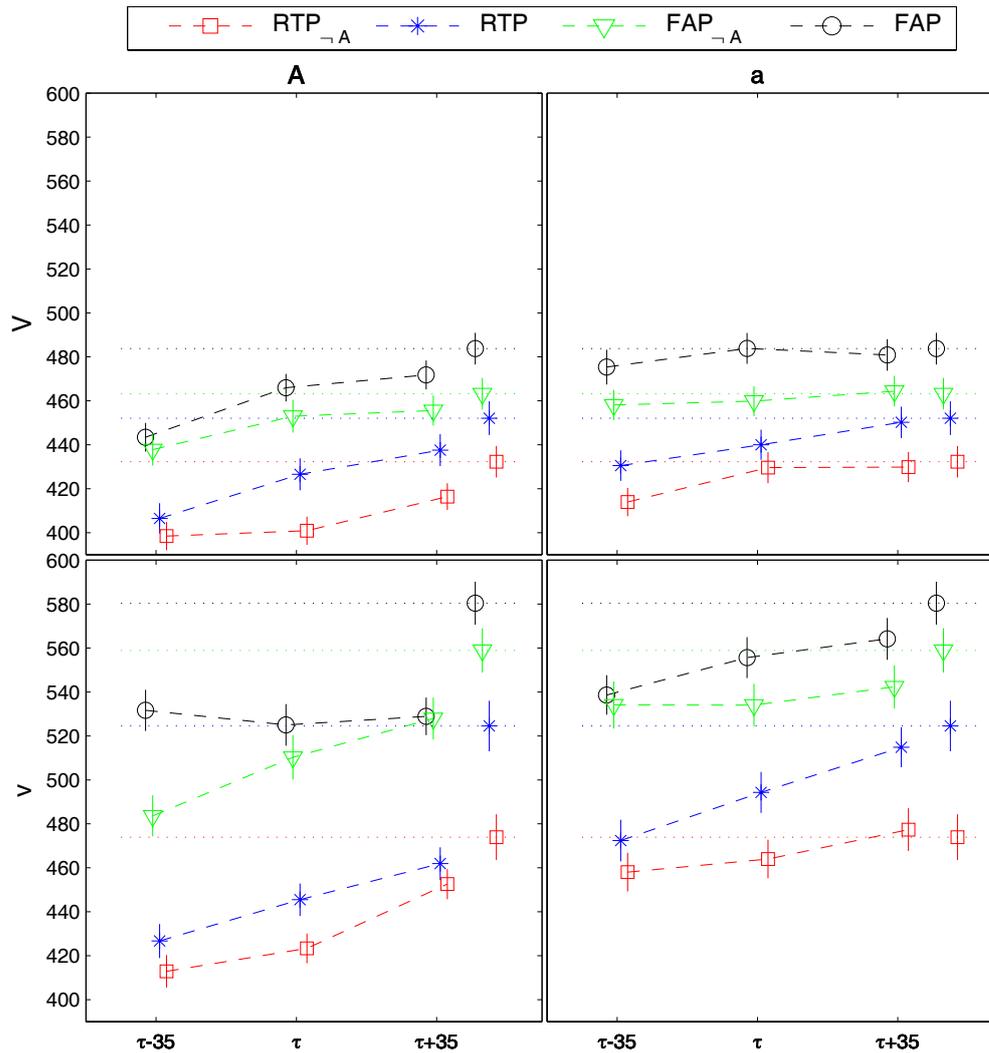


Figure 4.2: Reaction time (RT) [ms] with standard errors (vertical lines) as a function of stimulus onset asynchrony, stimulus intensity, and experimental paradigm. Panels representing four different intensity conditions are arranged as follows: intensity of the visual stimulus was high in the upper row and low in the bottom row; intensity of the auditory stimulus was high in the left column and low in the right column. Horizontal dotted lines represent responses to unimodal visual stimuli; dashed lines represent responses to audio-visual stimuli. Please see text for further details.

3. RTs from FAP (where catch trials are presented) are slower than those from FAP_{-A} (without catch trials).

A comparison of the intensity conditions revealed that the fastest RTs were observed for intensity condition VA (upper left panel), and subsequently slower RTs were observed for Va (upper right panel), vA (lower left panel), and va (lower right panel) where the slowest RTs were observed.

Systematical differences between the experimental paradigms can be found for RT. For the two redundant target tasks, RTP and RTP_{-A}, crossmodal RT was about the same for the intensity conditions VA, Va, and vA, and was considerably slower for the intensity condition va. This suggest that the response was determined by the stimulus with the higher intensity. For the focused attention tasks, FAP and FAP_{-A}, RT was about the same only for the intensity conditions VA and Va, and was considerably slower for vA and va. This finding suggests that the response was determined mainly by the intensity of the visual target stimulus.

4.3.3 Overall performance in terms of inverse efficiency scores

It has been shown that multisensory interaction effects can influence RT as well as DR (cf. Chapter 3). To quantify overall performance from both RT and DR, we calculated inverse efficiency scores (IES, Townsend & Ashby, 1983), a measure combining accuracy and RT (in a choice task) by dividing mean RT by the percentage of correct responses. With this correction, 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 et al., 2007; Shore et al., 2006; Spence et al., 2001).

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

$$\text{IES} = \frac{\text{RT}}{\text{DR}} \quad (4.1)$$

IESs calculated from RT and DR are summarized in Fig. 4.3. Different experimental conditions are coded as follows: RTP_{-A} is represented by red squares, RTP by blue asterisks, FAP_{-A} by green triangles, and FAP by black circles. For all conditions, horizontal dotted lines represent responses to unimodal visual stimuli; dashed lines represent responses to audio-visual stimuli for three different SOA ($\tau - 35$, τ , $\tau + 35$).

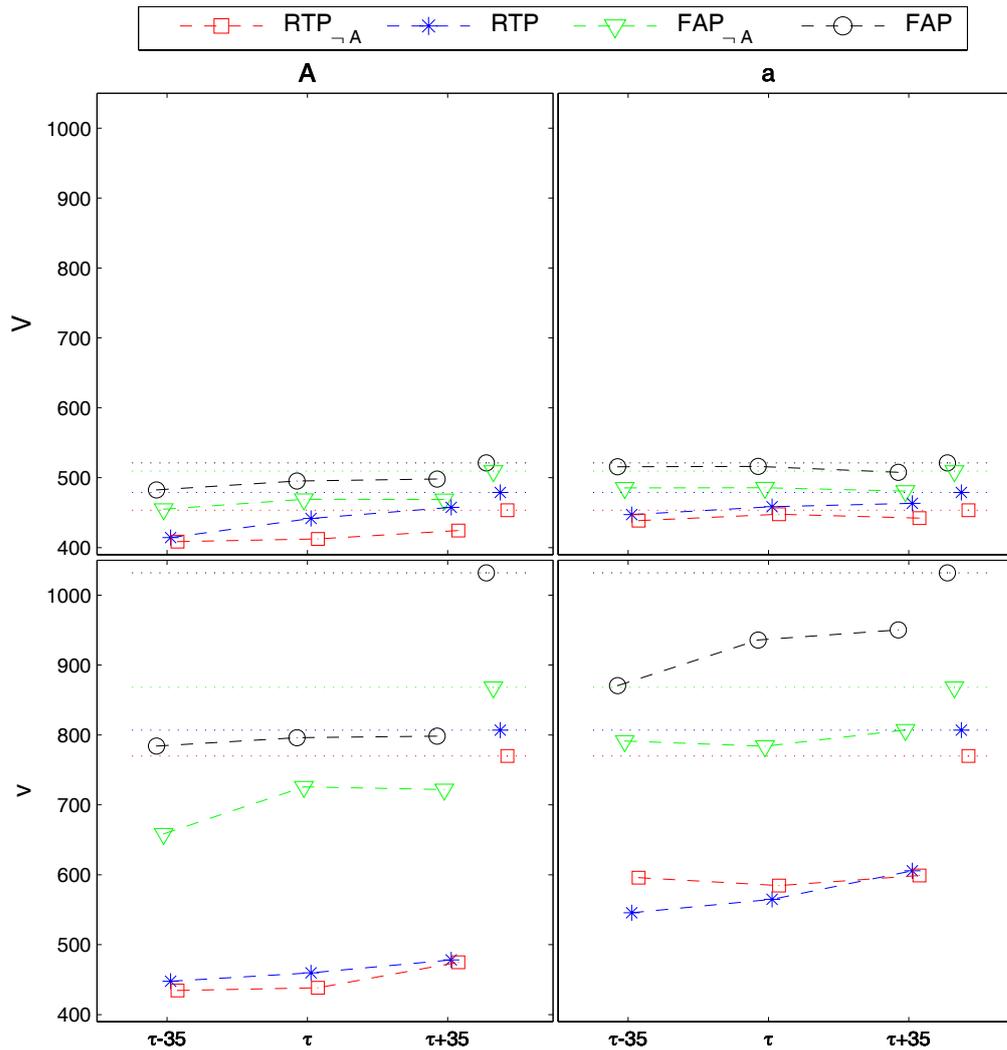


Figure 4.3: Inverse efficiency scores (IES) as a function of stimulus onset asynchrony, stimulus intensity, and experimental paradigm. Panels representing four different intensity conditions are arranged as follows: intensity of the visual stimulus was high in the upper row and low in the bottom row; intensity of the auditory stimulus was high in the left column and low in the right column. Horizontal dotted lines represent responses to unimodal visual stimuli; dashed lines represent responses to audio-visual stimuli. Please see text for further details.

By comparing IESs to the RTs given in Fig. 4.2, the effects of combining RT and DR can be evaluated. For all experimental paradigms, the difference between IES and RT is rather small for all conditions with visual stimuli of high intensity (VA and vA; compare the upper rows of Figs. 4.3 and 4.2), because DR was close to one for the respective conditions. When the visual stimulus was of low intensity and the auditory was of high intensity, the difference between IES and RT is only small

for the bimodal conditions of RTP and RTP_{-A}. For the unimodal visual conditions of RTP and RTP_{-A} and for all conditions of FAP and FAP_{-A}, IES is much higher than RT (compare the lower left panels of Figs. 4.3 and 4.2). When both stimuli are of weak intensity, IES is much larger than RT for all conditions (compare the lower right panels of Figs. 4.3 and 4.2). Albeit the differences between IES and RT, the patterns present in the RTs as described above, are also present in the IESs and are even accentuated:

1. IESs from the redundant target type tasks, RTP_{-A}, and RTP, are faster than both FAP_{-A} and FAP_{-A} (possibly caused by the difference in complexity between detection and discrimination tasks and statistical facilitation which can only occur in redundant target tasks);
2. unimodal IESs from RTP are slower than those from RTP_{-A}, whereas the difference between the bimodal IESs is smaller (possibly caused by MSE);
3. IESs from FAP are slower than those from FAP_{-A} (possibly caused by catch trials).

Furthermore, the afore-mentioned paradigm-specific influence of stimulus intensity on RT can also be found in the IES data. For RTP and RTP_{-A} the response appears to be determined by the stimulus with the higher intensity, whereas for FAP and FAP_{-A}, the response appears to be determined mainly by the intensity of the visual target stimulus.

