Integration auditiv-visueller Reizinformation

Inaugural-Dissertation zur Erlangung der Doktorwürde der Fakultät für Psychologie, Pädagogik und Sportwissenschaft der Universität Regensburg

vorgelegt von Steven Blurton aus Regensburg

Regensburg, 2014

Gutachter (Betreuer):	Prof. Dr. Mark W. Greenlee (Universität Regensburg)
Gutachter:	Prof. Dr. Karl-Heinz Bäuml (Universität Regensburg)

Nam impossibile est res huius mundi sciri, nisi sciatur mathematica.

– Roger Bacon, *Opus maius*

Dank

Ich möchte mich ganz herzlich bei all denen bedanken, die mich während meiner Zeit als Doktorand unterstützt und damit wesentlich zum Gelingen dieser Dissertation beigetragen haben.

Ganz besonders danke ich Prof. Dr. Matthias Gondan, der mich über den gesamten Zeitraum der Promotion in schier unermüdlichem Einsatz als Projektleiter aber vor allem als Mentor unterstützt hat und damit diese Dissertation überhaupt erst ermöglicht hat. Er ließ mir jedwede Unterstützung zukommen, die man sich als Doktorand nur wünschen kann. Ebenso möchte ich Prof. Dr. Mark W. Greenlee danken, der mich ebenfalls fortwährend wissenschaftlich und in jeglicher praktischer Hinsicht unterstützt hat.

Mein herzlicher Dank gilt auch den Mitarbeitern der Arbeitsgruppe am Lehrstuhl für Allgemeine Psychologie und Methodenlehre. Eure Unterstützung und vor allem das angenehme Arbeitsumfeld haben erheblich zu dieser Dissertation beigetragen. In diesem Zusammenhang gilt mein Dank auch den studentischen Hilfskräften Julia Wenzl, Anja Wienbreyer und Anna-Maria Wirth, die mich bei den langwierigen Datenerhebungen unterstützt haben.

Auch bei der Deutschen Forschungsgemeinschaft möchte ich mich an dieser Stelle für die gewährte finanzielle Unterstützung (GO 1855/1-1 und GR 988/20-2) bedanken.

Schließlich möchte ich meinen Freunden und meiner Familie danken, die mir während meiner Zeit als Doktorand viel Geduld und Verständnis entgegenbrachten und ich mich in jeder Phase unterstützt haben.

Inhalt

Zusammenfassung

Die in dieser Dissertation zusammengefassten Arbeiten befassen sich mit der Integration audiovisueller Reizinformation. Von entscheidendem Interesse war dabei die Frage, *wie* Information aus verschiedenen Modalitäten zu einem Perzept integriert werden.

Dazu wurde ein Reaktionszeitmodell untersucht, dem eine explizite Annahme zur Art und Weise der Integration zugrunde liegt – das Prinzip der additiven Superposition von Reizinformation. Das Modell trifft Aussagen über beobachtbares Verhalten in einem Redundanzexperiment, bei dem einer Testperson unimodale oder bimodale Reize dargeboten werden. Dieses Modell, und damit die Annahme der additiven Superposition, wurde im Rahmen dieser Dissertation einer eingehenden empirischen Überprüfung unterzogen, wobei zentrale Fragestellungen bezüglich der Integration multisensorischer Reize untersucht wurden.

Dies wären zum einen Redundanzeffekte unter verschiedenen Aufgabenanforderungen. Bisherige Untersuchungen stützten sich vor allem auf Ergebnisse zu Einfachreaktionsaufgaben. Wie aber läuft die Integration redundanter Reizinformation bei komplexeren Aufgaben ab? Die reine Stimulusdetektion beschreibt lediglich einen Aspekt menschlichen Erlebens; oft ist entscheidend, in angemessener Weise auf distale Reize zu reagieren. In der ersten Arbeit (Blurton, Greenlee & Gondan, 2014) wurden die bei einer Go/No-go-Aufgabe und einer Wahlreaktionsaufgabe anfallenden Redundanzgewinne untersucht. Die Ergebnisse weisen in vielfacher Weise darauf hin, dass die Integration redundanter Reizinformation ähnlich abläuft, wie bei Einfachreaktionsaufgaben.

Im Rahmen dieser Untersuchung wurde das Reaktionszeitmodell derart erweitert, dass es auch auf komplexere Aufgabenstellungen angewendet werden kann. In einem weiteren Artikel wurden mathematische Eigenschaften dieses und verwandter Modelle untersucht (Blurton, Kesselmeier & Gondan, 2012).

Eine weitere in diesem Kontext interessante Fragestellung ist der Einfluss

von Aufmerksamkeit auf die Integration redundanter Reizinformation. Dieser Fragestellung wurde in zwei weiteren Arbeiten nachgegangen. In einer Arbeit wurde die Ausrichtung räumlicher Aufmerksamkeit mit Hinweisreizen manipuliert und der Effekt dieser Manipulationen auf die Detektion von multisensorischen Reizen untersucht (Blurton, Greenlee & Gondan, in Begutachtung). Die Ergebnisse legen nahe, dass der Mechanismus der Integration multisensorischer Reizinformation unabhängig von der Ausrichtung räumlicher Aufmerksamkeit ist: Die Verarbeitung – und damit die Stimulusdetektion – ist zwar effektiver, wenn die Aufmerksamkeit auf die Reizposition gerichtet ist, die Ergebnisse lassen jedoch darauf schließen, dass die Art und Weise der Informationsintegration unabhängig von Aufmerksamkeitsverschiebungen ist.

In einer weiteren Arbeit wurde der Einfluss selektiver und räumlicher Daueraufmerksamkeit untersucht (Gondan, Blurton, Hughes & Greenlee, 2011). Auch die Ergebnisse dieser Experimente unterstützen die Auffassung, dass die Stimulusdetektion zwar von einer dauerhaften Ausrichtung der Aufmerksamkeit profitieren kann, die Art und Weise der Informationsintegration jedoch nicht davon abhängt.

Zusammenfassend lässt sich feststellen, dass das Prinzip der additiven Superposition multisensorischer Reizinformation ein eher generelles Integrationsprinzip darstellt, das unabhängig vom Aufgabentyp und der Ausrichtung von Aufmerksamkeit zu sein scheint.

Im folgenden Kapitel werden die theoretischen Grundlagen eingehender dargelegt; anschließend folgt ein kurzer Überblick über die erwähnten Arbeiten, die dieser Dissertation zugrunde liegen. Der Rest der Dissertation sind Nachdrucke von veröffentlichten Zeitschriftsartikeln sowie ein Manuskript, das zur Veröffentlichung eingereicht wurde.

1 Einleitung

1.1 Multisensorische Integration

Die Untersuchung multisensorischer Wahrnehmungsleistungen beschäftigt sich mit der Frage, wie Sinneseindrücke der verschiedenen Sinnessystems zu einem einheitlichen Perzept integriert werden (Welch & Warren, 1986). Die Integration ist notwendig, um von redundanten Reizinformationen zu profitieren, aber auch um widersprüchliche Informationen in Einklang zu bringen. Es ist bekannt, dass Reizinformation verschiedener Sinnessysteme in verschiedenen neuronalen Strukturen verarbeitet wird, daher stellt sich die Frage, wie und wo Information einzelner Sinnessysteme integriert werden. Als Zentren multisensorischer Integration wurden wiederholt die Colliculi superiores diskutiert (Meredith & Stein, 1986), jedoch scheinen auch kortikale Strukturen bei der Integration von multimodalen Sinneseindrücken eine Rolle zu spielen (Farah, Wong, Monheit & Morrow, 1989; Macaluso, Frith & Driver, 2000; McDonald, Teder-Sälejärvi, Di Russo & Hillyard, 2003; McDonald, Teder-Sälejärvi, Di Russo & Hillyard, 2005; Beer, Plank & Greenlee, 2011). In der Vergangenheit sind bei diesen Arbeiten zur multisensorischen Integration vor allem superadditive Zellantworten untersucht worden (Meredith & Stein, 1986), neuere Publikationen zu diesem Thema heben jedoch die Bedeutung additiver Zellantworten auf Informationen aus verschiedenen Sinnessystemen hervor (Stanford, Quessy & Stein 2005).

In der vorliegenden Arbeit wurde die Annahme der additiven Integration von sinnesspezifischer Aktivierung einer eingehenden empirischen Prüfung unterzogen. Dazu wurde mehrere Experimente zur Integration audiovisueller Reizinformation durchgeführt. Von visuell-vestibulären Interaktionen abgesehen, ist die Integration audiovisueller Reizinformation die mit Abstand am häufigsten untersuchte Form multisensorischer Wahrnehmung (z.B. Hershenson, 1962; Miller, 1982, 1986; Diederich & Colonius, 1987; Giray & Ulrich, 1993). Dies dürfte zum einen

daran liegen, dass auditive und visuelle Stimulation im Labor relativ einfach zu realisieren sind. Zum anderen dürfte auch ausschlaggebend sein, dass audiovisuelle Integration vielen kognitiven Funktionen des täglichen Lebens zugrunde liegt, zum Beispiel Sprache (McGurk & MacDonald, 1976; Sumby & Pollack, 1954).

Die Integration audiovisueller Reizformation wird häufig anhand des Redundanzeffektes untersucht. In dieser Versuchsanordnung werden in verschiedenen Bedingungen entweder Reize einer Sinnesmodalität (unimodale Bedingungen) oder redundante Zielreize verschiedener Sinnesmodalitäten dargeboten. Die Aufgabe ist meist, so schnell wie möglich auf alle Reize zu reagieren, es kommen jedoch auch Aufgaben des Go/No-go Typs zum Einsatz, bei der nur auf Zielreize reagiert werden soll, nicht aber auf Distraktoren. Dies ist vor allem bei Experimenten mit Ableitung des Elektroenzephalogramms (EEG) der Fall (z.B. bei Talsma, Doty & Woldorff, 2007). Unabhängig vom Aufgabentyp beobachtet man in der Regel, dass Reaktionen auf redundante Reize im Mittel schneller sind als in den unimodalen Bedingungen (Redundanzeffekt, engl.: *redundant signals effect* oder *redundant target effect*). Der Redundanzeffekt sagt zunächst noch nicht viel über den Mechanismus der Reizintegration aus; es ist seit langem bekannt, dass auch Modelle, die keine Integration im eigentlichen Sinne annehmen, Redundanzeffekte allein anhand stochastischer Gegebenheiten erklären können (Raab, 1962). Mit der Formulierung der als Wettlaufmodelle (*race model*) bezeichneten Modelle erschienen Redundanzexperimente zunächst als ungeeignet, um Fragestellungen zur multisensorischen Integration zu untersuchen.

Es gelang Miller (1982, 1986) jedoch, eine empirisch testbare Vorhersage für Wettlaufmodelle herzuleiten. Diese Vorhersage gibt eine obere Grenze für den Redundanzeffekt an, den Wettlaufmodelle (bzw. allgemeiner, Modelle mit separater Reizaktivierung in beiden Sinneskanälen) erklären können. Ist sie verletzt, scheiden Modelle separater Aktivierung als Erklärung aus, Miller (1982) sprach in diesem Fall von *Koaktivierung*. Koaktivierung sagt nichts über die genaue Art

der Integration aus; sie liegt dann vor, wenn der Redundanzeffekt größer ist, als von Modellen separater Aktivierung erklärt werden kann. Wenn separate Aktivierung ausgeschlossen wird, muss die Information aus den sinnesspezifischen Kanälen demnach an einer Stelle in einem gemeinsamen Kanal integriert werden. Auch wenn in zahlreichen Arbeiten belegt werden konnte, dass die Redundanzgewinne nur durch Koaktivierung erklärt werden können, und obwohl zahlreiche explizite Modelle zur Erklärung des Redundanzeffektes bei Koaktivierung vorgeschlagen wurden, beschränken sich die meisten empirischen Arbeiten im Bereich multisensorischer Integration auf eine Testung der Vorhersage des Wettlaufmodells (Abschnitt 1.2).

An dieser Stelle setzen die hier beschriebenen Experimente an: Um die Art und Weise der Integration genauer zu analysieren, wurde ein Reaktionszeitmodell mit einem explizit formulierten Integrationsmechanismus an die erhobenen Daten angepasst. Diesem in Abschnitt 1.3 genauer beschriebenen Koaktivierungsmodell liegt die Annahme einer additiven Superposition

$$\mathbf{X}_{\mathrm{AV}}(t) = \mathbf{X}_{\mathrm{A}}(t) + \mathbf{X}_{\mathrm{V}}(t)$$

der sinnesspezifischen Aktivierung im auditiven Verarbeitungskanal $\mathbf{X}_{\mathrm{A}}$ bzw. im visuellen Kanal $\mathbf{X}_{\mathrm{V}}$ zugrunde (Schwarz, 1994). Das Modell konnte die Daten unter verschiedenen Aufgabenstellungen (Kapitel 3) und auch unter verschiedenen Aufmerksamkeitsbedingungen (Kapitel 5 und 6) gut erklären. Im Folgenden werden zunächst die theoretischen Grundlagen der Arbeiten dargelegt – die Wettlaufungleichung und das verwendete Reaktionszeitmodell (Diffusions-Superpositionsmodell). Anschließend folgt eine kurze Übersicht über die Arbeiten, um sie in einen gemeinsamen Zusammenhang zu setzen.

1.2 Das Wettlaufmodell

Der Redundanzeffekt kann durch eine Reihe von Modellen beschrieben werden, diese lassen sich in zwei Klassen unterteilen: Modelle separater Aktivierung und Koaktivierungsmodelle (Miller, 1982). Bei Modellen separater Aktivierung wird angenommen, dass die Reizinformation in sinnesspezifischen Kanälen verarbeitet wird und dass die Aktivitäten dieser Kanäle nicht integriert werden. In diese Klasse fällt unter anderen das Wettlaufmodell. Beim Wettlaufmodell wird parallele Verarbeitung der Reizinformationen angenommen. Sobald ein Kanal die Verarbeitung abgeschlossen hat, erfolgt eine Reaktion (*first-terminating processing*). Der Redundanzeffekt ergibt sich in diesem Modell als statistische Erleichterung (*statistical facilitation*): bei bimodalen Reizen sind im Gegensatz zu unimodalen Reizen zwei Sinneskanäle aktiv, so dass langsame Detektionszeiten in einem Kanal durch schnellere Detektionszeiten im anderen Kanal kompensiert werden können. In Wettlaufmodell wird die Reizinformation nicht integriert, allein die stochastische Annahme sich überlappender Dichtefunktionen der sinnesspezifischen Detektionszeiten genügt, um den Redundanzeffekt erklären zu können.

Allerdings ist das Ausmaß des Redundanzeffektes begrenzt, das durch dieses Modell erklärt werden kann. Dadurch, dass der Redundanzeffekt durch statistische Erleichterung zustande kommen muss, muss die folgende Ungleichung bei separaten Aktivierungsmodellen immer gelten (Miller, 1982):

$$F_{\mathrm{R}}(t) \leq F_{\mathrm{S1}}(t) + F_{\mathrm{S2}}(t), \text{ für alle } t. \tag{1}$$

Der Wert der Verteilungsfunktion $F_{\mathrm{R}}(t)$ der Reaktionszeiten auf redundante Reize (R) darf nach Ungleichung 1 für keinen Zeitpunkt t größer sein, als die Summe der Verteilungsfunktionen der Reaktionszeiten auf die Einzelreize (S_1 und S_2). Da für $t \to \infty$ der rechte Teil von Ungleichung (1) gegen zwei und der linke Teil gegen eins geht, sind Verletzungen nur für kleine t zu erwarten (Abbildung 1.1).

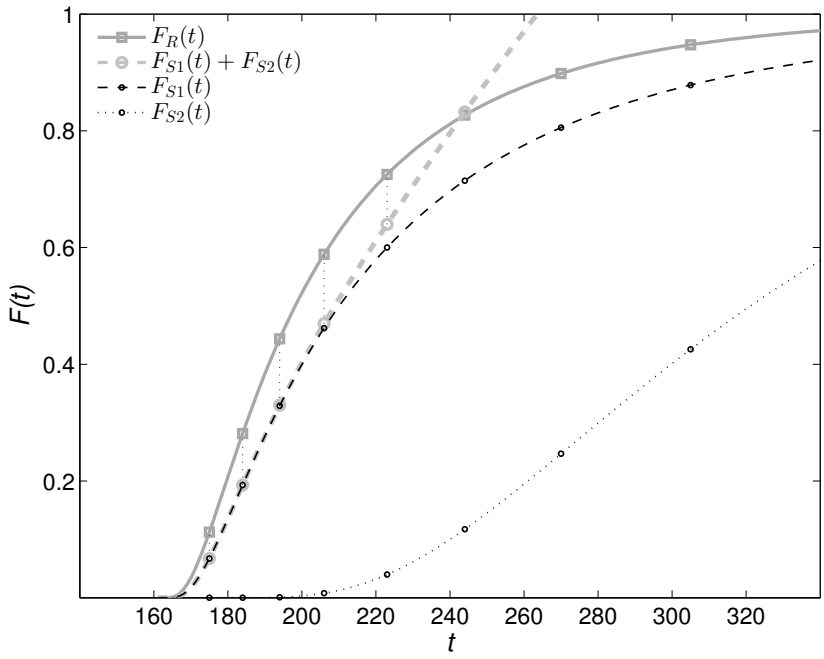


Abbildung 1.1: Theoretischer Redundanzeffekt auf Ebene der Verteilungsfunktionen. Die Verteilungsfunktion der Redundanzbedingung $F_{\mathrm{R}}(t)$ ist für kleine t größer als die Summe der Verteilungsfunktionen $F_{\mathrm{S1}}(t) + F_{\mathrm{S2}}(t)$ – eine Verletzung der Vorhersage für Wettlaufmodelle. Für große t strebt die Summe gegen 2 (nicht dargestellt). Die Verteilungsfunktionen sind theoretische Vorhersagen eines Koaktivierungsmodells (Schwarz, 1994, Tabelle 1), angepasst an die Daten von Miller (1986, Proband BD).

Ungleichung (1) wird allgemein als Wettlaufungleichung bezeichnet (*race-model-inequality*, Miller, 1982, 1986). Ist sie verletzt, dann ist der beobachtete Redundanzeffekt größer, als von Modellen separater Aktivierung erklärt werden kann. Die Verteilungsfunktionen in (1) werden durch die beobachteten Reaktionen geschätzt, deswegen können Verletzungen durch zufällige Stichprobenvariation zustande kommen. Um sich dagegen abzusichern, wurden zahlreiche (nonparametrische) Tests vorgeschlagen (Miller, 1986), denen unterschiedliche Annahmen zugrunde liegen (Maris & Maris, 2003) und die zum Teil unterschiedliche Hypothesen testen (Colonius & Diederich, 2006; Gondan, 2010). Es konnte außerdem gezeigt werden, dass Antizipationen die Tests der Wettlaufungleichung konservativer machen (Gondan & Heckel, 2008). Deswegen wurde vorgeschlagen, eine sog.

„Kill-the-twin“ Prozedur (KTT, Eriksen, 1988) vor dem Test der Wettlaufungleichung durchzuführen. Durch die KTT-Prozedur wird für jede Reaktion auf einen Durchgang ohne Zielreiz (*catch-trial*) eine Reaktion aus der Reaktionzeitverteilung der Zielreize entfernt, das heißt auf unendlich gesetzt. Dabei wird diejenige Reaktion auf einen Zielreiz gewählt, die von der Latenz her der Antizipation am nächsten ist. Die Reaktionszeitverteilungen werden dadurch von Antizipationen annähernd gesäubert; formal betrachtet ergibt sich die Wettlaufungleichung in diesem Fall zu

$$F_{\mathrm{R}}(t) + F_{\mathrm{C}}(t) \leq F_{\mathrm{S1}}(t) + F_{\mathrm{S2}}(t), \text{ für alle } t, \tag{2}$$

wobei als Antizipationen C alle Reaktionen auf *Catch-trials* gewertet werden. Vorhersage (2) wurde in den drei empirischen Arbeiten getestet.

Verletzungen der Wettlaufungleichung wurden oft berichtet, vor allem in Experimenten mit audiovisuellen Reizen und Einfachreaktionsaufgaben oder Go/ No-go Aufgaben. Interessant in diesem Zusammenhang ist die Beziehung zur *feature-integration*-Theorie (Treisman & Gelade, 1980; Treisman, 1986). So konnte belegt werden, dass Koaktivierung nur unter bestimmten Bedingungen auftritt: Feintuch und Cohen (2002) untersuchten Redundanzeffekte bei visuellen Reizen unter verschiedenen Aufmerksamkeitsbedingungen. Nur wenn es sich um Eigenschaften eines gemeinsamen Objekts handelte, wurde die Wettlaufungleichung verletzt, ansonsten lagen die Redundanzgewinne im Geltungsbereich des Wettlaufmodells. Es konnte außerdem belegt werden, dass ein Zusammenfassen zweier separater visueller Reize durch eine Umrahmung wiederum Redundanzgewinne erbrachte, die die Wettlaufungleichung verletzten. Was den Reaktionsmodus betrifft, wurden von Grice, Canham und Boroughs (1984) bzw. von Mordkoff und Miller (1993, Exp. 2) Verletzungen der Wettlaufungleichung bei Wahlreaktionen bzw. Go/No-go Aufgaben berichtet. Weitgehend unklar ist dagegen, welche Rolle Aufmerksamkeit bei

der Bearbeitung multisensorischer redundanter Reize zukommt. Studien zur Lokalisation von auditiven Reizen in Anwesenheit von visuellen Reizen (ventriloquism) weisen darauf hin, dass diese Form der Integration unabhängig von der Aufmerksamkeit der Testperson stattfindet (Bertelson, Vroomen, de Gelder & Driver, 2000, Vroomen, Bertelson & de Gelder, 2001). Es stellt sich die Frage, ob die Integration audiovisueller Reizinformation ebenfalls unabhängig von der Aufmerksamkeit der Testperson ist. Ist dies der Fall, müssten Verletzungen der Wettlaufungleichung unabhängig von der Ausrichtung der Aufmerksamkeit nachweisbar sein. Ist Aufmerksamkeit dagegen für eine effektive Integration notwendig, so wäre zu erwarten, dass Koaktivierungseffekte nur unter Beachtung der audiovisuellen Reize zu finden sind (vgl. Feintuch & Cohen, 2002). Dieser Frage wurde in zwei Arbeiten nachgegangen, in denen die Wettlaufungleichung unter verschiedenen Aufmerksamkeitsbedingungen überprüft wurde. In einer weiteren Arbeit wurden Redundanzgewinne unter verschiedenen Aufgabenstellungen (Go/No-go und Wahlreaktionsaufgabe) untersucht. In allen Arbeiten wurde außerdem ein Koaktivierungsmodell an die Daten angepasst und getestet, um den Integrationsmechanismus explizit zu untersuchen.

1.3 Das Diffusions-Superpositionsmodell

Wie bereits dargelegt, sind Koaktivierungsmodelle als Gegensatz zu Modellen separater Aktivierung zu verstehen. Kennzeichnend für alle Koaktivierungsmodelle ist, dass die Aktivität der sinnesspezifischen Kanäle in einem gemeinsamen Kanal integriert wird (vgl. Miller, 1982; Schwarz, 1989, 1994; Miller & Ulrich, 2003). Grundlage der vorliegenden Arbeit ist das Diffusions-Superpositionsmodell (DSM) von Schwarz (1994). Als Integrationsmechanismus wird in diesem Modell der wohl einfachste gewählt – laut Modell wird die Aktivität der sinnesspezifischen Känale in einem gemeinsamen Kanal aufaddiert. Im Folgenden soll, wie bei Schwarz (1994), der Redundanzeffekt im Kontext auditiver und visueller Reize betrachtet werden.

Laut Modell kann das Aktivierungsniveau des auditiven (A) und des vi-

suellen (V) Kanals als latente Variable durch Wiener Prozesse $\mathbf{X}_{\mathrm{A}}(t)$ bzw. $\mathbf{X}_{\mathrm{V}}(t)$ $[\mathbf{X}(0) = 0]$ beschrieben werden. Diese besitzen die Parameter μ_{A} und σ^2_{A} bzw. μ_{V} und σ^2_{V} und beschreiben die Informationsakkumulation nach Darbietung eines (unimodalen) Reizes. Der Wiener Prozess ist ein zeitstetiger Markov-Prozess mit stationären Inkrementen, seine Zuwächse $\mathrm{d}x$ bestehen aus normalverteilten $N[\mu\,\mathrm{d}t, \sigma^2\,\mathrm{d}t]$ und unabhängigen Inkrementen. Der Parameter $\mu \geq 0$ repräsentiert die Drift im Diffusionsprozess und der Parameter $\sigma > 0$ das zufällige Ausmaß der Diffusion (Diffusionskonstante). Die Partikeldichte $P(x,t)$ des Wiener Prozesses ergibt sich als Lösung der partiellen Differenzialgleichung (Fokker-Planck-Gleichung)

$$\frac{\partial}{\partial t}P(x,t) = \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2}P(x,t) - \mu\frac{\partial}{\partial x}P(x,t) \tag{3}$$

Ursprünglich wurde der Wiener-Prozess als Modell zur Beschreibung der Brownschen Molekularbewegung entwickelt (Einstein, 1905), fand in der Folgezeit jedoch Eingang in viele weitere Naturwissenschaften, u.a. in die Neurowissenschaften (Brunel & Wang, 2001).

Gemäß der Annahme additiver Superposition sinnesspezifischer Aktivierung kann die Aktivität des gemeinsamen Kanals $\mathbf{X}_{\mathrm{AV}}(t)$ als Summe der Aktivität des auditiven Kanals und des visuellen Kanals beschrieben werden: $\mathbf{X}_{\mathrm{AV}}(t) = \mathbf{X}_{\mathrm{A}}(t) + \mathbf{X}_{\mathrm{V}}(t)$. Dadurch ergibt sich die Aktivität des gemeinsamen Kanals ebenfalls zu einem Wiener Prozess mit den Parametern $\mu_{\mathrm{AV}} = \mu_{\mathrm{A}} + \mu_{\mathrm{V}}$ und $\sigma^2_{\mathrm{AV}} = \sigma^2_{\mathrm{A}} + \sigma^2_{\mathrm{V}} + 2\rho_{\mathrm{AV}}\sigma_{\mathrm{A}}\sigma_{\mathrm{V}}$. Dieser Prozess beschreibt die Informationsakkumulation nach Präsentation eines redundanten Reizes (AV). Eine weitere zentrale Annahme dieses Modells ist, dass dieser Prozess andauert, bis ein Antwortkriterium $c > 0$ erreicht ist. Erreicht der Prozess dieses Kriterium, ist die Reizdetektion abgeschlossen und es erfolgt eine Reaktion. Formal betrachtet handelt es sich bei c um einen absorbierenden Zustand, d.h. der Prozess verlässt diesen nicht mehr, wenn er ihn einmal

erreicht hat. Dazu muss die Fokker-Planck-Gleichung (3) zusätzlich unter der Nebenbedingung $P(c,t) = 0$ gelöst werden. Die in diesem Zusammenhang relevante Lösung lautet

$$P(x,t) = \frac{1}{\sigma\sqrt{2\pi t}} \left\{ \exp\left[-\frac{(x-\mu t)^2}{2\sigma^2 t} \right] - \exp\left[\frac{2\mu c}{\sigma^2} - \frac{(x-2c-\mu t)^2}{2\sigma^2 t} \right] \right\}. \quad (4)$$

Die Zeit, die es dauert, bis das Kriterium c zum ersten Mal erreicht und der Prozess absorbiert wird, lässt sich durch die Erstpassagezeit (*first-passage time*) $\mathbf{T}$ angeben (Schrödinger, 1915; von Smoluchowski, 1915; Fürth, 1917). Eine Möglichkeit, die Dichte- und Verteilungsfunktion für $\mathbf{T}$ zu berechnen, ist über

$$F(t \mid c, \mu, \sigma^2) = 1 - \int_{-\infty}^{c} P(x,t)\ \mathrm{d}x \quad (5)$$

die Verteilungsfunktion zu bestimmen (Schrödinger, 1915, Gl. 17–19), und über die Ableitung nach t die Dichtefunktion zu errechnen. Sie lautet (z.B. in Cox & Miller, 1965, S. 221):

$$f(t \mid c, \mu, \sigma^2) = \frac{c}{\sigma\sqrt{2\pi t^3}}\ \exp\left[-\frac{(c-\mu t)^2}{2\sigma^2 t} \right]. \quad (6)$$

Alternativ berechnet man die Dichtefunktion $f(t \mid c, \mu, \sigma^2)$ aus $\frac{1}{2}\sigma^2 \left(\frac{\partial P(x,t)}{\partial x} \right)_{x=c}$ und die Verteilungsfunktion anschließend durch Integration der Dichtefunktion über t (vgl. von Smoluchowski, 1915, Gl. 7 & 8; Fürth, 1917, S. 180).

Ist der Driftparameter darüber hinaus positiv ($\mu > 0$), lautet der Erwartungswert $E(\mathbf{T}) = c/\mu$ und die Varianz $Var(\mathbf{T}) = c \cdot \sigma^2/\mu^3$ (z.B. in Cox & Miller, 1965, S. 221–222). Diese Verteilung ist auch als *inverse Normalverteilung* (oder *inverse Gauß-Verteilung* oder *Wald-Verteilung*) bekannt.

Im Falle unimodaler oder synchroner bimodaler Reizdarbietung können die Modellvorhersagen für Mittelwerte und Varianz direkt aus diesen Beziehungen abgeleitet werden. Werden die bimodalen Reizkomponenten zeitversetzt dargeboten,

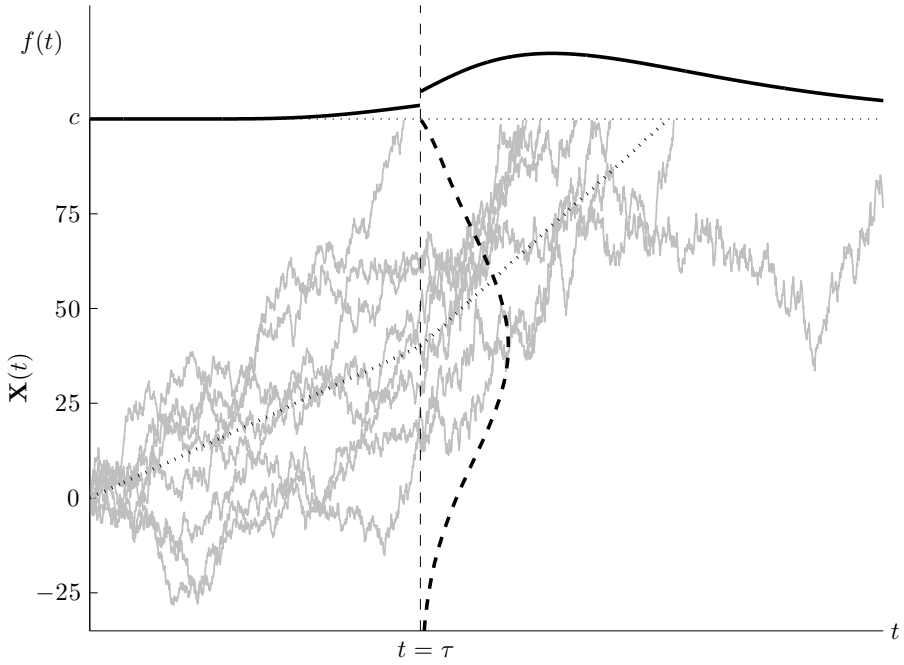


Abbildung 1.2: Additive Superposition der Diffusionsprozesse. Bis zum Zeitpunkt $\tau = 100$ ist lediglich ein Kanal aktiv. Nach der Präsentation des zweiten Reizes zum Zeitpunkt $t = \tau$ ist auch der zweite Kanal aktiv. Der Beitrag von beiden aktiven Kanäle (hier beide: $\mu = 0.4, \sigma = 3.0$) wird addiert, so dass das Kriterium c im Mittel schneller erreicht wird. Dies ist vor allem an der mittleren Position eines Prozesses ersichtlich (gestrichelte Linie). Zum Zeitpunkt $t = \tau$ hat der Prozess einen Zustand $x < c$ erreicht, die Wahrscheinlichkeitsdichte $P(x, \tau)$ ist durch die vertikale Verteilung gekennzeichnet (vgl. Gleichung 4). Die resultierende Dichtefunktion $f(t)$ ist im oberen Abschnitt dargestellt.

gestaltet sich die Herleitung der Modellvorhersagen erheblich schwieriger. In diesem Fall ist zunächst nur ein Kanal aktiv, der zweite ist bis zur Präsentation der zweiten Reizkomponente inaktiv. Nach Präsentation der zweiten Reizkomponente zum Zeitpunkt τ sind wiederum beide Kanäle aktiv und die Aktivität des gemeinsamen Kanals $\mathbf{X}_{\mathrm{AV}}(t)$ soll durch den oben beschriebenen Prozess repräsentiert werden (Abbildung 1.2). Dadurch verliert der Wiener Prozess eine seiner komfortablen Eigenschaften: die Stationarität. Es müssen daher die zwei Fälle $\mathbf{D} \leq \tau$ und $\mathbf{D} > \tau$ getrennt betrachtet werden; der gemeinsame Erwartungswert $E(\mathbf{D})$ ergibt sich als

Summe dieser Erwartungswerte. Für den ersten Fall lautet der Erwartungswert

$$E(\mathbf{D}|\mathbf{D} \leq \tau) \cdot P(\mathbf{D} \leq \tau) = \int_0^\tau t \times f(t \,|\, c, \mu, \sigma^2)\, \mathrm{d}t.$$

Nach der Zeit τ ist der Prozess entweder absorbiert, oder befindet sich in einem Zustand $\mathbf{X}(\tau) = x < c$. Von diesem Zustand aus benötigt der Prozess im Mittel $E[\mathbf{D}|\mathbf{X}(\tau) = x] = \frac{c-x}{\mu}$, um das Kriterium c zu erreichen. Die Wahrscheinlichkeit $P\left[(\mathbf{X}(\tau) = x\right]$ wird durch $P(x, \tau)$ [vgl. (4)] repräsentiert. Für den zweiten Fall $\mathbf{D} > \tau$ berechnet sich der Erwartungswert nach

$$E(\mathbf{D}|\mathbf{D} > \tau) \cdot P(\mathbf{D} > \tau) = \int_{-\infty}^{c} P(x, \tau) \cdot \left(\tau + \frac{c - x}{\mu}\right) \mathrm{d}x.$$

In seiner Arbeit hat Schwarz (1994, Gl. 10, Anhang A) analytische Lösungen für beide Ausdrücke hergeleitet, neben denen für die Erwartungswerte auch die entsprechenden Vorhersagen für die Varianz *Var*(**D**) (Schwarz, 1994, Anhang B).

Das Modell für die Detektionszeit **D** besitzt 4 Parameter, denn das Kriterium c dient lediglich als Skalierungsfaktor. Für die Vorhersage der mittleren Reaktionszeit **T** (**T** = **D** + **M**) muss mit μ_M für die mittlere Latenz der restlichen Prozesse **M** („motorische Prozesse“, vgl. Luce, 1986) ein weiterer Parameter geschätzt werden. Für die Vorhersage der Mittelwerte müssen folglich fünf freie Parameter aus den Daten geschätzt werden: μ_A, σ^2_A, μ_V, σ^2_V und μ_M. Soll zudem die Varianz der Reaktionszeiten vorhergesagt werden, kommt mit σ^2_M ein weiterer Parameter hinzu. Nimmt man korrelierte Reizkanäle an, muss außerdem ρ_AV geschätzt werden; mit der Korrelation zwischen Detektionszeit und Latenz der residualen Prozessen zusätzlich der Parameter ρ_DM. Für das Gesamtmodell mit Varianzvorhersage müssen zusätzlich zu den fünf oben genannten Parametern σ^2_M und ggf. ρ_AV sowie ρ_DM, insgesamt sechs bis maximal acht freie Parameter geschätzt werden.

Das DSM sagt Mittelwerte und Varianzen von Einfachreaktionen auf redun-

dante Reize voraus. Es konnte zum Beispiel die von Miller (1986) berichteten Daten sehr gut erklären. Eine deutliche Erweiterung dieses Modells zur Beschreiung der Integration trimodaler multisensorischer Reize wurde von Diederich (1995) vorgeschlagen. Dieses Modell umfasst neben einem dritten Aktivierungskanal ferner eine zweite, reflektierende Barriere, so dass der Informationskumulationsprozess im Gegensatz zum Schwarz-DSM nicht ins Negative abdriften kann (vgl. Abbildung 1.2). Zudem wird in diesem Modell nicht der Wiener Prozess, sondern der flexiblere Ohrnstein-Uhlenbeck-Prozess (Uhlenbeck & Ornstein, 1930) verwendet. Der Ohrnstein-Uhlenbeck-Prosss ist ebenfalls ein Diffusionsprozess, er beschriebt wiederum die Aktivierung der sinnesspezifischen Kanäle. Allerdings sind die involvierten Zustandsgleichungen derart komplex, dass für dieses Modell nur eine numerische Näherung gegeben werden konnte.

Eine weitere Verallgemeinerung auf komplexere Aufgabentypen als Einfachreaktionsaufgaben wurde im Zuge dieser Dissertation erarbeitet (s. Kapitel 3). Bei dieser Verallgemeinerung kommt eine zweite, absorbierende Barriere hinzu, so dass beispielsweise auch die Daten von Wahlreaktionsexperimenten mit zwei Antwortalternativen durch das DSM erklärt werden können (Abbildung 1.3). Dies ist inbesondere deswegen reizvoll, weil dieses Modell nicht nur an Erwartungswerte und Varianzen beider Antwortalternativen angepasst werden kann, sondern gleichzeitig an den Anteil korrekter Antworten. Im Gegensatz zum Wiener Prozess mit Drift und einer absorbierenden Barriere, bei der die Absorption an dieser Barriere ein sicheres Ereignis ist, ist das Erreichen einer der beiden Barrieren im Zwei-Barrieren-Fall eine (binäre) Zufallsvariable. Die Wahrscheinlichkeit für die Absorption an einer der Barrieren hängt von den Parametern μ, σ und der Platzierung der Barrieren u und $-\ell$ ab. Die Wahrscheinlichkeit für die Absorption an der oberen Barriere u ist durch

$$P_{\mathrm{u}} = 1 - \exp(-2\mu\ell/\sigma^2)/[\exp(-2\mu u/\sigma^2) - \exp(2\mu\ell/\sigma^2)]$$

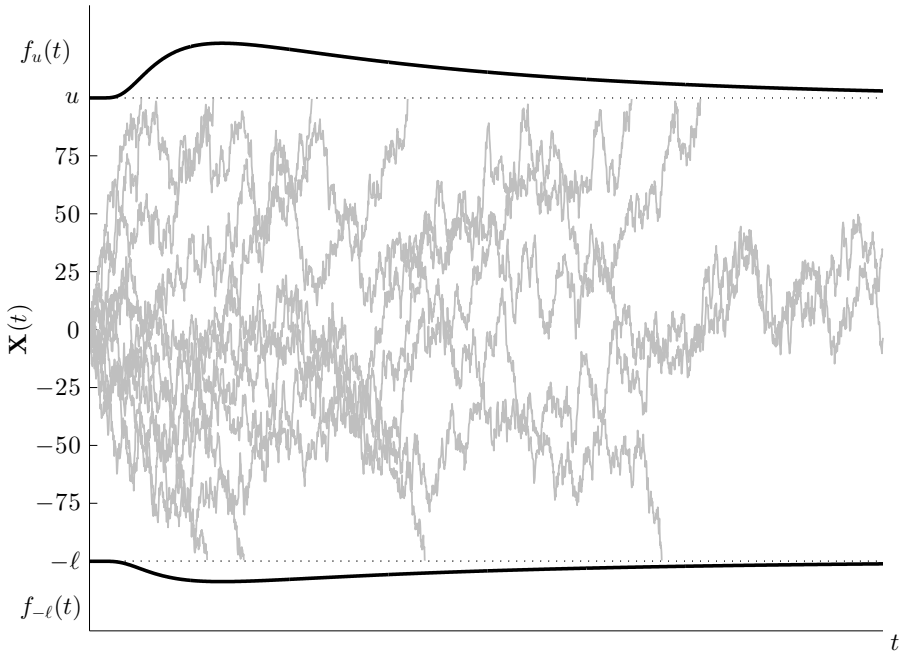


Abbildung 1.3: Der Wiener Prozess ($\mu = 0.4; \sigma = 9$) zwischen zwei absorbierenden Barrieren u und ℓ. Diese beiden Zustände können ebenfalls als Antwortkriterien für zwei Antwortalternativen aufgefasst werden. An den beiden Kriterien sind die respektiven Dichtefunktionen der Erstpassagezeit dargestellt. Die Wahrscheinlichkeit für die Absorption an der jeweiligen Barriere und damit die Wahrscheinlichkeit für die jeweilige Antwortalternative entspricht der Fläche unterhalb der jeweiligen Dichtefunktion. Dieser Prozess ist die Grundlage des Diffusionsmodells von Ratcliff (1978); das Superpositionsprinzip lässt sich auch auf diesen Wiener Prozess mit zwei Antwortkriterien anwenden.

gegeben. Dadurch, dass die Absorption an der unteren oder der oberen Barriere stattfinden kann, werden für die Vorhersage der Erwartungswerte und Varianzen der Detektionszeit die bedingten Erwartungswerte (Horrocks & Thompson, 2004) und Varianzen (Grasman, Wagenmakers & van der Maas, 2009) benötigt. Abgesehen davon kann das Superpositionsprinzip analog zum Ein-Barrieren-Fall angewendet werden, auch für den komplizierteren Fall zeitversetzter Reizdarbietung. Numerisch gestaltet sich die Berechnung der Modellvorhersagen in diesem Fall erheblich schwieriger, da für keine der benötigten Funktionsgleichungen (vgl. Gleichung 4, 5 und 6) geschlossene Ausdrücke bekannt sind.

Das DSM von Schwarz (1994) und die hier beschriebene Erweiterung stellen den theoretischen Rahmen dieser Arbeit dar. In bisherigen Arbeiten wurde lediglich das Wettlaufmodell als Erklärung für die beobachteten Redundanzeffekte ausgeschlossen; der genaue Integrationsmechanismus war bisher selten Gegenstand wissenschaftlicher Forschung (siehe aber Diederich, 1995). In den wenigen Arbeiten, in denen die Art der Integration genauer untersucht wurde, wurden Experimente zu Einfachreaktionen durchgeführt. Auch ist unklar, welche Rolle der Aufmerksamkeit bei der Integration zukommt. Im Zuge dieser Dissertation wurde das Integrationsprinzip auch bei komplexeren Aufgaben sowie der Einfluss von Aufmerksamkeit auf die Integration mulitsensorischer Reizinformation untersucht.

2 Übersicht der Originalarbeiten

Diese Dissertation ist als kumulative Dissertation ausgelegt; die zugrunde liegenden Arbeiten wurden in internationalen Zeitschriften veröffentlicht oder zur Veröffentlichung eingereicht. Im Folgenden sollen die Arbeiten kurz vorgestellt und insbesondere auf Untersuchungsgegenstand und Ergebnisse eingegangen werden, um einen Überblick über deren Zusammenhang geben zu können.

Multisensory processing of redundant information in go/no-go and choice responses

In der ersten Arbeit (S. 27–86) ging es um die Frage, inwiefern Reaktionszeiten bei Aufgaben, die komplexer sind als Einfachreaktionen, durch das Diffusions-Superpositionsmodell (vgl. Abschnitt 1.3) beschrieben werden können. Dazu wurden zwei Experimente durchgeführt: In Experiment 1 sollten die Probanden so schnell wie möglich auf bestimmte Zielreize reagieren und bei Distraktoren die Reaktion unterdrücken (Go/No-go-Aufgabe). In Experiment 2 sollten die Probanden auf bestimmte Reize so schnell und so genau wie möglich mit einer Antwortalternative, auf andere mit einer zweiten Antwortalternative reagieren (Wahlreaktionsaufgabe).

Das Design wurde in beiden Experimenten so gewählt, dass eine Unterscheidung von Reizenergie und Reizinformation möglich war. In klassischen Untersuchungen zum Redundanzeffekt war diese Unterscheidung nicht möglich, da Reizinformation und -energie konfundiert waren: in unimodalen Reizdurchgängen wurden stets weniger Informationen *und* Reizenergie präsentiert als in redundanten Reizdurchgängen (Miller, 1986). Die Erweiterung bestand darin, auditive und visuelle Neutralreize (oder: nicht-informative Reize; A_0 und V_0) zu verwenden, die mit keiner Antwort assoziiert waren und zusammen mit Go-Reizen sowie No-go-Reizen (A_+, V_+, A_-, V_-, Experiment 1) oder zusammen mit Zielreizen einer der beiden Antwortalternativen (A_L, V_L, A_R, V_R, Experiment 2) präsentiert wurden.

So wurden zum Beispiel auditive Go-Reize mit visuellen Neutralreizen kombiniert (A_+V_0); die korrekte Entscheidung hing in diesem Fall allein vom auditiven Zielreiz ab (vgl. Grice et al., 1984). Während sich die Reize hinsichtlich ihrer physikalischen Eigenschaften kaum voneinander unterschieden (z.B. Gabor-Reize leicht unterschiedlicher Ortsfrequenz), war die Relevanz für die Aufgabenstellung sehr unterschiedlich. Es ließen sich damit Rückschlüsse ziehen, ob der Diffusionsprozess im DSM tatsächlich die Antworttendenz repräsentiert, oder aber, wie von Schwarz (1994) vorgeschlagen, Stimulusenergie darstellen. In diesem Fall sollten die physikalisch recht ähnlichen Neutral- und Zielreize (bzw. Distraktoren in Experiment 1) im Modell sehr unterschiedliche Driftparameter besitzen. Die Neutralreize besaßen geringe Evidenz für die Entscheidung, die Driftparameter wären demnach ebenfalls gering. Für die Zielreize bzw. die Distraktoren wären (betragsmäßig) deutlich höhere Driftraten zu erwarten, die sich je nach Antwortalternative im Vorzeichen unterscheiden.

