



FAKULTÄT FÜR
NATURWISSENSCHAFTEN

Audiovisuelle Bewegungsvorhersage im dreidimensionalen Raum

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

genehmigt durch die Fakultät für Naturwissenschaften
der Otto-von-Guericke-Universität Magdeburg

von Sandra Dittrich, M. Sc.

geboren am 22. September 1987 in Bernburg

Gutachter:

Prof. Dr. Tömme Noesselt

Prof. Dr. Uta Noppeney

eingereicht am: 14. Juni 2018

verteidigt am: 22. Januar 2019

Inhalt

Zusammenfassung	8
Abstract	10
1 Gegenstand der Arbeit	11
2 Theoretischer Hintergrund	12
2.1 Audiovisuelle Bewegungswahrnehmung	12
2.2 Bewegungsvorhersage und verdeckte Bewegungswahrnehmung	14
2.2.1 Unimodale visuelle Bewegungsvorhersage und verdeckte Bewegungs- wahrnehmung	14
2.2.2 Bimodale audiovisuelle Bewegungsvorhersage und verdeckte Bewe- gungswahrnehmung	17
2.3 Einfluss der Aufgabenstellung	19
2.4 Verwendung von quasi-natürlichem Stimulusmaterial	22
3 Ziele der Arbeit	25
3.1 Fragestellungen	25
3.2 Übersicht Experimente	26
4 Allgemeine Methoden	28
4.1 Probanden	28
4.2 Design	28
4.3 Geräte und Stimuli	31
4.4 Durchführung	32
4.5 Auswertung	33
4.5.1 Datenaufbereitung	33
4.5.2 Reaktionszeit-Analyse	33
5 Zeitliche Bewegungsvorhersage (Detektion und Extrapolation)	34
5.1 Experiment 1: Detektion	34
5.1.1 Methoden	34
5.1.2 Ergebnisse	34
5.1.3 Diskussion	35
5.2 Experiment 2: Extrapolation	37
5.2.1 Methoden	37
5.2.2 Ergebnisse	38
5.2.3 Post-hoc-Betrachtung interindividueller Unterschiede	38
5.2.4 Diskussion	45
5.3 Kontrollexperiment: Bestimmbarkeit der Tonrichtung	47
5.3.1 Methoden	47

5.3.2	Ergebnisse	49
5.3.3	Diskussion	49
6	Variation der visuellen Stimulation (Diskrimination)	50
6.1	Experiment 3: Diskrimination des Bewegungsverlaufs	50
6.1.1	Methoden	50
6.1.2	Ergebnisse	50
6.1.3	Diskussion	51
6.2	Experiment 4: Erhöhung der visuellen Schwierigkeit	53
6.2.1	Methoden	53
6.2.2	Ergebnisse	53
6.2.3	Diskussion	54
6.3	Experiment 5: Vereinfachtes zweidimensionales Stimulusmaterial	57
6.3.1	Methoden	57
6.3.2	Ergebnisse	57
6.3.3	Diskussion	59
6.4	Post-hoc-Betrachtung der audiovisuellen Diskriminationsexperimente	60
6.4.1	Bayes-Faktor-Analyse	60
6.4.2	Interindividuelle Unterschiede in der Nutzung visueller und auditorischer Informationen	60
6.4.3	Interindividuelle Unterschiede in der Stabilität der Informationsnutzung	71
6.4.4	Zwischenfazit	72
7	Einfluss von Aufgabenstellung und visueller Gestaltung ohne Tonstimulation (Experiment 6)	75
7.1	Methoden	75
7.2	Ergebnisse	75
7.3	Diskussion	77
8	Allgemeine Diskussion	78
8.1	Zusammenfassung der wichtigsten Ergebnisse	78
8.2	Einfluss audiovisueller Stimulation bei zeitlicher Bewegungsvorhersage	78
8.3	Einfluss audiovisueller Stimulation bei räumlicher Diskrimination	82
8.4	Einfluss der Aufgabenstellung	88
8.5	Individuelle Unterschiede	91
8.6	Quasi-natürliches Stimulusmaterial und ökologische Validität	93
9	Ausblick	98
10	Schlussfolgerungen	99
	Literatur	100
A	Reaktionszeit-Differenzen für die Nutzergruppeneinteilung	121

B	Statistische Kennwerte für die Varianzanalysen der audiovisuellen Diskriminationsexperimente	133
C	Analyse korrekter Antworten der audiovisuellen Diskriminationsexperimente	140
D	Blockweise Betrachtung der Reaktionszeit-Differenzen	147
E	Vergleich korrekter Antworten der Nutzergruppen der audiovisuellen Diskriminationsexperimente	151
	Eigenständigkeitserklärung	153

Abbildungsverzeichnis

1	Übersicht über die sechs durchgeführten Hauptexperimente.	29
2	Visuelle Bahnverläufe, auditorische Stimulation und inkongruente Bedingungen	30
3	Experimentelles Set-Up im Labor	32
4	Beispielhaftes Bildmaterial aus Experiment 1 und 3.	35
5	Reaktionszeit in Abhängigkeit von der visuellen und auditorischen Stimulation sowie der Präsentationsdimension und visuellen Stimulation in Experiment 1 . .	36
6	Beispielhaftes Bildmaterial aus Experiment 2.	38
7	Absolute Reaktionszeit-Abweichung in Abhängigkeit von der Schätzdistanz und der auditorischen Stimulation in Experiment 2	39
8	Reaktionszeit-Abweichung sowie die Streuung der Reaktionszeit-Abweichung in Abhängigkeit von der Schätzdistanz und der auditorischen Stimulation in Experiment 2 für Nahtonutzer	42
9	Reaktionszeit-Abweichung sowie die Streuung der Reaktionszeit-Abweichung in Abhängigkeit von der Schätzdistanz und der auditorischen Stimulation in Experiment 2 für Ferntonnutzer	44
10	Abfragedisplays für die Richtung der Tonbewegung im Kontrollexperiment . . .	48
11	Reaktionszeit in Abhängigkeit von der visuellen und auditorischen Stimulation in Experiment 3	52
12	Beispielhaftes Bildmaterial aus Experiment 4	54
13	Reaktionszeit in Abhängigkeit von der visuellen und auditorischen Stimulation in Experiment 4	55
14	Beispielhaftes Bildmaterial aus Experiment 5 und 6	58
15	Reaktionszeit in Abhängigkeit von der visuellen und auditorischen Stimulation in Experiment 5	59
16	Reaktionszeit in Abhängigkeit von der visuellen und auditorischen Stimulation für die Beid- und Bildnutzer in Experiment 3 (Post-hoc-Betrachtung)	68
17	Reaktionszeit in Abhängigkeit von der visuellen und auditorischen Stimulation für die Beid- und Bildnutzer in Experiment 4 (Post-hoc-Betrachtung)	69
18	Reaktionszeit in Abhängigkeit von der visuellen und auditorischen Stimulation für die Beid- und Bildnutzer in Experiment 5 (Post-hoc-Betrachtung)	70
19	Reaktionszeit in Abhängigkeit von der visuellen Stimulation sowie der Präsentationsdimension für Detektion und Diskrimination in Experiment 6	76

Tabellenverzeichnis

1	Anzahl der Probanden für die jeweilige Nutzergruppeneinteilung für gerade und ungerade Blöcke in Experiment 2	40
2	Bayes-Faktoren für die audiovisuellen Diskriminationsexperimente	60
3	Anzahl der Probanden für die jeweilige Nutzergruppeneinteilung für gerade und ungerade Blöcke in den audiovisuellen Diskriminationsexperimenten	65
4	Vergleich der Ergebnisse für die audiovisuellen Diskriminationsexperimente für die Analyse gerader und ungerader Blöcke	66
A1	Auditorische Reaktionszeit-Differenzen für beide Schätzdistanzen in Experiment 2 auf Grundlage gerader Blöcke	121
A2	Auditorische Reaktionszeit-Differenzen für beide Schätzdistanzen in Experiment 2 auf Grundlage ungerader Blöcke	122
A3	Auditorische und visuelle Reaktionszeit-Differenzen in Experiment 3 auf Grundlage gerader Blöcke	123
A4	Auditorische und visuelle Reaktionszeit-Differenzen in Experiment 3 auf Grundlage ungerader Blöcke	124
A5	Auditorische und visuelle Reaktionszeit-Differenzen in Experiment 4 auf Grundlage gerader Blöcke	125
A6	Auditorische und visuelle Reaktionszeit-Differenzen in Experiment 4 auf Grundlage ungerader Blöcke	126
A7	Auditorische und visuelle Reaktionszeit-Differenzen in Experiment 5 auf Grundlage gerader Blöcke	127
A8	Auditorische und visuelle Reaktionszeit-Differenzen in Experiment 5 auf Grundlage ungerader Blöcke	128
A9	Vergleich der Einteilungskriterien für die Nutzergruppen in Experiment 2 für die Kategorisierung auf Grundlage gerader vs. ungerader Blöcke	129
A10	Vergleich der Einteilungskriterien für die Nutzergruppen in Experiment 3 für die Kategorisierung auf Grundlage gerader vs. ungerader Blöcke	130
A11	Vergleich der Einteilungskriterien für die Nutzergruppen in Experiment 4 für die Kategorisierung auf Grundlage gerader vs. ungerader Blöcke	131
A12	Vergleich der Einteilungskriterien für die Nutzergruppen in Experiment 5 für die Kategorisierung auf Grundlage gerader vs. ungerader Blöcke	132
B1	Statistische Kennwerte der Varianzanalyse Dimension \times Visuell \times Auditorisch für die Beidnutzer in Experiment 3	133
B2	Statistische Kennwerte der Varianzanalyse Dimension \times Visuell \times Auditorisch für die Bildnutzer in Experiment 3	134
B3	Statistische Kennwerte der Varianzanalyse Dimension \times Visuell \times Auditorisch für die Beidnutzer in Experiment 4	135
B4	Statistische Kennwerte der Varianzanalyse Dimension \times Visuell \times Auditorisch für die Bildnutzer in Experiment 4	136

B5	Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Beidnutzer in Experiment 5	137
B6	Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Bildnutzer in Experiment 5	138
B7	Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch × Experiment für die drei audiovisuellen Diskriminationsexperimente	139
C1	Vergleich der Ergebnisse für die audiovisuellen Diskriminationsexperimente für die Analyse gerader und ungerader Blöcke (Auswertung korrekter Antworten) .	140
C2	Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Beidnutzer in Experiment 3 (Korrekte Antworten)	141
C3	Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Bildnutzer in Experiment 3 (Korrekte Antworten)	142
C4	Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Beidnutzer in Experiment 4 (Korrekte Antworten)	143
C5	Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Bildnutzer in Experiment 4 (Korrekte Antworten)	144
C6	Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Beidnutzer in Experiment 5 (Korrekte Antworten)	145
C7	Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Bildnutzer in Experiment 5 (Korrekte Antworten)	146
D1	Anzahl Blöcke mit positiver bzw. negativer auditorischer Reaktionszeit-Differenz für nahe und ferne Distanzen in Experiment 2	147
D2	Anzahl Blöcke mit positiver bzw. negativer auditorischer und visueller Reaktionszeit-Differenz in Experiment 3	148
D3	Anzahl Blöcke mit positiver bzw. negativer auditorischer und visueller Reaktionszeit-Differenz in Experiment 4	149
D4	Anzahl Blöcke mit positiver bzw. negativer auditorischer und visueller Reaktionszeit-Differenz in Experiment 5	150
E1	Antwortraten für die vier Nutzergruppen in den drei audiovisuellen Diskriminationsexperimenten	151
E2	Statistische Kennwerte des Kruskal-Wallis-Tests für den Vergleich der Nutzergruppen in den audiovisuellen Diskriminationsexperimenten	152

Zusammenfassung

Die Fähigkeit zur Bewegungsvorhersage ist essentiell für unseren Alltag, z. B. um uns sicher im Straßenverkehr zu bewegen. Die bisherige Ergebnislage dazu gestaltet sich inkonsistent, wobei der Fokus auf die visuelle Modalität, verschiedene Aufgabenstellungen sowie divergentes Stimulusmaterial Gründe dafür darstellen können. Daher erfolgte in dieser Arbeit die systematische Untersuchung potentieller Einflussfaktoren der Bewegungsvorhersage. In sechs Experimenten bewegte sich eine Kugel in einer Box auf den Betrachter zu, die Kugel wurde kurzzeitig von einer Brücke verdeckt und erschien im Anschluss auf einer visuell passenden oder nicht passenden Bahn wieder. Die Kugelbewegung konnte unimodal (visuell) präsentiert oder von einem in der Tiefe moduliertem weißen Rauschen begleitet werden, das sich entweder in die gleiche oder entgegengesetzte Richtung bewegte. Die Probanden reagierten mit Tastendruck, sobald die Kugel wiedererschien (zeitliche Detektion, Experiment 1); schätzten, wann die Kugel eine bestimmte Position erreichen würde (zeitliche Extrapolation, Experiment 2) oder unterschieden, ob die Kugel nach der Verdeckung einer passenden oder nicht passenden Bahn folgte (räumliche Diskrimination, Experimente 3, 4 und 5). Für die Untersuchung des Einflusses realitätsnaher Stimuli wurde in den Experimenten 1 bis 4 eine dreidimensionale (3D) Darstellung (echt-3D, mit Disparitätsinformationen) mit einer pseudo-3D-Version (ohne Disparitätsinformationen) verglichen. Im 5. Experiment wurde der echt-3D-Variante eine vereinfachte zweidimensionale (2D) Version (einfach-2D) gegenübergestellt, die weder Disparitäts- noch Schatten- oder Perspektivenhinweise enthielt. In Experiment 6 erfolgte ein Aufgabenvergleich (Detektion vs. Diskrimination) für die einfach-2D- und echt-3D-Präsentation ohne Tonstimulation.

