

EBERRHARD-KARLS-UNIVERSITÄT TÜBINGEN
Naturwissenschaftliche Fakultät

Institut für Neurobiologie
Lehrstuhl für Kognitive Neurowissenschaften

Bachelorarbeit

**Multisensorische Integration im Arbeitsgedächtnis
- Überprüfung abstrakt kongruenter Stimuli**

Marc Fischer

April 2019

Gutachter: Dr. rer. nat. Gregor Hardiess
Institut für Neurobiologie
Kognitive Neurowissenschaften

Universität Tübingen
Auf der Morgenstelle 28
72076 Tübingen

Anerkennung

An erster Stelle möchte ich mich bei meinem Betreuer Dr. rer. nat. Gregor Hardiess für all die Zeit, Geduld und Unterstützung während der Arbeit an meiner Bachelorthesis bedanken. Besonders für die vielen Ratschläge und nützlichen Verweise, welche mir viel Zeit und Mühe erspart haben und ohne die die Arbeit jetzt nicht in ihrer Form vorliegen würde.

Ebenso möchte ich mich bei der gesamten Arbeitsgruppe der Kognitiven Neurowissenschaft der Eberhard-Karls-Universität Tübingen bedanken. Ich habe mich jederzeit willkommen und gut aufgehoben gefühlt und in meiner Zeit hier viele interessante und hilfreiche Anregungen bekommen. Speziell hervorheben möchte ich dabei die Unterstützung von Prof. Dr. Hanspeter A. Mallot sowie Niclas Renner, die mir in mehreren Treffen die Thematik mit mir besprochen haben und mich auf etwaige Probleme hingewiesen haben.

Auch bei allen Teilnehmern meiner Untersuchung, welche sich aus Freunden und Bekannten zusammensetzt, möchte ich mich recht herzlich bedanken. Für diesen nicht selbstverständlichen Aufwand wurde keine Entlohnung vergeben. Ohne ihre Hilfe wäre die gesamte Untersuchung nicht so reibungslos vonstattengegangen.

Zu guter Letzt möchte ich mich auch bei Maja Römer, Paul Jens und Lara Reichart für das Korrekturlesen und sehr hilfreiche Anregungen bezüglich der Verständlichkeit und Lesbarkeit der Arbeit bedanken.

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich diese schriftliche Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe und alle wörtlich oder sinngemäß aus anderen Werken übernommenen Aussagen als solche gekennzeichnet habe.

Datum, Ort, Unterschrift

Abstract

Die Verbindung zwischen dem Menschen und seiner Umwelt beruht weitgehend auf den von den verschiedenen Sinnen wahrgenommenen Informationen. Ohne diese wäre es nicht möglich, gerichtet mit der Umgebung zu interagieren oder sich als Individuum in dieser einzuordnen. Dementsprechend ist der Mensch sowie die meisten Lebewesen mit mehreren Sinnesorganen ausgestattet, welche jeweils auf das Aufnehmen unterschiedlicher Informationen spezialisiert sind. Die Frage nach der genauen Art und dem Zeitpunkt des Zusammensetzens der unterschiedlich aufgenommenen Informationen ist dabei von besonderem Interesse. Demzufolge ist das Thema der multisensorischen Integration bereits in mehreren Arbeiten untersucht worden. Dies geschieht zumeist in dem Vergleich multisensorischer Reize mit ihren Komponenten. So zeigte sich, dass die jeweiligen in Untersuchungen verwendeten Stimuli einen großen Einfluss auf die erzielten Ergebnisse haben können. Die folgende Arbeit befasst sich in Anlehnung an bereits durchgeführte Untersuchungen zum Thema der multisensorischen Integration mit dem Finden geeigneter audiovisueller Stimuli und einer Überprüfung ihrer Eignung für Untersuchungen des Arbeitsgedächtnisses sowie eines wahrnehmungstechnischen Vorteils. Die dabei schlussendlich entworfenen und experimentell auf ihre Eignung überprüften Stimuli beruhen auf einer, aus drei Tonanschlägen basierenden, unterschiedlichen Rhythmik. Diese ist visuell durch einen zeitlich kongruent entstehenden schwarzen Kreis dargestellt. Um die Signalstärke zwischen den Bedingungen in unterschiedlichen Modalitäten möglichst vergleichbar zu halten, wurden neben den unisensorischen und der aus diesen zusammengesetzten multisensorischen Darstellungsweisen zusätzliche Bedingungen eingeführt, in welchen die unimodale Darstellung zusätzlich mit einem Rauschen in der Nichtzielmodalität versehen wurde. Die Ergebnisse einer two-back-Aufgabe zeigten mit einer deutlich über dem Rateniveau liegenden Leistung eine Eignung der Stimuli für eine Untersuchung multisensorischer Integration im Arbeitsgedächtnis. Des Weiteren wurde deutlich, dass speziell das hinzugefügte auditive Rauschen eine negative Auswirkung auf die Leistung der Teilnehmer hatte. Es konnten in der dabei durchgeführten Aufgabe keine der bereits in anderen wissenschaftlichen Arbeiten zur multisensorischen Integration aufgetretenen Befunde repliziert werden.

Inhaltsverzeichnis

Anerkennung	0
Eigenständigkeitserklärung	2
Abstract	3
Abbildungsverzeichnis	5
1. Einleitung	6
1.1 Multisensorische Integration	6
1.2 Das Arbeitsgedächtnis	9
1.3 Multisensorische Integration und das Arbeitsgedächtnis	11
2. Stimulus Wahl	14
2.1 Anforderungen.....	14
2.2 Erster Ansatz: zufällig generierte Tonfolgen.....	15
2.3 Zweiter Ansatz: gebrochene Dreiklänge	16
2.4 Stimulus: Rhythmische Tonanschläge	17
3. Methode.....	21
3.1 Versuchsteilnehmer	21
3.2 Materielles Setup.....	21
3.3 Versuchsaufbau	21
3.4 Versuchsablauf	23
3.5 Auswertung	24
4. Resultate	Fehler! Textmarke nicht definiert.
4.1 Sensitivitätsindex d'	25
4.2 Sensitivitätsindex Konfidenzintervalle.....	26
4.3 Varianzanalyse	27
4.4 Reaktionszeitanalyse	28
4.5 Fragebogenauswertung.....	29
5. Diskussion	30
5.1 Eignung der verwendeten Stimuli	30
5.2 Vergleich der Modalitäten.....	31
5.3 weiteres Vorgehen.....	32
5.4 Konklusion	33
Literaturverzeichnis.....	35
Appendix	37
Fragebogen.....	37

Abbildungsverzeichnis

Abbildung 1: Unterschiedliche neuronale Integrationsformen. Abbildung aus Stein und Stanford(2008)	8
Abbildung 2: Mehrkomponentenmodell. Aus Zimmer (2008).....	10
Abbildung 3: Performance in der bimodalen Kategorisierungsaufgabe. Abbildung aus Niclas Renner (2018)	13
Abbildung 4: Sensitivitätsindex Leistung in Arbeitsgedächtnisaufgabe mit eingezeichnetem Standardfehler(SEM) Abbildung von Niclas Renner (2018)	13
Abbildung 5: psychometrische Funktionen. Abbildung aus Niclas Renner (2018)	14
Abbildung 6: Stimulus 1-3-5 in Wellenform.....	18
Abbildung 7: Stimulus 1-2-5 in Wellenform.....	18
Abbildung 8: Stimulus 1-4-5 in Wellenform.....	18
Abbildung 9: Stimulus 2-3-4 in Wellenform.....	19
Abbildung 10: Farbwerte der visuellen Darstellung.....	20
Abbildung 11: Rauschen in der auditiven Modalität.....	22
Abbildung 12: Ablauf eines Versuchsdurchgangs.	23
Abbildung 13: Boxplot der Sensitivitätsindex Werte.	25
Abbildung 14: 95% Konfidenzintervalle der Sensitivitätsindex Werte für.....	26
Abbildung 15: Vergleich der Mittelwerte zwischen der unisensorisch visuellen Bedingung und der auditiven mit visuellem Rauschen.....	27
Abbildung 16: Vergleich der Mittelwerte zwischen der audiovisuellen Bedingung und der visuellem Bedingung, welche von auditivem Rauschen begleitet wird.	28
Abbildung 17: Vergleich der Mittelwerte der visuellen Präsentation mit auditivem Rauschen mit der bimodalen– und auditiven Bedingung, wobei zweitens ebenfalls von einem Rauschen in der anderen unisensorischen Modalität begleitet wird.	28
Abbildung 18: Boxplot der Reaktionszeiten.....	28

1. Einleitung

Im alltäglichen Leben des Menschen kommt es fortlaufend zu Interaktionen mit der Umwelt sowie der Einordnung des eigenen Seins in jene. Beides ist nur aufgrund eines beständigen Stroms an zuvor aufgenommenen Informationen über den aktuellen Zustand der Umgebung möglich. Die Information ist hierbei eine Voraussetzung für das gerichtete Interagieren mit der Umwelt und bewirkt ebenfalls ein bewusstes Wahrnehmen ihrer aktuellen Zustände im Umfeld des Individuums. Somit ist die Aufnahme und Verarbeitung von Umgebungsinformationen eine grundlegende Notwendigkeit, um viele evolutionär wichtige Funktionen wie das Erkennen von Gefahr oder das Beschaffen von Versorgung und Schutz zu gewährleisten. Dementsprechend besitzt der menschliche Körper eine Vielzahl an Sensoren - die unterschiedlichen Sinnesorgane -, von welchen jedes mit mehreren spezialisierten Hirnregionen verknüpft ist, um die gewonnenen Umwelteindrücke wahrzunehmen und zu interpretieren. Hierbei gibt jedes einzelne Sinnesorgan aufgrund der Spezialisierung ein unterschiedlich komplexes und informatives Abbild seiner Umgebung wieder, dessen Gewichtung je nach Lebenssituation variiert. Dabei wird erst durch die Menge an kombinierten Eindrücken der einzelnen Sinne – im Vergleich zu den einzelnen sinnesspezifischen Informationen das bestmögliche Verständnis der eigenen Umgebung ermöglicht. Diese Interaktion der unterschiedlichen Sinne, welche mit einer Verbindung und Erhöhung des Informationsgehalts einhergeht, wird als multisensorische Integration verstanden. Dabei ist unter anderem die Fähigkeit von großer Bedeutung, redundante Information, welche in mehreren unterschiedlichen Modalitäten vorliegt, aber von der gleichen Referenz verursacht wird (zum Beispiel ein bellender Hund), in einem multisensorischen Event festzuhalten und ihrer räumlichen Quelle zuzuordnen.

1.1 Multisensorische Integration

Bei Untersuchungen zum Thema der multisensorischen Integration wird zumeist eine intermodale Stimulus Kombination mit ihren einzelnen Komponenten im Hinblick auf ihren Einfluss auf eine messbare Verhaltensänderung oder einer neuronalen Erregung verglichen.

So ließ sich zum Beispiel bereits vor über 50 Jahren ein Effekt der Stimulus-Synchronizität von Licht und Ton auf die Reaktionszeit von Menschen feststellen (Hershenson, 1962). In einem einfachen experimentellen Setup, in welchem die Versuchspersonen möglichst zeitnah auf die Präsentation eines Lichts, eines Tons oder einer bimodalen Darbietung der beiden, mit unterschiedlichen Abspielzeitpunkten der einzelnen Reize, reagieren sollten, zeigte sich, dass für nur kleine Asynchronitäten die Reaktionszeit bei gleichzeitiger Darstellung beider Modi signifikant kürzer als die der einzelnen Reize war. Diese kürzere Reaktionszeit in der bimodalen Bedingung, der sogenannte redundancy effect, ließ sich seitdem in vielen Untersuchungen zu dem Unterschied in der Verarbeitung zwischen multisensorischen und unisensorischen Reizen wiederfinden (Juan et al., 2017; Schröter, Frei, Ulrich, Miller, 2009). Dies zeigt auf eine sehr anschauliche Art und Weise, dass multisensorische Integration in der Lage ist die sensorische Verarbeitungsgeschwindigkeit zu erhöhen. Bei multimodal dargestellten Stimuli wird die für eine Reaktion benötigte Erregungsgrenze früher erreicht, was zu einer kürzeren Verzögerung im Vergleich zu unimodalen Stimuli führt. Zur Erklärung hierfür liegen, wie von Gondan und Minakata (2016) in ihrer Arbeit beschrieben, zwei unterschiedliche Modelansätze vor.

Dem Coactivation-Modell folgend, beruht die verbesserte Reaktionszeit bei einem, gleichzeitig in unterschiedlichen Modalitäten, gezeigten Stimulus auf einer Integration in der Verarbeitung der einzelnen Reize. Dem Race-Modell folgend, lässt sich die kürzere Verarbeitungszeit auch mit zwei unterschiedlichen Verarbeitungskanälen gemäß der beteiligten Sinnesmodalitäten erklären. Da in diesem Fall lediglich der schnellere von beiden teilhabenden Kanälen die entsprechende Reaktion auslöst, führt dieser Mechanismus aber im Durchschnitt auch zu einer kürzeren Reaktion bei in beiden Modalitäten vorliegender Information. Somit lässt sich eine kürzere Reaktionszeit nicht unbedingt mit einer Integration der einzelnen Sinneswahrnehmungen erklären. Vielmehr ist es notwendig zu zeigen, dass diese kürzere Reaktionszeit nicht einfach durch die Summe der Wahrscheinlichkeiten einer ausreichenden Aktivierung durch einer der unabhängig voneinander beteiligten Kanäle erklärbar ist (Treisman, 1998; Gondan und Minakata, 2016).

Ein weiteres in Untersuchungen zu der Thematik aufgetretenes Phänomen ist der positive Effekt semantischer Kongruenz dargestellter multisensorischer Stimuli auf Identifikationsaufgaben. So können zum Beispiel verrauschte Videos in einer audiovisuell kongruenten Bedingung wesentlich schneller identifiziert werden (Cox und Hong, 2015). Ebenso ist auch die Identifizierung natürlicher visuell dargestellter Stimuli schneller, wenn diesen ein kongruenter auditiver Prime vorangeht (Schneider und Engel, 2008).

Um mehr Klarheit über die Prozesse der multisensorischen Integration zu bekommen, werden auch die unterschiedlichen Erregungsmuster einzelner multisensorisch aktiver Neuronen betrachtet. Diese werden hierbei nach dem Prinzip der inversen Effektivität von unterschiedlichen Reizkombinationen und unterschiedlicher Reizqualität beeinflusst, wie unter anderem von Stein und Stanford in ihrer Review (2008) aufgezeigt wird. Ist die multimodal gegebene redundante Information bereits durch einen der beteiligten Sinne für ein klares Identifizieren und Einordnen der Umwelt ausreichend, so verschafft die Integration der redundanten Information keinen wahrnehmungstechnischen Vorteil. Sie bildet lediglich einen zusätzlichen Berechnungsaufwand.

Ist die von den Sinneseindrücken aufgenommene Information hingegen unvollständig oder in ihrer Qualität für ein schnelles Erkennen und Einordnen der Umwelt nicht ausreichend, liegt in der Integration beider Informationsquellen das Potential einer deutlichen Steigerung der Wahrnehmungsleistung. Ein einfaches Beispiel für eine solche Situation ist das Erkennen eines um Hilfe suchenden Menschen im Meer bei Nebel und Rauschen der Wellen. Auch wenn das Ziel in einer Modalität nicht sauber zu erkennen ist, ermöglicht multisensorische Integration in diesem Fall das Erkennen dieser Person durch das Zusammensetzen der jeweiligen Information der Sinnesorgane.