4.3.4 Relative amount of response enhancement in terms of MRE*

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. RTs and DRs for unimodal auditory stimuli are only available for one paradigm (RTP), because they were either not presented (RTP_{-A} and FAP_{-A}) or participants were instructed not to ignore auditory stimuli

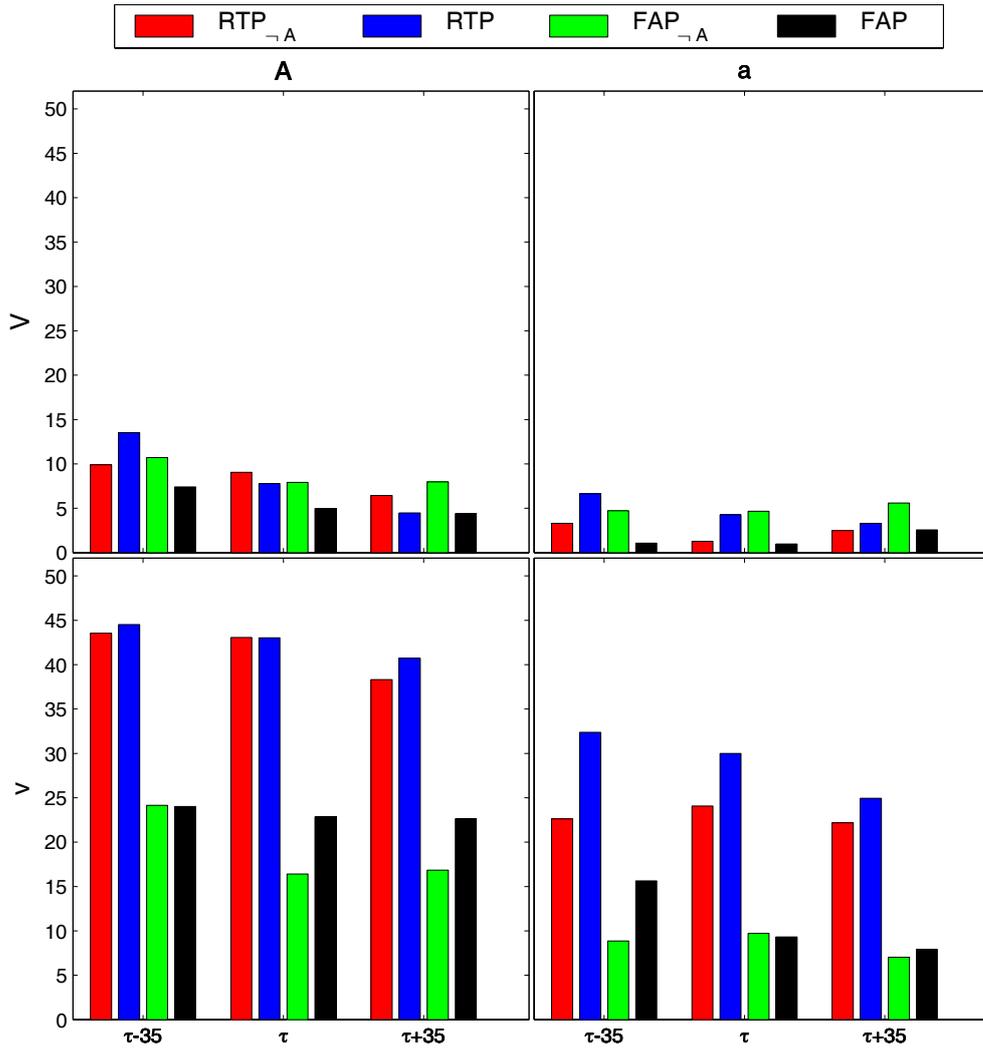


Figure 4.4: Multisensory response enhancement calculated from inverse efficiency scores (MRE*) [%] as a function of stimulus onset asynchrony, stimulus intensity, and experimental paradigm. Panels representing four different intensity conditions are arranged as follows: intensity of the visual stimulus was high in the upper row and low in the bottom row; intensity of the auditory stimulus was high in the left column and low in the right column. Please see text for further details.

(FAP). Therefore, only data from visual and audio–visual conditions were used to quantify the magnitude of interaction effects.

We used IES to calculate MRE*, a descriptive measure relating the response on visual conditions (IES_V) to that on bimodal audio–visual conditions (IES_{VA}) that is defined as

$$\text{MRE}^* = \frac{\text{IES}_V - \text{IES}_{VA}}{\text{IES}_V} \times 100 \quad (4.2)$$

where IES_V and IES_{VA} denote the transformed mean RTs in the visual and the audio-visual condition.

Note the following properties of MRE^* : (1) performance in the bimodal condition is compared to the performance in the visual conditions; (2) increasing performance in the bimodal condition (*i.e.*, a smaller IES) leads to larger positive values; (3) decreasing performance in the bimodal condition (*e.g.*, larger IES) is 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 performance in the visual condition.

MRE^* s are summarized in Fig. 4.4 (see also Table 4.5 in the appendix). Different experimental conditions are coded as follows: RTP_{-A} is represented by red bars, RTP by blue bars, FAP_{-A} by green bars, and FAP by black bars.

Interestingly, the relative amount of response enhancement is about the same for all experimental paradigms if only stimuli of high intensity are presented (see condition VA in the upper left panel), although RT differed considerably between those conditions.

The influence of stimulus intensity on the relative amount of response enhancement can be evaluated by comparing the upper left panel (intensity condition VA) and the lower right panel (intensity condition va) of Fig. 4.4. For both condition, the SOA was equal to the difference between the RTs to the unimodal stimuli (see Subsection Stimuli for details), possibly leading to physiological synchronicity and therefore optimal interaction in both conditions. MRE^* is larger for low stimulus intensities (va) compared high stimulus intensities (VA) for three out of four paradigms (RTP_{-A} , RTP, and FAP), providing strong evidence in favor of the principle of inverse effectiveness. For FAP_{-A} , MRE^* is about the same for high intensity and low intensity stimuli.

For the intensity conditions Va and vA, the difference in RT between the uni-

modal conditions was not compensated by adjusted SOA, making optimal interaction due to physiological synchronicity less likely. Nevertheless, the amounts of response enhancement in the intensity condition vA were larger than those in the intensity condition va (compare lower left and lower right panel). Since the overall stimulus intensity of the bimodal stimulus complex is lower for the va condition, the principle of inverse effectiveness would predict the largest amounts of response enhancement for this condition.

4.4 Summary & Discussion

An experiment was conducted to investigate the influence of different experimental paradigms on multisensory interaction. Participants were presented either with visual and audio–visual stimuli or with visual, audio–visual, and auditory stimuli, and were instructed either to respond to any stimulus they perceive or to respond only to visual stimuli (see Table 4.1 for the resulting conditions). RT and DR were recorded as dependent measures and were used to calculate IES, a descriptive measure of overall performance. The relative amount of response enhancement was quantified from IES by calculating MRE*, a descriptive measure that relates the performance in the unimodal visual condition to the performance on bimodal conditions.

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 (see Introduction). The amount of multisensory interaction measured in terms of MRE* 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 principle of inverse effectiveness. Contrary to the predictions of the principle of inverse effectiveness, MRE* for the presentation of two weak stimuli (va) was exceeded by MRE* in conditions

where a weak visual stimulus was paired with a high intense auditory stimulus (vA), although the overall intensity of the bimodal stimulus was higher in the latter condition. However, it is not clear whether this represents a multisensory effect or an artifact due to the calculation of MRE*, which did not take into account performance on unimodal auditory stimulation. When comparing conditions va and vA, the change of stimulus intensity is accounted for in the bimodal performance, but not in the unimodal (baseline) condition, although the intensity of one of the stimuli is changed (see Bernstein et al., 1970, p. 196, for a related discussion).

Across all intensity conditions, overall performance (measured in terms of IES) was highest if participants were presented with visual and audio–visual stimuli and were instructed to respond to any stimulus they perceive (RTP). Performance in the unimodal visual condition was slightly lower when unimodal auditory trials were presented additionally (RTP_{-A}), suggesting evidence for modality switch costs, which are expected to only influence the performance in unimodal conditions. If participants were instructed to respond only to the visual stimulus and to ignore any other stimulus (FAP and FAP_{-A}), overall performance was lower compared to conditions where they were instructed to respond to any stimulus. This decrease in performance might have been caused by an increase in task complexity: to respond only to the visual stimulus it was necessary to discriminate stimulus modalities. The lowest performance in terms of IES was observed when participants were instructed to respond only to visual stimuli and auditory (catch) trials were presented in addition to visual and audio–visual trials.

4.5 Appendix

Modality	SOA	Paradigm			
		RTP _{-A}	RTP	FAP _{-A}	
v		0.616 (0.035)	0.650 (0.033)	0.644 (0.033)	0.563 (0.037)
V		0.953 (0.012)	0.944 (0.013)	0.909 (0.017)	0.928 (0.015)
a		-	0.647 (0.033)	-	0.084 (0.053)
A		-	0.919 (0.016)	-	0.241 (0.049)
VA	$\tau - 35$	0.975 (0.009)	0.981 (0.008)	0.963 (0.011)	0.919 (0.016)
	τ	0.972 (0.009)	0.966 (0.01)	0.966 (0.010)	0.941 (0.014)
	$\tau + 35$	0.981 (0.008)	0.956 (0.012)	0.972 (0.009)	0.947 (0.013)
Va	$\tau - 35$	0.944 (0.013)	0.963 (0.011)	0.944 (0.013)	0.922 (0.016)
	τ	0.959 (0.011)	0.959 (0.011)	0.947 (0.013)	0.938 (0.014)
	$\tau + 35$	0.972 (0.009)	0.972 (0.009)	0.966 (0.010)	0.947 (0.013)
vA	$\tau - 35$	0.950 (0.013)	0.953 (0.012)	0.734 (0.029)	0.678 (0.032)
	τ	0.966 (0.010)	0.969 (0.010)	0.703 (0.030)	0.659 (0.033)
	$\tau + 35$	0.953 (0.012)	0.966 (0.010)	0.731 (0.029)	0.663 (0.032)
va	$\tau - 35$	0.769 (0.027)	0.866 (0.020)	0.675 (0.032)	0.619 (0.035)
	τ	0.794 (0.025)	0.875 (0.020)	0.681 (0.032)	0.594 (0.036)
	$\tau + 35$	0.797 (0.025)	0.850 (0.022)	0.672 (0.032)	0.594 (0.036)

Table 4.3: Detection rate as a function of intensity, stimulus onset asynchrony, and experimental paradigm.