Dazu wurde das Modell von Schwarz (1994) zunächst um ein zweites Antwortkriterium erweitert (vgl. Abbildung 1.3). Diese Erweiterung folgt im Wesentlichen der Herleitung des Originalmodells, nur dass in diesem Fall die entsprechenden Ausdrücke für einen Wiener Prozess zwischen zwei absorbierenden Barrieren verwendet wurden. Das resultierende Modell wurde an die mittleren Reaktionszeiten, die Standardabweichungen der Reaktionszeiten der Reaktionen auf Go-Reize (Experiment 1) bzw. beider Antwortalternativen (Experiment 2) und an den Anteil korrekter Antworten angepasst. Die entsprechenden Ausdrücke sind deutlich komplexer als im Fall eines Antwortkriteriums. Eine Darstellung des Modells, auch von dessen Vorhersagen für zeitversetzt präsentierte Reize, ist im Anhang A (S. 69 ff) der hier beschriebenen Veröffentlichung zu finden.

Die Ergebnisse beider Experimente legen nahe, dass durch das Modell tatsächlich die Antworttendenz beschrieben wird. Das DSM konnte die mittleren Reaktionszeiten, die Reaktionszeitvarianz und den Anteil korrekter Antworten

in beiden Experimenten sehr genau vorhersagen und zwar nicht nur in den klassischen Bedingungen aus unimodalen und redundanten Zielreizen, sondern auch in den bimodalen, nicht-redundanten Bedingungen mit einer Neutralreizkomponente. Dies stützt zum einen die Annahme der additiven Superposition, zum anderen unterstützten die geschätzten Driftparameter des Modells die Annahme, dass durch diese tatsächlich der Beitrag der präsentierten Reize zum Entscheidungsprozess repräsentiert wurde. Sie war für die Neutralreize betragsmäßig deutlich geringer als für die Zielreize (Experiment 1 und 2) oder die Distraktoren (Experiment 1). Zudem wurde die Annahme paralleler Verarbeitung nach dem Wettlaufmodell auf verschiedene Weise widerlegt. Sowohl die klassische Analyse der Verteilungsfunktionen (race model inequality), als auch Kapazitätsüberlegungen (*capacity coefficient*: $C(t)$, Townsend & Nozawa, 1995) und eine explizite Modellierung durch ein Diffusions-Wettlaufmodell wiesen darauf hin, dass das Wettlaufmodell die Daten nicht oder nur unzureichend erklären konnte.

Der Beitrag der Koautoren bei dieser Veröffentlichung umfasst Überarbeitungen des ursprünglichen Manuskripts bis hin zum endgültigen, publizierten Artikel. Bei der Datenerhebung wurde ich von zwei studentischen Hilfskräften unterstützt.

Fast and accurate calculations for cumulative first-passage distributions in Wiener diffusion models

In dieser theoretischen Arbeit (S. 87–112) untersuchten wir mathematische Eigenschaften der Verteilungsfunktionen $F(t)$ der Erstpassage im Wiener Diffusionsmodell mit zwei absorbierenden Barrieren. Deren theoretische Form ist seit langer Zeit bekannt (z.B. bei Feller, 1968), darüber hinaus existieren zwei äquivalente Repräsentation dieser Funktion (Feller, 1968). Beide Repräsentationen basieren jedoch auf einer unendlichen Reihe. In der Vergangenheit wurden beide Repräsentationen verwendet (Ratcliff, 1978; Van Zandt, Colonius & Proctor, 2000), da keine

generell bessere Konvergenzeigenschaften besitzt. Allerdings unterscheiden sich die beiden Repräsentationen systematisch in ihrer Konvergenzeigenschaft bezüglich der Variable t: die eine konvergiert sehr gut für kleine Werte von t, die andere sehr gut für große Werte t.

Die Berechnung erfordert demnach die Auswertung einer unendlichen Reihe. Für die konkrete Anwendung in einem Diffusionsmodell muss die Berechnung an einer Stelle abgebrochen werden. Der Fehler, der durch den Abbruch entsteht, ist unbekannt. In dieser Arbeit konnte gezeigt werden, dass sich eine *obere Grenze* für diesen Fehler angeben lässt, wenn die Summation nach K Summmanden abgebrochen wird.

In der Arbeit wurden für beide Repräsentationen Abschätzungen des maximalen Fehlers hergeleitet, der sich durch den vorzeitigen Abbruch der Auswertung der jeweiligen Reihe bei Abbruch nach K Summanden ergibt. Durch Auflösen nach K ermöglicht diese Abschätzung eine Bestimmung der Anzahl der Summanden, die benötigt werden, um eine vorher definierte Genauigkeit zu erreichen. Dadurch war es außerdem möglich, die auf diese Weise bestimmte Anzahl von benötigten Summanden dazu heranzuziehen, diejenige der beiden Repräsentationen zu wählen, die bei einer gegebenen Parametrisierung für geforderte t die besseren Konvergenzeigenschaften besitzt. Eine derartige Vorgehensweise ermöglicht demnach nicht nur eine effektive Kontrolle des Berechnungsfehlers, sondern auch eine komputational effiziente Berechnung der Erstpassagezeit im Wiener Diffusionsmodell. Diese beiden Vorteile sind vor allem dann wichtig, wenn die Berechnung von Modellvorhersagen mittels numerischer Integration erfolgt (Ratcliff, 1978), was eine häufige Auswertung der Zielfunktionen erforderlich macht, und die Parameterschätzung mittels iterativer Verfahren durchgeführt wird, für die eine möglichst glatte Oberfläche der Zielfunktion hilfreich ist. Bei der Schätzung der Modellparameter des Zwei-Barrieren-Modells (siehe oben) wurde diese Berechnungsmethode angewendet. Die Veröffentlichung umfasste zudem einsatzbereiten Quelltext, der mittler-

weile in der Skriptsprache *R* zu einem Gesamtpaket zusammengefasst wurde und so für alle Benutzer zur Verfügung gestellt worden ist.

Der Beitrag der Koautoren dieser Veröffentlichung ist wie folgt: Der Artikel wurde von mir und Matthias Gondan verfasst; Miriam Kesselmeier wirkte bei allen Überarbeitungen bis hin zum publizierten Artikel mit. Die zugrunde liegenden Berechnungen wurden von allen drei Autoren durchgeführt, die Skripten im Anhang wurden von Matthias Gondan (*R*-Skript) und mir (MATLAB-Skript) angefertigt.

Cross-modal cueing effects on multisensory integration

In dieser Arbeit (S. 113–147) wurde der Einfluss visuell-räumlicher Hinweisreize auf die Integration redundanter Zielreize untersucht. Die zu erwartenden Cueing-Effekte sollten in einem gemeinsamen Modell mit einem zusätzlichen Parameter beschrieben werden. In wiederum zwei Experimenten wurden zum einen endogene räumliche Hinweisreize (Experiment 1) und zum anderen exogene räumliche Hinweisreize verwendet (Experiment 2). Endogene Hinweisreize zeichnen sich dadurch aus, dass sie symbolisch (z.B. durch Pfeile) den zu erwartenden Zielreizort anzeigen. Der darauf folgende Zielreiz muss anschließend mit größerer Wahrscheinlichkeit dort erscheinen, damit der Hinweisreiz effektiv ist; erscheint der Zielreiz an anderer Stelle, so ist die Reaktion im Mittel langsamer (Posner, 1980). Der Zugewinn eines gültigen Hinweisreizes und die Kosten eines ungültigen Hinweisreizes werden als Aufmerksamkeitseffekte ausgelegt (Posner, Snyder & Davidson, 1980; Bashinski & Bacharach, 1980): im Fall gültiger Hinweisreize erwartet der Proband den Reiz an dieser Stelle und richtet seinen Aufmerksamkeitsfokus auf diesen Ort aus. Erscheint der Reiz anschließend tatsächlich an dieser Stelle, ist die Reizverarbeitung effizient, d.h. die Extraktion der Reizinformation geschieht schneller. Erscheint der Zielreiz nach einem ungültigen Hinweisreiz dagegen an einer anderen Stelle, die nicht im Aufmerksamkeitsfokus liegt, ist die Verarbei-

tung dementsprechend ineffizienter. Diese Unterschiede führen zu Differenzen in der mittleren Reaktionszeit und werden als Cueing-Effekte bezeichnet.

Im Modell dieser Arbeit sollten die Cueing-Effekte auf die Integration audiovisueller Reize in einem gemeinsamen Modell dargestellt werden. Eine interessante Frage in diesem Zusammenhang war jedoch auch, ob beide Modalitäten gleichermaßen auf die Aufmerksamkeitsmanipulation ansprechen. Ist dies der Fall, spricht dies für ein supramodales Aufmerksamkeitssystem, in dem die räumliche Aufmerksamkeit modalitätsunabhängig ausgerichtet werden kann (Farah et al., 1989). Um diese Vorhersage empirisch zu testen, spezifizierten wir ein restriktives Modell mit einem Skalierungsfaktor $g \geq 1$ für modalitätsinvariante Aufmerksamkeitseffekte. Dieser Skalierungsfaktor wurde mit den Drifts und Diffusionsparametern μ_{A}, σ_{A}, μ_{V} und σ_{V} multipliziert. Er beschreibt somit die effizientere Informationsverarbeitung im auditiven und zugleich im visuellen Kanal. Dies ist eine starke Annahme, deren Plausibilität durch ein alternatives, liberaleres Modell mit modalitätsspezifischen Aufmerksamkeitseffekten untersucht wurde. Um modalitätsspezifische Aufmerksamkeitseffekte darstellen zu können, wurde ein zweiter Skalierungsfaktor als freier Parameter eingeführt. In diesem Modell wurde die Diffusionsparameter μ_{A} und σ_{A} bzw. μ_{V} und σ_{V} mit zwei separaten Skalierungsfaktoren g_{A} und g_{V} multipliziert. Die beiden Modelle waren hierarchisch geschachtelt, denn das restriktivere, modalitätsinvariante Modell ergibt sich aus dem liberaleren, modalitätsspezifischen Modell, wenn man im letzteren die Restriktion einführt, dass die beiden Skalierungsfakoren g_{A} und g_{V} gleich sein sollen. In diesem Fall ergibt sich das modalitätsinvariante Modell als Spezialfall. Diese Verschachtelung nutzten wir für die empirische Überprüfung beider Modelle gegeneinander mittels χ^2-Tests. Die Ergebnisse dieses Tests waren sehr aufschlussreich und zeigten einen deutlichen Unterschied zwischen den Aufmerksamkeitseffekten bei endogenen und exogenen Hinweisreizen. Die Verhaltensdaten bei endogenen Hinweisweizen ließen sich sehr gut mit dem Modell für modalitätsinvariante Aufmerksamkeitseffekte be-

schreiben und die Hinzunahme eines weiteren Parameters für modalitätsspezifische Aufmerksamkeitseffekte brachte keine substanzielle Verbesserung der Anpassung an die Daten. Ein gänzlich anderes Bild ergab sich bei der Anpassung des DSM an die Daten des Experimentes mit exogenen Hinweisreizen. Hier war die Anpassung des Modells mit modalitätsinvarianten Aufmerksamkeitseffekten nicht nur signifikant schlechter als jenes mit modalitätsinvarianten Aufmerksamkeitseffekten, es zeigte sich außerdem, dass die modalitätsspezifischen Effekte der Hinweisreize hauptsächlich darauf zurückzuführen waren, dass die Verarbeitung auditiver Reize im Gegensatz zur visuellen Reizverarbeitung nahezu unbeeinflusst von der Gültigkeit der Hinweisreize war ($g_{\mathrm{A}} \approx 1$).

Die Ergebnisse bei willentlicher Aufmerksamkeitsverschiebung mit endogenen Hinweisreizen (Experiment 1) standen damit im Einklang mit der Annahme eines supramodalen Aufmerksamkeitssystems, in dem beide Modalitäten gleichermaßen von der Gültigkeit des (visuellen) Hinweisreizes profitierten. Die Ergebnisse des Experimentes mit exogenen Hinweisreizen (Experiment 2) erweitern dagegen die Erkenntnisse aus Studien zur Aufmerksamkeitsmanipulation bei denen wiederholt Nulleffekte bei auditiven Zielreizen durch visuelle exogene Hinweisreize berichtet wurden (z.B. bei Spence & Driver, 1997). Dies weist auf eine Asymmetrie bei auditiven und visuellen Hinweisreizen hin, denn auditive Hinweisreize können sehr wohl die Verarbeitung von visuellen Zielreizen beschleunigen (Spence & Driver, 1994) – eine klare Verletzung der Supramodalitätsannahme. Mit der Anwendung von Reaktionszeitmodellen, denen unterschiedliche Annahmen bezüglich des Aufmerksamkeitsmechanismus zugrunde lagen, erbrachten wir starke empirische Belege für supramodale Aufmerksamkeitssteuerung im Falle endogener Hinweisreize und einer hauptsächlich visuell gesteuerten räumlichen Aufmerksamkeitsverschiebung im Falle exogener Hinweisreize. Diese Erkenntnis erweitert nicht nur bisherige Ergebnisse, sondern sie ist auch empirisch gefestigter, da bisherige Ergebnisse und

deren Interpretation oftmals auf den signifikanten oder nicht-signifikanten statistischen Testergebnissen einer einzigen experimentellen Bedingung fußten.

Die Koautoren dieses Manuskripts überarbeiteten frühere Versionen bis hin zum eingereichten Manuskript. Bei der Datenerhebung wurde ich von einer studentischen Hilfskraft unterstützt.

Effects of spatial and selective attention on basic multisensory integration

In der letzten Arbeit (S. 149–175) wurde untersucht, inwiefern das DSM Effekte einer dauerhaften Aufmerksamkeitsausrichtung abbilden kann. In Experimenten zur räumlichen Aufmerksamkeit (Experiment 1) und selektiven Aufmerksamkeit (Experiment 2) wurde der Frage nachgegangen, ob identische Zielreize, die an identischen Positionen präsentiert werden, unterschiedlich effizient wahrgenommen werden, und zwar abhängig davon, ob zusätzlich Reize an weiteren Positionen dargeboten werden (räumliche Aufmerksamkeit), oder abhängig davon, ob auf diese Zielreize nur reagiert werden soll, wenn sie alleine erscheinen, nicht jedoch wenn an alternativen Positionen weitere Distraktoren dargeboten werden (selektive Aufmerksamkeit).

Die Effekte der unterschiedlichen Aufmerksamkeitsbedingungen wurden – analog zur vorherigen Studie – durch eine Skalierung der Diffusionsparameter berücksichtigt. Diese Skalierung wurde anders als in der vorherigen Arbeit jedoch nicht durch einen expliziten Faktor modelliert, sondern implizit durch ein variables Antwortkriterium. Wie bereits erwähnt, ist das Antwortkriterium c im DSM kein freier Parameter, sondern ein fester Skalierungsfaktor. Dieser Umstand wurde in dieser Arbeit genutzt, um unterschiedliche Verarbeitungseffizienz in den experimentellen Bedingungen zu erfassen. Eine effizientere Verarbeitung ist mathematisch äquivalent mit einem geringeren Antwortkriterium, eine ineffizientere Verarbeitung mit einem höheren Antwortkriterium. Im ersten Experiment dieser Arbeit erwarteten wir, dass in der Bedingung mit nur einer Reizposition ein en-

gerer Aufmerksamkeitsfokus gewählt wurde, als wenn Reize an mehreren Stellen präsentiert werden. Unterschiedlich große Aufmerksamkeitsfoci sind in der Literatur mit unterschiedlicher Effizienz belegt, je größer der Aufmerksamkeitsfocus desto ineffizienter die Reizverarbeitung. Genau dies konnte in dieser Arbeit experimentell bestätigt werden: in der Bedingung, in der nur Reize in der Mitte präsentiert wurden, waren die mittleren Reaktionszeiten niedriger, als auf identische Reize in der anderen Bedingung, in der zusätzlich an zwei weiteren Positionen Reize erschienen, auf die ebenfalls reagiert werden sollte. Dieser Unterschied sollte durch ein gemeinsames Modell mit variablem Antwortkriterium beschrieben werden. Bei der Schätzung der Modellparameter des gemeinsamen Modells für beide Bedingungen erhielten wir tatsächlich unterschiedliche Antwortkriterien, die residualen Prozesse waren dagegen kaum betroffen. Wir interpretierten diese Ergebnisse als Aufmerksamkeitseffekte: in der Bedingung, in der ein engerer Aufmerksamkeitsfokus gewählt werden konnte, fokussierten die Probanden ihre Aufmerksamkeit auf die einzige Zielposition und die Verarbeitung der dort präsentierten Reize war effektiver, als wenn die Probanden ihre Aufmerksamkeit auf mehrere Positionen aufteilen mussten. Wie im ersten Experiment der dritten Arbeit waren die Aufmerksamkeitseffekte supramodal, d.h. beide Sinnesmodalitäten profitierten gleichermaßen von einer effizienteren Verarbeitung in der Bedingung mit engem Aufmerksamkeitsfokus.

Im zweiten Experiment unterschieden sich die beiden Bedingungen nicht nur in den Anforderungen an die Steuerung der Aufmerksamkeit, sondern auch in der Aufgabe der Probanden: in einer Bedingung sollte auf alle Reize reagiert werden, die, wie im ersten Experiment, lediglich an einer Position erschienen. In der zweiten Bedingung sollten die Probanden *nur* auf diese Reize reagieren und bei begleitenden Distraktoren ihre Antwort unterdrücken. Auch hier stellten sich die erwarteten Aufmerksamkeitseffekte ein: das Antwortkriterium war höher, d.h. die Reizverarbeitung weniger effektiv, wenn mehrere Reizpositionen beachtet werden

mussten, als wenn die Aufmerksamkeit selektiv auf die zentrale Position gerichtet werden konnte. Der Unterschied im Reaktionsmodus zwischen beiden Bedingungen führte außerdem zu einer vergrößerten Latenz der Residualprozesse bei der Go/No-go-Aufgabe (Gondan, Götze & Greenlee, 2010).

Die Veröffentlichung wurde von Matthias Gondan und mir erstellt und von allen Koautoren überarbeitet. Die Auswertung stammt zum überwiegenden Teil von Matthias Gondan. Die Datenerhebung wurde von Flavia Hughes, sowie einer studentischen Hilfskraft und mir durchgeführt.

3 Multisensory processing of redundant information in go/no-go and choice responses

Steven P. Blurton[a], Mark W. Greenlee[a], and Matthias Gondan[b]

[a]Institute for Psychology, University of Regensburg, Germany

[b]Department of Psychology, University of Copenhagen, Denmark

Abstract[1]

In multisensory research, faster responses are commonly observed when multimodal stimuli are presented, as compared to unimodal target presentations. This so-called *redundant-signals effect* can be explained by several frameworks, including separate-activation and coactivation models. The redundant-signals effect has been investigated in a large number of studies; however, most of those studies have been limited to the rejection of separate-activation models. Coactivation models have been analyzed in only a few studies, primarily using simple response tasks. Here, we investigated the mechanism of multisensory integration underlying go/no-go and choice responses to redundant auditory–visual stimuli. In the present study, the mean and variance of response times, as well as the accuracy rates of go/no-go and choice responses, were used to test a coactivation model based on the linear superposition of diffusion processes (Schwarz, 1994) within two absorbing barriers. The diffusion superposition model accurately describes the means and variances of response times as well as the proportions of correct responses observed in the two tasks. Linear superposition thus seems to be a general principle in the integration of redundant information provided by different sensory channels, and is not restricted to simple responses. The results connect existing theories of multisensory integration with theories on choice behavior.

Keywords: Multisensory processing; Math modeling; Decision making

[1]erschienen als: Blurton, S. P., Greenlee, M. W., & Gondan, M. (2014). Multisensory processing of redundant information in go/no-go and choice responses. *Attention, Perception, & Psychophysics*, *76*, 1212–1233.

Everyday perception simultaneously involves several senses. The sensory information relayed by different perceptual subsystems has to be integrated in order to take advantage of redundant and ancillary multisensory information, as well as to solve the problem of contradictory information. In speech perception, for example, lip reading can substantially improve the intelligibility of speech, especially in a noisy environment (Sumby & Pollack, 1954). Contradictory information from audition and vision, on the other hand, can evoke a percept presented neither verbally nor visually (McGurk & MacDonald, 1976).

Elementary multisensory perception is often studied with redundant signals from different senses—mostly audition, vision, and/or touch (Diederich, 1995; Giray & Ulrich, 1993; Miller, 1982, 1986). In redundant-signals experiments, the stimuli are presented either alone (unimodal conditions) or together (bimodal condition), and observers are usually asked to respond as quickly as possible to either signal (simple response task). A common observation is that response times (RTs) are, on average, lower in bimodal than in unimodal conditions. Different theoretical accounts can explain this so-called *redundant-signals effect*; these make different assumptions about the possible mechanism of integration of information from the different senses. The simplest form of multisensory "integration" for redundant signals is the separate- or parallel-activation account: Information from different sensory modalities is processed in separate channels, and the channel that first detects the stimulus triggers the response. According to this model, the redundant-signals effect is due to statistical facilitation; that is, RTs to redundant signals are, on average, lower because prolonged detection times in one channel are compensated for by the other channel (Raab, 1962).

Miller (1982, 1986) derived an upper bound for the redundancy gains that can be explained by separate-activation models. If the channel-specific processing time distributions are invariant across experimental conditions (i.e., context

invariance; see, e.g., Luce, 1986), Miller showed that the cumulative response time distribution [CDF; $F(t) = P(\mathbf{T} \leq t)$] for redundant audiovisual (AV) stimuli can never exceed the sum of the CDFs for the unimodal auditory (A) and visual (V) stimuli:

$$F_{\mathrm{AV}}(t) \leq F_{\mathrm{A}}(t) + F_{\mathrm{V}}(t), \text{ for all } t. \tag{1}$$

If Inequality 1 (the "race model inequality") is violated for any t, context-invariant separate processing is ruled out (see, e.g., Mordkoff & Yantis, 1991; Townsend & Nozawa, 1995, for a thorough discussion of alternatives). In multisensory studies on the redundant-signals effect, violations of Inequality 1 have often been reported (e.g., Diederich & Colonius, 1987; Miller, 1982, 1986; Mordkoff & Yantis, 1993).

In contrast to separate-activation models, coactivation models assume that information from different sensory systems is pooled into a common channel (e.g., Diederich, 1995; Miller, 1982, Appx. A; Miller, 1986; Miller & Ulrich, 2003; Schwarz, 1989, 1994). Detection is finished when activation in this common channel exceeds a certain level. Because coactivation models assume true intersensory facilitation rather than statistical facilitation, they can explain redundancy gains that exceed the upper bound imposed by the race model inequality (see Colonius & Townsend, 1997, for a theoretical framework of coactivation models).

A prominent framework of coactivation models is based on diffusion processes. The basic idea of diffusion models is that information accrual following stimulus presentation can be described by a noisy stochastic (diffusion) process in the presence of one absorbing barrier (Diederich, 1995; Schwarz, 1994). As soon as this barrier is reached, a response is initiated (Fig. 3.1a). Diffusion models are also used to describe human two-choice decisions (e.g., Busemeyer & Townsend, 1993; Diederich, 1997). To this end, diffusion processes between two absorbing

barriers are assumed, and absorption at one barrier corresponds to the choice of one of two alternatives (Fig. 3.1b). An important feature of diffusion models is that they predict not only the choice probability, but also the time that it takes to reach that decision. Also, trial-to-trial variability can be elegantly explained, because the accumulation process is inherently noisy. Diffusion models have been successfully demonstrated to explain data from experiments on memory retrieval (Ratcliff, 1978), lexical-decision tasks (Gomez, Ratcliff, & Perea, 2007), temporal-order judgments (Schwarz, 2006), numerical distance effects (Schwarz & Ischebeck, 2003), and decision making (Ratcliff & McKoon, 2008).

In the present study, we examined the performance of a two-choice diffusion model to explain choice behavior in a redundant-signals task. To this end, we combined the modeling approaches for the redundant-signals experiment (Diederich, 1995; Schwarz, 1994) with two diffusion models that have been proposed to describe human two-choice decisions (Busemeyer & Townsend, 1993; Diederich, 1997; Ratcliff, 1978). We present a model that can explain multisensory integration in tasks that are more complex than simple responses, but also, more generally, behavior in two-choice decisions, if stimuli are presented with an onset asynchrony. In two ways, this is a generalization of existing models for choice behavior and the processing of redundant targets. First, choice models are generalized by adding a second stimulus component that is relevant for the decision and that is possibly presented with a stimulus asynchrony. Second, existing models of multisensory integration with redundant signals are generalized for two response alternatives, so that multisensory integration can be tested with more complex, and arguably more ecologically valid, tasks than the mere detection of a stimulus.

Diffusion superposition model

The existing diffusion superposition models (DSM) for the redundant-signals effect have been used to explain data from simple response tasks (Diederich, 1995;

Schwarz, 1994), and, more precisely, tasks with one response alternative (Gondan, Götze, & Greenlee, 2010). In this model, the information accrual process is assumed to be a time-homogeneous diffusion process (i.e., a Wiener process). In a diffusion model for a binary choice response task, the absorbing barriers u and $-\ell$ represent the response criteria for the two response alternatives (Diederich, 1997; Ratcliff, 1978). Reaching one of the barriers corresponds to the choice between the possible alternatives (Busemeyer & Townsend, 1993).

In the context of the multisensory integration of audiovisual redundant signals, the channel-specific activity is described by diffusion processes with the parameters μ_{A}, σ^2_{A} and μ_{V}, σ^2_{V} for the auditory and the visual channels, respectively. Those parameters denote the drift and variance of both processes. This process evolves within two absorbing barriers that represent the response criteria for the two response alternatives. In the unimodal conditions, only one modality contributes to the accumulation of evidence. In the bimodal condition, the contributions of both stimuli superimpose additively; thus, when stimuli are simultaneously presented, the drift is $\mu_{\mathrm{AV}} = \mu_{\mathrm{A}}+\mu_{\mathrm{V}}$, and the variance is $\sigma^2_{\mathrm{AV}} = \sigma^2_{\mathrm{A}}+\sigma^2_{\mathrm{V}}+2\rho_{\mathrm{AV}}\sigma_{\mathrm{A}}\sigma_{\mathrm{V}}$. The parameter ρ_{AV} represents the correlation between activation in the sensory channels (Miller, 1991). In bimodal conditions, the drift is thus increased, so that the absorbing barrier $u > 0$ is, on average, reached earlier than in the unimodal case. This is most obvious in the case of simultaneously presented redundant signals, but the same principle applies to redundant signals presented with a stimulus onset asynchrony (SOA). If the targets are presented asynchronously, the process changes to a Wiener process with a new drift and new variance when the second stimulus is presented (Ratcliff, 1980; Schwarz, 1994). Again, the new drift is the sum of the single stimulus drifts (additive superposition). The process is then a compound of a first part, in which only the first stimulus is active, and a second part, in which the underlying buildup of evidence is governed by both stimuli (Fig. 3.1b), as in the case of simultaneous presentation. A detailed description of the

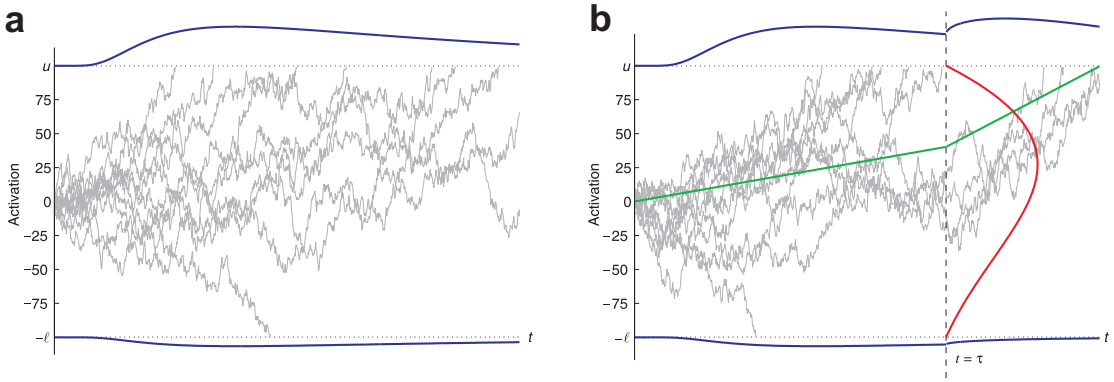


Figure 3.1: **a** Wiener process ($\mu = 0.6$, $\sigma^2 = 81$) between two absorbing barriers ($u = 100$, $-\ell = -100$) starting at $\mathbf{X}(0) = 0$. The curves at the top and bottomrepresent the first-passage densities $f_u(t)$ and $f_{-\ell}(t)$ for these parameters (on same arbitrary scale, for the purpose of comparison). The curves between the barriers represent random instances of the process. Due to positive drift, the probability of reaching the upper criterion first is greater than that of reaching the lower criterion. Absorption at either criterion corresponds to one of two response alternatives. **b** Wiener process ($\mu = 0.6$, $\sigma^2 = 81$) between two absorbing barriers ($u = 100$, $-\ell = -100$) starting at $\mathbf{X}(0) = 0$, with changing drift ($\mu = 1.8$) after time τ. The average position of a random process is displayed as a straight line to illustrate the increasing drift after τ. At time τ, a process may have already been absorbed or may still continue. The increasing drift affects the first-passage densities $f_u(t)$ and $f_{-\ell}(t)$ after time τ. The position density for processes that are not absorbed prior to time τ is displayed vertically.

model is given in the Method section, and explicit expressions are provided in Appendix A.

Most publications on the redundant-signals effect have used complex tasks but tested for violations of the race model inequality only. For example, neurophysiological studies have made frequent use of more-complex tasks, such as the go/no-go task (e.g., "oddball detection"; Senkowski, Talsma, Grigutsch, Herrmann, & Woldorff, 2007; Teder-Sälejärvi, McDonald, Di Russo, & Hillyard, 2002), and demonstrated that the race model cannot explain the redundancy gain, but they have not tested alternative models that could explain the redundancy gains

observed in the data. The limited number of studies that have actually tested alternative models have focused on simple response tasks (e.g., Diederich, 1995; Miller, 1982, Appx. A; Miller, 1986; Miller & Ulrich, 2003; Schwarz, 1994; but see Gondan et al., 2010).

In the present study, we investigated redundancy gains in choice and go/no-go tasks, and explicitly tested whether a coactivation model can account for the RTs and accuracy observed in these tasks. We then present a DSM with two absorbing barriers to test whether this model can account for the RTs and response accuracy observed in redundant-signals experiments with a go/no-go task (Exp. 1) and a two-alternative choice RT task (Exp. 2).

3.1 Method

In Experiment 1, participants made speeded responses in an audiovisual go/no-go task; in Experiment 2, they made speeded responses in a perceptual choice task. The visual stimuli were three different Gabor patches; the auditory stimuli were tones of three different frequencies. In addition to our go and no-go stimuli, we employed noninformative stimuli that contained no task-relevant information, but allowed for bimodal combinations of the stimuli without response conflicts. In Experiment 2, the experimental setup was largely the same as in Experiment 1, with a slight modification of the visual stimuli: The participants were given the task to choose between two response alternatives instead of refraining from responding on some trials.

Participants. Eight right-handed students from the University of Regensburg (all female; mean age 23.3 years) and one author participated in Experiment 1. Seven students (six female, one male; mean age 22.5 years) of the University of Regensburg and one author participated in Experiment 2. All of the participants had normal or corrected-to-normal vision and reported normal hearing. Students

were naive with respect to the purpose of the experiment and the employed stimulus conditions, and they were either paid (7€/h) or received course credit for participation. Prior to the start of the experiment, the participants gave their informed consent, and the experiment was conducted according to the Declaration of Helsinki.

Apparatus and stimuli. The experiments took place in a light- and sound-attenuated room (Industrial Acoustics Company GmbH, Niederkrüchten, Germany), which was dimly illuminated from behind and above. The participants sat on a chair in front of the computer screen, which was placed on a desk at a distance of 60 cm. Loudspeakers were placed at the left and right of the screen. Stimulus presentation was controlled with a standard IBM-compatible PC running Presentation software (Neurobehavioral Systems, Albany, California).

In Experiment 1, the visual stimuli were Gabor patches (sigma = 0.8 deg, 1.8 cycles/deg) presented centrally on a gray background (18 cd/m2). These Gabor patches (26 cd/m^2) were either horizontal (left-to-right grating) or tilted by 45 deg to the left or the right. In Experiment 2, we used three Gabor patches (sigma = 0.8 deg) with different spatial frequencies (1.1, 1.8, and 2.5 cycles/deg). Tones of three different frequencies were employed as auditory stimuli [370, 392, and 415.3 Hz at 35 dB(A), representing the tones F4#, G4, and G4# on a chromatic scale]. The Gabor patches and tones were presented alone (V, A) or together (AV). The intertrial interval (ITI) between two successive trials was randomly chosen from an exponential distribution with an expected value of 1,000 ms. To avoid either too-short or too-long ITIs, a fixed time interval (600 ms) was added to each randomly assigned ITI and the longest 5 % of intervals in the exponential distribution were truncated, respectively. Participants responded to these stimuli by pressing a response button placed at their dominant hand (Exp. 1) or buttons at both hands (Exp. 2).

Experimental tasks. In Experiment 1, the participants were asked to respond as quickly as possible if the high-frequency tone (A^+), the right-tilted Gabor patch (V^+), or both together (A^+V^+) were presented (go stimuli). The participants had 1.5 s to respond before the program proceeded with the next trial. Alternatively, the participants were asked to withhold a response if the low-frequency tone (A^-), the left-tilted Gabor-patch (V^-), or both together (A^-V^-) were presented (no-go stimuli). To avoid response conflicts in the bimodal stimulus conditions, the go-stimulus components were never presented together with no-go components (Table 1). Instead, we combined go and no-go components only with an intermediate-frequency tone (A^0) or a horizontal Gabor patch (V^0). These were regarded as noninformative or neutral stimuli, since they contained no information for the decision of whether to press the button or withhold a response in a given trial (see Grice et al., 1984, for a similar experimental setup). Critically, noninformative stimuli with go targets were presented with the same probability as noninformative stimuli paired with no-go stimuli (Table 1). All bimodal go stimuli were therefore either redundant (e.g., high-pitch tone and right-tilted Gabor patch) or nonredundant (a visual or auditory noninformative stimulus together with an auditory or visual go stimulus: A^+V^0 or A^0V^+). Together with the corresponding no-go trials (A^-V^-, A^-V^0, and A^0V^-) and four unimodal conditions (A^+, V^+, A^-, and V^-) there were ten different target conditions (cf. Table 3.1). Regardless of redundancy, bimodal stimuli consisted of one visual and one auditory stimulus component. All bimodal combinations V(τ)A were presented with the following SOAs: V33A, V67A, V100A, AV, A33V, A67V, and A100V (e.g., A67V means that the tone preceded the Gabor patch by 67 ms). In the following presentation, negative τ shall denote conditions A(τ)V.

In Experiment 2, participants were asked to choose as quickly as possible which of two buttons to press, depending on the stimulus. They should press the left button if a low-spatial-frequency Gabor patch (V_L) or a low-frequency

Table 3.1: *Experiment 1—Trial number of experimental conditions*

		Auditory			
		Go	Neutral	No-go	Null
Visual	Go	**7**	**7**	–	1
	Neutral	**7**	–	**7**	–
	No-go	–	**7**	**7**	1
	Null	1	–	1	–

Six bimodal conditions (bold numbers, each with seven SOAs) and four unimodal conditions were created. Each condition was employed once per block so that trial frequencies can be directly derived from the number of conditions. Trial frequencies with neutral stimuli were equal across targets (go) and nontargets (no-go). This scheme also applies to Experiment 2 if one replaces "go" and "no-go" with the two response alternatives.

tone (A_L) was presented. If the high-spatial-frequency Gabor patch (V_R) or the high-frequency tone (A_R) was presented, the participant should press the right response button. As in Experiment 1, left-response stimulus components were never combined with right-response stimulus components, but on some trials with noninformative stimulus components instead. The target stimuli (A_L, V_L, A_R, and V_R) were combined with Gabor patches (V_0) or tones (A_0) of intermediate frequency that were not associated with any alternative, and all combinations were equally likely. Bimodal stimuli for each side were either redundant (A_LV_L or A_RV_R) or contained either a noninformative auditory stimulus (A_0V_L or A_0V_R) or a noninformative visual stimulus (A_LV_0 or A_RV_0). Together with the unimodal conditions, we thus had ten different response relevance classes, five for each response alternative. Again, the intermediate stimuli (A^0, V^0) were regarded as noninformative, since they contained no information about the required response and were uncorrelated with the correct response. Due to SOA variation, bimodal

stimulus trials were more frequent than unimodal stimulus trials, but the frequencies of trials with bimodal stimuli were the same for all possible combinations of response relevance (Table 3.1). The same SOA variation as in Experiment 1 was used for the bimodal stimuli.

Procedure. Each participant was tested in nine sessions, each of which lasted about an hour. At the beginning of each session, all six stimulus components were presented, and the participants were reminded of the relevance of the stimuli for the respective tasks. A single session consisted of 12 blocks, with a break after the first six blocks. Each block was preceded by a task instruction and contained one trial of each bimodal condition [V(τ)A in the go and no-go combinations] and one trial of each of the unimodal conditions, resulting in a total of 46 trials. For Experiment 1, each participant was tested in 102 blocks, so that the number of replications was 102 in all 46 stimulus conditions.

The procedure in Experiment 2 was roughly the same as in Experiment 1. Each block comprised one repetition of each condition [$V^L(\tau)A^L$, $V^L(\tau)A^0$, $V^0(\tau)A^L$, $V^R(\tau)A^R$, $V^R(\tau)A^0$, $V^0(\tau)A^R$, A^L, V^L, A^R, and V^R], and thus, also 46 trials. The trial frequencies can be deduced from Table 1, if one replaces "go" and "no-go" with "left" and "right", respectively. Each condition had 108 replications. We ran a test block at the beginning of each experiment to ensure that the required task had been understood, especially the response relevance of our bimodal targets. In this test block, participants had more time to respond and were given visual feedback ("correct" or "false" displayed on the screen). The data from this block were discarded.

Tests of race model predictions. We tested the race model predictions on two grounds: the race model inequality (Eq. 1) and capacity considerations. The latter involved testing an index for process capacity $C(t)$ that had been introduced by Townsend and colleagues (e.g., Houpt & Townsend, 2012; Townsend & Altieri,

2012; Townsend & Nozawa, 1995). The capacity index is a measure that provides a time-dependent assessment of processing capacity and can be calculated from the estimated cumulative hazard function $H(t)$,

$$C(t) = H_{\mathrm{AV}}(t)/[H_{\mathrm{A}}(t) + H_{\mathrm{V}}(t)]. \tag{2}$$

If the assumption of parallel first-terminating processing with unlimited capacity and independent racers holds, this measure (Eq. 2) equals unity. In other words, the independent race model serves as a benchmark to which the observed performance is compared. If $C(t) > 1$ (supercapacity), capacity increases with increasing load. According to the same reasoning, Townsend and Altieri (2012) recently generalized the capacity index for more-complex tasks, such as the two-choice task used in Experiment 2. Hence, this generalized assessment index $A(t)$ was used as the capacity measure in Experiment 2.

In both experiments, erroneous responses might occur because the participant made guesses instead of informed responses. The test of the race model has been shown to be refined if the "twins" of these guessing times are subtracted from the RT distribution (Eriksen, 1988; Gondan & Heckel, 2008). We determined the "twins" separately for each SOA and response condition. For Experiment 1, this meant that for each response to a redundant V(τ)A no-go trial (guess), one response to a redundant V(τ)A go was cancelled out (set to infinity). The correspondence was established by choosing the response that was as close as possible to that of the guess, regarding its latency. For auditory unimodal targets, twins were determined by the auditory no-go condition, and for the visual unimodal target condition by the visual unimodal no-go condition. For condition V(τ)A, the race model then predicts

$$[F_{\mathrm{V+(\tau)A+}}(t) - F_{\mathrm{V-(\tau)A-}}(t)] \leq [F_{\mathrm{V+}}(t) - F_{\mathrm{V-}}(t)] + [F_{\mathrm{A+}}(t-\tau) - F_{\mathrm{A-}}(t-\tau)], \text{ for all } t. \tag{3}$$

In Experiment 2, the RT distributions of one alternative were corrected by incorrect responses to the other alternative. Thus, for example, "left" RT distributions were corrected for (incorrect) "left" responses to stimuli for which a "right" response would have been correct.

Miller (1986) suggested quantifying the amount of violation of this inequality by the positive area Δ_τ enclosed by the observed AV distribution and the sum of the observed A and V distributions. We pooled data from all sessions so that responses were included in each CDF. The violation areas of all SOA conditions were measured and aggregated to a single violation area Δ by a weighted sum. The weights denote a triangular weighting function; we assigned weights of 1, 2, 3, 4, 3, 2, and 1 to conditions A100V, A67V, A33V, AV, V33A, V67A, and V100A, respectively (a "symmetric umbrella"; Gondan, 2009). This means, for example, that the violation area in condition A33V was weighted three times as much as that of condition A100V. Though many studies have reported redundancy gains to be most pronounced when the visual precedes the auditory signal by some moderate SOA (e.g., Miller, 1986; Senkowski et al., 2007), we used a symmetric function because according to the average RTs to unimodal A and V trials, we expected violations of the race model inequality to be most pronounced in the synchronous stimulus presentation condition. We then applied Miller's (1986) bootstrap test to assess whether the violation area for a given participant reflected coactivation or was due to sampling error. Bootstrap samples of the unimodal RTs were drawn from the observed RT distributions; bimodal RTs were bootstrapped from the distribution of minima of the unimodal RTs, adjusted for SOA and assuming a maximally negative channel correlation between A and V (Ulrich & Giray, 1986). In each simulation, the aggregate violation area Δ^* was determined, and this simulation was repeated 10,001 times, resulting in a reference distribution of Δ under the race model assumption. The race model is significantly violated at $\alpha = .05$ (one-tailed) if the observed Δ is greater than 95 % of the simulated Δ^*.

The average amount of race model violation (supercapacity) was quantified by the capacity index (Exp. 1) and the assessment index (Exp. 2). The statistical significance of race model violations and supercapacity was tested at the group level using a permutation test (Gondan, 2010). For these tests, we used the 5th through 30th RT percentiles, in steps of 5 %. Since the respective indices should be one for all time points, under standard race model assumptions, the capacity estimates of the RT percentiles were aggregated using a $T_{\max}$ statistic (cf. Gondan, 2010).

The diffusion superposition model with two absorbing barriers. Standard diffusion models on choice behavior assume that the information accumulation process can be described by a time-homogeneous diffusion process (i.e., Wiener process) $\mathbf{X}(t)$ between two absorbing barriers u and $-\ell$. The process has drift μ and variance $\sigma^2 > 0$ and starts at $\mathbf{X}(0) = 0$. In the following discussion, absorption at the upper barrier and the lower barrier corresponds to the mutually exclusive events $\mathbf{C} = 1$ and $\mathbf{C} = 0$, respectively. For redundant signals, additive superposition of channel activity is assumed: $\mathbf{X}_{\mathrm{AV}}(t) = \mathbf{X}_{\mathrm{A}}(t) + \mathbf{X}_{\mathrm{V}}(t)$, resulting in a new Wiener process with drift $\mu_{\mathrm{AV}} = \mu_{\mathrm{A}} + \mu_{\mathrm{V}}$ and variance $\sigma^2_{\mathrm{AV}} = \sigma^2_{\mathrm{A}} + \sigma^2_{\mathrm{V}} + 2\rho_{\mathrm{AV}}\sigma_{\mathrm{A}}\sigma_{\mathrm{V}}$. The decision time $\mathbf{D}$ corresponds to the first-passage time; that is, as soon as either barrier is reached, a response is produced. To predict the mean detection times and response accuracy, thus, six parameters are needed ($-\ell$, μ_{A}, σ^2_{A}, μ_{V}, σ^2_{V} , and ρ_{AV}). The upper barrier u can be fixed (e.g., at 100) because it is only a scaling parameter. Model predictions for mean and variance of detection time can be derived from the first-passage distributions and moments of a Wiener process between two absorbing barriers (e.g., Cox & Miller, 1965). Expressions for the expected values of $\mathbf{D}$ are summarized in Horrocks and Thompson (2004, their Eqs 6–10), and those for the variance in Grasman et al. (2009, their Eq. 14). Reaching the upper criterion occurs with some probability $P(\mathbf{C} = 1)$,

and the complementary event—that is, reaching the lower criterion—occurs with a probability of $P(\mathbf{C} = 0) = 1 - P(\mathbf{C} = 1)$. Expressions for the probabilities of reaching the lower and upper criteria are also summarized in Horrocks and Thompson (2004, their Eqs. 5 and 9). It is commonly assumed (Luce, 1986) that the observed response time $\mathbf{T}$ is the sum of the signal detection $\mathbf{D}$ and condition-invariant residual processes $\mathbf{M}$ (e.g., finger movement). Thus, another parameter (μ_{M}) is necessary, which represents the mean latency of all residual processes $\mathbf{M}$ (Townsend & Honey, 2007). The prediction of the RT variance requires two additional parameters—namely, the variance of the latency of residual processes (σ^2_{M}) and the correlation between residual processes latency $\mathbf{M}$ and detection time $\mathbf{D}$ (ρ_{DM}). The correlation parameter ρ_{DM} did not substantially improve the model fits, so it was fixed at zero in the models presented.