Diesem Prinzip folgend, müsste also die schwächste Stimulus Kombination zu der größten Erhöhung der Ausgangsaktivitäten innerhalb eines multisensorischen Neurons führen. Dementsprechend ordneten auch Stein und Stanford in ihrer Arbeit die Integrationsmuster der multisensorischen Integration nach der Qualität der Eingangsreize. Liegt die Information lediglich in einem schwachen Eingangssignal vor, ist das nach der Interaktion entstandene Ausgangssignal größer als die Summe seiner Eingangssignale. Dies wird von ihnen als Superadditivität bezeichnet. Ist die Qualität der Eingangsreize hingegen sehr groß, ist das aus ihnen entstandene Ausgangssignal kleiner als die Summe seiner einzelnen integrierten Bestandteile (Subadditivität). Liegen die Eingangssignale hingegen in einer mittleren Stärke vor, so kommt es zu einem additiven Integrationsmuster. Das Ausgangssignal ist also so stark wie die Summe der einzelnen Eingangssignale (siehe Abbildung 1).

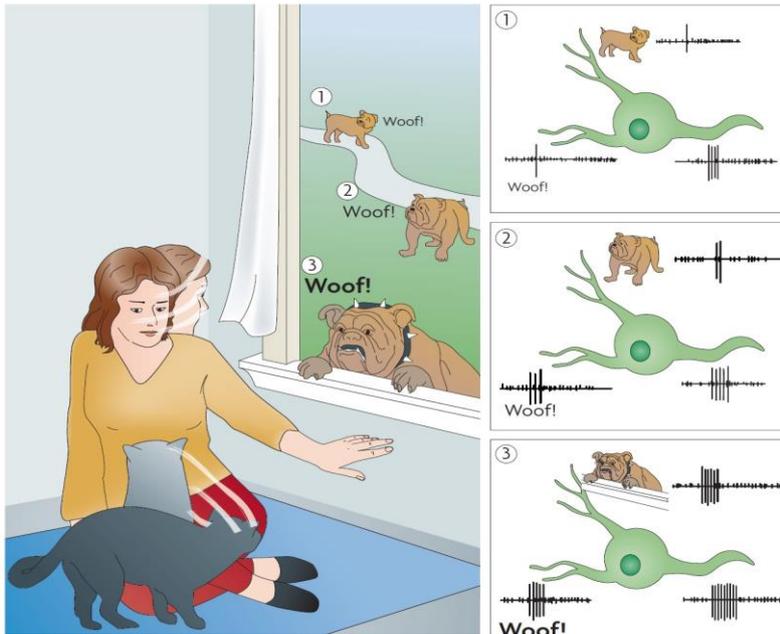


Abbildung 1: Unterschiedliche neuronale Integrationsformen abhängig vom Eingangssignal. 1.) Superadditivität bei schwachem Eingangssignal. 2.) Additivität bei durchschnittlich starkem Eingangssignal. 3.) Subadditivität bei starkem Eingangssignal. Abbildung aus Stein und Stanford(2008)

Ein ähnliches Muster zeigte sich auch in den Untersuchungen von Werner und Noppeney (2009). Sie führten eine Detektions- und Kategorisierungsaufgabe mit auditiven, visuellen und in beiden Modalitäten dargestellten Musikinstrumenten sowie Werkzeugen durch. Hierbei wurde zusätzlich der Grad an Informationen, den die Stimuli enthalten, gemäß dem Prinzip der inversen Effektivität variiert. Die Stimuli waren entweder intakt, leicht bis zur Detektionsschwelle degradiert oder vollständig von Rauschen überlagert, sodass ein Erkennen des Stimulus nicht möglich ist.

Sie beobachteten bei intakten Stimuli subadditive Interaktionsformen im STS (Sulcus temporalis superior), bei welchen die Reaktion auf bimodal dargestellte Stimuli deutlich schwächer ausfiel, als die Reaktion auf den stärksten unimodal präsentierten Stimulus. Bei Darstellung der degradierten Stimuli ließen sich hingegen additive Interaktionsformen feststellen.

Dies zeigte sich auch in einer besseren Leistung in der Kategorisierungsaufgabe bei degradierten, bimodalen Stimuli gegenüber der besten unimodalen Bedingung. Eine weitere Auffälligkeit in ihren Ergebnissen bildet hierbei der Umstand, dass bei vielen Teilnehmern der bimodale Vorteil ausreichend groß ist, um über der mithilfe der aufaddierten Wahrscheinlichkeiten berechneten Obergrenze unabhängiger Verarbeitung zu liegen. Somit liegt ein klares Indiz für eine tatsächliche multisensorische Integration der Information vor.

Bei Untersuchungen zur multisensorischen Integration mithilfe von bildgebenden Verfahren zeigten sich des Weiteren Einflüsse der Integration auf die Aktivität von ursprünglich als unisensorisch betrachteten Domänen. So konnten Giard und Peronnet (1999) in ihrer Arbeit einen Nachweis für eine audiovisuelle Interaktion im menschlichen Kortex bereits 50 Millisekunden nach der Stimulus Präsentation erbringen. Auch Foxe et al. (2002) konnten in ihrer Arbeit Integrationseffekte zwischen auditiven und haptischen Stimuli in einer zuvor unisensorisch betrachteten Region des auditiven Kortex nachweisen. Hierbei ist die entsprechende Hirnregion stärker aktiviert, als sich durch die Summation der ausgelösten taktilen und auditiven Erregung schätzen ließe.

In ihrer Arbeit zeigen Wahn und König (2017) zudem auf, dass gemäß dem Charakter einer Aufgabe sowie der beteiligten Modalitäten die aufmerksamkeitsbedingte Ressourcenverteilung zwischen den einzelnen Modalitäten entweder geteilt oder unabhängig voneinander arbeitet. Werden diese unabhängig voneinander rekrutiert, besitzt somit jede Modalität ihr eigene Aufmerksamkeitskapazität. Bei objektorientierten Aufgaben werden die benötigten attentionalen Ressourcen in der visuellen und auditiven Modalität getrennt voneinander rekrutiert, während sie zwischen haptischen und visuellen Sinnen teilweise geteilt eingezogen werden. Bei Aufgaben mit einem räumlichen Bezug oder bei Aufgaben mit einem räumlichen und objektorientierten Bezug zeigt sich ebenso, dass die Modalitäten teilweise ihre Ressourcen teilen.

Da die Kapazität menschlicher Aufmerksamkeit mit der Funktion des Arbeitsgedächtnisses in engem Zusammenhang stehen (Oliver, 2008), werfen die obigen Befunde die Frage auf, in welcher Form die redundant multimodal vorliegende Information in diesem gespeichert vorliegt. Diese könnte entweder separat in der jeweiligen Modalität vorliegen oder bereits in einem frühen Verarbeitungsschritt zu einer gemeinsamen Repräsentation zusammengefügt werden.

1.2 Das Arbeitsgedächtnis

Das zumeist herangezogene Arbeitsgedächtnis-Modell ist das von Baddeley und Hitch 1974 eingeführte Mehrkomponentenmodell. Dieses beruht auf der Beobachtung getrennter Ressourcen zwischen der visuellen und auditiven Modalität und bestand anfangs aus drei Komponenten. Die phonologische Schleife speichert sprachliche Informationen in Lautsprache und besteht aus den Bestandteilen des passiven phonologischen Speichers, welcher aufgenommene Informationen für etwa zwei Sekunden speichern kann, sowie dem artikulatorischen Kontrollprozess, der eine aktivere Rolle bei der Sprachproduktion spielt. Der räumlich-visuelle Notizblock auf der anderen Seite speichert vorübergehend visuelle und räumliche Informationen, wobei die beiden Subtypen in diesem nicht stark in ihrem Verarbeitungstyp interferieren, sodass eine Störung der räumlichen Wahrnehmung durch visuelle Aufgaben kaum beobachtet wird. Dies zeigt sich unter anderem in einer Aktivierung unterschiedlicher Hirnregionen je nach Art der gestellten Aufgabe (Smith et al., 1995). Die dritte Komponente des 1974 entwickelten Modells bildet die zentrale Exekutive. Ihre wichtigsten Funktionen bestehen aus der Verbindung des Arbeits- mit dem Langzeitgedächtnis sowie der Modulation von aufmerksamkeitsbasierten Prozessen.

Um unter anderem die im Laufe der Zeit aufkommenden Belege für frühe multimodale Interaktionen ebenfalls erklären zu können, wurde dem Modell noch ein zusätzliches Subsystem hinzugefügt. (Baddeley, 2000) Der episodische Buffer bildet hierbei ein multimodales Speichersystem, welches dazu in der Lage ist Informationen aus beiden anderen Subsystemen zu kombinieren und zeitlich einander zuzuordnen. Siehe Abbildung 2 für eine Übersicht des Mehrkomponentenmodells.

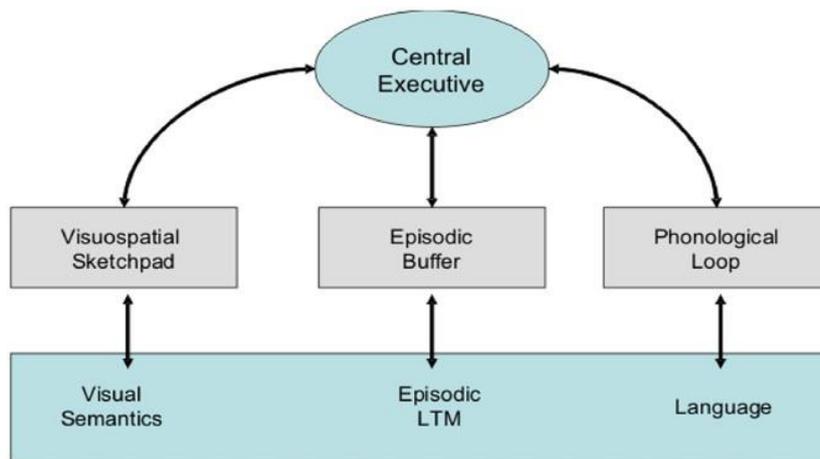


Abbildung 2: Mehrkomponentenmodell. Aus Zimmer (2008)

Die Kapazität des Arbeitsgedächtnisses beträgt hierbei insgesamt ungefähr vier Objekte, unabhängig von der Anzahl ihrer Individuellen Merkmalen (Luck & Vogel, 1997), wobei die individuellen Unterschiede zwischen einer Kapazität von 1,5 bis 5 Objekten liegen können (Vogel & Machizawa, 2004).

Dem Arbeitsgedächtnismodell von Baddeley und Hitch folgend, liegen die modalitätsspezifischen Informationen zunächst immer in ihrer entsprechenden Subdomäne vor. Eine Integration multisensorischer Stimuli kann erst in einem späteren Prozess mithilfe des episodischen Buffers vollführt werden.

1.3 Multisensorische Integration und das Arbeitsgedächtnis

Auch wenn das Arbeitsgedächtnis historisch gesehen hauptsächlich unisensorisch aufgefasst wurde, führten aktuellere Befunde über die letzten 20 Jahre zu einer genaueren Untersuchung des Zusammenhangs zwischen der Verarbeitung multisensorisch dargestellter Stimuli und dem Arbeitsgedächtnis. So konnten unter anderem Thompson und Paivio (1994) bereits eine Verbesserung der free-recall-Leistung bei audiovisuell präsentierten Reizen gegenüber ihren einzelnen Komponenten nachweisen. Sie konnten durch die Untersuchung unimodaler Bild-Bild- und Ton-Ton-Kombinationen zusätzlich auch das lediglich redundante Vorliegen der Information als Ursache ausschließen. Dieses Phänomen wurde von Goolkasian und Foos (2005) weiter untersucht. Auch ihre Ergebnisse zeigten höhere free-recall-Leistungen für audiovisuell dargestellte Stimuli bestehend aus Wörtern oder Bildern sowie dem gesprochenen zugehörigen Wort. Sie schlossen ebenfalls, dass die verbesserte Gedächtnisleistung auf das Vorhandensein der Information in zwei unterschiedlichen Modalitäten zurückzuführen ist, nicht aber auf die Redundanz der Information selbst. Diesen Ergebnissen schließt sich auch die Arbeit von Delogu et al (2009) an. Sie legten den Fokus auf den Vergleich nonverbaler und verbaler Bedingungen bezüglich ihres recall-Verhaltens mit audiovisuellen, visuellen und auditiv vorliegenden Stimuli. Durch das Verhindern der Artikulation von Versuchspersonen während der Aufgabe verschlechterte sich ihre Erinnerungsleistung sowohl für verbale als auch nonverbale Stimulus Präsentationen. Daraus schlussfolgerten sie, dass unabhängig von der Präsentationsform jeder Stimulus in einer verbalen Form kodiert vorliegt. Die Form, in welcher eine Information aufgenommen wird, gibt also nur bedingt Aufschluss darüber, in welcher Repräsentation diese im Arbeitsgedächtnis gespeichert wird. Hierbei ist zu beachten, dass eine solche Transformation der Repräsentation immer auf den semantischen Zusammenhang des Langzeitgedächtnisses angewiesen ist.

Es scheint also ein Leistungsunterschied in der Erinnerung zwischen multimodal kongruenten und unimodalen Reizen zu bestehen. So ließ sich unter anderem ebenfalls nach Sauls and Cowan (2007) eine größere Arbeitsgedächtniskapazität für audiovisuelle Stimuli gegenüber modalitätsspezifischen Stimuli nachweisen. Auch wenn ähnliche Resultate von Fougny und Marois (2011) in ihrer kritischen Auseinandersetzung mit der Arbeit von Sauls and Cowan erzielt wurden, existiert zurzeit keine allgemein akzeptierte Erklärung für diese Beobachtung. Es ist noch unklar, ob die Ursache der Verbesserung der Gedächtnisleistung in einer Integration der Information beider Modalitäten zu einer amodalen allgemeinen Repräsentation begründet liegt, oder ob einfach die voneinander unabhängige Speicherung additiv und ohne Interferenzen zu einer Verbesserung der Leistung führt.

Auch zuvor beobachtete Verhaltensvorteile der multisensorischen Integration lassen sich teilweise in Arbeitsgedächtnisaufgaben wiederfinden. So konnten Brunetti, Indraccolo, Mastroberardino, Spence, und Santangelo (2017) einen bimodalen Reaktionszeitvorteil in einer two-back-Aufgabe nachweisen. Die Teilnehmer sind hierbei dazu angewiesen immer zu überprüfen, ob der gegebene Stimulus aus einer Reihe an Stimulus Präsentationen mit demjenigen zwei Stellen zuvor übereinstimmt. Dabei ist in ihrem Ansatz zwischen einer Zielmodalität, in welcher die Aufgabe ausgeführt wird, und der zweiten Modalität, in welcher entweder der entsprechend kongruente Stimulus oder ein inkongruenter Distraktor mit präsentiert wird, zu unterscheiden. Sie verwendeten hierbei zwei Sets an Stimuli, eines bestehend aus einer Anzahl von Objekten sowie Tönen und eines aus gesprochenen sowie gezeigten Ziffern. Auffällig ist, dass sich der positive Kongruenzeffekt auf die Reaktionszeit in der auditiven Zielbedingung nur für Anzahl an Objekten und in der visuellen Bedingung nur für Ziffern zeigt.

Hardiess, Erhardt und Mallot zeigten 2015 zusätzlich auch einen positiven Effekt bimodaler Präsentation auf die Leistung des Arbeitsgedächtnisses. Sie verwendeten hierzu eine two-back-Aufgabe, in welcher die Stimuli entsprechend der Bedingung visuell, auditiv oder in bimodaler Form vorlagen. Die visuellen Stimuli waren hierbei 6 zufällige Muster bestehend aus 6 schwarzen Punkten. Die auditiven Stimuli bildeten sich aus 6 Klavier-Akkorden. In der bimodalen Bedingung setzten sich die Stimuli aus sechs unterschiedlichen Mustern sowie sechs Gitarrenakkorden zusammen.