Modality	SOA	Paradigm						
		RTP _{-A}		RTP		FAP _{-A}		FAP
v		473.94 (10.44)		524.54 (11.47)		558.92 (10.10)		580.40 (9.75)
V		432.28 (7.17)		452.10 (7.65)		463.26 (7.16)		483.74 (7.20)
a		-	-	450.14 (13.37)		-	-	441.30 (23.36)
A		-	-	397.09 (9.49)		-	-	400.62 (15.62)
VA	$\tau - 35$	398.45 (6.42)		406.50 (6.94)		437.77 (7.28)		443.40 (6.51)
	τ	400.88 (6.43)		426.55 (7.24)		453.01 (7.44)		465.96 (6.26)
	$\tau + 35$	416.41 (6.06)		437.57 (7.24)		455.59 (6.86)		471.77 (6.55)
Va	$\tau - 35$	413.93 (6.46)		430.53 (6.95)		458.12 (6.86)		475.35 (7.89)
	τ	429.60 (7.11)		440.00 (6.84)		459.83 (6.81)		483.84 (7.01)
	$\tau + 35$	429.83 (6.89)		450.23 (7.14)		464.39 (6.98)		480.86 (7.15)
vA	$\tau - 35$	412.90 (7.38)		426.67 (7.72)		483.64 (9.35)		531.66 (9.36)
	τ	423.31 (6.73)		445.47 (7.36)		510.25 (10.20)		525.01 (9.46)
	$\tau + 35$	452.57 (6.83)		461.90 (7.43)		527.90 (9.55)		528.92 (8.56)
va	$\tau - 35$	458.02 (8.77)		472.39 (9.46)		534.10 (10.70)		538.65 (8.99)
	τ	463.96 (8.77)		494.32 (9.31)		534.00 (9.71)		555.63 (9.33)
	$\tau + 35$	477.37 (9.75)		514.87 (9.10)		542.37 (9.89)		564.19 (9.53)

Table 4.4: Reaction time [ms] as a function of intensity, stimulus onset asynchrony, and experimental paradigm.

Modality	SOA	Paradigm			
		RTP $_{\neg A}$	RTP	FAP $_{\neg A}$	FAP
VA	$\tau - 35$	9.9	13.5	10.7	7.4
	τ	9.1	7.8	7.9	5.0
	$\tau + 35$	6.4	4.5	8.0	4.4
Va	$\tau - 35$	3.3	6.6	4.7	1.1
	τ	1.3	4.3	4.7	1.0
	$\tau + 35$	2.5	3.3	5.6	2.6
vA	$\tau - 35$	43.5	44.5	24.1	24.0
	τ	43.1	43.0	16.4	22.8
	$\tau + 35$	38.3	40.7	16.9	22.6
va	$\tau - 35$	22.6	32.4	8.9	15.6
	τ	24.1	30.0	9.7	9.3
	$\tau + 35$	22.2	24.9	7.0	7.9

Table 4.5: Multisensory response enhancement calculated from inverse efficiency scores (MRE*) [%] as a function of intensity, stimulus onset asynchrony, and experimental paradigm.

5 The race model inequality for censored reaction time distributions¹

Abstract

The race model inequality (RMI) introduced in Miller (1982) puts an upper limit on the amount of reaction time facilitation within the redundant signals paradigm that is consistent with a race model. Here, it is shown through theoretical analysis and numerical simulation that inferences from the RMI test may become invalid when the experimenter misses a proportion of the responses by limiting the recording interval (right-censoring) or excluding outliers from analysis (left- and/or right-censoring). Moreover, a correction of the inequality test for right-censored reaction time distributions is proposed.

Keywords redundant target paradigm · race model inequality · crossmodal interaction · censored distribution · reaction time

5.1 Introduction

In the redundant signals paradigm for simple reaction time (RT), the observer must initiate a response as quickly as possible following the detection of any stimulus onset. A typical finding is a redundancy gain: responses are faster, on average, when two or more signals are presented simultaneously than when a single signal appears. This *redundant signals effect* (RSE) has often – although not always – been replicated under different experimental settings, *e.g.* comparing uni- *vs.* multimodal stimulation (Diederich, 1995; Diederich & Colonius, 1987; Gielen et al., 1983; Miller, 1982, 1986; Molholm et al., 2004), single *vs.* multiple stimuli within the same modality (*e.g.*, Schwarz & Ischebeck, 1994), or comparing monocular *vs.* binocular stimulation (Hughes & Townsend, 1998; Westendorf & Blake, 1988), and also for specific populations (*e.g.*, Corballis, 1998; Reuter-Lorenz et al., 1995; Savazzi & Marzi, 2004; Miller, 2004, for split-brain individuals; and Marzi et al., 1996, for

¹This chapter was accepted for publication in *Attention, Perception, & Psychophysics* on November, 14th, 2009 as: Rach, S., Diederich, A., Steenken, R., & Colonius, H. The race model inequality for censored reaction time distributions. (Scheduled to be published in volume 72, issue 3, April 2010.) The copyright is held by Psychonomic Society Publications. This document may not exactly correspond to the final published version. Psychonomic Society Publications disclaims any responsibility or liability for errors in this manuscript.

hemianopics).

Raab (1962) proposed a *race model* for simple RT postulating that (i) each individual stimulus elicits a (normally distributed) detection process performed in parallel to the others, and (ii) the winner's time determines the observable reaction time. The race model opens up the possibility that the RSE is generated by statistical facilitation: If detection latencies are interpreted as (nonnegative) random variables, the time to detect the first of several redundant signals is faster, on average, than the detection time for any single signal. Testing the race model amounts to probing whether an observed RT speed-up is too large to be attributable to statistical facilitation (*viz.* probability summation), no matter which distributional assumptions have been made.

A test of general race models was developed by Miller (1978, 1982) showing that

$$\Pr(RT_{XY} \leq t) \leq \Pr(RT_X \leq t) + \Pr(RT_Y \leq t) \quad (5.1)$$

must hold for all $t \geq 0$. This *race model inequality* (RMI) follows from

$$\Pr_{XY}(\min(X, Y) \leq t) \leq \Pr_X(X \leq t) + \Pr_Y(Y \leq t), \quad (5.2)$$

for any pair of random variables (X, Y) with a joint probability distribution based on \Pr_{XY} and with its marginal distributions identical to \Pr_X and \Pr_Y . Thus, as observed in Luce (1986, p. 130), the RMI test requires that the RT distributions in the single signal conditions are identical to the corresponding (marginal) RT distributions in the redundant signals condition (*cf.* Colonius, 1990). Note that, for fixed t , Inequality (5.2) corresponds to the wellknown *Boole's inequality* (*e.g.* Billingsley, 1979). Neglecting possible additional components (like motor time), the inequality stipulates that the RT distribution function for redundant stimuli is never larger than the sum of the RT distributions for the single stimuli. A violation of this inequality is interpreted as an indication of an underlying coactivation mechanism or some other strong form of non-independence.

Miller's test has become a standard tool in numerous empirical reaction time studies [see references above]. Moreover, it has been the subject of various theoretical and methodological studies as well (Ashby & Townsend, 1986; Colonius, 1990, 1999; Colonius & Ellermeier, 1997; Colonius & Townsend, 1997; Colonius & Vorberg, 1994; Diederich, 1992b; Miller, 1986, 1991, 2004; Miller & Ulrich, 2003; Mordkoff & Yantis, 1991; Townsend & Nozawa, 1995, 1997; Townsend & Wenger, 2004; Ulrich & Giray, 1986; Ulrich et al., 2007; Ulrich & Miller, 1997). For recent approaches toward a Kolmogorov-Smirnov-type test of the inequality, see Maris & Maris (2003) and Vorberg (2008). It has been observed by several investigators that the diagnostic power of the RMI test may be compromised under certain experimental conditions. Specifically, when subjects are not discouraged to make so-called "fast guesses", *i.e.*, responses given without processing the stimulus and assumed to be much faster than regular responses, or if display conditions are such that stimuli are sometimes not detected at all, it has been shown (Eriksen, 1988; Gondan & Heckel, 2008; Miller & Lopes, 1991; Miller & Ulrich, 2003) that the power of the RMI test to detect violations of the race model diminishes. Here we show that this holds not only for the case when subjects are missing a proportion of the stimuli but also when an experimenter excludes a proportion of the responses, by censoring the RT distributions from the left and/or the right, *i.e.*, excluding responses below and/or above a certain RT value because they are considered as anticipations or outliers (e.g., Giray & Ulrich, 1993; Leo et al., 2008; Miller, 2007b; Savazzi & Marzi, 2008; Schwarz, 2006). Furthermore, the experimenter may miss a proportion of the responses by not registering reactions given after some arbitrary upper bound; up to 8% of responses have sometimes been eliminated this way (cf. Gondan et al., 2007; Miller, 2007a,b,c).

5.2 Race model inequality for right-censored RT distributions

For concreteness, in the following we consider a simple RT experiment where subjects have to react by pressing a button upon detecting a stimulus, where on some of the trials more than one stimulus is presented (redundant signals paradigm). Let X, Y denote the random latencies to process two stimuli presented in a redundant trial, with distribution functions F_X, F_Y , respectively. When no responses are registered t_R [ms], say, after stimulus onset, the (right-censored) distribution function for the observed reaction times is

$$\begin{aligned} F_X(t|t_R) &= \Pr(X \leq t | X \leq t_R) \\ &= \Pr(\{X \leq t\} \cap \{X \leq t_R\}) / \Pr(X \leq t_R), \end{aligned}$$

by definition of conditional probability and assuming $\Pr(X \leq t_R) > 0$. It follows that

$$F_X(t|t_R) = \begin{cases} F_X(t)/F_X(t_R), & \text{if } t \leq t_R; \\ 1, & \text{otherwise.} \end{cases}$$

Analogous equations hold for the random variables Y and $\min(X, Y)$. Thus, the RMI that is actually tested in an experiment where no responses longer than t_R are recorded, is not

$$F_{XY}(t) \leq F_X(t) + F_Y(t) \tag{5.3}$$

but rather,

$$F_{XY}(t|t_R) \leq F_X(t|t_R) + F_Y(t|t_R) \tag{5.4}$$

or, equivalently,

$$F_{XY}(t)/F_{XY}(t_R) \leq F_X(t)/F_X(t_R) + F_Y(t)/F_Y(t_R),$$

writing F_{XY} for the distribution function in the redundant signals condition. Multiplying both sides of this inequality by $F_{XY}(t_R)$ yields

$$F_{XY}(t) \leq F_X(t) \times F_{XY}(t_R)/F_X(t_R) + F_Y(t) \times F_{XY}(t_R)/F_Y(t_R). \quad (5.5)$$

Note that under the race model assumption, by the monotonicity of probability,

$$\max[\Pr(X \leq t), \Pr(Y \leq t)] \leq \Pr(\{X \leq t\} \cup \{Y \leq t\})$$

for all t . Thus, in particular,

$$\max[F_X(t_R), F_Y(t_R)] \leq F_{XY}(t_R). \quad (5.6)$$

This inequality, sometimes called *Grice-inequality* (Grice et al., 1984), implies that the ratios $F_{XY}(t_R)/F_X(t_R) \equiv r_X$ and $F_{XY}(t_R)/F_Y(t_R) \equiv r_Y$ in Inequality (5.5) are larger than or equal to one. Thus, testing Inequality (5.4) is equivalent to testing

$$F_{XY}(t) \leq r_X \times F_X(t) + r_Y \times F_Y(t). \quad (5.7)$$

Since the right-hand side of this inequality may be larger than $F_X(t) + F_Y(t)$ for certain values of t , it may happen that the RMI for right-censored RT, Inequality (5.4), is satisfied for all t while the “true” (i.e., uncensored) RMI, Inequality (5.3), is violated for certain values of t . Thus, for right-censored distributions the RMI test becomes less diagnostic: it may indicate no violation of the race model although, in fact, RT facilitation is stronger than predicted by the race model. In other words, if the RMI test for right-censored data indicates a violation of the race model, then it is safe to assume that the redundancy gain is in fact large enough to suggest some type of coactivation mechanism.