Die Daten sprechen dafür, dass in den zeitlichen Aufgaben (Detektion und Extrapolation) zusätzliche Toninformationen genutzt werden, jedoch offenbarte sich die Tonnutzung während der Extrapolation erst in einer crossvalidierten Post-hoc-Analyse der Daten. Dabei unterschieden sich die Probanden darin, zu welchem Zeitpunkt der Schätzung sie von den Tönen profitieren konnten. Für die drei audiovisuellen Diskriminationsexperimente ergab die klassische Overall-Analyse diffuse bis keine signifikanten Effekte. Bayes-Analysen und ebenfalls crossvalidierte Post-hoc-Analysen deuten auf interindividuelle Unterschiede in der Nutzung visueller Informationen in der räumlichen Diskrimination hin, wobei die Mehrheit der Probanden von visuell passenden Bahnen profitierte. Für die Nutzung auditorischer Informationen konnten hingegen keine Hinweise auf ein konstantes Nutzungsverhalten gefunden werden, vielmehr scheinen hier zumeist Fluktuationen innerhalb der Probanden vorzuliegen. Die Präsentationsdimension beeinflusste das Verhalten der Probanden dahingehend, dass während der Detektion der visuelle Erleichterungseffekt ohne Disparitätsinformationen größer ausgeprägt war, in Experiment 6 hingegen war dieser für die echt-3D-Präsentation größer als in der einfach-2D-Darbietung. In den audiovisuellen Diskriminationsexperimenten konnte kein Einfluss der Präsentationsdimension gefunden werden, wobei dies vermutlich auf die Tonstimulation zurückzuführen ist. Die inkohärenten Toneinflüsse deuten auf eine nicht erfolgte Integration von Ton- und Bildinformationen hin, die dazu führte, dass die Probanden die Töne

eher als störend empfanden und die Tonstimulation dadurch ein Rauschen in den Daten erzeugte. Gründe für diese ausbleibende Integration werden diskutiert. Insgesamt verdeutlichen die Ergebnisse, dass Aufgabenanforderungen, individuelle Unterschiede und der Realitätsgrad des Stimulusmaterials die audiovisuelle Bewegungsvorhersage modulieren. In zukünftigen Untersuchungen sollten daher nicht nur Unterschiede zwischen den Probanden bedacht, sondern auch die Interaktionen mit einer alltagsnahen Stimulusgestaltung detaillierter untersucht werden.

Abstract

Predicting motion is an essential skill in our everyday life (e. g. in road traffic). Previous studies failed to find consistent results. Possible reasons might be that they focused predominantly on the visual modality, used different tasks as well as divergent stimulus material. Therefore, potential factors governing motion prediction were tested in this study. In a series of six psychophysical experiments a ball followed a trajectory towards the observer and temporarily disappeared behind an occluding element. The ball could reappear on a visually congruent or incongruent path and its movement could be accompanied by a continuously moving white noise. The additional sound moved in the same or in a different direction than the ball movement. A unimodal visual-only condition was also presented. Participants indicated when the ball reappeared after occlusion (temporal detection, experiment 1), or extrapolated when the ball would have reached a specified target position (temporal extrapolation, experiment 2). In experiments 3, 4, and 5 they discriminated whether the ball reappeared on a visually congruent or incongruent path (spatial discrimination). All experiments were conducted varying the factor visual depth to compare the influence of realistic stimulus material. Stimuli were presented on a stereoscopic screen with disparity cues in three-dimensional (3D) space (real-3D) and were contrasted with a pseudo-3D version which contained no disparity information (experiment 1 – 4). For experiment 5 a simplified two-dimensional (2D) version with missing disparity as well as no cast shadows and no perspective cues served as a comparison. In experiment 6 two tasks (detection vs. discrimination) were compared for real-3D and simplified-2D presentation without auditory stimulation.

Participants benefitted from additional sounds during detection. A post-hoc cross-validation procedure revealed that this was also the case for temporal extrapolation. However, participants differed in the point of time at which they profited from sounds. Most of them performed better in auditory conditions for one extrapolation distance only while their performance was impaired for the second distance. For audiovisual discrimination experiments solely diffuse or non-significant effects were found during classical overall analyses. Bayesian analyses and similar post-hoc cross-validation procedures like in experiment 2 revealed that most participants profited from visually congruent paths. However, auditory stimulation still modulated behaviour non-systematically. Auditory impact seemed to fluctuate during experimental sessions. Furthermore, dimensional presentation affected visual facilitation in two experiments. During experiment 1 visual facilitation was more pronounced during pseudo-3D than real-3D sessions, whereas in experiment 6 this effect was more distinctive for real-3D than simplified-2D stimulation. No effect of stimulus material was found in audiovisual discrimination experiments which might be due to a missing integration of visual and auditory input. Therefore, participants may have misinterpreted auditory stimuli as non-useful noise and could probably be distracted by sounds. Reasons for non-existing multimodal integration are discussed. Together, the results of these experiments indicate that task demands, individual differences, and a realistic experimental environment modulate audiovisual motion prediction and need to be considered in future research.

1 Gegenstand der Arbeit

Die Vorhersage von Bewegungen spielt nicht nur in der Natur eine Rolle, wenn das Naherungsverhalten von Fressfeinden eingeschatzt werden soll. Auch in unserem alltaglichen Umfeld ist sie bedeutsam, beispielsweise wenn wir einen Gegenstand fangen oder im Straenverkehr antizipieren, wie sich die Verkehrsteilnehmer verhalten werden. Zu welchem Zeitpunkt man die Spur auf der Autobahn wechseln sollte oder wann die Strae uberquert werden kann, sind dabei nur zwei aufkommende Fragen. Die Fahigkeit zur Bewegungsvorhersage ist somit essentiell im Alltag: Sie ist nicht nur omniprasent, sondern ebenso von hoher Relevanz fur die eigene Sicherheit, um z. B. Kollisionen und Unfalle vermeiden zu konnen.

Bisherige wissenschaftliche Studien, in denen dieses Thema untersucht wurde, legten das Hauptaugenmerk vor allem auf die visuelle Modalitat und zeichnen sich zudem durch die Nutzung diverser Aufgabenstellungen und sehr unterschiedlichem Stimulusmaterial aus. Somit ist die Vergleichbarkeit der Befunde und die Vereinheitlichung von daraus resultierenden Erkenntnissen erschwert. Daher soll in dieser Arbeit systematisch untersucht werden, welche Faktoren die Bewegungsvorhersage in welcher Form beeinflussen. Da Bewegungsinformationen in unserem Alltag nur selten rein visueller Natur sind, wird ein multisensorischer Ansatz gewahlt und die visuelle Bewegung von zusatzlichen Tonen begleitet.

Als Einflussfaktoren werden folgende Aspekte systematisch im Detail empirisch betrachtet: 1) die Aufgabenstellung, 2) der Einfluss visueller und auditorischer Informationen und 3) die Realitatsnaher der visuellen Stimulusgestaltung. Bei der Aufgabenstellung wird zwischen Detektion, Extrapolation und Diskrimination variiert, wobei zusatzlich der Fokus zwischen zeitlichen und raumlichen Aspekten bei der Bewegungsvorhersage wechselt. Fur den Einfluss visueller und auditorischer Informationen wird zwischen unimodal visuell und bimodal audiovisuell sowie kongruenten und inkongruenten Informationen unterschieden. Daruber hinaus wird untersucht, ob sich Unterschiede fur visuelle Stimuli mit variierendem Ausma an Realitatsnaher finden lassen, um der Frage nachzugehen, inwieweit okologisch valide Schlussfolgerungen aus Experimenten zur Bewegungsvorhersage gezogen werden konnen. Dafur wurde ein besonderes Augenmerk auf die Gestaltung der Tiefendimension gelegt und Stimulusmaterial erstellt, welches sich durch diverse Tiefenhinweise wie Perspektive, Schatteneffekte und Disparitat (in Form von stereoskopischer Prasentation) auszeichnet.

Es wurden insgesamt sechs Experimente durchgefuhrt. Diese werden nach der Darlegung des theoretischen Hintergrunds einzeln vorgestellt sowie deren Ergebnisse abschlieend diskutiert und die gewonnenen Erkenntnisse in die bestehende Literatur eingeordnet.

2 Theoretischer Hintergrund

2.1 Audiovisuelle Bewegungswahrnehmung

Selten werden Informationen über Bewegungen in nur einer Modalität übertragen, somit ist die multisensorische Untersuchung der Bewegungswahrnehmung für eine realitätsnahe Forschung unerlässlich. Im Mittelpunkt dieser Arbeit steht die audiovisuelle Bewegung. Daher soll zunächst darauf eingegangen werden, inwieweit die visuelle Modalität die Wahrnehmung einer auditorischen Bewegung beeinflussen kann und umgekehrt.

Laut bisherigen Studien zur gegenseitigen Beeinflussung visueller und auditorischer Bewegung profitieren Probanden zumeist von einer bimodalen Bewegungspräsentation, indem sie schneller auf audiovisuelle Bewegungsstimuli als auf unimodale Bewegungsinformationen reagieren (Cappe et al., 2009; Harrison et al., 2010; Tyll et al., 2013). Außerdem verbessert sich die Diskriminationsleistung für Bewegungen bei visuell schwierig zu erkennenden Stimuli, wenn synchron dazu ein Ton präsentiert wird (Lewis & Noppeney, 2010). Ebenso ist die Sensitivität für die Detektion von Bewegungen, die audiovisuell präsentiert werden, erhöht gegenüber der unisensorischen Darstellung, wenn sich beide Modalitäten gleichzeitig in dieselbe Richtung bewegen (Meyer et al., 2005). Visuelle Reize können zudem die Wahrnehmung von auditorischen Bewegungsreizen vermehrt in eine bestimmte Richtung lenken (Jain et al., 2008; Soto-Faraco et al., 2002; Soto-Faraco, Spence & Kingstone, 2004). Das heißt, die Probanden nehmen eine Tonbewegung in diesen Fällen als sich in die gleiche Richtung bewegend wie die visuelle Bewegung wahr. Dies trifft umgekehrt auch auf visuelle Bewegungen mit uneindeutiger Bewegungsrichtung zu: Präsentiert man den Probanden gleichzeitig einen sich bewegendem Ton mit eindeutigen Richtungsinformationen, tendieren die Versuchspersonen verstärkt dazu, die visuelle Bewegung in die gleiche Richtung wahrzunehmen (Alink et al., 2012; Jain et al., 2008; Maeda et al., 2004; Meyer & Wuerger, 2001). Auch für biologische Bewegungen konnten verbesserte Leistungen bei audiovisuellen Stimuli gefunden werden (Arrighi et al., 2009; A. Brooks et al., 2007).

Darüber hinaus beeinflussen Töne die sogenannte Stream-and-Bounce-Illusion. Bewegen sich zwei visuelle Objekte aufeinander zu, überlagern und bewegen sich anschließend weiter, kann diese Konstellation so wahrgenommen werden, dass die beiden Objekte sich aneinander vorbeibewegen und am Ende ihre Positionen getauscht haben (Streaming) oder dass sie zum Zeitpunkt des Überlagerns kollidieren und sich zu ihrem jeweiligen Ausgangspunkt zurückbewegen (Bouncing). Ein zum Zeitpunkt der visuellen Überlagerung präsentierter auditorischer Stimulus erhöht die Wahrscheinlichkeit der Bouncing-Wahrnehmung (Sanabria et al., 2004, 2007). Ähnliche Ergebnisse fanden Ecker und Heller (2005): Diese präsentierten ihren Probanden eine Kugel, die sich durch eine Box bewegte, wobei visuell sowohl die Wahrnehmung einer vertikal rollenden als auch einer hüpfenden Bewegung möglich war. Wurde die visuelle Stimulation von einem sich entsprechend bewegendem Ton begleitet, wurde häufiger angegeben, die jeweils passende visuelle Variante der Bewegung im Display zu sehen.

Ebenso erhöht sich die Wahrscheinlichkeit eine visuelle Scheinbewegung in eine bestimmte Richtung wahrzunehmen, wenn gleichzeitig ein entsprechender Ton zu hören ist. Den

Probanden werden zwei visuelle Stimuli mit leichtem Versatz präsentiert. Wird dies mit zusätzlichen Tönen kombiniert, dann erhöht sich die Angabe der Probanden, eine Bewegung in eine bestimmte Richtung wahrzunehmen (Freeman & Driver, 2008; Getzmann, 2007). Dies gilt umgekehrt für die auditorische Richtungswahrnehmung von Tönen bei gleichzeitiger visueller Stimulation, die sich in eine bestimmte Richtung bewegt (Mays & Schirillo, 2005; Soto-Faraco, Spence & Kingstone, 2004). Darüber hinaus kann sowohl ein sich bewegendes Ton die Wahrnehmung einer Bewegung für stationäre visuelle Reize auslösen (Hidaka et al., 2009, 2011; Teramoto, Manaka et al., 2010) als auch umgekehrt ein sich bewegendes visueller Reiz dazu führen, die Bewegung eines stationären Tons wahrzunehmen (Kitagawa & Ichihara, 2002).

Ein essentielles Phänomen bei der audiovisuellen Integration ist der multimodale Erleichterungseffekt (auch Redundante-Signale-Effekt). Das heißt, die Probanden reagieren schneller bei Präsentation kongruenter bimodaler Stimuli als bei der unimodalen Variante. Dieser Effekt konnte nicht nur für einfache Stimuli (Leone & McCourt, 2013; Miller, 1982; Molholm et al., 2002; Teder-Sälejärvi et al., 2002), sondern ebenso für komplexe Objekt- und Sprachstimuli gefunden werden (Besle et al., 2004; Suied et al., 2009). Auch für bimodale Bewegungsstimuli zeigen Versuchspersonen eine verbesserte Leistung gegenüber der unimodalen Stimulation, wenn die Bewegung für beide Modalitäten kongruent präsentiert wird. Bei Harrison et al. (2010) zeigte sich für die Probanden eine signifikante Reaktionszeit-Erleichterung bei der Richtungsdiskrimination für schwelennahe audiovisuelle Bewegungsstimuli gegenüber unimodalen Varianten. Wurden die unimodalen Stimuli individuell in ihrer Stärke angepasst (sodass die Reaktionszeiten auf unimodale Präsentationen für beide Modalitäten ähnlich ausfielen), ergab sich auch für überschwellige Stimuli ein bimodaler Vorteil. Die audiovisuellen Erleichterungseffekte traten in weiteren Experimenten dieser Studie auch für die Bewegungsdetektion sowie eine Bewegt-Stationär-Diskrimination auf. In einer ähnlichen Untersuchung von Wuerger et al. (2003) wurden den Probanden zwei Intervalle mit Random-Dot-Kinematogrammen sowie auditorischem weißen Rauschen entweder einzeln oder gemeinsam präsentiert. Die Versuchspersonen gaben an, in welchem dieser beiden Intervalle eine Bewegung enthalten war. Auch bei ihnen war die Bewegungsidentifikation verbessert, wenn in beiden Modalitäten eine Bewegung präsentiert wurde. Bei der Lokalisation von sich bewegendes Targets ist die Trefferrate in audiovisuellen Durchgängen verglichen zu unimodaler Präsentation ebenfalls erhöht (Hofbauer et al., 2004). Die Versuchspersonen gaben in dieser Studie an, wo sich ein Target zu einem bestimmten Zeitpunkt befindet, wobei eine Antwort innerhalb eines bestimmten Fensters als Treffer definiert wurde. Neben der höheren Trefferrate war ebenso die Abweichung der Antworten in der audiovisuellen Bedingung geringer als in unimodalen Durchgängen.