Einen weiteren Blick auf die Auswirkung bimodaler Stimulus Präsentationen im Arbeitsgedächtnis wirft Niclas Renner in seiner Masterarbeit am Lehrstuhl für Kognitive Neurowissenschaft in Tübingen. In dieser untersucht er im Besonderen, inwiefern sich eine verbesserte perzeptionelle Leistung auch auf das Arbeitsgedächtnis auswirkt. Hierzu variiert er wie bereits Werner und Noppeney (2009) gemäß dem Prinzip der inversen Effektivität die Qualität des Informationsgehalts in welcher die Stimuli präsentiert werden. Die verwendeten 3 Stimuli waren dabei visuell unterschiedlich orientierte Gabor Patches und auditiv ein von links, rechts (im Winkel von 45 Grad) oder zentral klingender C-Dur Gitarrenakkord. Hierbei sind jeweils ein Patch und Akkord über die Orientierung paarweise kongruent. Die Studie bestand aus zwei geteilten Versuchen. Zuerst vollführten die Versuchsteilnehmer zur Detektionsschwellschätzung eine unimodale Kategorisierungsaufgabe, in welcher entsprechend der Modalität ein gezeigtes Bild oder Ton einem der drei möglichen Richtungsstimuli zugeordnet werden sollte. Während dieser Aufgabe wurde zunächst das Signal-zu-Rausch Verhältnis einer weißen gaußschen Maske und dem Stimulus zur Schätzung einer psychometrischen Funktion für die unterschiedlichen Detektionsraten variiert. Anhand dieser ließen sich die Stimuli nun für jeden Versuchsteilnehmer in die Faktorstufen nach den Erkennungsraten von 50, 75 und 99 Prozent verrauschen. Mit diesen Erkennungsraten ließ sich in einer erneuten Durchführung der Kategorisierungsaufgabe die Frage eines multisensorischen Wahrnehmungsvorteils untersuchen. Hierbei sind die Degradierungsstufen des Stimulus von besonderer Bedeutung, da sie erst das Erkennen eines multisensorischen Wahrnehmungsvorteils ermöglichen. Bei leicht zu identifizierenden und damit kategorisierenden Reizen ist eine Verbesserung der Wahrnehmungsleistung nicht möglich, weil diese bereits optimal funktioniert. Mittels der von Treisman (1998) eingeführten Methode der Probability-Summation wird dieser perzeptionelle Vorteil anschließend mit der Wahrscheinlichkeit einer besseren Zuordnungs- und Identifikationsrate aufgrund einer an das Race-Modell angelehnten Erklärung verglichen. So wird überprüft, inwiefern dieser lediglich auf einer multisensorischen Integration als Erklärung beruhen kann.

Anschließend wurde in einer two-back-Aufgabe untersucht, auf welche Weise sich ein möglicher wahrnehmungsbasierter Vorteil in der Arbeitsgedächtnisleistung wiederfindet. In dieser wurden in zwei Sitzungen blockweise die neun unterschiedlichen Faktorstufen (Detektionsschwelle*Modalität) variiert.

Die Ergebnisse zeigten jedoch keine der erwarteten und gemäß dem Prinzip der inversen Effektivität vermuteten Leistungsverbesserungen bei einer multisensorischen Präsentation der Stimuli. In der bimodalen Kategorisierungsaufgabe ließ sich feststellen, dass die korrekten Zuordnungsraten nicht über der mithilfe der unisensorischen Raten berechneten Probability-Summation Schwelle liegen. Die Rate richtiger Zuordnung überschreitet teilweise ebenso nicht den Mittelwert der unisensorischen Performances (siehe Abbildung 3). Auch nach der Betrachtung des Sensitivitätsindex (d' -prime) und der Reaktionszeit der Arbeitsgedächtnisaufgabe ließ sich kein signifikanter bimodaler Vorteil im Vergleich zu der visuellen Bedingung nachweisen. Lediglich die Leistung bei auditiven Stimuli lag deutlich unter der in den anderen beiden Bedingungen (Siehe Abbildung 4).

Die Ursache hierfür zeigt sich bei der Betrachtung der psychometrischen Funktion für die auditiven Erkennungsraten. Hierbei wird deutlich, dass intrapersonell sehr unterschiedliche richtungsabhängige Schwellenwerte existieren, welche zusammengesetzt zu keiner guten Schätzung führen. Eine weitere Ursache bilden teilweise von der auf einem bayesianischen Ansatz beruhenden adaptiven psi-marginal Methode angenommene nicht optimale Prior zum Anfitzen der psychometrischen Funktion (Siehe Abbildung 5). Ebenso zeigte sich bei beiden Aufgaben zur Schwellenbestimmung im Falle von Unsicherheit ein deutlicher Bias zur mittleren Orientierung.

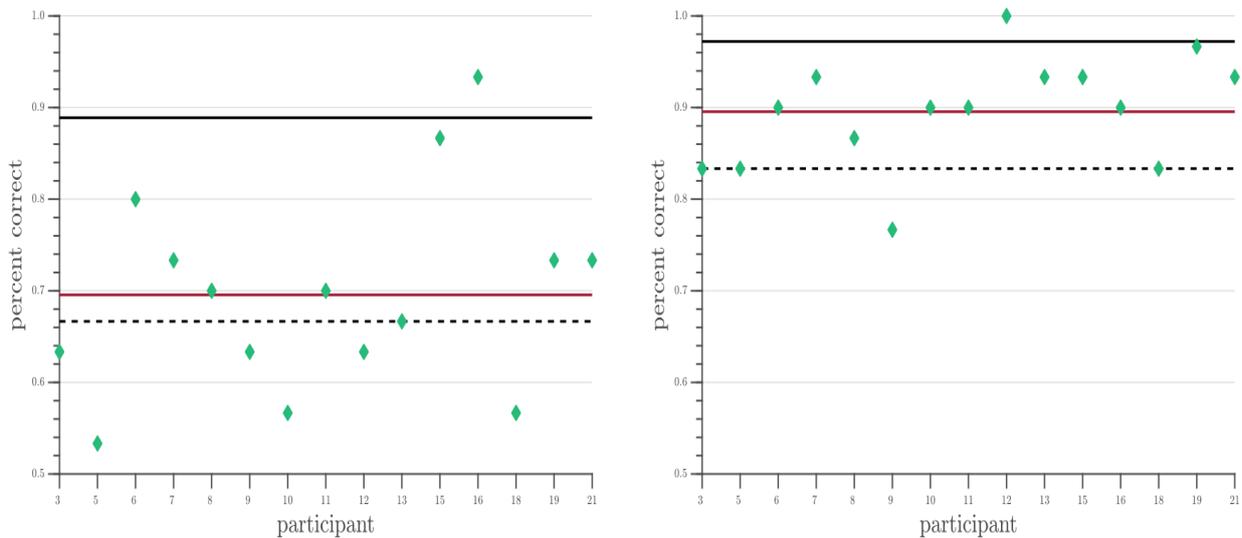


Abbildung 3: Performance in der bimodalen Kategorisierungsaufgabe. Die Grünen Punkte zeigen hierbei die individuellen Raten der einzelnen Versuchspersonen, die rote Linie den daraus berechneten Mittelwert, die gestrichelte Linie zeigt die unisensorischen Erfolgsraten (links: 66%; rechts: 83%) Die schwarze Linie bildet die mittels Probability Summation geschätzte Erfolgsrate aufgrund einer unabhängigen Verarbeitung. Abbildung aus Niclas Renner (2018)

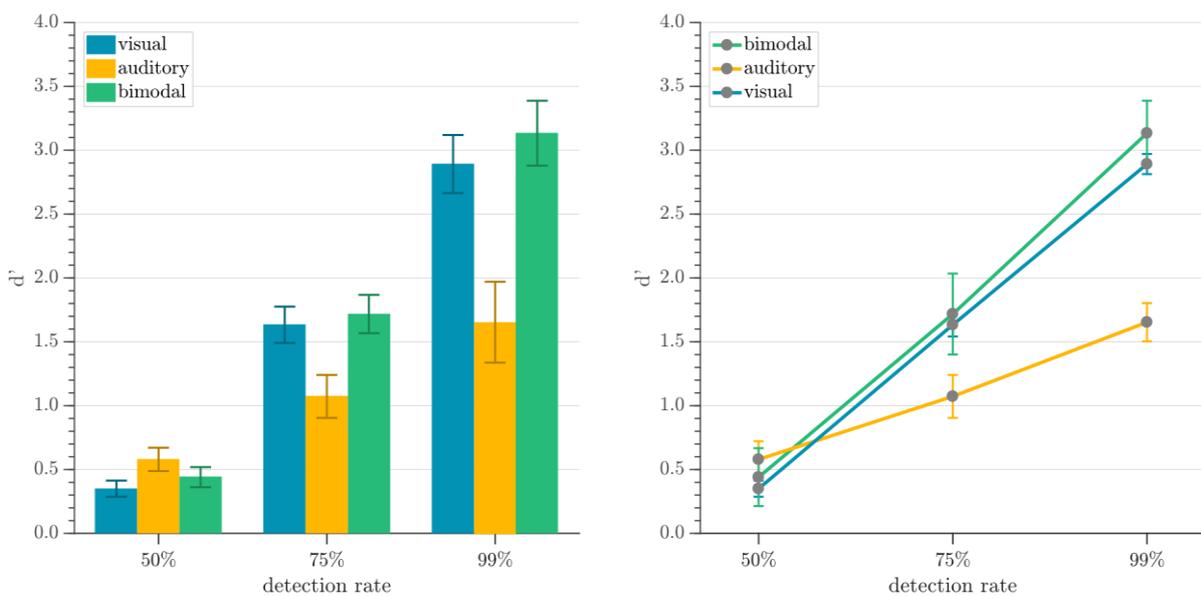


Abbildung 4: Sensitivitätsindex Leistung in Arbeitsgedächtnisaufgabe mit eingezeichnetem Standardfehler(SEM) Abbildung von Niclas Renner (2018)

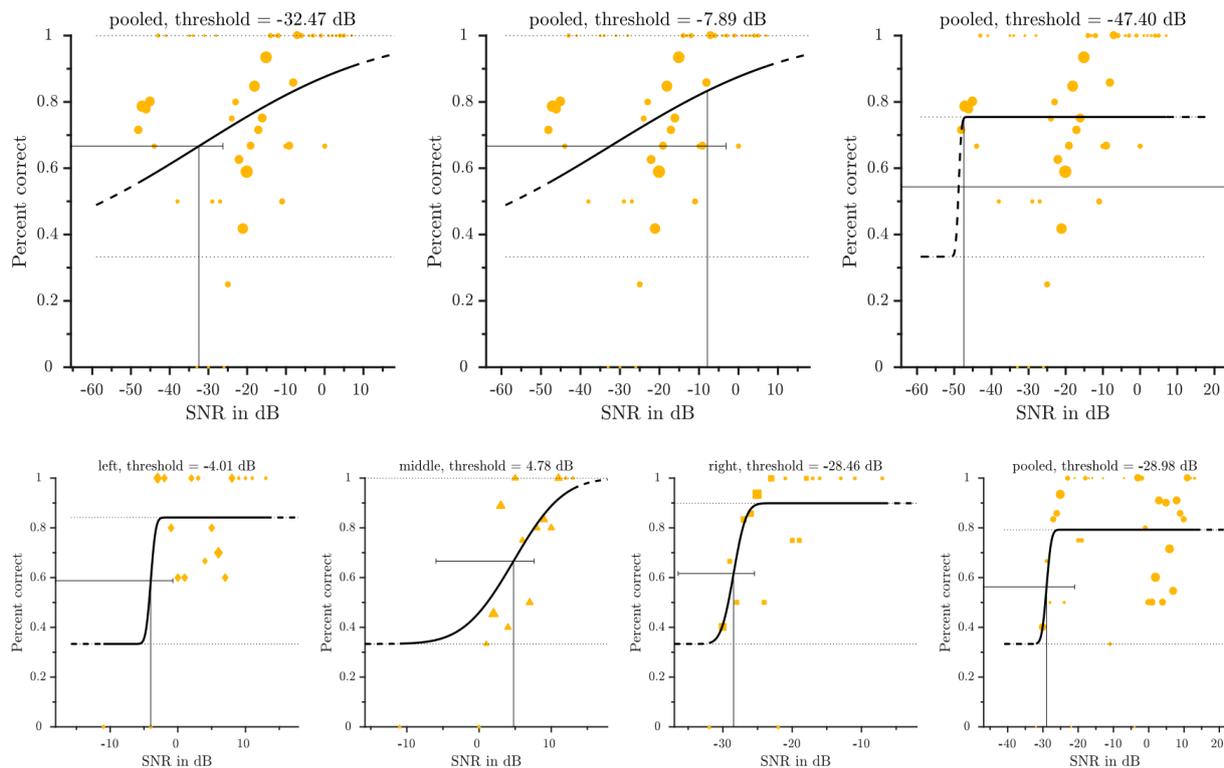


Abbildung 5: 1. Reihe: psychometrische fits für Erkennungsraten von 50%, 75%, 99%. Ein unsensibler Prior führt zu einem niedrigeren Signal-zu-Rausch Verhältnis bei 99% als bei 50%. 2. Reihe: große Unterschiede in den geschätzten richtungsabhängigen Schwellen und psychometrischen Funktionen führt zu einem schlechten Fit der Gesamtdaten. Abbildung aus Niclas Renner (2018)

2. Stimulus Wahl

In meiner Arbeit möchte ich nun auf den Ergebnissen von der Masterthese von Niclas Renner (2018) aufbauen und eine Lösung für die aufgetretenen Probleme finden. Hierbei ist vor allem das Finden und Überprüfen alternativer Stimuli von großer Wichtigkeit, um ein ungenaues Schätzen der psychometrischen Funktionen aufgrund der Richtungsabhängigkeit im Hörvermögen der einzelnen Versuchspersonen zu vermeiden.

2.1 Anforderungen

Die allerwichtigste Eigenschaft multisensorischer Stimuli bildet, dass die beiden einzelnen unimodalen Bestandteile aus Zuordnungszwecken zueinander kongruent sind. Um unter diesem Aspekt eine möglichst große Ähnlichkeit zu ermöglichen, ist es sinnvoll bei audiovisuellen Stimuli, wie von Werner und Noppeney (2008) vorgeschlagen, die zeitlich dynamische Natur des Schalls auch auf die visuelle Präsentationsform zu übertragen. Hierfür bildet die einfachste Lösung ein kurzes Video. Ein weiterer

interessanter Aspekt bei der Stimulus-Wahl ist die genaue Umsetzung der Kongruenz zwischen den unterschiedlichen Modalitäten. Die naheliegendste und am weitesten verbreitete Form ist hierbei die auch unter anderem von Werner und Noppeny (2008) genutzte Form der semantischen Kongruenz. Die Versuchsteilnehmer haben bereits gelernt das Geräusch seinem entsprechenden Abbild zuzuordnen und umgekehrt. Problematisch bei der Untersuchung von Prozessen des Arbeitsgedächtnisses ist aber hierbei, dass diese semantische Verbindung beider Reize nur mithilfe einer Langzeitgedächtnisbeteiligung möglich ist. Auch wenn dies in dem klassischen Verständnis des Arbeitsgedächtnisses durchaus möglich ist, könnte diese besondere Form der Integration einen Einfluss auf die Verarbeitung der beiden präsentierten Reize besitzen. So müsste zuvor überprüft werden, ob die Versuchsteilnehmer nicht unterschiedlich gut ausgeprägte Langzeitgedächtnisassoziationen zu den semantischen kongruenten Stimuli besitzen. Um dies sicher auszuschließen, sollten die verwendeten Stimulus-Paare dementsprechend abstrakt kongruent sein. Eine weitere Auffälligkeit in der Arbeit von Niclas Renner ist eine nahezu fehlerfreie Leistung bei der two-back-Aufgabe in der Bedingung mit einer Erkennungsrate von 99%. Die einfachste Möglichkeit um eine derartige Überperformance zu verhindern, besteht in einem Erschweren der Aufgabe. Eine einfache Lösung wäre hierfür das Einführen eines zusätzlichen Stimulus-Paares. Ein weiterer wichtigster Aspekt, der bei der Wahl der verwendeten Stimuli zu beachten ist, ist das Ausschließen eines stärkeren Einflusses einzelner Stimuli auf die Arbeitsgedächtnis- und Detektionsleistung. Um dies zu gewährleisten, muss die Reizenergie aller Stimuli innerhalb einer Modalität möglichst ähnlich sein. Zusammenfassend werden in dieser Arbeit also vier abstrakt kongruente Stimulus-Paare gesucht, welche über eine vergleichbare Stimulus- oder Signalenergie verfügen. Dabei bietet sich die Darstellungsform eines kurzen Videos in der visuellen Modalität an.