Effect of right-censoring due to missing trials: Numerical example

This effect of right-censoring is illustrated here assuming gamma distributions (see appendix) for all three latencies, $X \sim \text{gamma}(2, 1/300)$, $Y \sim \text{gamma}(2, 1/250)$, and $XY \sim \text{gamma}(2, 1/140)$, where XY denotes the redundant signals latency. To see that this model does not satisfy the race model inequality, consider the function

$$R(t) = F_{XY}(t) - \min\{F_X(t) + F_Y(t), 1\}. \quad (5.8)$$

Obviously, the RMI is violated whenever this function takes on positive values. The upper left panel in Fig. 5.1 shows violations with the chosen gamma distributions when t is small enough. Taking the shaded area above the abscissa defined by $R(t)$ as an index of the magnitude of violation (Colonius & Diederich, 2006; Miller, 1986), the other panels of the figure illustrate how this area shrinks to zero with decreasing t_R , i.e., with amplifying censoring, making it more and more difficult to detect any violation of the race model inequality.

5.3 Race model inequality for left-censored RT distributions

Under the same experimental setting, if responses faster than or equal to t_L [ms] are considered anticipations and therefore excluded from analysis, the *left-censored* probability distribution function for the remaining reaction times is

$$\begin{aligned} F_X(t|t_L) &= \Pr(X \leq t | X > t_L) \\ &= \Pr(t_L < X \leq t) / [1 - \Pr(X \leq t_L)], \end{aligned}$$

by definition of conditional probability and assuming $\Pr(X > t_L) > 0$. It follows that

$$F_X(t|t_L) = \begin{cases} [F_X(t) - F_X(t_L)] / [1 - F_X(t_L)] & \text{for } t > t_L; \\ 0, & \text{otherwise.} \end{cases}$$

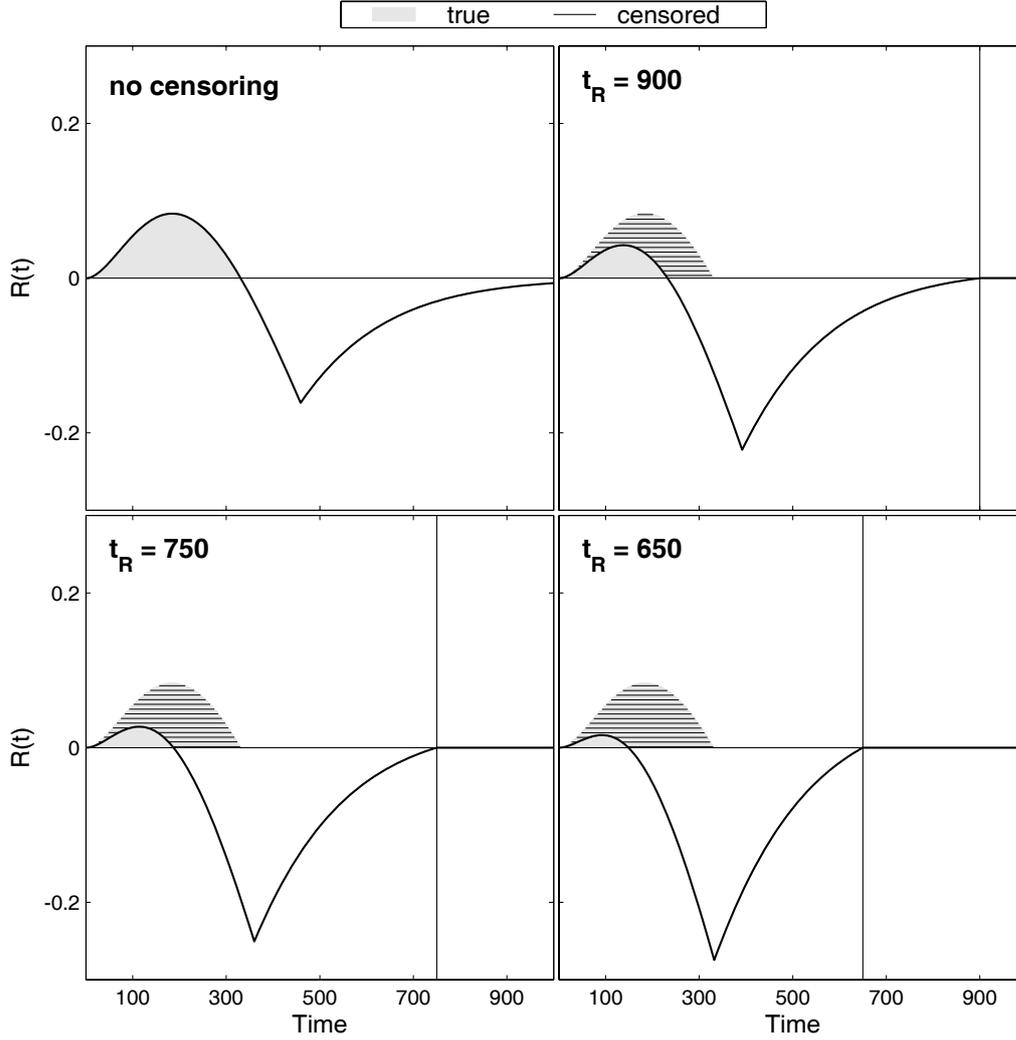


Figure 5.1: Positive values of the function $R(t)$ indicate violations of the race model inequality. Underlying distributions are gamma with a shape parameter of 2 and scale parameters of $1/300$ (X), $1/250$ (Y), and $1/140$ (redundant). The gray area below $R(t)$ is a measure of the amount of violation. As censoring is increased (by shifting the vertical line indicating the cut point t_R to the left), this area shrinks (horizontal hatching indicates the amount of decrease). Different subplots present results for truncation times $t_R = +\infty, 900, 750$, and 650 ms.

Analogous equations hold for the random variables Y and $\min(X, Y)$. Thus, the RMI actually tested in an experiment where no responses shorter than or equal to t_L are recorded is not

$$F_{XY}(t) \leq F_X(t) + F_Y(t)$$

but rather

$$F_{XY}(t|t_L) \leq F_X(t|t_L) + F_Y(t|t_L), \quad (5.9)$$

or equivalently, for $t > t_L$,

$$\frac{F_{XY}(t) - F_{XY}(t_L)}{1 - F_{XY}(t_L)} \leq \frac{F_X(t) - F_X(t_L)}{1 - F_X(t_L)} + \frac{F_Y(t) - F_Y(t_L)}{1 - F_Y(t_L)},$$

assuming non-zero denominators. Multiplying both sides of this inequality by $1 - F_{XY}(t_L)$ and adding $F_{XY}(t_L)$ yields

$$\begin{aligned} F_{XY}(t) &\leq F_{XY}(t_L) + [F_X(t) - F_X(t_L)] \cdot \frac{1 - F_{XY}(t_L)}{1 - F_X(t_L)} \\ &\quad + [F_Y(t) - F_Y(t_L)] \cdot \frac{1 - F_{XY}(t_L)}{1 - F_Y(t_L)} \\ &\equiv H(t). \end{aligned} \tag{5.10}$$

From Equation (5.6), both $[1 - F_{XY}(t_L)]/[1 - F_Y(t_L)]$ and $[1 - F_{XY}(t_L)]/[1 - F_X(t_L)]$ are less than or equal to 1, implying

$$\begin{aligned} H(t) &\leq F_{XY}(t_L) + [F_X(t) - F_X(t_L)] + [F_Y(t) - F_Y(t_L)] \\ &= F_X(t) + F_Y(t) + F_{XY}(t_L) - [F_X(t_L) + F_Y(t_L)] \\ &= F_X(t) + F_Y(t) + c, \end{aligned}$$

where $c \equiv F_{XY}(t_L) - [F_X(t_L) + F_Y(t_L)]$ is smaller than or equal to zero if RMI holds at t_L . Therefore, under the race model assumption we have

$$F_{XY}(t) \leq H(t) \leq F_X(t) + F_Y(t) + c, \tag{5.11}$$

with the first inequality representing the test performed under left-censoring. Obviously, for negative c the upper bound $H(t)$ constitutes a test stronger than that required by the race model when no left-censoring has been done. In other words, under left-censoring the inequality test of Eq. (5.9) may signal violations of the race model although it actually holds, i.e., left-censoring may generate "false alarms".

On the other hand, it is important to realize that at the same time, under left-censoring, an experimenter using the test of Eq. (5.9) will also miss all potential

violations of the race model occurring before the censoring time t_L .

Effect of left-censoring due to missing trials: Numerical example

Assuming gamma distributions for all three latencies, $X \sim \text{gamma}(2, 1/300)$, $Y \sim \text{gamma}(2, 1/250)$, and $XY \sim \text{gamma}(2, 1/140)$, the effect of left-censoring is illustrated in Fig. 5.2. Again, the shaded area above the abscissa defined by $R(t)$ is an index of the magnitude of violation (cf. Eq. (5.8)). The upper left panel in Figure 5.2 shows violations with the chosen gamma distribution when t is large enough. The other panels of the figure illustrate how this area is influenced by increasing t_L . With amplifying left-censoring, the amount of true violations not detected increases (horizontal hatching), whereas the amount of violations detected in the censored distributions but not present in the original distributions increases (vertical hatching).

5.4 Correcting for right-censoring

The results obtained so far clearly suggest that one should avoid censoring whenever possible (cf. Ulrich & Miller, 1994). Nevertheless, there are situations where censoring is a consequence of the experimental setup or paradigm. For instance, in simple response tasks with weak stimuli, a response is recorded as missing whenever the participant fails to respond within a certain time period (e.g., Gondan et al., 2007; Miller, 2007a,b,c), which amounts to right-censoring by the end of the recording interval t_R . The proportion of trials, where an experimenter fails to record a reaction time in a given condition because no response was given, is usually an observable quantity. Let p_X, p_Y, p_{XY} denote the corresponding probabilities for single stimulus and redundant stimulus presentations, respectively, with $p_X = 1 - F_X(t_R)$, $p_Y = 1 - F_Y(t_R)$, $p_{XY} = 1 - F_{XY}(t_R)$. Then ², Inequality (5.4) can be transformed back into the uncensored Inequality (5.3) by multiplying each of the distributions

²In order to check for independence of these "miss rates", the different tests suggested in Shaw (1982, pp. 367–369) may be considered.