Insgesamt deuten die Ergebnisse bisheriger Studien zur audiovisuellen Bewegungswahrnehmung darauf hin, dass beide Modalitäten sich gegenseitig beeinflussen können. Dabei können sowohl quantitative Veränderungen für verschiedene Leistungsmaße als auch qualitative Wahrnehmungsveränderungen auftreten.

2.2 Bewegungsvorhersage und verdeckte Bewegungswahrnehmung

2.2.1 Unimodale visuelle Bewegungsvorhersage und verdeckte Bewegungswahrnehmung

Bewegungen in unserem Alltag werden oftmals zwischenzeitlich verdeckt, z. B. verschwindet ein Auto im Straßenverkehr kurzzeitig hinter einem Lastkraftwagen oder einer Baumgruppe. Daher sollten auch verdeckte Bewegungen bzw. Bewegungsvorhersagen untersucht werden, um die Forschung im Bereich der Bewegungswahrnehmung möglichst realitätsnah zu gestalten.

In den Experimenten in diesem Forschungsbereich bewegt sich ein Objekt auf seiner Bahn in eine bestimmte Richtung. Während dieser Bewegung wird es kurzzeitig von einem weiteren Objekt (Okkluder) verdeckt bzw. verschwindet vom Display. Die Versuchspersonen haben die Aufgabe, entweder vorherzusagen, zu welchem Zeitpunkt das Objekt einen bestimmten Ort erreichen würde (Time-to-Arrival-/Time-to-Contact-Schätzung, z. B. DeLucia, 2004b; Landwehr et al., 2013; Lughtigheid & Welchman, 2011) oder anzugeben, an welchem Ort sich das Objekt zu einem bestimmten Zeitpunkt befinden würde (Predicted Motion, z. B. Prime & Harris, 2010). Auch eine Reaktion auf das Objekt bei Wiedererscheinen ist möglich (z. B. Diskrimination), wobei hier oftmals die räumliche bzw. zeitliche Kontinuität der Bewegungsbahn oder die Eigenschaften des Objekts vor und nach der Verdeckung variiert werden (z. B. Flombaum & Scholl, 2006; Kawachi & Gyoba, 2006). Müssen die Versuchspersonen die Bewegung eines Objekts vorhersagen, wird ihre Leistung mit zunehmender Dauer der Verdeckung bzw. des Zeitraums, in dem das Objekts nicht sichtbar ist, schlechter. Je kürzer der Zeitraum, für den eine Schätzung erfolgen soll, umso bessere Schätzleistungen werden erzielt (Caird & Hancock, 1994; Lyon & Waag, 1995; Schiff & Oldak, 1990).

Ein entscheidendes Konstrukt in diesem Zusammenhang ist die Objektkonstanz. Diese liegt vor, wenn eine Repräsentation von Teilen des visuellen Felds als dasselbe Objekt erstellt und erhalten wird – auch über die Zeit hinweg, wenn das Objekt sich bewegt, sich seine Merkmale verändern oder dessen Input, z. B. durch Verdeckung, unterbrochen wird (Scholl, 2007). Der Beobachter trifft hiernach bei der Betrachtung die Entscheidung, ob er das aktuell wahrgenommene Objekt bereits zuvor wahrgenommen hat und es sich um dasselbe oder ein neues Objekt handelt. Entsprechend dieser Entscheidung wird eine bereits existierende Objektrepräsentation erneuert oder eine neue angelegt (Flombaum et al., 2009).

Eine entscheidende Bedingung für das Auftreten von Objektkonstanz nach einer Verdeckung ist der Prozess des Verschwindens und Wiedererscheinens. Dieser muss mit einer sukzessiven Verkleinerung (beim Verschwinden) bzw. Vergrößerung (beim Wiedererscheinen) des Objekts an der Kante des Okkluders einhergehen. Ist der Verdeckungsprozess hingegen durch ein abruptes Stoppen vor dem Okkluder, gefolgt von einem sofortigen Verschwinden und einem ebenfalls abrupten vollständigen Wiedererscheinen des Objekts auf der Gegenseite gekennzeichnet, dann wird das Objekt eher nicht als dasselbe vor und nach der Verdeckung wahrgenommen (Kawachi & Gyoba, 2006; Scholl & Pylyshyn, 1999). Dies ist ebenfalls der Fall, wenn das Objekt vor dem Verschwinden kleiner wird (Implosionsbedingung) und nach der Verdeckung wieder an Größe gewinnt (Explosionsbedingung, Scholl & Pylyshyn, 1999). In

diesen Fällen des unrealistischen Verdeckungsprozesses reagieren die Probanden langsamer auf das Objekt bzw. detektieren Veränderungen schlechter (Kawachi & Gyoba, 2006; Scholl & Pylyshyn, 1999).

Neben der Gestaltung des Verdeckungsprozesses spielt der raumzeitliche Verlauf der Bewegungsbahn eine entscheidende Rolle. Versuchspersonen profitieren dabei von einem räumlich und zeitlich kontinuierlichen Bahnverlauf. Sie reagieren schneller auf ein Objekt bzw. detektieren Veränderungen von Symbolen auf diesen besser, wenn es nach der Verdeckung zum erwarteten Zeitpunkt am erwarteten Ort wiedererscheint. Diese beiden Faktoren spielen eine größere Rolle als die Tatsache, dass das Objekt während der Verdeckung seine visuellen Eigenschaften beibehält (Flombaum & Scholl, 2006; Flombaum et al., 2009; Kawachi & Gyoba, 2006). Dieser Reaktionszeitvorteil, der bei visuell kongruenten Bahnen vor und nach der Verdeckung entsteht, wird als visueller Erleichterungseffekt bezeichnet (Kawachi & Gyoba, 2006). Die daraus resultierende Wahrnehmung als dasselbe Objekt zeigt sich auch in den neuronalen Korrelaten der Objektkonstanz. Werden Probanden Bilder von Gesichtern oder Häusern gezeigt und währenddessen sukzessive von einem Okkluder verdeckt, unterscheiden sich die objektspezifischen Aktivierungen (Fusiform Face Area für Gesichter, lateraler okzipitaler Komplex für Häuser) nur geringfügig in der sichtbaren und verdeckten Bedingung. Dies spricht dafür, dass die Aktivierungen in den objektbezogenen Arealen nicht die Wahrnehmung dieser, sondern eher deren Repräsentation, die auch während der Verdeckung bestehen bleibt, widerspiegeln (Hulme & Zeki, 2006).

Für die Vorhersage von Objektbewegungen existieren zwei Erklärungsmechanismen, mit welchen die dahinter liegenden Prozesse genauer charakterisiert werden. Der erste Mechanismus beruht auf der Tau-Theorie (H. Hecht & Savelsbergh, 2004, siehe auch D. N. Lee, 1976). Nach dieser nutzen Beobachter für die Vorhersage von Bewegungen optische Variablen. Im Vordergrund steht hierbei Tau, wobei dieses definiert wird über die aktuelle Veränderungsrate des retinalen Winkels zwischen Pfad und Objektzentrum. Da Tau ausreicht, um die Ankunftszeit des bewegten Objekts an einer bestimmten Position exakt zu berechnen, sind weitere Informationen wie z. B. die Objektgröße oder die Distanz zwischen Objekt und Beobachter für die Vorhersage zukünftiger Positionen nicht notwendig. Die Beobachter extrahieren entsprechend zu Bewegungsbeginn Tau und nutzen im Anschluss einen einfachen Abzählmechanismus, um zu dem Zeitpunkt, der initial über Tau geschätzt wurde, eine Reaktion auszuführen („clocking process“, Tresilian, 1995). Diese Theorie gilt in dieser reduzierten Variante jedoch mittlerweile als überholt, da zahlreiche empirische Ergebnisse gegen sie sprechen. So ist die Vorhersageleistung von Probanden oftmals schlechter als von Tau vorhergesagt, zum anderen fällt sie ebenfalls gut aus, wenn experimentell keine Tau-Informationen zur Verfügung gestellt werden. Außerdem interferiert die Vorhersageleistung mit zahlreichen Informationsquellen, die theoretisch nicht notwendig sind (z. B. mit Objekteigenschaften wie Größe oder Kontrast). Daher wird davon ausgegangen, dass weitere kognitive Prozesse bei der Vorhersage eine Rolle spielen müssen (für einen Überblick hierzu siehe DeLucia, 2004a; H. Hecht & Savelsbergh, 2004; Tresilian, 1999).

Der zweite Erklärungsmechanismus stellt diese kognitiven Prozesse bei der Bewegungsvorhersage direkt in den Vordergrund. Er wurde in verschiedenen Studien unter unterschiedlichen Bezeichnungen eingeführt, z. B. als „Tracking-Mechanismus“ (Lyon & Waag, 1995), „kognitive Bewegungsextrapolation“ (DeLucia & Liddell, 1998) oder „mentale Imagery Strategie“ (Huber & Krist, 2004). Generell wird bei diesem Mechanismus davon ausgegangen, dass der Beobachter bei der Bewegungsvorhersage eine kognitive Repräsentation des bewegten Objekts anlegt, auf die er während einer nicht sichtbaren Phase zurückgreift und sich so die Weiterbewegung mental vorstellt. Eine Reaktion erfolgt laut diesem Mechanismus, sobald das vorgestellte Objekt den Zielort erreicht (DeLucia & Liddell, 1998).

Für diesen Mechanismus sprechen zahlreiche empirische Befunde. So kommt es in einer Vorhersageaufgabe, in der die Versuchspersonen zu Beginn nicht wissen, für welche Dauer ein Objekt nicht sichtbar sein wird, zu ähnlichen Vorhersagefehlern wie in einer klassischen Time-to-Contact-Aufgabe, bei der zu Beginn eines Durchgangs feststeht, welche Länge das Verdeckungsintervall einnimmt (DeLucia & Liddell, 1998). Dieser Befund und die Tatsache, dass eine klassische TTC-Aufgabe mit einer gleichzeitigen variablen Vorhersageaufgabe interferiert, sprechen dafür, dass beiden Aufgaben ein ähnlicher Prozess zugrunde liegt. Dieser kann kein einfacher Abzählprozess sein, da dieser in einer Aufgabe mit variablen Verdeckungsintervall nicht verwendet werden kann (DeLucia & Liddell, 1998). Zudem verschlechtert sich die Vorhersageleistung, wenn während des Vorhersageprozesses auf der nicht sichtbaren Bewegungsbahn des verdeckten Objekts Distraktoren eingeblendet werden. Dies spricht dafür, dass die Probanden sich vorstellen, das Objekt bewegte sich weiter (Lyon & Waag, 1995). Die Ergebnisse einer Untersuchung von de'Sperati und Deubel (2006) unterstützen diese Ansicht. Bei ihnen sollten sich die Probanden die Fortbewegung eines Targets auf einer Kreisbahn vorstellen. Während dieser Vorstellung blitzte auf dem Display ein Lichtblitz auf und die Probanden sollten eine Sakkade in Richtung des Lichtreizes ausführen. Dabei waren sie schneller, wenn der Lichtreiz auf der nicht sichtbaren Bewegungsbahn des Targets auftauchte. Dies spricht laut den Autoren dafür, dass die mentale Vorstellung der Fortbewegung des verdeckten Objekts mit einer Verlagerung der visuell-räumlichen Aufmerksamkeit einhergeht. Auch die Augenbewegungen von Probanden während einer verdeckten Bewegung sprechen für ein mentales Verfolgen des nicht sichtbaren Objekts (Huber & Krist, 2004; Jonikaitis et al., 2009). Dass die Probanden die vorgestellte Bewegung mit ihren Augenbewegungen verfolgen, zeigen auch Untersuchungen, in denen sich die Schätzleistungen, in denen die Probanden eine bestimmte Stelle fixieren mussten, von denen, in denen freie Augenbewegungen erlaubt waren, unterschieden (Bennett et al., 2010; Makin & Poliakoff, 2011).

Zusätzlich wird die Annahme einer vorgestellten Fortbewegung während der Vorhersage von Studien untermauert, in denen die neuronalen Korrelate der Bewegungsvorhersage untersucht wurden. Bei einer Zellableitung von Neuronen im posterioren Parietalkortex (v. a. Sulcus intraparietalis und Sulcus temporalis superior) beim Rhesusaffen waren diese bei der Betrachtung einer Bewegung aktiv. Wurde das Objekt während der Bewegung kurzzeitig verdeckt, wurde dennoch Aktivität in den Neuronen gefunden, wobei ungefähr die Hälfte der Neuronen signifikant aktiver war, wenn eine Fortbewegung des Stimulus während der Verdeck-

ckung angenommen werden konnte verglichen zu einer stationär verdeckten Bedingung (Assad & Maunsell, 1995). Ein ähnliches Aktivierungsmuster im Sulcus temporalis superior konnte bei Makaken während der Betrachtung verdeckter Objekte gefunden werden (Baker et al., 2001). Auch im Menschen aktiviert die Vorstellung einer Kugelbewegung ein Netzwerk aus parieto-frontalen und lateral okzipitalen Regionen (inklusive hMT und V5+) und damit ähnliche Areale wie bei der visuellen Bewegungswahrnehmung (Kaas et al., 2010). In einer Untersuchung von Olson et al. (2003) wurde den Probanden direkt die Bewegung eines Objekts gezeigt, welche kurzzeitig durch einen Okkluder verdeckt wurde. Hier fanden sich ebenfalls Aktivierungen in Sulcus intraparietalis sowie in mittleren temporalen Regionen während der Verdeckung. Im Experiment von Shuwairi et al. (2007) ergaben sich in der verdeckten Bewegungsbedingung zudem Aktivierungen im lateralen okzipitalen Komplex, was erneut die Objekt Konstanz (s. o.) von sich kontinuierlich bewegenden Objekten während der Verdeckung bestätigt und für eine mentale Repräsentation der vorgestellten Bewegung während dieser spricht. Bei konkreter Untersuchung räumlicher und zeitlicher Aspekte der Bewegungsvorhersage konnte gefunden werden, dass die parietalen Aktivierungen eher mit den räumlichen Aspekten bzw. der raumzeitlichen Integration verschiedener Informationen assoziiert werden (Assmus et al., 2003, 2005; Beudel et al., 2009). Zeitliche Aspekte der Vorhersage scheinen hingegen eher mit Regionen im Kleinhirn verknüpft zu sein (Beudel et al., 2009; Deluca et al., 2014; O'Reilly et al., 2008).