2.2 Erster Ansatz: zufällig generierte Tonfolgen

Um den Aspekt der möglichst gleichen Reizenergie zu gewährleisten, waren die ersten im Rahmen dieser Arbeit erstellten Stimuli aus drei zufälligen Sinustönen erstellte Tonfolgen. Hierzu wurden mithilfe von Matlab R2018b (Mathworks, Inc.) jeweils hundert unterschiedliche Tonfolgen erstellt und anschließend gemäß ihrer mittleren Reizenergie verglichen. Diese wurde mittels einer einfachen Formel für die durchschnittliche Reizenergie $\frac{Amplitude^2}{2}$ berechnet, was sich aus einer allgemeinen Definition der elektrischen Signalstärke $Power = \frac{Voltage^2}{Resistance}$ ableiten lässt. Die einzelnen Töne einer jeden Folge konnten hierbei eine sinusoidale Amplitude mit Werten zwischen 100 und 1000 aufweisen, was einer Spanne von 40db -80db gleichkommt. Die Frequenz, also die Klangfarbe eines jeden der drei Töne, wurde zufällig einem Wert zwischen 200Hz und 2000Hz zugeordnet. Da das menschliche Gehör am empfindlichsten für Töne mit einer Frequenz um 3500Hz ist, war die Spanne möglicher Frequenzen zunächst breiter, wurde dann aber aufgrund eines unangenehm schrillen Hörerlebnisses weiter eingeschränkt. Die geplante visuelle Umsetzung bestand hierbei aus einem Kreis, welcher gemäß den Eigenschaften des kongruenten Tones seine Beschaffenheit verändert. Die Frequenz, also der Ton, veränderte die Kreisfarbe in Graustufen, während die Amplitude (Lautstärke) einen Einfluss auf die Größe des Kreises nahm.

Bei genauerer Betrachtung der auditiven Stimuli und einer Planung des weiteren Vorgehens kamen jedoch einige Probleme auf, was schlussendlich zum Verwurf dieser Idee führte. Das offensichtlichste

Problem bildete hierbei das Ordnen der Töne nach einer möglichst gleichen Reizenergie, da diese lediglich von der Amplitude abhängig ist, welche gleichzeitig aber ein zu variierendes Merkmal bildete. Dementsprechend bestanden die Tonfolgen mit der ähnlichsten Reizenergie immer aus vergleichbar lautstarken einzelnen Tönen oder Tönen mit Lautstärken an beiden extremen Enden des Spektrums, was zu einem sehr anstrengenden Klangeindruck führte. Ein weiteres Problem wird offensichtlich, wenn man sich der Verrauschung des visuellen Stimulus zuwendet. Da sich in dieser Darstellungsform sowohl die Helligkeit des Kreises als auch seine Größe verändert, werden sich auch Schwellenwerte des Signal-zu-Rausch Verhältnis für das Erkennen des jeweiligen Merkmals unterscheiden. Eine bewegte Kante, wie es eine Vergrößerung des Kreises nach sich zieht, wird immer wesentlich schwerer durch ein Störreiz zu überlagern sein, als eine Veränderung der Helligkeit innerhalb eines Farbtons.

2.3 Zweiter Ansatz: gebrochene Dreiklänge

Um den Klangcharakter der auditiven Reize zu verbessern, wurden die Sinustöne in einem zweiten Ansatz nicht mehr zufällig erstellt, sondern in ihrer Frequenz an Ganztönen orientiert. Hierbei wurde zusätzlich immer der Grundton mit zwei Obertönen zu einem Tonika-Dreiklang seiner jeweiligen Tonart vervollständigt. Um auch einen harmonischen Klangcharakter der Tonfolge zu garantieren, spielte diese immer einen gebrochenen C-Dur-Dreiklang mit unterschiedlichen Anfangstönen und in unterschiedlichen Tonlagen. Anstatt durch die Amplitude der einzelnen Töne unterschieden sich die unterschiedlichen Tonfolgen nun zusätzlich noch anhand ihrer Rhythmik.

Leider stellte sich auch hierbei heraus, dass trotz unterschiedlicher Töne und Tonlagen die mithilfe von Matlab R2018b erstellten Stimuli nur schwer differenzierbar sind. Da das subjektive Lautstärkeempfinden des Weiteren neben der Amplitude auch noch von der Tonhöhe (Frequenz) abhängig ist, zeigten sich diesbezüglich auch Unterschiede zwischen den Stimuli. Dies ließe sich zwar durch ein Orientieren am tatsächlichen Lautheitswert in Sone ausgleichen, wurde aber aufgrund eines Verwerfens des Stimulus nicht mehr durchgeführt. Eine weitere Sorge bestand erneut in der Deutlichkeit der visuellen Darstellung, da die Dimension der Rhythmik sich lediglich über den Zeitpunkt des Farbwechsels in Graustufen bestimmen lässt. Dies ist wesentlich leichter durch ein Rauschen zu überdecken und grundsätzlich schwerer in einem kurzen Zeitintervall wahrzunehmen als eine dauerhafte Farbveränderung. Da die Töne jedoch aufgrund ihrer Erstellungsmethode einen ähnlichen Klangcharakter aufweisen, wären die unterschiedlichen Rhythmen vermutlich ein entscheidender Aspekt der auditiven Differenzierung.

Um die aufgetretenen Probleme der visuellen Darstellung zu umgehen, bietet es sich an die Stimuli nur noch in einer Dimension zu variieren. Hierbei fiel die Entscheidung auf die Rhythmik, da dies ein einfaches Orientieren der Kreisgröße gemäß der aktuellen Lautstärke des Stimulus ermöglicht.

2.4 Stimulus: Rhythmische Tonanschläge

Die Stimuli, welche letztendlich die geforderten Ansprüche ohne größere Schwierigkeiten bei der Umsetzung erfüllten, bilden unterschiedliche Rhythmen eines Tonanschlags. Diese werden visuell über einen Kreis in Graustufen gemäß der Tonamplitude dargestellt. Nach einigem Ausprobieren mit kurzen harmonischen Gitarren oder Klavierakkorden, fiel die Wahl des zu verwendenden Tons schlussendlich auf einem einzelnen kurzen Basston. Die Vorteile sind hierbei eine auch nach vielen Wiederholungen angenehme Klangfarbe sowie hauptsächlich die zeitliche Kürze eines jeden Anschlags. Diese ermöglichte ein Aufteilen von drei Tonanschläge auf die Stimuluspräsentationsdauer von 1,5 Sekunden. Hierbei ist es wichtig, dass sich der Stimulus aus drei oder mehr einzelnen Anschlägen zusammensetzt, da so erst vier unterschiedliche Rhythmen entstehen können. Zur Erstellung und für das Schneiden der auditiven Stimuli wurde das Programm Audacity Portable verwendet. Die Verteilung der Tonanschläge über die 1.5 Sekunden erfolgt hierbei über ein Einteilen der Präsentationsdauer in 5 Intervalle. Diese orientiert sich an der Länge eines Tonanschlags von 0.25 Sekunden. Am Anfang und Ende der Tonspur wird hierbei nicht unmittelbar ein Tonanschlag gesetzt, weil dies unweigerlich zu einem abgehackten Eindruck führt. Die ersten drei zu unterscheidenden Rhythmen haben dabei die Form 1-3-5 (Abbildung 6), 1-2-5 (Abbildung 7) sowie 1-4-5 (Abbildung 8). Die Benennung bezieht sich hierbei stärker auf die zwischen den einzelnen Anschlägen liegenden Pausen, als die exakten Intervalle, in welchen diese erfolgen. Die ‚1‘ symbolisiert einen ersten Anschlag im Zeitraum 0.05 Sekunden bis 0.30 Sekunden, die ‚2‘ einen unmittelbar darauffolgenden zwischen 0.30 Sekunden und 0.55 Sekunden. Ebenso verhält es sich mit den Anschlagpunkten, welche mit 4 und 5 codiert sind. Diese symbolisieren zusammen zwei Anschläge in einem Zeitraum von 0.9 sec bis 1.4 sec. Um die Gleichmäßigkeit des Rhythmus bei dem Stimulus 1-3-5 zu gewährleisten, wird der Anschlagzeitpunkt, welcher mit der 3 gekennzeichnet ist, exakt in die Mitte zwischen die anderen beide gesetzt. Der Basston nimmt also in diesem Fall einen Zeitraum zwischen 0.6 sec und 0.85 sec in der Tonspur ein. Hierbei ist des Weiteren von Bedeutung, dass die Pausen zwischen den einzelnen Tönen möglichst groß sind. Denn um gemäß den Anforderungen einen vierten, von den anderen differenzierbaren, Rhythmus aus drei Tönen zu generieren bleibt, nach der Verwendung von langen Pausen zwischen allen Tönen, ein Stimulus, in welchem die drei Anschläge direkt hintereinander abgespielt werden. Dies findet bei diesem in der Mitte der Tonspur zwischen 0.375 Sekunden und 1.125 Sekunden statt. Hierbei ist zu beachten, dass sich dieser Stimulus in seinem Klangcharakter nur über das Fehlen von Pausen gegenüber dem Stimulus 1-3-5 absetzt, weswegen der mögliche Unterschied zwischen diesen diesbezüglich vollständig ausgenutzt wird, auch wenn dies die Intervalleinteilung verletzt. Um nichtsdestotrotz eine einheitliche Benennung anhand der Pausenposition zu gewährleisten, wird der vierte Stimulus mit 2-3-4 (Abbildung 9) bezeichnet. Die Tonspur eines jeden Stimuli ist in Stereo und mit 8000Hz gesampelt.

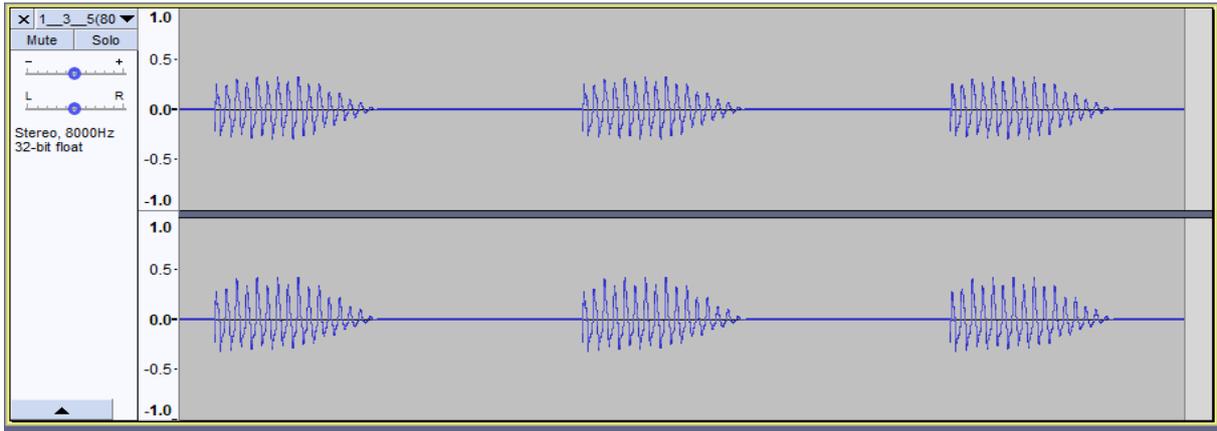


Abbildung 6: Stimulus 1-3-5 in Wellenform.

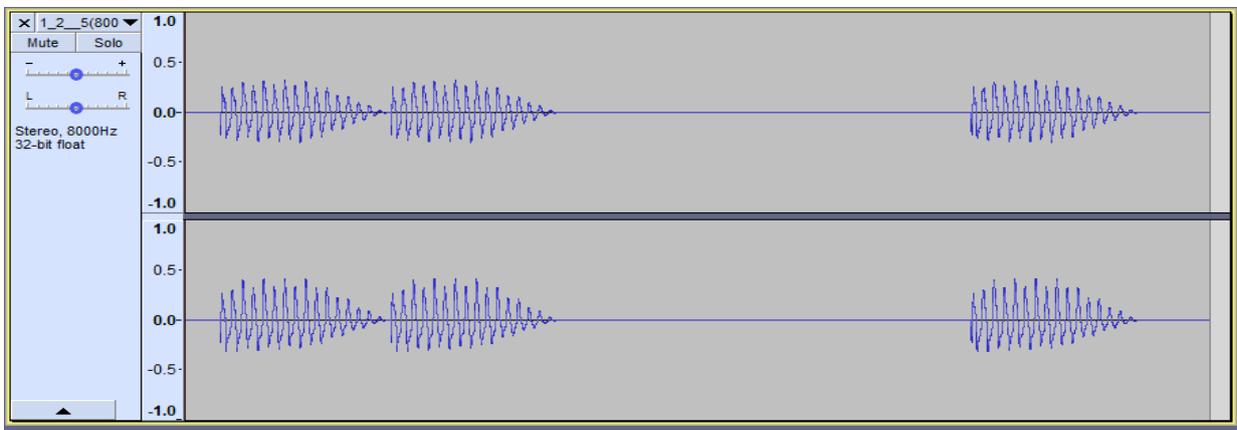


Abbildung 7: Stimulus 1-2-5 in Wellenform

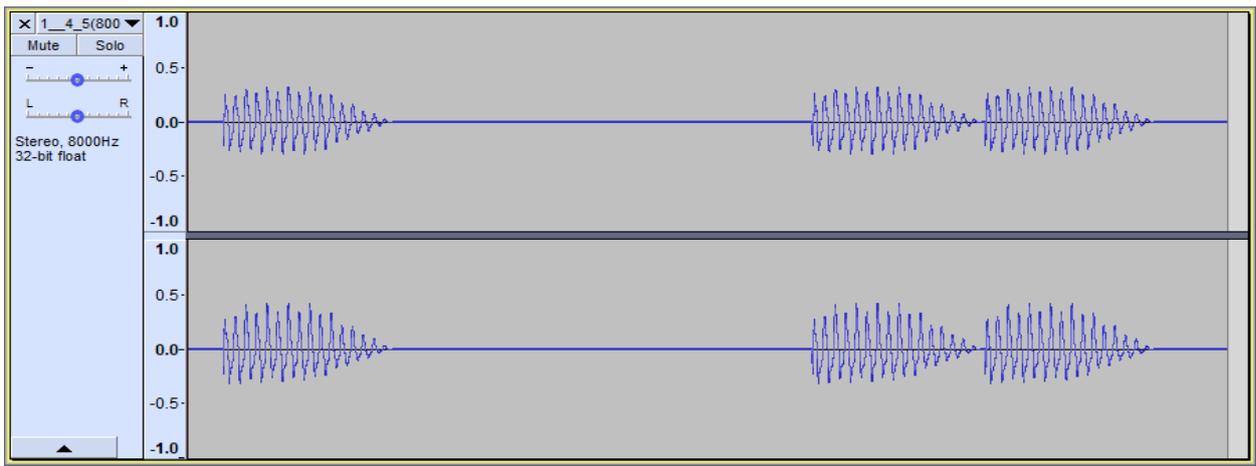


Abbildung 8: Stimulus 1-4-5 in Wellenform

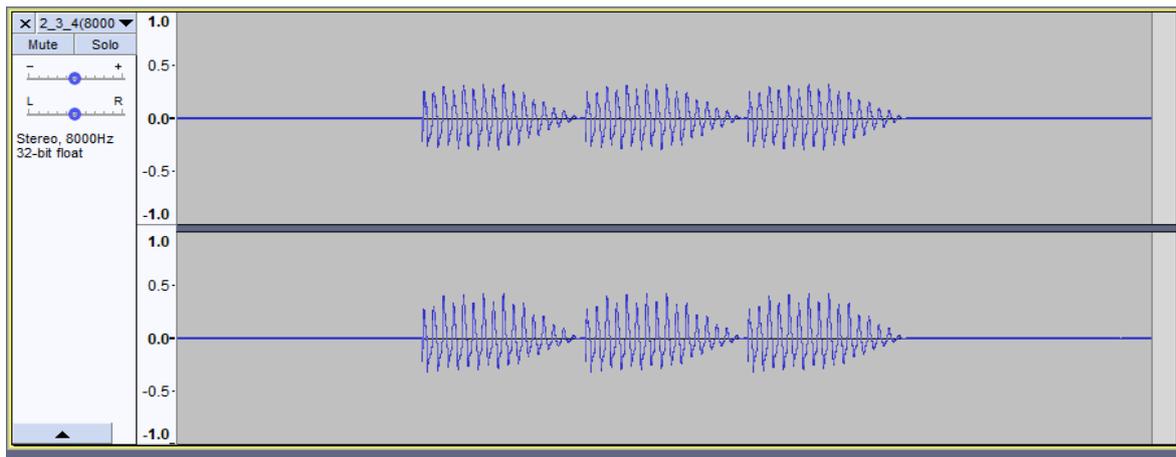


Abbildung 9: Stimulus 2-3-4 in Wellenform.