With all parameters specified, the two-barrier model predicts mean RTs (i.e., $E[\mathbf{T}|\mathbf{C} = 1]$ and $E[\mathbf{T}|\mathbf{C} = 0]$), variance ($Var[\mathbf{T}|\mathbf{C} = 1]$ and $Var[\mathbf{T}|\mathbf{C} = 0]$) and response probabilities ($P[\mathbf{C} = 1]$ and $P[\mathbf{C} = 0]$) in a go/no-go or a binary choice RT task. For stimuli presented with onset asynchrony τ, the probability of reaching the upper criterion first, $P(\mathbf{C} = 1|\tau)$, the expected value of the time of first passage $E(\mathbf{T}|\mathbf{C} = 1, \tau)$ and its variance $Var(\mathbf{T}|\mathbf{C} = 1, \tau)$, have to be calculated conditional on the SOA. The derivations of SOA-dependent expected values $E(\mathbf{T}|\mathbf{C} = 1, \tau)$ and $E(\mathbf{T}|\mathbf{C} = 0, \tau)$, of the variances $Var(\mathbf{T}|\mathbf{C} = 1, \tau)$ and $Var(\mathbf{T}|\mathbf{C} = 0, \tau)$ of the first-passage times, and of the probabilities $P(\mathbf{C} = 1|\tau)$ and $P(\mathbf{C} = 0|\tau)$ of reaching either barrier first are described in Appendix A. This two-choice diffusion model, even though it is able to predict also response accuracy, needs only one parameter more than the original DSM (Schwarz, 1994)—namely, the second absorbing barrier $-\ell$. In the case of an unbiased model, the barriers are equidistant from the origin. On the other hand, by setting this parameter far apart from the origin ($\ell >> u$), the probability of absorption at the upper barrier tends to unity, and the two-barrier model predictions for the expected value and

variance of **D** correspond to those of the one-barrier DSM.

Test of the diffusion model. We employed the two-barrier model to predict mean and variance of go trials and the amount of incorrect responses observed in the no-go stimulus conditions of Experiment 1 and both response alternatives of Experiment 2. In a go/no-go task, the lower barrier is assumed to be implicit, because absorption at the lower barrier does not lead to an observable response. An observable response is only initiated if the upper barrier is hit. On this note, a false-positive response to a no-go stimulus corresponds to reaching the *upper* barrier u in a no-go trial (i.e., $\pi = P[\mathbf{C} = 1|u, -\ell, \mu_{\mathrm{A}-}, \sigma^2_{\mathrm{A}-}, \mu_{\mathrm{V}-}, \sigma^2_{\mathrm{V}-}, \rho_{\mathrm{AV}}]$). To predict the mean and standard deviation of RTs of the nonredundant stimuli with a noninformative auditory or visual component, additional diffusion parameters were necessary for the noninformative components ($\mu_{\mathrm{A}0}$, $\sigma^2_{\mathrm{A}0}$ and $\mu_{\mathrm{V}0}$, $\sigma^2_{\mathrm{V}0}$ for the noninformative auditory and visual components, respectively). In order to obtain predictions for the number of false alarms in (redundant and nonredundant) no-go trials, diffusion parameters for these stimuli were also needed: $\mu_{\mathrm{A}-}$, $\sigma^2_{\mathrm{A}-}$ for the auditory no-go component, and $\mu_{\mathrm{V}-}$, $\sigma^2_{\mathrm{V}-}$ for the visual no-go component. Note that, contrary to the test of the race model inequality, no false alarms were removed from the set of observed RTs, because false alarms can be explicitly predicted by the model. This is a direct consequence of the noise in the diffusion processes, which is also responsible for trial-to-trial variation. With ρ_{DM} set to zero, 16 parameters are needed to derive 69 predictions for the mean and variance of RTs of all go trials and the percentage of responses to no-go trials.

The model fit for Experiment 2 was analogous, except that we added two parameters for the residual process mean latency and its variance of the second response alternative (μ_{ML}, σ^2_{ML}). The parameter for the lower criterion turned out to be approximately the same as that for the fixed upper barrier in most participants of Experiment 2, so it was also fixed at $-\ell = -100$, yielding equal decision criteria for left and right responses (unbiased model; Wagenmakers, van der Maas, &

Grasman, 2007). This resulted in a model with 17 parameters to predict the mean and variance of RTs and the response accuracy of both response alternatives—that is, 138 predictions in total.

The model fits were assessed by a goodness-of-fit statistic that was the sum of squares of the normalized differences between the DSM model predictions and the observed data—that is, the observed mean and variance of RTs, as well as the observed response frequencies in no-go trials (Exp. 1) and for both response alternatives (Exp. 2):

$$X^2_{\mathrm{m}} = \sum_{\tau} \{m_{\mathrm{V}(\tau)\mathrm{A}} - E[\mathbf{T}_{\mathrm{V}(\tau)\mathrm{A}}]\}^2 / [\hat{s}^2_{\mathrm{V}(\tau)\mathrm{A}} / N_{\mathrm{V}(\tau)\mathrm{A}}], \tag{4}$$

$$X^2_{\mathrm{s}} = \sum_{\tau} \{\hat{s}_{\mathrm{V}(\tau)\mathrm{A}} - \sqrt{Var[\mathbf{T}_{\mathrm{V}(\tau)\mathrm{A}}]}\}^2 / \{\hat{s}^2_{\mathrm{V}(\tau)\mathrm{A}} / [2 \times N_{\mathrm{V}(\tau)\mathrm{A}}]\}, \tag{5}$$

$$X^2_{\mathrm{p}} = \sum_{\tau} N_{\mathrm{V}(\tau)\mathrm{A}} \times \left\{p_{\mathrm{V}(\tau)\mathrm{A}} - \pi_{\mathrm{V}(\tau)\mathrm{A}}\right\}^2 / \{p_{\mathrm{V}(\tau)\mathrm{A}} \times [1 - p_{\mathrm{V}(\tau)\mathrm{A}}]\}, \tag{6}$$

$$\tau \in \{0, \pm 33, \pm 67, \pm 100, \pm \infty\}.$$

Negative τ denotes conditions A(τ)V, and $\tau = \pm\infty$ denotes the unimodal visual and auditory stimuli. Thus, the goodness-of-fit statistic $X^2 = X^2_{\mathrm{m}} + X^2_{\mathrm{s}} + X^2_{\mathrm{p}}$ was the unweighted sum of normalized differences between predicted and observed means and the standard deviations of RTs for go trials (Eqs. 4 and 5), plus the normalized differences between the predicted and observed response probabilities (Eq. 6). We fitted the model to the means and standard deviations of RTs and not to the complete distribution, because the distribution of residual processes $\mathbf{M}$ was not known. Fitting the first moments of this distribution, on the other hand, only required the first moments of the base time distribution, which we included as free parameters.

If the model holds, the goodness-of-fit statistics (Eqs. 4–6) asymptotically follow χ^2 distributions. Note that the distribution of the sum of Eqs. 4–6, though

it is a sum of squared normalized differences, does not correspond to a standard χ^2 statistic, because the means (Eq. 4), variances (Eq. 5), and response accuracy (Eq. 6) are not independent. For Experiment 1, using the proportion of responses to no-go stimuli extracted the best information available from go and no-go trials, since the accuracy rate in go trials was at 100 % in most participants. However, in order to prevent implausible predictions for the numbers of correct responses in go trials, we included the percentage of responses in go trials as a side condition. Though the percentages of responses to go trials do not contribute to the X^2 statistics reported in Table 3.3 below, they were added as a further constraint in the fitting procedure. In order to fit the model to response frequencies predicted to be one or zero, the variance was calculated as if the response frequency were .01 or .99, respectively. In this way, we prevented the variance from becoming zero in the extreme cases. The functions were implemented in MATLAB (MathWorks, Natick, MA), and the parameters were estimated using the optimizer fminsearchbnd (downloadable from MATLAB Central) to search for the minimum X^2 value.

The diffusion race model with two absorbing barriers. The DSM can be directly compared to a diffusion race model (DRM). In the DRM, the assumption of time-homogeneous channel-specific diffusion processes is kept, but the combined activity $\mathbf{X}_{\mathrm{AV}}(t)$ does not correspond to $\mathbf{X}_{\mathrm{A}}(t) + \mathbf{X}_{\mathrm{V}}(t)$. Rather, it is assumed that the first process to reach the criterion initiates the response—that is, a parallel independent race model. The resulting RT distribution is then a mixture distribution determined partly by the auditory and partly by the visual channel. The derivation is outlined in Appendix B. The DRM was fitted analogously to the DSM by minimizing the same goodness-of-fit statistic X^2—that is, the sum of Eqs. 4 to 6.

3.2 Experiment 1

In Experiment 1, participants made speeded responses in an audiovisual go/no-go task with redundant signals. Response latencies and task performance weremeasured and used to test whether separate-activation models or the diffusion superposition model could better account for the observed data.

3.2.1 Results

Response times and task performance. The SOA-dependent mean RTs exhibited the typical wing-shaped pattern, with the mean RTs to synchronous stimuli being lower than those to any stimuli presented with an SOA (Fig. 2, upper panel). Task performance, as measured by false alarms to no-go stimuli, varied substantially between participants. Four participants (i.e., 2, 3, 4, and 9) made very few errors. Responses to no-go stimuli were less than 10 % in these participants, regardless of the stimulus combination (redundant or nonredundant no-go) and SOA. In the remaining participants, we observed effects of SOA and the modality of the target stimulus component (A^+V^+, A^0V^+,A^+V^0) on the observed error rates in bimodal nonredundant trials (for participant-specific results, please see Supplemental Fig. 3.1 in the online supplemental materials). If the noninformative stimulus preceded the no-go stimulus, higher SOAs led to higher response frequencies. Error rates were also lower for bimodal redundant signals than for nonredundant signals, with no visible effect of SOA. The observed standard deviations were largely independent of SOA, but they differed between response conditions (bimodal redundant, bimodal nonredundant). A speed–accuracy tradeoff was also evident: Regarding their mean RTs, Participants 1 and 8 were fastest, and both showed the highest error rates, which amounted to 30 % and more in some conditions. On the other hand, four participants (2, 3, 4, and 9) did not show any, or only very few, errors, and three of those four participants have the slowest mean RTs of all participants (i.e., Participants 2, 3, and 4). Participant 9

made very few errors together with relatively fast mean RTs. Errors (lapses) in go trials were rare, amounting to about 2 % for all participants, with the exception of Participants 2, who missed 9 % of the targets in one condition, and 4, who shows frequent misses (up to 15 %) in the auditory go conditions (unimodal and bimodal nonredundant targets).

Tests of race model predictions. In a first analysis, we calculated the violation area Δ as the weighted sum of the violation areas of all SOA conditions (Gondan, 2009) and determined the participant-specific p values by the bootstrap procedure described in Miller (1986). The race model inequality was significantly violated in every single participant, thereby ruling out the classical race model (i.e., context-invariant parallel processing) as an explanation of the redundancy gains in the go/no-go task (Table 3.2).

Table 3.2: *Experiment 1—Result of race model inequality tests*

	Participant								
	1	2	3	4	5	6	7	8	9
Δ	96.8	103.2	146.1	62.8	60.4	151.6	138.0	72.7	75.8
p value	$< .001$	$< .001$	$< .001$	.039	.035	$< .001$	$< .001$	$< .001$	$< .001$

Δ: violation area up to the 20th percentile; p value: proportion of simulations (under race model assumptions) that showed violation areas greater than the observed violation area Δ.

In line with this finding, the capacity coefficient $C(t)$ indicated supercapacity over a large range of t (Supplemental Fig. 3.2A). The group test for the capacity index $C(t)$ at early t yielded significant supercapacity ($T_{\max} = 7.96$, $T_{0.95} = 2.27$; $p < .001$). The individual time courses of processing capacity are available in the supplemental material (Supplemental Fig. 3.2A). Average violations of the standard race model inequality across all participants (cf. Mordkoff & Danek, 2011) are displayed in Supplemental Fig. 3.2B.

Table 3.3: *Experiment 1—Parameter estimates and model fit of the diffusion superposition model with two criteria for mean and variance of response times of go trials together with amount of responses to no-go trials*

		Participant								
Parameter		1	2	3	4	5	6	7	8	9
Go	μ_{V+}	0.56	0.76	0.77	0.85	0.69	0.48	0.48	0.54	0.81
responses	σ^2_{V+}	4.4	11.3	1.0	1.6	14.3	1.2	1.0	16.5	12.5
	μ_{A+}	0.78	0.47	0.66	0.55	1.08	0.45	0.44	0.39	0.76
	σ^2_{A+}	9.7	6.3	1.0	15.9	49.1	4.3	1.8	3.0	12.6
	μ_M	293.4	448.8	642.0	530.1	416.8	360.4	350.0	238.6	354.3
	σ^2_M	2,339	3,472	11,917	11,271	1,451	6,466	11,457	1,380	5,448
	ρ_{DM}	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
	u	(100)	(100)	(100)	(100)	(100)	(100)	(100)	(100)	(100)
Noninfor-	μ_{V0}	0.15	0.09	0.24	0.13	0.11	0.12	0.16	0.16	0.01
mative	σ^2_{V0}	9.3	8.4	14.4	3.3	6.1	6.1	5.4	12.1	4.4
Stimli	μ_{A0}	0.18	0.15	0.21	0.14	0.29	0.19	0.15	0.29	0.07
	σ^2_{A0}	8.3	5.8	1.0	1.3	7.9	8.3	6.5	2.7	9.1
No-go	μ_{V-}	–1.07	–0.28	–1.68	–1.43	–2.31	–1.83	–0.60	–1.42	–2.00
responses	σ^2_{V-}	80.3	1.0	64.9	59.5	95.6	82.4	25.4	69.0	66.9
	μ_{A-}	–0.39	–5.35	–3.55	–0.22	–0.30	–0.26	–0.49	–0.50	–2.00
	σ^2_{A-}	4.9	243.9	2.8	1.0	1.0	1.1	25.7	36.2	72.4
	$-\ell$	–111	–69	–63	–64	–512	–88	–401	–76	–101
	ρ_{AV}	0.95	0.26	0.20	0.83	0.81	0.90	0.04	0.15	0.18
GOF X^2	DSM	128.0	110.5	90.7	104.4	87.9	89.0	79.7	84.4	144.7
	DRM	148.3	113.9	102.0	115.5	93.0	124.8	91.3	98.5	147.6

μ_*, σ^2_*: drift and variance for visual (auditory) diffusion process for go stimuli (μ_{V+}, σ^2_{V+}, μ_{A+}, σ^2_{A+}), noninformative stimuli (μ_{V0}, σ^2_{V0}, μ_{A0}, σ^2_{A0}), and no-go stimuli (μ_{V-}, σ^2_{V-}, μ_{A-}, σ^2_{A-}); u: evidence criterion for response; $-\ell$: evidence criterion for no response; ρ_{AV}: correlation between $\mathbf{D}_A$ and $\mathbf{D}_V$; μ_M, σ^2_M: mean and variance of residual component duration $\mathbf{M}$; ρ_{DM}: correlation between $\mathbf{D}$ (detection time) and $\mathbf{M}$ (residual component duration). Parameters in parentheses are fixed parameters. GOF X^2 : goodness-of-fit statistic (higher results indicate worse fit).

Diffusion superposition model. A diffusion superposition model with two absorbing barriers (u fixed at 100) and 16 free parameters was fitted to the mean RTs and variances of RTs in a go/no-go task (69 predictions in total). Figures that display participant-specific performance and DSM predictions are available as supplemental materials (Supplemental Fig. 3.1). Except for one participant, the model fit was acceptable to excellent in all participants. The diffusion superposition model captured all aspects of the go/no-go data well. This was primarily the wing-shaped pattern of the SOA-dependent mean RTs (Fig. 3.2; circles) in the redundant-signals conditions. Additionally, the model had to explain the data of the nonredundant bimodal conditions with a noninformative stimulus. This included the rising mean RTs in the nonredundant noninformative stimulus-first conditions (Fig. 3.2; auditory targets shown by upward pointing triangles, visual targets by downward pointing triangles). Responses to redundant no-go stimuli were rare (lower panel of Fig. 3.2; circles). Responses to no-go stimuli accompanied by noninformative stimuli were more frequent in most participants (e.g., Participants 1, 4, 5, 6, 7, and 8; see Supplemental Fig. 3.1) and were generally increased with the delay of the no-go stimulus (Fig. 3.2; auditory no-go stimuli shown by upward pointing triangles, visual no-go stimuli by downward pointing triangles). Standard deviations, which were rather constant across all conditions, were also predicted well by the model. This was also true for the numbers of correct responses in all go conditions, which attained 100 % in most participants and conditions (Fig. 3.2, top of the lower panel). Only the model fit for the data of Participant 9 exhibited frequent deviations from the observed mean RTs and standard deviations, especially in the bimodal nonredundant conditions $\mathrm{V}^{+}(\tau)\mathrm{A}^{0}$ and $\mathrm{V}^{0}(\tau)\mathrm{A}^{+}$.

As can be seen in Table 3.3, the drifts of both noninformative stimulus components ($\mu_{\mathrm{A}0}$, $\mu_{\mathrm{V}0}$) were generally lower than those of the target stimulus components ($\mu_{\mathrm{A}+}$, $\mu_{\mathrm{V}+}$). The reverse is true for the drifts of the no-go stimuli

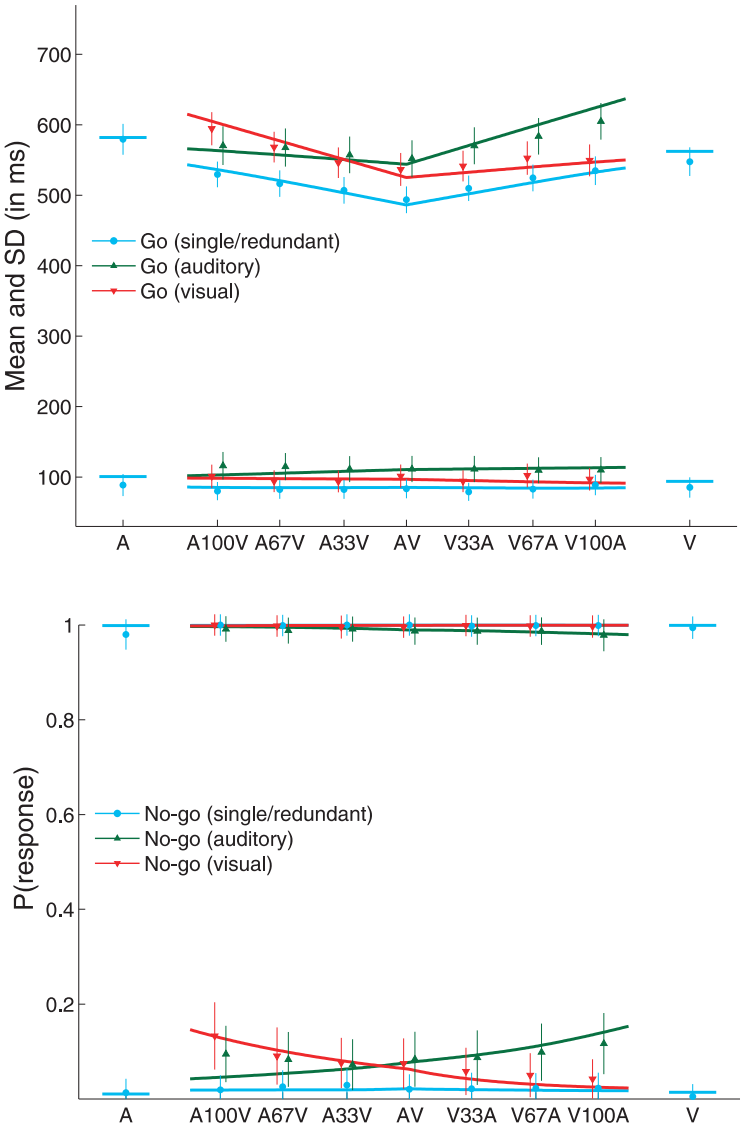


Figure 3.2: Averaged mean response times and standard deviations of go trials (*upper panel*), as well as proportions of responses to go and no-go trials (*lower panel*), together with model predictions from the diffusion superposition model for all nine participants in Experiment 1. The curves represent averaged model predictions for unimodal/bimodal redundant, nonredundant auditory-target, and nonredundant visual-target stimulus conditions. Data from the redundant bimodal conditions and unimodal conditions are displayed as circles, and data from the bimodal nonredundant conditions with either visual or auditory noninformative stimuli are displayed as upward-pointing triangles or downward-pointing triangles, respectively. On the x-axis, conditions are ordered from left to right according to stimulus onset asynchrony, and error bars denote 95% confidence interval based on averaged standard errors. Participant-specific data and model fits are available as online supplemental materials (Supplemental Fig. 3.1).

(μ_{A-}, μ_{V-}), which were negative in all participants and generally below the noninformative target drifts. The parameter estimates for the lower barrier differed considerably from those for the upper barrier (100) in most participants, suggesting a response bias in those participants. This is also evident in estimates for drifts of the noninformative stimuli; when these were fixed to zero, the model fit was considerably worse. In three participants, the estimate for the lower barrier is farther away from the origin than that for the upper barrier (Participants 1, 5, and 7), and in five participants, the estimated lower barrier is closer to the origin than the upper barrier (Participants 2, 3, 4, 6, and 8).

Diffusion race model. The diffusion race model (DRM) failed to explain the mean and standard deviation of the observed RT (Fig. 3.3). This is evident in the conditions with noninformative stimuli, and especially in the comparison of redundant and unimodal conditions. One theoretical premise is that redundant targets' mean RTs approach the unimodal RT means when the SOA becomes infinite. This is the case for both models, and the DSM nicely fits the observed mean RTs. However, the DRM clearly cannot explain the observed redundancy gain, even on the level of mean RTs.

Like those for mean RTs, the DRM predictions for response frequencies in the nonredundant no-go conditions were overestimated (Fig. 3.3, lower panel). Thus, the observed mean RTs in go conditions with noninformative stimuli and the response frequency to no-go stimuli with the same noninformative stimuli were faster and more accurate, respectively, than those predicted by the DRM.

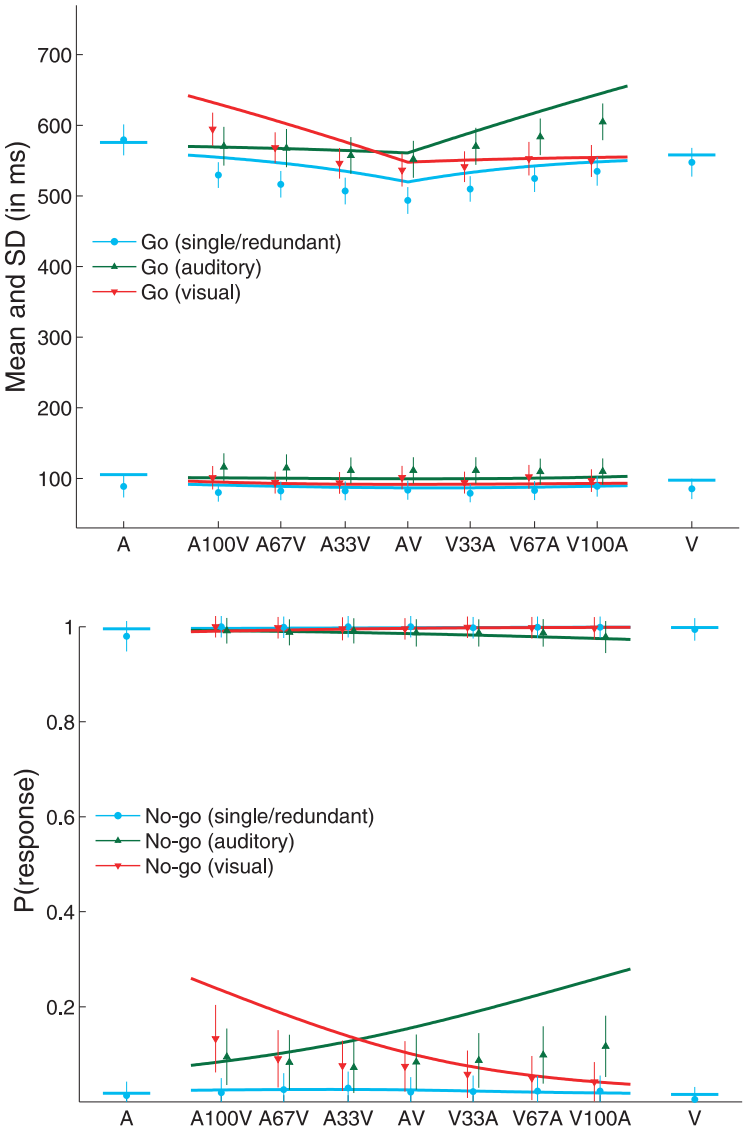


Figure 3.3: Averaged mean response times and standard deviations of go trials (*upper panel*) and proportions of error responses on no-go trials (*lower panel*), together with averaged model predictions from the diffusion racemodel, for all nine participants in Experiment 1. The data are the same as in Fig. 3.2, but the model predictions are derived from the diffusion race model. The coding of conditions is the same as in Fig. 3.2.

3.2.2 Discussion

The goal of Experiment 1 was to investigate the mechanism of integration in a redundant-signals experiment with a task more complex than a simple response task (e.g., Miller, 1986). For this purpose, we tested nine participants in a go/no-go task and asked them to respond as quickly as possible to predefined audiovisual target stimuli and to refrain from responding to no-go stimuli (Miller, 1982). A diffusion superposition model with two absorbing barriers was fitted to the means and variances of the RTs of 23 go conditions, as well as to the accuracy rates of 23 no-go conditions. The race model inequality was significantly violated in all participants; thus, the standard race model with independent channels and unlimited capacity does not seem to be a plausible candidate to explain the observed data. Likewise, the capacity coefficient (Townsend & Nozawa, 1995) indicated supercapacity across multiple time points in all participants. Processing was considerably faster than was predicted by standard parallel processing (race model), suggesting interactive channels or coactive processing. On the other hand, the diffusion superposition model with two absorbing barriers could explain the mean RTs, standard deviations, and accuracy rates well in all but one participant; the fit of the model was acceptable to excellent in most participants (Table 3.3). The parameter estimates for the lower barrier suggest that the participants adopted two different strategies in the go/no-go task. This is indicated by a response bias, represented by the relative distances of the upper and lower barriers from the origin. In Participants 5 and 7, the lower criterion is farther away from the origin than the (fixed) upper criterion. This, in turn, implies that those participants were prepared to respond and withheld the response if a no-go stimulus was presented. Participants 2, 3, 4, 6, and 8 seem to have chosen a different strategy: These participants expected a no-go stimulus and, thus, tended not to respond. The parameter estimates for the lower barrier lie closer to the origin than the upper

criterion. The other two participants do not seem to have adopted a strategy with any response bias at all.

That the model explained the observed data has two interesting implications. First, we provided evidence that the same principles that govern multisensory integration in simple RTs also apply to a more complex task, such as the go/no-go task. To this end, we not only tested separate-activation models, but also an alternative model. This has been done before (Gondan et al., 2010), but the results of that study were subject to response conflicts that participants might have had problems resolving. We avoided response conflicts by adding nontargets: Rather than directly combining go and no-go audiovisual stimulus components, we used a third class of noninformative stimulus that could be presented with go and no-go stimulus components (Grice et al., 1984). Whereas in Gondan et al. (2010) it seemed that some participants resolved possible response conflicts by serial processing of the stimuli, the present results suggest coactive processing of the bimodal information when the provided information is free of response conflicts.

Second, using a DSM with two absorbing barriers allowed for description not only of the mean RTs and variance, but also of response accuracy in the go/no-go task (Rach, Diederich, & Colonius, 2011). This approach is not new (Gomez et al., 2007), but for the first time we used it to model the data of a go/no-go task in an experiment with redundant signals presented with an SOA. There is also a conceptual difference between our diffusion-modeling approach and that based on Ratcliff's (1978) diffusion model. The models based on Ratcliff's diffusion model allow for different modulations of the parameters for example, a randomly distributed starting point or trial-to-trial variations in drift rates. The diffusion superposition models for the redundant-signals effect (Diederich, 1995; Schwarz, 1994) and the model that we outlined here currently do not allow for these variations. Instead, the diffusion process is divided into two (or possibly more) parts that may have different drifts and variances (see Fig. 3.1b). This

approach is analytically tedious for two parts (Ratcliff, 1980; Schwarz, 1994) and requires numerical approximation for three or more subdivisions (Diederich, 1995), but it has proven to be successful in modeling the mean RTs and variances of a redundant-signals task with stimuli presented with onset asynchronies. Therefore, we followed the same approach in the derivation of the diffusion model with two absorbing barriers. Note that the application of such a compound model is not restricted to the redundant-signals task. Similar models have been used to explain temporal-order judgments (Schwarz, 2006) and could be applied to data from any experimental setup inwhich the rate of accumulation of information is expected to change at some time. This includes not only stimuli presented with an onset asynchrony, but any experimental setup with an interstimulus interval (e.g., a sequential comparison task or cueing tasks).

3.3 Experiment 2

In Experiment 2, participants made choice responses to audiovisual redundant targets in a speeded response task with two alternatives. Analogously to Experiment 1, we compared the model fits of the two diffusion models to the two-choice data and investigated process capacity in a two-choice task.

3.3.1 Results

RTs and task performance. The mean RTs of redundant stimuli exhibited a roughly symmetric wing-shaped pattern for both alternatives and in all participants (Fig. 3.4; circles in upper panels). Responses to the bimodal redundant targets (A^+V^+) were the fastest in all conditions. For the bimodal nonredundant stimuli (A^+V^0 and A^0V^+), a similar pattern was observed, but the shape was clearly asymmetric: The mean RTs of the conditions in which the noninformative stimulus followed the target stimulus were nearly constant across SOA conditions (Fig. 3.4; left half of upward-pointed triangles and right half of downward trian-

gles). If the noninformative stimulus preceded the target by some SOA, we found an increase in mean RTs that was dependent on the SOA (right half of upward triangles and left half of downward triangles). Similar patterns can be observed for the numbers of correct responses. The frequency of correct responses was highest for the bimodal redundant and unimodal conditions (lower panels of Fig. 3.4, same condition-coding as above) and lower for the bimodal nonredundant signals. This was especially evident in conditions in which the noninformative stimulus was presented before the target stimulus. In some participants, task performance approached chance level with increasing SOAs between the noninformative and target stimuli. That is, some participants tended to guess the correct response, but nevertheless, the mean response latency increased with decreasing performance. Standard deviations, on the other hand, were largely constant across all conditions in all observers (Fig. 3.4, upper panels). Figures of the participant-specific results are available as supplemental materials (Supplemental Fig. 3.3).

Tests of the race model predictions. The race model was tested for the redundant-stimulus conditions of both response alternatives by calculating the observed violation areas for each response alternative. The race model inequality for the redundant-stimulus conditions was significantly violated in six participants (i.e., 1, 3, 4, 5, 6, and 8), for either left or right responses. The remaining two participants do not show significant violations of the race model inequality in either left or right responses (Table 3.4). In Experiment 2, the generalized assessment index (Townsend & Altieri, 2012) was used to estimate processing capacity. Again, under the canonical parallel independent race model $A(t) = 1$, we tested the assessment index against unity and focused on the correct and fast case. Contrary to the capacity index and standard race model tests, the generalized capacity index $A(t)$ also explicitly takes into account the observed erroneous responses. This yields four cases, depending on response correctness (correct/incorrect) and response speed

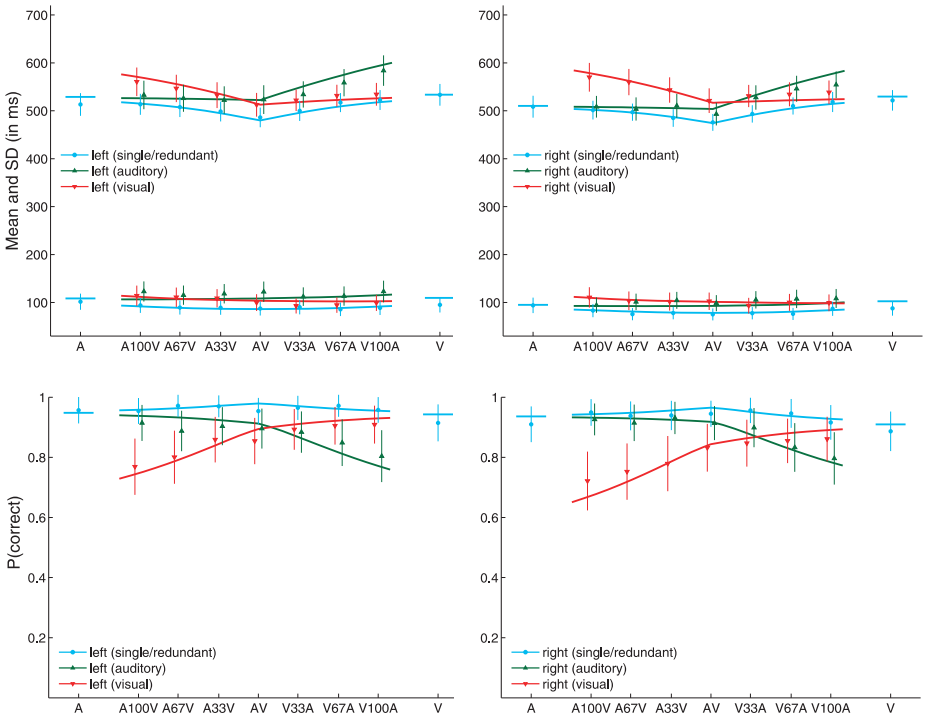


Figure 3.4: Averaged mean response times and standard deviations of choice responses (*upper panels*) and the corresponding proportions of correct responses (*lower panels*), together with averaged model predictions from the diffusion superposition model for all eight participants in Experiment 2. For each observer, the left panels display the data and model predictions from one response alternative (*left buttonpress*), and the right panels display the data and model predictions from the other response alternative (*right buttonpress*). The curves represent the averaged model predictions for unimodal/bimodal redundant, nonredundant auditory-target, and nonredundant visual-target stimulus conditions. Data from the redundant bimodal conditions and unimodal conditions are displayed as *circles*, and data from the bimodal nonredundant conditions with either visual or auditory noninformative stimuli are displayed in contrasting ways (auditory target: upward-pointing triangles; visual target: downward-pointing triangles). On the x-axis, conditions are ordered from left to right according to stimulus onset asynchrony, and error bars denote 95 % confidence interval based on averaged standard errors. Participant-specific data and model fits are displayed in Supplemental Fig. 3.3.

Table 3.4: *Experiment 2—Result of race model inequality tests for both response alternatives*

		Participant							
		1	2	3	4	5	6	7	8
Left	Δ	50.1	0.4	77.6	57.1	14.9	45.3	24.0	6.6
responses	p value	.350	.986	.127	$< .001$	.678	.025	.511	.799
Right	Δ	107.6	27.1	106.4	32.1	35.6	39.5	27.5	42.8
responses	p value	$< .001$	.617	$< .001$	.040	.040	.064	.141	.007

Δ: violation area up to the 20th percentile; p value: proportion of simulations (under race model assumptions) that showed violation areas greater than the observed violation area Δ.

(fast/slow). In the correct and fast case, capacity was significantly greater than was predicted by parallel independent processing in the right-response conditions ($T_{\max} = 5.66$, $T_{0.95} = 2.26$, $p < .001$), and, to a similar extent in the left-response conditions ($T_{\max} = 5.69$, $T_{0.95} = 3.05$, $p = .003$). The assessment index showed supercapacity for either left or right responses in all but one participant (Participant 5), whose data for both response alternatives were largely consistent with parallel independent processing (Supplemental Fig. 4A). The lower boundary was violated only in Participant 3, whose left responses showed severely limited capacity. The time course of $A(t)$, together with the upper and lower capacity bounds, are displayed in Supplemental Fig. 3.4A. Violations of the race model inequality for all participants are shown in Supplemental Fig. 3.4B. In general, race model inequality tests seem to be less sensitive than capacity analyses regarding violations of race model predictions for two-choice response data.

Diffusion superposition model. A two-barrier DSM was fitted to the mean RTs and proportions of correct responses for both alternatives. Except for Participant 5, the model fit was acceptable in all participants (see Fig. 3.4 and Table 3.5); some

Table 3.5: *Experiment 2—Parameter estimates and model fits of the two-barrier diffusion superposition model for the mean and variance of response times of both response alternatives, together with the respective proportions of correct responses*

		Participant							
Parameter		1	2	3	4	5	6	7	8
Left	μ_{V}	0.76	0.63	1.01	0.85	0.97	0.71	0.77	0.73
responses	σ^2_{V}	74.5	56.1	115.0	67.9	33.4	54.6	15.7	43.7
	μ_{A}	0.71	0.74	1.59	0.87	0.86	0.70	0.73	0.88
	σ^2_{A}	75.3	73.0	106.5	66.1	30.1	43.3	43.3	40.6
	μ_{M}	438.5	441.8	427.0	455.7	471.1	371.3	436.8	331.1
	σ^2_{M}	9,764	10,886	4,342	2,605	4,257	2,202	5,413	7,520
	ρ_{DM}	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
	u	(100)	(100)	(100)	(100)	(100)	(100)	(100)	(100)
Noninfor-	μ_{V0}	0.05	–0.08	–0.10	–0.03	–0.07	0.05	–0.08	–0.02
mative	σ^2_{V0}	64.9	35.6	94.7	65.3	6.1	33.5	19.6	46.9
Stimli	μ_{A0}	0.05	0.05	–0.08	0.08	–0.01	0.05	0.04	0.20
	σ^2_{A0}	73.6	68.8	82.7	59.2	23.7	55.7	52.1	49.0
Right	μ_{V}	–0.41	–0.70	–1.02	–1.00	–0.77	–0.64	–0.77	–1.19
responses	σ^2_{V}	75.0	74.9	95.3	79.7	19.9	61.7	17.9	70.3
	μ_{A}	–0.78	–0.78	–2.13	–0.97	–0.86	–1.10	–0.73	–1.26
	σ^2_{A}	100.6	90.6	111.1	59.8	37.3	59.0	50.3	77.2
	μ_{M}	482.5	434.3	385.8	437.9	444.7	385.3	420.2	364.2
	σ^2_{M}	7,297	6,359	3,029	2,611	2,424	3,070	3,976	9,511
	ρ_{DM}	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
	$-\ell$	(–100)	(–100)	–144	(–100)	(–100)	(–100)	(–100)	(–100)
	ρ_{AV}	–0.33	–0.22	–0.27	–0.49	0.28	–0.24	–0.04	–0.41
GOF X^2	DSM	156.1	268.3	255.6	204.9	445.6	214.0	206.8	127.2
	DRM	376.1	195.6	357.0	339.1	429.9	284.0	253.2	180.0

μ_*, σ^2_*: drift and variance for visual (auditory) diffusion process for target (μ_{V}, σ^2_{V}, μ_{A}, σ^2_{A}) and noninformative (μ_{V0}, σ^2_{V0}, μ_{A0}, σ^2_{A0}) stimuli; u: evidence criterion for "left" response; $-\ell$: evidence criterion for "right" response; ρ_{AV}: correlation between $\mathbf{D}_{\mathrm{A}}$ and $\mathbf{D}_{\mathrm{V}}$; μ_{M}, σ^2_{M}: mean and variance of residual process latency; ρ_{DM}: correlation between $\mathbf{D}$ (detection time) and $\mathbf{M}$ (residual component duration). Parameters in parentheses are fixed parameters. GOF X^2: goodness-of-fit statistic (higher results indicate worse fit).

participants (i.e., 1 and 8) showed excellent model fit. Again, drift estimates are ordered; thus, drifts for stimuli requiring a left buttonpress direct the process to the upper barrier associated with this response. By contrast, drifts for stimuli requiring a right buttonpress tend toward the lower barrier, and parameter estimates for drifts for the noninformative stimuli are intermediate and vary around zero. The estimates for the lower barrier were very close to those for the upper barrier, so we fitted a model to the data in which the parameter $-\ell$ was fixed to -100. This unbiased model (Wagenmakers et al., 2007) fit approximately as well as the more liberal model in all but one participant (Participant 3). In this participant, the model fit of the more restrictive model was considerably worse, so we kept the liberal model. In the remaining participants (i.e., 2, 4, 6, and 7), the model fit well to the observed means and standard deviations of RTs and the proportions of correct responses. Systematic deviations between model predictions and the data were obtained in only one participant (Participant 5). This is evident in the proportion of correct responses, which are systematically overestimated by the model, and in the mean RTs of the unimodal stimuli, which are also overestimated (Supplemental Fig. 3.4). Further systematic deviations are observable in standard deviations; thus, for this participant the model had to be rejected. In the majority of participants, however, the model yielded a good approximation of the mean and variance of their RTs and numbers of correct responses in the perceptual-choice task.

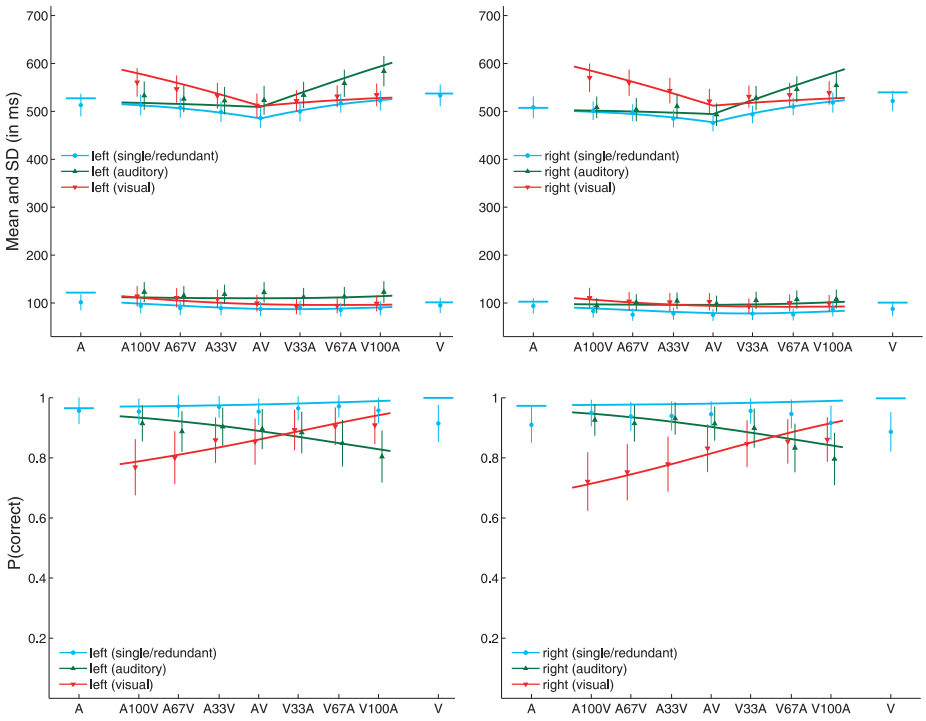


Figure 3.5: Averaged mean response times and standard deviations of both response alternatives (*upper panels*) and proportions of correct responses for both response alternatives (*lower panels*), together with averaged model predictions from the diffusion race model (*curves*), for all eight participants in Experiment 2. The data are the same as in Fig. 3.4, and the same coding of conditions was used as in that figure.

Diffusion race model. By contrast, a race of independent diffusion processes (DRM) again failed to make adequate predictions in most participants (Table 3.5). Interestingly, the deviations were most pronounced in the response accuracy of the unimodal and redundant conditions, whereas predictions of the means and standard deviations of RTs were generally acceptable (Fig. 3.5, lower panel). In particular, the model was not able to explain response accuracy in the unimodal and redundant conditions for the right response alternative. The decrease in the accuracy rate of the unimodal visual-target condition (Fig. 3.5, lower panels) is not captured by the model, and the accuracy rates in the right-response redundant-target conditions are generally overestimated (Fig. 3.5, lower right panel). Because

the mean RT and response accuracy for the right response alternative share the same parameters (μ_{A-}, σ^2_{A-}, μ_{V-}, and σ^2_{V-}), the responses are again either faster or more accurate than the DRM would predict. On a quantitative basis, the DRM fit worse than the DSM in all but one participant (Table 3.5), if one disregards Participant 5, whose data were not explained satisfactorily by either model. But in Participant 2, the model fit is better than that of the DSM, and in line with this observation, this is also the participant with the fewest violations of the race model inequality (Table 3.4). So, although the DRM is certainly only one instance of all possible race models, the fact that the same assumptions, differing only with respect to how information is integrated across channels, lead to such differences in quantitative and qualitative model comparisons further endorses the DSM.

3.3.2 Discussion

The goal of Experiment 2 was to test separate-activation models and to test whether a diffusion superposition model with two absorbing barriers can also be applied to the mean and variance of RTs and the proportions of correct responses from a perceptual-choice task with redundant stimuli. To this end, we fitted the DSM to the data of eight participants. For the first time, we demonstrated that a DSM can describe well the mean RTs observed in a redundant-signals experiment with a perceptual-choice task. The DSM itself is rather restrictive; with only two additional parameters (relative to the model for the go/no-go task), we were able to derive accurate predictions for the mean and variance of RTs and the response accuracy of both response alternatives. Each diffusion parameter contributed to the predictions of mean RTs, standard deviations, and proportions correct, and, albeit using different stimulus parameters, the data from both response alternatives could be explained by a single model. This includes the observed decrease in performance with increasing SOAs in the nonredundant distractor-first conditions, which is also captured well by the model. Apart from one participant, the

model fit was acceptable to excellent; thus, the model fits support the idea of an additive superposition of channel-specific activity also in a perceptual two-choice RT task. Given the stimulus materials and parameter estimates, it seems safe to conclude that the diffusion process in the DSM represents stimulus information, not stimulus energy. Hence, the same principles that underlie models for human decision making (Busemeyer & Townsend, 1993; Diederich, 1997) can be applied to a multisensory choice RT task in a redundant-signals experiment.