2.2.2 Bimodale audiovisuelle Bewegungsvorhersage und verdeckte Bewegungswahrnehmung

Da ein Großteil der Studien zur Bewegungsvorhersage sich bisher auf die visuelle Modalität konzentrierte, soll in der vorliegenden Arbeit die Untersuchung der Bewegungsvorhersage um den Aspekt der multimodalen Präsentation erweitert werden. In diesem Bereich existieren einige Studien aus der Säuglingsforschung. So übt ein zusätzlicher, ebenfalls dynamischer, auditorischer Stimulus bei vier Monate alten Säuglingen einen positiven Einfluss auf die wahrgenommene Trajektorienkontinuität aus, während das sich bewegende Objekt zeitweise von einem Okkluder verdeckt wird (Bremner et al., 2012). Ebenso zeigen Säuglinge in dieser Altersgruppe eine verstärkte okulomotorische Antizipation bei gleichzeitiger Präsentation dynamischer Tonreize während der Verdeckung (Kirkham et al., 2012).

Auch in der Erwachsenenforschung wurden für die audiovisuelle Bewegungsvorhersage erst wenige Studien durchgeführt, die zudem inkonsistente Ergebnisse hervorbrachten. Einige Studien zum Representational-Momentum (die Tendenz, sich bei der Positionsangabe für ein sich bewegendes Objekt in dessen Bewegungsrichtung zu verschätzen) konnten zeigen, dass Töne sich auf den Prozess, die Objektposition vorherzusagen, auswirken. Präsentiert man einen Ton bevor ein sich bewegender Punkt verschwindet, und die Probanden sollen dessen letzte Position angeben, dann verschiebt sich die Ortsangabe entgegengesetzt zur visuellen Bewegungsrichtung, womit sich der Representational-Momentum-Bias verkleinert (Chien et al., 2013). Ebenso kann die Dislokation verkleinert werden, wenn ein begleitender Ton kurz vor dem Verschwinden des visuellen Objekts endet, umgekehrt vergrößert sie sich, wenn der Ton über das Verschwinden hinaus vorhanden ist (Teramoto, Hidaka et al., 2010).

Eine systematischere Untersuchung zum Zusammenspiel von visuellen und auditorischen Informationen bei der Positionsvorhersage eines Objekts wurde von Prime und Harris (2010) durchgeführt. Sie verglichen die Genauigkeit, mit der die Probanden den Ort eines sich bewegenden audiovisuellen Stimulus vorhersagen konnten, nachdem dieser zu einem bestimmten Zeitpunkt der Bewegung verdeckt wurde (der Ton war weiterhin zu hören) und die Ton- und Bildinformationen zeitlich kongruent oder inkongruent präsentiert wurden. Die Probanden zeigten den typischen Representational-Momentum-Effekt und überschätzten die Targetposition in den kongruenten Bedingungen in Bewegungsrichtung. In den inkongruenten Bedingungen, in denen der Ton für die visuelle Bewegung entweder zu schnell oder zu langsam sein konnte, verschob sich die Dislokation in Richtung des Tons. Die Autoren folgerten daraus, dass die Versuchspersonen sich bei der Vorhersage nicht ausschließlich auf die visuellen Informationen verließen, sondern ebenso die Töne nutzten.

Demgegenüber brachten Studien, in denen die Time-to-Arrival-Aufgabe genutzt wurde, einen eher geringen Einfluss auditorischer Informationen auf die Bewegungsvorhersage hervor. In den Experimenten von Hofbauer et al. (2004) mussten die Probanden angeben, wann ein sich bewegendes audiovisuelles Objekt eine bestimmte Position erreichen würde. Dabei konnte das Objekt wahlweise bis zur Reaktion sichtbar sein oder einen bestimmten Zeitraum vor Erreichen des Zielorts verdeckt werden. Die Probanden profitierten nur in der nicht verdeckten Bedingung von audiovisuellen gegenüber unimodalen Informationen. In der verdeckten Bewegungsvariante unterschied sich die Leistung der Probanden zwischen den audiovisuellen und visuellen Bedingungen nicht. Die Autoren sahen darin den Beleg dafür, dass eine effektive Integration auditorischer und visueller Informationen für die Angabe der aktuellen, nicht aber bei der Vorhersage der zukünftigen Position von sich bewegenden Objekten stattfindet. Bei letzterer entsteht durch eine auditorische Stimulation ein nur geringfügiger zusätzlicher Nutzen. In einer weiteren Studie konnte die Arbeitsgruppe jedoch zeigen, dass bei Time-to-Arrival-Schätzungen mit audiovisuellen Stimuli die Variabilität der Antworten gegenüber den unimodalen Bewegungen verringert ist (Wuerger et al., 2010).

Auch DeLucia et al. (2016) untersuchten die audiovisuelle Informationsnutzung für Time-to-Contact-Aufgaben. Bei ihnen bewegte sich ein Objekt auf den Betrachter zu, verschwand und die Probanden gaben an, wann das Objekt sie erreicht hätte. Die Bewegung konnte entweder visuell, auditorisch oder audiovisuell präsentiert werden. Die Versuchspersonen waren schlechter in ihrer Schätzung, wenn sie nur auditorische Informationen erhalten hatten, sie unterschieden sich jedoch nicht in ihren Schätzungen für visuelle und audiovisuelle Durchgänge. Die Ergebnisse einer multiplen Regressionsanalyse zeigten hingegen eine Nutzung sowohl visueller als auch auditorischer Informationen bei der Vorhersage, welche anscheinend jedoch nicht in gleichen Maßen genutzt wurden. Die Probanden gewichteten die visuellen Informationen stärker und profitierten dadurch in audiovisuellen gegenüber rein visuellen Bewegungen nicht zusätzlich. Zu einer ähnlichen Schlussfolgerung gelangten L. Zhou et al. (2007). Bei ihnen bewegte sich ein computeranimiertes Auto auf den Beobachter zu. Nach zwei Bewegungsdurchgängen gaben die Probanden an, in welchem davon das Objekt sie eher erreichen würde (relative Time-to-Contact-Schätzung). Auch hier war ein Großteil der

Probanden (8 von 9) vor allem sensitiv für die visuellen Bewegungsinformationen. Die Schätzleistung war für visuelle und audiovisuelle Bewegungen besser als für auditorische. Visuelle und audiovisuelle Bedingungen unterschieden sich jedoch nicht voneinander.

Andere Forschergruppen nutzten für die Untersuchung der Bewegungsvorhersage alltagsnähere Szenarios aus dem Straßenverkehr. Bei Schiff und Oldak (1990) sahen die Probanden Videoaufnahmen von Fahrzeugen, die sich auf sie zubewegten, hörten das dabei entstehende Geräusch oder waren beiden Informationskanälen gleichzeitig ausgesetzt. Die Probanden waren in ihren Schätzungen schlechter, wann das Fahrzeug bei ihnen ankommen würde, wenn sie die Fahrzeugbewegung nur gehört hatten verglichen zur Stimulation mit Bild- und Bild-/Tonmaterial. Der Unterschied zwischen visueller und audiovisueller Stimulation war nicht signifikant. Auch bei der Entscheidung, ob man eine Straße überqueren kann während sich ein Fahrzeug nähert, bringen audiovisuelle gegenüber rein visuellen Informationen wenig Vorteile (Hassan, 2012). Ebenso fanden Keshavarz et al. (2017) in einer Time-to-Contact-Untersuchung mit Fahrzeugen in einer virtuellen Umgebung, dass visuelle Informationen in audiovisuellen Durchgängen bei der Schätzung der Ankunftszeit dominieren.

Dass Probanden durchaus auditorische Informationen für die Bewegungsvorhersage nutzen, konnten Gordon und Rosenblum (2005) demonstrieren. Bei ihnen kam es zu keinem Leistungsabfall, wenn während des Vorhersageprozesses das Signal in einer der beiden Modalitäten unterbrochen wurde. Die Probanden waren anscheinend in der Lage, flexibel ihre Nutzung an die zur Verfügung stehenden Informationsquellen anzupassen. Des Weiteren waren die Probanden in den Untersuchungen von Bernhard et al. (2011) in einer virtuellen Kreuzungsüberquerungsaufgabe besser, wenn sie zum visuellen Input gleichzeitig räumlich modulierte auditorische Bewegungsinformationen der Autos hörten.

Zusammenfassend lässt sich die Untersuchung der visuellen Bewegungsvorhersage als recht umfassend betrachten. Es liegen zahlreiche Studien zu diesem Thema inklusive damit assoziierten Erklärungsmechanismen sowie neuronalen Korrelaten vor. Die Studienlage für multimodale Untersuchungen in diesem Bereich gestaltet sich hingegen spärlich. Die Nutzung unterschiedlicher Aufgabenstellungen sowie verschiedenartiges Stimulusmaterial könnten dabei zu den uneinheitlichen Ergebnissen beitragen. Daher soll im Folgenden auf diese beiden Einflussfaktoren genauer eingegangen werden.

2.3 Einfluss der Aufgabenstellung

Für die Untersuchung der Bewegungsvorhersage sind zwei Aspekte der Aufgabenstellung von Relevanz. Zunächst sollte gemäß der Modalitäts-Angemessenheits-Hypothese (Modality Appropriateness Hypothesis, Welch & Warren, 1980; Welch et al., 1986) eine Abhängigkeit vom Fokus der Aufgabenstellung gegeben sein. Laut dieser verwenden Probanden in Aufgaben, in denen räumliche Aspekte im Vordergrund stehen, vermehrt visuelle Informationen, während sie in zeitlich orientierten Aufgaben eher auf auditorische Informationen zurückgreifen. Dies hängt mit den verschiedenen Auflösungen der sensorischen Verarbeitung für räumliche und zeitliche Aspekte zusammen. Die Auflösung für räumliche Informationen ist im visuellen System größer,

im auditorischen System ist die zeitliche Auflösung überlegen. Die Versuchsperson wählt die Modalität, die für den jeweiligen Aufgabenaspekt am geeignetsten erscheint.

Den zweiten großen Einflussfaktor stellt die Art der Aufgabenstellung dar. Die beiden in psychologischen Untersuchungen sehr häufig verwendeten Aufgaben Detektion und Diskrimination unterscheiden sich in ihrer kognitiven Beanspruchung: Die Detektion (Einfachreaktion auf einen Reiz) ist als weniger beanspruchend anzusehen als die Diskrimination, bei der eine Wahlreaktion aus verschiedenen Antwortalternativen getroffen werden muss (Bertelson & Boons, 1960; Luck & Hillyard, 1995). In Studien zur Bewegungsvorhersage wird häufig eine weitere Aufgabenstellung benutzt: Die Probanden sagen die Bewegung eines Objekts vorher (Extrapolation, siehe Abschnitt 2.2.1). Wie bereits im vorherigen Kapitel dargelegt, wurden zahlreiche verschiedene Aufgabenstellungen verwendet, um die Bewegungsvorhersage zu untersuchen. Eine vergleichende Analyse verschiedener Aufgabenstellungen bei der Bewegungsvorhersage mit identischem Stimulusmaterial fehlt jedoch bisher.

Allerdings konnten Untersuchungen zu anderen Phänomenen aus dem Bereich der visuellen Aufmerksamkeit einen Einfluss der Aufgabenstellung nahelegen. So fällt der Temporal-Orienting-Effekt (schnellere Reaktion, wenn ein Target nach einem zuvor gecueten Zeitintervall erscheint) in Detektions- und Diskriminationsaufgaben unterschiedlich groß aus, zudem ist sein Auftreten während einer Diskrimination an bestimmte Bedingungen geknüpft (Correa et al., 2004). Auch der Inhibition-of-Return-Effekt (verlangsamte Reaktion auf ein Target an einem Ort, der zuvor bereits gecuet wurde) wird von der Aufgabenstellung beeinflusst (Chica et al., 2006; Lupiáñez et al., 2001), ebenso wie das Entdecken impliziter Veränderungen (Triesch et al., 2003) oder das Auftreten bzw. die Größe des Attentional Blinks (verschlechterte Reaktion auf ein zweites Target kurz nach der Reaktion auf ein erstes Target, T. Stein et al., 2009). Die Aufgabenstellung kann des Weiteren andere Maße wie die Augenbewegungen der Probanden beeinflussen (Boutet et al., 2017).