Die für Audiodateien relativ niedrige Samplingrate von 8000Hz liegt hierbei in der Vermeidung von Informationsverlust bei der Erstellung des visuellen Stimulus aus der Tonspur begründet. So wird die Tonspur in Matlab 2018b auf eine Rate von 60Hz für eine Darstellung als Video reduziert. Dabei wird der ursprüngliche Ton von Stereo auf eine Tonspur reduziert und die negativen Amplituden werden ins Positive umgekehrt. Nach einer Normierung der Werte auf das Intervall zwischen 255 und 0, was den Graustufen in Matlab entspricht, entsteht so ein Vektor, welcher für jeden der 90 Frames eines 1.5 sekundigen Videos (Abspielrate von 60Hz) den zur auditiven Darstellungsform kongruenten Farbwert enthält. Dieser ermöglicht mithilfe von Matlab die dynamische Erstellung eines Kreises für die visuelle Darstellungsform parallel zum Abspielen der Tonspur, welcher mit jedem Anschlag im Rhythmus eine schwarze Farbe annimmt und so scheinbar entsteht. Dabei ist der niedrigste Farbwert des Kreises mit 240 bewusst nicht auf ein reines weiß festgesetzt, damit immer eine Kontur zur Orientierung erkennbar bleibt. Abbildung 10 verschafft hierbei einen Eindruck über die Farbe des Kreises zu unterschiedlichen Zeitpunkten der Tonspur. Wichtig ist dabei zu beachten, dass die Relation zwischen Kreisbild und Rahmen nicht den wirklichen Versuchsbedingungen entspricht. Die scheinbare Größe des Kreises beträgt unter diesen 0.118 Radiant und wurde mit der Formel $alpha = 2 * \arctan\left(\frac{g}{2r}\right)$ berechnet, wobei g den Durchmesser des Kreises und r den Abstand der Versuchsperson zum Bildschirm enthält.

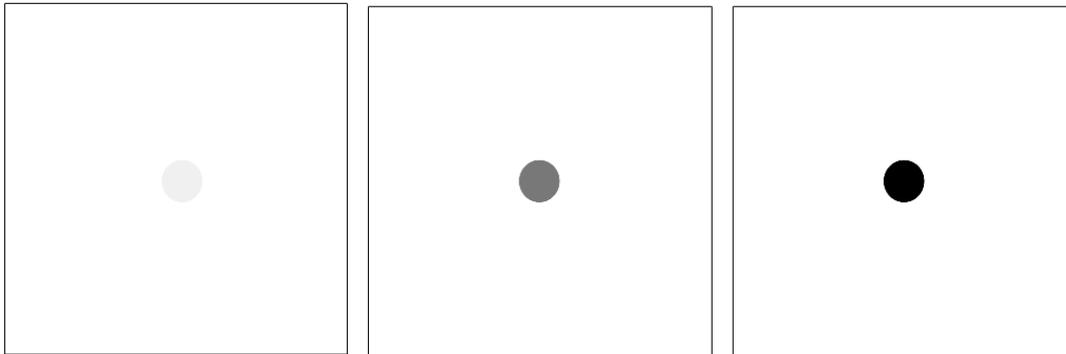


Abbildung 10: Farbwerte der visuellen Darstellung. Links: Kreisfarbe in den Pausen zwischen den Tonanschlägen (Farbwert:240). Mitte: Kreisfarbe bei Anschwellen des Tons (Farbwert: 120). Rechts: Kreisfarbe bei maximaler Lautstärke (Farbwert: 0)

Es wirkt, als würde sich zeitgleich mit jedem Anschlag eines Basstons ein schwarzer Kreis bilden und wieder verschwinden. Durch diese zeitliche Kongruenz sind die beiden unterschiedlichen Darstellungsformen auf eine abstrakte Art und Weise einander zuordenbar. Da sich jeder der Stimuli aus den gleichen Bestandteilen zusammensetzt und sich hauptsächlich in der Platzierung und Länge der Pausen unterscheidet, ist somit auch die Reizenergie vergleichbar. Jedoch stellt sich bei der Verwendung der Stimuli die Frage, ob sich die unterschiedlichen Rhythmen bei der Präsentation einer Stimulusfolge nicht zu einem schwammigen Gesamteindruck vermischen.

Des Weiteren ist es sinnvoll beim weiteren Vorgehen ein Auge auf die scheinbar größere Ähnlichkeit der Stimuli 1-3-5 und 2-3-4 zu legen. Diese könnte gerade bei einer klassischen Arbeitsgedächtnisaufgabe wie einem two-Back-Ansatz zu einem ungewollten Einfluss auf die Leistung führen. Um die Stimuli auf ihre Eignung zu überprüfen, wurde im Folgenden eine einfache Arbeitsgedächtnisaufgabe durchgeführt.

3. Methode

3.1 Versuchsteilnehmer

Die 11 am Versuch teilnehmenden Personen bilden sich aus Studenten und Angestellten der Eberhard-Karl-Universität Tübingen. Alle am Versuch teilnehmenden Personen waren rechtshändig. Von den elf Teilnehmern waren 3 weiblich. Der Altersdurchschnitt liegt bei 23 Jahren, bei einer Spanne von 19-31 Jahren. Alle Teilnehmer verfügten über ein normales oder korrigiertes Sehvermögen. Die Versuchspersonen nahmen freiwillig und ohne finanziellen Ausgleich an dem Experiment teil und unterzeichneten zuvor diesbezüglich eine Einverständniserklärung. Sie wurden angehalten, bei Müdigkeit oder Bedarf jederzeit eine Pause zu machen. Hierfür war im experimentellen Aufbau zusätzlich nach jedem Block eine Pause eingebaut, sodass die Teilnehmer selbst die Geschwindigkeit des Voranschreitens bestimmen konnten. Alle Probanden waren zuvor naiv, also ohne Vorwissen, gegenüber dem experimentellen Setup sowie der Zielsetzung des Experiments.

3.2 Materielles Setup

Der Versuch fand in einem schwach beleuchteten Raum an einem Computer statt. Das hierbei verwendete Betriebssystem ist Windows 10. Zum Programmieren des Experiments wurden, wie schon zuvor bei der visuellen Darstellung der Stimuli, Matlab R2018b (Mathworks, Inc.) mit Einbindung der Psychophysics Toolbox Version 3 (Brainard und Vision, 1997) und der Signal Processing Toolbox verwendet. Die visuellen Reize wurden über einen Fujitsu B22t-7 proGreen Bildschirm bei einer Helligkeit von 100 und Kontrastwerten sowie einem Schwärze-Level von 50 präsentiert. Die auditiven Reize wurden über Kopfhörer (Sennheiser HD 280 Pro) abgespielt. Die Antworten der Teilnehmer wurden über eine HP Smartcard CCID Tastatur aufgenommen. Die Versuchsperson saß ohne eine Kopffixierung in etwa 55 cm vom Bildschirm entfernt.

3.3 Versuchsaufbau

Als Arbeitsgedächtnisaufgabe zur Überprüfung der Eignung der Stimuli wurde wie bereits in der Arbeit von Niclas Renner (2018) ein two-back-Ansatz gewählt. Hierbei wurde blockweise die Modalität, in welcher die Stimuli dem Teilnehmer präsentiert wurden, variiert. Dies beugt einem Effekt durch Modalitätswechsel innerhalb eines Blocks vor, wie er unter anderem von Gonden und Lange (2004) auf die Reaktionszeit bei bimodalen und unimodalen Darstellungen untersucht wird. Die Bedingungen, in welchen die Stimuli präsentiert wurden, waren unisensorisch visuell, unisensorisch auditiv und eine multisensorisch kongruente Darstellungsweise. Zusätzlich wurden zwei weitere Bedingungen überprüft, in welchen, wie von Werner und Noppeney (2009) vorgeschlagen, die unisensorischen Zielstimuli in der jeweils anderen Modalität von einem Rauschen begleitet werden. Dies hält die Reizenergie zwischen den Bedingungen in welchen die Information unisensorisch und multisensorisch

präsentiert werden ansatzweise gleich. Das auditive Rauschen wurde hierbei mithilfe von Audacitys bereitgestellter Funktion für weißes gaußsches Rauschen erstellt. Anschließend wurde die Lautstärke des Rauschens auf die des auditiven Stimulus normalisiert, um die Reizenergie zwischen Reiz und Rauschen vergleichbar zu halten. Für ein angenehmeres Hörempfinden wurde das Rauschen zusätzlich mit einer Ein- und Ausblende versehen (siehe Abbildung 10). Das Rauschen in der visuellen Bedingung wurde in Matlab mithilfe der Funktion `wgn`, welche eine weißes gaußsches Rauschen Matrix entsprechenden Größe und Stärke erstellt, generiert. Hierbei ist zu beachten, dass diese mit jedem Frame neu aufgerufen wurde, was zu einem zufälligen dynamischen Rauschen führt.

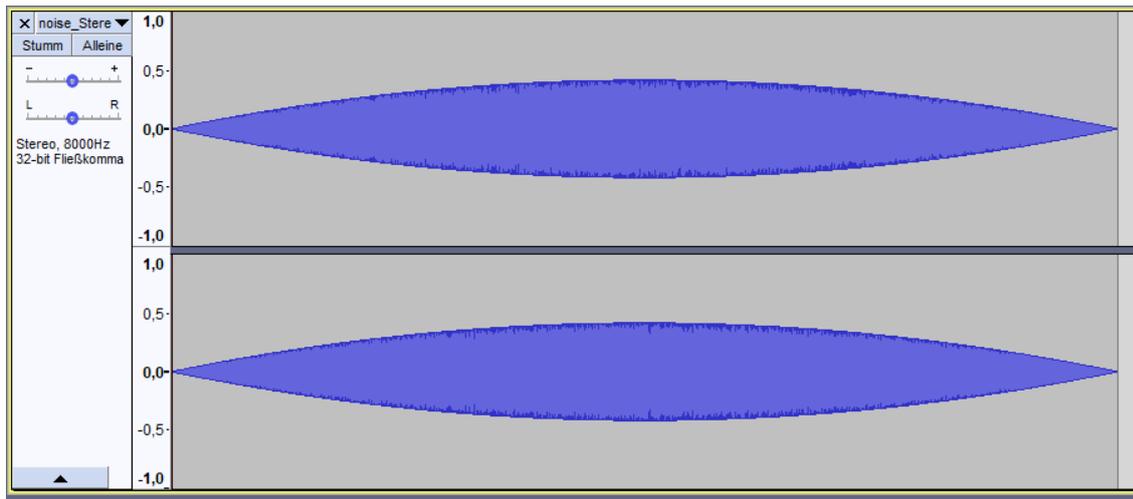


Abbildung 11: Rauschen in der auditiven Modalität, besteht aus ein- und ausgeblendetem white Gaussian noise. Die Abbildungen 6-10 sind aus Audacity portable entnommen

Jeder der 5 Versuchs-Blöcke bestand jeweils aus 48 Stimulus Präsentationen. Davon waren 12 Target-Durchgänge, bei welchen der gezeigte Stimulus mit dem zwei Stellen zuvor gezeigtem übereinstimmte, und 36 no-Target-Durchgänge, bei welchen die zu vergleichenden Stimuli unterschiedlich waren. Die 12 Target-Durchgänge teilten sich wiederum in 4 für jeden unterschiedlichen Stimulus auf. Insgesamt wurde in jedem einzelnen Block jede Form des Stimulus genau 12-mal dem Versuchsteilnehmer präsentiert. Die Reihenfolge, in welcher die Blöcke von unterschiedlichen Versuchspersonen bearbeitet wurden, wurde in dem Versuchsaufbau nicht variiert, da das Ziel hauptsächlich in der Untersuchung der Nutzbarkeit der gewählten Stimuli lag. Die Blöcke waren immer in der Reihenfolge visuell, auditiv, audiovisuell, visuell mit auditivem Rauschen und auditiv mit visuellem Rauschen. Vor jedem Block wurde dem Versuchsteilnehmer die Zielmodalität in einem kurzen Text mitgeteilt. Vor dem ersten Block erfolgte eine ausführlichere Erklärung der Aufgabe. Zusätzlich wurde vor Beginn einer jeden Untersuchung die Aufgabe der Versuchsperson in einer Anleitung erklärt, welche auch mit der Versuchsleitung durchgesprochen wurde, wobei etwaige Fragen beantwortet werden konnten. Vor Beginn eines jeden Blocks wurde für 2 Sekunden ein Fixationskreuz im Zentrum der visuellen Darstellungsform präsentiert.

3.4 Versuchsablauf

Nach der Einführung eines Teilnehmers wurde diesem zunächst ein Testdurchlauf des Experiments vorgelegt, damit dieser alle unterschiedlichen Stimuli kennenlernen und sich an die Aufgabe gewöhnen konnte. Hierbei bestand jeder Block aus 16 Durchgängen, in denen jeder Stimulus vier Mal präsentiert wurde. Das Training bot hierbei auch die Möglichkeit bei Bedarf aufkommende Fragen zu klären. In der eigentlichen Aufgabe wurde dem Teilnehmer anschließend in jedem Block eine Folge aus 48 Stimuli präsentiert. Vor der Darstellung eines Stimulus wurde vor jedem Durchgang für 0.5 Sekunden ein Fixationskreuz im Zentrum des Bildschirms gezeigt. Daraufhin folgt die 1.5 sekündige Stimulus Präsentation mit einer anschließend 1 sekündigen Blank Phase. Während dieser 2.5 Sekunden wurde die Antwort des Versuchsteilnehmers aufgezeichnet. Erkannte dieser den gezeigten Stimulus als den gleichen wie den zwei Stellen zuvor präsentierten, so sollte dies mit ‚y‘ für yes bestätigt werden. War dies nicht der Fall, war die geforderte Eingabe ‚n‘ für no. Die Versuchspersonen nutzen zur Eingabe gleichmäßig immer die entsprechenden Zeigefinger (‚y‘ linke Hand, ‚n‘ rechte Hand). Kam es zu zwei Eingaben innerhalb dieser Zeit, so wurde keine von beiden aufgezeichnet. Das Ende des Antwortintervalls wurde anschließend durch eine erneute Präsentation des Fixationskreuzes für 0.5 Sekunden markiert. Dies diente gleichzeitig der Ankündigung der nächsten Stimulus Präsentation und fixierte den Blick der Teilnehmer auf die Mitte der Darstellung (siehe Abbildung 12)

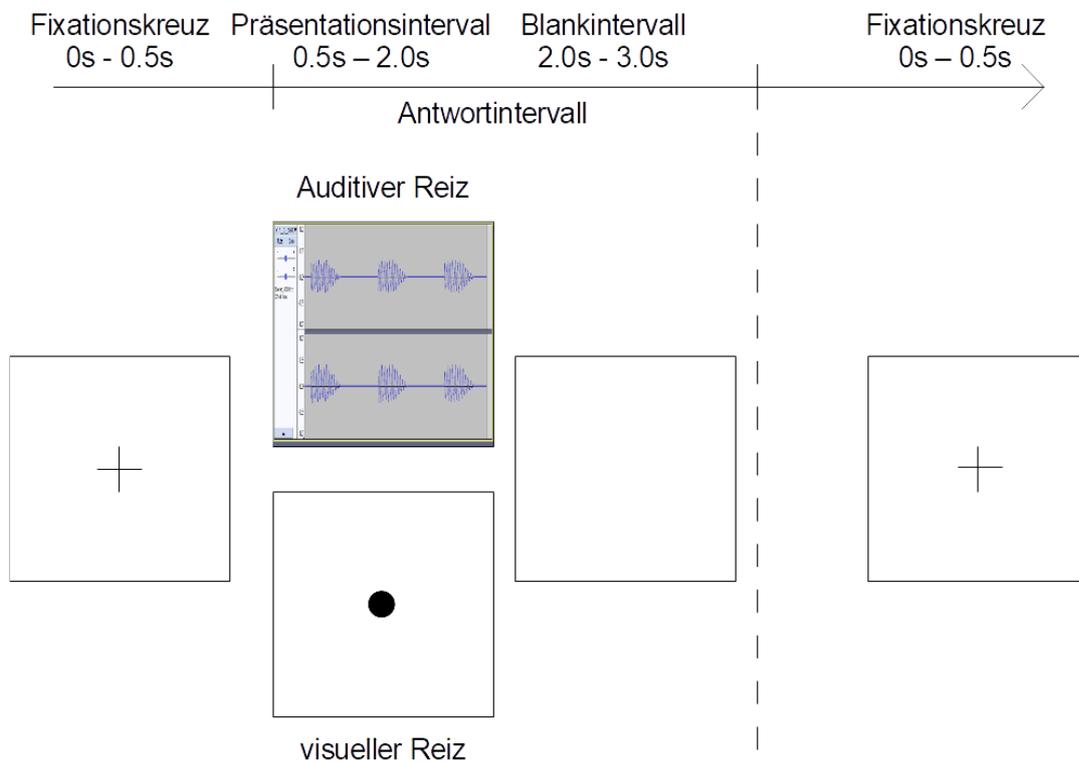


Abbildung 12: Ablauf eines Versuchsdurchgangs. Die gestrichelte Linie markiert das Ende des Antwortintervalls und den Start eines neuen Durchgangs.