As an alternative model, we used a diffusion race model so as to compare the model fits of very similar models that differ with respect to the type of channel integration (parallel first-terminating vs. parallel coactive). With that in mind, it is striking to see the differences in the abilities of both models to make predictions for the observed data. We are aware that the DRM is just one of many possible race models, but it is in closest correspondence to the DSM, which in turn provides an adequate model fit. Apart from small deviations, both mean RTs and choice probabilities could be explained well by the DSM. The observed standard deviations of some conditions deviated from model predictions, probably because the estimate for standard deviation was highly susceptible to outliers. We also tested for violations of the race model inequality and found frequent violations, which seemed to be less pronounced in Experiment 2 than in Experiment 1. This result is in line with previous studies on the redundant-signals effect with choice RT tasks (e.g., Grice et al., 1984; Grice & Reed, 1992; Grice & Canham, 1990). For instance, the base time seems to be increased and more variable in the two-choice task (Townsend & Honey, 2007), as can be inferred from the DSM parameter estimates (Table 3.5). It has been demonstrated that base time can play a critical role in tests of race model violations (Townsend & Honey, 2007). However, using the assessment index as a benchmark for parallel independent processing, we found strong evidence for supercapacity—that is, violations of the parallel-independent-processing assumption. A tentative explanation would be that by taking response

accuracy into account, the assessment index provided a more sensitive measure of parallel independent processing in a two-choice task than did the race model inequality. At any rate, erroneous responses seem to play a critical role when testing race model assumptions in a choice RT task. This can be seen in both the results of our capacity analysis and the model comparison of the DSM with the DRM, which differed mostly in their ability to explain the observed response accuracy. Altogether, parallel race models did not seem to provide a good explanation for the observed results in the majority of participants.

3.4 General discussion

3.4.1 Multisensory integration

The question of which cognitive architecture underlies the processing of redundant multisensory signals has been thoroughly addressed, sometimes with mixed results, and often on theoretical grounds. Although the empirical evidence seems to favor the coactivation account if tests of the race model inequality are used, ongoing discussion concerns whether the race model inequality is an appropriate tool to decide this question at all. Specifically, the validity of race model tests relies on the assumption of selective influence,whichmight or might not hold in a given experiment (e.g., Luce, 1986; Townsend & Nozawa, 1995). Alternative measures have been proposed to decide this question, but selective influence is a silent or explicit assumption in all of these measures.

In this study, we tested whether a model with a simple additive superposition of diffusion processes could also be applied to data from a redundant-signals experiment with tasks more complex than simple RT. For this purpose, we conducted two experiments with different response requirements: In Experiment 1, participants were asked to perform an audiovisual go/no-go task, and in Experiment 2, a multisensory choice task. So far, most studies on multisensory integration

with go/no-go (Miller, 1982) and choice RT tasks have only tested for violations of the race model inequality (and frequently found them), but have not tested alternative models. On the other hand, explicit models for the redundant-signals effect have been presented and tested mainly using data from simple response experiments (Diederich, 1995; Miller, 1986; Schwarz, 1994). By contrast, most neurophysiological studies have made use of the go/no-go, or oddball, task and of choice RT tasks (e.g., Talsma, Kok, Slagter, & Cipriani, 2008; Woldorff et al., 1993). It has also been put into question whether the simple RT task is suitable to study certain multisensory aspects at all (Spence & Driver, 1997). We provided explicit model predictions for diffusion models based on channel superposition (DSM) or the race-type integration of information (DRM) for two-choice data obtained in multisensory redundant-signals experiments. The integration mechanism according to the DSM is the additive superposition of modality-specific activation. This is in line with studies that have reported approximately additive neural responses to bimodal stimuli (Stanford, Quessy, & Stein, 2005).

In the two experiments reported here, we not only employed unimodal and bimodal audiovisual stimuli, but also used noninformative ("neutral") stimuli to combine the two tasks with a redundant-signals experiment (Grice et al., 1984). Apart from solving possible response conflicts, the noninformative stimuli gave rise to additional bimodal conditions. The bimodal conditions could therefore be either redundant or nonredundant, depending onwhether a noninformative stimulus was presented in one channel or both channels contained targets. The results show that the DSM can make adequate predictions for those conditions, also under the assumption of additive superposition, if the respective stimuli are modeled with separate parameters (Tables 3.3 and 3.5). The estimates suggest that these stimuli were indeed conceived as neutral, since the drift rates for the noninformative stimuli were approximately zero in all participants. In general, the estimates were smaller in absolute values than the estimates for the two response-relevant stimuli

(Exp. 2) or the go/no-go stimuli (Exp. 1). Therefore, the diffusion process of the DSM is likely to represent the response tendency associated with a stimulus, and not simply stimulus energy. This distinction is hard to make in simple RT tasks.

3.4.2 RT and response accuracy

One major advantage of two-choice diffusion models is the joint analysis of RT data (mean, variance, or distribution) and response accuracy (e.g., Rach et al., 2011). Most publications on the redundant-signals effect have involved testing the race model inequality—that is, RT distributions. But comparing the results of the two experiments reported here, we can observe a striking difference regarding response accuracy: In the choice experiment, the numbers of correct responses were roughly symmetric across the two response alternatives in all participants (Fig. 3.4 and Supplemental Fig. 3.3). In the go/no-go task, however, omissions in go stimuli (i.e., lapses) were far less frequent than false alarms to no-go stimuli (Fig. 3.2 and Supplemental Fig. 3.1). This could be explained by differences between the estimates for the lower barrier in both experiments. In the choice response task, the estimates for the lower barrier are identical or roughly identical to the fixed value of the upper barrier. Therefore, we further restricted the model so that both barriers had to be equally distant from the origin (i.e., $u = \ell = 100$). This model fit nearly as well as the model with the lower barrier as a free parameter in all but one participant (Participant 3). However, in the go/no-go task, the estimates differed considerably in most participants. Therefore, a considerable response bias seemed to be present in Experiment 1 (go/no-go task), but not in Experiment 2 (binary choice task). Note that slightly different visual stimuli were used in the two experiments, which could confound comparison of the two experiments. However, a response bias in go/no-go but not in two-choice tasks as a reason for the different results in the two tasks has also been discussed earlier (Gomez et al., 2007). Also note the clear difference between the estimates for the channel correlation ρ_{AV} in

both experiments (Tables 3.3 and 3.5). In the DSM, the ρ_{AV} parameter reflects the correlation of the buildup of evidence in the two modality-specific channels. Overall, fluctuations of attention might result in a positive ρ_{AV} (Exp. 1) because the buildup of evidence in both channels is then either high in both channels (if the participant concentrates on the task) or low in both channels (if the participant is distracted). On the other hand, attentional fluctuations between channels might result in a negative channel correlation and negative ρ_{AV} (see, e.g., the discussion in Miller, 1982): In certain trials, attention might be focused on the auditory modality, so that the buildup of evidence occurs faster in the auditory channel but slower in the visual channel, and vice versa when attention is focused on the visual channel. Especially in the more-demanding task of Experiment 2, the latter fluctuations might be prompted by unimodal stimuli (including bimodal stimuli with long SOAs) that lead to a single percept and direct attention to a certain modality. Except for the response criterion in Experiment 1, the estimates of most parameters for the DSM are largely consistent within each data set, even though the participants' data varied considerably in both response latency and response accuracy (Tables 3.3 and 3.5, Supplemental Figs. 3.1 and 3.3). The greatest variations of parameter estimates are consequently found in the estimates of μ_M and σ^2_M. This implies that the variability between our participants (e.g., motivation) largely affected the residual, or motor, component.

Apart from the aforementioned asymmetry, error rates seem to be generally lower in the go/no-go task than in the two-choice task (Perea, Rosa, & Gómez, 2002). The DRM exhibited difficulties in explaining the error rate of Experiment 2, whereas the mean RTs were relatively well predicted by the model. As opposed to Experiment 1, responses were less correct (or slower) than the DRM predicted. In Experiment 2, lesser violations of the race model inequality were observed, and the DRM showed considerably better fits to the RT data. Response accuracy, however, could not be explained by the model. This fits nicely with the results of the

test of the assessment index for supercapacity (Townsend & Altieri, 2012). The assessment index explicitly takes errors into account and reveals strong evidence for violations of parallel-independent-processing assumptions. In Experiment 1, on the other hand, frequent violations of standard parallel-processing assumptions were observed, and the DRM largely failed to predict RT measures. The DSM, on the other hand, could adequately describe both RTs and response accuracy and allowed for interesting implications. According to our data, a considerable increase in mean RTs and a decrease in response accuracy take place between bimodal redundant and bimodal nonredundant-signal trials, and the decrease is clearly more pronounced in the choice RT task. All of these aspects are captured well by the DSM. This demonstrates the necessity to analyze and, in this case, model response accuracy along with response latency (Rach et al., 2011; Townsend & Altieri, 2012).

3.4.3 Possible applications of the two-choice diffusion model

Using go/no-go and two-choice RT tasks, we demonstrated that the scope of the DSM (Schwarz, 1994) is not restricted to simple response tasks. This is not only interesting for modeling purposes; rather, two important implications arise from this result. First, many studies on multisensory perception make frequent use of the tasks that we used here—that is, the go/no-go, or oddball, task and the choice RT task. Second, and more importantly, the results of the bimodal response conditions support the notion that not energy summation, but the coactivation of response tendencies, accounts for the observed redundancy gains. According to recent neuroscience studies (e.g., Angelaki, Gu, & DeAngelis, 2009; Ma, Beck, Lantham, & Pouget, 2006; Stanford et al., 2005), the additive superposition of channel-specific activation seems to be a more general principle in the human brain. With purely visual stimuli, several studies have also successfully linked behavioral diffusion models to neuropsychological data (e.g., Ratcliff, Cherian, & Segraves,

2003; Ratcliff, Philiastides, & Sajda, 2009). On this note, an accumulation of evidence in the prefrontal cortex seems to be consistent with the predictions of the two-barrier diffusion model, as has been confirmed with electroencephalographic data (Ratcliff et al., 2009).

Diffusion processes, thus, seem to be a promising approach to model a variety of human behaviors (Smith, 2000). This encompasses not only the integration of evidence from multisensory sources (Diederich, 1995; Gondan et al., 2010), but also attentional effects (Gondan et al., 2011), memory retrieval (Ratcliff, 1978), masking effects (Ratcliff & Rouder, 2000), and simple two-choice decisions (e.g., Krajbich, Armel, & Rangel, 2010; Usher & McClelland, 2001). Here, we showed that the abundant research on diffusion models of decision making might be connected to diffusion models that account for the redundant-signals effect (Diederich, 1995; Schwarz, 1994). The same principles that underlie models for decision making in other experimental tasks, such as the numerical-distance effect or perceptual masking, are also applicable to model the redundant-signals effect with more-complex tasks than the simple RT task. Even more general, the model itself can be applied to any experimental data from a decision task in which the rate of information processing is assumed to change after some time τ.

Author note Supported by Grant Nos. GO 1855/1-1 and GR 998/20-2 from the German Research Foundation (Deutsche Forschungsgemeinschaft). We also thank James Townsend, Durk Talsma and an anonymous reviewer for their help on an earlier version of this manuscript.

3.5 Appendix A. Predictions of the diffusion superposition model with two barriers

In order to derive predictions for the mean and the variance of RTs and the amount of responses in the SOA-condition, these have to be conditioned on the cases $\mathbf{D} \leq \tau$ and $\mathbf{D} > \tau$ (cf. Diederich, 1995; Ratcliff, 1980; Schwarz, 1994). Technically, this model is a generalization of the model with one absorbing barrier (Schwarz, 1994). Schwarz derived analytical solutions for mean RTs and standard deviations of asynchronously presented stimuli. Here, we outline the derivation of the prediction functions for expected time, variance and amount of passages at the upper barrier ($\mathbf{C} = 1$) only, however, predictions for absorption at the lower barrier can be readily obtained using the same functions by changing the sign of μ and interchanging the barriers. It is assumed that for $t \leq \tau$, drift and variance of the process correspond either to μ_{A} and σ^2_{A} or to μ_{V} and σ^2_{V}, depending on which stimulus has been presented first. This process can reach one of the barriers within the interval $0 < t \leq \tau$. If this has not happened before the onset of the second stimulus, both stimuli contribute to the process that then has drift $\mu_{\mathrm{A}} + \mu_{\mathrm{V}}$ and variance $\sigma^2_{\mathrm{A}} + \sigma^2_{\mathrm{V}} + 2\rho_{\mathrm{AV}}\sigma_{\mathrm{A}}\sigma_{\mathrm{V}}$. For the probability of a first passage of the upper criterion $P(\mathbf{C} = 1)$, one has to calculate the sum of the first-passage probabilities of both cases:

$$P(\mathbf{C} = 1) = P(\mathbf{C} = 1 \cap \mathbf{D} \leq \tau) + P(\mathbf{C} = 1 \cap \mathbf{D} > \tau). \tag{A1}$$

The first summand can be easily obtained using the subsurvivor function $S_u(t|u, -\ell, \mu, \sigma^2)$ of passages through the upper barrier, which is presented by Horrocks and Thompson (2004):

$$P(\mathbf{C} = 1 \cap \mathbf{D} \leq \tau) = P_u - S_u(\tau),$$

where P_u is the well-known probability of absorption at the upper barrier (e.g., Horrocks & Thompson, 2004, their Eq. 9; Ratcliff, 1978, his Eq. A8). There exist two representations for the subsurvivor function S_u, both involving infinite series, but with different convergence behaviors. For speed reasons and effective control of the error tolerance, we calculate the terms needed to achieve a predefined error tolerance in advance and used whichever representation converged faster (Blurton et al., 2012). For the second summand of Eq. A1, one has to consider that at time τ the process starts anywhere within the range $-\ell < x < u$. The density of x at the start of the second part of the process is given by an infinite series $g(x, \tau|u, -\ell, \mu, \sigma^2)$, which is the density of a Wiener process between two absorbing barriers (Cox & Miller, 1965, p. 222, their Eq. 78). A second representation of this function is also an infinite series (Cox & Miller, 1965, p. 222, their Eq. 81). Again, we implement both representations and use the representation that converge faster for any given x, τ, parameters and specified error tolerance. Then,

$$P(\mathbf{C} = 1 \cap \mathbf{D} > \tau) = \int_{-\ell}^{u} P_x(\mathbf{C} = 1) \times g(x, \tau|u, -\ell, \mu, \sigma^2)\mathrm{d}x$$

is the probability of a first passage at the upper barrier, given that no absorption occurred before τ. $P_x(\mathbf{C} = 1)$ indicates that barriers u and $-\ell$ are replaced by the modified barriers $(u - x)$ and $-(\ell + x)$ in $P_u = P(\mathbf{C} = 1)$, making the expression dependent on x.

Likewise, the expected time of a first passage of a Wiener process between two barriers, given that absorption occurs at the upper barrier, is

$$\begin{aligned} E(\mathbf{D}|\mathbf{C} = 1) = \frac{E(\mathbf{D}|\mathbf{C} = 1, \mathbf{D} \leq \tau) \times P(\mathbf{C} = 1|\mathbf{D} \leq \tau)}{P(\mathbf{C} = 1)} + \\ + \frac{E(\mathbf{D}|\mathbf{C} = 1, \mathbf{D} > \tau) \times P(\mathbf{C} = 1|\mathbf{D} > \tau)}{P(\mathbf{C} = 1)}. \end{aligned} \tag{A2}$$

The first part of the numerator in Eq. A2 is the expected value for the first

passage of the upper barrier, given that the upper barrier is reached first and that this occurs before the second stimulus is presented. This is

$$E(\mathbf{D}|\mathbf{D} \leq \tau) \times P(\mathbf{C} = 1, \mathbf{D} \leq \tau) = \int_0^\tau t \times f_u(t|u, -\ell, \mu, \sigma^2)\mathrm{d}t, \tag{A3}$$

where $f_u(t|u, -\ell, \mu, \sigma^2)$ is the subdensity of first passage times at the upper barrier, a computationally efficient MATLAB implementation of which can be found in Navarro and Fuss (2009). The integral can then be determined numerically.

For the second summand of the numerator in Eq. A3, one needs the expected value of $\mathbf{D}$, conditioned on a first passage at the upper barrier. This term, $E(\mathbf{D}|\mathbf{C} = 1)$, is also given in Horrocks and Thompson (2004, their Eq. 10). Again, the process does not necessarily start at the origin, but rather at any value within the interval $-\ell < x < u$. Thus, one has to insert $(u - x)$ and $-(\ell + x)$ instead of u and $-\ell$, respectively, making $E(\mathbf{D}|\mathbf{C} = 1)$ dependent on x. Then,

$$\begin{aligned} &E(\mathbf{D}|\mathbf{C} = 1, \mathbf{D} > \tau) \times P(\mathbf{C} = 1, \mathbf{D} > \tau) = \\ &\quad = \int_{-\ell}^{u} [\tau + E_x(\mathbf{D}|\mathbf{C} = 1)] \times g(x, \tau|u, -\ell, \mu, \sigma^2)\mathrm{d}x \end{aligned} \tag{A4}$$

represents the second part of the numerator in Eq. A3. Again, $E_x(\mathbf{D}|\mathbf{C} = 1)$ denotes that $E(\mathbf{D}|\mathbf{C} = 1)$ is used with the modified barriers $(u - x)$, and $-(\ell + x)$ and represents the expected value of $\mathbf{D}$ for the second part of the process $(t > \tau)$. Thus, adding τ and numerically integrating over x from $-\ell$ to u yields the expected value of first-passage times for the case $\mathbf{D} > \tau$.

The variance predictions of RT are derived as follows: The variance is the second (central) moment minus the squared first moment, so only the derivation of the former is needed. Again, we condition of the two cases $\mathbf{D} \leq \tau$ and $\mathbf{D} > \tau$:

$$E(\mathbf{D}^2|\mathbf{C}=1) = \frac{E(\mathbf{D}^2|\mathbf{C}=1,\mathbf{D}\leq\tau)\times P(\mathbf{C}=1|\mathbf{D}\leq\tau)}{P(\mathbf{C}=1)} + \frac{E(\mathbf{D}^2|\mathbf{C}=1,\mathbf{D}>\tau)\times P(\mathbf{C}=1|\mathbf{D}>\tau)}{P(\mathbf{C}=1)} \quad \text{(A5)}$$

The first summand is derived similar to the first moment (Eq. A3), except that t^2 is multiplied with $f(t)$ before numerical integration. The second summand of Eq. A5 is more problematic, but the derivation follows that for the first central moment (Eq. A4). Thus, the second central moment conditioned on a particular value of x is needed. Differentiating between the overall decision time $\mathbf{D}$ and the decision time $\mathbf{D}^*$ for the second part after τ —that is, after the onset of the second stimulus—we obtain

$$E(\mathbf{D}^2|\mathbf{C}=1,\mathbf{X}(\tau)=x) = E_x[(\tau+\mathbf{D}^*)^2] = \tau^2 + 2\tau E_x(\mathbf{D}^*) + E_x[(\mathbf{D}^*)^2] \quad \text{(A6)}$$

The second moment $E[(\mathbf{D}^*)^2]$ can be derived from Grasman et al. (2009, their Eq. 14). As for the first central moment, Expression A6 with the modified barriers $(u-x)$ and $-(\ell+x)$ instead of u and $-\ell$, respectively, has to be multiplied with the density $g(x,\tau|u,-\ell,\mu,\sigma^2)$ and integrated over all values of x for $-\ell < x < u$. Finally, the response time $\mathbf{T}$ is conceived as the sum of the decision time $\mathbf{D}$ and residual processes $\mathbf{M}$, so the model predictions for the mean and variance of response time $\mathbf{T}$ are

$$E(\mathbf{T}) = E(\mathbf{D}+\mathbf{M}) = E(\mathbf{D}) + \mu_{\mathrm{M}} \quad \text{(A7)}$$

$$Var(\mathbf{T}) = Var(\mathbf{D}) + \sigma^2_{\mathrm{M}} + 2\rho_{\mathrm{DM}}\sigma_{\mathrm{D}}\sigma_{\mathrm{M}}. \quad \text{(A8)}$$

Altogether, seven free parameters (μ_{A}, σ^2_{A}, μ_{V}, σ^2_{V}, ρ_{AV}, $-\ell$, and μ_{M}) are needed to derive the predictions for mean RTs $E(\mathbf{T})$, and nine parameters (i.e.,

ρ_{DM} and σ^2_{M} added) are necessary for variance predictions $Var(\mathbf{T})$. In the case of an unbiased model, the numbers of parameters reduce to six and eight parameters for mean RT and the variance of RT, respectively.

3.6 Appendix B. Predictions of a diffusion race model with two barriers

In this appendix, we outline the derivation of the diffusion race model tested in this study. This model is based on two diffusion processes (for auditory and visual channels), which are again assumed to be Wiener processes between two absorbing barriers. In accordance with standard race model assumptions, and in contrast to the DSM, processing is complete when either racer has finished, irrespective of the place of absorption. We assume independent racers ($\rho_{\mathrm{AV}} = 0$) because, for the general case with correlated processes A and V, the two-dimensional distribution of the first-passage time of a Wiener process between two absorbing barriers is needed, which we believe is yet unknown. For the special case of two uncorrelated racers, the density of the minimum of absorption time $f_{\mathrm{D}}(t)$ [$\mathbf{D} = \min(\mathbf{D}_{\mathrm{A}}, \mathbf{D}_{\mathrm{V}})$] is derived. This is a mixture distribution of the first-passage times of either racer, given that this racer finished first. Also, we have to consider two cases for each racer—that is, being absorbed at either the upper or lower barrier. Thus, we have four cases (cf. Townsend & Altieri, 2012). We consider one case (racer A finishes first at the upper barrier), and the remaining cases are treated analogously. The probability that racer A, denoted by the process $\mathbf{X}_{\mathrm{A}}(t)$, will win the race against racer V, denoted by the process $\mathbf{X}_{\mathrm{V}}(t)$, by absorption at the upper barrier ($\mathbf{A} \cap \mathbf{C} = 1$) is

$$P_{\mathrm{AV}}(\mathbf{A} \cap \mathbf{C} = 1) = \int_0^\infty P[\mathbf{X}_{\mathrm{A}}(t) \geq u] \times P[-\ell < \mathbf{X}_{\mathrm{V}}(t) < u]\mathrm{d}t,$$

which reduces to

$$\int_0^\infty f_{u,\mathrm{A}}(t) \times S_{\mathrm{V}}(t)\mathrm{d}t. \tag{B1}$$

Given that processes $\mathbf{X}_{\mathrm{A}}(t)$ and $\mathbf{X}_{\mathrm{V}}(t)$ are terminated at first passage, there must not be a prior passage of either u or $-\ell$. Therefore, the appropriate density function $f_{u,\mathrm{A}}(t)$ is used for racer A, and the survivor function $S_{\mathrm{V}}(t)$ for racer V in Eq. B1. This yields the probability that racer A finishes the race in the case of two synchronously presented targets. For the mean RT prediction, one is usually not interested in which racer produced the response, but whether a response was a correct response. We therefore add the probabilities that either A or V produced a response at the upper barrier, whatever the assignment of the barriers is (correct or incorrect):

$$\begin{aligned} P_{\mathrm{AV}}(\mathbf{C}=1) &= P(\mathbf{A}\cap\mathbf{C}=1) + P(\mathbf{V}\cap\mathbf{C}=1) \\ &= \int_0^\infty f_{u,\mathrm{A}}(t) \times S_{\mathrm{V}}(t)\mathrm{d}t + \int_0^\infty f_{u,\mathrm{V}}(t) \times S_{\mathrm{A}}(t)\mathrm{d}t. \end{aligned} \tag{B2}$$

Addition is valid here, because we have mutually exclusive events. Then the expected value of $\mathbf{D}$, conditional on upper absorption is

$$\begin{aligned} E_{\mathrm{AV}}(\mathbf{D}|\mathbf{C}=1) &= \frac{E(\mathbf{D}|\mathbf{A}\cap\mathbf{C}=1)\times P(\mathbf{A}\cap\mathbf{C}=1)}{P_{\mathrm{AV}}(\mathbf{C}=1)} \\ &\quad + \frac{E(\mathbf{D}|\mathbf{V}\cap\mathbf{C}=1)\times P(\mathbf{V}\cap\mathbf{C}=1)}{P_{\mathrm{AV}}(\mathbf{C}=1)} \\ &= \frac{\int_0^\infty t \times f_{u,\mathrm{A}}(t) \times S_{\mathrm{V}}(t)\mathrm{d}t + \int_0^\infty t \times f_{u,\mathrm{V}}(t) \times S_{\mathrm{A}}(t)\mathrm{d}t}{P_{\mathrm{AV}}(\mathbf{C}=1)} \end{aligned} \tag{B3}$$

For variance predictions, we calculate the second central moment, which is the expression above, but with t in the integrals of Eq. B3 replaced by t^2. These expressions are the predictions of redundant signals without onset asynchrony. For SOA-dependent predictions, one simply has to shift the time scale of that process

by some time τ. We define positive SOAs $\tau > 0$ as the cases in which V is presented before A. Then, the expressions B2 and B3 above generalize to

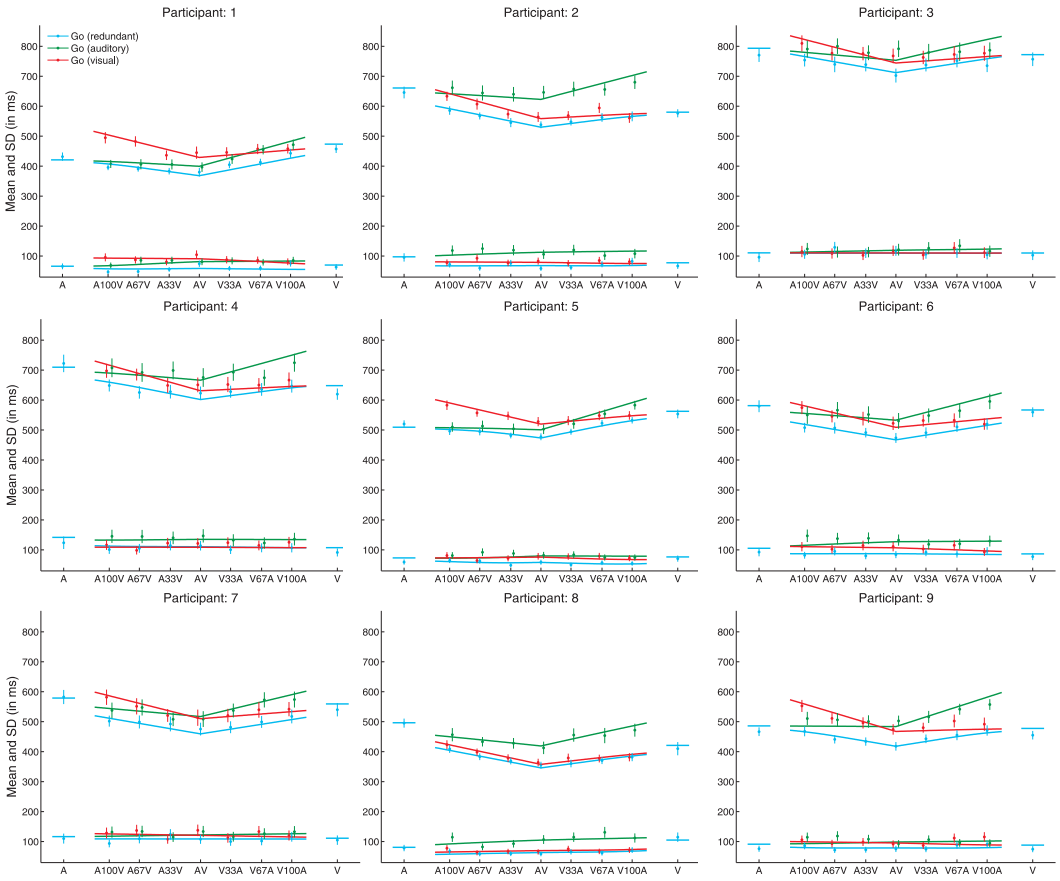
$$\begin{aligned} P_{\mathrm{AV}}(\mathbf{C}=1) &= P(\mathbf{A}\cap\mathbf{C}=1) + P(\mathbf{V}\cap\mathbf{C}=1) \\ &= \int_0^\infty f_{u,\mathrm{A}}(t-\tau)\times S_{\mathrm{V}}(t)\mathrm{d}t + \int_0^\infty f_{u,\mathrm{V}}\times S_{\mathrm{A}}(t-\tau)\mathrm{d}t \end{aligned} \tag{B4}$$

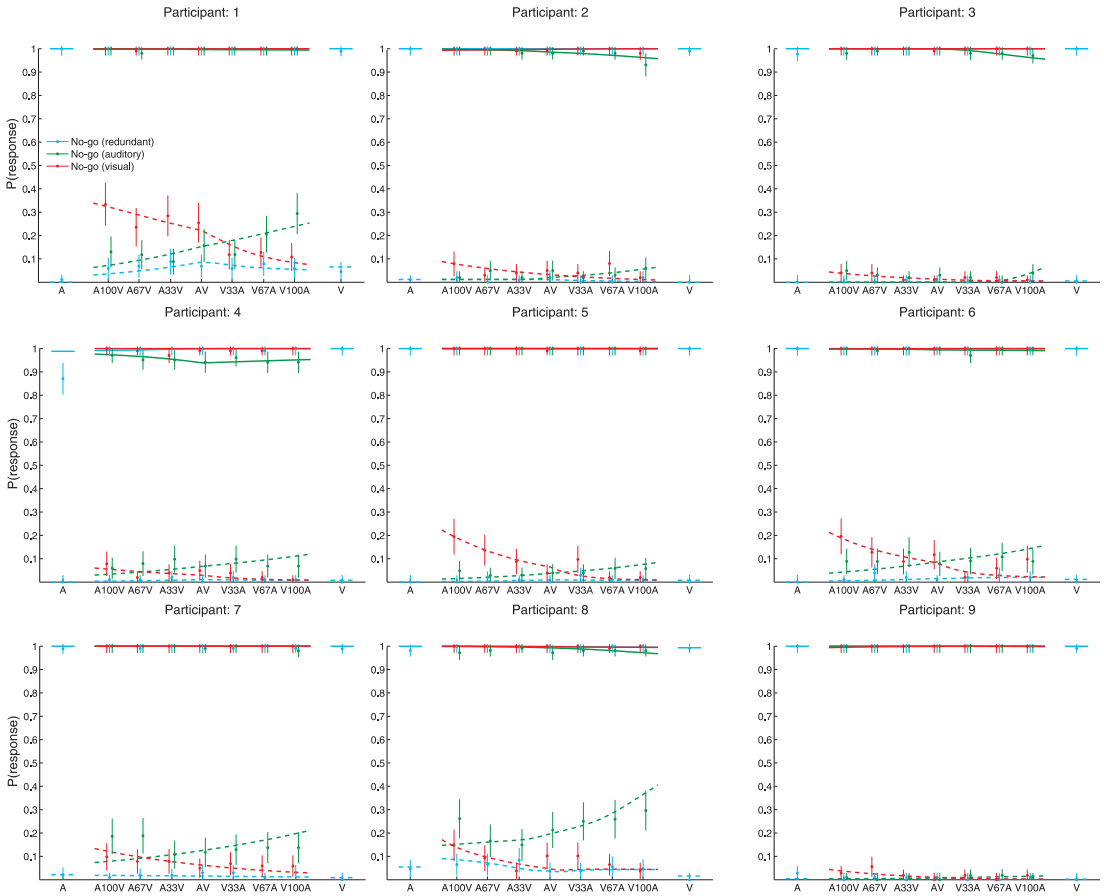
and

$$E(\mathbf{D}|\mathbf{C}=1) = \frac{\int_0^\infty t\times f_{u,\mathrm{A}}(t-\tau)\times S_{\mathrm{V}}(t)\mathrm{d}t + \int_0^\infty t\times f_{u,\mathrm{V}}(t)\times S_{\mathrm{A}}(t-\tau)\mathrm{d}t}{P_{\mathrm{AV}}(\mathbf{C}=1)} \tag{B5}$$

The case of $\tau < 0$ is derived likewise; in that case, all functions in Eqs. B4 and B5 with subscript V are delayed by τ. Clearly, predictions for response accuracy, mean RTs and variance in the unimodal cases are derived from the same expressions as in the coactivation model (DSM). Like the coactivation model, these results are moments of $\mathbf{D}$, not $\mathbf{T}$, so that residual parameters μ_{M} and σ^2_{M} are added to mean detection time and variance predictions, respectively (cf. Eqs. A7 and A8).

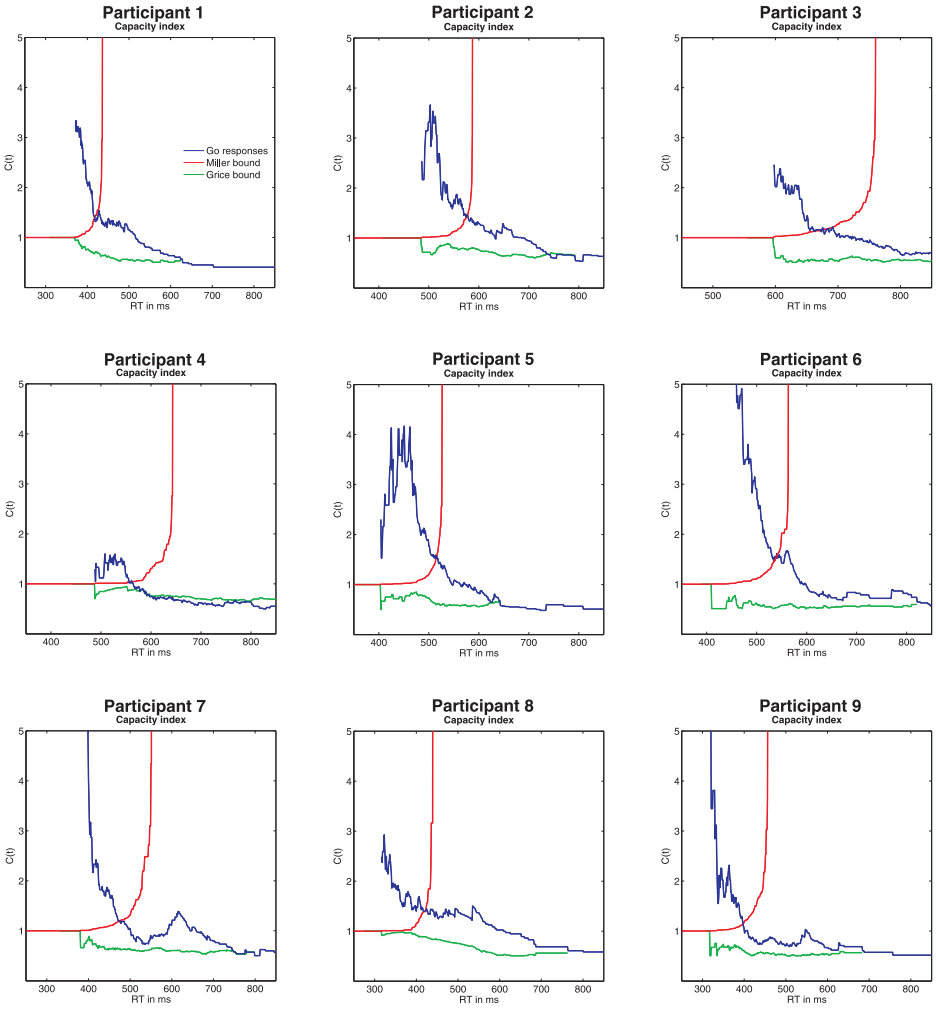
3.7 Supplemental Material



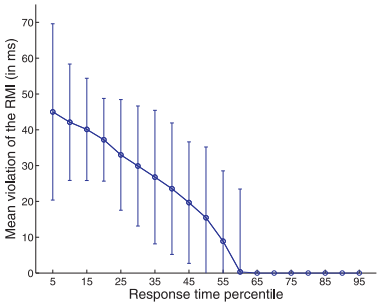


Supplemental Figure 3.1: Participant-specific mean and standard deviations of response times together with percent correct response observed in Experiment 1 with model predictions of the diffusion superposition model (DSM). Data (circles) and model predictions (curves) for the unimodal and bimodal redundant go conditions are displayed in blue, non-redundant stimulus conditions with non-informative visual or non-informative auditory stimulus are displayed in green or red, respectively. The same color coding was used for response frequency to no-go trials. The average of these data and predictions is shown in Figure 3.2 of the main text.

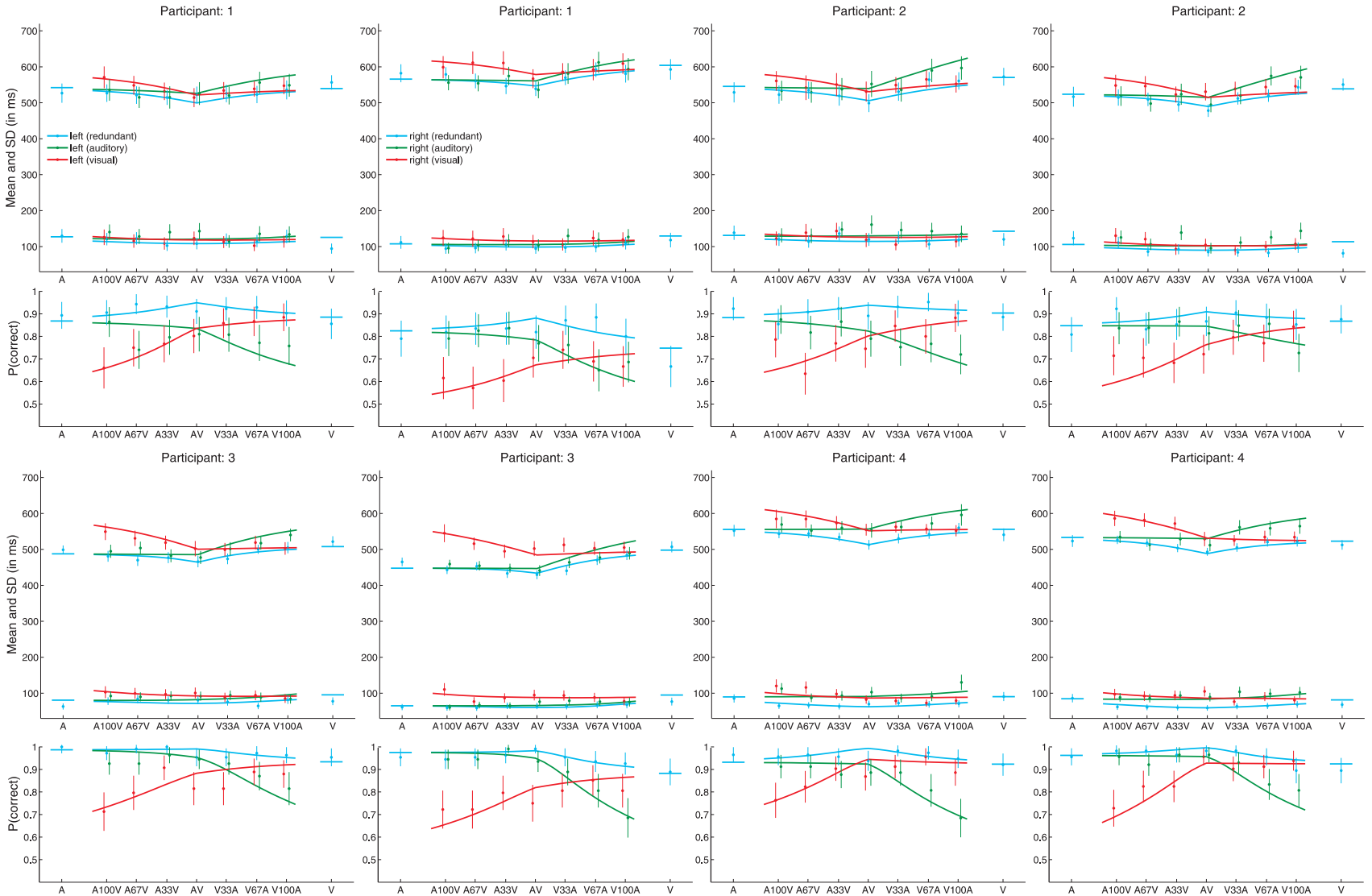
a


Participant 1
Capacity index
Participant 2
Capacity index
Participant 3
Capacity index
Participant 4
Capacity index
Participant 5
Capacity index
Participant 6
Capacity index
Participant 7
Capacity index
Participant 8
Capacity index
Participant 9
Capacity index
C(t)
RT in ms
Go responses
Miller bound
Grice bound

b


Mean violation of the RMI (in ms)
Response time percentile

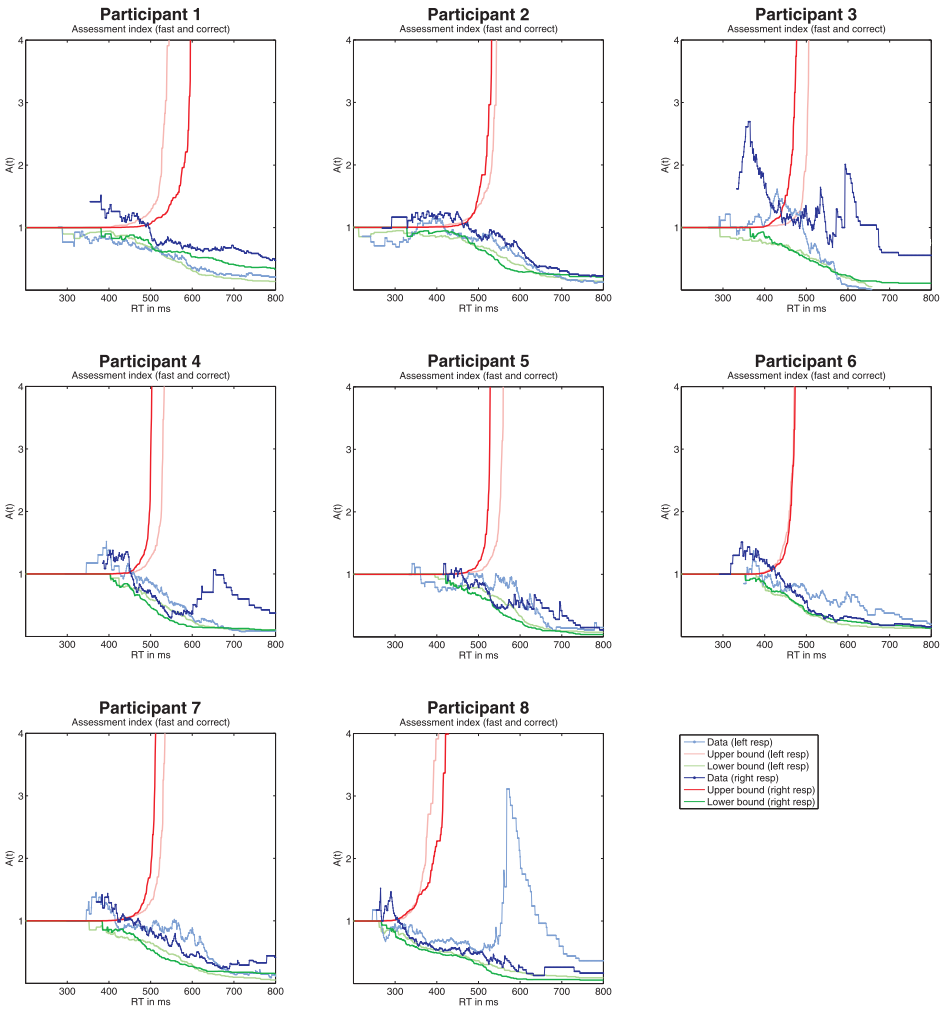
Supplemental Figure 3.2: **a**: Estimated capacity index $C(t)$ of Experiment 1 (blue lines) together with the Miller bound (red), indicative of supercapacity and Grice bound (green), indicative of very limited processing capacity (Townsend & Nozawa, 1995). The estimation of $C(t)$ is based on unimodal and bimodal redundant targets (no SOA). In every single participant, the Miller bound is violated at low t, suggesting supercapacity at those time points. The group test showed significant violations for these time points. **b**: Amount of violation of the RMI observed in unimodal and bimodal redundant (no SOA) conditions for all participants. Error bars represent 95 % confidence intervals and are only displayed for race model inequality violations (see Mordkoff & Danek, 2011).


Participant: 1
Participant: 1
Participant: 2
Participant: 2
Participant: 3
Participant: 3
Participant: 4
Participant: 4
Mean and SD (in ms)
P(correct)
left (redundant)
left (auditory)
left (visual)
right (redundant)
right (auditory)
right (visual)
A
A100V
A67V
A33V
AV
V33A
V67A
V100A
V

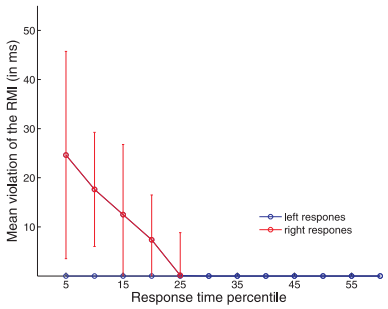
Participant: 5
Participant: 6
Participant: 7
Participant: 8
Mean and SD (in ms)
P(correct)
A
A100V
A67V
A33V
AV
V33A
V67A
V100A
V

Supplemental Figure 3.3: Participant-specific mean and standard deviations of response times together with percent correct response observed in Experiment 2 with model predictions of the diffusion superposition model (DSM). Data (circles) and model predictions (curves) for the unimodal and bimodal redundant conditions are displayed in blue, non-redundant stimulus conditions with auditory target or visual target are displayed in green or red, respectively. The averages of these predictions and data are displayed in Figure 3.4 of the main text.

a


Participant 1
Participant 2
Participant 3
Participant 4
Participant 5
Participant 6
Participant 7
Participant 8
Assessment index (fast and correct)
A(t)
RT in ms
Data (left resp)
Upper bound (left resp)
Lower bound (left resp)
Data (right resp)
Upper bound (right resp)
Lower bound (right resp)

b


Mean violation of the RMI (in ms)
Response time percentile
left respones
right respones

Supplemental Figure 3.4: **a**: Estimated assessment index $A(t)$ (Townsend & Altieri, 2012) of Experiment 2 for left response conditions (green circles) and right response conditions (red circles). The estimation of $A(t)$ is based on unimodal and bimodal redundant targets (no SOA) for the correct and fast case. Again, the respective upper bound (solid lines), indicative of super capacity, and the lower bound (dashed lines), indicative of very limited capacity, are shown for comparison. In all participants except Participant 5, capacity exceeds the upper bound (supercapacity) at low t for right correct and fast responses (red dots). For left responses, the upper bound is exceeded only by 5 participants and to a lesser amount compared to right responses. **b**: Amount of violation of the RMI observed in unimodal and bimodal redundant (no SOA) conditions for all participants and both response alternatives. Error bars represent 95% confidence intervals and are only displayed for race model inequality violations (Mordkoff & Danek, 2011) that were only found for right responses.