In multisensorischen Untersuchungen konnte ebenfalls ein Einfluss der Aufgabenstellung aufgezeigt werden. Das Temporal-Binding-Window (Zeitfenster, in dem zwei präsentierte Reize als ein Objekt wahrgenommen werden) im Stream-and-Bounce-Paradigma (s. o.) weist eine unterschiedliche Breite auf, je nachdem, ob die Probanden angeben sollten, ob sie ein Streaming- bzw. Bouncing-Ereignis wahrgenommen haben oder ob sie beurteilen sollten, ob die visuelle Überschneidung der Stimuli und der Ton gleichzeitig präsentiert wurden (Donohue et al., 2015). Auch bei einfacher Präsentation kurzer visueller und auditorischer Reize erfolgt eine Anpassung des Temporal-Binding-Windows in Abhängigkeit davon, ob eine Reihenfolgenbeurteilung (Temporal-Order-Judgement) oder eine Reaktion auf den ersten präsentierten Stimulus verlangt wird (Mégevand et al., 2013). Gleichfalls konnten Stevenson und Wallace (2013) für verschiedene Stimulusarten (einfache Scheibenstimuli/Pieptöne, audiovisuelle Sprachstimuli, audiovisuelle Videosequenzen zur Werkzeugnutzung) einen Einfluss der Aufgabenstellung auf die Ausprägung des Temporal-Binding-Windows belegen. Dieses wies für die Reihenfolgenbeurteilung und bei der Angabe, welche von zwei hintereinander dargebotenen audiovisuellen Stimulationen synchron war, eine engere Breite auf. Bei der Angabe, ob die audiovisuellen Stimuli zur gleichen Zeit präsentiert wurden und ob die Präsentation als ein einziges Ereignis

wahrgenommen wurde, war das Temporal-Binding-Window breiter ausgeprägt. Beim direkten Vergleich von Detektion und Diskrimination fanden Gleiss und Kayser (2013) eine Verbesserung der visuellen Detektionsleistung in der Peripherie bei gleichzeitiger Tonpräsentation, die sich jedoch nicht in einer verbesserten Diskrimination niederschlug. D. Hecht et al. (2008) konnten hingegen zeigen, dass die multisensorischen Integrationsgewinne (Erleichterungseffekte) bei Wahlreaktionsaufgaben größer ausfallen als bei Einfachreaktionen.

In einer Untersuchung zum Zusammenspiel von Aufgabenstellung und Raum-Zeit-Anforderung profitierten die Probanden von einer visuotaktilen gegenüber einer unimodalen Stimulation, wenn sie nur auf einen Stimulus reagieren sollten, unabhängig davon, wo dieser präsentiert wurde (Girard et al., 2011). War jedoch eine Diskrimination erforderlich und die Probanden sollten nur auf die rechtsseitig präsentierten Stimuli reagieren, dann riefen räumlich inkongruente bimodale Stimuli höhere Reaktionszeiten hervor als gleichseitige Stimuli. Die Autoren folgerten daraus, dass das Auftreten multisensorischer Effekte von den Aufgabenanforderungen abhängig ist. Dass überdies gegensätzliche multisensorische Effekte für das gleiche experimentelle Set-Up auftreten können, wurde von Sinnott et al. (2008) demonstriert. Sie ließen ihre Probanden in visuellen und auditorischen Streams auf visuelle, auditorische oder audiovisuelle Targets in einer Detektions- und einer Diskriminationsaufgabe reagieren. Es kam in der Detektion zu einem multisensorischen Erleichterungseffekt (Redundante-Target-Effekt), in der Diskrimination hingegen zu einem visuellen Dominanzeffekt und damit einer multisensorischen Inhibition.

Im Bereich der Extrapolationsaufgabe(n) wurden zum Einfluss der Aufgabenstellung bisher sehr wenige Studien und diese mit ausschließlich visuellem Material durchgeführt. In den Untersuchungen von Makin und Bertamini (2014) führten die Probanden verschiedene Schätzaufgaben aus. Neben einer klassischen Time-to-Arrival-Schätzung mussten sie in anderen Aufgaben angeben, wann eine sich mit Gabor-Patches füllende Box vollständig mit diesen bedeckt ist und wann ein sich auf einer Kreisbahn bewegendes Balken eine komplette Rotation zurückgelegt hat. Zwischen den verschiedenen Schätzaufgaben ergaben sich hohe Korrelationen in den Leistungen der Probanden, was laut den Autoren auf einen übergeordneten Schätzmechanismus für verschiedene Aufgabentypen hindeutet. In einer weiteren Studie wurden die absolute und relative Time-to-Arrival-Schätzung verglichen (Lugtigheid & Welchman, 2011). Die Autoren kamen zu dem Ergebnis, dass die Tendenz die Ankunftszeit des Objekts zu über- oder unterschätzen in der relativen Aufgabe geringer ausfiel. In einem älteren Übersichtsartikel arbeitete Tresilian (1995) Unterschiede zwischen verschiedenen Extrapolationsaufgaben heraus. Er folgerte, dass Time-to-Arrival-Aufgaben eher auf einen kognitiven Prozess, relative Varianten der Schätzung hingegen eher auf die Nutzung von Heuristiken zurückzuführen sind. Außerdem stellte er Übung sowie die Dauer der sichtbaren und unsichtbaren Bewegung als essentielle Einflussfaktoren auf die Bewegungsvorhersage heraus.

Dass die verwendete Aufgabenstellung Einfluss auf das Verhalten der Probanden nehmen kann, konnte durch zahlreiche Studien belegt werden. Da bisherige Studien zur Bewegungswahrnehmung verschiedene Aufgabenstellungen verwendeten, erscheint eine Untersuchung

verschiedener Aufgaben für das gleiche experimentelle Set-Up sinnvoll, um etwaige Zusammenhänge zwischen Vorhersageleistung und Aufgabenstellung aufzudecken.

2.4 Verwendung von quasi-natürlichem Stimulusmaterial

Neben den verschiedenen verwendeten Aufgabenstellungen kann ebenso das sehr divergente Stimulusmaterial eine Ursache für die inkonsistente Ergebnislage zur audiovisuellen Bewegungsvorhersage darstellen. Insbesondere die Verwendung einer vereinfachten Gestaltung ist für die Generalisierbarkeit der Ergebnisse problematisch. In vielen Untersuchungen zur audiovisuellen Bewegungswahrnehmung bzw. -vorhersage wurden größtenteils stark vereinfachte und somit realitätsferne Stimuli genutzt, wie z. B. Lichtblitze (Harrison et al., 2010; Hidaka et al., 2009; Hofbauer et al., 2004; Wuerger et al., 2010), Gitterstimuli (Jain et al., 2008; Maeda et al., 2004), Random-Dot-Kinematogramme (Alais & Burr, 2004a; Baumann & Greenlee, 2007; Gleiss & Kayser, 2014; Meyer & Wuerger, 2001) oder einfache geometrische Formen (Bremner et al., 2012; Chien et al., 2013; Freeman & Driver, 2008; Hidaka et al., 2009; Kirkham et al., 2012; Prime & Harris, 2010). Diese wurden in der auditorischen Domäne kombiniert mit Pieptönen (Freeman & Driver, 2008), Klickgeräuschen (Hofbauer et al., 2004; Mays & Schirillo, 2005; Wuerger et al., 2010), White-Noise-Bursts (Harrison et al., 2010; Hidaka et al., 2009; Teramoto, Manaka et al., 2010) oder lautstärkemodulierten Reintönen aus zwei Lautsprechern (Prime & Harris, 2010).

Um Untersuchungen realitätsnäher zu gestalten, nutzten einige Forschergruppen Looming-Stimuli. Diese vergrößern ihre Grundfläche auf dem Display und erzeugen so den Eindruck, sich auf den Betrachter zuzubewegen. Auf diese Weise soll eine höhere Verhaltensrelevanz erzielt werden, da ein sich auf den Beobachter zu bewegendes Objekt potentiell eine Reaktion auf den Stimulus nach sich ziehen könnte. Dem gegenüber stehen Receding-Stimuli, die sich durch eine Verringerung der Grundfläche scheinbar vom Betrachter entfernen. Im audiovisuellen Bereich wurden in verschiedenen Studien multisensorische Interaktionseffekte mit Looming-Stimuli untersucht. In den Experimenten von Kitagawa und Ichihara (2002) sahen Probanden einen quadratischen Looming- oder Receding-Stimulus. Nachdem sie auf die Bewegung des visuellen Stimulus adaptiert hatten, gaben die Probanden vermehrt an, einen gleichzeitig präsentierten stationären Ton ebenfalls als bewegt wahrzunehmen. Bei Cappe et al. (2009) zeigten die Probanden präferiert für bimodale Looming-Stimuli Integrationseffekte, sie erzielten in diesen Bedingungen bessere Leistungen als in den unisensorischen Durchgängen oder den statischen audiovisuellen Vergleichsbedingungen. Auch in der Untersuchung von Tyll et al. (2013) reagierten die Probanden schneller auf bimodale Looming-Stimuli.

In einem nächsten Schritt lässt sich experimentell eine Bewegung im Raum nicht nur durch Looming-/Receding-Stimuli andeuten, sondern durch Verwendung von Disparitätshinweisen, und damit einer dreidimensionalen (3D) Darstellung, tatsächlich erzeugen. Disparität ist ein binokularer Tiefenhinweis, der durch den horizontalen Abstand beider Augen beim Betrachten einer Szene entsteht. Dadurch unterscheiden sich die retinalen Positionen eines Bildpunkts in beiden Augen voneinander und dem Betrachter wird so das Stereosehen ermöglicht (Mather, 2009). Disparitätshinweise wurden beispielsweise bei Harrison et al. (2015) verwendet. Sie

präsentierten ihren Versuchspersonen visuelle und auditorische Stimuli im zweidimensionalen (2D) und 3D-Raum. Die Probanden gaben an, in welche Richtung sich der auditorische Stimulus bewegte. Sie reagierten akkurater in kongruenten audiovisuellen Bedingungen als in inkongruenten. Dieser Kongruenzeffekt fiel bei 3D-Stimulation größer aus als bei 2D-Stimuli, zudem waren die kortikalen Antworten im Elektroenzephalogramm (EEG) für 3D-Stimuli stärker. Auch bei einer visuellen Suchaufgabe im 3D-Raum, bei der sich die Stimuli auf die Versuchspersonen zu- oder von ihnen wegbewegten, fand eine erleichterte Targetdetektion statt, wenn diese von passenden Tonbewegungen (lauter bzw. leiser werdend) begleitet wurden (Zannoli et al., 2012).

Beim direkten Vergleich von 2D- und 3D-Material konnten Unterschiede für auditorisches, visuelles und audiovisuelles Material gefunden werden. Getzmann und Lewald (2010) präsentierten ihren Probanden auditorische Bewegungen in verschiedenen Komplexitätsstufen: von einfachen, über Pegelunterschiede erzeugte, Tonbewegungen hin zu komplexen 3D-Freifeldtönen. Dabei führten die 3D-Töne zu kürzeren Reaktionszeiten und früheren EEG-Antworten als einfache 2D-Töne. Beim Betrachten von den gleichen Videoclips in 2D und 3D ergaben Nachbefragungen der Probanden, dass die 3D-Variante als intensiver durch ein größeres „Eintauchgefühl“ erlebt wurde. Dies ging in einer multivariaten Analyse der funktionellen Magnetresonanztomographie-(fMRT)-Daten mit höheren Intersubjektkorrelationen der kortikalen Netzwerke einher (Gaebler et al., 2014). Ebenso zeigen sich sowohl stärkere Aktivierungen in objekt- und bewegungsspezifischen Hirnarealen für visuelle und auditorische 3D-Bewegungen (Ogawa & Macaluso, 2013) als auch spezifische Aktivierungen für 3D-Anteile in Filmen (Ogawa et al., 2013).

Nicht nur über eine 3D-Darstellung im Raum, sondern ebenso über den Grad der Realitäts-treue des Stimulusmaterials lassen sich Experimente alltagsnäher gestalten. Die oben bereits vorgestellten Studien zur audiovisuellen Bewegungswahrnehmung, in denen realitätsnahes Video- und Tonmaterial verwendet wurde (Gordon & Rosenblum, 2005; Roper & Hassan, 2014; Schiff & Oldak, 1990), haben im Gegenzug ihre Probleme in einer schwächeren experimentellen Kontrollierbarkeit durch Faktoren wie unterschiedliche visuelle Merkmale der dargestellten Fahrzeuge oder auch eine geringere zeitliche Genauigkeit, vor allem bei der audiovisuellen Synchronität. Dass die Nutzung alltagsnaher Stimuli für valide Aussagen unerlässlich ist, zeigen Experimente, in denen verschiedene Materialarten verglichen wurden. Beim Vergleich von Random-Dot-Kinematogrammen kombiniert mit einfachen amplituden- und frequenzmodulierten Tönen mit natürlicheren Shepard-Stimuli fiel der gefundene Looming-Bias (geringere Reaktionszeiten) stärker für die Shepard-Stimuli aus (Conrad et al., 2013). In einer weiteren Studie kam es bei realitätsnahen Videos von Tropf-Sequenzen zu früheren multisensorischen Integrationseffekten und unterschiedlichen EEG-Aktivierungsmustern im Vergleich zu abstrakten Bewegungsstimuli (Senkowski et al., 2007). Zudem konnten Snow et al. (2011) demonstrieren, dass die typischen Repetitionseffekte im fMRT-Signal bei der Betrachtung von Bildern schwächer ausfallen oder nicht auftreten, wenn man Probanden statt 2D-Bildern ihre realen 3D-Objekt-Pendants präsentiert.

Die in diesem Abschnitt vorgestellte Literatur legt einen Einfluss der Stimulusmaterialgestaltung sowohl auf Verhaltensmaße als auch assoziierte neuronale Prozesse nahe. Insbesondere vor dem Hintergrund der Übertragbarkeit von Studienergebnissen auf den Alltag stellt der Vergleich verschieden gestalteter Materialien, v. a. bezüglich der Tiefendimension, einen wichtigen Untersuchungsaspekt dar.

3 Ziele der Arbeit

3.1 Fragestellungen

Aufgrund der uneinheitlichen Ergebnislage und der Vielzahl potentieller Einflussfaktoren für die audiovisuelle Bewegungsvorhersage im Raum, soll in dieser Arbeit eine systematische Untersuchung dieser Zusammenhänge erfolgen. Im Vordergrund steht ein alltagsnahes Experimentalset-Up, das vor allem durch eine multisensorische Stimulation und eine räumliche Gestaltung der visuellen und auditorischen Bestandteile gekennzeichnet ist. Es wird das Zusammenspiel kongruenter und inkongruenter visueller und auditorischer Informationen bei der Bewegungsvorhersage betrachtet. Im ersten Experimentaltteil soll die Abhängigkeit dieses Zusammenspiels für zeitliche Aspekte der Bewegungsvorhersage exploriert werden, im zweiten Experimentaltteil stehen räumliche Aspekte im Vordergrund.