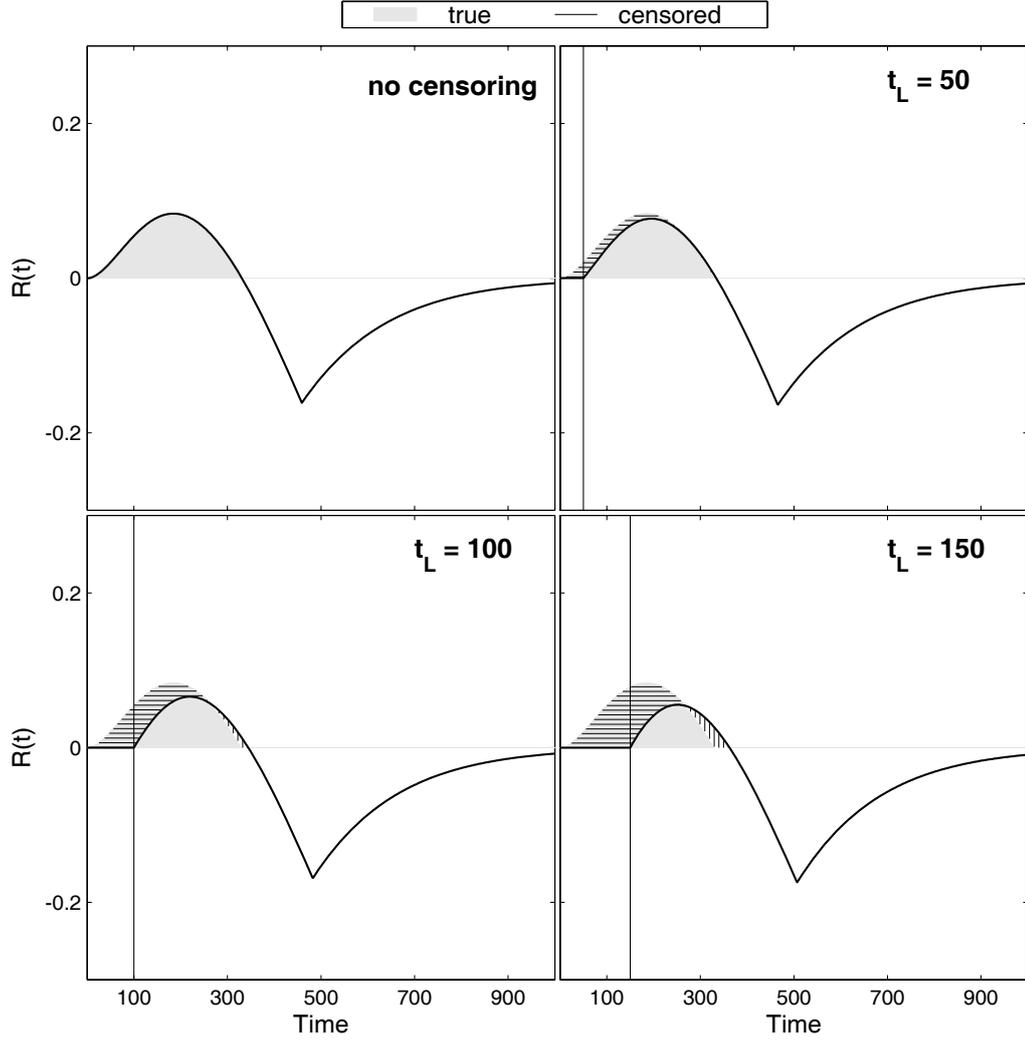


Figure 5.2: Positive values of the function $R(t)$ indicate violations of the race model inequality. Underlying distributions are gamma with a shape parameter of 2 and scale parameters of $1/300$ (X), $1/250$ (Y), and $1/140$ (redundant). The gray area below $R(t)$ is a measure of the amount of violation. As censoring is increased (by shifting the vertical line indicating the cut point t_L to the right), the amount of true violations not detected increases for the censored distributions (horizontal hatching), whereas the amount of violations detected in the censored distributions but not present in the original distributions increases (vertical hatching). Different subplots present results for truncation times $t_L = 0, 50, 100,$ and 150 ms.

by the corresponding proportion,

$$F_{XY}(t|t_R) \times (1 - p_{XY}) \leq F_X(t|t_R) \times (1 - p_X) + F_Y(t|t_R) \times (1 - p_Y). \quad (5.12)$$

In practice, when only estimates of the proportions of missing reactions (and of all distribution functions) are available, the above inequality will only approximate the

“true” inequality (5.3). An implementation of this correction is straightforward. First, choose the replacement time t_{max} to equal the largest RT in the whole data set increased by 1. Second, for every missing trial, set the corresponding RT to t_{max} . The corrected data set can then be subjected to common algorithms for the RMI test without further modifications (e.g., Ulrich et al., 2007). An analogous correction formula for simultaneous left- and right-censoring is given in the appendix.

5.5 Amount of inferential errors caused by censoring: Simulation results

As shown above, censoring may lead to two different kinds of inferential error: (i) a violation may occur in the original data but not in the censored data (i.e., a *miss*), or (ii) no violation occurs in the original data but it is indicated in the censored data (i.e., a *false alarm*).

To investigate the amount of inferential errors that may be caused by censoring, a simulation study for simple RT experiments was conducted. RTs were assumed to be Weibull-distributed for both single and redundant targets conditions. The Weibull distribution (see appendix) has previously been used to describe RT distributions in a variety of situations (cf., Colonius, 1995; Hsu, 2005; Logan, 1992; Maloney & Wandell, 1984; Marley, 1989; Marley & Colonius, 1992; McGill, 1963). An additional constant base time of 200 ms was assumed. The shape parameter of the Weibull was set to $\alpha = 1.5$ for all conditions. The scale parameters for the distributions of the single target conditions, X and Y , were set to $\lambda_X = \lambda_Y = 0.01$. To study the influence of the magnitude of coactivation gain (that is, the RT gain exceeding the race model prediction), three different scale parameters were assumed for the redundant targets condition, Z . The scale parameter was varied as follows: no coactivation gain, that is, a race model (gain = 0 ms; $\lambda_Z = (\lambda_X^\alpha + \lambda_Y^\alpha)^{1/\alpha} = 0.0159$), a small coactivation gain (gain = 5 ms; $\lambda_Z = 0.0174$), and a large coactivation gain (gain = 10 ms; $\lambda_Z = 0.0193$). The effect of the level of censoring was studied by excluding 1, 2.5, 5, 7.5, 10, or 15 % of the original data set. Furthermore, the direction of

censoring was varied by excluding data from the left, from the right, or from both directions³, resulting in a total of 18 censoring conditions (6 (levels) \times 3 (directions)). To study the appropriateness of the proposed correction procedures, the uncensored distributions were reconstructed using Eq. (5.12) for right-censored data, Eq. (5.A2) with $t_R = \infty$ for left-censored data, and Eq. (5.A2) for data censored from both directions.

Under these conditions, 10,000 simple RT experiments were simulated as follows. For each of 20 (virtual) participants, samples of 40 RTs were drawn from each of X , Y , and Z . From these samples, the cumulative distribution functions (CDFs) F_X , F_Y , and F_Z were estimated.⁴ Furthermore, the right-hand side of RMI, $F_S(t)$, was computed as follows:

$$F_S(t) = \begin{cases} F_X(t) + F_Y(t), & F_X(t) + F_Y(t) \leq 1 \\ 1, & \text{otherwise.} \end{cases} \quad (5.13)$$

From F_Z and F_S , the percentiles ($\hat{z}_{0.05}, \dots, \hat{z}_{0.5}$), and ($\hat{s}_{0.05}, \dots, \hat{s}_{0.5}$) were computed. Testing the RMI for percentiles above the 50% percentile seems not sensible because violations against the RMI are harder to detect as $F_X(t) + F_Y(t)$ becomes large relative to $F_Z(t)$ (Miller, 1982). For each of the percentiles, paired t-tests were computed across all participants with α adjusted to the number of percentiles tested (Bonferroni correction).⁵ Then, for each of the 18 censoring conditions, the original samples drawn from X , Y , and Z were censored and the resulting samples were subjected to the same testing procedure as the original samples (described above). Finally, the CDFs from the censored samples were corrected using the formulas Eq. (5.12) and Eq. (5.A2), and the resulting data were also subjected to the testing procedure. For each of these 10,000 simulated experiments, it was thus determined whether or not a violation occurred in the original, the censored, or the corrected sample.

³Excluding $x\%$ of the data by censoring from both directions was done by censoring $(x/2)\%$ from the left and $(x/2)\%$ from the right.

⁴CDFs and percentiles were estimated using the MatLab routine provided by Ulrich et al. (2007).

⁵Paired t-tests were computed using the MatLab routine provided by Cardillo (2006).

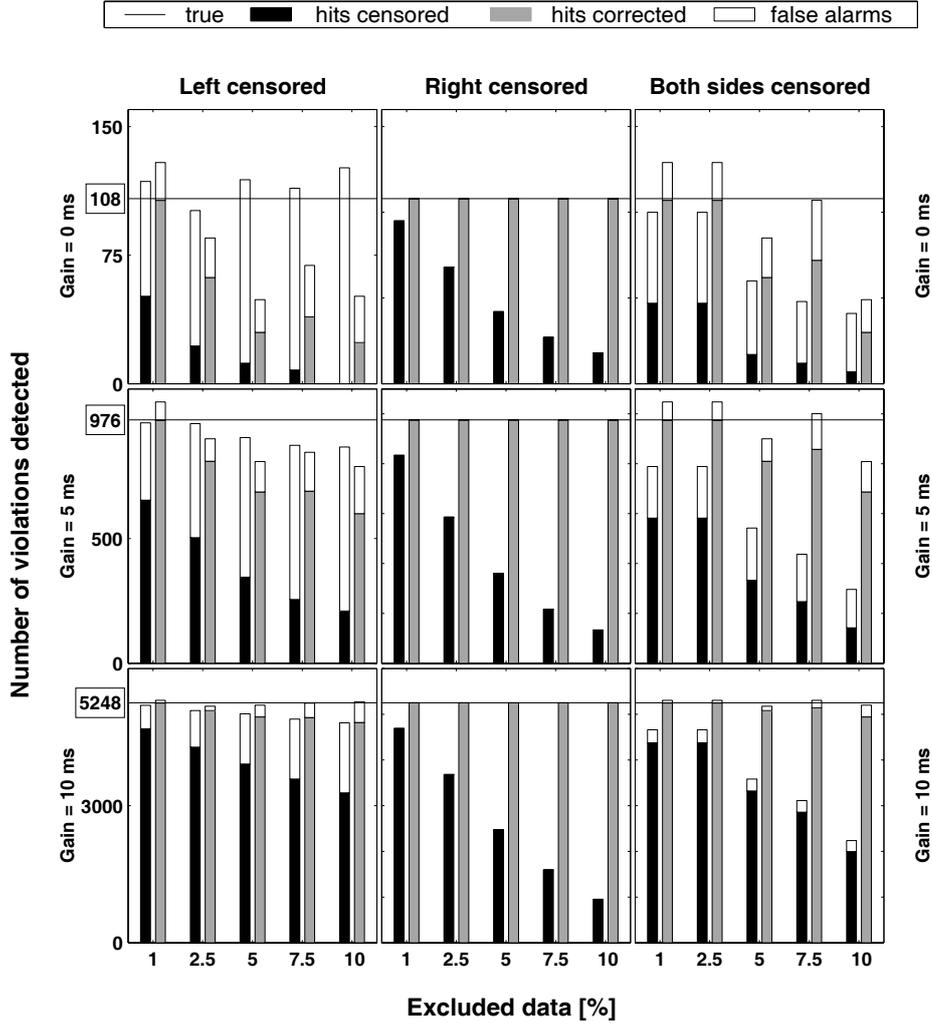


Figure 5.3: Influence of censoring on the diagnosticity of the race model inequality (RMI) as a function of coactivation gain [ms], censoring method, and percentage of excluded RTs. Columns present data for different censoring methods (from left to right: censoring from the left, censoring from the right, and censoring from both sides simultaneously). Rows present simulations for different coactivation gains (that is, the RT gain exceeding the race model prediction; from top to bottom: 0 ms, 5 ms, 10 ms). In each plot, the horizontal black lines indicate the total number of violations against the RMI detected in the original data in 10,000 simulated experiments. Black (gray) bars indicate the number of violations detected in the censored (corrected) data, when a violation was detected in the original data for the same experiment. White bars stacked on top of the black and gray bars indicate the number of false alarms, that is, experiments where a violation was detected in the modified data, but none was detected in the original data. See text for further details.