4 Fast and accurate calculations for cumulative first-passage time distributions in Wiener diffusion models

Steven P. Blurton[a], Miriam Kesselmeier[b], and Matthias Gondan[b]

[a]Department of Psychology, University of Regensburg, Germany

[b]Institute of Medical Biometry and Informatics, University of Heidelberg, Germany

Abstract[2]

We propose an improved method for calculating the cumulative first-passage time distribution in Wiener diffusion models with two absorbing barriers. This distribution function is frequently used to describe responses and error probabilities in choice reaction time tasks. The present work extends related work on the density of first-passage times [Navarro, D.J., Fuss, I.G. (2009). Fast and accurate calculations for first-passage times in Wiener diffusion models. Journal of Mathematical Psychology, 53, 222–230]. Two representations exist for the distribution, both including infinite series. We derive upper bounds for the approximation error resulting from finite truncation of the series, and we determine the number of iterations required to limit the error below a pre-specified tolerance. For a given set of parameters, the representation can then be chosen which requires the least computational effort

Keywords: Diffusion model; Wiener process; First passage times; Response times

[2]erschienen als: Blurton, S. P., Kesselmeier, M. & Gondan, M. (2012). Fast and accurate calculations for cumulative first-passage time distributions in Wiener diffusion models. *Journal of Mathematical Psychology*, *56*, 470–475. ©2012, reprinted with permission from Elsevier. The final publication is available at www.journals.elsevier.com/journal-of-mathematical-psychology (doi: 10.1016/j.jmp.2012.09.002).

4.1 Introduction

In the presence of two mutually exclusive competing risks, event times can often be described by a stochastic process drifting between two absorbing barriers. Typical examples include sequential sampling models of human decision making (e.g., Busemeyer & Townsend, 1993; Diederich, 1997; Ratcliff, 1978; Ratcliff & McKoon, 2008), or length of stay in hospital (with the two outcomes death and healthy discharge, e.g., Horrocks and Thompson (2004)). The central assumption of these models is that a hidden underlying state randomly moves between two alternatives until eventually one of two criteria is reached (so called absorbing barriers). The appeal in those models lies in the possibility to derive predictions not only for the probabilities for the two outcomes, but also for the time it takes until the barrier is hit. Often, this process is assumed to be continuous, resulting in the well-known diffusion models for which the time-homogeneous Brownian motion process (Wiener process, Fig. 4.1) is most popular.

The Wiener process $\mathbf{X}(t|v, \sigma^2)$ is described by two parameters v and $\sigma^2 > 0$ representing the drift and variance (noise) of the process. The process is temporally and spatially homogeneous, that is, drift and variance neither depend on the current state nor the time elapsed (e.g., Smith, 2000). In the two-alternative choice model, the process is assumed to start at $\mathbf{X}(0) = z$, and two absorbing barriers are assumed at zero and a, representing the two outcomes, $0 < z < a$. Despite the relative simplicity of the process, it is hard to derive expressions for the density and distribution of the firstpassage times in the two-barrier situation. One must rather rely on infinite series (Wald, 1947). Of course, the evaluation of infinite series can only involve a finite number of terms. The series, however, are known to converge, and it is possible to estimate the error that results when calculation is stopped at a certain number of steps. The usual approach is to terminate the calculation when a desired level of accuracy is met, for example, if the absolute error is lower than some tolerance $\varepsilon > 0$. This limit can be reached after evalua-

tion of very few terms when convergence is good at the point where the function is evaluated. On the other side, at critical points, sufficient accuracy requires the calculation of several hundred terms or even more.

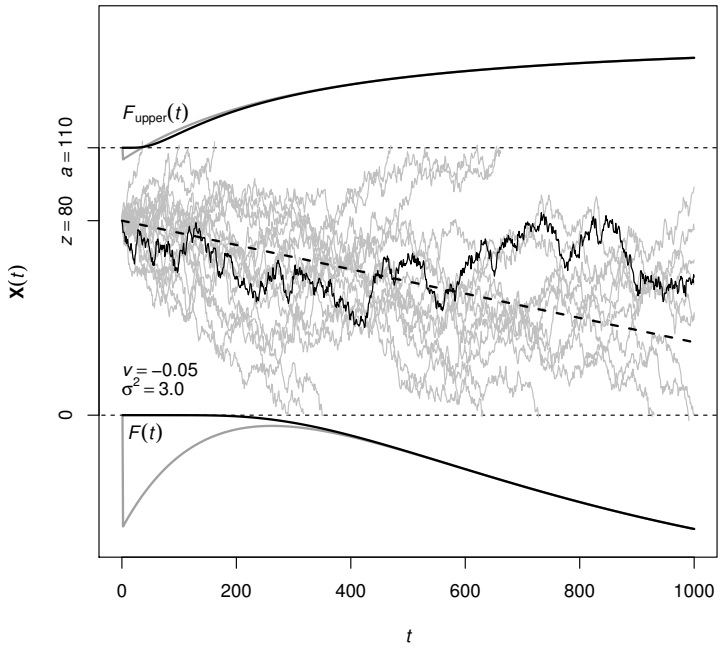


Figure 4.1: Realizations of a Wiener process with variance $\sigma^2 = 3.0$ and drift $v = -0.05$ (dashed line) starting at $z = 80$ between two absorbing barriers at $a = 110$ and zero. The black curves indicate the cumulative first-passage time distributions $F(t|v/\sigma, a/\sigma, w)$, $w = z/a$, at the lower barrier and at the upper barrier [the latter is determined via $F_{\text{upper}}(t \mid v/\sigma, a/\sigma, w) = F(t \mid -v/\sigma, a/\sigma, 1-w)$]. The gray lines show the approximation error resulting from early truncation at $K = 2$.

Two representations exist for the first-passage time density of a Wiener process between two absorbing barriers. These representations show different convergence behavior: While one representation converges quickly for small values of t, the other representation converges quickly for large values of t. Navarro and Fuss (2009) exploited these properties and provided a decision rule when to use the one or the other representation. The decision rule depends on the number of

terms needed to achieve a predefined level of accuracy. Based on this idea, we propose a computationally efficient way to compute the cumulative first-passage time distribution of a Wiener process between two absorbing barriers.

4.2 First-passage time density and distribution

We consider a Wiener Process with drift v, starting at $\mathbf{X}(0) = z$. Without loss of generality, the variability can be fixed at $\sigma^2 = 1$, since it only scales the other parameters. A lower and an upper absorbing barrier is fixed at zero and a, with $0 < z < a$. Setting, for convenience, $w = z/a$, the density $f(t)$ of first absorption at the lower barrier is described by the two series (e.g., Feller, 1968, p. 359 and p. 370),

$$f^{\ell}(t \mid v, a, w) = \frac{\pi}{a^2} \exp\left(-vaw - \frac{v^2 t}{2}\right) \times \sum_{k=1}^{\infty} k \sin(\pi k w) \exp\left[-\frac{1}{2}\left(\frac{k\pi}{a}^2\right) t\right],$$

$$f^{s}(t \mid v, a, w) = \frac{a}{\sqrt{t^3}} \exp\left(-vaw - \frac{v^2 t}{2}\right) \times \sum_{k=-\infty}^{\infty} (2k + w) \varphi\left(\frac{2k + w}{\sqrt{t}} a\right),$$

with $\varphi(x)$ denoting the standard normal density. Because absorption can occur at both the upper and the lower barrier, f^{ℓ} and f^{s} are, in fact, subdensities and do not fully integrate to one. Whereas f^{ℓ} converges quickly for large t, f^{s} converges quickly for small t (Horrocks & Thompson, 2004; Van Zandt, Colonius, & Proctor, 2000). First absorption at the upper barrier is described by $f(t \mid -v, a, 1 - w)$; the lower density is given by $f(t \mid v/\sigma, a/\sigma, w)$.

Navarro and Fuss (2009) investigated the numerical properties of the two representations truncated at some K. In particular, they provide upper bounds for the error which results when the series f^{ℓ}, f^{s} are evaluated for $k = 1, \ldots, K$ and $k = -K, \ldots, K$, respectively. They derived expressions for the required number of summands K which limit the truncation error $|f(t) - f_K(t)|$ below a certain

criterion $\varepsilon > 0$ for each representation. For a prespecified set of parameters, the representation which is computationally least demanding can then be chosen.

In the present note we consider the cumulative first-passage time distribution, that is, the probability of absorption between time zero and some t, which is given by the integral of the density between zero and t. This distribution is again described by two alternative series with different convergence properties. Similar to Navarro and Fuss (2009) we derive upper bounds for the number of iterations K required to limit the truncation error below a certain tolerance $\varepsilon > 0$ for both representations. For a given set of parameters, the distribution is then determined using the representation which requires least computational effort. The decision is, thus, based on the number of iterations K, multiplied by the time it takes for each iteration.

4.3 Large-time representation

The large-time representation of the subdistribution of firstpassage times (e.g., Ratcliff and Tuerlinckx (2002), Eq. B1; Ratcliff (1978), Eq. A12) is obtained by integration of the large-time density $f^{\ell}(t)$ over $[0, t]$. Equivalently, the integral of $f^{\ell}(t)$ over $[t, \infty]$ is subtracted from the total probability P of absorption at the lower barrier

$$F^{\ell}(t) = P - \frac{2\pi}{a^2} \exp\left(-vaw - \frac{v^2 t}{2}\right) \times \sum_{k=1}^{\infty} \frac{k \sin(\pi k w)}{v^2 + (k\pi/a)^2} \exp\left[-\frac{1}{2}\left(\frac{k\pi}{a}\right)^2 t\right] \quad (1)$$

with

$$P = \begin{cases} \frac{1-\exp[-2va(1-w)]}{\exp(2vaw)-\exp[-2va(1-w)]}, & \text{for } v \neq 0 \\ 1 - w, & \text{for } v = 0. \end{cases}$$

This workaround is necessary because term-wise integration of the infinite series

$f^\ell(t)$ over $[0,t]$ requires uniform convergence of $f^\ell(t)$ within the range of integration, which can be demonstrated for positive t only (Appendix A).

When determining $F^\ell(t)$, the series must be truncated at some $K \geq 1$. The number of summands K should be chosen such that the truncation error $|F^\ell(t) - F^\ell_K(t)|$ is below some tolerance $\varepsilon > 0$, that is,

$$\left| \frac{2\pi}{a^2} \exp\left(-vaw - \frac{v^2 t}{2}\right) \times \sum_{k=K+1}^{\infty} \frac{k \sin(\pi k w)}{v^2 + (k\pi/a)^2} \exp\left[-\frac{1}{2}\left(\frac{k\pi}{a}\right)^2 t\right] \right| \leq \varepsilon. \quad (2)$$

To this end, two conditions for K must be satisfied,

$$K^2 \geq \frac{1}{t}\left(\frac{a}{\pi}\right)^2, \text{ and}$$

$$K^2 \geq -\frac{2}{t}\left(\frac{a}{\pi}\right)^2 \left\{ \log\left[\frac{\varepsilon \pi t}{2}\left(v^2 + \frac{\pi^2}{a^2}\right)\right] + vaw + \frac{v^2 t}{2} \right\}. \quad (3)$$

A detailed derivation is found in Appendix A. Briefly, the expression in (2) is simplified by omitting the sine and limiting k at 1 in the denominator of the fraction behind the Σ. The exponential series is then replaced by an integral representing its upper bound and solved for K. As expected, the number of required terms increases monotonically with ε and decreases with t—hence the name of (1), "large-time representation". For small t, K tends to infinity (Fig. 4.2).

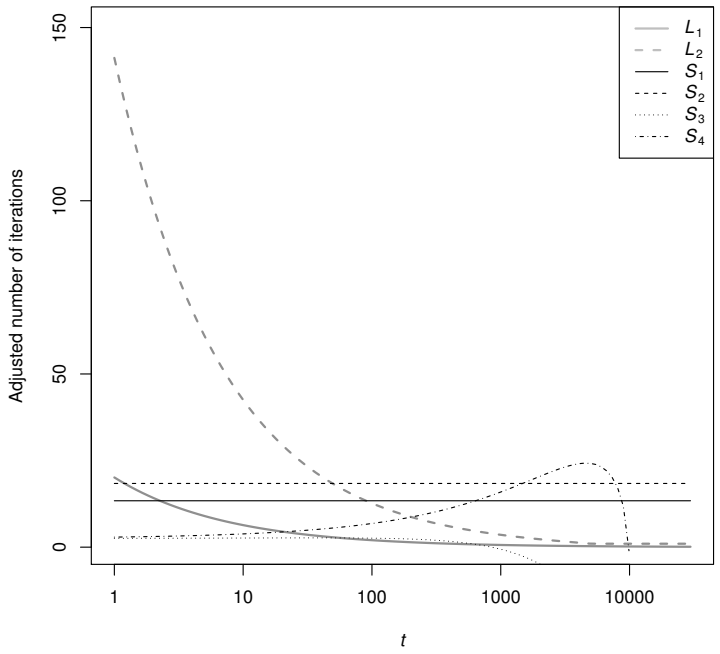


Figure 4.2: Convergence of the two representations of the cumulative first-passage time distribution for $v = -0.06$, $a = 63.2$, $w = 0.5$, and $\varepsilon = 1.5 \times 10^{-8}$ which is the square root of floating point precision with doubles. $L1$, $L2$: Criteria for the large-time representation (3), with the subscripts denoting the square root of the criteria given in Appendix A. The required number of iterations is given by the ceiling of the maximum of $L1$ and $L2$. $S1$, $S2$, $S3$, $S4$: Criteria for the small-time representation (5). The subscripts denote the respective expressions of Appendix B. The adjusted computational effort is again given by the ceiling of the maximum of $S1$, $S2$, $S3$ and $S4$, multiplied by 10 to account for the increased computing demands of the small-time representation.

4.4 Small-time representation

The second representation of the cumulative first-passage time distribution is obtained by integration of the small-time density (e.g., Horrocks & Thompson, 2004):

$$F^s(t) = P - \text{sgn } v \cdot \sum_{k=-\infty}^{\infty} \left[\exp(-2vak - 2vaw) \times \Phi \left(\text{sgn } v \frac{2ak + aw - vt}{\sqrt{t}} \right) \right.$$

$$- \exp(2vak)\, \Phi\left(\operatorname{sgn} v \frac{-2ak - aw - vt}{\sqrt{t}}\right)\Bigg], \tag{4}$$

with P defined as in (1) and $\Phi(x)$ denoting the cumulative standard normal distribution. As before, the series in (4) describes the survivor function, that is, the probability for absorption between t and infinity, such that the result is again subtracted from the probability P of absorption at the lower barrier. Although this representation is undefined for $t = 0$, $\lim_{t\to 0} F^s(t)$ can be shown to be zero, and the series shows good convergence for small $t > 0$. Despite the name, convergence is acceptable for large t; the series is computationally expensive, however, for drift rates near zero. We first consider negative drift $v < 0$ (denoted by an additional superscript), that is, we are interested in a process with drift towards the lower barrier. Truncation of $F^s.(t)$ at some $K \geq 1$ yields a truncation error $|F^{s-}(t) - F^{s-}(t)_K|$ which should again be below $\varepsilon > 0$,

$$\left| \sum_{|k|=K+1}^{\infty} \Bigg[\exp(2vak)\, \Phi\left(\frac{2ak + aw + vt}{\sqrt{t}}\right) - \exp(-2vak - 2vaw)\, \Phi\left(\frac{-2ak - aw + vt}{\sqrt{t}}\right)\Bigg] \right| \leq \varepsilon.$$

As shown in Appendix B, three conditions must be satisfied for K,

$$K \geq w - 1 + \frac{1}{2va} \log\left\{\frac{\varepsilon}{2}[1 - \exp(2va)]\right\},$$

$$K \geq \frac{0.535\sqrt{2t} + vt + aw}{2a}, \text{ and}$$

$$K \geq \frac{w}{2} - \frac{\sqrt{t}}{2a}\, \Phi^{-1}\left[\frac{\varepsilon a}{0.3\sqrt{2\pi t}} \exp\left(\frac{v^2 t}{2} + vaw\right)\right]. \tag{5}$$

As illustrated in Fig. 4.2, the first requirement dominates the criteria over a large range of t. Expression (4) is computationally more complex than the large-time representation (2). For a fixed K, repeated evaluation with different parameters showed $F^s(t)$ to be about ten times slower than $F^\ell(t)$.

For positive drift,

$$F^{s+}(t) = P - \sum_{k=-\infty}^{\infty} \left[\exp(-2vaw - 2vak)\, \Phi\left(\frac{2ak + aw - vt}{\sqrt{t}}\right) - \exp(2vak)\, \Phi\left(\frac{-2ak - aw - vt}{\sqrt{t}}\right) \right].$$

As shown in Appendix B, the number of required summands can be determined using the criteria (5), with $v' = v$ instead of v and a modified tolerance criterion $\varepsilon' = \varepsilon \exp(-2vaw)$. In the zero drift case, the series simplifies to

$$F^{s0}(t) = 2 \sum_{k=0}^{\infty} \left[\Phi\left(\frac{-2ak - aw}{\sqrt{t}}\right) - \Phi\left(\frac{-2ak - 2a + aw}{\sqrt{t}}\right) \right],$$

and evaluation of $K \geq \frac{w}{2} - \frac{\sqrt{t}}{2a} \Phi^{-1}\left(\frac{\varepsilon}{2-2w}\right)$ terms guarantees a finite truncation error below ε.

4.5 Discussion

The present paper provides finite approximations of the cumulative first-passage times in the two-barrier diffusion model that controls the approximation error below a pre-specified tolerance. By comparing the required number of iterations in the two representations (2) and (4), and adjusting for the time necessary to evaluate a single summand of the series, the representation which requires least computational effort can be chosen. The present approach is to be preferred over *ad hoc* methods in which evaluation of the series is stopped when a single term is below the tolerance: When truncation is based on the absolute value of a single

summand, the truncation error might be larger than expected because an infinite number of summands is dropped. To overcome this limitation, current implementations of the method sometimes evaluate a much larger number of summands than necessary. Here we propose to control the truncation error of the entire set of truncated summands. Precision is, therefore, controlled uniformly for all parameter combinations, which yields a smooth surface for numerical likelihood maximization (e.g., Horrocks & Thompson, 2004).

In some applications, other parameter estimation procedures might be more suitable. For example, for the well known diffusion model (Ratcliff, 1978), an algorithm for calculation of the distribution function has been proposed by Voss and colleagues (Voss, Rothermund, & Voss, 2004; Voss & Voss, 2007[3], Eq. A9). The approach of Voss and colleagues is similar to ours, but they derive an expression for the required number of steps using the large-time representation (1) only. In the general case, this threshold is far too conservative, especially for small error bounds. Alternatively, discrete approximations (e.g., random walks) to continuous diffusion processes offer more complex, yet more flexible implementations of diffusion processes (Diederich & Busemeyer, 2003).

Applications of the proposed method arise in fitting Ratcliff's (1978) diffusion model to observed response times, for example, from two-alternative choice tasks. Several methods have been proposed for this purpose, none of which can be said to be uniformly superior to the other methods (Ratcliff & Tuerlinckx, 2002, pp. 443f). The so-called chi-square fitting method and the weighted least squares fitting method make heavy use of the cumulative first-passage time distribution $F(t)$. In contrast, likelihood maximization primarily uses the density $f(t)$ of the absorption times. In the latter approach, the distribution $F(t)$ is still needed in the presence of censored observations. Censoring occurs, for example, when the observer is unable to decide between two alternatives within a reasonable amount of

[3]Die Referenz in der Originalarbeit ist nicht korrekt; die korrekte Referenz lautet: Voss & Voss (2008)

time, or when responses are registered during a short time window in fixed stimulation protocols (e.g., in fMRI experiments). Then, absorption might be known to have occurred at the upper barrier, but it is only known to have occurred later than some t ("misses"). A diffusion model with a deadline parameter could account for this, making use of the distribution $F(t)$, because the likelihood contribution then corresponds to the upper subsurvivor function at t. If absorption is only known to have occurred later than some t, and the outcome is unknown because no response has been given, the likelihood contribution corresponds to the sum of the upper and the lower subsurvivor function at t. The present method, thus, complements Navarro and Fuss' (2009) work on the density representation and will allow for the efficient parameter adjustment of diffusion models of competing risks even in the presence of censored observations.

4.6 Appendix A. Integral and convergence of the large-time representation

By collapsing the two exponentials, the density $f^{\ell}(t)$ is restated as a series of exponentials of t,

$$f^{\ell}(t) = \frac{\pi}{a^2} \exp(-vaw) \times \sum_{k=1}^{\infty} k \sin(\pi k w) \exp\left\{ -\frac{1}{2} \left[v^2 + \left(\frac{k\pi}{a} \right)^2 \right] t \right\}.$$

Summand-wise integration of $f^{\ell}(t)$ over the interval $[\tau, \infty]$, $\tau > 0$ requires uniform convergence of $f^{\ell}(t)$ within that interval. This can be shown, for example, by the so-called majorant criterion (Weierstrass M-test). To this end, we define an upper bound M_{τ} for $f^{\ell}(t)$ with a small $\tau > 0$, and drop the sine. Because $\sin x$ cannot exceed 1 and $\exp(-ct)$, $c > 0$, monotonically decreases in t,

$$\left|f^{\ell}(t)\right| \leq M_{\tau} = \frac{\pi}{a^2}\exp(-vaw) \times \sum_{k=1}^{\infty} k \exp\left\{-\frac{1}{2}\left[v^2 + \left(\frac{k\pi}{a}\right)^2\right]\tau\right\}, \text{ for } t \geq \tau.$$

The series $f^{\ell}(t)$ then converges if M_{τ} converges. Convergence of M_{τ} can be shown by the integral test because M_{τ} is positive-valued and strictly monotonically decreasing in k. As the integral $\int_1^{\infty} k \exp\left[-\frac{\tau}{2}\left(\frac{\pi}{a}\right)^2 k^2\right] \mathrm{d}k$ exists and is finite, M_{τ} converges, such that $f^{\ell}(t)$ is uniformly convergent within $[\tau, \infty]$. Because all summands are exponentials of t, the antiderivative of $f^{\ell}(t)$, $t \geq 0$, is easily found:

$$\int_{\tau}^{t} f^{\ell}(s)\mathrm{d}s = -\frac{2\pi}{a^2}\exp(-vaw) \times \sum_{k=1}^{\infty} \frac{k\sin(\pi k w)}{v^2 + \left(\frac{k\pi}{a}\right)^2} \exp\left\{-\frac{1}{2}\left[v^2 + \left(\frac{k\pi}{a}\right)^2\right]t\right\}\Bigg|_{\tau}^{t}.$$

The distribution function $F^{\ell}(t)$ is then obtained by subtracting the integral of $f^{\ell}(t)$ for $[\tau, \infty]$ from the total proportion P of absorptions at the upper barrier,

$$F^{\ell}(t) = P - \frac{2\pi}{a^2}\exp\left(-vaw - \frac{v^2 t}{2}\right) \times \sum_{k=1}^{\infty} \frac{k\sin(\pi k w)}{v^2 + \left(\frac{k\pi}{a}\right)^2} \exp\left[-\frac{1}{2}\left(\frac{k\pi}{a}\right)^2 t\right]. \quad \text{(A1)}$$

What happens if the series in (A1) is truncated after evaluations of K terms? In order to guarantee that the approximation error is below a certain tolerance $\varepsilon > 0$, the absolute difference between the full series $F^{\ell}(t)$ and the truncated series $F_K^{\ell}(t)$ must be kept below the tolerance,

$$\left|F^{\ell}(t) - F_K^{\ell}(t)\right| =$$

$$\left|\frac{2\pi}{a^2}\exp\left(-vaw - \frac{v^2 t}{2}\right) \times \sum_{k=K+1}^{\infty} \frac{k\sin(\pi k w)}{v^2 + (\frac{k\pi}{a})^2} \exp\left[-\frac{1}{2}\left(\frac{k\pi}{a}\right)^2 t\right]\right| \leq \varepsilon.$$

To determine the conditions for K some estimations have to be carried out. A first upper bound for $\left|F^{\ell}(t) - F_K^{\ell}(t)\right|$ is obtained by noting that $0 \leq |\sin x| \leq 1$ and by setting $k = 1$ in the denominator as a lower bound for the fraction before the exponential. Thus, $\left|F^{\ell}(t) - F_K^{\ell}(t)\right| \leq \varepsilon$ if

$$\frac{2\pi}{a^2} \exp\left(-vaw - \frac{v^2 t}{2}\right) \times \sum_{k=K+1}^{\infty} \frac{k}{v^2 + \left(\frac{\mathbf{1}\cdot\pi}{a}\right)^2} \exp\left[-\frac{1}{2}\left(\frac{k\pi}{a}\right)^2 t\right] \leq \varepsilon.$$

The factor before the sum is positive. Truncation of the series should, thus, be limited to those K for which the elements decrease in k. The first derivative of the function $h(k) = k \exp\left[-\frac{1}{2}\left(\frac{k\pi}{a}\right)^2 t\right]$ must, therefore, be negative, which is guaranteed if K^2 is greater than

$$L_1 = \frac{1}{t}\left(\frac{a}{\pi}\right)^2. \tag{A2}$$

Since the elements decrease, an upper bound for the error series $\sum_{k=K+1}^{\infty} h(k)$ is given by the integral of $h(k)$ within K and infinity, so that

$$\left|F^{\ell}(t) - F_K^{\ell}(t)\right| \leq \frac{2\pi}{a^2} \exp\left(-vaw - \frac{v^2 t}{2}\right) \times \frac{1}{v^2 + (\pi/a)^2} \int_K^{\infty} h(k)\mathrm{d}k,$$

which is below ε if $K^2 \geq L_2$ with

$$L_2 = -\frac{2}{t}\left(\frac{a}{\pi}\right)^2 \left\{\log\left[\frac{\varepsilon\pi t}{2}\left(v^2 + \frac{\pi^2}{a^2}\right)\right] + vaw + \frac{v^2 t}{2}\right\}. \tag{A3}$$

For large t, the condition silently holds. In the other cases, K is set to the ceiling of $\max(\sqrt{L_1}, \sqrt{L_2})$.

In the zero drift case $v = 0$, Expression (A1) simplifies to

$$F^{\ell 0}(t) = P - \frac{2}{\pi} \sum_{k=1}^{\infty} \frac{\sin(\pi k w)}{k} \exp\left[-\frac{1}{2}\left(\frac{k\pi}{a}\right)^2 t\right].$$

The truncation error is controlled if the series is evaluated until K^2 is above $L_1^0 = L1$ and

$$L_2^0 = -\frac{2}{t}\left(\frac{a}{\pi}\right)\log\left(\frac{\varepsilon\pi^3 t}{2a^2}\right).$$

4.7 Appendix B. Convergence of the small-time representation

We first consider negative drift ($v < 0$), denoted by an additional superscript. The truncation error of the small-time version of the subdistribution is most easily controlled by decomposing $F^{s-}(t)$ (which is known to be finite) into three distinct series:

$$\begin{aligned} F^{s-}(t) = \; P - \sum_{k=0}^{\infty} \Bigg[& \exp(2vak)\,\Phi\left(\frac{2ak + aw + vt}{\sqrt{t}}\right) \\ & - \exp(-2vak - 2vaw)\,\Phi\left(\frac{-2ak - aw + vt}{\sqrt{t}}\right)\Bigg] \\ - \sum_{k=1}^{\infty} & \exp(-2vak)\,\Phi\left(\frac{-2ak + aw + vt}{\sqrt{t}}\right) \\ + \sum_{k=1}^{\infty} & \exp(2vak - 2vaw)\,\Phi\left(\frac{2ak - aw + vt}{\sqrt{t}}\right). \end{aligned} \tag{B1}$$

All series are positive, the truncation error of the sum is, thus, guaranteed to be below the error tolerance ε if the approximation error of each summand is controlled at $\varepsilon/2$.

Denoting the inverse Gaussian distribution by $W(t \,|\, c, \mu) = \Phi\left(\frac{\mu t - c}{\sqrt{t}}\right) + \exp(2c\mu)\Phi\left(\frac{-\mu t - c}{\sqrt{t}}\right)$, the first series in Expression (B1) can be rewritten as $\sum_{k=0}^{\infty}\exp(2vak)[1 - W(t \,|\, 2ak + aw, -v)]$. Again, we truncate after K summands have been evaluated. Because $W(t)$ is bounded between 0 and 1, and v is negative,

$\exp(2vak)$ is recognized as a decreasing geometric series:

$$\sum_{k=K+1}^{\infty} \exp(2vak)[1-W(t\,|\,2ak+aw,-v)] \leq \sum_{k=K+1}^{\infty} \exp(2vak) = \frac{\exp[2va(K+1)]}{1-\exp(2va)}.$$

The result is below the tolerance $\varepsilon/2$ for K greater than

$$S_1 = -1 + \frac{1}{2va} \log\left\{\frac{\varepsilon}{2}[1-\exp(2va)]\right\},$$

independent of t. Similarly, the last series in (B1) has converged for K above

$$S_2 = w + S_1,$$

which, of course, includes $S1$.

In the second series in (B1), large exponentials are multiplied with tiny $\Phi(-x)$, such that the product is finite. An upper bound for $\Phi(-x)$ is given by Ermolova and Haggman (2004),

$$\Phi(-x) = \frac{1}{2}\mathrm{erfc}\left(\frac{x}{\sqrt{2}}\right) \leq 0.3\,\exp\left(-1.01\frac{x^2}{2}\right) \leq 0.3\,\exp\left(-\frac{x^2}{2}\right).$$

The Ermolowa–Haggman bound requires the argument of $\mathrm{erfc}(\cdot)$ to be greater than 0.535, which is satisfied if K is greater than

$$S_3 = \frac{0.535\sqrt{2t} + vt + aw}{2a}.$$

Application of the bound to the second series in (B1) yields exponentials decreasing in k, for which an upper bound is given by their integral. This integral is then recognized as a normal distribution:

$$\begin{aligned}
&\sum_{k=K+1}^{\infty} \exp(-2vak)\Phi\left(\frac{-2ak+aw+vt}{\sqrt{t}}\right) \\
&\qquad \le 0.3\exp\left(-vaw-\frac{v^2t}{2}\right)\sum_{k=K+1}^{\infty}\exp\left[-\frac{(2ak-aw)^2}{2t}\right] \\
&\qquad \le 0.3\exp\left(-vaw-\frac{v^2t}{2}\right)\int_K^{\infty}\exp\left[-\frac{(2ak-aw)^2}{2t}\right]\mathrm{d}k \\
&\qquad = \frac{0.3}{2a}\sqrt{2\pi t}\exp\left(-vaw-\frac{v^2t}{2}\right)\Phi\left(\frac{aw-2aK}{\sqrt{t}}\right).
\end{aligned}$$

The result is below $\varepsilon/2$ if

$$\Phi\left(\frac{aw-2aK}{\sqrt{t}}\right) \le \frac{\varepsilon a}{0.3\sqrt{2\pi t}}\exp\left(\frac{v^2t}{2}+vaw\right).$$

If the right hand side is larger than one, the condition silently holds. In the other cases, standard approximations for the quantile function of the normal distribution can be used to solve for K which must be greater than

$$S_4 = \frac{w}{2}-\frac{\sqrt{t}}{2a}\Phi^{-1}\left[\frac{\varepsilon a}{0.3\sqrt{2\pi t}}\exp\left(\frac{v^2t}{2}+vaw\right)\right].$$

For positive drift $v > 0$,

$$\begin{aligned}
F^{s+}(t) = P - \sum_{k=-\infty}^{\infty}\Bigg[&\exp(-2vak-2vaw)\times\Phi\left(\frac{2ak+aw-vt}{\sqrt{t}}\right) \\
&-\exp(2vak)\Phi\left(\frac{-2ak-aw-vt}{\sqrt{t}}\right)\Bigg],
\end{aligned}$$

with truncation error

$$\left|F^{s+}(t) - F_K^{s+}(t)\right| = \sum_{k=K+1}^{\infty} \left[\exp(-2vak - 2vaw) \times \Phi\left(\frac{2ak + aw - vt}{\sqrt{t}}\right) - \exp(2vak)\,\Phi\left(\frac{-2ak - aw - vt}{\sqrt{t}}\right) \right].$$

The truncation error for positive drift, thus, corresponds to $\exp(2vaw)$ times the error for negative drift, $\exp(2vaw)\,|\,F^{s+}(t\,|\,v,a,w) - F_K^{s+}(t\,|\,v,a,w)| = |F^{s-}(t\,|\,-v,a,w) - F_K^{s-}(t\,|\,-v,a,w)|$. The required number of iterations for $v > 0$ can, therefore, be determined using the expressions for $v' = -v$ with a stricter criterion $\varepsilon' = \varepsilon \exp(-2vaw)$.

In the special case of zero drift, the series reduces to

$$F^{s0}(t) = 2\sum_{k=0}^{\infty} \left\{ \Phi\left[\frac{-2ak - a + a(1-w)}{\sqrt{t}}\right] - \Phi\left[\frac{-2ak - a - a(1-w)}{\sqrt{t}}\right] \right\}.$$

The expression can be illustrated as series of bands of width $2(1-w)$ along the negative tail of a normal distribution (e.g., Fig 2.2 in Horrocks, 1999) with mean zero and variance t/a^2,

$$F^{s0}(t) = 2\sum_{k=0}^{\infty} \int_{-2k-1-(1-w)}^{-2k-1+(1-w)} \varphi\left(x \,\middle|\, 0, \frac{t}{a^2}\right) \mathrm{d}x,$$

such that the truncation error $|F^{s0}(t) - F_K^{s0}(t)|$ is below $2(1-w) \times \Phi(-2K + w\,|\,0, t/a^2)$. The latter expression fulfills the tolerance criterion ε if $K \geq \frac{w}{2} - \frac{\sqrt{t}}{2a} \cdot \Phi^{-1}\left(\frac{\varepsilon}{2-2w}\right)$.

4.8 Appendix C. Supplementary data

Supplementary material related to this article can be found online at `http://dx.doi.org/10.1016/j.jmp.2012.09.002`. This online supplementary material

contains scripts written in R statistical language (R core team, 2012) and Matlab which contain seven helper functions and a main function:

- exp_pnorm(a, b): In $F^s(t)$, large positive a coincide with large negative b in $\exp(a) \cdot \Phi(b)$. This raises numerical issues because values near infinity are multiplied with values near zero. In such cases, we approximated the normal distribution by an exponential (Kiani, Panaretos, Psarakis & Saleem, 2008) such that the product is determined via $\exp(a + b)$ which is numerically feasible.

- K_large(t, v, a, w, epsilon): number of summands needed for the large-time representation of the upper subdistribution $F^\ell(t|v, a, w)$. The parameters denote the time, the drift of the process, the upper barrier a, the relative start point w and the tolerance bound epsilon, respectively. Time t can be a vector.

- K_small(t, v, a, w, epsilon): same for small-time representation $F^s(t|v, a, w)$.

- Pu(v, a, w): calculates the probability of absorption at the lower barrier.

- Fl_lower(t, v, a, w, K): calculates the lower subdistribution using the large-time representation $F^\ell(t|v, a, w)$. The number of summands is given by K.

- Fs_lower(t, v, a, w, K): same for small-time representation $F^s(t|v, a, w)$.

- Fs0_lower(t, a, w, K): same for small-time representation $F^s(t|v = 0, a, w)$.

- F_lower(t, v, a, w, sigma2, epsilon): This is the main function for determining the cumulative first-passage time distribution at the lower barrier for a Wiener process with drift mu and variance sigma2 between two absorbing barriers at 0 and a > 0. The function invokes K_small and K_large to determine the number of summands required to attain precision epsilon > 0. The time points t and parameters can be given as vectors. For each element of

the vectors t and the parameters, the function automatically selects the representation which requires less terms. Negative drifts and non-unit variances sigma2 are handled.

- F_upper(t, v, a, w, sigma2, epsilon): First-passage time distribution at the upper barrier a.

Example usage

The cumulative first-passage time distributions shown in Figure 4.1 of the main article have been determined by F_lower(t=1:1000, v=-0.05, a=110, w=80/110, sigma2=3, epsilon=1.5e-8) and F_upper(...) with the same arguments.

R script

```
# Calculates exp(a) * pnorm(b) using an approximation by Kiani et al.  (2008)
exp_pnorm = function(a, b)
{
r = exp(a) * pnorm(b)
d = is.nan(r) & b < -5.5
r[d] = 1/sqrt(2) * exp(a - b[d]*b[d]/2) * (0.5641882/b[d]/b[d]/b[d] - 1/b[d]/sqrt(pi))
r
}
# Number of terms required for large time representation
K_large = function(t, v, a, w, epsilon)
{
sqrtL1 = sqrt(1/t) * a/pi
sqrtL2 = sqrt(pmax(1, -2/t*a*a/pi/pi * (log(epsilon*pi*t/2 * (v*v + pi*pi/a/a)) +
v*a*w + v*v*t/2)))
ceiling(pmax(sqrtL1, sqrtL2))
}
```

```
# Number of terms required for small time representation
K_small = function(t, v, a, w, epsilon)
{
if(abs(v) < sqrt(.Machine$double.eps)) # zero drift case
return(ceiling(pmax(0, w/2 - sqrt(t)/2/a * qnorm(pmax(0, pmin(1, epsilon/(2-2*w)))))))
if(v > 0) # positive drift
return(K_small(t, -v, a, w, exp(-2*a*w*v)*epsilon))
S2 = w - 1 + 1/2/v/a * log(epsilon/2 * (1-exp(2*v*a)))
S3 = (0.535 * sqrt(2*t) + v*t + a*w)/2/a
S4 = w/2 - sqrt(t)/2/a * qnorm(pmax(0, pmin(1, epsilon * a / 0.3 / sqrt(2*pi*t) *
exp(v*v*t/2 + v*a*w))))
ceiling(pmax(S2, S3, S4, 0))
}

# Probability for absorption at upper barrier
Pu[4] = function(v, a, w)
{
e = exp(-2 * v * a * (1-w))
if(e == Inf)
return(1)
if(abs(e - 1) < sqrt(.Machine$double.eps)) # drift near zero or w near 1
return(1 - w)
(1 - e) / (exp(2*v*a*w) - e) # standard case
}
```

[4] Die Beschreibung lässt darauf schließen, dass die Funktion die Wahrscheinlichkeit der Absorption an der oberen Barriere zurückgibt, tatsächlich gibt sie (korrekterweise) die Wahrscheinlichkeit der Absorption an der unteren Barriere zurück (vgl. Beschreibung auf Seite 104).

```
# Large time representation of lower subdistribution
Fl_lower = function(t, v, a, w, K)
{
F = numeric(length(t))
for(k in K:1)
F = F - k / (v*v + k*k*pi*pi/a/a) * exp(-v*a*w - 1/2*v*v*t - 1/2*k*k*pi*pi/a/a*t)
* sin(pi*k*w)
Pu(v, a, w) + 2*pi/a/a * F
}

# Small time representation of the upper subdistribution
Fs_lower = function(t, v, a, w, K)
{
if(abs(v) < sqrt(.Machine$double.eps)) # zero drift case
return(Fs0_lower(t, a, w, K))
S1 = S2 = numeric(length(t))
sqt = sqrt(t)
for(k in K:1)
{
S1 = S1 + exp_pnorm(2*v*a*k, -sign(v)*(2*a*k+a*w+v*t)/sqt) -
exp_pnorm(-2*v*a*k - 2*v*a*w, sign(v)*(2*a*k+a*w-v*t)/sqt)
S2 = S2 + exp_pnorm(-2*v*a*k, sign(v)*(2*a*k-a*w-v*t)/sqt) -
exp_pnorm(2*v*a*k - 2*v*a*w, -sign(v)*(2*a*k-a*w+v*t)/sqt)
}
Pu(v, a, w) + sign(v) * ((pnorm(-sign(v) * (a*w+v*t)/sqt) -
exp_pnorm(-2*v*a*w, sign(v) * (a*w-v*t)/sqt)) + S1 + S2)
}
```

```
# Zero drift version
Fs0_lower = function(t, a, w, K)
{
F = numeric(length(t))
for(k in K:0)
F = F - pnorm((-2*k - 2 + w)*a/sqrt(t)) + pnorm((-2*k - w)*a/sqrt(t))
2*F
}

# Lower subdistribution
F_lower = function(t, v, a, w, sigma2, epsilon)
{
a = a / sqrt(sigma2)
v = v / sqrt(sigma2)
K_l = K_large(t, v, a, w, epsilon)
K_s = K_small(t, v, a, w, epsilon)
F = numeric(length(t))
i = (K_l < 10*K_s)
if(any(i)) F[i] = Fl_lower(t[i], v, a, w, max(K_l[i]))
if(any(!i)) F[!i] = Fs_lower(t[!i], v, a, w, max(K_s[!i]))
F
}
# Upper subdistribution
F_upper = function(t, v, a, w, sigma2, epsilon)
{
F_lower(t, -v, a, 1-w, sigma2, epsilon)
}
```

Matlab script

```
% Calculates exp(a) * pnorm(b) using an approximation by Kiani et al.  (2008)

function res = exp_pnorm(a, b)

res = exp(a) .* erfc(-b/sqrt(2))/2;

d = isnan(res) & b < -5.5;

if(any(d))

res(d) = 1 ./ sqrt(2) .* exp(a - b(d) .* b(d) ./ 2) .* (0.5641882 ./ b(d) ./ b(d)

./ b(d) - 1 ./ b(d) / sqrt(pi));

end

return


% Number of terms required for large time representation

function K = K_large(t, v, a, w, epsilon)

sqrtL1 = sqrt(1./t) * a / pi;

sqrtL2 = sqrt(max(1, -2./t*a*a/pi/pi .* (log(epsilon*pi*t/2.* (v*v + pi*pi/a/a))

+ v*a*w + v*v*t/2)));

K = ceil(max(sqrtL1, sqrtL2));

return

% Number of terms required for small time representation

function K = K_small(t, v, a, w, epsilon)

if((abs(v) < sqrt(eps))) % drift near zero

K = ceil(max(0, w/2 + sqrt(t/2) / a * erfcinv(2*max(0, min(1,epsilon / (2-2*w))))));

return

end

if(v > 0) % positive drift

K = K_small(t, -v, a, w, exp(-2*a*w*v)*epsilon);

return

end
```

```
S2 = zeros(1, length(t))+w - 1 + 1/2/v/a * log(epsilon/2 * (1-exp(2*v*a)));
S3 = (0.535 * sqrt(2*t) + v*t + a*w)/2/a;
S4 = w/2 + sqrt(t/2)/a .* erfcinv(2*max(0, min(1, epsilon * a / 0.3 ./ sqrt(2*pi*t)
.* ...
exp(v*v*t/2 + v*a*w))));
K = ceil(max(0, max(vertcat(S2, S3, S4))));
return
% Probability for absorption at upper barrier[5]
function P = Pu(v, a, w)
e = exp(-2*v*a*(1-w));
if(e == Inf)
P = 1;
elseif(abs(e - 1) < sqrt(eps))
P = 1 - w; % drift near zero or w near 1
else
P = (1 - e) / (exp(2*v*a*w) - e); % standard case
end
return
% Large time representation of lower subdistribution
function Fl = Fl_lower(t, v, a, w, K)
Fl = zeros(1, length(t));
for k = K : -1 :  1
Fl = Fl - (k ./ (v*v + pi*pi*k*k/a/a) * exp(-v*a*w - 1/2*v*v*t - 1/2*pi*pi*k*k/a/a*t)
* sin(pi*k*w));
end
Fl = Pu(v, a, w) + 2*pi/a/a * Fl;
return
```

[5] siehe oben.

```
% Small time representation of the upper subdistribution
function Fl = Fs_lower(t, v, a, w, K)
if(abs(v) < sqrt(eps))
Fl = Fs0_lower(t, a, w, K);
return
end
S1 = zeros(1, length(t));
S2 = zeros(1, length(t));
sqt = sqrt(t);
for k = K : -1 :  1
S1 = S1 + exp_pnorm(2*v*a*k, -sign(v)*(2*a*k+a*w+v*t)./sqt) - ...
exp_pnorm(-2*v*a*k - 2*v*a*w, sign(v)*(2*a*k+a*w-v*t)./sqt);
S2 = S2 + exp_pnorm(-2*v*a*k, sign(v)*(2*a*k-a*w-v*t)./sqt) - ...
exp_pnorm(2*v*a*k - 2*v*a*w, -sign(v)*(2*a*k-a*w+v*t)./sqt);
end
Fl = Pu(v, a, w) + sign(v) * ((1/2 * erfc(sign(v) * (a*w+v*t)./sqt/sqrt(2)) - ...
exp_pnorm(-2*v*a*w, sign(v) * (a*w-v*t)./sqt)) + S1 + S2);
return

% Small time representation of the upper subdistribution (zero drift)
function Fl = Fs0_lower(t, a, w, K)
Fl = zeros(1, length(t));
for k = K : -1 :  0
Fl = Fl - erfc((2*k + 2 - w)*a./sqrt(2*t)) + erfc((2*k + w)*a./sqrt(2*t));
end
return
```

```
% Lower subdistribution
function Fl = F_lower(t, v, a, w, sigma2, epsilon)
a = a / sqrt(sigma2);
v = v / sqrt(sigma2);
K_l = K_large(t, v, a, w, epsilon);
K_s = K_small(t, v, a, w, epsilon);
Fl = zeros(1,length(t));
i = (K_l < 10*K_s);
if(any(i)); Fl(i) = Fl_lower(t(i), v, a, w, max(K_l(i))); end
if(any( i)); Fl( i) = Fs_lower(t( i), v, a, w, max(K_s( i))); end
return

% Upper subdistribution
function Fu = F_upper (t, v, a, w, sigma2, epsilon)
Fu = F_lower(t, -v, a, 1-w, sigma2, epsilon);
return
```

References

- Kiani, M., Panaretos, J., Psarakis, S., & Saleem, M. (2008). Approximations to the Normal distribution function and an extended table for the mean range of the Normal variables. *Journal of the Iranian Statistical Society*, *7*, 57–72.
- R Core Team (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

5 Cross-modal cueing effects on audiovisual spatial attention

Steven P. Blurton[a], Mark W. Greenlee[a], and Matthias Gondan[b]

[a]Institute for Pschology, University of Regensburg, Germany

[b]Department of Psychology, University of Copenhagen, Denmark

Abstract[6]

Previous research with visuospatial cues has demonstrated that we can shift our focus of visuospatial attention without overt eye movements. The shift can be stimulus-driven (exogenous cues), or voluntarily (endogenous cues). In both cases, visual processing is facilitated at the location of the attentional focus. Visual cues can also effectively influence auditory target detection. This type of cross-modal cueing, however, seems to depend on the type of the employed visual cue: facilitation effects have been frequently reported for endogenous visual cues while exogenous cues seem to be mostly ineffective. In two experiments, we investigated cueing effects on the processing of audiovisual redundant signals. In Experiment 1 we used endogenous cues to investigate their effect on the detection of auditory, visual, and audiovisual targets presented with onset asynchrony. Consistent cueing effects were found in all target conditions. In Experiment 2 we used exogenous cues and found only small effects of cue validity, not only in unimodal auditory targets, but in *all* audiovisual targets. We adjusted a response time model to the data of both experiments and tested whether the observed cueing effects were modality-dependent. While the results observed with endogenous cues imply that the integration of multisensory signals is modulated by a single system operating in a top-down manner, bottom-up control of attention, as observed in the exogenous cueing task, mainly exerts its influence through modality-specific subsystems.