3.5 Auswertung

Während der two-back-Aufgabe wurde die Reaktionszeit sowie die Antwort der Teilnehmer für jede Darstellung ((y', n')) aufgezeichnet. Aus letzteren lässt sich die Wahrscheinlichkeit für korrekt erkannte two-back-Matches (Hit Rate) sowie die Wahrscheinlichkeit eines False Alarms erschließen. Als False Alarm bezeichnet man das Bestätigen eines two-back-Matches ohne dass ein solches vorliegt. Durch das Berechnen des Sensitivitätsindex mittels $d' = z(\text{hit Rate}) - z(\text{false alarm Rate})$ lässt sich die Leistung der Versuchsteilnehmer und damit auch die Verwendbarkeit der Stimuli für eine Arbeitsgedächtnisaufgabe überprüfen. z' bezeichnet hierbei die Umkehrfunktion der gaußschen Verteilungsfunktion. Hierbei wurden Raten von genau Null oder Eins korrigiert, da diese nach der Umkehrung zu z Werten von negativ unendlich und unendlich führen können (Stanislaw, Todorov, 1999). Der Sensitivitätsindex ist das hierbei gewählte Performancemaß in Detektionsaufgaben (two-back-match Detektion), da es anders als zum Beispiel die einfache Hitrate in Prozent auch das fehlerhafte Erkennen angeblicher target-Durchläufe mit einberechnet. Das verringert den Einfluss des richtigen Bestimmens der two-back-matches basierend auf Unwissen, was sonst zu überverhältnismäßig guten Ergebnissen führen kann. Um diesbezüglich einen Vergleich zwischen den Bedingungen zu ermöglichen, wird der Sensitivitätsindex für jeden Block und die Gesamtheit der Daten berechnet.

Da nicht viele Daten erhoben wurden und die Genauigkeit der erzielten Sensitivitätsindex Ergebnisse demzufolge sehr gering ist, werden des Weiteren mithilfe der Antwortraten der einzelnen Bedingungen und der Bootstrap Methode 95% Konfidenzintervalle erstellt. Dies liefert einen guten Einblick in die mögliche Varianz der Ergebnisse und ermöglicht sinnvolle Vergleiche zwischen den einzelnen Bedingungen. Dabei ist zu beachten, dass diese Intervalle aus einem in 100.000 simulierten Experimenten berechneten Datensatz erstellt werden und dementsprechend einen breiteren Einblick in die mögliche Varianz der Sensitivitätsindex-Werte vermitteln als die tatsächlich gemessenen Daten. Um weitere Vergleiche zwischen den Bedingungen zu ziehen wird eine einfaktorielle Varianzanalyse mit ihnen als Faktor durchgeführt. Zum Nachweisen eines möglichen bimodalen Vorteils auf die Reaktionsgeschwindigkeit wird die bimodale Reaktionszeit wie in der Arbeit von Juan et al (2017) über die Formel $RT_{gain} = \frac{U_{max} - AV}{U_{max}} * 100$ verglichen. U_{max} ist hierbei die Reaktionszeit in der besten unisensorischen Bedingung, AV die der audiovisuellen. Dies geschah zunächst für alle Versuchsteilnehmer. Anschließend wurden der Mittelwert für einen insgesamt Überblick herangezogen. Dabei ist zu beachten, dass doppelte Antworten oder fehlende Antworten aus der Analyse ausgeschlossen werden. Die Reaktionszeit wurde ab Beginn der 1,5 sekündigen Stimulus Präsentation gemessen.

In einem Versuch Informationen über die Unterscheidbarkeit der Stimuli sowie weitere subjektive Eindrücke der Teilnehmer bezüglich der Eignung der gewählten Stimuli aufzunehmen, wurde diese gebeten im Anschluss an das Experiment einen Fragebogen auszufüllen. Die Fragen waren dabei folgende: Schätzen Sie bitte die Schwierigkeit der heutigen Arbeitsgedächtnisaufgabe ein? Verwenden Sie eine besondere Taktik? Empfanden Sie ein oder mehrere der unten genannten Bedingungen als leichter? Konzentrierten Sie sich in dem bimodalen Block stärker auf die Bilder oder auf die Töne? In welcher Modalität wurden Sie am stärksten durch das Rauschen eingeschränkt? Nach drei Versuchspersonen wurde der Fragebogen zusätzlich noch um Fragen bezüglich der Differenzierbarkeit der unterschiedlichen Stimuli innerhalb einer Darstellungsform sowie das Ausmaß der auf dem Rauschen beruhenden Erschwernis erweitert.

4. Ergebnisse

4.1 Sensitivitätsindex d'

Der Sensitivitätsindex über alle Bedingungen und Versuchspersonen hinweg beträgt in der Untersuchung 1.44. In der visuellen Bedingung beträgt der Mittelwert über alle Teilnehmer 1.26, in der auditiven Bedingung 1.60, in der bimodalen Bedingung 1.93, in der visuellen Bedingung mit auditivem Rauschen 1.13 und in der auditiven Bedingung mit visuellem Rauschen 2.09. Der Wert Null bildet hierbei als Referenz den Fall ab, dass die Versuchsperson rät, ob der gezeigte Stimulus der selbige ist wie zwei zuvor.

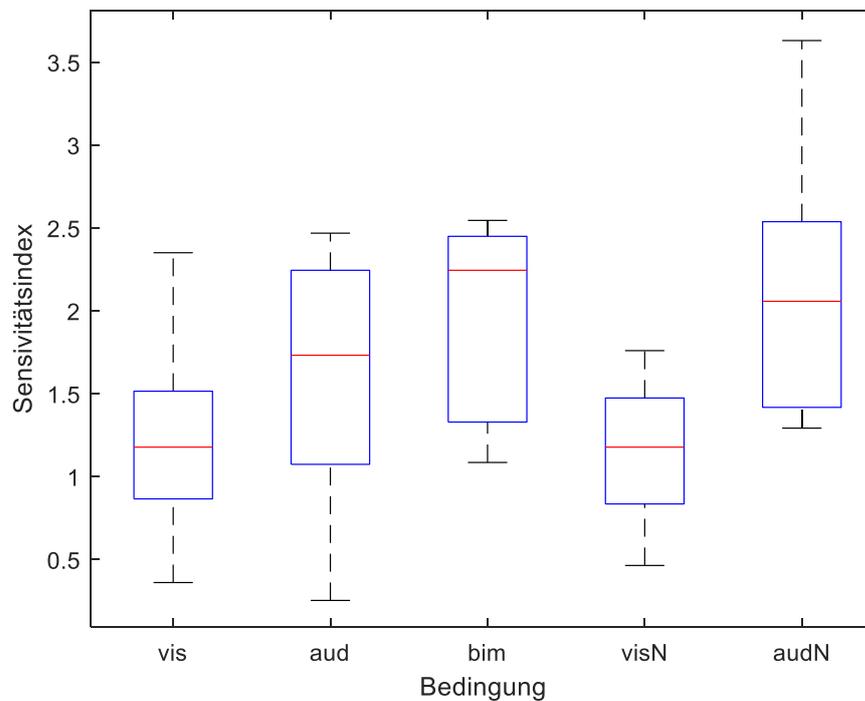


Abbildung 13: Boxplot der Sensitivitätsindex Werte. Die rote Linie bildet den Median, die Blaue Box gibt die 25% und 75% Perzentile an. Extremwerte sind über die gestrichelten Linien gekennzeichnet. Die Bedingungen sind wie folgt abgekürzt. Vis steht für unisensorisch visuell, aud für unisensorisch auditiv, bim für bimodale Darstellung, visN für eine visuelle Darstellung begleitet von einem auditiven Rauschen und audN für eine auditive Darstellung begleitet von einem visuellen Rauschen.

4.2 Sensitivitätsindex Konfidenzintervalle

Da die mithilfe der Antwortraten berechneten Sensitivitätsindex-Werte nur bedingt genau sind und keine Vergleiche untereinander zulassen, wurden, um ein besseres Verständnis für die mögliche Varianz der Werte zu gewinnen, 95% Konfidenzintervalle erstellt. Diese geben an, dass in 95 Prozent der erstellten Intervalle der wahre Wert in den Intervallgrenzen liegt. Die zur Erstellung verwendete Methode ist als Bootstrap Methode bekannt und beruht auf dem Resampling der beobachteten Ergebnisse (Efron und Tibshirani 1993). Die genaue Größe der Intervalle und ihre Relation zueinander ist in Abbildung 12 und der darauffolgenden Tabelle abzulesen. Auffällig ist, dass sich alle Intervalle zumindest teilweise überschneiden sowie über Null liegen.

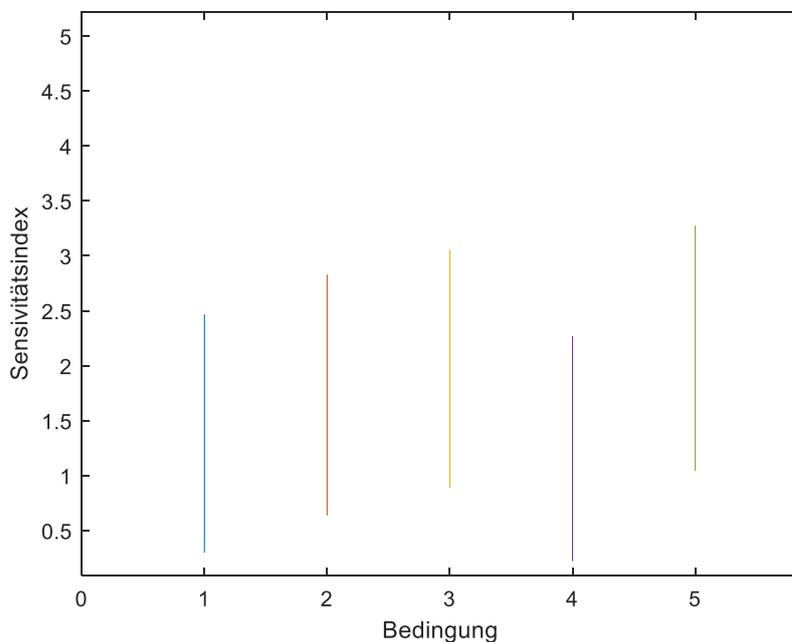


Abbildung 14: 95% Konfidenzintervalle der Sensitivitätsindex Werte für 1.) unisensorisch visuell, 2.) unisensorisch auditiv, 3.) bimodale Darstellung, 4.) visuelle Darstellung begleitet von einem auditiven Rauschen, 5.) auditive Darstellung begleitet von einem visuellen Rauschen.

Tabelle 1: Genaue Größe der 95% Intervalle(Bootstrap).

Perzentile	Visuell	Auditiv	Bimodal	Visuell+Rauschen	Auditiv+Rauschen
2,5%	0.30	0.64	0.89	0.22	1.04
97,5%	2.47	2.83	3.06	2.27	3.27

4.3 Varianzanalyse

Eine einfaktorielle Varianzanalyse der Sensitivitätsindex-Werte mit den Faktorstufen visuell, auditiv, bimodal, visuell mit Rauschen und auditiv mit Rauschen zeigt einen signifikanten Unterschied zwischen den einzelnen Bedingungen ($F = 4.75, p < 0.005$). Hierbei unterscheiden sich die Faktoren visuell und auditiv mit Rauschen, bimodal und visuell mit Rauschen sowie visuell mit Rauschen und bimodal, auditiv mit Rauschen über einen signifikanten Mittelwertsunterschied. Eine Übersicht hierüber liefern die Abbildungen 15 bis 17. Eine einfaktorielle Varianzanalyse der Reaktionszeiten zeigt hingegen keine signifikanten Unterschiede. ($F=1.4, p= 0.2469$)

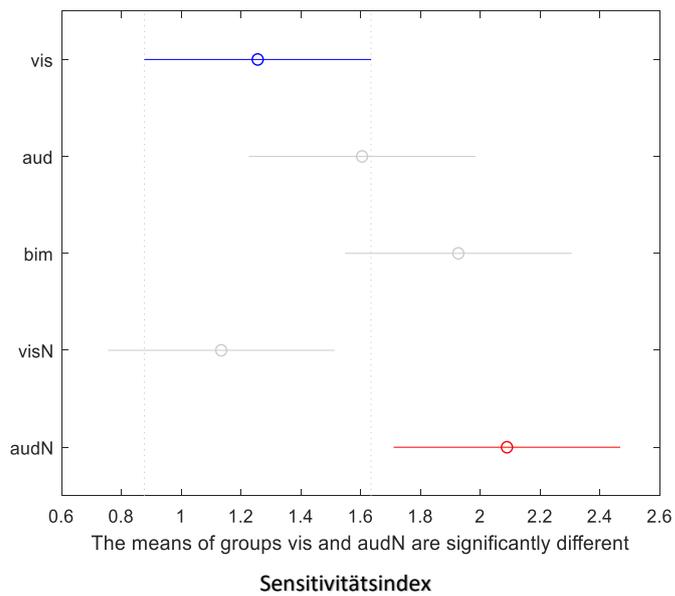


Abbildung 15: Vergleich zwischen der unisensorisch visuellen Bedingung und der auditiven Bedingung mit visuellem Rauschen. Der Punkt gibt den Mittelwert an. Die horizontalen Linien veranschaulichen 95% Konfidenzintervalle.