The results of the simulation are given in Fig. 5.3, presenting the influence of censoring on the diagnosticity of the RMI test as a function of the coactivation gain [ms], censoring method, and level of censoring. Columns present data for the

different censoring cases (from left to right: censoring from the left, censoring from the right, and censoring from both sides simultaneously). Rows present simulations for different coactivation gains (from top to bottom: 0 ms, 5 ms, 10 ms). In each plot, the horizontal black lines indicate the total number of violations of the RMI detected in the original data in 10,000 simulated experiments (i.e., the “true” number of violations). Black (respectively, gray) bars indicate the number of violations detected in the censored (respectively, corrected) data, when a violation was detected in the original data for the same experiment (i.e., a hit). Black (or gray) bars smaller than the black horizontal line indicate conditions where the RMI is less sensitive for the respective samples, that is, where violations are detected in the original sample but none are detected in the censored (respectively, corrected) sample. White bars stacked on top of the black and gray bars indicate the number of false alarms, that is, experiments where a violation was detected in the modified data, but none was detected in the original data. For the “no coactivation gain” condition (upper left panel), for instance, violations against the RMI were detected for 108 out of 10,000 simulated experiments (represented by the horizontal line). When 1% of the data were discarded by left-censoring, violations were detected for only 51 out of these 108 experiments (leftmost black bar). In addition, violations were detected for 67 experiments for which no violation was indicated in the original data (i.e., false alarms, white bar stacked on top of the leftmost black bar). The black and the white bar together cutting across the horizontal line indicates that left-censoring led to an RMI test stronger than that required by the race model when no left-censoring has been done. When the censored data were corrected, 107 out of 108 “true” violations were detected (leftmost gray bar) and, in addition, 22 false alarms occurred (white bar stacked on top of the leftmost gray bar). The RMI test for the corrected data is also stronger than required by the race model for the original data, as the gray and the white bar together cut across the horizontal line.

Fig. 5.3 clearly indicates that censoring can dramatically decrease the diagnosticity of the RMI, as the number of hits lies systematically below the “true” number of violations for almost all censored conditions (see the black bars), and the situa-

tion is worsening with increasing proportions of excluded RTs. The right-censoring results replicate the previous findings with the gamma distribution. Furthermore, for left-censoring and censoring from both directions (leftmost and rightmost column), the number of false alarms increases with increasing proportions of excluded RTs, whereas no false alarms occurred for right-censored data (middle column). The gray bars indicate that the suggested correction for right-censored data is capable of restoring the diagnosticity of the RMI. However, for left-censoring and censoring from both directions, the correction is not capable of countermanding the errors due to censoring, as, for most of the conditions in the leftmost and rightmost column, the gray bars fall below the horizontal lines and false alarms occur very frequently.

Results similar to the Weibull and the gamma distribution example studied here should occur with many other distribution families. In particular, we would expect the effect of censoring to become more severe when reaction time distributions are heavy-tailed, like the log-normal or the Pareto distribution.

5.6 Concluding Remarks

We have studied the effect of censoring reaction time distributions on the performance of the race model inequality (RMI) test. Theoretical analysis and numerical simulations revealed that the performance of the RMI test may be severely compromised under both left- and right-censoring. In particular, simulation results indicate that even moderate censoring of less than 5% of the total data set may strongly inflate the number of RMI violations not detected (misses) or may cause false alarms when the race model is in fact valid. Our recommendation, therefore, is that censoring should be avoided whenever possible.

Our analysis was based on the assumption that the rationale for censoring the distributions is somewhat pragmatic: right-censored distributions come about because the missing trials are indeed very late responses occurring after the end of the recording interval. One might substantiate this assumption, for example, by applying an experimental design that involves more than one stimulus intensity, in order

to allow for the affirmation of an inverse relation between stimulus intensity and the percentage of missing trials. On the other hand, if one has reasons to assume that the observed reaction times distributions are actually a probability mixture resulting from different underlying processes like fast guesses or lapses of attention, our recommendation of avoiding censoring or for correcting the censored distributions may no longer be valid.

5.7 Appendix

Gamma and Weibull distribution

Random variable X is distributed as *gamma* (α, λ) if its probability density is $f(x) = \lambda \exp[-\lambda x](\lambda x)^{\alpha-1}/\Gamma(\alpha)$ for $x > 0$ and $\alpha, \lambda > 0$.

Random variable X is distributed as *weibull* (α, λ) if its (cumulative) distribution function is $F(x) = 1 - \exp[-(\lambda x)^\alpha]$ for $x > 0$ and $\alpha, \lambda > 0$. In both cases, α is referred to as *shape* parameter and λ as *scale* parameter.

Correcting the RMI for left- and right-censoring

Let $t_L < t_R$ be the cutoffs where the RT distribution is censored from the left and right, respectively. The (left- and right-censored) distribution function for the observed reaction times is

$$\begin{aligned} F_X(t|t_L, t_R) &= \Pr(X \leq t | t_L \leq X \leq t_R) \\ &= \Pr(t_L \leq X \leq t) / [\Pr(X \leq t_R) - \Pr(X \leq t_L)]. \end{aligned}$$

The cumulative distribution function of the censored distribution is then given by

$$F_X(t|t_L, t_R) = \begin{cases} 0 & t < t_L, \\ [F_X(t) - F_X(t_L)] / [F_X(t_R) - F_X(t_L)] & \text{for } t_L \leq t \leq t_R, \\ 1 & \text{otherwise.} \end{cases} \quad (5.A1)$$

With $p_{X,L} = F_X(t_L)$, $p_{Y,L} = F_Y(t_L)$, $p_{XY,L} = F_{XY}(t_L)$ for the proportion of left-censored data, as well as $p_{X,R} = 1 - F_X(t_R)$, $p_{Y,R} = 1 - F_Y(t_R)$, $p_{XY,R} = 1 - F_{XY}(t_R)$ for the proportion of right-censored data, Inequality (5.A1) can be transformed back

to Inequality (5.3) by

$$\begin{aligned} & F_{XY}(t|t_L, t_R) \times (1 - p_{XY,R} - p_{XY,L}) + p_{XY,L} \leq \\ & F_X(t|t_L, t_R) \times (1 - p_{X,R} - p_{X,L}) + p_{X,L} + F_Y(t|t_L, t_R) \times (1 - p_{Y,R} - p_{Y,L}) + p_{Y,L}. \end{aligned} \tag{5.A2}$$

Note that Eq. (5.A2) simplifies to Eq. (5.12) for $t_L = 0$. Setting $t_R = \infty$ leads to a correction for left censored data.

6 General summary and discussion

Four experiments were conducted to investigate multisensory interaction effects of weak stimuli. In two experiments, visual, tactile, and visuo-tactile stimuli with varying durations were presented and saccadic reaction times (RTs) were recorded as dependent measure. Participants were instructed to respond only to visual stimuli and to ignore tactile stimuli (Focused Attention Paradigm, FAP). It was shown that the relative amount of response enhancement elicited by very short stimuli exceeded that elicited by longer stimuli.

In the third experiment, visual, auditory, and audio-visual stimuli with different intensities were presented and RTs and detection rates (DRs) were recorded. Participants were instructed to respond to any stimulus they detect (Redundant Target Paradigm, RTP). For the quantification of multisensory interaction effects, we developed two different measures of overall performance that integrate RTs and DRs. In the first approach, multisensory interaction effects were quantified from inverse efficiency scores (IES), a descriptive measure that weights RT by DR. In the second approach, interaction effects were quantified from parameters provided by sequential sampling models which were fitted to RTs and DRs. Bootstrap analyses revealed that both measures indicated the largest amount of response enhancement for the weakest stimulus conditions.

In the fourth experiment, visual and auditory stimuli of different intensities were presented in a detection task. Two different experimental instructions (RTP vs. FAP) and two presentation schemes (visual and audio-visual stimuli vs. visual, auditory, and audio-visual stimuli) were combined, resulting in four different experimental conditions. Overall performance quantified from IES was larger for RTP task than for FAP tasks. Presenting a unimodal auditory stimulus resulted in a decrease of performance in both RTP and FAP tasks. However, in RTP tasks, only the performance in the unimodal condition decreased, whereas the performance in both the unimodal visual and the audio-visual condition decreased in the FAP

task. In three out of four experimental conditions, the presentation of weak stimuli resulted in larger amounts of response enhancement compared to the presentation of high intense stimuli.

All four experiments provided evidence in favor of the principle of inverse effectiveness (POIE), which predicts that weaker stimuli are likely to elicit relatively larger amounts of response enhancement compared to stronger stimuli. This is in accordance with studies that used the POIE to describe results obtained by varying the intensity of stimuli (*e.g.*, Diederich & Colonius, 2004; Hecht et al., 2008) or by varying the signal to noise ratio (Corneil et al., 2002). Furthermore, the POIE accounted for patterns observed in neural responses recorded in midbrain structures of primates and cats (*e.g.*, Perrault et al., 2005; Rowland et al., 2007; Stein & Meredith, 1993; Wallace et al., 1998).

However, recently, the validity of the POIE has been discussed controversially (Holmes, 2007, 2008, 2009; Ma et al., 2009; Ross et al., 2007; Stein et al., 2009). Holmes (2007, 2008, 2009) suggested that – at least for some of the data – inverse effectiveness can be explained by regression towards the mean (RTTM). RTTM is a statistical principle describing the phenomenon that, for two measurements, the second measurement tends to be closer to the population mean than the first. Holmes illustrated his argument with a study from Serino et al. (2007, Experiment 1) who tested for inverse effectiveness by dividing their group of participants into good and bad performers based on results in the unimodal task (“neutral condition”), and then compared the performance of these groups in the bimodal task (“hand condition”). Serino et al. observed that the improvement in the bimodal condition was larger for the bad performers and interpreted this finding as evidence in favor of the POIE. Holmes (2007) argued that the larger gain of the bad performers can be explained in terms of RTTM: bad and good performers were selected for their extreme scores in the unimodal condition (*i.e.*, post hoc) and therefore are expected to have less extreme scores when tested again in the bimodal condition. Although Holmes’ argument might appropriately address a weak point of the Serino study, we do not think that the results presented in this thesis can be explained in terms of

RTTM, because this phenomenon arises from the post hoc assignment of intensity labels to experimental conditions based on the results in unimodal conditions. In the experiments presented here, the stimulus intensities were assigned to the experimental conditions a priori, *i.e.*, before the experiment was conducted. Since the chance of drawing an extreme random sample should be the same for all unimodal and bimodal conditions in the experiment, RTTM is expected to cancel out across the course of the experiment.