Dafür wurde folgendes Grundexperiment in verschiedenen Variationen durchgeführt: Eine Kugel bewegt sich in einer 3D-Box auf den Betrachter zu. Dabei wird sie kurzzeitig von einer Brücke verdeckt. Die Kugel kann im Anschluss räumlich passend zu ihrer bisherigen Bewegungsbahn vor der Verdeckung wiedererscheinen oder auf einer davon abweichenden Bahn. Außerdem kann die Bewegung der Kugel unimodal (nur visuell) präsentiert werden oder von einem sich ebenfalls auf den Hörer zu bewegendem Ton begleitet werden, wobei die Tonrichtung der Bewegungsrichtung der Kugel (links nach rechts oder rechts nach links) entspricht oder sich entgegengesetzt bewegt.

Für die Untersuchung der zeitlichen Dimension wurden die Aufgaben Detektion und Extrapolation gewählt. Die Probanden mussten angeben, wann die Kugel wiedererschien oder wann sie eine bestimmte Position nach der Verdeckung erreichen würde. Hierbei wurde zunächst davon ausgegangen, dass in der Detektionsaufgabe ein visueller Erleichterungseffekt auftritt. Ebenso sollten die Probanden in der Extrapolation für ein kurzes Schätzintervall bessere Leistungen erzielen als für einen längeren Zeitraum. Da in diesen beiden Experimenten zeitliche Aspekte im Vordergrund standen, sollten zusätzliche auditorische Informationen einen relevanten Einfluss ausüben. Hierbei stellte sich die Frage, inwieweit crossmodale Interaktionen in Form von besseren Leistungen in auditorisch kongruenten Bedingungen auftreten. Außerdem wurde untersucht, ob gefundene Effekte ggf. stärker ausfallen, wenn sie alltagsnäher, d. h. mit zusätzlichen Disparitätsinformationen, präsentiert werden. Des Weiteren wurde in diesem Experimentaltteil die Aufgabenstellung variiert, um einen potentiellen Einfluss verschieden starker Aufgabenanforderungen zu untersuchen.

Im zweiten Teil der Untersuchungen stand der räumliche Aspekt im Vordergrund, die Versuchspersonen unterschieden zwischen einem passenden und nicht passenden Bahnverlauf nach der Verdeckung (Diskrimination). Daher wurde davon ausgegangen, dass in diesen Experimenten visuelle Informationen gegenüber auditorischen bevorzugt werden und somit wurde hier erneut das Zusammenspiel zwischen visueller und auditorischer Präsentation betrachtet. Außerdem sollten auch in diesen Experimenten Vorteile für visuell passende Bahnverläufe auftreten. Da die Aufgabenstellung in diesem Teil der Arbeit konstant gehalten wurde, stand stattdessen die Variation des visuellen Stimulusmaterials im Vordergrund. Dafür wurde in ei-

nem Experiment die visuelle Erkennbarkeit der Kugel reduziert, sodass eine Nutzung visueller Informationen erschwert werden sollte und die Probanden verstärkt auf auditorische Stimuli während der Vorhersage zurückgreifen sollten. In einem weiteren Experiment erfolgte der Vergleich von 3D-Material mit vereinfachtem 2D-Material, um eine Aussage über den Beitrag verschiedener Gestaltungsformen für Tiefenhinweise bei der Bewegungswahrnehmung treffen zu können. In den übrigen vier Experimenten wurde eine reale 3D-Variante des Experiments (echt-3D, mit Disparitätsinformationen) mit einer pseudo-3D-Variante (ohne Disparitätsinformationen) verglichen. Die pseudo-3D-Variante erzeugte einen Tiefeneindruck über Hinweise wie Schattenwurf und die Verwendung einer zentralperspektivischen Darstellung. So konnte untersucht werden, ob die Verwendung zusätzlicher Disparitätsinformationen einen Einfluss auf die Bewegungsvorhersage der Probanden ausübt.

In einem letzten Experiment wurde der Frage nachgegangen, wie sich die Daten gestalten, wenn keine zusätzlichen Töne präsentiert werden. Dafür wurden sowohl eine Detektions- als auch Diskriminationsaufgabe ohne auditorische Stimulation durchgeführt, um diese beiden Aufgabenstellungen für die visuelle Bewegungsvorhersage vergleichen zu können. Auch hier wurde zwischen echt-3D und einfach-2D unterschieden. Im Fokus dieser Untersuchung standen vor allem die Vergleiche mit den Ergebnissen der bisherigen Experimente, in denen zusätzlich auditorische Stimuli präsentiert wurden.

3.2 Übersicht Experimente

Experiment 1 In diesem Experiment wurde das Grundexperiment mit einer Detektionsaufgabe durchgeführt. Die Probanden gaben per Tastendruck an, wann die Kugel nach der Verdeckung wiedererschien. Im Mittelpunkt der Aufgabenbearbeitung stand die zeitliche Dimension, womit die Nutzung von Tönen für die Bearbeitung der Aufgabenstellung vorteilhaft sein sollte. Zudem erfolgte ein Vergleich zwischen der pseudo-3D-Variante und der echt-3D-Variante des Experiments, um den Einfluss zusätzlicher Disparitätsinformationen zu überprüfen.

Experiment 2 Für dieses Experiment wurde eine anspruchsvollere Extrapolationsaufgabe gewählt. Die Kugel tauchte nach dem Verschwinden unter der Brücke nicht wieder auf und die Versuchspersonen gaben mit einem Tastendruck an, wann die Kugel eine bestimmte Position (markiert mit einer im Bild eingefügten Stange) erreichen würde. Auch hier standen zeitliche Aspekte und die Untersuchung des Toneinflusses im Vordergrund. Außerdem wurde erneut der Vergleich zwischen den beiden 3D-Präsentationsarten gezogen.

Experiment 3 Im 3. Experiment wurde das Grundexperiment aus Experiment 1 in gleicher Weise wiederholt, allerdings trafen die Probanden nach dem Wiedererscheinen der Kugel die Entscheidung, ob sich die Kugel passend oder nicht passend zu ihrer Bewegungsbahn vor der Verdeckung weiterbewegte (Diskrimination). Somit war diese Aufgabe in der räumlichen Dimension fordernd. Hierbei lag das Augenmerk vor allem darauf, ob die gleichzeitig präsentierten Töne in einer räumlichen Aufgabe weniger relevant sind. Auch hier erfolgte der Vergleich der pseudo- und echt-3D-Version des Experiments.

Experiment 4 Hier wurde erneut eine Diskriminationsaufgabe verwendet, wobei gleichzeitig das visuelle Material verändert wurde: Das Muster der Kugel wurde an das Muster des Boxbodens angepasst, um so die visuelle Schwierigkeit durch eine geringere Erkennbarkeit der Kugel zu erhöhen und damit die Reliabilität der visuellen Modalität herabzusetzen. Damit sollte eine Nutzung visueller Informationen verringert werden bei gleichzeitig größer werdendem Einfluss der auditorischen Stimulation. Ebenso wurde erneut zwischen pseudo- und echt-3D-Stimulation verglichen.

Experiment 5 Im 5. Experiment kam wieder die Diskriminationsaufgabe zum Einsatz, wobei der Vergleich zwischen der echt-3D-Stimulation und einer vereinfachten 2D-Darstellung im Vordergrund stand. Auf diese Art und Weise wurde untersucht, inwieweit sich eine realitätsnahe Experimentalgestaltung von einer vereinfachten Darstellungsvariante unterscheidet.

Experiment 6 Im letzten Experiment wurden aufgrund der Post-hoc-Betrachtung der Ergebnisse aus den Diskriminationsexperimenten die Töne während des Experiments nicht präsentiert, um zu untersuchen, wie sich die Bewegungsvorhersage ohne auditorische Stimulation während Detektion und Diskrimination verhält. Hier wurde zudem wie in Experiment 5 zwischen einer echt-3D- und einer einfach-2D-Darstellung unterschieden, um den Einfluss ökologisch validen Materials in der visuellen Modalität genauer betrachten zu können.

Anmerkung Eine Übersicht für alle sechs Experimente mit den verwendeten Aufgabenstellungen, der Designzusammensetzung sowie einem Beispielbild für das visuelle Stimulusmaterial ist im Abschnitt Allgemeine Methoden in Abbildung 1 dargestellt (siehe Seite 29).

Die Ergebnisse von Experiment 1 und 2 sind in Auszügen erschienen in: Dittrich & Noesselt (2018). Temporal Audiovisual Motion Prediction in 2D- vs. 3D-Environments. *Frontiers in Psychology*, 9, 368. doi: 10.3389/fpsyg.2018.00368. In der vorliegenden Arbeit wurden adaptierte Varianten von dort veröffentlichten Abbildungen verwendet. Dies betrifft die Abbildungen 2, 3, 4, 6 und 7 sowie Teile der Abbildungen 5, 8 und 9.

4 Allgemeine Methoden

Die hier aufgeführten Informationen betreffen die Methodik für alle sechs durchgeführten Hauptexperimente. Etwaige Abweichungen und/oder Ergänzungen sind im Methodenteil des jeweiligen Experiments dargelegt. Das grundlegende experimentelle Set-Up bestand aus einer Abwandlung des Ball-in-a-Box-Paradigmas von Kersten et al. (1997): Eine Kugel bewegte sich durch eine Box hindurch auf den Betrachter zu und wurde währenddessen kurzzeitig von einer Brücke verdeckt.

4.1 Probanden

Die Probanden aller Experimente waren normalsichtig, d. h. es lag keine Sehschwäche, kein Strabismus und keine Farbsehstörung vor. Die Fähigkeit des 3D-Sehens wurde mithilfe des Lang-Stereotests II (Lang-Stereotest AG, Schweiz) vor den Testungen sichergestellt. Keiner der Teilnehmer berichtete ein eingeschränktes Hörvermögen oder eine vorhandene neurologische bzw. psychiatrische Störung. Die Händigkeit der Versuchspersonen wurde mit einer ins Deutsche übersetzten Variante des Edinburgh-Handedness-Inventory erfasst (Oldfield, 1971). Für ihre Teilnahme wurden die Probanden wahlweise monetär oder mit Versuchspersonenstunden im Rahmen ihres Psychologie-Studiums vergütet. Alle Probanden nahmen auf freiwilliger Basis an den Untersuchungen teil und erklärten schriftlich ihr Einverständnis. Alle Experimente wurden im Einklang mit den Richtlinien des lokalen Ethikkomitees durchgeführt.

4.2 Design

Jedes Experiment wurde, mit Ausnahme von Experiment 6, mit Hilfe eines 2x2x3-Within-Designs realisiert. Der erste Faktor „Dimension“ kennzeichnete die visuelle Darstellungsart im Raum. Es wurde zwischen pseudo-3D und echt-3D (Experimente 1 bis 4) oder einfach-2D und echt-3D (Experiment 5) variiert. Der zweite Faktor „visuelle Stimulation“ definierte, ob die Kugel auf einer visuell passenden oder nicht passenden Bahn nach der Verdeckung (Experimente 1, 3, 4 und 5) wiedererschien. In Experiment 2 (Extrapolation) wurde dieser Faktor in die beiden verschiedenen Schätzdistanzen (nah und fern) aufgeteilt. Der dritte Faktor „auditorische Stimulation“ betraf die Variation des Tons. Die Bewegung der Kugel konnte unimodal (ohne Ton), mit einem zur visuellen Bewegungsrichtung passenden Ton (kongruent) oder einem Ton mit entgegengesetzter Richtung (inkongruent) präsentiert werden. In Experiment 6 erfolgte keine auditorische Stimulation, hier wurden nur der Faktor Dimension (einfach-2D vs. echt-3D) und Visuell (passend vs. nicht passend) variiert. Eine Übersicht mit den in allen sechs Experimenten verwendeten Aufgabenstellungen, der konkreten Zusammensetzung des Designs sowie einem Beispielbild für die visuelle Stimulation ist in Abbildung 1 zu sehen. Abbildung 2 zeigt schematisch die visuellen und auditorischen Bahnen für die Experimente sowie Beispiele für inkongruente Bedingungen.