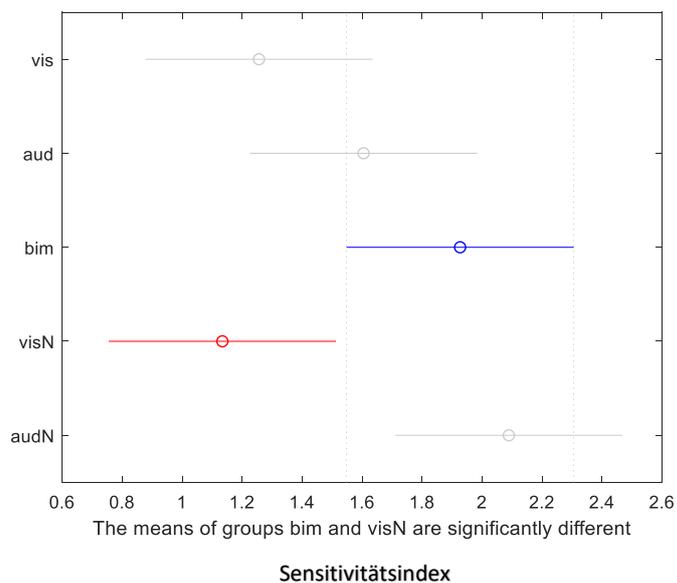


Abbildung 16: Vergleich zwischen der audiovisuellen Bedingung und der visuellen Bedingung, welche von auditivem Rauschen begleitet wird. Der Punkt gibt den Mittelwert an. Die horizontalen Linien veranschaulichen 95% Konfidenzintervalle

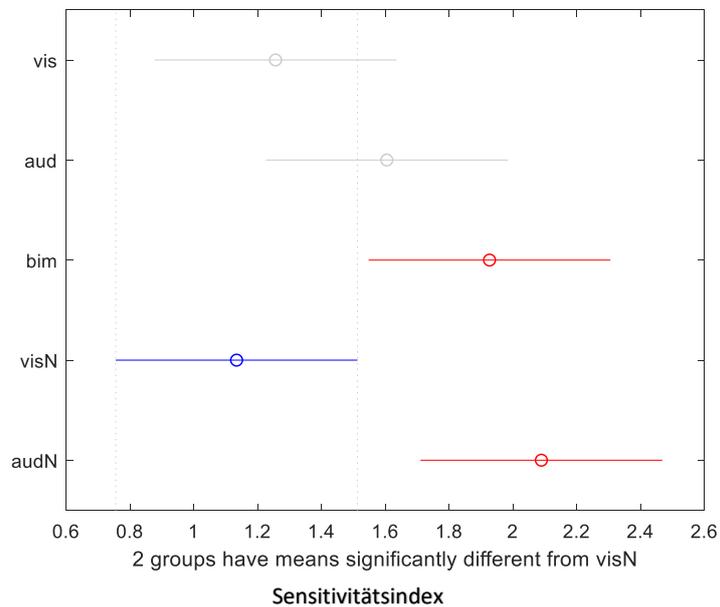


Abbildung 17: Vergleich der Bedingung, welche aus der visuellen Präsentation mit auditivem Rauschen besteht mit der bimodalen- und auditiven Bedingung, die ebenfalls von einem Rauschen in der anderen unisensorischen Modalität begleitet wird. Der Punkt gibt den Mittelwert an. Die horizontalen Linien veranschaulichen 95% Konfidenzintervalle

4.4 Reaktionszeitanalyse

Auch der bedingungsweise Vergleich der Reaktionszeiten liefert bei einer weiteren Betrachtung keine offensichtlichen Unterschiede. So beträgt die Reaktionszeit gemittelt über sämtliche Durchgänge 1.79 Sekunden. In der unisensorisch visuellen Bedingung beträgt diese 1.87, in der unisensorisch auditiven 1.76, in der bimodalen 1.77, in der visuellen mit Rauschen 1.82 und in der auditiven mit visuellem Rauschen 1.71.

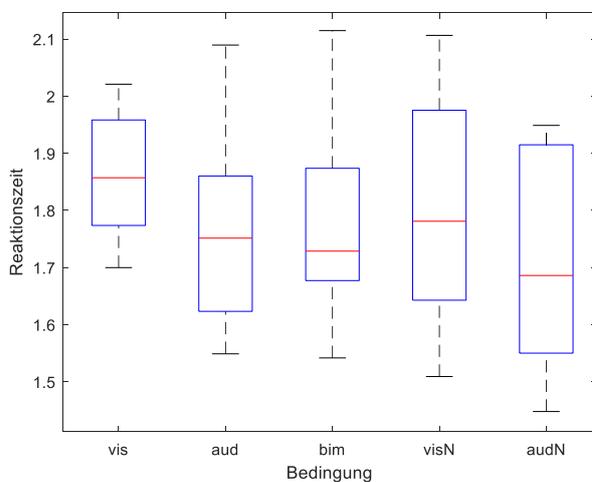


Abbildung 18: Boxplot der Reaktionszeiten. Die rote Linie zeigt den Median. Die blauen Linien zeigen die 25%, 75% Perzentile. Die gestrichelten Linien geben Extremwerte an. Für eine Beschreibung der Bedingungen siehe Abbildung 13

Betrachten wir nun den, wie in den Methoden erklärt, berechneten möglichen bimodalen Reaktionszeitvorteil, stellt sich heraus, dass der Wert des multisensorischen Vorteils mit -1.31 unter null liegt. Somit ist die Reaktionszeit in der bimodalen Bedingung tendenziell langsamer als in der besten unimodalen Bedingung. Wirft man allerdings einen Blick auf die Ergebnisse der einzelnen Versuchspersonen vor einer Mittelung dieser fällt auf, dass sechs der elf beteiligten Teilnehmer über einen bimodalen Reaktionszeitvorteil mit Werten zwischen 0.46 und 5.32 verfügen. Der negative Gesamtwert entsteht aufgrund von wesentlich ausgeprägteren negativen Vorteilen zwischen -3.16 und -6.81 für die Teilnehmer, bei welchen die Reaktionszeit der besten unisensorischen Bedingung die der multisensorischen Reaktionszeit übersteigt.

4.5 Fragebogenauswertung

Bei der Betrachtung der Ergebnisse des Fragebogens zur Einschätzung der subjektiven Wahrnehmung der Stimuli stechen einige Auffälligkeiten deutlich sichtbar hervor. So wird die Aufgabe mit einem Schnitt von $4,18$ auf einer Skala von eins bis fünf als durchaus schwierig wahrgenommen. Werfen wir einen Blick auf den Unterschied zwischen den einzelnen Bedingungen, zeigt sich, dass alle Teilnehmer des Experiments entweder die auditive und / oder bimodale Bedingung als leichter empfanden. So gaben drei Personen die unisensorisch auditive Stimulus Darstellung als einfachsten Teil der Aufgabe an. Ebenso war es für drei Teilnehmer mit der multisensorischen Präsentationsweise. Besonders hervorzuheben ist aber, dass für alle weiteren Teilnehmer der einfachere Aufgabenteil sowohl die auditive als auch die bimodale Bedingung enthält. Keiner der Probanden empfand die visuelle Darstellungsweise als einfachste für die gegebene Aufgabe.

Ähnliche Ergebnisse zeigen sich auch bei der Betrachtung der Differenzierbarkeit der unterschiedlichen Stimuli in den unisensorischen Modalitäten. Der Durchschnitt bezüglich der Differenzierbarkeit der visuellen Stimuli hebt sich mit $2,63$ mit acht Befragten deutlich von den 4.63 bezüglich der auditiven Differenzierbarkeit ab. Das Antwortintervall, in welchem die subjektive Unterscheidbarkeit bewertet wurde, liegt auch zwischen eins und fünf. Dies zeigt sich auch in der Wahl der Modalität, auf welche sich in der audiovisuellen Darstellung fokussiert wurde. So geben sieben Probanden an, sich vermehrt an den Tönen orientiert zu haben, zwei Teilnehmer geben an, ihre Aufmerksamkeit gleichermaßen verteilt zu haben und zwei Probanden konzentrierten sich stärker auf die visuelle Darstellungsform der multisensorischen Stimuli.

Ähnliche Beobachtungen lassen sich ebenso nach einer Betrachtung der Antworten auf die Fragen bezüglich des Rauschens in der nicht Zielmodalität machen. Die Versuchsteilnehmer fühlten sich durch dieses durchschnittlich mit einem Wert von 3.18 (1-5) eingeschränkt. Die Frage nach der Modalität, in welcher das Rauschen als stärker störend wahrgenommen wurde, zeigt eine sehr eindeutige Präferenz. So ist diese für zehn von elf teilnehmenden Personen das Rauschen in seiner auditiven Form. Lediglich ein Teilnehmer empfand das Rauschen als in beiden Bedingungen gleichermaßen störend. Die Frage nach dem Vorhandensein einer Taktik wird von sieben teilnehmenden Personen bejaht. Bei vier von diesen bezieht sich die Strategie auf die Übertragung der rhythmisch tonalen Veränderung der Stimuli auf die visuelle Darstellungsweise. Hierbei wurde unter anderem spezifisch das rhythmische Bewegung des Fußes oder ein Merken der Bilder als Töne genannt. Letzterer Ansatz wurde auch von zwei unterschiedlichen Versuchspersonen angewendet und vermerkt.

5. Diskussion

5.1 Eignung der verwendeten Stimuli

Die Zielsetzung des durchgeführten Experiments ist es in erster Linie zu überprüfen, inwiefern die gewählten Stimuli für eine Arbeitsgedächtnisaufgabe und weitere Untersuchungen der Thematik angelehnt an die Masterthese von Niclas Renner (2018) geeignet sind. Eine hierbei aufkommende Fragestellung war, inwieweit die über Pausen definierten unterschiedlichen Stimuli aufgrund der zügigen Aneinanderreihung, wie in einer n-back-Aufgabenstellung üblich, ineinander verschwimmen oder aber die eindimensional dynamische Art der Stimuli diese grundsätzlich nur schwer unterscheidbar oder merkbar macht.

Der durchschnittliche Sensitivitätsindex d' von 1,44 verdeutlicht, dass die Leistung der Versuchspersonen eindeutig über einem Rateniveau von 50% liegt. Die Aufgabe ist also mit den verwendeten Stimuli durchaus lösbar und es sollte bei ihrer Verwendung nicht zu Höchstwerteffekten (ceiling-Performance) kommen, wie sie in der Arbeit von Niclas Renner in der Detektionsratenbedingung von 99% beobachtet werden konnten. Dies wird weiter durch die berechneten Konfidenzintervalle bestätigt. Auch wenn diese, mit der Bootstrap Methode in 1000000 mithilfe von Simulationen berechneten Versuchsansätzen, geschätzten Intervalle auf den - aufgrund der geringen Datenmenge - ungenauen Ratenschätzern beruhen, geben sie doch einen willkommenen Überblick über die Varianz der Daten. Die große Spanne der Intervalle liegt in der geringen Anzahl an Versuchsdurchläufen in jeder Simulation (48) sowie der Anzahl an Simulationen begründet. Hervorzuheben ist dabei, dass trotzdem keines der gesampelten Intervalle den Wert Null miteinschließt. In 95% aller Intervallberechnungen liegt die wahre Hit-, False Alarm Rate demzufolge innerhalb des Intervalls und ist somit größer als das Rateniveau.

Die einzige Frage, welche in der durchgeführten Untersuchung bezüglich der Eignung der Stimuli auftritt, betrifft die unterschiedliche Differenzierbarkeit der unisensorischen Stimuli untereinander. Dies ist unter anderem anhand der Ergebnisse des Fragebogens verdeutlicht (2,63 in der visuellen Darstellung und 4,63 in der auditiven Darstellungsform). Dementsprechend wurden die schlechtesten Sensitivitätsindex-Leistungen auch in den visuellen Bedingungen erbracht. Dies legt nahe, dass ein Unterschied in der Leistung von Arbeitsgedächtnis und Identifikationsaufgaben bei den unisensorischen Stimuli in deren unterschiedlicher Natur begründet sein kann. Da das Ziel der Arbeit aber nicht in einem Nachweisen von Unterschieden in der Verarbeitung einzelner Sinnesmodalitäten sondern in der Untersuchung der multisensorischen Integration dieser liegt, stellt dieser Umstand nur ein geringes Hindernis da. Ein Integrationsvorteil wird, wie unter anderem in der Reaktionszeituntersuchung gemäß Juan et al (2017), üblicherweise mit einem Vergleich zu der besten unimodalen Bedingung untersucht. Auch die - um die Art der Interaktion zu unterscheidende - oftmals durchgeführte Probability Summation ist bei einer modalitätsabhängigen Differenz unisensorischer Leistungen noch aussagekräftig und anwendbar. Dabei ist es nichtsdestotrotz wichtig darauf zu achten, dass mögliche Effekte zwischen den unisensorischen Bedingungen auf der unterschiedlichen Darstellungsform der Stimuli beruhen können.

5.2 Vergleich der Modalitäten

Vor der Interpretation scheinbarer Unterschiede zwischen den einzelnen Bedingungen ist zuerst hervorzuheben, dass die Ergebnisse auf einer sehr kleinen Datenmenge beruhen und somit individuelle Leistungen einen großen Gesamteinfluss nehmen. Dies zeigt sich zum Beispiel in der Betrachtung der möglichen bimodalen Reaktionszeitvorteile. Es ist des Weiteren hervorzuheben, dass aufgrund der fehlenden Randomisierung der Blöcke Lerneffekte nicht ausgeschlossen werden können. Dieses Problem wird sehr deutlich, wenn die Leistung in dem letzten durchgeführten Block betrachtet wird. In diesem wurde den Versuchsteilnehmern der Stimulus auditiv bei einem gleichzeitigen visuellen Rauschen präsentiert. In dieser Bedingung wurde mit einem Sensitivitätsindex von 2.09 die beste Leistung im Vergleich mit den anderen Bedingungen abgerufen. Die gemessene Leistung in der unimodal auditiven Bedingung betrug 1.60. Auch wenn die Unterschiede zwischen den Sensitivitätsindex-Mittelwerten, wie sich in der Überlappung der Konfidenzintervalle zeigt, keine fundierten Schlüsse zulassen, erscheint gerade dieser Unterschied doch, außerhalb von auf Lerneffekten und Ungenauigkeiten basierenden Erklärungen, verwunderlich, da die Bedingung, welche das Rauschen enthält, als tendenziell schwerer einzustufen ist.

Die in den Fragebögen auftretende, deutlich schwierigere Differenzierbarkeit innerhalb der visuellen Darstellung bietet eine mögliche Erklärung für die Ergebnisse der Varianzanalyse. Dort wird deutlich, dass gerade die Leistungen in den unimodal visuellen Bedingungen sich signifikant von denen der anderen abheben. So ist diese in der unisensorisch visuellen Bedingung deutlich hinter der in der besagten letzten Bedingung zurückgeblieben. Wurde zusätzlich zur visuellen Darstellung noch ein auditives Rauschen abgespielt fallen die erbrachten Leistungen auch signifikant unter die der bimodalen Bedingung. Diese besonders schlechte Leistung bei einem gleichzeitigen Abspielen des Rauschens ist dabei besonders interessant, da dieses ursprünglich zu einer Vergleichbarkeit der Reizenergie zwischen uni- und multisensorischen Bedingungen eingeführt wurde (Werner und Noppeney 2008). Weiterhin ist hierbei auffällig, dass es bei Hinzufügen des visuellen Rauschens nicht zu einem Leistungsabfall kommt. Eine ähnliche Beobachtung machten bereits Delogu et al. (2009) in ihrer Arbeit. Sowohl wenn ein Stimulus in der verbalen oder nonverbalen Bedingung präsentiert wurde, kam es zu einer schlechteren free recall Leistung bei einem Artikulationsverbot. Unabhängig von der Darstellungsform wurde die Information in diesem Fall in einer verbalen Form im Gedächtnis gehalten.