The second focal point of this thesis was the development of methods for the quantification of interaction effects in the vicinity of detection thresholds. We demonstrated that, for very weak stimuli, multisensory interaction effects can simultaneously influence RT and DR and that separately analyzing these measures can lead to contradictory conclusions: for some conditions of the third experiment, response enhancement was found when performance was quantified in terms of RT, but none was found in the same condition when performance was quantified in terms of DR, and vice versa. This result is in line with a study by Arieh & Marks (2008) who demonstrated that a gain in RT in crossmodal conditions can be due to speed-accuracy tradeoff, rather than due to multisensory interaction: participants responded faster in crossmodal conditions, but, at the same time, made more errors. We presented two indices that both integrate RT and DR into a single measure of overall performance. Multisensory response enhancement calculated from IES (MRE*) relates IES in crossmodal conditions to IES in unimodal conditions. MRE* is easy to compute and it can be subjected to non-parametric bootstrap to allow for statistical conclusions. However, not much is known about the assumptions that underly the calculations of IES and therefore the theoretical foundation of MRE* is not very elaborated. The second index, multisensory performance enhancement (MPE), is calculated from parameters of sequential sampling models that have been fitted to the data. MPE can be subjected to parametric bootstrap to allow for statistical conclusions. The sequential sampling models used in this approach are very elaborated and there is much evidence indicating their appropriateness for the integration of speed and accuracy measures. However, the parametric bootstrap is

computationally extensive; it requires computational power and time. Therefore, we provided two tools that might be useful in different situations: MRE* can be used for quick analyses to get a first impression of the data, while MPE provides the theoretical foundation for elaborated evaluations.

The idea of quantifying overall performance from multiple response measures is not limited to behavioral experiments. Recently, Bell et al. (2006) reported findings from neurophysiological studies on primates demonstrating that variations of stimulus intensity do not only affected a neuron’s activity in terms of spike count, but also its neuronal response onset latencies. Specifically, responses to more efficient stimuli did evoke larger spike counts *and* earlier response onset latencies. Furthermore, Rowland et al. (2007) reported that physiological response latencies can be shortened by multisensory integration. In this context, accurate response quantification requires an integration of spike count and response onset latency (*e.g.*, Rowland & Stein, 2008) and sequential sampling models might provide a useful tool for such analyses.

Furthermore, we demonstrated that the utilization of very weak stimuli can also influence the diagnosticity of the race model inequality (RMI), a common tool to test for multisensory response enhancement in RT. Often researchers exclude RTs above or below a particular criterion from further analyses because they are considered as anticipation errors or misses, respectively. We presented proofs showing that the diagnosticity of the RMI can be severely compromised if responses are missed because they are slower than a certain censoring time or excluded because they are considered as outliers. Correction formulas for censored data were derived for censoring from the left, from the right, and from both directions. However, a simulation study revealed that only the correction for censoring from the right seems feasible, as the remaining two corrections are likely to introduce false positives, that is, they indicate violations against the RMI for the censored data, when no violations are present for the “true” data. These results suggest that censoring should be avoided whenever possible. The correction for censoring from the right can be of special interest for experiments with stimuli near detection thresholds, because responses

to very weak stimuli can occur later than the termination of the recording interval, resulting in a right-censored RT distribution and the above-mentioned consequences for the RMI test.

Bibliography

- 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.
- Ashby, F. G. & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154–179.
- Becker, W. (1991). Saccades. In R. H. S. Carpenter (Ed.), *Eye Movements*, volume 8 of *Vision and visual dysfunction* chapt. 5, (pp. 95–137). Houndmills, Basingstoke, Hampshire and London: The Macmillan Press Ltd.
- Bell, A. H., Meredith, M. A., Opstal, A. J. V., & Munoz, D. P. (2005). Cross-modal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements. *Journal of Neurophysiology*, *93*, 3659–3673.
- Bell, A. H., Meredith, M. A., Opstal, A. J. V., & Munoz, D. P. (2006). Stimulus intensity modifies saccadic reaction time and visual response latency in the superior colliculus. *Experimental Brain Research*, *174*, 53–59.
- 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., & Edelman, B. A. (1969). Effects of an auditory signal on visual reaction time. *Journal of Experimental Psychology*, *80*(3), 567–569.

- Bernstein, I. H. & Edelman, 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.
- Billingsley, P. (1979). *Probability and measure*. New York: John Wiley & Sons.
- 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.
- Cardillo, G. (2006). Student t-test for unpaired or paired samples. <http://www.mathworks.com/matlabcentral/fileexchange/12699>.
- Colonus, H. (1990). Possibly dependent probability summation. *Journal of Mathematical Psychology*, *34*, 253–275.
- Colonus, H. (1995). The instance theory of automaticity: Why the weibull? *Psychological Review*, *102*, 744–750.
- Colonus, H. (1999). A theorem on parallel processing models with a generalized stopping rule. *Mathematical Social Sciences*, *38*, 247–258.
- Colonus, H. & Arndt, P. A. (2001). A two-stage model for visual-auditory interaction in saccadic latencies. *Perception & Psychophysics*, *63*(1), 126–147.
- Colonus, H. & Diederich, A. (2002). A maximum-likelihood approach to modeling multisensory enhancement. In T. G. Dietterich, S. Becker, & Z. Ghahramani (Eds.), *Advances in neural information processing systems 14*. Cambridge: The MIT Press.
- Colonus, H. & Diederich, A. (2004). Multisensory interaction in saccadic reaction time: A time-window-of-integration model. *Journal of Cognitive Neuroscience*, *16*(6), 1–10.

- Colonus, H. & Diederich, A. (2006). Race model inequality: Interpreting a geometric measure of the amount of violation. *Psychological Review*, *113*(1), 148–154.
- Colonus, H. & Ellermeier, W. (1997). Distribution inequalities for parallel models of reaction time with an application to auditory profile analysis. *Journal of Mathematical Psychology*, *41*, 19–27.
- Colonus, H. & Townsend, J. T. (1997). Activation-state representation of models for the redundant signals effect. In A. A. J. Marley (Ed.), *Choice, Decision, and Measurement: Essays in Honor of R. Duncan Luce* (pp. 245–254). Mahwah, NJ: Erlbaum.
- Colonus, H. & Vorberg, D. (1994). Distribution inequalities for parallel models with unlimited capacity. *Journal of Mathematical Psychology*, *38*, 35–58.
- Corballis, M. C. (1998). Interhemispheric neural summation in the absence of the corpus callosum. *Brain*, *121*, 1795–1807.
- 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. (1992a). *Intersensory facilitation: Race, superposition, and diffusion models for reaction time to multiple stimuli* (Dissertation). Frankfurt: Verlag Peter Lang.
- Diederich, A. (1992b). Probability inequalities for testing separate activation models of divided attention. *Perception & Psychophysics*, *52*, 714–716.
- 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.
- Diederich, A. & Colonius, H. (2007). Modeling spatial effects in visuotactile saccadic reaction time. *Perception & Psychophysics*, *69*(1), 56–67.
- Diederich, A. & Colonius, H. (2008). Crossmodal interaction in saccadic reaction time: Separating multisensory from warning effects in the time window of integration model. *Experimental Brain Research*, *186*(1), 1–22.
- Diederich, A., Colonius, H., Bockhorst, D., & Tabeling, S. (2003). Visual-tactile spatial interaction in saccade generation. *Experimental Brain Research*, *148*, 328–337.
- Efron, B. & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, *1*, 54–75.
- Eriksen, C. W. (1988). A source of error in attempts to distinguish coactivation from separate activation in the perception of redundant targets. *Perception & Psychophysics*, *44*, 191–193.

- 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.
- Ghazanfar, A. A. & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*, 278–285.
- 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.
- Gondan, M. & Heckel, A. (2008). Testing the race inequality: A simple correction procedure for fast guesses. *Journal of Mathematical Psychology*, *52*, 322–325.
- Gondan, M., Lange, K., Rösler, F., & Röder, B. (2004). The redundant target effect is affected by modality switch costs. *Psychonomic Bulletin & Review*, *11*(2), 307–313.
- Gondan, M., Vorberg, D., & Greenlee, M. W. (2007). Modality shift effects mimic multisensory interactions: an event-related potential study. *Experimental Brain Research*, *182*, 199–214.
- Grice, G. R., Canham, L., & Boroughs, J. M. (1984). Combination rule for redundant information in reaction time tasks with divided attention. *Perception & Psychophysics*, *35*, 451–463.

- Harrington, L. K. & Peck, C. K. (1998). Spatial disparity affects visual-auditory interactions in human sensorimotor processing. *Experimental Brain Research*, *122*, 247–252.
- Hecht, D., Reiner, M., & Karni, A. (2008). Enhancement of response times to bi- and tri-modal sensory stimuli during active movements. *Experimental Brain Research*, *185*, 655–665.
- Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *Journal of Experimental Psychology*, *63*(3), 289–293.
- Hilgard, E. R. (1933). Reinforcement and inhibition of eyelid reflexes. *Journal of General Psychology*, *8*, 85–113.
- Holmes, N. P. (2007). The law of inverse effectiveness in neurons and behavior: Multisensory integration versus normal variability. *Neuropsychologica*, *45*, 3340–3345.
- Holmes, N. P. (2008). The seemingly inviolable principle of inverse effectiveness: In search of a null hypothesis. Paper presented at the 9th International Multisensory Research Forum, Hamburg, Germany, July, 17th, 2008.
- Holmes, N. P. (2009). The principle of inverse effectiveness in multisensory integration: Some statistical considerations. *Brain Topography*, *21*, 168–176.
- Hsu, Y.-F. (2005). On measuring the minimum detection time: a simple reaction time study in the time estimation paradigm. *British Journal of Mathematical and Statistical Psychology*, *58*(2), 259–284.
- Hughes, H. C. & Townsend, J. T. (1998). Varieties of binocular interactions in human vision. *Psychological Science*, *9*, 53–60.
- Kadunce, D. C., Vaughan, W. J., Wallace, M. T., & Stein, B. E. (2003). The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. *Experimental Brain Research*, *139*, 303–310.