Keywords: multisensory processes; math modeling; spatial attention; cross-modal cueing

[6]Dieses Manuskript ist zur Veröffentlichung im *Journal of Experimental Psychology: Human Perception and Performance* eingereicht worden und wird zur Zeit begutachtet.

5.1 Introduction

Everyday perception commonly involves several senses (Welch & Warren, 1986). Multisensory research deals with their interplay, how information obtained by the senses is integrated and how attention can be directed to an object of interest, irrespective of whether we see or hear it. Attention can be directed overtly (e.g., by eye, head or body movements) or covertly, that is, without directing the peripheral receptors to the object of interest. Covert orienting of visuo-spatial attention has been investigated in a large number of studies with spatial cues (e.g., Posner, Nissen, & Ogden, 1978; Posner, Snyder, & Davidson, 1980; Jonides, 1981; Theeuwes, 1991; Carrasco, Ling, & Read, 2004; Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010).

Two main classes of cues and their effect on attention are distinguished in the research on covert attention shifts with the spatial cueing setup: endogenous and exogenous cues (e.g., Jonides, 1981; Theeuwes, 1991). The properties of the cues play a key role in the modulation of attention. An endogenous cue, that is, a central, symbolic cue (e.g., an arrowhead) is supposed to influence attention in a top-down manner. The cued location has to be derived indirectly from the direction of the arrow, and attention is then deliberately directed to the cued location. For effective voluntary attentional control the cue must be predictive, that is, the (posterior) probability that a target appears at the cued location must be greater than the (prior) probability that a target appears at that location anyway.

Beside this voluntary control of attention, there is a simpler, automated form of attentional modulation. If a cue suddenly appears in the periphery, it automatically attracts visuospatial attention. This is a form of bottom-up control of attention; effective cues are therefore, in general, non-symbolic and are presented in the periphery at or near the possible location of a subsequent target. Such peripheral cues do not even have to be predictive; those cues can direct attention

without any validity (Yantis & Jonides, 1984). Both types of cues can effectively draw attention to the cued location. If a subsequent target is presented at that location, responses are faster and more accurate than responses to targets at uncued locations. These cueing effects are typically interpreted as attention effects, that is, spatial attention facilitates the perception of stimuli presented at the attended location. While some studies focused on response speed (Posner, 1980; Posner et al., 1980; Jonides, 1981), other studies used perceptual decision tasks to test response accuracy (e.g., Bashinski & Bacharach, 1980; Dosher & Lu, 2000; Carrasco et al., 2004; Liu, Fuller, & Carrasco, 2006; Fairhall & Macaluso, 2009).

The initial studies on the spatial cueing task focused on the visual modality (but see Lansman, Farr, & Hunt, 1984). Subsequent research also investigated the effects of secondary tasks such as language processing (Posner, Inhoff, Friedrich, & Cohen, 1987) on detection times in the spatial cueing task. After links between audition and vision had been established, multisensory attention was further investigated to identify the extent of links between attention of the different senses and whether there is a supramodal attentional system (Farah, Wong, Monheit, & Morrow, 1989; Eimer & Schröger, 1998; Eimer, 2001; Eimer & Driver, 2001; Talsma & Kok, 2002; Talsma et al., 2010). For example, Buchtel and Butter (1988) used visual and auditory exogenous cues and measured the response time to visual and auditory targets in a simple reaction time task. The main finding of their study was that visual—but not auditory—targets were affected by the preceding visual cues. Using a perceptual two-choice task, Spence and Driver (1996) thoroughly investigated the effects of endogenous visual cues on visual and auditory perception. Critically, they used an orthogonal cueing technique (Spence & Driver, 1994); an improved method for cueing in perceptual decision tasks that precludes response priming (Ward, 1994; Spence & Driver, 1997). Spence and Driver (1996) reported cueing effects on both visual and auditory targets, that is, visual cues directed cross-modal attention and thus also speeded up the perceptual decision in pure

auditory targets. Response accuracy was less affected by the attentional manipulation in either modality (Spence & Driver, 1996, Table 1). Spence and Driver (1996) argued that auditory perception is affected by attentional mechanisms, but only through late processes that do contribute to perceptual decisions but not to simple reaction time. Several EEG studies, however, found that very early processes were equally affected by attentional manipulations (Eimer, 2001; Talsma & Woldorff, 2005).

The results of cross-modal cueing in exogenous attention were different: Spence and Driver (1997) conducted a series of experiments in which they demonstrated that exogenous visual cues did not influence the perception of auditory signals (cf. Buchtel & Butter, 1988). The relationship between audition and vision in exogenous spatial attention is asymmetric—that is, exogenous auditory cues influence visual perception (Spence & Driver, 1994) but visual cues do not seem to affect auditory perception (Driver & Spence, 1998). These results are challenged by those of Ward (1994) and of Ward, McDonald, and Lin (2000) who reported exactly the opposite asymmetry, albeit in different experimental setup with rather complex cueing and a non-spatial go/no-go task. The non-spatial response task was subsequently identified as the important difference between those two studies (Spence, McDonald, & Driver, 2004; Koelewijn, Bronkhorst, & Theeuwes, 2010).

Although these studies investigated cross-modal cueing effects, they provide little insight into the interplay between the processing of multisensory information and attention shifts—unless one takes up the position that multisensory integration *is* cross-modal attention (cf. Spence, McDonald, & Driver, 2004, p. 306ff). Like in spatial cueing experiments, it has been demonstrated that exogenous and endogenous multisensory cues have quite different effects on multisensory integration (see Talsma et al., 2010, for a review). Endogenous cues seem to facilitate multisensory integration of subsequent percepts (Fairhall & Macaluso, 2009), especially if the perceptual load is high (Talsma & Woldorff, 2005). Studies with

exogenous cues, on the other hand, support the notion that those cues induce a spread of spatial attention across modalities that seems to be location unspecific (Talsma et al., 2010). For example, a spatially uninformative auditory cue can enhance target detection in a visual search task, whereas temporal, but spatially uninformative visual cues had no effect on performance in the visual search task (Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008). It remains open, though, if this cross-modal spread of attention is symmetric, that is, if visual cues without spatial information can facilitate performance in an auditory task (e.g., auditory discrimination task). The results from Spence and Driver (1997) would suggest that this is not generally the case.

A common experimental paradigm to study multisensory integration is the redundant signals setup. In this setup, participants are provided with information in different modalities, for example, vision and audition. Participants are usually given the task to respond as quickly as possible to any stimulus. The participants have to divide their attention across the modalities, this task has therefore also been called divided attention task (Miller, 1982; Diederich & Colonius, 1987). The typical finding is that if signals from both modalities are present (redundant signals, AV), responses are faster than to targets from any single modality (unimodal targets), that is, faster than responses to auditory (A) or visual (V) targets. A more precise analysis of the integration of redundant signals requires the extension of this basic setup to redundant stimuli presented with onset asynchrony (e.g., Miller, 1986). Then, one can not only investigate the presence or absence of redundancy gains, but rather the mechanism of multisensory integration with fine-grained resolution (e.g., Miller, 1986, Eq. 3). The unimodal conditions A and V can then be considered as end points of a continuum where the second stimulus follows with infinite SOA. Whereas attentional effects on visual perception have been investigated with redundant signals (Mordkoff & Yantis, 1993; Feintuch & Cohen, 2002; Miller, Beutinger, & Ulrich, 2009; Mordkoff & Danek, 2011), the

effects of spatial cues on the integration of multisensory redundant signals are largely unknown. There is only one study that investigated differences in multisensory processing of redundant signals under different spatial attention conditions (Gondan, Blurton, Hughes, & Greenlee, 2011). In that study, the effect of a different attention focus size was compared in two conditions (wide focus vs. narrow focus of sustained attention). As one would expect, a narrow focus facilitated the detection of audiovisual signals and their integration into a single percept. However, the question of cross-modal interactions in spatial attention is left open in that study. It remains also unclear, whether transient attention exerts the same influence on multisensory processing as sustained attention (Eimer, 1996).

In this study, we investigated the effect of visual spatial cues on multisensory perception of auditory, visual and audiovisual targets at different peripheral locations. In Experiment 1, we used central, symbolic, and informative (endogenous) visual cues and in Experiment 2 we employed peripheral and non-informative (exogenous) visual cues. A diffusion superposition model (Schwarz, 1994; Diederich, 1995) was fitted to observed response times to assess the effects of those cues on multisensory integration of redundant signals. Using comparisons of hierarchically nested models we tested (i) for the presence of cueing effects at all and (ii) whether these effects were modality-invariant or modality-specific.

5.2 Experiment 1

5.2.1 Methods

In Experiment 1, participants made speeded responses in an audiovisual detection task with central visual predictive cues. In addition to unimodal auditory and visual target conditions, we employed bimodal target conditions in which bimodal stimuli were presented with onset asynchrony.

Participants Sixteen students (mean age: 23.9 years, four male, and one left-handed) of the University of Regensburg participated in the first experiment. All had normal or corrected-to-normal visual acuity and were naïve with regards to the purpose of the experiment. They were either paid for participation (7 € per hour) or received course credit. Before participation, they gave written informed consent and the study was conducted in accordance with the Declaration of Helsinki.

Apparatus and stimuli The experiment took place in a light and sound attenuated room, that was dimly illuminated from behind and above (Industrial Acoustics GmbH, Niederkrüchten). The participants viewed the screen with their head resting on a chin rest 70 cm in front of the screen that had a size of 57 cm × 72 cm (54.4 deg) and was projected on by a projector (NEC V230X, NEC Corporation, Minato) from outside the cabin. The loudspeakers were mounted adjacent to the screen right at the left and right edges of the screen inside the cabin. The visual target was a Gabor patch (sigma = 0.8 deg, 1.8 cycles/deg, Michelson contrast: $L_{\mathrm{M}} = 0.969$, size: 6 deg) presented either in the left or right periphery (24 deg eccentricity) on a uniform gray background with a white fixation cross (0.8 deg) at the center of the screen. The auditory target was white noise (45 dBA), presented over loudspeakers placed left and right to the screen. The speakers were connected to a low-latency sound card (Soundblaster Audigy 2 ZS, Creative Technology, Singapore) that was installed in a standard IBM-compatible PC running Presentation (Neurobehavioral Systems, Berkeley, CA). The target stimuli were generated with MATLAB (MathWorks, Natik, MA). The spatial cues were arrowheads (1 deg) presented adjacent to the central fixation cross. The arrows were presented for 67 ms and directed either to the left side or to the right side. The cueing interval, that is, the time between cue offset and target onset was 300 ms. Each trial started with a cue followed by a single target (unimodal visual or auditory condition), by an audiovisual target (redundant visual-auditory condition) or no target (catch

trials). To discourage from anticipatory responses we used a rate of one catch trial for six target trials. In target trials, targets were presented at the cued location in 75% of the trials (cue validity). Between two trials we used a variable inter-trial-interval that had a base time of 600 ms, plus an exponentially distributed random duration. To avoid too long intervals, the exponential distribution was truncated at 95%. Participants should press a response button placed at their dominant hand.

Experimental tasks Participants were instructed to respond as fast as possible when a target appeared either at the expected (cued) location or at the unexpected (uncued) location and to withhold responses in trials with only cues but no targets (simple speeded response). They were instructed to fixate the central position of the screen which was marked by a cross during the whole experiment. On each trial they had one second to respond before the trial ended and stimulation resumed with the next trial. The redundant targets were presented with stimulus onset asynchrony (SOA). The auditory target preceded the visual target by 33 ms, 67 ms, or 100 ms, or followed the visual target by the same amount of time. Thus, redundant targets were presented with seven SOAs: A33V, A67V, A100V, V33A, V67A, V100A, and the synchronous target condition, AV. Together with the unimodal conditions A and V we thus had nine SOA conditions.

Procedure Each participant was tested in three separate sessions and each session lasted about one hour. In each session, participants were tested with two blocks of 336 trials each. A short break was given between the blocks. For each participant we obtained data from 48 trials with an invalid cue from 144 trials with a valid cue. These numbers are pooled for left/right target locations but apply to all SOA conditions employed. Thus we have tested each participant in 1,728 cue/target trials and 288 catch trials.

Analysis of response times The high number of catch trials allows for a detailed analysis of empirical response time (RT) distributions. We used the so called "kill-the-twin procedure" (Eriksen, 1988) combined with the Kaplan-Meier estimate to retrieve the RT distribution swept from anticipations (Koch et al., 2013). The reasoning behind this procedure is that the real empirical distributions of target RTs are contaminated by guesses and anticipatory responses. In a first step, we determined those guesses by eliminating responses from the target RT distribution for each response to a catch trial. In the kill-the-twin procedure this is done by right-censoring six responses in closest temporal correspondence to a given catch trial response, thereby accounting for the fact that for each catch trial 4.5 valid and 1.5 invalid trials were presented, respectively. With this procedure we aimed at controlling for anticipations. In an analogous manner, we also censored all RTs greater than 1000 ms (less than 0.5 % of RT in all participants) to correct for attentional lapses ("misses"). No further correction was applied to the RT data. Thus, no response was removed from RT distribution and it can be argued that the assumption of an uninformative censoring mechanism yields a conservative bound for correction of the response time distribution (Gondan & Fimm, 2013). After pooling responses from left and right targets, the Kaplan-Meier estimate $\hat{G}(t)$was used to estimate the response time distribution; the area under $1 - \hat{G}(t)$ was then used to calculate mean RT. Afterwards, we fitted the Diffusion Superposition Model (DSM, Schwarz, 1994) to the mean reaction times of all conditions.

The race model inequality Several models have been proposed to account for the redundant signals effect; the most prominent model classes are the race models (Raab, 1962) and coactivation models (Miller, 1982). In the former, it is assumed that detection of signals can be conceived as a race between active channels in which the winner of the race determines the detection time. The redundant signals effect is then a consequence of *statistical facilitation*: if the latency distributions

of detection times of the two channels overlap, detection of redundant signals is, on average, faster, because in that case slow detection times in one channel are compensated by faster detection in the other channel. It has been shown that, under certain assumptions, the redundancy gain by the race model has an upper bound (Miller, 1982, 1986). Because information is never integrated in the sense that it is pooled into a common channel, the response time distribution function of redundant target, $G_{\mathrm{AV}}(t)$, cannot exceed the sum of the distribution functions of the unimodal targets:

$$G_{\mathrm{AV}}(t) \leq G_{\mathrm{A}}(t) + G_{\mathrm{V}}(t) \text{ , for all } t. \tag{B1}$$

Violations of this upper bound have often been reported and are usually interpreted as evidence for coactivation (e.g., Miller, 1982, 1986; Schröter, Ulrich, & Miller, 2007) but it should be noted that the above inequality only holds under the assumption that channel-specific processing times are invariant across the experimental conditions (Luce, 1986). We tested the race model inequality with the permutation test described by (Gondan, 2010). For the permutation test, we used an aggregated test statistic, which was the weighted sum $\Sigma_\tau w_\tau \Delta_\tau$ of SOA-dependent violations Δ_τ of the race model inequality (Gondan, 2009), for example $\Delta_\tau = \max[0, \hat{G}_{\mathrm{A}(\tau)\mathrm{V}}(t) – \hat{G}_{\mathrm{A}}(t) – \hat{G}_{\mathrm{V}}(t – \tau)]$ for auditory-first stimuli. The weighting function had the form of a shifted umbrella; specifically, we assigned weights w_τ = 2, 3, 4, 3, 2, 1, and 1 to conditions V100A, V67A, V33A, AV, A33V, A67V and A100V, respectively. The weight function was shifted, because from the difference in mean reaction time of single target conditions we expected race model violations to be most pronounced when the visual stimulus preceded the auditory stimulus by moderate SOA (psychological synchrony)—that is, V33A. For each condition, we calculated violation statistics for six RT percentiles (5th, 10th, ..., 30th percentile) and standard T-statistics for each percentile across participants.

To avoid multiple tests, the six T-statistics were aggregated into a $T_{\max}$ statistic using the max function across the six percentiles. The distribution of this $T_{\max}$ statistic under the null hypothesis (i.e., the race model holds) was retrieved by a permutation procedure: Under the null hypothesis, observed deviations from inequality (1) should be random. Therefore, we randomly ($p = 1/2$) assigned signs to the observed violation statistics of each participant and calculated the resulting T statistic. This permutation procedure was repeated 10,001 times to obtain the reference distribution under the null hypothesis. The p values were then calculated as the proportion of permutations that yielded greater T than the observed $T_{\max}$. The race model assumptions were conceived as significantly violated, if the p value was less than 5%—that is, if $T_{\max}$ was greater than $T_{0.95}$.

Diffusion Superposition Model We used the Diffusion Superposition Model (DSM Schwarz, 1994) to assess redundancy gains and attention effects on response times. This computational model describes response times of a redundant signals experiment and can readily be applied to data from two or more experimental manipulations, like spatial cueing. The DSM assumes an information accumulation process that can be described by a time-homogenous diffusion process (Wiener process) with two parameters, drift $\mu > 0$ and variance $\sigma^2 > 0$. In a redundant signals task, each channel is assumed to represent such a process, that is, a Wiener process with parameters μ_{A} and σ^2_{A} for the auditory processing channel and parameters μ_{V} and σ^2_{V} for the visual processing channel. This process describes the detection process as an information accumulation process (sequential sampling). The process continues until a certain response criterion c is reached for the first time. The detection time $\mathbf{D}$ is the first passage time to c, it follows an inverse Gaussian distribution and with expected value $E(\mathbf{D}) = c/\mu$ (cf. Cox & Miller, 1965, p. 222). When redundant (audiovisual) targets are presented, both channels are active and it is assumed that the information of both channels is pooled

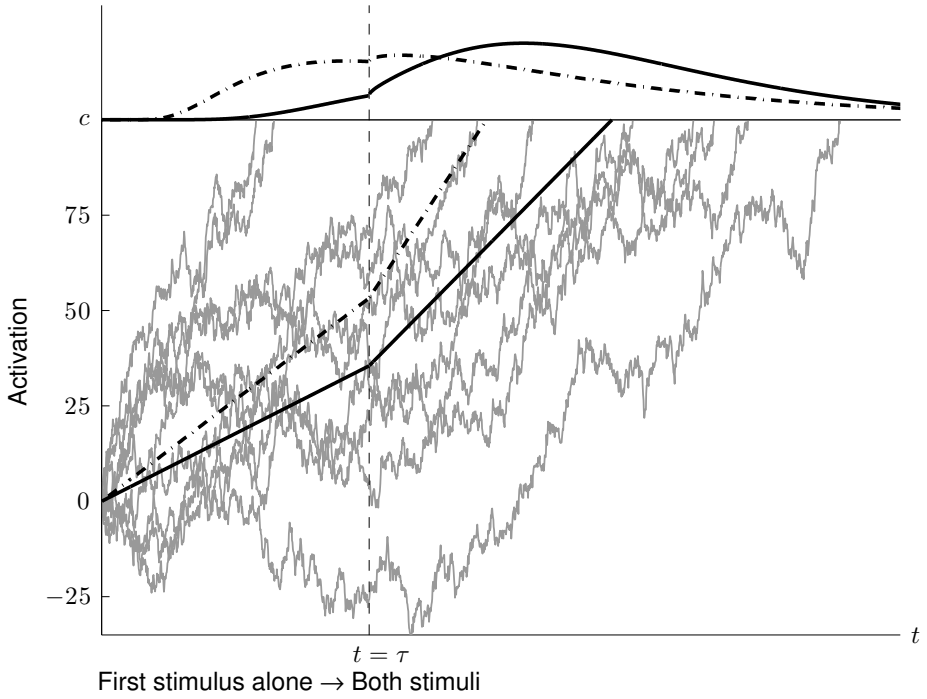


Figure 5.1: The signal detection process in the common channel according to the Diffusion Superposition Model. At stimulus onset, a Wiener process with drift starts at $\mathbf{X}(0) = 0$ and evolves over time, until the criterion c is hit for the first time. Displayed are ten realizations of a Wiener processes with parameters $\mu = 0.53$ and $\sigma = 4.3$ and response criterion $c = 100$. When redundant targets are presented with some SOA (here: $\tau = 67$), the process contains two parts: at the beginning, only one channel contributes to the activity of the common channel. At the onset of the second stimulus component the process has attained a state $\mathbf{X}(t = \tau) < c$, if the criterion has not already been reached. Afterwards, the second channel (here: $\mu = 0.53$ and $\sigma = 4.3$) is also active and its contribution is added into the common channel. The effects of the additive superposition can be seen in the average position of a process (i.e., the solid linear trend) and in the solid upper curve that represents the resulting first-passage time density $f(t|c, \mu, \sigma)$ that is arbitrarily scaled for display purposes. The dashed linear trend and the dashed first-passage time $f(t|c, \mu \times g, \sigma \times g)$ are displayed to demonstrate the influence of the attention scaling factor g ($g = 1.5$).

into a common channel (additive superposition). The DSM is thus an instance of a coactivation model; the redundant signals effect is explained to be due to the faster buildup of evidence in the case of both channels being active. The contributions of both processes are additively superimposed, so that the drift of the common process is $\mu_{\mathrm{AV}} = \mu_{\mathrm{A}} + \mu_{\mathrm{V}}$. Because the drifts are added, the response criterion c is reached faster when two signals are present than when only one signal was presented (Figure 1, solid lines). This is evident in the expressions for the expected time of $\mathbf{D}$, which are given by $E(\mathbf{D}_{\mathrm{A}}) = c/\mu_{\mathrm{A}}$ and $E(\mathbf{D}_{\mathrm{V}}) = c/\mu_{\mathrm{V}}$ for the auditory and visual channel, respectively, and the expected time in case of redundant targets, $E(\mathbf{D}_{\mathrm{AV}}) = c/(\mu_{\mathrm{A}} + \mu_{\mathrm{V}})$. From these expressions it is easily seen that the expected value for $\mathbf{D}$ decreases, if either the criterion c is decreased or, equivalently, the modality-specific drifts are multiplied by a constant $g > 1$. This decrease affects the mean detection times of all conditions, including those of redundant signals presented with SOA (see Schwarz, 1994, Eq. 10 for details).

Because the diffusion process is supposed to describe only the detection time and response time is usually assumed to be compound of an information accumulation part and residual process ("motor processes", Luce, 1986), the mean latency of all those processes (μ_{M}) is added to the above expression to derive predictions for the observed mean response times. The response criterion c is fixed at 100 because it is only a scaling factor. The DSM has therefore 5 free parameters (μ_{A}, σ^2_{A}, μ_{V}, σ^2_{V}, μ_{M}) that need to be estimated from data in order to fit the model to a set of observed mean response times.

The model fit was assessed by a X^2-statistic that was the sum of normalized differences between observed and predicted mean response times. This statistic is asymptotically χ^2–distributed with *df* determined by the number of predicted experimental conditions minus the number of free parameters. We fitted three models to the data that differed only with respect to the assumptions of spatial cueing effects. The first and most restrictive model had five parame-

ters: four diffusion parameters (μ_A, σ_A^2, μ_V, σ_V^2) and the mean latency of residual processes μ_M. This model served as a null-model, because it contained no free parameter for cueing effects and, thus, did not predict any effects of attention shifts. The second and more plausible model had one additional free parameter, a separate attention factor g_{valid} for the processing of validly cued targets (i.e., $\mu_{A,valid} = g_{valid} \times \mu_A$, same for σ_A, μ_V, σ_V). With this additional parameter, cueing effects could be modelled by an increase in spatial attention (g_{valid}) for validly cued trials. This model was still rather restrictive, because the experimental manipulation of cue validity of all nine target conditions was allowed to affect only this parameter. If g_{valid} is unity, the predictions for targets at expected locations are identical to those presented at unexpected locations (i.e., the null-model). If g_{valid} is greater than unity than it follows for the RT model predictions that $c/\mu = E(\mathbf{D}|\text{invalid cue}) > E(\mathbf{D}|\text{valid cue}) = c/(\mu \times g_{valid})$ for both auditory and visual targets (Figure 1, dashed lines). Thus, this model captures possible attention effects in a single parameter, which is modality invariant as it equally affects all conditions of validly cued trials. To test whether this assumption was justified by the data, we fitted a third model with an additional free parameter to allow for modality-specific cueing effects ($g_{A,valid}$ and $g_{V,valid}$). If, for example, auditory targets are presented at expected locations, $g_{A,valid}$ should be greater than unity and the model predicts an decreased processing time in the auditory channel due to more efficient information retrieval. This is again best seen in the predictions for the detection of unimodal targets, for example, auditory targets, which is given by $E(\mathbf{D}_A|\text{valid cue}) = c/(\mu_A \times g_{A,valid})$. The same holds for the processing of visual targets at expected and unexpected locations. In that case, the respective parameters μ_V and $g_{V,valid}$ for the visual channel must be used instead of μ_A and $g_{A,valid}$.

In both experiments, we fitted the model to response times observed in 18 experimental conditions. The degrees of freedom for each participant's model fit

was $df = 18 - 5 = 13$ for the null-model (predicting no cueing effects), $df = 12$ for the model with modality-invariant cueing effects and $df = 11$ for the model with modality-specific cueing effects. Thus, the DSM for modality-invariant cueing effects needs one additional parameter (g_{valid}) compared to the null-model. The necessity for this parameter was tested by a nested model comparison. The observed difference ΔX^2 of the X^2 goodness-of-fit statistics for the null-model and the model with modality-invariant cueing effects is asymptotically χ^2–distributed with one degree of freedom. Thus, the attention factor g_{valid} was deemed to be significantly different from unity, if the observed difference ΔX^2 was greater than the 95 %-percentile of the $\chi^2(1)$ distribution. This test represents a test for an overall cueing effect. The same test was used to test for possible modality-specific cueing effects. As noted above, the assumption of modality-specific effects of cueing adds another free parameter to the model. The effects of spatial cueing were deemed to be significantly different across auditory ($g_{\text{A,valid}}$) and visual ($g_{\text{V,valid}}$) modality, if the observed difference ΔX^2 was greater than the 95%-percentile of the $\chi^2(1)$ distribution.

For clarity, we mainly focus on the aggregate model fits of all participants. The aggregate model fit of the null-model with five free parameters contained $N \times 18 = 234$ mean response times of $N = 13$ participants from which $N \times 5 = 65$ free parameters were estimated, resulting in 234 – 65 = 169 degrees of freedom (df). The aggregate model fit of the modality-invariant cueing effect model had 234 – 78 = 156 degrees of freedom and the fit of the modality-specific cueing effects model had 234 – 91 = 143 degrees of freedom. Likewise, the influence of the additional parameters were judged to be significant on group-level, if the observed difference ΔX^2 was greater than the 95 %-percentile of the $\chi^2(13)$ distribution for each parameter added.

5.2.2 Results

Analysis of response times and anticipations Despite the high number of catch trials, responses to catch trials were quite frequent. Most participants responded on 5 % up to 20 % of the catch trials. Three participants (i.e., Participants 1, 2 and 7) were removed from further analyses because they responded in more than 25 % of catch trials, which would have required us to censor a large amount of the RT distributions of correct responses. As expected, responses in all target conditions were faster, on average, if the target was preceded by a valid spatial cue than by an invalid cue. This cueing effect was found in the unimodal auditory and visual targets as well as all audiovisual targets (Table 5.1). Validly cued auditory targets ($M = 227 \pm SD = 20.0$ ms) were not only faster than invalidly cued auditory targets ($M = 246 \pm 28.6$ ms), but also faster than validly cued visual targets ($M = 269 \pm 28.9$ ms; modality effect). Invalidly cued auditory ($M = 246 \pm 28.6$ ms) and visual targets ($M = 324 \pm 30.1$ ms) showed the same modality effect. Regarding the redundant targets, we found the typical wing-shaped patterns of mean RT in both the valid cue condition and the invalid cue condition. Irrespective of cue validity condition, responses were fastest in the (synchronous) redundant target condition 5.1. With increasing delay of either target component, mean response times approached the respective mean response time of the unimodal targets.

Race model inequality The race model inequality was significantly violated on group level in the invalid cue condition ($T_{\max} = 2.60$; $T_{0.95} = 2.38$; $p = .032$) but not in the valid cue condition ($T_{\max} = 0.543$; $T_{0.95} = 2.41$; $p = .523$). Redundancy gains were therefore greater than predicted by the race model in the invalid cue conditions. In the valid cue condition, observed redundancy gains were smaller and compatible with race model assumptions.

Table 5.1: *Experiment 1—Mean RT (in ms) to auditory (A), visual (V), and audiovisual (AV) targets presented with onset asynchrony after valid cues and invalid cues (N=13 Participants).*

	Target condition								
Cue condition	A	A100V	A67V	A33V	AV	V33A	V67A	V100A	V
valid	227	219	216	211	206	229	244	255	269
invalid	246	241	233	235	225	248	270	290	324

Note—Participant specific mean RTs were retrieved from RT distributions corrected for anticipations (see Methods). For valid cues, these distributions included 144 responses per participant; for invalid cues 48 responses (75 % cue validity). This data corresponds to the circles (valid cues) and squares (invalid cues) in Figure 2.

Response time model All response conditions were used to fit the diffusion superposition model to the observed mean response times. Different model assumptions were made to account for different cueing effects. On group level, the null-model, assuming no effects of spatial cueing, was clearly rejected ($X^2 = 1481$; $df = 169$; $p < .001$). By contrast, both the model with modality-independent cueing effects ($X^2 = 169.2$; $df = 156$; $p = .222$; Figure 2) and the model with modality-dependent cueing effects ($X^2 = 157.0$; $df = 143$; $p = .200$) both provided good fits to the observed data. The comparison of the null-model with the modality-invariant cueing effect model yielded a significant difference ($\Delta X^2 = 1311$; $df = 13$; $p =< .001$) that is, we obtained significant cueing effects. However, the additional parameter for modality specific attention effects increased the fit only marginally ($\Delta X^2 = 12.26$; $df = 13$; $p = .506$). In other words, the attention scaling factors $g_{\text{A,valid}}$ and $g_{\text{V,valid}}$ were not significantly different from each other. With the attention scaling factors fixed to be equal in both modalities, the model very well predicted the observed patterns of response times (Figure 5.2). Consistent with faster responses to auditory targets, drifts and variance of the auditory processing channel were somewhat greater than drift and variance of the visual

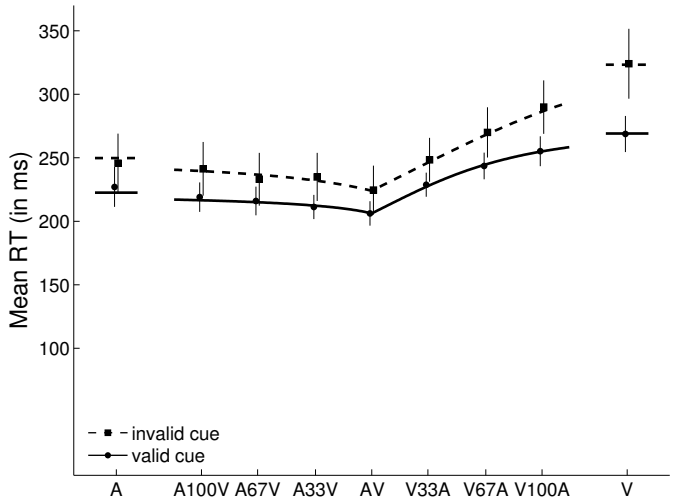


Figure 5.2: Experiment 1—Mean response times observed in valid (dots) and invalid (squares) cue trials averaged across all participants with error bars denoting estimates of the 95% confidence intervals. The lines represent model predictions of the DSM with a modality-invariant cueing effect for the valid (solid line) and invalid (dashed line) cueing condition. Model predictions are also averaged across all participants. On group level and in most participants, the model predictions are in very close agreement with the observed data.

processing channel (Table 5.2), which largely resembled results that were reported earlier in simple-response divided attention tasks (cf. Schwarz, 1994).

The model fits for single participants provided similar results. With the attention facilitation factor fixed to be unity in both cue validity conditions (null-model), the model was rejected in all but one participant. When this constraint was dropped, the model fit was considerably better in all participants. Thus, the cueing parameter g_{valid} for invalidly cued targets was significantly smaller than for validly cued targets (i.e., unity) in all participants. Finally, adding a parameter for modality-specific attention effects did not significantly improve the model fit in any participant. The DSM predictions with a modality-invariant cueing effect were in very good agreement with the observed data. This model produced the same wing-shaped pattern as observed in the data and, more importantly, predicts a decrease of response times for targets presented at expected locations (valid cue).

Table 5.2: *Experiment 1—Mean (± S.D.) parameter estimates and model fit of the Diffusion Superposition Model with modality invariant cueing effect.*

	μ_A	1.60 (± 0.81)
	σ_A	26.5 (± 25.8)
	μ_V	0.71 (± 0.17)
	σ_V	3.6 (± 2.0)
	μ_M	175.5 (± 15.7)
	g_{valid}	1.71 (± 0.61)
Overall fit	$X^2(156)$	169.2
	p value	.222
	$\Delta X^2(13)$	12.26
	p value	.506

Note—μ_A, σ_A^2, μ_V, σ_V^2: drift and variance for auditory and visual diffusion processes, μ_M: mean latency of residual processes, g_{valid}: attention factor for validly cued targets. Overall fit X^2: goodness-of-fit statistic (higher results indicate worse fit), ΔX^2: difference in GOF compared to the model with modality-specific cueing effects ($g_{A,valid} \neq g_{V,valid}$).

5.2.3 Discussion

In Experiment 1, we extended the cross-modal spatial attention setup (Spence & Driver, 1996) to a redundant signals setup with visual spatial cues. This provided a more detailed picture of the effects of those types of cues on multisensory perception. The motivation for this extension was two-fold: firstly, on empirical side, we obtained information about multimodal targets and, on this note, unimodal targets can be viewed as extremes of the continuum of bimodal targets (namely, with SOA = ∞). The redundant signals setup is then a natural extension of previous experimental designs investigating cross-modal attention. In this way we could also test race model predictions and found significant violations of the race

model inequality, at least for invalidly cued targets. The redundancy gain observed in the valid cue condition was smaller and not significant. Besides general issues of power in testing the race model inequality we do not have a straightforward explanation of this finding. However, it seems implausible that the cue validity changed the mode of processing between parallel coactive (invalid cues) and parallel first-terminating (valid cues). Rather, the facilitation effects observed with valid cues could have led to such increased processing speed that the second stimulus component itself could not have substantially sped up the response. By analogy, if targets are presented at different eccentricities, the redundancy gain is smaller for targets presented at more central locations (i.e., if the target falls into the small receptive fields of the fovea) than for targets presented at more peripheral locations (see, e.g., Schwarz, 2006). Therefore, the more efficiently a signal is processed the smaller is the expected redundancy gain.

Sophisticated response time models are available that can explain the redundant signals effect (e.g. Schwarz, 1994; Diederich, 1995). These models provide further insights that cannot be obtained by standard analyses of mean RT (i.e., ANOVA). By application of this model we demonstrated that endogenous spatial cues not only affected response times to visual and auditory targets but rather to the whole SOA-dependent mean RT curve of multisensory integration (Table 5.1 and Figure 5.2). Regarding the unimodal targets, we replicated findings that were reported earlier. On the one hand, this was the effect of visual cues on visual targets, as described by (Posner, 1980). On the other hand, we found an effect of cue validity on auditory targets (Spence & Driver, 1996). These results were in line with previous studies (Farah et al., 1989), but by fitting a diffusion model to data, we revealed that the observed cueing effects could be modeled by a single parameter to describe the effects of attention in both modalities. The validity of this assumption was indicated by the good to excellent agreement of model predictions and data. Under the assumption of modality-invariant cueing, the cue was

allowed to affect an attentional scaling factor which in turn affected processing in both modalities to an equal extent. The model with modality-specific attention effects did not provide a better account for the observed data, so, by parsimony, the results from Experiment 1 support the assumption of a modality invariant attention effect[1].

The attention factor describes the increased efficiency for targets presented at expected locations, that is, when they fall into the spatial attention focus in the valid cue conditions. This leads to faster responses, as less time is needed, on average, to reach the response criterion. Alternatively, the higher efficiency could reflect a lower response criterion because, a lower response criterion is mathematically equivalent to increased drifts and variance of both auditory and visual stimuli, hence the modality-invariance. The interpretation of information sampling efficiency fits well with the notion of attention as a signal-to-noise modulator: If the cue pointed to the correct location, participants had enough time to shift their attention to that location so that upon stimulus onset information about the stimulus was extracted more efficiently. Vice versa, if the stimulus appeared at the unexpected location, information about that stimulus would have been obtained less efficiently, as indicated by the estimate of the attention factor. Different efficiency could also be the reason for better performance and increased sensitivity, for example, in detection tasks (Bashinski & Bacharach, 1980; Dosher & Lu, 2000; Carrasco et al., 2004). The interpretation that mainly perceptual processes were facilitated was further supported by the results that indicate that the model could well explain the observed data with constant residual processes latency. The inspection of parameter estimates leads to interesting implications: if endogenous attention is shifted by cues in a way that perceptual processes are modulated independent of modality, this is indicative of a single, a supramodal spatial attention

[1]Of course, non-significant results should not be taken as evidence for the absence of an effect, it might also be the case that there were modality-specific effects that were just too weak to reach statistical significance.

system (Farah et al., 1989). Even more, if objects appear within this focus, all its features benefit from the increased efficiency, independent of their modality. This is exactly what we observed in Experiment 1. However, pre-attentive (early) integration of the auditory and visual targets could also explain the observed modality-invariant attention shift. According to this view, multisensory objects are integrated into a single percept at an early stage so that the whole percept benefits from attentional resources. This distinction between pre-attentive integration and cross-modal attention (Spence et al., 2004) is not possible by means of the applied response time model. Either way, the results demonstrate that the integration of audiovisual stimuli seems to follow the principle of additive superposition and that, if the stimuli fall into the attention focus, all features benefit from attention in the same manner; not only qualitatively, but also quantitatively.

5.3 Experiment 2

In Experiment 2, we studied the effect of exogenous cues on multisensory integration, using the same experimental setup. Unlike Experiment 1 we used peripheral, non-predictive cues with a shorter cueing interval.

5.3.1 Methods

Nine new participants (mean age: 29.6 years, six female, and two left-handed) were tested in Experiment 2. Before participation, they gave written informed consent and all were naïve regarding the purpose of the experiment.

Apparatus and Stimuli The experiment was performed with the same apparatus as Experiment 1. The target stimuli were also identical. In Experiment 2, we used red square frame (6 deg) as cues that were presented at the possible target locations. After cue presentation, the target followed after 100 ms. The cueing interval was shorter than in Experiment 1 to avoid so-called inhibition of return

effects (Posner & Cohen, 1984). Contrary to Experiment 1, these cues were non-informative, that is, in about 43 % (3/7) of the trials the subsequent target was presented at the cued location and with the same frequency at the opposite location. In about 14 % (1/7) of cued trials there was no subsequent target (catch trial). The frequency of catch trials was the same as in Experiment 1.

Experimental task Participants were instructed to respond as fast as possible when a target appeared at either location and to avoid anticipations, that is, responding to the cue alone. Again, participants were instructed to maintain their fixation which was indicated by a cross at the center of the screen.

Procedure As in Experiment 1, each participant was tested in three separate sessions and each session lasted about one hour. The total number of trials for each participant is 1,728 target trials and 288 catch trials. The number of trials is thus the same as in Experiment 1; however, due to the change in cue validity we obtained 96 replications for each condition, regardless of whether the cue was valid or invalid. In one half of the trials a cue in the left periphery was displayed, in the other half of trials a cue appeared in the right periphery. Target location was also split evenly between left and right peripheral positions.

Statistical analyses We employed the same guessing correction as in Experiment 1, that is, we determined contaminants in the RT distribution by the kill-the-twin procedure and obtained estimates of RT mean and variance with the Kaplan-Meier estimate. The only difference is that due to the changed cue validity we censored three responses for every response to a catch trial for both valid and invalid cue conditions. Attentional lapses (RT $>$ 1000) were also censored, this affected less than 1 % of all RT in all but one participant (Participant 4: 4.2 %). The analysis was analogous to Experiment 1: We tested fits of the diffusion superposition model with different underlying assumptions to the data. This was

(i) the null-model, (ii) the model with modality invariant cueing effects, and (iii) the model with modality-specific cueing effects. Thus, we fitted the same models with the same parameters as before by using a X^2 goodness-of-fit statistic. This statistic is again asymptotically χ^2–distributed with the degrees of freedom given in Experiment 1. For the aggregate-model statistics ($N = 9$ participants), the degrees of freedom were $df = 117$ for the null-model (i), $df = 108$ for the model with modality-invariant attention effects (ii) and $df = 99$ for the model with modality-specific attention effects (iii). The permutation test of the race model inequality was analogous to Experiment 1.

5.3.2 Results

Response times and anticipations In line with previous studies we found a cueing effect of visual exogenous cues on the mean response times of visual targets (Table 5.3): mean response time (± SD) in the visual target condition decreased from 328 (± 36.7) ms to 291 (± 33.4) ms, if targets were presented at the cued location rather than the opposite location. However, unlike Experiment 1, there was no cueing effect for the auditory targets: the mean RT from validly cued auditory targets (M = 305 ± 34.9 ms) was slightly higher than from invalidly cues auditory targets (M = 300 ± 33.5 ms). In general, the mean RT curve of SOA-dependent targets followed again a wing-shaped pattern in both validity conditions. Responses to synchronous redundant targets were the fastest, irrespective of cue condition (Table 5.3). With increasing SOA, mean RT increased and approached the unimodal target mean RT at both ends. Unlike Experiment 1, the cueing effects were not smallest in the redundant synchronous condition with cueing effects increasing with SOA. Here, the cueing effect varied systematically with the onset of the visual target: The cueing effect was most pronounced in the visual unimodal condition and became smaller, the more the visual target component was delayed

(with respect to the auditory target component) and completely disappeared in the auditory unimodal condition.

Table 5.3: *Experiment 2–Mean RT (in ms) to auditory (A), visual (V), and audiovisual (AV) targets presented with onset asynchrony after valid cues and invalid cues (N = 9 Participants).*

	Target condition								
Cue condition	A	A100V	A67V	A33V	AV	V33A	V67A	V100A	V
valid	305	265	256	246	233	246	256	267	291
invalid	300	271	268	251	241	257	271	288	328

Note–Participant specific mean RTs were retrieved from RT distributions corrected for anticipations (see Methods). For valid cues, these distributions included 96 responses per participant; for invalid cues also 96 responses (50 % cue validity). This data corresponds to the circles (valid cues) and squares (invalid cues) in Figures 5.3a and 5.3b.

Race model inequality The observed redundancy gains were again significantly greater than a race model would predict in the invalid cue conditions ($T_{\max} = 2.71$; $T_{0.95} = 2.19$; $p = .027$), but in Experiment 2 also in the valid cue conditions ($T_{\max} = 2.75$; $T_{0.95} = 2.67$; $p = .028$). Thus, unlike Experiment 1, significant violations of race model predictions were obtained for both valid and invalid cue conditions.