Aufgrund der tonalen rhythmischen Natur der Stimuli könnten diese während der Untersuchung in einer auditiven Repräsentation gemäß der phonologischen Schleife abgespeichert werden, sodass sie mit dem abgespielten auditiven Rauschen interferieren. Dies würde die gemessene Verschlechterung der Leistung sowie die Einvernehmlichkeit der Fragebogenantworten erklären, welche zeigt, dass die Teilnehmer sich wesentlich stärker von dem auditiven Rauschen eingeschränkt fühlten. Auch die mehrfach vorhandene Taktikbeschreibung, die Stimuli über ihren tonalen Charakter zu behalten, lässt sich gut mit diesem Erklärungsansatz in Einklang bringen. Auch wenn die Untersuchung dem zufolge ein Indiz für den von Delogu et al. (2009) verwendeten Erklärungsansatz bietet, ist die durchgeführte Aufgabenstellung nicht auf eine Untersuchung der Repräsentationsart, in welcher unisensorische Informationen im Arbeitsgedächtnis zwischengespeichert werden, ausgelegt. Hierzu wäre eine Überprüfung mit zwei unterschiedlichen Konstellationen von Stimuli sinnvoll. Eine Konstellation wären die hier gewählten Paare, deren intrinsische Natur einen eher tonalen Charakter widerspiegelt. Diese

werden in einer Arbeitsgedächtnisaufgabe mit einem Stimuluspaar verbalen Ursprungs unter Rauschen in der anderen Modalität verglichen.

Betrachten wir nun das mögliche Vorhandensein eines multisensorischen Integrationsvorteils. Hierfür lässt sich in der durchgeführten Untersuchung kein direkter Beleg, wie er zum Beispiel von Hardiess, Erhardt und Mallot (2015) erbracht wurde, finden. So ist die Leistung der Versuchspersonen in der audiovisuellen Bedingung nicht signifikant besser als die in einer der unisensorischen Bedingungen. Ebenso lässt sich auch der bereits von Brunetti, Indraccolo, Mastroberardino, Spence, und Santangelo (2017) gefundene multisensorische Reaktionszeitvorteil nicht belegen. Ursache hierfür ist vermutlich die zu kleine Datenmenge, da jede Bedingung nur 48-mal bei elf Versuchspersonen abgefragt wurde. Ebenso ist hervorzuheben, dass in dem durchgeführten Experiment noch keine degenerierten Stimuli gemäß dem Prinzip der inversen Effektivität verwendet wurden, sodass es nur schwer möglich ist, unter dem verwendeten Versuchsansatz Integrationsprinzipien gemäß der Superadditivität und Additivität zu belegen. Des Weiteren lässt sich somit der Einfluss eines potentiellen multisensorischen Wahrnehmungsvorteils, welcher von den Degradierungsstufen abhängig ist, auf das Arbeitsgedächtnis nicht untersuchen.

5.3 Ausblick

Um diese Phänomene der multisensorischen Integration weiter zu untersuchen, empfiehlt es sich, erneut einen ähnlichen Versuchsaufbau wie ihn Niclas Renner (2018) in seiner Arbeit nutzte zu verwenden. Um die Stimuli in Degradierungsstufen einzuteilen, müssen diese erst in einer Detektionsaufgabe zur Schätzung der Detektionsraten mittels einer psychometrischen Funktion ermittelt werden. Hierbei ist die Orientierung an den Erkennungsraten von 50%, 75% und 99% als durchaus sinnvoll. Dabei bietet sich eine wahre two-alternative-forced-choice-Aufgabenstellung an, um den zuvor beobachteten Bias bei Unwissenheit zu vermeiden. In dieser werden dem Versuchsteilnehmer in jedem Durchgang zwei unterschiedliche Reize präsentiert. Einer enthält den verwendeten Stimulus in unterschiedlichen Signal-zu-Rausch Verhältnissen, während der andere präsentierte Reiz nur Rauschen enthält. Die Aufgabe ist es hierbei den Stimulus enthaltenden Reiz zu identifizieren. Dabei ist zu beachten, dass in der visuellen Darstellung, angepasst an die auditive Präsentationsform, die zwei Reize ebenfalls zeitlich hintereinander abgespielt werden müssen. Mit diesem Ansatz lassen die sich Schwellen in jeweils einem Versuch für die unisensorischen Bedingungen schätzen. Dabei bietet es sich an, anstelle der zuvor verwendeten adaptiven psy-marginal Methode, eine konstante Methode mit einer festen Anzahl an unterschiedlichen Signal-zu-Rausch Leveln zu nehmen. Dies dient der Stabilisation des Verfahrens zum Erstellen der psychometrischen Funktion. Dabei wäre es interessant zu beobachten, wie stark die intrapersonellen und interpersonellen Unterschiede bei den einzelnen Teilnehmern ins Gewicht fallen. Dieses Wissen könnte nämlich weitere Untersuchungen dieser Art vereinfachen.

Mit den auf diese Weise geschätzten Detektionsrate lässt sich anschließend erneut in einer two-alternative-forced-choice Aufgabe das Vorhandensein eines perzeptuellen multisensorischen Integrationsvorteils untersuchen. Dabei wird die audiovisuelle Wahrnehmungsleistung mit der besten

unisensorischen verglichen. Außerdem bietet sich hierbei die Möglichkeit, wie in der vorigen Arbeit von Niclas Renner, die Natur eines möglichen Wahrnehmungsvorteils weiter zu beleuchten.

Unter den angenommenen Voraussetzungen, dass die zwei Modalitäten der Stimuli getrennt voneinander mit der gleichen Gewichtung bearbeitet werden sowie, dass das richtige Erkennen des Stimulus in einer Modalität für eine richtige Antwort ausreicht, lässt sich mittels einer Form der Probability Summation (Treisman 1998) die bimodale Erkennungsrate auf Basis einer unabhängigen Verarbeitung ihrer Bestandteile berechnen. Die hierbei verwendete Formel ist $P_{va} = P_v + P_a - P_v \times P_a$. Liegt die gemessene Wahrnehmungsleistung über dieser Schwelle, lässt sich die Annahme tätigen, dass die perzeptuelle Verbesserung tatsächlich auf eine frühe Form der multisensorischen Integration zurückzuführen ist und nicht mithilfe eines Race-Model Ansatzes erklärbar ist.

Nach den Aufgaben zur Detektionsratenbestimmung und zum Untersuchen eines Wahrnehmungsvorteils lässt sich mit den Stimuli erneut ein potentieller Vorteil multisensorischer Darstellung im Arbeitsgedächtnis untersuchen. Hierzu bietet sich gemäß der Arbeit von Niclas Renner (2018) erneut eine two-back-Aufgabe an. Besonders, da die Eignung der verwendeten Stimuli bei einer solchen bereits überprüft wurde. Hierbei sind die Degradierungsstufen der Stimuli in das Versuchsdesign mit einzubeziehen, sodass ein 3*3 Faktor-Design entsteht. Die unterschiedlichen Faktorstufen bilden sich aus den Erkennungsraten sowie den verwendeten Modalitäten. Da eine Annäherung der Reizenergie zwischen den uni- und multisensorischen Bedingungen mithilfe eines zusätzlichen, parallel abgespielten Rauschens zur unimodalen Stimulus Präsentation die Aufgabe subjektiv in einer Bedingung deutlich gestört hat, ist diese Methode – zumindest ohne weitere Untersuchungen bezüglich der Natur des störenden Einflusses - vorerst nicht zu empfehlen.

Da bei der Überprüfung der Stimuli auf eine Eignung für eine Arbeitsgedächtnisaufgabe die einzelnen Reize nicht verrauscht vorlagen und kein Höchstwerteffekt (Decken Performance) beobachtet werden konnte, scheint die Aufgabe mit der Verwendung der Stimuli schwerer geworden zu sein. Dies zeigt sich unter anderem in der Bewertung der Schwierigkeit durch die Versuchspersonen von 4,18. Demzufolge bietet sich die Überlegung an, die verwendeten Stimuli auf 1-3-5, 1-2-5 und 1-4-5 zu reduzieren. Dies würde einem möglichen Einfluss der subjektiv größeren Ähnlichkeit zwischen den Stimuli 1-3-5 und 2-3-4 vorbeugen, während der Schwierigkeitsgrad der Aufgabe etwas leichter wird. Dieser wird sich durch die unterschiedlichen Degenerierungsgrade der Stimuli weiter erschweren. Somit könnte eine möglicherweise zu schwere und für die Teilnehmer frustrierende Aufgabe verhindert werden. Wenn nicht erscheint es sinnvoll die Stimuli zusätzlich nochmal auf ihre Differenzierbarkeit zu untersuchen. So ließe sich zum Beispiel ein Einfluss der scheinbaren Ähnlichkeit zwischen den Stimuli 1-3-5 und 2-3-4 mithilfe eines Betrachtens der False Alarm Rate in Durchläufen, bei welchen der jeweils andere zwei Stellen zuvor präsentiert wurde, betrachten.

5.4 Konklusion

Der Zweck dieser Arbeit war das Finden und Überprüfen von abstrakt kongruenten Stimulus-Paaren, welche sich für weitere Untersuchungen bezüglich multisensorischer Integrationsvorteile im Arbeitsgedächtnis eignen. Da bei der angedachten weiteren Untersuchung der Thematik zusätzlich, gemäß dem Prinzip der inversen Effektivität, diese mit unterschiedlichen Signal-zu-Rausch Verhältnissen degeneriert werden, müssen die verwendeten Stimuli anhand von einfachen und markanten Eigenschaften differenzierbar sein.

Die Ergebnisse zeigen, dass die in dieser Arbeit gefundenen und überprüften Stimuli gut für die Verwendung in einer Arbeitsgedächtnisaufgabe geeignet sind. Des Weiteren sind sie derartig konzipiert, dass auch eine Detektionsaufgabe zur Berechnung unterschiedlicher Wahrnehmungsschwellen gut umsetzbar scheint. Dies wurde jedoch noch nicht explizit empirisch überprüft. Da das Angleichen der Signalstärke zwischen den unisensorischen und multisensorischen Bedingungen die Aufgabe auffällig erschwerte, sollte dies in den weiteren Experimenten mit Vorsicht bedacht werden.

Es konnte in der Untersuchung keines der zuvor beobachteten Phänomene multisensorischer Integration wiedergefunden werden. Hierbei ist hervorzuheben, dass das experimentelle Setup dafür aufgrund eines Fehlens der Degradierungsstufen sowie dem Absehen von einer Randomisierung der Blöcke nicht ausgelegt war. Ebenso ist die erhobene Datenmenge zu gering, um sichere Schlüsse zuzulassen.

Literaturverzeichnis

- Angelaki, D. E., Gu, Y., & DeAngelis, G. C. (2009). Multisensory integration: psychophysics, neurophysiology, and computation. *Current opinion in neurobiology*, 19 (4), 452–458.
- Alais, D., Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, 14(3), 257–262,
- Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in cognitive sciences*, 4 (11), 417–423.
- Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial vision*, 10, 433–436
- Brunetti, R., Indraccolo, A., Mastroberardino, S., Spence, C., & Santangelo, V. (2017). The impact of cross-modal correspondences on working memory performance. *Journal of Experimental Psychology: Human Perception and Performance*, 43 (4), 819.
- Colonus, H., Wolff, F. H., Diederich, A. (2017), Trimodal Race Model Inequalities in Multisensory Integration: 1. Basics. *Frontiers in Psychology*, 8, 1141
- Cox, D., Hong, S. W. (2015). Semantic-based crossmodal processing during visual suppression. *Frontiers in Psychology*, 6,722.
- Delogu F., Raffone A., Belardinelli M. O. (2009). Semantic encoding in working memory: is there a (multi) modality effect?. *Memory*, 17, 655–663
- Efron, B. Tibshirani, R. (1993). An Introduction to the Bootstrap. Chapman and Hall/CRC
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415 (6870), 429.
- Foxe J. J., Wylie G. R., Martinez A., Schroeder C. E., Javitt D. C., Guilfoyle D., et al. . (2002). Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J. Neurophysiol.* 88, 540–543.
- Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of cognitive neuroscience*, 11 (5), 473–490.
- Gondan, M., Minakata, K., A tutorial on testing the race model inequality. *Attention, Perception, & Psychophysics*, 78 (3), 723–735.
- Goolkasian, P., & Foos, P. W. (2005). Bimodal format effects in working memory. *The American journal of psychology*, 61–78.
- Hardiess, G., Erhardt, S., & Mallot, H. (2015). Bimodal perceptual integration facilitates representation in working memory. *Perception*, 44, 341–341
- Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *Journal of Experimental Psychology*, 63(3), 289-293
- Juan, C., Cappe, C., Alric, B., Roby, B., Gilardeau, S., Barone, P., Girard, P. (2017). *Plos one*, 12(2), e0172480. doi: 10.1371.

- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390 (6657), 279.
- Mastroberardino, S., Santangelo, V., Botta, F., Marucci, F. S., Belardinelli, M. O. (2008). How the bimodal format of presentation affects working memory: an overview. *Cognitive Processing*, 9(1), 69-76.
- Oliver, C. N. L. (2008). Interactions between visual working memory and visual attention. *Frontiers in Bioscience*, 13, 1182-91
- Quak, M., London, R. E., & Talsma, D. (2015). A multisensory perspective of working memory. *Frontiers in human neuroscience*, 9, 197.
- Renner, N. (2018). Multisensory integration in the working memory - the interplay between perception & working memory in n-back performance (Masterthesis)
- Saults J. S., Cowan N. (2007). A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *J. Exp. Psychol. Gen.* 136, 663–684
- Schneider, T. R., Engel, A. K., Debener, S. (2008). Multisensory identification of natural objects in a two-way crossmodal priming paradigm. *Experimental Psychology*, 55(2), 121-32.
- Schröter, H., Frei, L. S., Ulrich, R., Miller, J. (2009). The auditory redundant signal effect: an influence of number of stimuli or number of percepts?. *Attention, perception & psychophysics*, 71(6), 1375-84.
- Smith, E. E., Jonides, J., Koeppe, R. K., Awh, E. (1995). Spatial versus Object Working Memory: PET Investigations. *Journal of Cognitive Neuroscience*, 7(3), 337-56.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers*, 31 (1), 137– 149.
- Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, 9 (4), 255–266.
- Thompson, V. A., & Paivio, A. (1994). Memory for pictures and sounds: Independence of auditory and visual codes. *Canadian Journal of Experimental Psychology*, 48 (3), 380.
- Treisman, M. (1998). Combining information: probability summation and probability averaging in detection and discrimination. *Psychological Methods*, 3, 252–265.
- Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428 (6984), 748.
- Wahn, B., & König, P. (2017). Is attentional resource allocation across sensory modalities task-dependent? *Advances in cognitive psychology*, 13 (1), 83.
- Werner, S., & Noppeney, U. (2009). Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. *Cerebral cortex*, 20 (8), 1829–1842.
- Xie, Y., Xu, Y., Bian, C., & Li, M. (2017). Semantic congruent audiovisual integration during the encoding stage of working memory: an erp and sloreta study. *Scientific reports*, 7 (1), 5112.
- Zimmer, H. D. (2008). Visual and spatial working memory: From boxes to network. *Neuroscience and Biobehavioral Reviews*, 32, 1373-1395.

Appendix

Fragebogen

Name:

Zum Abschluss würde ich Sie bitten noch einen kurzen Fragebogen auszufüllen.

1. Schätzen Sie bitte die Schwierigkeit der heutigen Arbeitsgedächtnisaufgabe ein?

leicht

schwer



**2. Verwendeten Sie eine besondere Taktik um die Aufgabe besonders gut zu bewältigen?
Wenn dies der Fall ist, beschreiben Sie diese bitte in einigen knappen Worten.**

Nein

ja _____

3. Empfanden Sie ein oder mehrere der unten genannten Bedingungen als leichter?

Alle Blöcke hatten die gleiche Schwierigkeit

Visuell

Auditiv

Bimodal

4. Konzentrierten Sie sich in dem bimodalen Block stärker auf die Bilder oder auf die Töne?

Nur Töne

vermehrt Töne

beides gleichmäßig

vermehrt Bilder

Nur Bilder



5. In welcher Bedingung wurden Sie am stärksten durch das Rauschen eingeschränkt?

- Visuell
- Auditiv
- In beiden Bedingungen in gleichen Maßen

6. Wie sehr hat das Rauschen in der jeweils anderen Modalität Ihre Aufgabe erschwert?

Gar nicht sehr stark



7. Wie gut konnten Sie die Stimuli in ihrer visuellen Darstellung unterscheiden ?

Nur schwer sehr gut



8. Wie gut konnten Sie die Stimuli in ihrer auditiven Darstellung unterscheiden ?

Nur schwer sehr gut