- King, A. J. & Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research*, *60*, 492–500.
- Kitagawa, N. & Spence, C. (2005). Investigating the effect of a transparent barrier on the crossmodal congruency effect. *Experimental Brain Research*, *161*, 62–71.
- Leo, F., Bertini, C., di Pellegrino, G., & Làdavas, E. (2008). Multisensory integration for orienting responses in humans requires the activation of the superior colliculus. *Experimental Brain Research*, *186*, 67–77.
- Logan, G. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *102*(18), 883–914.
- 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.
- Ma, W. J., Zhou, X., Ross, L. A., Foxe, J. J., & Parra, L. C. (2009). Lip-reading aids word recognition most in moderate noise: A bayesian explanation using high-dimensional feature space. *PLoS ONE*, *4*(3), e4638.
- Maloney, L. T. & Wandell, B. A. (1984). A model of a single visual channels response to weak test lights. *Vision Research*, *24*, 633–640.
- Maris, G. & Maris, E. (2003). Testing the race model inequality: A nonparametric approach. *Journal of Mathematical Psychology*, *47*, 507–514.
- Marley, A. A. J. (1989). A random utility family that includes many of the classical models and has closed form choice probabilities and choice reaction times. *British Journal of Mathematical and Statistical Psychology*, *42*, 13–36.

- Marley, A. A. J. & Colonius, H. (1992). The horse race random utility model for choice probabilities and reaction times, and its competing risks interpretation. *Journal of Mathematical Psychology*, *36*, 1–20.
- Marzi, C. A., Smania, N., Martini, M. C., Gambina, G., Tomerelli, G., Palamara, A., Alessandrini, F., & Prior, M. (1996). Implicit-redundant target effects in visual extinction. *Neuropsychologia*, *34*, 9–22.
- McGill, W. J. (1963). Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*, volume 1 (pp. 309–360). New York: Wiley.
- 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. (1978). Multidimensional same-different judgments: Evidence against independent comparisons of dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 411–422.
- 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. (1991). Channel interaction and the redundant targets effect in bimodal divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 160–169.
- Miller, J. O. (2004). Exaggerated redundancy gain in the split brain: A hemispheric coactivation account. *Cognitive Psychology*, *49*, 118–154.
- Miller, J. O. (2007a). Contralateral and ipsilateral motor activation in visual simple reaction time: a test of the hemispheric coactivation model. *Experimental Brain Research*, *176*, 539–558.

- Miller, J. O. (2007b). Effects of response task and accessory stimuli on redundancy gain: Tests of the hemispheric coactivation model. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 829–844.
- Miller, J. O. (2007c). Interhemispheric interactions and redundancy gain: tests of an interhemispheric inhibition hypothesis. *Experimental Brain Research*, *180*, 389–413.
- Miller, J. O. & Lopes, A. (1991). Bias produces by fast guessing in distribution-based tests of race models. *Perception & Psychophysics*, *50*, 584–590.
- Miller, J. O. & Ulrich, R. (2003). Simple reaction time and statistical facilitation: A parallel gains model. *Cognitive Psychology*, *46*, 101–151.
- Molholm, S., Ritter, W., Javitt, D., & Foxe, J. J. (2004). Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. *Cerebral Cortex*, *14*, 452–465.
- Mordkoff, J. T. & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 520–538.
- Morrell, E. K. (1967). Intersensory facilitation of reaction time. *Psychonomic Science*, *8*(2), 77–78.
- Morrell, E. K. (1968). Temporal characteristics of sensory interaction in choice reaction times. *Journal of Experimental Psychology*, *77*(1), 14–18.
- Munoz, D. P. & Wurtz, R. H. (1995). Saccade-related activity in monkey superior colliculus. II. Characteristics of burst and buildup cells. *Journal of Neurophysiology*, *73*(6), 2313–2333.
- 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 Neurophysiol*, *93*, 2575–2586.

- Populin, L. C. & Yin, T. C. T. (2002). Bimodal interactions in the superior colliculus of the behaving cat. *Journal of Neuroscience*, *22*(7), 2826–2834.
- Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Science*, *24*, 574–590.
- 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.
- Reuter-Lorenz, P. A., Nozawa, G., Gazzaniga, M. S., & Hughes, H. C. (1995). Fate of neglected targets: A chronometric analysis of redundant target effects in the bisected brain. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 211–230.
- Röder, B., Kusmirek, A., Spence, C., & Schicke, T. (2007). Developmental vision determines the reference frame for the multisensory control of action. *PNAS*, *104*, 4753–4758.
- Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. J. (2007). Do you see what I am saying? exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex*, *17*, 1147–1153.
- Rowland, B. E., Quessy, S., Stanford, T. R., & Stein, B. E. (2007). Multisensory integration shortens physiological response latencies. *Journal of Neuroscience*, *27*, 5879–5884.
- Rowland, B. E. & Stein, B. E. (2008). Temporal profiles of response enhancement in multisensory integration. *Frontiers in Integrative Neuroscience*, *2*, 218–224.
- Savazzi, S. & Marzi, C. A. (2004). The superior colliculus subserves interhemispheric neural summation in both normals and patients with a total section or agenesis of the corpus callosum. *Neuropsychologia*, *42*, 1608–1618.

- Savazzi, S. & Marzi, C. A. (2008). Does the redundant signal effect occur at an early visual stage? *Experimental Brain Research*, *184*, 275–281.
- Schwarz, W. (1994). Diffusion, superposition, and the redundant–targets effect. *Journal of Mathematical Psychology*, *38*, 504–520.
- Schwarz, W. (2006). On the relationship between the redundant signals effect and temporal order judgments: Parametric data and a new model. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 558–573.
- Schwarz, W. & Ischebeck, A. (1994). Coactivation and statistical facilitation in the detection of lines. *Perception*, *23*, 157–168.
- Serino, A., Farn‘e, A., Rinaldesi, M. L., Haggard, P., & L’adavas, E. (2007). Can vision of the body ameliorate impaired somatosensory function? *Neuropsychologica*, *45*, 1101–1107.
- Shaw, M. L. (1982). Attending to multiple sources of information: I. The integration of information in decision making. *Cognitive Psychology*, *14*, 353–409.
- Shore, D. I., Barnes, M. E., & Spence, C. (2006). Temporal aspects of the visuotactile congruency effect. *Neuroscience Letters*, *392*, 96–100.
- Spence, C. & Driver, J. (1997). On measuring selective attention to an expected sensory modality. *Perception & Psychophysics*, *59*, 384–403.
- Spence, C., McGlone, F. P., Kettenmann, B., & Kobal, G. (2001). Attention to olfaction: A psychophysical investigation. *Experimental Brain Research*, *138*, 432–437.
- Spence, C., Nicholls, M. E., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, *63*(2), 330–336.
- Stanford, T. R. & Stein, B. E. (2007). Superadditivity in multisensory integration: putting the computation in context. *Neuroreport*, *18*(8), 787–792.

- Steenken, R., Colonius, H., Diederich, A., & Rach, S. (2008). Visualauditory interaction in saccadic reaction time: Effects of auditory masker level. *Brain Research, 1220*, 150–156.
- Stein, B. E. & Meredith, M. A. (1993). *The merging of the senses*. London, Cambridge: The MIT Press.
- Stein, B. E. & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience, 9*, 255–266.
- Stein, B. E., Stanford, T. R., Ramachandran, R., Perrault, T. J., & Rowland, B. A. (2009). Challenges in quantifying multisensory integration: alternative criteria, models, and inverse effectiveness. *Experimental Brain Research, 198*, 113–126.
- Todd, J. W. (1912). Reaction to multiple stimuli. In R. S. Woodworth (Ed.), *Archives of Psychology, No.25. Columbia Contributions to Philosophy and Psychology*, Vol. XXI, No. 8. New York: The Science 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.
- Townsend, J. T. & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology, 39*, 321–359.
- Townsend, J. T. & Nozawa, G. (1997). Serial exhaustive models can violate the race model inequality: Implications for architecture and capacity. *Psychological Review, 104*, 595–602.
- Townsend, J. T. & Wenger, M. J. (2004). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review, 111*, 1003–1035.

- Ulrich, R. & Giray, M. (1986). Separate-activation models with variable base time: Testability and checking of cross-dependency. *Perception & Psychophysics*, *36*, 545–558.
- Ulrich, R., Miller, J., & Schröter, H. (2007). Testing the race model inequality: An algorithm and computer programs. *Behavior Research Methods*, *39*, 291–302.
- Ulrich, R. & Miller, J. O. (1994). Effects of truncation on reaction time analysis. *Journal of experimental psychology. General.*, *123*, 34–80.
- Ulrich, R. & Miller, J. O. (1997). Tests of race models for reaction time experiments with asynchronous redundant targets. *Journal of Mathematical Psychology*, *41*, 367–381.
- Vorberg, D. (2008). Exact statistical tests of the race-model and related inequalities. Paper presented at the 39th Meeting of the European Mathematical Psychology Group, University of Graz, Austria, September 7 - 11, 2008.
- 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.
- Wallace, M. T., Meredith, M. A., & Stein, B. E. (1998). Multisensory integration in the superior colliculus of the alert cat. *Journal of Neurophysiology*, *80*(2), 1006–1010.
- Wallace, M. T., Wilkinson, L. K., & Stein, B. E. (1996). Representation and integration of multiple sensory inputs in primate superior colliculus. *Journal of Neurophysiology*, *76*, 1246–1266.
- Welch, R. B. & Warren, D. H. (1986). Intersensory interactions. In K. R. Boff, L. Kaufmann, & J. P. Thomas (Eds.), *Handbook of perception and human performance, Vol.1: Sensory processes and perception* chapter 25, (pp. 25.1–25.36). New York: Wiley.

- Westendorf, D. H. & Blake, R. (1988). Binocular reaction times to contrast increments. *Vision Research*, *28*, 355–359.
- 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.

Acknowledgements

This thesis was supported by a tuition waiver from Jacobs University Bremen.

I am grateful to my supervisors Adele Diederich and Hans Colonius for their valuable advice, reliable support, and patience.

I would like to thank Claus Hilgetag, Thomas Münte, and Benjamin Godde for joining my dissertation committee.

In Fall 2005, I accidentally deleted many data sets from one of the experiments presented in this thesis, annihilating the work of almost 3 months. Thanks to the help of Annette Schomburg, I was able to restore the majority of the lost files. Furthermore, the pre-existing experimental setup utilized for the experiments was built by Annette Schomburg. Annette, many thanks for helping with technical expertise and – even more – for many enjoyable conversations.

I am thankful to Felix Schmitz-Justen for many enjoyable lunch breaks.

I would like to thank Margaret Parks, Angie Köllner (former Nichols), Dennis Küster, and the volunteers who participated in my experiments.

I am thankful to Michael O. Akinde and Robert Osterlund for proofreading parts of this thesis.

I am very grateful to Rike Steenken for helpful discussions and an open ear whenever I needed one. Rike, thank you for being a friend.

The love of a family is life's greatest blessing.

Unknown

Declaration by word of honor

I hereby certify on my honor that this thesis is my own work and that I have completed it without undue help from third parties and without the use of any material other than permitted. Any thoughts and ideas taken directly or indirectly from others are highlighted as such. Neither this work in its present form nor any other work of its contents has been submitted to another German or foreign board of examiners so far.

Oldenburg, 22. September 2009

Stefan Rach