Diffusion Superposition Model The aggregate model fit on group level revealed that the modality-specific attention model fitted best to data ($X^2 = 92.36$; $df = 99$; $p = .668$, Figure 5.3a). But in clear contrast to Experiment 1, the restriction that the attention effect is modality-independent lead to considerably increased X^2, that is, worse model fits ($X^2 = 143.3$; $df = 108$; $p = .013$, Figure 5.3b). Both models correctly described the asymmetric wing-shaped pattern of mean RT in both the valid and the invalid cue condition, but it became evident that the modality-invariant model systematically underestimated cueing effects in visual

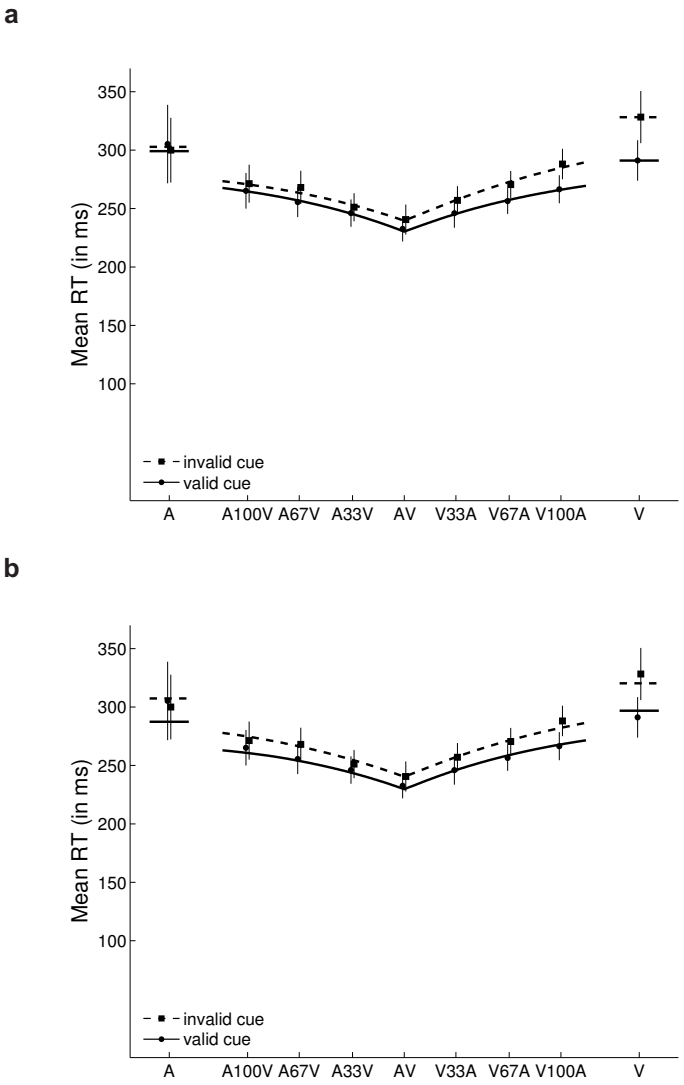


Figure 5.3: Experiment 2—**a**: Mean response times observed in valid and invalid cue trials averaged across all participants together with the DSM predictions of the model with modality-specific cueing effects. Coding of valid and invalid cue conditions is the same as in Figure 5.2. Error bars denote estimates of the 95% confidence intervals. This was the only model which was able to explain the observed patterns of results in Experiment 2. Note that not only there was no cueing effect in the auditory unimodal condition (A), but that the cueing effect systematically varies with the SOA: the more the visual target is delayed (negative SOA) the less effect of cue validity was observed. **b**: Same data as in **a**, but with averaged model predictions from the more restrictive modality-invariant cueing model that very well explained observed data in Experiment 1. The model fit was poor on quantitative grounds, it is easily seen that cueing effects are systematically underestimated in the visual unimodal (V) and overestimated in the auditory unimodal (A) condition. The model was also violated on group level ($p < .001$) and in most participants.

targets and overestimated cueing effects in auditory targets (Figure 5.3b). By contrast, the modality-specific attention model well captured the difference in the cueing effect between auditory and visual unimodal target trials (Figure 5.3a). The model comparison thus yielded a significant difference ($\Delta X^2 = 50.97$; $df = 10$; $p < .001$), that is, the attention factors were significantly different across modalities (Table 5.4): the effect of valid cues on auditory processing ($g_{\mathrm{A,valid}} = 1.03$) was much lower than on visual processing ($g_{\mathrm{V,valid}} = 1.31$). Post-hoc we tested the auditory attention parameter $g_{\mathrm{A,valid}}$ against unity and obtained non-significant results ($\Delta X^2 = 4.03$; $df = 9$; $p = .909$). This implies that there was a negligible cueing effect on processing in the auditory channel. Unlike Experiment 1 the estimated drift for processing in the auditory channel was similar to that for visual processing (Table 5.4).

Therefore, we report the results of the model with modality-specific cueing effects for participant-specific model assessment. The fit of this model was good in all participants. In line with the averaged modality-specific attention parameters, the factor to describe the processing facilitation in the auditory modality was close to unity in about half of participants, the maximum estimate for $g_{\mathrm{A,valid}}$ was 1.13. The difference between the modality-specific cueing effects model fit and the modality-specific cueing effects model fit was significant in about one half of the participants. More importantly, the modality-invariant model produced the same bias in predicted cueing effects in the unimodal conditions as reported above: cueing effects for auditory targets were over-estimated and cueing effects for visual targets were underestimated by the model for modality-invariant cueing effects. The effect of cueing on the latency of non-perceptual processes was again negligible. The ability of the modality-invariant model to explain the observed data was rather poor and the model was rejected based on both qualitative and quantitative assessment. Therefore, the observed data in Experiment 2 was best described by the model, which assumes modality-specific cueing effects.

Table 5.4: *Experiment 2—Mean (± S.D.) parameter estimates and model fit of the Diffusion Superposition Model with modality-specific cueing effect for mean reaction times in valid and invalid cue conditions.*

		μ_A	0.76 (± 0.16)
		σ_A	12.7 (± 8.8)
		μ_V	0.64 (± 0.12)
		σ_V	7.0 (± 3.1)
		μ_M	166.5 (± 21.0)
	g_{valid}	auditory	1.03 (± 0.05)
		visual	1.31 (± 0.15)
Overall fit		$X^2(99)$	92.4
		p value	.668
		$\Delta X^2(9)$	50.97
		p value	$< .001$

Note—μ_A, σ_A^2, μ_V, σ_V^2: drift and variance for auditory and visual diffusion processes, μ_M: mean latency of residual processes, g_{valid}: scale factor (facilitation) for validly cued targets. Overall fit X^2: goodness-of-fit statistic (higher results indicate worse fit), ΔX^2: difference in GOF compared to the base model with modality-invariant cueing effects ($g_{A,valid} = g_{V,valid}$).

5.3.3 Discussion

In Experiment 2 we investigated the effects of exogenous cues on the performance in a multisensory redundant signals task. The effects were similar to endogenous cueing effects, but also differed in several important aspects. As expected, the cues produced significant cueing effects in the visual unimodal targets (Jonides, 1981; Müller & Rabbitt, 1989) and non-significant cueing effects for auditory targets (Buchtel & Butter, 1988; Spence & Driver, 1997). Here, we also extended the

cross-modal cueing setup to a redundant signals setup and used a computational model to assess redundancy gains and cue validity effects in a single model. The race model was rejected because redundancy gains in both the valid and the invalid cue condition was greater than predicted by parallel-first terminating processing. Therefore, we applied a coactivation model to the data and tested for different assumptions regarding the mechanism of attentional modulation by exogenous cues. In Experiment 1, the modality-invariant cueing model best described both redundancy gains and cueing effects. By contrast, in Experiment 2, the modality-specific cueing model accounted much better for the observed patterns of results. Smallest cue validity effects were found in the auditory single target condition.

Considering the parameter estimates and data, the attention effects on the processing in the auditory channel were rather small. This becomes evident in the comparison of validly and invalidly cued auditory targets that showed no cueing effect. Even more, if one compares the model parameters estimated from the data of Experiment 2 with those obtained in Experiment 1, it is striking to see how inefficient auditory processing became with the same targets. The only difference is that exogenous instead of endogenous cues were used in Experiment 2. The perception of both the auditory stimuli (unimodal targets) and the auditory stimulus components (bimodal targets) did hardly benefit from being attended to. To some extent, this is quite the opposite of what we observed in Experiment 1. Whereas endogenous cues lead to facilitation effects of the whole multisensory percept, exogenous cues facilitate only visual targets or the visual target component. The perception of an accompanying auditory target component is not facilitated (or attenuated). Therefore, in stimulus-driven attention visual cues seem to facilitate mainly, or, even exclusively, the visual perception of audiovisual signals and their integration into a common percept. Alternatively, the spatial properties of visual exogenous cues do not seem to play a significant role in auditory perception. So while auditory cues can facilitate visual processing (Van der Burg et al., 2008) even

if they are location-unspecific through a cross-modal spreading of spatial attention (Talsma et al., 2010), this seemingly does not apply to visual cues in an auditory target detection task. This interpretation would explain the reported null-effects of visual cues on auditory targets (Buchtel & Butter, 1988; Spence & Driver, 1997). However, in addition to these studies that reported null-effects of visual exogenous cues on auditory target detection (Buchtel & Butter, 1988) or auditory target elevation discrimination (Spence & Driver, 1997), here we demonstrated that this is also the case for auditory target components in redundant bimodal targets. This point is critical because the interpretation of similar results obtained so far often hinges on the non-significant test result of a single experimental condition. The model parameters on which we base our interpretation incorporate data of not less than 8 conditions (all invalid conditions but the visual unimodal target condition that is independent of the auditory attention factor).

It is also worth noting that in all models the mean latency of residual processes was kept constant, irrespective of modality, redundancy, and cue validity. Thus, we conclude that exogenous cues influence (visual) perception, at least much more than non-perceptual processes, in the integration of audiovisual signals.

5.4 General Discussion

While many studies have reported cross-modal cueing effects of visual cues on auditory perception in endogenous cueing experiments (Farah et al., 1989; Spence & Driver, 1996; Eimer & Schröger, 1998), the results of studies involving exogenous cross-modal cueing are rather inconclusive (Buchtel & Butter, 1988; Spence & Driver, 1997; Ward et al., 2000). In two experiments we employed both types of cues in an audiovisual redundant signals task. This extension aimed at investigating the effects of spatial attention on the integration of audiovisual signals (e.g. Bertelson et al., 2000; Vroomen et al., 2001). With redundant audiovisual signals one cannot only investigate cross-modal cueing effects but also how the processing

of auditory and visual stimulus component of a redundant signal takes place and how this processing is modulated by varying attention conditions. We addressed this question with the application of a diffusion model to test explicitly for different cross-modal cueing effects. The model has repeatedly shown to successfully predict redundancy gains in bimodal divided attention tasks (Schwarz, 1994; Diederich, 1995; Gondan et al., 2010, 2011; Blurton et al., 2014). The inspection of the model parameters on the one hand supports previous interpretations and extends results in endogenous cueing tasks, but, on the other hand, allows for interesting new interpretations of exogenous cross-modal cueing.

As expected, we observed robust cueing effects under voluntary control of attention. The interesting point in this case is that this applies not only for auditory targets. Rather, all audiovisual redundant targets exhibited effects of valid and invalid cueing. Across all experimental conditions, cueing effects were according to the assumptions of a modality-invariant influence of (supramodal) attention: They were not only cross-modal in nature, but even equal in strength across modalities. This notion of attentional cueing effects is compatible with modality-specific differences in mean RTs for unimodal auditory and visual targets. This modality effect could be interpreted as modality-specific attention effects. The application of the diffusion model and the interpretation of its parameter estimates suggest otherwise. In this model, perceptual facilitation is operationalized by attention parameters that were chosen to capture the more efficient processing in both channels for target stimuli that fall into the focus of attention. According to this model assumption, the difference in cueing effects between audition and vision can be explained quite easily due to the fact that auditory signals are processed faster than visual signals. Corresponding to the model prediction, this simply attenuates the effect of increased efficiency in the auditory channel compared to the visual channel. Top-down control of spatial attention seems therefore largely compatible with a common, supramodal attention system (Talsma et al., 2010;

Koelewijn et al., 2010). The current finding of a modality-invariant cueing effect is in line with previous studies on voluntarily controlled attention and its effect on multisensory processing (Farah et al., 1989; Talsma et al., 2010). The results thereby add to those of sustained spatial attention effects on multisensory integration of audiovisual signals (Gondan et al., 2011). In that study, redundancy gains were investigated under two conditions, one with a single (central) possible target location and one with multiple (central and two peripheral) possible target locations. Both conditions were presented block-wise (sustained attention) and participants were a priori informed about the condition. The comparison of centrally presented targets in both conditions demonstrated that targets are more efficiently (i.e., faster) processed in the condition with a single target location (i.e., narrow attention focus) as compared to targets that appeared at the same location but could also have appeared at other locations (i.e., wide attention focus). Most notably, this facilitation effect was also modality-invariant and residual processes latency was also not affected, only perceptual processes (Gondan et al., 2011).

Based on the data of this study it cannot be decided whether voluntarily controlled attention facilitates the perception of already integrated multisensory objects (early integration), or, if attention facilitates the integration of unimodal percepts into a single object (late integration). Neurophysiological studies support both notions: If one interprets the effects of attention on early processes as a consequence of pre-attentive integration, a number of studies provide evidence for early integration (Eimer, 2001). But it has also been demonstrated, that attention modulates event-related scalp potentials of redundant signals at various stages, including late processes (Talsma & Woldorff, 2005). Given that we employed a simple response task and observed rather short response latencies, one could interpret the results of the first experiment as evidence for early integration.

It has been argued (Spence & Driver, 1997) that cross-modal attention of visual cues on auditory targets depends on late processes that are not available in

simple response tasks. This argument addresses the representation of auditory and visual information in the brain. Vision is spatiotopic from the beginning; different neuronal populations encode information of different spatial locations. Audition, on the other hand, is not spatiotopic but tonotopically organized. Spatial location has to be extracted from different cues, for example, the interaural-time or level differences, the latter related to sound attenuation caused by the skull, whereas the former by different distance between the sound source and the left or right ears. This extraction takes place at a rather late stage of auditory processing and that this information might therefore not be available in a simple response task. This difference seems not to be crucial for endogenous cross-modal cueing effects, though, as endogenous visual-on-auditory cueing effect have been reported in a perceptual decision tasks (Spence & Driver, 1996), a detection task (Farah et al., 1989) and now in a redundant signals task. For stimulus-driven attention, this difference in perceptual latency could be a critical point and is discussed below. The results of the present study demonstrate that the facilitation effects of attention on the processing of objects with features from different senses are largely independent of modality.

The results of Experiment 2, with stimulus driven control of attention suggest quite the contrary and, at first glance, rather add to the ambiguity created by results of studies on exogenous cross-modal (i.e., visual-to-auditory) cueing. Some have found a cueing effect while others do not and methodical differences in those studies make it hard to draw a clear cut conclusion (see Spence et al., 2004, for a review). The consensus is that visual cues can influence auditory perception, but only under certain conditions that are not yet fully understood. A number of arguments have been put forward that aimed at explaining the conditions under which cross-modal cueing occurs. The arguments include response task, differences in spatial resolution between modalities, and criterion shifts, amongst others (Spence et al., 2004; Koelewijn et al., 2010). It has been argued, for example, that

criterion shifts rather than facilitation by attention account for the faster responses to targets presented at the pre-cued location. This concern applies not specifically to cross-modal attention studies, but to all studies that are based upon the original spatial cueing task (Posner et al., 1978; Posner, 1980; Jonides, 1981). A perceptual decision task can help to distinguish criterion shifts from attention effects, for example, by application of the signal detection model (Green & Swets, 1966). Here, we applied an alternative, but related model (Wagenmakers et al., 2007) and obtained evidence that it was indeed perceptual processes that were facilitated or attenuated, depending on the validity of the cue. If criterion shifts accounted for the results, one would expect that also the latency of non-perceptual processes were affected by cue validity. Rather, the redundancy gains observed with exogenous cues indicate that it was the processing of visual features that was facilitated when they were attended. The processing of auditory features was largely the same, irrespective of the preceding cue. In a recent review, Talsma et al. (2010) integrated the diverse literature of cross-modal cueing effects on multisensory integration into a common framework. They reviewed studies involving top-down as well as bottom-up control of cross-modal attention and propose models for both forms of attentional control. According to this model, exogenous cues lead to a spatial unspecific spread of attention across the senses, at least in the audition-to-vision-direction. This conclusion was based on the finding that auditory cues can produce location unspecific facilitation effects on visual processing (Van der Burg et al., 2008). On this note, the observed data in Experiment 2 could be the result of visual cues leading to cross-modal facilitation of (auditory) perception, irrespective of the location of the subsequent (auditory) target. This would not only explain the null effects of visual cues on the latency of auditory target perception, but also the null effect on all audiovisual targets that we observed in the redundant signals conditions. The comparison of response times and the estimated model parameter from that data suggest an alternative interpretation.

Both mean response times as well as parameter estimates support the notion that auditory targets were never attended to, regardless of whether the preceding visual cue was pointing to the correct location or the opposite location. On this note, the visual modality became dominant during the integration of redundant signals. This dominance could have been induced by the preceding visual cue. However, if one considers visual dominance as a consequence of visual cues, this also leads to another implication. If visual processing dominated the perception of redundant signals after exogenous (but not endogenous) cues, the greater resolution of the visual system might have facilitated the effects of spatially separated visual targets and auditory targets (Spence et al., 2004; Koelewijn et al., 2010). This would explain why the processing of the same auditory targets presented at the same locations was effectively facilitated in Experiment 1, but not in Experiment 2. Of course, this interpretation needs to be confirmed by data as, to our knowledge, the criticality of spatial separation of auditory and visual targets in stimulus-driven (but not voluntary attention) has often been discussed, but never been directly investigated. The results from the redundant audiovisual conditions support existing theories on endogenous cross-modal interactions and shed new light on exogenous cross-modal interactions. Upcoming studies on exogenous links between vision and audition can help to further refine the relationship between the two systems.

6 Effects of Spatial and Selective Attention on Basic Multisensory Integration

Matthias Gondan[a,b], Steven P. Blurton[b], Flavia Hughes[b], and Mark W. Greenlee[b]

[a]Institute for Medical Biometry and Informatics, University of Heidelberg

[b]Department of Psychology, University of Regensburg

Abstract[7]

When participants respond to auditory and visual stimuli, responses to audiovisual stimuli are substantially faster than to unimodal stimuli (redundant signals effect, RSE). In such tasks, the RSE is usually higher than probability summation predicts, suggestive of specific integration mechanisms underlying the RSE. We investigated the role of spatial and selective attention on the RSE in audiovisual redundant signals tasks. In Experiment 1, stimuli were presented either centrally (narrow attentional focus) or at 1 of 3 unpredictable locations (wide focus). The RSE was accurately described by a coactivation model assuming linear superposition of modality-specific activation. Effects of spatial attention were explained by a shift of the evidence criterion. In Experiment 2, stimuli were presented at 3 locations; participants had to respond either to all signals regardless of location (simple response task) or to central stimuli only (selective attention task). The RSE was consistent with task-specific coactivation models; accumulation of evidence, however, differed between the 2 tasks.

Keywords: multisensory processes, spatial attention, selective attention

[7]erschienen als: Gondan, M., Blurton, S. P., Hughes, F., & Greenlee, M. W. (2011). Effects of spatial and selective attention on basic multisensory integration. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1887–1897.

Visuospatial attention is often seen in an analogy to a spotlight which means that stimuli falling into this spotlight are processed more efficiently at the cost of those that do not (Posner et al., 1980; Cave & Bichot, 1999). When attention is zoomed in, perception of stimuli at the center of focus becomes more efficient; the greater the distance to this focus, the less efficient perception becomes (Castiello & Umiltà, 1990). Traditionally, selective spatial attention has been studied with auditory stimuli (e.g., Broadbent, 1952) or with visual stimuli (Posner, 1980). However, everyday perception is mostly multisensory in nature. As different sensory systems provide complementary, redundant, or conflicting information (Welch & Warren, 1986), effective behavior requires integration of the sensory signals provided by the different senses. Thus, research effort has been directed to investigate the role of attention in multisensory, especially audiovisual (Driver, 1996; Spence & Driver, 1997) and visuotactile perception (Macaluso et al., 2000). One of the most fundamental questions that arise with spatial attention is whether there exists a common, supramodal attentional system or several, independent subsystems. With cueing paradigms, (Spence & Driver, 1997) demonstrated that visuospatial attention can be directed by auditory cues. On the other hand, some degree of independence has also been found between attentional resources of different sensory systems (Alais, Morrone, & Burr, 2006). While it is widely accepted that spatial attention is a multisensory phenomenon, the exact role of attention in multisensory perception still remains unclear.

In research on multisensory processes, the most basic experimental setup is the bimodal redundant signals paradigm: Participants are asked to respond in the same way to stimuli of two different modalities (e.g., auditory and visual, A, V). In some trials, both stimuli are presented (AV), and this stimulus combination is referred to as the redundant signals condition. In the redundant signals condition, responses are usually substantially faster than in the single target conditions (e.g., Raab, 1962). At first glance, this so-called redundant signals effect (RSE, e.g.,

Kinchla, 1974) might be taken as sufficient evidence for the existence of genuine multisensory integration mechanisms. However, different processing architectures can account for redundancy gains in such tasks; the most important model classes are race models (or, more generally, separate activation models, e.g., Raab (1962) and coactivation models (Miller, 1982). In the race model, both components of a bimodal stimulus are processed in parallel channels; the overall processing time is determined by the channel which has first finished processing (e.g., having reached a threshold first). This mechanism eliminates slow processing times from the modality-specific distributions, which, on average, results in faster responses to bimodal events. The redundancy gain of the race model has an upper limit, though; this upper limit is known as the race model inequality (Miller, 1982):

$$F_{\mathrm{AV}}(t) \leq F_{\mathrm{A}}(t) + F_{\mathrm{V}}(t), \text{ for all } t, \tag{1}$$

with $F(t) = P(\mathbf{T} \leq t)$ denoting the probability for a response latency $\mathbf{T}$ within t ms. In bimodal divided attention, response times for AV have often been observed to violate Inequality 1.

Instead, coactivation models have been proposed which specify a more or less explicit integration mechanism (Miller, 1982, App. A; Miller, 1986, Eq. 3; Schwarz, 1989, 1994; Diederich, 1995; Miller & Ulrich, 2003). Several coactivation models assume linear superposition of modality-specific information (Schwarz, 1989, 1994; Diederich, 1995; Miller & Ulrich, 2003). Let $\mathbf{X}_{\mathrm{A}}(t)$, $\mathbf{X}_{\mathrm{V}}(t)$ denote the stochastic processes describing the buildup of evidence in the auditory and visual channel, respectively. Superposition models assume that the activation of the combined channels corresponds to the sum of the two sensory-specific channels, $\mathbf{X}_{\mathrm{AV}}(t) = \mathbf{X}_{\mathrm{A}}(t) + \mathbf{X}_{\mathrm{V}}(t)$. Detection occurs whenever an evidence criterion c is surpassed for the first time. For time-homogenous diffusion processes underlying the channel-specific buildup of evidence $\mathbf{X}_{\mathrm{A}}(t)$, $\mathbf{X}_{\mathrm{V}}(t)$, Schwarz (1994) derived pre-

dictions for the mean and the variance of the detection time $\mathbf{D}$ for unimodal and bimodal stimuli presented simultaneously, or with onset asynchrony τ:

$$
\begin{aligned}
E[\mathbf{D}_{\mathrm{A}}] &= c/\mu_{\mathrm{A}}, \\
E[\mathbf{D}_{\mathrm{V}}] &= c/\mu_{\mathrm{V}}, \\
E[\mathbf{D}_{\mathrm{AV}}] &= c/(\mu_{\mathrm{A}} + \mu_{\mathrm{V}}), \\
\mathrm{E}\big[\mathbf{D}_{\mathrm{A}(\tau)\mathrm{V}}\big],\ \mathrm{E}\big[\mathbf{D}_{\mathrm{V}(\tau)\mathrm{A}}\big] &\text{: see Schwarz (1994, Eq. 10),}
\end{aligned}
\tag{2}
$$

with μ_{A}, μ_{V} denoting the drift rates of the modality-specific diffusion processes, and $c > 0$ denoting an absorbing barrier (i.e., the evidence criterion). Assuming a SOA invariant μ_{M} summarizing the mean duration of processes not described by the model, Schwarz (1994, Figure 1) demonstrated that the diffusion superposition model well described the mean response Times $E(\mathbf{T}) = E(\mathbf{D}) + \mu_{\mathrm{M}}$ observed by Miller (1986) in a simple speeded response task with audiovisual stimuli presented at different onset asynchronies (for the standard deviations, see Schwarz, 1994, Figure 2).

The present study investigates the integration of redundant signals under different attentional conditions on the basis of the diffusion superposition model. It is usually assumed that "integration" requires spatial attention; for example, to solve the binding problem in visual object perception as posited by the feature integration theory (Treisman & Gelade, 1980; Treisman, 1986). In the dimensional action model, which incorporates many ideas of the feature integration theory, attention plays a central role (Cohen & Shoup, 1997). The model assumes that properties of a visual stimulus are decomposed into a number of dimensional modules (e.g., for form, color, orientation, etc.). Each dimensional module detects the presence or absence of features in its respective dimension. The activation elicited by these features is then transmitted to dimension-specific response selection processors (Cohen & Feintuch, 2002). For redundant signals of the same dimension,

the model predicts only limited redundancy gains, because both stimuli activate only a single response selector (e.g., Miller et al., 2009). In redundant signals of different dimensions, however, two response selectors are simultaneously active, yielding especially fast responses. Indeed, Feintuch and Cohen (2002) observed that response time distributions for redundant color orientation targets violate the race model inequality (1), however, only if the stimulus components were presented in close spatial proximity such that participants could direct spatial attention to the location of the target.

By analogy, one would expect that spatial attention is necessary for multisensory coactivation, as well. However, the role of attention in multisensory integration remains controversial (e.g., Navarra, Alsius, Soto-Faraco, & Spence, 2010). It has been argued that connections between auditory and visual cortices are so abundant that multisensory integration processes do not require spatial attention (Bertelson et al., 2000). Moreover, multisensory processing can precede attentional allocation (Driver, 1996). On the other hand, spatial attention has been shown to affect the earliest multisensory components of the event-related potential (Talsma et al., 2007), which at least suggests that attention is involved in multisensory integration processes.

Concerning the redundant signals effect, the exact role of attention in effective integration is not yet fully understood (Miller et al., 2009). It is known that under some circumstances, attention can be assigned to more than one location at a time (Castiello & Umiltà, 1992; McCormick, Klein, & Johnston, 1998; Dubois, Hamker, & VanRullen, 2009). This split of spatial attention is advantageous in the sense that a series of locations do not have to be attended to in a serial manner. The advantage comes at a cost; however, splitting spatial attention means dividing resources, leaving less beneficial effects of attention the more locations one is assigning attention to (Castiello & Umiltà, 1992). It remains unclear whether the attentional spot is enlarged or simply fewer focused to encompass all target

locations (e.g., those indicated by cues), or if the attentional system is capable of dividing the attentional focus to several locations simultaneously (Bichot, Cave, & Pashler, 1999).

In the present study, we directly compared two conditions of spatial attention (narrow focus, wide focus) in a redundant signals experiment using audiovisual stimuli with varying onset asynchronies between the auditory and visual stimulus components (Experiment 1). The obtained mean response times were then modeled with a diffusion model of the redundant signals effect (Schwarz, 1994), separately for both attentional conditions. In line with earlier results, we expected that the diffusion superposition model can describe the redundancy gains in the narrow focus condition. We additionally applied a common model to identify similarities and differences in audiovisual perception for the different levels of spatial attention. Whereas elementary modality-specific perception to the same stimuli can be expected to be similar, it is unclear whether the same superposition mechanism holds for the two attentional conditions.

In a second experiment, we used the same diffusion model approach to compare mean reaction times of simple responses with those of selective attention. In this experiment, stimulation was exactly the same under both conditions, but for the selective attention condition, observers were instructed to attend to only one location and to disregard stimuli at the other locations. Earlier results (Gondan et al., 2010) obtained from a Go/No-go task indicate that coactivation cannot be taken for granted in tasks more complex than speeded responses. The main question addressed by Experiment 2 was, therefore, whether the superposition model can explain redundancy gains in a selective attention task, and if so, whether modality-specific processing is qualitatively the same for the different tasks.

6.1 Experiment 1

In Experiment 1, participants made speeded responses to auditory, visual, and audiovisual signals presented with onset asynchrony. In the narrow focus condition, stimuli were presented at a single central position only; here participants could concentrate on the source of stimulation. In the wide focus condition, stimuli were presented randomly at one of three possible locations (left, right, center), such that participants had to enlarge their attentional spotlight in order to attend to all three locations.

6.1.1 Methods

Participants. Four right-handed volunteers (3 students from the University of Regensburg, one male, two female, mean age 22 years, and one author) participated in the experiment. All reported normal hearing and normal or corrected-to-normal visual acuity with an intact field of view. Except for the coauthor, the participants were naive regarding the purpose of the experiment and the stimulus conditions employed. Informed consent was obtained from all participants prior to participation. Results were stored in anonymous form. Participants received course credit or payment (7€ per hour) for their participation. The experiment was conducted in accordance to the standards laid down in the Declaration of Helsinki.

Apparatus. The experiment was conducted in a light- and sound-proof room (Industrial Acoustics Company GmbH, Niederkrüchten, Germany), which was dimly illuminated from behind and above. The participants were directly facing the stimulation device, which was placed on a desk at a distance of 60 cm. The device consisted of a projection screen for the visual stimuli and three mobile loudspeakers for the auditory stimuli, which had been placed on elevated platforms in the central position and the outer left and outer right side of the desk. Stimulus

presentation and response time recording was controlled by a standard personal computer running "Presentation" (Neurobehavioral Systems, Albany, California).

Stimuli. A Gabor-patch was projected at three different positions on a uniform gray background using a luminance-calibrated liquid crystal display projector: the center, the outer left, and the outer right (angle 30° each) side of the screen. White noise (50 dB) served as the auditory stimulus and was emitted via three loudspeakers in the same central, left, and right positions (not visible to the participant). Audiovisual signals were presented in spatial correspondence, at 13 stimulus onset asynchronies (SOAs, cf. Miller, 1986)(SOAs, cf. Miller, 1986): A, A167V, A133V, A100V, A67V, A33V, AV, V33A, V67A, V100A, V133A, V167A, and V (SOA in ms). Catch trials (C, i.e., trials in which no stimulus appeared at the usual stimulus onset) were embedded in the experimental procedure to discourage anticipatory guesses. The interstimulus interval varied uniformly between 2100 ms and 3000 ms. Participants had to respond to the stimuli by pressing a response button with their dominant hand.

Experimental tasks. In the wide focus condition (WID), the participants were told to respond as quickly as possible to any detected signal at any possible position (left, center, right). Because the participants did not know the position at which the stimulus would appear, they had to spread their attention over all three locations. In the narrow focus condition (NAR), stimuli appeared only in the center of the screen or from the central loudspeaker. Here, participants knew the position at which the stimuli would appear, so they could concentrate on this position.

Procedure. Each participant was tested in three sessions of about 3 hours each. At the start of the experiment, the participants were instructed and a training block was run with the same stimulus protocol as used in the main session. The

main session was divided into 13 blocks of 10 min each. Breaks were made on request of the participants; usually participants requested a break of 20 to 30 min at the middle of the session. Each block comprised both experimental tasks, so each block started with a screen indicating the current experimental condition, the WID or NAR. The participants had to fixate a plus (+) sign which appeared at the center of the screen during the interstimulus intervals. In both tasks, the stimuli were presented in a randomized sequence. Each of the 3 (WID)/1 (NAR) locations SOA stimulus conditions appeared three times within each block, yielding a maximum of 110 replications per experimental condition.

Test of the race model. The race model inequality was tested after cleaning the response time distributions of the different conditions using the "kill-the-twin" procedure (Eriksen, 1988). In the kill-the-twin procedure, the response time distribution for catch trials $F_{\mathrm{C}}(t) = P(\mathbf{T}_{\mathrm{C}} \leq t)$ is subtracted from the response time distributions of all SOA-specific conditions (Gondan & Heckel, 2008). For Condition V(τ)A, the modified inequality, thus, reads as

$$\left[F_{\mathrm{V}(\tau)\mathrm{A}}(t) - F_{\mathrm{C}}(t)\right] \leq \left[F_{\mathrm{V}}(t) - F_{\mathrm{C}}(t)\right] + \left[F_{\mathrm{A}}(t-\tau) - F_{C}(t-\tau)\right], \text{ for all } t. \quad (3)$$

Miller (1986) suggested to measure the amount of violation of Inequality 3 by the positive area enclosed by the AV curve and the summed A and V curves. We used a nonparametric variant of this area based on the rank-transformed data,

$$\Delta_\tau = \sum_{i=1}^{N} \max\{0, [F_{\mathrm{V}(\tau)\mathrm{A}}(t_i) - F_{\mathrm{C}}(t_i)] - [F_{\mathrm{V}}(t_i) - F_{\mathrm{C}}(t_i)] - [F_{(\tau)\mathrm{A}}(t_i) - F_{(\tau)\mathrm{C}}(t_i)]\},$$

with $F_{\tau\mathrm{A}}$ denoting the shifted response time distribution for auditory stimuli. This measure is scale-invariant and robust with regard to outliers. The violation area was measured in all SOA conditions and collapsed into an aggregate violation area by a weighted sum $\Delta = \sum_\tau \lambda(\tau) \times \Delta_\tau$, with $\lambda(\tau) > 0$ denoting a triangular

weighting function assigning weights 1, 2, 3, 4, 5, 6, 5, 4, 3, 2, 1 to Conditions A167V, A133V, A100V, A67V, A33V, AV, V33A, V67A, V100A, V133A, V167A, respectively ("symmetric umbrella," Gondan, 2009). To test whether $\Delta > 0$ observed in a given participant reflects true coactivation or is due to sampling error, 10,000 computer simulations were performed (Miller, 1986). In each simulation, bootstrap samples of the unimodal response times were drawn from the observed response time distributions, bimodal response times were bootstrapped from the distribution of minima of the unimodal response times, adjusted for SOA and assuming a maximally negative channel correlation between A and V (Ulrich & Giray, 1986). In each simulation, the aggregate violation area Δ^* was determined, resulting in 10,000 simulated Δ^*. The race model is rejected at $p < .05$, if the observed Δ is greater than 95% of the Δ^* values under the race model assumption.

Diffusion superposition model. The diffusion superposition model predicts the mean response times for audiovisual stimuli with given SOA τ using five free parameters: drift and variance of the auditory (μ_A, σ^2_A) and the visual process (μ_V, σ^2_V), and a mean residual μ_M summarizing everything not described by the model. In the task-specific models, the barrier c was fixed at 100, since it only scales the other parameters.

For the model fit, trimmed mean response times were used, excluding the upper and the lower 2% of the response times, separately for each condition. Goodness-of-fit χ^2 was calculated by adding up the squared standardized differences χ^2_τ between model prediction $\mathrm{E}[\mathbf{T}_\mathrm{V}(\tau)\mathrm{A}]$ and observed mean response time $m_{\mathrm{V}(\tau)\mathrm{A}}$ for each SOA:

$$\chi^2 = \sum_\tau X^2_\tau = \sum_\tau \{m_{\mathrm{V}(\tau)\mathrm{A}} - \mathrm{E}[\mathbf{T}_{\mathrm{V}(\tau)\mathrm{A}}]\}^2 / \{\hat{s}^2_{\mathrm{V}(\tau)\mathrm{A}}/n_{\mathrm{V}(\tau)\mathrm{A}}\}, \tag{4}$$

with $\hat{s}^2_{\mathrm{V}(\tau)\mathrm{A}}/n_{\mathrm{V}(\tau)\mathrm{A}}$ denoting the square of the observed standard error. If the model holds, the means $m_{\mathrm{V}(\tau)\mathrm{A}}$ are approximately normally distributed around $\mathrm{E}[\mathbf{T}_{\mathrm{V}(\tau)\mathrm{A}}]$,

and, thus, for large $n_{V(\tau)A}$, the squared standardized means converge to an approximate χ^2_1 distribution. As the five model parameters are adjusted to the means observed in 13 SOAs, the sum of the SOA-specific χ^2_τ values approximately follows a χ^2 distribution with $13 - 5 = 8$ degrees of freedom. The model was adjusted to the observed mean response times by minimizing 4(4) using the constrOptim command of the R statistical language (R Development Core Team, 2010) with restrictions $0.1 \leq \mu_A$, $\mu_V \leq 4$, $10 \leq \sigma^2_A$, $\sigma^2_V \leq 10,000$, $100 \leq \mu_M \leq 1000$.

In a first step, separate models were adjusted to the mean response times observed in the two tasks. As stated above, for the WID, only the responses to central stimuli were analyzed. In a second step, an aggregate model was adjusted to mean response times of both the NAR and the WID (central stimuli only). In this aggregate model, diffusion parameters were assumed to be equal for both tasks; task-specific processing and attentional demands were taken into account by allowing different evidence criteria c_{WID} and c_{NAR} for the two tasks ($c_{WID} = 100$, c_{NAR}: variable).

6.1.2 Results

In the NAR, participants knew that stimuli were presented at the central location only; whereas in the WID, stimulus presentation was randomized, with one third of the stimuli presented left or right or at the central position. Direct comparison of the response times in the two tasks is, thus, most informative for centrally presented stimuli. For these stimuli, mean response times for Participants 1, 2, 3, and 4 were lower in the NAR than in the WID (averaged over SOA, 17, 31, 14, 13 ms for Participant 1, 2, 3, and 4, respectively). Within both tasks, mean response times showed a wing-shaped pattern (see Figure 6.1), replicating the usual relationship between SOA and mean response time observed in redundant signals tasks with asynchronous targets (Ulrich & Miller, 1997; Schwarz, 1994). Omissions and re-

sponses to catch trials were extremely rare (false alarm rate 0 %, omission rate around 1 %) and were, thus, not further analyzed.

Race model. After correction for fast guesses using the kill-the-twin procedure (Eriksen, 1988), violation areas of the race model inequality were obtained for each SOA and added up using a symmetric umbrella weighting function (Gondan, 2009). These observed summed violation areas were compared with their bootstrap distribution under the race model assumption (Miller, 1986). Consistent violations of the race model inequality were observed for the WID (p values of bootstrap test .026, .040, .014, $<$ .001 for Participants 1, 2, 3, and 4, respectively). In the NAR, violations of the race model inequality reached statistical significance for three participants (p = .003, .180, $<$ .001, $<$ .001 for Participants 1, 2, 3, and 4, respectively).

Diffusion superposition model. In a first step, task-specific diffusion superposition models were adjusted to the mean response times observed in the WID and NAR. Fitted parameters for these models are shown in Table 6.1. In line with the substantially lower mean reaction times for auditory stimuli, the drift rates for the auditory process are higher than for the visual process. The task-specific models show acceptable goodness-of-fit in all participants (summarized goodness-of-fit statistic for the WID: $\chi^2 = 44.92$, $df = 32$, $p = .064$; for NAR: $\chi^2 = 20.98$, $df = 32$, $p = .932$). In a second step, an aggregate model was adjusted to the mean response times for both tasks, with common parameters describing the diffusion processes and the residual, but different evidence barriers ($c_{\mathrm{WID}} = 100$ fixed, $c_{\mathrm{NAR}} < 100$) for the two tasks. Figure 6.1 illustrates the good agreement between predicted and observed mean response times in the two tasks. The aggregate model adequately describes the observed mean response times ($\chi^2 = 106.55$, $df = 80$, $p = .025$); however, model fit is poor in one participant (right panel of Figure 6.1).

Table 6.1: *Diffusion Superposition Model for the WID and the NAR, and the Common Model for Both Tasks of Experiment 1*

	WID				NAR				Common model			
Participant	1	2	3	4	1	2	3	4	1	2	3	4
μ_V	0.86	0.72	1.14	1.13	0.79	0.57	0.79	1.13	0.83	0.63	0.94	1.13
σ_V^2	71.1	26.9	10.0	10.0	47.5	24.4	25.4	25.3	60.3	23.3	31.9	16.1
μ_A	2.52	1.25	2.16	2.03	1.89	1.19	1.39	1.90	2.22	1.14	1.79	1.98
σ_A^2	10.0	256.4	3157	10.0	10.0	530.3	10.0	37.8	10.0	345.0	209.2	18.5
c_{WID}	(100)	(100)	(100)	(100)	—	—	—	—	(100)	(100)	(100)	(100)
c_{NAR}	—	—	—	—	(73.1)	(73.8)	(81.0)	(77.6)	73.1	73.8	81.0	77.6
μ_M	200.2	268.9	197.2	130.8	190.7	260.3	173.4	130.4	195.6	261.6	185.5	130.7
GOF χ^2	9.00	9.12	11.07	15.73	4.73	0.94	9.52	5.79	15.49	19.91	45.27	25.88
df	8	8	8	8	8	8	8	8	20	20	20	20
P	.343	.333	.197	.046	.785	.999	.301	.670	.748	.463	.001	.170
Summary	$\chi^2_{(32)} = 44.92$, $p = .064$				$\chi^2_{(32)} = 20.98$, $p = .932$				$\chi^2_{(80)} = 106.55$, $p = .025$			

Note. μ_V, σ_V^2 (μ_A, σ_A^2): drift and variance for visual (auditory) diffusion process, c: evidence barrier for the WID and the NAR (fixed values in parentheses), μ_M: mean duration of residual component. GOF χ^2, *df*, *p*: goodness-of-fit statistic (significant results indicate bad fit). Summary is based on the sum of the participant-specific GOF statistics.

6.1.3 Discussion

Our goal of the first experiment was to investigate whether and how redundancy gains are affected by spatial attention. Abstract artificial stimuli were used which are known to be very effective in their respective modalities. Participants had to perform a simple response task to audiovisual target stimuli under two attentional conditions. In the NAR (narrow focus), stimuli were presented at a constant predictable location in the center of fixation. In the WID (wide focus), stimuli were presented randomly at one of three different locations. The central stimuli used in the two attentional conditions had the same physical properties. In the NAR, participants could concentrate on the central location, whereas in the WID, participants had to attend to all three locations simultaneously.

Indeed, mean response times were lower in the NAR as compared with the same stimuli in WID. The magnitude of the attention effect was, however, small in terms of absolute reaction times (see Figure 6.1). Although the attentional requirements differed for the NAR and the WID, the participants were asked to fixate the central location in both conditions. Therefore, it cannot be ruled out that participants actually concentrated more on the central position in the WID than on the two peripheral positions. This would explain the rather small attention effect in Experiment 1. Except for one participant, the test of the race model inequality revealed significant violations of the race model inequality in both the NAR and the WID, indicative of coactive processing of the redundant information.

In the first step of the main analysis, task-specific diffusion superposition models (Schwarz, 1994) were adjusted to the mean
response times observed in the two tasks. Good agreement between model and data is evident in Figure 6.1 and Table 6.1. Replicating earlier results by Schwarz (1994), Diederich (1995), and Gondan et al. (2010), mean response times for asynchronous audiovisual stimuli can be well described by a model assuming linear superposition of channel-specific activity. In simple tasks, multisensory "integration" can thus be

reduced to a simplistic additive channel summation mechanism, without necessity of superadditive neural circuitry (e.g., Stanford et al., 2005).

In summary, Experiment 1 supports the superposition model for simple responses to audiovisual stimuli for different levels of spatial attention. What then is the role of attention in multisensory integration? Does integration occur in a basic bottom-up manner, or is it necessary to direct spatial attention to the location of the stimuli in order to effectively integrate them (e.g., Feintuch & Cohen, 2002)? We addressed this question by fitting an aggregate superposition model to the response times observed for the central stimuli common to both tasks. Assuming that the stimulus-specific diffusion processes describe elementary perceptual processes common to the two conditions, the diffusion parameters μ_{A}, σ^2_{A}, μ_{V}, σ^2_{V}, were constrained to be equal in the two tasks. Moreover, as both tasks were simple response tasks (Type A, Donders, 1868/1969), residual processes described by μ_{M} were assumed to be equal in the two tasks as well. Goodness-of-fit of this aggregate model was acceptable in three participants. Different evidence barriers were allowed in the two tasks: An increased absorbing barrier in the WID, or, equivalently, a reduced criterion in the NAR, reflecting the improvement in stimulus detection when spatial attention is directed to the source of stimulation (e.g., Eimer & Driver, 2000, Figure 1; Hillyard, Hink, Schwendt, & Picton, 1973).