Aufgabe	Design	Bildbeispiel																									
<p>Experiment 1/3</p> <p>Detektion/ Diskrimination</p>	<table border="1"> <tr> <td colspan="2"></td> <td colspan="3">auditorische Stimulation</td> </tr> <tr> <td>Dimension</td> <td></td> <td>ohne Ton</td> <td>kongruent</td> <td>inkongruent</td> </tr> <tr> <td rowspan="4">visuelle Stimulation</td> <td rowspan="2">kongruente Bahn</td> <td>pseudo-3D</td> <td>pseudo-3D</td> <td>pseudo-3D</td> </tr> <tr> <td>echt-3D</td> <td>echt-3D</td> <td>echt-3D</td> </tr> <tr> <td rowspan="2">inkongruente Bahn</td> <td>pseudo-3D</td> <td>pseudo-3D</td> <td>pseudo-3D</td> </tr> <tr> <td>echt-3D</td> <td>echt-3D</td> <td>echt-3D</td> </tr> </table>			auditorische Stimulation			Dimension		ohne Ton	kongruent	inkongruent	visuelle Stimulation	kongruente Bahn	pseudo-3D	pseudo-3D	pseudo-3D	echt-3D	echt-3D	echt-3D	inkongruente Bahn	pseudo-3D	pseudo-3D	pseudo-3D	echt-3D	echt-3D	echt-3D	
		auditorische Stimulation																									
Dimension		ohne Ton	kongruent	inkongruent																							
visuelle Stimulation	kongruente Bahn	pseudo-3D	pseudo-3D	pseudo-3D																							
		echt-3D	echt-3D	echt-3D																							
	inkongruente Bahn	pseudo-3D	pseudo-3D	pseudo-3D																							
		echt-3D	echt-3D	echt-3D																							
<p>Experiment 2</p> <p>Extrapolation</p>	<table border="1"> <tr> <td colspan="2"></td> <td colspan="3">auditorische Stimulation</td> </tr> <tr> <td>Dimension</td> <td></td> <td>ohne Ton</td> <td>kongruent</td> <td>inkongruent</td> </tr> <tr> <td rowspan="4">visuelle Stimulation</td> <td rowspan="2">nahe Schätzung</td> <td>pseudo-3D</td> <td>pseudo-3D</td> <td>pseudo-3D</td> </tr> <tr> <td>echt-3D</td> <td>echt-3D</td> <td>echt-3D</td> </tr> <tr> <td rowspan="2">ferne Schätzung</td> <td>pseudo-3D</td> <td>pseudo-3D</td> <td>pseudo-3D</td> </tr> <tr> <td>echt-3D</td> <td>echt-3D</td> <td>echt-3D</td> </tr> </table>			auditorische Stimulation			Dimension		ohne Ton	kongruent	inkongruent	visuelle Stimulation	nahe Schätzung	pseudo-3D	pseudo-3D	pseudo-3D	echt-3D	echt-3D	echt-3D	ferne Schätzung	pseudo-3D	pseudo-3D	pseudo-3D	echt-3D	echt-3D	echt-3D	
		auditorische Stimulation																									
Dimension		ohne Ton	kongruent	inkongruent																							
visuelle Stimulation	nahe Schätzung	pseudo-3D	pseudo-3D	pseudo-3D																							
		echt-3D	echt-3D	echt-3D																							
	ferne Schätzung	pseudo-3D	pseudo-3D	pseudo-3D																							
		echt-3D	echt-3D	echt-3D																							
<p>Experiment 4</p> <p>Diskrimination</p>	<table border="1"> <tr> <td colspan="2"></td> <td colspan="3">auditorische Stimulation</td> </tr> <tr> <td>Dimension</td> <td></td> <td>ohne Ton</td> <td>kongruent</td> <td>inkongruent</td> </tr> <tr> <td rowspan="4">visuelle Stimulation</td> <td rowspan="2">kongruente Bahn</td> <td>pseudo-3D</td> <td>pseudo-3D</td> <td>pseudo-3D</td> </tr> <tr> <td>echt-3D</td> <td>echt-3D</td> <td>echt-3D</td> </tr> <tr> <td rowspan="2">inkongruente Bahn</td> <td>pseudo-3D</td> <td>pseudo-3D</td> <td>pseudo-3D</td> </tr> <tr> <td>echt-3D</td> <td>echt-3D</td> <td>echt-3D</td> </tr> </table>			auditorische Stimulation			Dimension		ohne Ton	kongruent	inkongruent	visuelle Stimulation	kongruente Bahn	pseudo-3D	pseudo-3D	pseudo-3D	echt-3D	echt-3D	echt-3D	inkongruente Bahn	pseudo-3D	pseudo-3D	pseudo-3D	echt-3D	echt-3D	echt-3D	
		auditorische Stimulation																									
Dimension		ohne Ton	kongruent	inkongruent																							
visuelle Stimulation	kongruente Bahn	pseudo-3D	pseudo-3D	pseudo-3D																							
		echt-3D	echt-3D	echt-3D																							
	inkongruente Bahn	pseudo-3D	pseudo-3D	pseudo-3D																							
		echt-3D	echt-3D	echt-3D																							
<p>Experiment 5</p> <p>Diskrimination</p>	<table border="1"> <tr> <td colspan="2"></td> <td colspan="3">auditorische Stimulation</td> </tr> <tr> <td>Dimension</td> <td></td> <td>ohne Ton</td> <td>kongruent</td> <td>inkongruent</td> </tr> <tr> <td rowspan="4">visuelle Stimulation</td> <td rowspan="2">kongruente Bahn</td> <td>einfach-2D</td> <td>einfach-2D</td> <td>einfach-2D</td> </tr> <tr> <td>echt-3D</td> <td>echt-3D</td> <td>echt-3D</td> </tr> <tr> <td rowspan="2">inkongruente Bahn</td> <td>einfach-2D</td> <td>einfach-2D</td> <td>einfach-2D</td> </tr> <tr> <td>echt-3D</td> <td>echt-3D</td> <td>echt-3D</td> </tr> </table>			auditorische Stimulation			Dimension		ohne Ton	kongruent	inkongruent	visuelle Stimulation	kongruente Bahn	einfach-2D	einfach-2D	einfach-2D	echt-3D	echt-3D	echt-3D	inkongruente Bahn	einfach-2D	einfach-2D	einfach-2D	echt-3D	echt-3D	echt-3D	
		auditorische Stimulation																									
Dimension		ohne Ton	kongruent	inkongruent																							
visuelle Stimulation	kongruente Bahn	einfach-2D	einfach-2D	einfach-2D																							
		echt-3D	echt-3D	echt-3D																							
	inkongruente Bahn	einfach-2D	einfach-2D	einfach-2D																							
		echt-3D	echt-3D	echt-3D																							
<p>Experiment 6</p> <p>Detektion/ Diskrimination</p>	<table border="1"> <tr> <td colspan="2"></td> <td colspan="2">Dimension</td> </tr> <tr> <td rowspan="4">visuelle Stimulation</td> <td>einfach-2D</td> <td>echt-3D</td> <td></td> </tr> <tr> <td>kongruente Bahn</td> <td>kongruente Bahn</td> <td></td> </tr> <tr> <td>einfach-2D</td> <td>echt-3D</td> <td></td> </tr> <tr> <td>inkongruente Bahn</td> <td>inkongruente Bahn</td> <td></td> </tr> </table>			Dimension		visuelle Stimulation	einfach-2D	echt-3D		kongruente Bahn	kongruente Bahn		einfach-2D	echt-3D		inkongruente Bahn	inkongruente Bahn										
		Dimension																									
visuelle Stimulation	einfach-2D	echt-3D																									
	kongruente Bahn	kongruente Bahn																									
	einfach-2D	echt-3D																									
	inkongruente Bahn	inkongruente Bahn																									

Abbildung 1. Übersicht über die sechs durchgeführten Hauptexperimente. Abgebildet sind die im jeweiligen Experiment verwendete Aufgabenstellung, die konkrete Faktorengestaltung des Designs sowie ein beispielhaftes visuelles Stimulusbild. Weitere Details zur verwendeten Methode der jeweiligen Experimente sind dem Haupttext zu entnehmen.

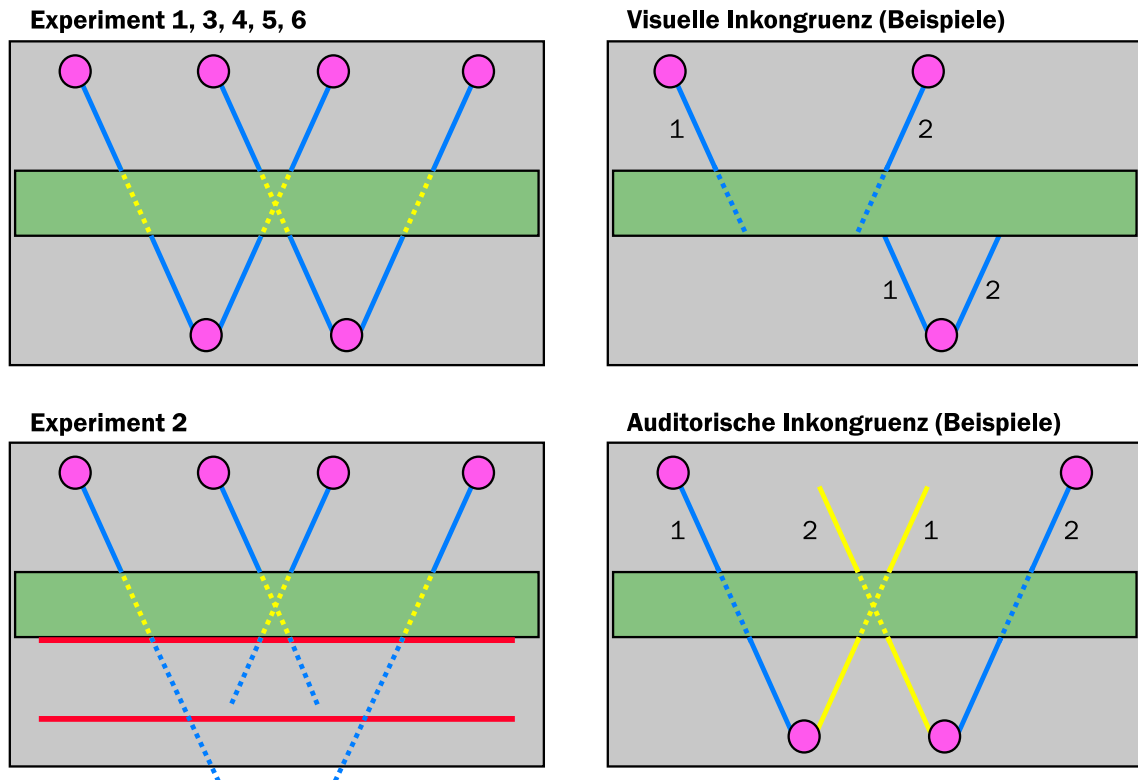


Abbildung 2. Schematische Übersicht (Draufsicht) zu den visuellen Bahnverläufen und der auditorischen Stimulation (links) sowie Beispiele für inkongruente Bedingungen (rechts). In den Experimenten 1, 3, 4, 5 und 6 (links oben) konnte sich die Kugel auf vier möglichen Bahnen bewegen: links, mittig links, mittig rechts und rechts (blaue Linien). Diese Bewegung wurde in bimodalen Bedingungen von einem Ton begleitet, der während der Verdeckung gedämpft wurde (gelbe Linien). Die Kugel war während der Verdeckung (grüner Balken) nicht sichtbar (gepunktete Linien, durchgezogene Linien kennzeichnen visuell sichtbare Bahnbewegungen). In Experiment 6 entfiel die auditorische Stimulation. In Experiment 2 (links unten) startete die Kugel auf einer der vier möglichen Bahnen, erschien aber nach der Verdeckung nicht wieder. Die Szene wurde ggf. inklusive Ton bis zu einem Zeitpunkt von 1533 ms nach dem tatsächlichen Erreichen der Markierung (rote Linie) präsentiert (hier beispielhaft dargestellt durch die verschiedenen Linienlängen nach der Verdeckung). In visuell inkongruenten Durchgängen (rechts oben) erschien die Kugel auf der jeweils anderen Bahn der Bewegungsrichtung wieder. Als Beispiele: Kugel 1 startet auf der linken Bahn und taucht auf der mittig linken wieder auf. Kugel 2 startet auf der mittig rechten Bahn und taucht auf der rechten wieder auf. Bei auditorisch inkongruenten Durchgängen (rechts unten) wurde die visuelle Bewegung mit einem Ton für die andere Bewegungsrichtung gepaart. Als Beispiele: Kugel 1 bewegt sich auf der visuell linken Bahn und wird mit einem mittig rechten Ton kombiniert. Kugel 2 bewegt sich auf der visuell rechten Bahn und wird mit einem mittig linken Ton kombiniert. Um die Abstände zwischen den Tonbewegungen für inkongruente Bedingungen konstant zu halten, wurde eine innere visuelle Bahn bei inkongruenter Paarung immer mit der äußeren auditorischen Bahn der Gegenrichtung und eine äußere visuelle Bahn mit der inneren auditorischen Bahn der Gegenrichtung kombiniert. Weitere Details zur verwendeten Methode der jeweiligen Experimente sind dem Haupttext zu entnehmen. Abbildung adaptiert nach: Dittrich und Noesselt (2018), S. 5.

4.3 Geräte und Stimuli

Alle Experimente wurden mit MATLAB (R2012b, 8.00.783, The MathWorks, USA) und der Erweiterung Psychophysics Toolbox 3.011 (Brainard, 1997; Kleiner et al., 2007) programmiert und präsentiert. Zur Erzeugung des echt-3D-Eindrucks wurde ein stereoskopischer Bildschirm (Planar SD2220W; Planar Systems, USA) verwendet. Dieser bestand aus zwei separaten Bildschirmen, wobei über einem standardgemäß aufgestellten Monitor in einem Winkel von 110 ° ein zweiter Monitor montiert war. Zwischen diesen beiden Bildschirmen befand sich ein halbdurchlässiger Spiegel, der das Bild des unteren Bildschirms durchlässt und das des oberen spiegelt, sodass visuelle Inhalte gleichzeitig von beiden Monitoren wahrgenommen werden können. Präsentiert man auf beiden Bildschirmen horizontal leicht versetzte Bilder, können die Probanden die visuelle Stimulation in 3D wahrnehmen, ähnlich wie während eines 3D-Kinofilms, wenn sie eine polarisierte 3D-Brille tragen.

Beide Bildschirme wiesen eine Bildwiederholungsrate von 60 Hz, eine Auflösung von 1920 x 1080 Pixeln und eine Größe von 21 Zoll auf. Der untere Bildschirm befand sich in einem Abstand von 50 cm zu einer Kinnstütze. Auf einem imaginären Rechteck zwischen dem unteren Bildschirm und der Kinnstütze befanden sich auf Kopfhöhe vier Lautsprecher (Ekulit ESM 50 A) an den jeweiligen Ecken des Rechtecks. Jeweils zwei Lautsprecher waren horizontal 60 cm bzw. in der Tiefe 35 cm voneinander entfernt. Das experimentelle Set-Up aus Bildschirm, Lautsprechern und Probandenposition kann Abbildung 3 entnommen werden. Die Antwortfassung erfolgte mittels Tastendruck auf einer USB-Tastatur (Damian, 2010).

Das visuelle Stimulationsmaterial wurde mit der 3D-Animationssoftware Blender (Version 2.75a, www.blender.org) erstellt. Eine hellbeige halboffene Box mit dunkelgrau-weißem Schachbrett-Boden befand sich vor einem einfarbigen grauen Hintergrund. Die Box war 23.94° bis 45.08° breit und 16.50° bis 20.96° tief. In der Mitte befand sich eine hellgrüne Brücke (Breite: 31.82°) mit eckiger Öffnung. Der Teil der Brücke, der die Kugelbewegung verdeckte, war 4.24° tief. Vom hinteren Ende der Box bewegte sich eine magentafarbene Kugel mit einem mittelblauen Querstreifen auf einer diagonalen Bahn auf den Betrachter zu. Sie hatte zu Beginn der Bewegung einen Durchmesser von 1.15°, bei Bewegungsende von 1.83°. Die Kugel konnte sich auf vier verschiedenen Bahnen auf den Betrachter zubewegen. Ihre Startpositionen waren bei $\pm 6.41^\circ$ bzw. $\pm 1.83^\circ$ von der Bildmitte am oberen Rand der Box lokalisiert, die Bewegung endete bei $\pm 2.86^\circ$ von der Bildmitte am unteren Rand der Box. Dadurch ergaben sich die vier Bewegungsbahnen links, mittig links, mittig rechts und rechts (siehe auch Abbildung 2). Die Endpositionen für die linke und mittig rechte sowie die rechte und mittig linke Bahn waren identisch. Die Kugel konnte sich nach der Brücke auf einer visuell passenden oder nicht passenden Bahn fortbewegen. Bei einer visuell nicht passenden Bahn folgte sie nach dem Verdecken der jeweils anderen Bahn für diese Bewegungsrichtung (z. B. mittig links nach zuvor links oder rechts nach zuvor mittig rechts, siehe auch Abbildung 2). Im 2. Experiment erschien die Kugel nach der Brücke nicht wieder (Details dazu siehe Abschnitt 5.2).