In Participant 3, fit of the common model was poor, although the results were qualitatively similar to the other participants. Closer inspection of Figure 6.1 suggests that in Participant 3, the attentional effect is limited to the visual modality—in the left (auditory) wing of the SOA-mean curve, an attentional effect is virtually absent. In the aggregate model, the barrier c corresponds to the evidence criterion common to both modalities. An increased barrier, thus, affects both modalities simultaneously (recall that the mean detection time for auditory and visual stimuli corresponds to c/μ_{A} and c/μ_{V}, respectively). This model prediction is in line with the supramodal nature of attentional effects observed in crossmodal

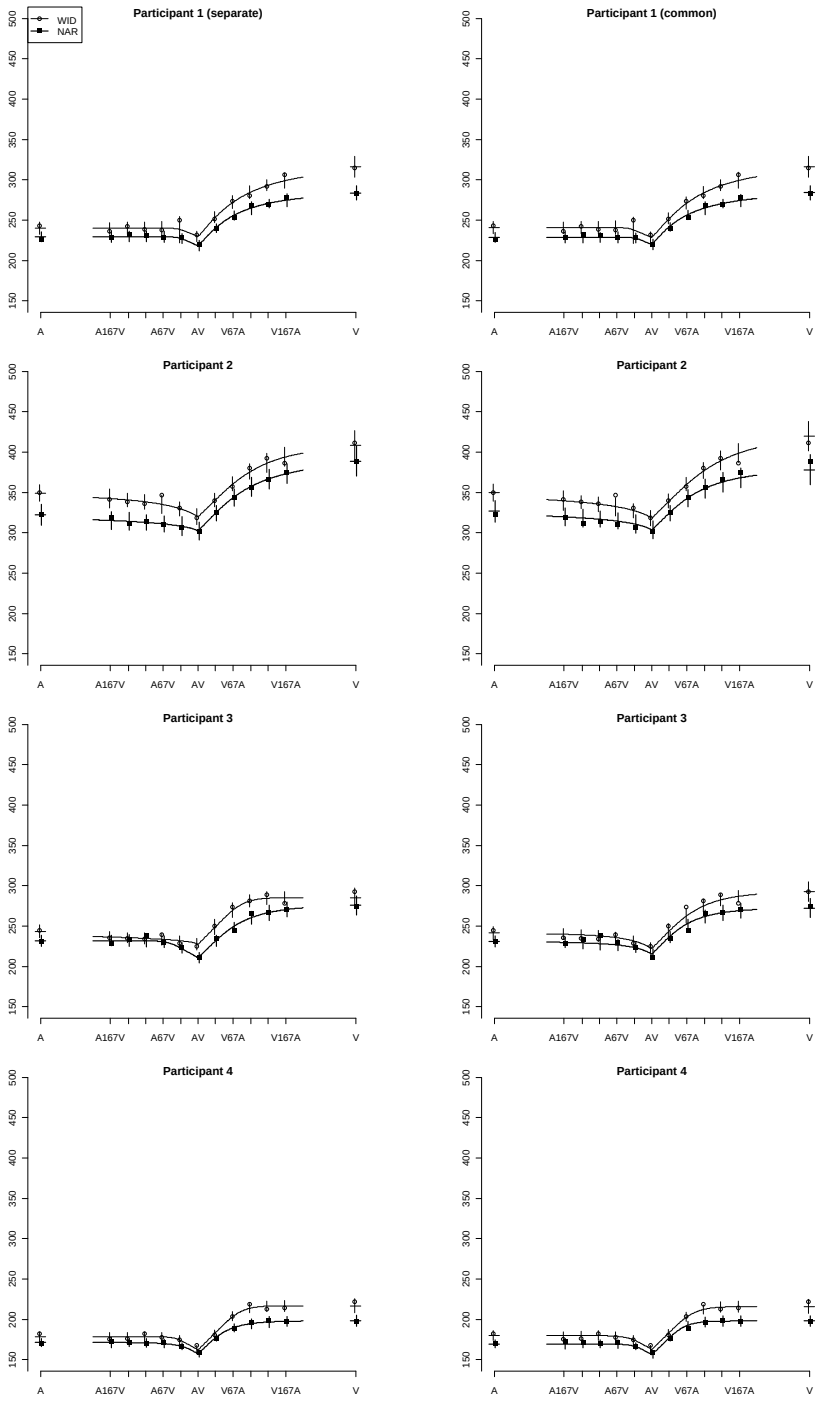


Figure 6.1: Experiment 1—SOA specific mean response times observed (dots) in the NAR and the WID (central stimuli only). Lines: Model prediction including 95% confidence intervals based on the observed standard deviation. Left: Task-specific models. Right: Common aggregate model with task-specific evidence barrier.

attention tasks (e.g., Eimer & Driver, 2000). Participant 3's results are incompatible with this supramodal notion of spatial attention: In this participant, effects of spatial attention were limited to the visual modality only (cf. Driver & Spence, 1998, Box 1).

6.2 Experiment 2

In Experiment 2, we introduced a Go/Nogo feature in order to investigate if the superposition model can describe the redundancy gains in selective attention tasks.

6.2.1 Methods

Participants. Seven new students from the University of Regensburg (one male, six female, mean age 24.2 years, one left-handed) participated in Experiment 2. All reported normal hearing and normal or corrected-to-normal visual acuity with an intact field of view. The participants were naive regarding the purpose of the experiment and the stimulus conditions employed. Informed consent was obtained from all participants prior to participation. They received course credit or payment for participation.

Experimental tasks. The apparatus and the stimulus conditions employed were identical to Experiment 1. The simple response task was the same as in the WID of Experiment 1: The participants were told to respond as quickly as possible with their dominant hand to any detected signal at any possible location (left, center, right).

The second task was a selective attention task (SEL): Stimuli were presented, in randomized order, at three locations, but participants were instructed to respond to the central stimuli only (Go trials), and refrain from responding to peripheral stimuli (Nogo).

Procedure. Due to the additional peripheral stimuli in the SEL, the entire experiment prolonged to about 12 hours per participant. Data acquisition was split again into three sessions. In each session, 10 blocks of about 15 min duration were conducted; each block comprised both attentional conditions. Again, the first block served as a training block and the data were not analyzed. Breaks were made on request of the participant. Each of the three locations × SOA stimulus conditions appeared three times within each block, yielding a maximum of 87 replications per experimental condition.

Race model test. For the WID, the race model inequality was tested in the same way as described for Experiment 1. For the SEL, a kill-the-twin correction was applied using the erroneous responses to peripheral stimuli:

$$\left[F_{\mathrm{V}(\tau)\mathrm{A}}(t) - F_{\mathrm{v}(\tau)\mathrm{a}}(t)\right] \leq \left[F_{\mathrm{V}}(t) - F_{\mathrm{v}}\right] + \left[F_{\mathrm{A}}(t-\tau) - F_{\mathrm{a}}(t-\tau)\right], \text{ for all } t, \quad (5)$$

with $F_{\mathrm{v}(\tau)\mathrm{a}}(t) = \max\left[F_{\mathrm{v}(\tau)\mathrm{a}}(t|\mathrm{left}),F_{\mathrm{v}(\tau)\mathrm{a}}(t|\mathrm{right})\right]$, $F_{\mathrm{v}}(t) = \min\left[F_{\mathrm{v}}(t|\mathrm{left}),F_{\mathrm{v}}(t|\mathrm{right})\right]$, $F_{\mathrm{a}}(t) = \min\left[F_{\mathrm{a}}(t|\mathrm{left}),F_{\mathrm{a}}(t|\mathrm{right})\right]$ denoting the false alarm distribution recorded for peripheral stimuli presented to the left and to the right location (see Gondan et al., 2010, for a similar procedure).

Superposition model. Again, separate diffusion superposition models were fitted to the mean response times observed in the WID and the SEL. In a second step, we tried to adjust a common model to the two tasks, with identical diffusion parameters describing perception of the same stimuli used in the two tasks, but different evidence barriers c_{WID} and c_{SEL} accounting for different attentional demands in the WID and the SEL. Whereas in Experiment 1, both tasks required simple responses, the SEL task of Experiment 2 requires Go/Nogo discrimination (Type C response, Donders, 1868/1969). This additional requirement was accounted for by allowing different residuals $\mu_{\mathrm{M,WID}}$ and $\mu_{\mathrm{M,SEL}}$ in the two tasks.

6.2.2 Results

As for Experiment 1, direct comparison of the response times observed in the WID and the SEL is most informative for centrally presented stimuli. Reflecting the increased control demands of the Go/Nogo responses, mean response times for Participants 1, 2, 3, 4, 5, 6, and 7 were substantially higher in the SEL than in the WID (130, 168, 157, 125, 507, 304, and 140 ms, respectively, see Figure 2). The relationship between SOA and mean response time followed the usual wing shape. Omission rate was below 1 % in the WID and below 2 % in the SEL (Participant 5: 5 %). In the SEL, responses to peripheral stimuli occurred in maximally 2 % of the stimuli. Misses and false alarms were, thus, not further analyzed (except for the kill-the-twin-correction).

Race model inequality. In the WID, violations of the race model inequality were observed for Participants 2, 3, 4, 6, and 7 (Part. 7: $p = .001$, others $p < .001$), whereas redundancy gains observed for Participants 1 and 5 were consistent with parallel processing ($p = .365$, .111, respectively). In SEL, coactivation effects were observed in all participants (all $p < .01$).

Diffusion superposition model. We first tried to adjust an aggregate model with identical diffusion parameters to the two tasks (Table 6.2, "Common Model"). The fit of this aggregate model was poor in all participants, and the model systematically underestimated the mean response times for auditory stimuli in the SEL, while auditory response times in the WID were systematically overestimated (summarized goodness-of-fit statistic: $\chi^2 = 1192$, $df = 133$, $p < .001$). Task-specific diffusion superposition models, however, can describe the mean response times recorded for the two tasks (see Figure 6.2). Parameters and goodness-of-fit statistics are summarized in Table 6.2 (columns WID and SEL). The model

fit is acceptable for the WID ($\chi^2 = 70.62$, $df = 56$, $p = .090$) and for the SEL ($\chi^2 = 74.08$, $df = 56$, $p = .053$).

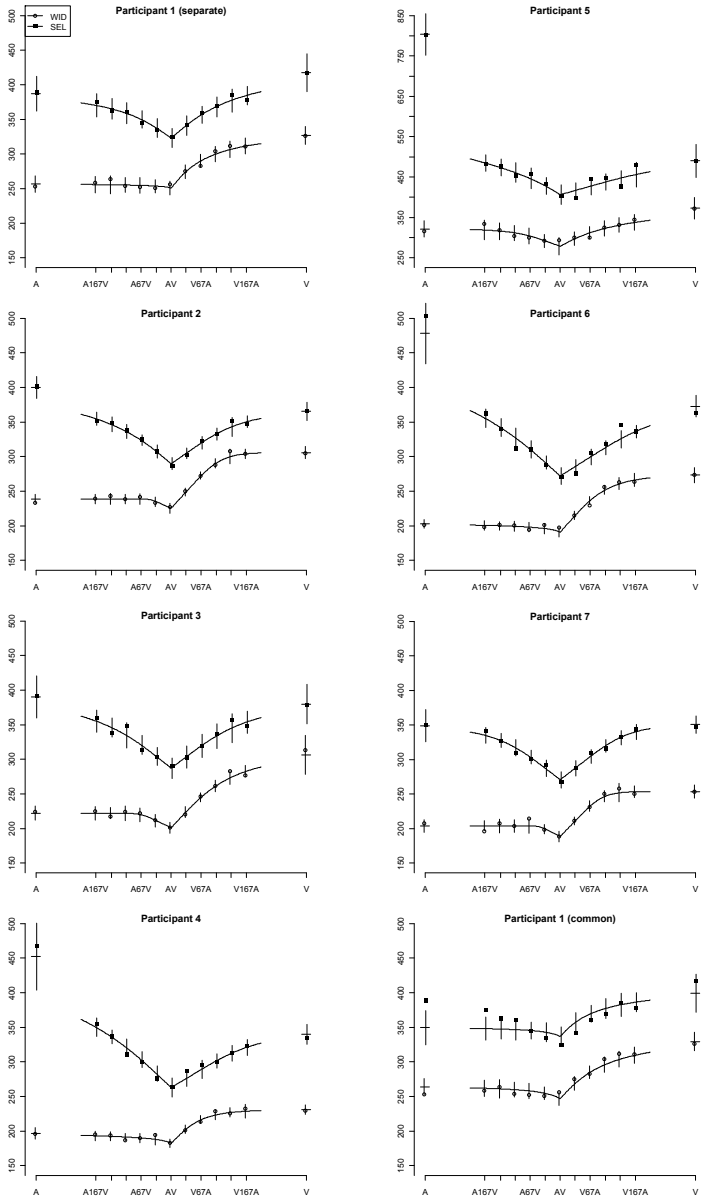


Figure 6.2: Experiment 2—Separate superposition models for the WID and SEL. A common model cannot be adjusted to the response times observed in the two tasks (lower right).

Table 6.2: *Diffusion Superposition Model for the WID and the SEL, and the Common Model for Both Tasks of Experiment 2*

	WID						
Participant	1	2	3	4	5	6	7
μ_{V}	1.04	0.88	0.66	1.33	0.63	0.88	1.04
σ^2_{V}	126.9	10.0	33.2	62.9	94.1	27.5	10.0
μ_{A}	3.80	2.16	1.46	2.42	0.94	2.33	2.14
σ^2_{A}	1997	10.0	10.0	1412	36.5	821.2	10.0
c_{WID}	(100)	(100)	(100)	(100)	(100)	(100)	(100)
c_{SEL}	—	—	—	—	—	—	—
$\mu_{\mathrm{M,WID}}$	230.1	192.4	153.8	155.5	214.5	159.6	156.9
$\mu_{\mathrm{M,SEL}}$	—	—	—	—	—	—	—
GOF χ^2	7.33	8.35	9.04	9.89	7.84	15.91	12.26
df	8	8	8	8	8	8	8
P	.502	.400	.339	.273	.450	.044	.140
Summary	$\chi^2_{(56)} = 70.62$, $p = .090$						
	SEL						
μ_{V}	0.58	0.60	0.53	0.51	0.38	0.41	0.63
σ^2_{V}	59.5	17.1	20.4	10.0	10.0	10.0	10.0
μ_{A}	0.70	0.50	0.50	0.32	0.17	0.29	0.64
σ^2_{A}	61.1	54.4	28.4	30.6	113.8	34.6	18.4
c_{WID}	—	—	—	—	—	—	—
c_{SEL}	(100)	(100)	(100)	(100)	(100)	(100)	(100)
$\mu_{\mathrm{M,WID}}$	—	—	—	—	—	—	—
$\mu_{\mathrm{M,SEL}}$	245.0	198.0	189.6	142.2	223.6	127.6	191.6
GOF χ^2	3.85	4.48	7.45	10.57	15.21	25.34	7.18
df	8	8	8	8	8	8	8
P	.871	.811	.489	.227	.055	.001	.518
Summary	$\chi^2_{(56)} = 74.08$, $p = .053$						

Note. μ_{V} (μ_{A}), σ^2_{V} (σ^2_{A}): drift and variance for visual (auditory) diffusion process, c: evidence barrier for the WID and the SEL (fixed values in parentheses), $\mu_{\mathrm{M,WID}}$, $\mu_{\mathrm{M,SEL}}$: mean duration of residual component in the WID and SEL. GOF χ^2, *df*, p: goodness-of-fit statistic (significant results indicate bad fit). Summary is based on participant-specific GOF statistics.

Table 6.2 (continued): *Diffusion Superposition Model for the WID and the SEL, and the Common Model for Both Tasks of Experiment 2*

	Common model						
Participant	1	2	3	4	5	6	7
μ_{V}	0.83	0.83	0.76	1.31	3.14	0.77	1.00
σ^2_{V}	88.4	10.0	10.0	40.7	408.2	33.2	10.0
μ_{A}	1.83	1.30	1.18	1.74	1.69	1.61	1.45
σ^2_{A}	337.4	10.0	56.1	10.0	8670	157.9	10.0
c_{WID}	(100)	(100)	(100)	(100)	(100)	(100)	(100)
c_{SEL}	75.5	150.1	143.4	215.2	999	78.1	147.1
$\mu_{\mathrm{M,WID}}$	209.3	176.3	148.6	144.7	282.1	145.9	145.7
$\mu_{\mathrm{M,SEL}}$	308.2	219.2	215.3	192.6	196.8	252.6	211.1
GOF χ^2	59.84	252.8	97.17	252.3	70.28	336.7	122.8
df	19	19	19	19	19	19	19
P	$< .001$	$< .001$	$< .001$	$< .001$	$< .001$	$< .001$	$< .001$
Summary	$\chi^2_{(133)} = 1192$, $p < .001$						

Note. μ_{V} (μ_{A}), σ^2_{V} (σ^2_{A}): drift and variance for visual (auditory) diffusion process, c: evidence barrier for the WID and the SEL (fixed values in parentheses), $\mu_{\mathrm{M,WID}}$, $\mu_{\mathrm{M,SEL}}$: mean duration of residual component in the WID and SEL. GOF χ^2, *df*, *p*: goodness-of-fit statistic (significant results indicate bad fit). Summary is based on participant-specific GOF statistics.

6.2.3 Discussion

In Experiment 2, we investigated the role of selective attention in a speeded response task with audiovisual stimuli presented at three different locations. In the first condition (WID), participants had to respond to stimuli presented at any of three locations, whereas in the second condition (SEL) they were asked to respond selectively to stimuli presented at the central location. Thus, the two conditions comprised the same stimulation but required different response types (simple vs. Go/Nogo). As for Experiment 1, response times in the two conditions were compared only for centrally presented stimuli. In line with the increased requirements of the Go/Nogo task, mean reaction times for the SEL were higher compared with the WID. Redundancy gains in the WID significantly violated the race model

inequality in only one participant, whereas clear evidence for coactivation was obtained in the SEL.

Task-specific superposition models for the mean reaction times showed an excellent fit for all participants, suggesting that linear superposition of channel specific diffusion processes (Schwarz, 1994) may well explain behavior in simple and more complex response paradigms (Figure 6.2, left column). Parameter estimates (see Table 2) for task-specific models indicate that selective attention affects both the diffusion and residual processes: (A) In the WID, drift rates turned out to be higher than in the SEL, whereas the mean residual turned out to be lower for the WID than for the SEL. Assuming that the overall response time $\mathbf{T} = \mathbf{D}+\mathbf{M}$ (Luce, 1986, ch. 3) decomposes into perception-related processes $\mathbf{D}$ being described by the diffusion model, whereas $\mathbf{M}$ summarizes everything else (e.g., motor preparation and execution), the manipulation of selective attention affects processing stages related to both $\mathbf{D}$ and $\mathbf{M}$: Increased drift rates estimated for the WID might reflect an increased buildup of evidence in this condition, but can, at the same time, indicate a lower amount of evidence necessary for stimulus detection. The higher μ_{M} observed in the SEL for Participants 1–5 and 7 is, presumably, due to the increased control demands of the Go/Nogo response selection processes.

The aggregate model incorporates responses of centrally presented stimuli in both tasks; it clearly fails to give a valid description of the mean reaction times in the two tasks. Model fit was poor in all participants ($p < .001$). A single scaling factor c is, thus, insufficient to describe the buildup of evidence under the two attentional conditions. Visual and auditory processing seems to be differentially affected when spatial attention is selectively assigned to one location as compared with the control condition with simple responses. Whereas auditory drift rates are about 2 or even 3 times greater than visual drift rates in the WID task, both drift rates are quite similar in the SEL task. Compared with the simple detection task, location discrimination might be substantially more difficult for auditory compared

with visual stimuli. This is reflected by the asymmetric effects of the attentional manipulation on the mean response times (see Figure 6.2). This differential effect cannot be accounted for by the aggregate model, which assumes that the auditory and the visual channels are both equally affected by the attentional manipulation. Interestingly, the task-specific models show good fit.

6.3 General Discussion

The goal of the present study was to investigate effects of different attentional conditions on mean response times in two redundant signals experiments. In Experiment 1 we compared two conditions of narrow and wide spatial attention; in Experiment 2 participants had to attend either selectively to a single spatial location or to three locations simultaneously. Crossmodal attention studies have provided abundant evidence for attentional mechanisms common to vision and audition (e.g., Spence & Driver, 1997) and vision and touch (e.g., Eimer & Driver, 2000), though there seems to exist some degree of independence between the different modalities (Alais et al., 2006).

What effect does attention have on audiovisual integration? While attention seems to be critical for early multisensory eventrelated potential interactions (Talsma et al., 2007), little is known about the effects of spatial attention on behavior, for example, audiovisual redundancy gains. In all conditions of the present experiments, mean response times were well described by a diffusion model based on linear superposition of modality-specific activation in the two channels (Schwarz, 1994). In Experiment 1, attentional modulation involved a change of the size of the attentional focus and the expected results were obtained. Attention specific benefits of focused spatial attention were observed (as compared with a control condition with a wide attentional focus); these benefits were well described by a model that asserts different evidence criteria for the two attentional conditions. The lower evidence criterion in the focused attention condition can be interpreted

in two ways: On one hand, it might reflect a lower amount of evidence necessary for stimulus detection; on the other hand it might reflect more effective accumulation of evidence in the focused attention condition. The diffusion superposition model is mute in this respect; neurophysiological evidence, however, suggests the latter interpretation (Hillyard et al., 1973; Talsma et al., 2007).

What effect does attention have on audiovisual integration? While attention seems to be critical for early multisensory eventrelated potential interactions (Talsma et al., 2007), little is known about the effects of spatial attention on behavior, for example, audiovisual redundancy gains. In all conditions of the present experiments, mean response times were well described by a diffusion model based on linear superposition of modality-specific activation in the two channels (Schwarz, 1994). In Experiment 1, attentional modulation involved a change of the size of the attentional focus and the expected results were obtained. Attentionspecific benefits of focused spatial attention were observed (as compared with a control condition with a wide attentional focus); these benefits were well described by a model that asserts different evidence criteria for the two attentional conditions. The lower evidence criterion in the focused attention condition can be interpreted in two ways: On one hand, it might reflect a lower amount of evidence necessary for stimulus detection; on the other hand it might reflect more effective accumulation of evidence in the focused attention condition. The diffusion superposition model is mute in this respect; neurophysiological evidence, however, suggests the latter interpretation (Hillyard et al., 1973; Talsma et al., 2007).

The spatial Go/Nogo task used in Experiment 2 of the present study differs from the discrimination task used in Gondan et al. (2010, Experiment 2). In Gondan et al. (2010), participants received combinations of audiovisual stimuli (both either targets or distractors). A response was required when either of the stimulus components was a target. For some participants, model fit returned seemingly implausible estimates for some parameters (namely, σ_{A}^2 and σ_{V}^2 were

close to zero, suggestive of a deterministic buildup of evidence). It turned out that such a special case of the diffusion superposition model mimics the predictions of a serial self-terminating model of information processing. We argued that these participants might have processed the redundant information serially, as a consequence of response competition induced by combinations of targets in one modality and nontargets in the other modality. In the selective attention task used in the present study, only conflict-free stimulus combinations were used, thereby avoiding response competition effects. The good agreement between model and data (Figure 6.2, Table 6.2) demonstrates that the superposition model can actually describe behavior in conflict-free audiovisual redundant signals experiments, even for the more demanding Go/Nogo task.

The two experiments, thus, show that the superposition model (Schwarz, 1989, 1994) can explain redundancy gains under different attentional conditions. Spatial attention, in our experimental setup, could be fully described by a shift of the evidence barrier (which we think is related to more efficient processing in the two sensory channels; Hillyard et al., 1973). Manipulations of selective attention affect the two modalities differentially, but audiovisual integration still follows the principle of linear, additive superposition of modality-specific activation (Stanford et al., 2005; Ma et al., 2006). The present study focuses on basic mechanisms of multisensory integration observable in rather simple experimental tasks. The stimuli used in the present study are, thus, rather abstract and somehow artificial, and we have chosen white noise and Gabor patches mainly because these stimuli are known to be effective in their respective modality (e.g., Watson, Barlow, & Robson, 1983). There is a growing number of studies using the redundant signals paradigm (e.g., about 400 citations of Miller, 1982, in Google Scholar in February, 2011), most of these studies limit their analysis to the test of the race model inequality. If the race model fails, separate activation is ruled out (Miller, 1982). However, without testing a specific coactivation model, little is known about the

specific mechanisms underlying the integration of the redundant information. The limited number of studies of formal coactivation models (e.g. Diederich, 1995; Miller & Ulrich, 2003; Schwarz, 1989, 1994) mainly describe redundancy gains observed in simple response tasks with beeps and flashes presented from a single source of stimulation (e.g., the setup used by Miller, 1986). We have shown that for the Go/Nogo task (i.e., a slightly more complex task than just simple responses) coactivation effects cannot be taken for granted and linear superposition does not always describe the observed redundancy gains (Gondan et al., 2010). Although the experimental setup is still far from being ecologically valid, the present study sheds light on the basic principles of multisensory processing and the role of attention therein.

Literatur

Alais, D., Morrone, C. C., & Burr, D. (2006). Separate attentional resources for vision and audition. *Proc. R. Soc. B*, *273*, 1339–1345. doi: 10.1098/rspb.2005.3420

Angelaki, D. E., Gu, Y., & DeAngelis, G. C. (2009). Multisensory integration: psychophysics, neurophysiology, and computation. *Current opinion in neurobiology*, *19*, 452–458. doi: 10.1016/j.conb.2009.06.008

Bashinski, H. S., & Bacharach, V. R. (1980). Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. *Perception & Psychophysics*, *28*(3), 241–248. doi: 10.3758/BF03204380

Beer, A. L., Plank, T., & Greenlee, M. W. (2011). Diffusion tensor imaging shows white matter tracts between human auditory and visual cortex. *Experimental Brain Research*, *213*, 299–308.

Bertelson, P., Vroomen, J., de Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics*, *62*, 321–332. doi: 10.3758/BF03205552

Bichot, N. P., Cave, K. R., & Pashler, H. (1999). Visual selection mediated by location: Feature-based selection of noncontiguous locations. *Perception & Psychophysics*, *61*, 403–423. doi: 10.3758/BF03211962

Blurton, S. P., Greenlee, M. W., & Gondan, M. (2014). Multisensory processing of redundant information in go/no-go and choice responses. *Attention, Perception, & Psychophysics*, *76*, 1212–1233.

Blurton, S. P., Kesselmeier, M., & Gondan, M. (2012). Fast and accurate calculations for cumulative first-passage time distributions in wiener diffusion models. *Journal of Mathematical Psychology*, *56*, 470–475. doi: 10.1016/j.jmp.2012.09.002

Broadbent, D. E. (1952). Listening to one of two synchronous messages. *Journal of Experimental Psychology*, *44*, 51–55. doi: 10.1037/h0056491

Brunel, N., & Wang, X.-J. (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *Journal of Computational Neuroscience*, *11*, 63–85.

Buchtel, H. A., & Butter, C. M. (1988). Spatial attentional shifts: Implications for the role of polysensory mechanisms. *Neuropsychologia*, *26*(4), 499–509. doi: 10.1016/0028-3932(88)90107-8

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*, 432–459. doi: 10.1037/0033-295X.100.3.432

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*(3), 308–313. doi: 10.1038/nn1194

Castiello, U., & Umiltà, C. (1990). Size of the attentional focus and efficiency of processing. *Acta Psychologica*, *73*, 195–209. doi: 10.1016/0001-6918(90)90022-8

Castiello, U., & Umiltà, C. (1992). Splitting focal attention. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 837–848. doi: 10.1037/0096-1523.18.3.837

Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin & Review*, *6*, 204–223. doi: 10.3758/BF03212327

Cohen, A., & Feintuch, U. (2002). Attention and Performance XIX: Common mechanisms in rerception and action. In W. Prinz & B. Hommel (Eds.), (pp. 587–608). Oxford: Oxford University Press.

Cohen, A., & Shoup, R. (1997). Perceptual dimensional constraints on response selection processes. *Cognitive Psychology*, *32*, 128–181. doi: 10.1006/cogp.1997.0648

Colonius, H., & Diederich, A. (2006). The race model inequality: interpreting a geometric measure of the amount of violation. *Psychological Review*, *113*(1), 148. doi: 10.1037/0033-295X.113.1.148

Colonius, 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). Erlbaum Mahwah, NJ.

Cox, D. R., & Miller, H. D. (1965). *The theory of stochastic processes.* New York: Wiley.

Diederich, A. (1995). Intersensory facilitation of reaction time: Evaluation of counter and diffusion coactivation models. *Journal of Mathematical Psychology*, *39*, 197–215. doi: 10.1006/jmps.1995.1020

Diederich, A. (1997). Dynamic stochastic models for decision making in an uncertain environment. *Journal of Mathematical Psychology*, *41*, 260–274. doi: 10.1006/jmps.1997.1167

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*(3), 304–322. doi: 10.1016/S0022-2496(03)00003-8

Diederich, A., & Colonius, H. (1987). Intersensory facilitation in the motor component? *Psychological Research*, *49*(1), 23–29. doi: 10.1007/BF00309199

Donders, F. C. (1868/1969). On the speed of mental processes. *Acta Psychologica*, *30*, 412–431. doi: 10.1016/0001-6918(69)90065-1

Dosher, B. A., & Lu, Z.-L. (2000). Noise exclusion in spatial attention. *Psychological Science*, *11*(2), 139–146. doi: 10.1111/1467-9280.00229

Driver, J. (1996). Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature*, *381*, 66–68. doi: 10.1038/381066a0

Driver, J., & Spence, C. (1998). Attention and the crossmodal construction of space. *Trends in Cognitive Sciences*, *2*(7), 254–262. doi: 10.1016/S1364-6613(98)01188-7

Dubois, J., Hamker, F. H., & VanRullen, R. (2009). Attentional selection of noncontiguous locations: The spotlight is only transiently "split". *Journal of Vision*, *9*, 3.1–11. doi: 10.1167/9.5.3

Eimer, M. (1996). ERP modulations indicate the selective processing of visual stimuli as a result of transient and sustained spatial attention. *Psychophysiology*, *33*(1), 13–21. doi: 10.1111/j.1469-8986.1996.tb02104.x

Eimer, M. (2001). Crossmodal links in spatial attention between vision, audition, and touch: evidence from event-related brain potentials. *Neuropsychologia*, *39*(12), 1292–1303. doi: 10.1016/S0028-3932(01)00118-X

Eimer, M., & Driver, J. (2000). An event-related brain potential study of cross-modal links in spatial attention between vision and touch. *Psychophysiology*, *37*, 697–705. doi: 10.1111/1469-8986.3750697

Eimer, M., & Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial attention: evidence from event-related brain potential studies. *Neuroscience & Biobehavioral Reviews*, *25*(6), 497–511. doi: 10.1016/S0149-7634(01)00029-X

Eimer, M., & Schröger, E. (1998). Erp effects of intermodal attention and cross-modal links in spatial attention. *Psychophysiology*, *35*(3), 313–327. doi: 10.1017/S004857729897086X

Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, *322*, 549–560. doi: 10.1002/andp.19053220806

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. doi: 10.3758/BF03208712

Ermolova, N., & Haggman, S. G. (2004). Simplified bounds for the complementary error function; application to the performance of signal processing systems. *Proceedings of the 12th Eurpean signal processing conference*, 1087–1090.

Fairhall, S., & Macaluso, E. (2009). Spatial attention can modulate audiovisual integration at multiple cortical and subcortical sites. *European Journal of Neuroscience*, *29*(6), 1247–1257. doi: 10.1111/j.1460-9568.2009.06688.x

Farah, M. J., Wong, A. B., Monheit, M. A., & Morrow, L. A. (1989). Parietal lobe mechanisms of spatial attention: Modality-specific or supramodal? *Neuropsychologia*, *27*(4), 461–470. doi: 10.1016/0028-3932(89)90051-1

Feintuch, U., & Cohen, A. (2002). Visual attention and coactivation of response decisions for features from different dimensions. *Psychological Science*, *13*, 361–369. doi: 10.1111/j.0956-7976.2002.00465.x

Feller, W. (1968). *An introduction to probability theory and its applications* (3rd ed., Vol. 1). New York: Wiley.

Fürth, R. (1917). Einige Untersuchungen über Brownsche Bewegung an einem Einzelteilchen. *Annalen der Physik*, *358*(11), 177–213.

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. doi: 10.1037/0096-1523.19.6.1278

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General*, *136*(3), 389–413. doi: 10.1037/0096-3445.136.3.389

Gondan, M. (2009). Testing the race model inequality in redundant stimuli with variable onset asynchrony. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 575–579. doi: 10.1037/a0013620

Gondan, M. (2010). A permutation test for the race model inequality. *Behavior Research Methods*, *42*(1), 23–28. doi: 10.3758/BRM.42.1.23

Gondan, M., Blurton, S. P., Hughes, F., & Greenlee, M. W. (2011). Effects of spatial and selective attention on basic multisensory integration. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(6), 1887–

1897. doi: 10.1037/a0025635

Gondan, M., & Fimm, B. (2013, August). Simple correction methods for task completion times contaminated by errors, outliers and omissions. In *Paper presented at the annual meeting of the society of mathematical psychology.* Potsdam, Germany.

Gondan, M., Götze, T., & Greenlee, M. W. (2010). Redundancy gains in simple responses and go/no-go tasks. *Attention, Perception, & Psychophysics*, *72*, 1692–1709. doi: 10.3758/APP.72.6.1692

Gondan, M., & Heckel, A. (2008). Testing the race inequality: A simple correction procedure for fast guesses. *Journal of Mathematical Psychology*, *52*, 323–325. doi: 10.1016/j.jmp.2008.08.002

Grasman, R. P., Wagenmakers, E.-J., & van der Maas, H. L. (2009). On the mean and variance of response times under the diffusion model with an application to parameter estimation. *Journal of Mathematical Psychology*, *53*(2), 55–68. doi: 10.1016/j.1498jmp.2009.01.006

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York, NY: Wiley.

Grice, G. R., & Canham, L. (1990). Redundancy phenomena are affected by response requirements. *Perception & Psychophysics*, *48*(3), 209–213. doi: 10.3758/BF03211520

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. doi: 10.3758/BF03203922

Grice, G. R., & Reed, J. M. (1992). What makes targets redundant? *Perception & Psychophysics*, *51*(5), 437–442. doi: 10.3758/BF03211639

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

Hillyard, S. A., Hink, R., Schwendt, V. L., & Picton, T. W. (1973). Electrical

signs of selective attention in the human brain. *Science*, *182*, 177–179. doi: 10.1126/science.182.4108.177

Horrocks, J. (1999). *Double barrier models for length of stay in hospital*. Unpublished doctoral dissertation, Waterloo, Canada.

Horrocks, J., & Thompson, M. E. (2004). Modeling event times with multiple outcome using the wiener process with drift. *Lifetime Data Analysis*, *10*, 29–49. doi: 10.1023/B:LIDA.0000019254.29153.1a

Houpt, J. W., & Townsend, J. T. (2012). Statistical measures for workload capacity analysis. *Journal of Mathematical Psychology*, *56*(5), 341–355. doi: 10.1016/j.jmp.2012.05.004

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In *Attention and Performance IX* (pp. 187–203). Hillsdale, NJ: Earlbaum Associates.

Kinchla, R. (1974). Detecting target elements in multielement arrays. a confusability model. *Perception & Psychophysics*, *15*, 149–158. doi: 10.3758/BF03205843

Koch, S. H., Weir, C., Westenskow, D., Gondan, M., Agutter, J., Haar, M., ... Staggers, N. (2013). Evaluation of the effect of information integration in displays for icu nurses on situation awareness and task completion time: A prospective randomized controlled study. *International Journal of Medical Informatics*, *82*(8), 665–675.

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychologica*, *134*(3), 372–384.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298. doi: 10.1038/nn.2635

Lansman, M., Farr, S., & Hunt, E. (1984). Expectancy and dual-task interference.

Journal of Experimental Psychology: Human Perception and Performance, *10*(2), 195–204. doi: 10.1037/0096-1523.10.2.195

Liu, T., Fuller, S., & Carrasco, M. (2006). Attention alters the appearance of motion coherence. *Psychonomic Bulletin & Review*, *13*(6), 1091–1096. doi: 10.3758/BF03213931

Luce, R. D. (1986). *Response times. their role in inferring mental organization.* New York, NY: Oxford University Press.

Ma, W. J., Beck, J. M., Lantham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, *9*, 1432–1438. doi: 10.1038/nn1790

Macaluso, E., Frith, C. D., & Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, *289*, 1206–1208. doi: 10.1126/science.289.5482.1206

Maris, G., & Maris, E. (2003). Testing the race model inequality: a nonparametric approach. *Journal of Mathematical Psychology*, *47*(5), 507–514. doi: 10.1016/S0022-2496(03)00062-2

McCormick, P. A., Klein, R. M., & Johnston, S. (1998). Splitting vs. sharing focal attention: Comment on castiello & umiltà. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 350–357. doi: 10.1037/0096-1523.24.1.350

McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., & Hillyard, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *Journal of Cognitive Neuroscience*, *15*(1), 10–19. doi: 10.1162/089892903321107783

McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., & Hillyard, S. A. (2005). Neural basis of auditory-induced shifts in visual time-order perception. *Nature Neuroscience*, *8*(9), 1197–1202. doi: 10.1038/nn1512

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*,

746–748. doi: 10.1038/264746a0

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*, 640–662.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*, 247–279. doi: 10.1016/0010-0285(82)90010-X

Miller, J. (1986). Timcourse of coactivation in bimodal redundant signals. *Perception & Psychophysics*, *40*, 331–343. doi: 10.3758/BF03203025

Miller, J. (1991). Channel interaction and the redundant-targets effect in bimodal divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(1), 160–169. doi: 10.1037/0096-1523.17.1.160

Miller, J., Beutinger, D., & Ulrich, R. (2009). Visuospatial attention and redundancy gain. *Psychological Research*, *73*, 254–262. doi: 10.1007/s00426-008-0212-0

Miller, J., & Ulrich, R. (2003). Simple reaction time and statistical facilitation. a parallel grains model. *Cognitive Psychology*, *46*, 101–151. doi: 10.1016/S0010-0285(02)00517-0

Mordkoff, J. T., & Danek, R. H. (2011). Dividing attention between color and shape revisited: redundant targets coactivate only when parts of the same perceptual object. *Attention, Perception, & Psychophysics*, *73*(1), 103–112. doi: 10.3758/s13414-010-0025-2

Mordkoff, J. T., & Miller, J. (1993). Redundancy gains and coactivation with two different targets: The problem of target preferences and the effects of display frequency. *Perception & Psychophysics*, *53*(5), 527–535.

Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 520–538. doi: 10.1037/0096-1523.17.2.520

Mordkoff, J. T., & Yantis, S. (1993). Dividing attention between color and shape: Evidence of coactivation. *Perception & Psychophysics*, *53*(4), 357–366. doi: 10.3758/BF03206778

Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(2), 315–330.

Navarra, J., Alsius, A., Soto-Faraco, S., & Spence, C. (2010). Assessing the role of attention in the audiovisual integration of speech. *Information Fusion*, *11*, 4–11. doi: 10.1016/j.inffus.2009.04.001

Navarro, D. J., & Fuss, I. G. (2009). Fast and accurate calculations for first-passage times in wiener diffusion models. *Journal of Mathematical Psychology*, *53*, 222–230. doi: 10.1016/j.jmp.2009.02.003

Perea, M., Rosa, E., & Gómez, C. (2002). Is the go/no-go lexical decision task an alternative to the yes/no lexical decision task? *Memory & Cognition*, *30*(1), 34–45. doi: 10.3758/BF03195263

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, *32*, 3–25. doi: 10.1080/00335558008248231

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and Performance X: Control of Language Processes* (pp. 531–556). London: Lawrence Erlbaum Associates Ltd.

Posner, M. I., Inhoff, A. W., Friedrich, F. J., & Cohen, A. (1987). Isolating attentional systems: A cognitive-anatomical analysis. *Psychobiology*, *15*(2), 107–121. doi: 10.3758/BF03333099

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In H. L. Pick (Ed.), *Modes of perceiving and processing information* (pp. 137–157). New York, NY: Wiley.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*, 160–174. doi: 10.1037/0096-3445.109.2.160

R Development Core Team. (2010). *R: A language and environment for statistical computing.* Vienna, Austria. Retrieved from `http://www.R-project.org` (ISBN 3-900051-07-0)

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, *24*(5 Series II), 574–590.

Rach, S., Diederich, A., & Colonius, H. (2011). On quantifying multisensory interaction effects in reaction time and detection rate. *Psychological Research*, *75*(2), 77–94. doi: 10.1007/s00426-010-0289-0

Ratcliff, R. (1978). A theory of memory retrieval. *Psycholgical Review*, *85*, 59–108. doi: 10.1037/0033-295X.85.2.59

Ratcliff, R. (1980). A note on modeling accumulation of information when the rate of accumulation changes over time. *Journal of Mathematical Psychology*, *21*(2), 178–184.

Ratcliff, R., Cherian, A., & Segraves, M. (2003). A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of simple two-choice decisions. *Journal of Neurophysiology*, *90*, 1392–1407. doi: 10.1152/jn.01049.2002

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922. doi: 10.1162/neco.2008.12-06-420

Ratcliff, R., Philiastides, M. G., & Sajda, P. (2009). Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proceedings of the National Academy of Sciences*, *106*(16), 6539–6544. doi: 10.1073/pnas.0812589106

Ratcliff, R., & Rouder, J. N. (2000). A diffusion model account of masking in

two-choice letter identification. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(1), 127–140. doi: 10.1037/0096-1523.26.1.127

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*, 438–481. doi: 10.3758/BF03196302

Schrödinger, E. (1915). Zur Theorie der Fall- und Steigversuche an Teilchen mit Brownscher Bewegung. *Physikalische Zeitschrift*, *16*, 289–295.

Schröter, H., Ulrich, R., & Miller, J. (2007). Effects of redundant auditory stimuli on reaction time. *Psychonomic Bulletin & Review*, *14*(1), 39–44. doi: 10.3758/BF03194025

Schwarz, W. (1989). A new model to explain the redundant signals effect. *Perception & Psychophysics*, *46*, 498–500. doi: 10.3758/BF03210867

Schwarz, W. (1994). Diffusion, superposition, and the redundant-targets effect. *Journal of Mathematical Psychology*, *38*, 504–520. doi: 10.1006/jmps.1994.1036

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*(3), 558–573. doi: 10.1037/0096-1523.32.3.558

Schwarz, W., & Ischebeck, A. (2003). On the relative speed account of number-size interference in comparative judgments of numerals. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(3), 507-522. doi: 10.1037/0096-1523.29.3.507

Senkowski, D., Talsma, D., Grigutsch, M., Herrmann, C. S., & Woldorff, M. G. (2007). Good times for multisensory integration: effects of the precision of temporal synchrony as revealed by gamma-band oscillations. *Neuropsycho-*

logia, *45*(3), 561–571. doi: 10.1016/j.neuropsychologia.2006.01.013

Smith, P. L. (2000). Stochastic dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology*, *44*, 408–463. doi: 10.1006/jmps.1999.1260

Spence, C., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(3), 555–574. doi: 10.1037/0096-1523.20.3 .555

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 1005–1030. doi: 10.1037/0096-1523.22.4.1005

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, *59*(1), 1–22. doi: 10.3758/BF03206843

Spence, C., McDonald, J., & Driver, J. (2004). Exogenous spatial cuing studies of human crossmodal attention and multisensory integration. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 277–320). New York, NY: Oxford University Press.

Stanford, T. R., Quessy, S., & Stein, B. E. (2005). Evaluating the operations underlying multisensory integration in the cat superior colliculus. *The Journal of Neuroscience*, *25*(28), 6499–6508. doi: 10.1523/JNEUROSCI.5095-04.2005

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, *26*(2), 212–215. doi: 10.1121/1.1907309

Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cerebral Cortex*, *17*(3), 679–690. doi: 10.1093/cercor/bhk016

Talsma, D., & Kok, A. (2002). Intermodal spatial attention differs between vision and audition: An event-related potential analysis. *Psychophysiology*, *39*(6),

689–706. doi: 10.1111/1469-8986.3960689

Talsma, D., Kok, A., Slagter, H. A., & Cipriani, G. (2008). Attentional orienting across the sensory modalities. *Brain and Cognition*, *66*(1), 1–10. doi: 10 .1016/j.bandc.2007.04.005

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, *14*(9), 400–410. doi: 10.1016/j.tics.2010.06.008

Talsma, D., & Woldorff, M. G. (2005). Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *Journal of Cognitive Neuroscience*, *17*(7), 1098–1114. doi: 10.1162/0898929054475172

Teder-Sälejärvi, W., McDonald, J., Di Russo, F., & Hillyard, S. (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (erp) recordings. *Cognitive Brain Research*, *14*(1), 106–114. doi: 10.1016/S0926-6410(02)00065-4

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics*, *49*(1), 83–90. doi: 10.3758/BF03211619

Townsend, J. T., & Altieri, N. (2012). An accuracy–response time capacity assessment function that measures performance against standard parallel predictions. *Psychological Review*, *119*(3), 500–516. doi: 10.1037/a0028448

Townsend, J. T., & Honey, C. J. (2007). Consequences of base time for redundant signals experiments. *Journal of Mathematical Psychology*, *51*(4), 242–265. doi: 10.1016/j.jmp.2007.01.001

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*(4), 321–359. doi: 10.1006/jmps.1995.1033

Treisman, A. M. (1986). Properties, parts, and objects. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of Perception and Human Performance*

(pp. 35/1–35/70). John Wiley & Sons.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136. doi: 10.1016/0010-0285(80)90005-5

Uhlenbeck, G. E., & Ornstein, L. S. (1930). On the theory of the Brownian motion. *Physical Review*, *36*(5), 823. doi: 10.1103/PhysRev.36.823

Ulrich, R., & Giray, M. (1986). Separate-activation models with variable base times: Testability and checking of cross-channel dependency. *Perception & Psychophysics*, *39*(4), 248–254. doi: 10.3758/BF03204931

Ulrich, R., & Miller, J. (1997). Tests of race models for reaction time in experiments with asynchronous redundant signals. *Journal of Mathematical Psychology*, *41*(4), 367–381. doi: 10.1006/jmps.1997.1181

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological Review*, *108*(3), 550–592. doi: 10.1037/0033-295X.108.3.550

Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review*, *7*, 208–256.

Van der Burg, E., Olivers, C. N., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1053–1065. doi: 10.1037/0096-1523.34.5.1053

von Smoluchowski, M. (1915). Notiz über die Berechnung der Brownschen Molekulargewegung bei der Ehrenhaft-Millikanschen Versuchsanordnung. *Physikalische Zeitschrift*, *17/18*, 318–321.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*(7), 1206–1220. doi: 10.3758/BF03196893

Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of

diffusion model parameters. *Journal of Mathematical Psychology*, *52*(1), 1–9. doi: 10.1016/j.jmp.2007.09.005

Vroomen, J., Bertelson, P., & de Gelder, B. (2001). The ventriloquist effect does not depend on the direction of automatic visual attention. *Perception & Psychophysics*, *63*(4), 651–659.

Wagenmakers, E.-J., van der Maas, H. L., & Grasman, R. P. (2007). An ez-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*(1), 3–22. doi: 10.3758/BF03194023

Wald, A. (1947). *Sequential analysis*. New York: Wiley.

Ward, L. M. (1994). Supramodal and modality-specific mechanisms for stimulus-driven shifts of auditory and visual attention. *Canadian Journal of Experimental Psychology*, *48*(2), 242–259.

Ward, L. M., McDonald, J. J., & Lin, D. (2000). On asymmetries in cross-modal spatial attention orienting. *Perception & Psychophysics*, *62*(6), 1258–1264. doi: 10.3758/BF03212127

Watson, A. B., Barlow, H., & Robson, J. G. (1983). What does the eye see best? *Nature*, *302*(5907), 419–422. doi: 10.1038/302419a0

Welch, R. B., & Warren, D. H. (1986). Intersensory interactions. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of Perception and Human Performance* (Vol. 25, pp. 1–36). New York: Wiley.

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences*, *90*(18), 8722–8726.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 601–621.