Die visuellen Stimuli in den Experimenten 1 bis 4 enthielten in der pseudo-3D Variante Tiefencues, vor allem eine perspektivische Darstellung und Schattenwurf standen im Vordergrund. Für Experiment 5 und 6 wurde das Material in der Vergleichsbedingung vereinfacht in

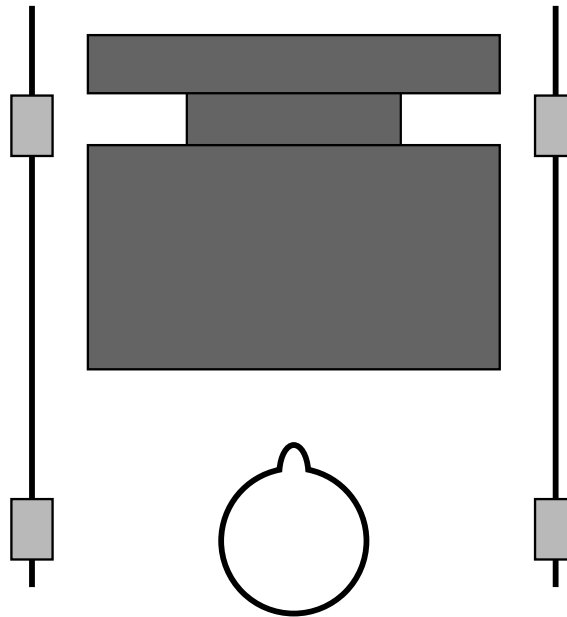


Abbildung 3. Schema des experimentellen Set-Ups (Draufsicht) bestehend aus 3D-Bildschirm (dunkelgrau) und Lautsprechern (hellgrau, montiert auf einem Metallgestell). Der Abstand des Probanden (weiß) zum unteren Bildschirm wurde über eine Kinnstütze konstant gehalten. Weitere Details zum experimentellen Set-Up können dem Haupttext entnommen werden. Abbildung adaptiert nach: Dittrich und Noesselt (2018), Supplementary Material, S. 1.

2D dargestellt (einfach-2D, Details siehe Abschnitt 6.3). Die Bilder für die echt-3D-Bedingungen hatten einen horizontalen Versatz zwischen den Bildern des oberen und unteren Bildschirms von 0.69° bis 2.98° . Der jeweils verwendete Versatz wurde für jeden Probanden in einem Vortest festgelegt, in dem dieser aus sechs Versatzvarianten wählen konnte (Details siehe Durchführung, Abschnitt 4.4).

Über die vier Lautsprecher wurde in den Experimenten 1 bis 5 eine Bewegung mittels verschiedenen Lautstärkepegeln erzeugt, die sich auf den Betrachter zubewegte. Die mittlere Lautstärke betrug 78 dB(A) . Sie wurde um durchschnittlich 5 dB(A) für den Zeitraum gedämpft, in dem die Kugel während der Bewegung durch die Brücke verdeckt wurde. Bei den Tönen handelte es sich um ein kontinuierliches weißes Rauschen mit einem Frequenzbereich von $1000 - 2500 \text{ Hz}$ und einer Sample Rate von 44100 Hz . Wurde ein zur visuellen Bewegung passender Ton präsentiert, bewegte sich der Ton in die gleiche Richtung (links nach rechts bzw. rechts nach links) auf die Versuchsperson zu. Bei einem nicht passenden Ton bewegte er sich entgegengesetzt zur visuellen Bewegungsrichtung (siehe auch Abbildung 2).

4.4 Durchführung

Das Experiment wurde jeweils zweimal direkt hintereinander durchgeführt. Einmal in der Variante pseudo-3D/einfach-2D und einmal in echt-3D. Die Reihenfolge beider Teile war zwischen den Probanden gemeinsam mit ihrem Geschlecht¹ ausbalanciert. Vor dem echt-3D-

¹In einigen Studien konnten Geschlechterunterschiede bei der Verwendung von Vorhersageaufgaben gefunden werden (z. B. Caird & Hancock, 1994; Schiff & Oldak, 1990). Da das Geschlecht als Einflussfaktor jedoch in dieser Arbeit kein zu untersuchender Bestandteil war, wurde es zwischen den Probanden ausbalanciert, um etwaige Konfundierungen zu vermeiden.

Teil konnte die Versuchsperson ein beispielhaftes Szenenbild mit den sechs verschiedenen zur Verfügung stehenden Versatzarten betrachten. Über die Tastatur konnte der Versatz zwischen den Bildern der beiden Monitore geändert werden. Die Probanden wählten den Versatz, bei dem sie den 3D-Eindruck am größten, aber auch natürlichsten empfanden. Beide Experimentalteile begannen mit einem Übungsblock bestehend aus 6 Durchgängen. Im Anschluss folgten 16 Experimentalblöcke mit jeweils 24 Durchgängen. Zwischen den Blöcken gab es jeweils eine von den Probanden selbst zu terminierende Pause. Die Aufgabenstellung bestand je nach Experiment aus einer Detektions-, Diskriminations- oder Extrapolationsaufgabe (Details siehe jeweiliges Experiment).

Zu Beginn eines Durchgangs ruhte die Kugel für 200 ms an ihrer Startposition, im Anschluss bewegte sie sich für 4167 ms auf den Betrachter zu (mit Ausnahme von Experiment 2, hier erschien die Kugel nach der Brücke nicht wieder). Während ihrer Bewegung wurde die Kugel für 567 ms von der Brücke verdeckt. Das Intertrialintervall betrug bei allen Experimenten 500 ms.

4.5 Auswertung

4.5.1 Datenaufbereitung

Vor der Auswertung wurden Trials entfernt, in denen die Flipzeiten durch den Monitor nicht gewährleistet wurden. So gingen in die Auswertung nur Durchgänge ein, in denen die Synchronität zwischen visueller und auditorischer Stimulation und/oder eine durchgängige visuelle Bewegung sichergestellt werden konnten. Zudem wurden Durchgänge ausgeschlossen, in denen nicht die vorgesehene(n) Taste(n) gedrückt wurde(n) sowie diejenigen Trials, deren Reaktionszeiten außerhalb des Intervalls von $\pm 2 SD$ auf Versuchspersonen-, Experimentaltteil- und Bedingungsebene lagen. Für die Diskriminationsexperimente (Experimente 3, 4 und 5 sowie dem Telexperiment in Experiment 6) wurden für die Analyse der Reaktionszeiten zusätzlich Trials mit nicht korrekten Antworten ausgeschlossen.

4.5.2 Reaktionszeit-Analyse

Es wurden Varianzanalysen (ANOVAs) mit Messwiederholung für die drei Faktoren Dimension, Visuell und Auditorisch (für Experiment 6 siehe Abschnitt 7) mit SPSS 23.0 (IBM, USA) durchgeführt. Für die Experimente wurde als abhängige Variable die Reaktionszeit betrachtet (in Experiment 2 handelte es sich hierbei um Differenzen, siehe Abschnitt 5.2). Für die Diskriminationsexperimente wurden ebenfalls die korrekten Antworten analysiert, um Hinweise auf einen Speed-Accuracy-Trade-Off in den Daten ausschließen zu können. Etwaige Post-hoc-Tests der ANOVAs wurden nach Bonferroni korrigiert sowie bei verletzter Sphärizitätsbedingung eine Freiheitsgradkorrektur nach Greenhouse-Geisser angewandt.

5 Zeitliche Bewegungsvorhersage (Detektion und Extrapolation)

5.1 Experiment 1: Detektion

In Experiment 1 wurde eine Detektionsaufgabe verwendet, bei der temporale Aspekte im Vordergrund standen. Die Probanden sollten daher vermehrt auditorische Informationen für die Aufgabenbearbeitung nutzen (Welch & Warren, 1980; Welch et al., 1986). Zudem sollten sie Vorteile für visuell kongruente Bahnverläufe zeigen, da diese den Erwartungen an die Bewegung entsprechen (Kawachi & Gyoba, 2006). Darüber hinaus wurde verglichen, ob sich Unterschiede zwischen der pseudo- und echt-3D-Präsentation ergeben. Wenn Disparitätsinformationen einen entscheidenden Faktor darstellen, sollte sich dies entsprechend z. B. in schnelleren Reaktionszeiten oder stärker ausgeprägten Erleichterungseffekten in der echt-3D-Darstellung widerspiegeln (Conrad et al., 2013; Harrison et al., 2015).

5.1.1 Methoden

An Experiment 1 nahmen insgesamt 45 Probanden teil, davon wurden acht von der Datenanalyse ausgeschlossen. Diese tendierten zu antizipierendem Verhalten: Ihre durchschnittliche Reaktionszeit für mindestens einen der beiden Experimentalteile lag unter 120 ms. Ein weiterer Proband wurde aufgrund von fehlendem Instruktionsverständnis ausgeschlossen. Die durchschnittliche Reaktionszeit beider Experimentalteile wurde direkt im Anschluss an das Experiment berechnet und die ausgeschlossenen Probanden entsprechend der Experimentalreihenfolge und dem Geschlecht ersetzt. Somit wurden die Daten von 36 Probanden (je 18 männlich/weiblich) analysiert. Diese waren im Mittel 22.50 Jahre ($SD = 2.49$) alt, zwei davon waren links-, einer beid- und 33 rechtshändig.

Die Probanden absolvierten in diesem Experiment eine einfache Detektionsaufgabe. Das heißt, sie drückten die linke Pfeiltaste der Tastatur, sobald die Kugel nach dem Verdecken durch die Brücke wiedererschien. Abbildung 4 zeigt beispielhaft die Start- und Endposition der Kugel für eine Bewegungsbahn in Experiment 1. Die übrigen Methodendetails können den Allgemeinen Methoden (siehe Abschnitt 4) entnommen werden.

5.1.2 Ergebnisse

Es wurden insgesamt 4.96 % der Trials von der Datenanalyse ausgeschlossen, wobei der Anteil aufgrund von Flipungenauigkeiten entfernter Durchgänge 0.12 % betrug.

Bei der Analyse der Reaktionszeiten ergaben sich zwei signifikante Haupteffekte der Faktoren Visuell, $F(1, 35) = 105.00$, $p < .001$, und Auditorisch, $F(1.25, 43.75) = 79.40$, $p < .001$. Der Haupteffekt Dimension wurde nicht signifikant, $F(1, 35) = 0.01$, $p = .917$. Es zeigte sich außerdem eine signifikante Interaktion zwischen Dimension und Visuell, $F(1, 35) = 4.91$, $p = .033$. Alle weiteren Effekte wurden nicht signifikant: Dimension \times Auditorisch, $F(1.46, 51.17) = 1.64$, $p = .208$; Visuell \times Auditorisch, $F(2, 70) = 0.89$, $p = .414$, und Dimension \times Visuell \times Auditorisch, $F(2, 70) = 2.16$, $p = .123$. Die Reaktionszeiten in Abhängigkeit von der visuellen und auditorischen Stimulation sind in Abbildung 5 a dargestellt.

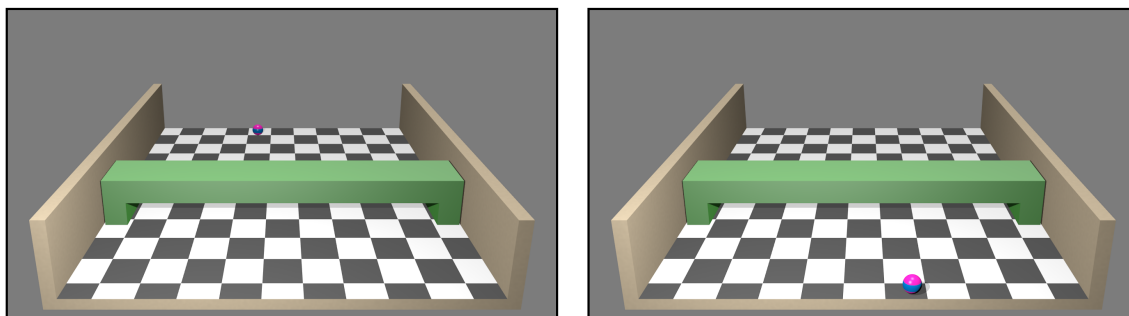


Abbildung 4. Beispielhaftes Bildmaterial aus Experiment 1 und 3. Links ist die Start-, rechts die Endposition der Kugelbewegung abgebildet. Die Kugel bewegte sich auf einer von vier möglichen Bahnen vom hinteren Rand der Box auf den Betrachter zu und wurde kurzzeitig durch die Brücke verdeckt. Sobald die Kugel auf der anderen Seite der Brücke wiedererschien, reagierten die Probanden mit Knopfdruck (Experiment 1) oder gaben an, ob sich die Kugel auf einer passenden Bahn nach der Verdeckung weiterbewegte (Experiment 3). Abbildung adaptiert nach: Dittrich und Noesselt (2018), S. 4.

Die Probanden reagierten schneller bei visuell passenden (241 ms) als bei visuell nicht passenden Bahnen (255 ms). Sie waren für auditorisch passende (241 ms) und nicht passende Töne (240 ms) schneller als in der unimodalen Bedingung (263 ms), $p < .001$. Die beiden Tonbedingungen unterschieden sich nicht voneinander, $p = 1.0$. Die Interaktion zwischen Dimension und visueller Stimulation zeigte sich in einem größeren visuellen Erleichterungseffekt (15 ms) während pseudo-3D-Stimulation verglichen zu den echt-3D-Bedingungen (12 ms), $t(35) = -2.13$, $p = .040$ (siehe Abbildung 5 b).

5.1.3 Diskussion

In Experiment 1 wirkte sich die zusätzliche Tonpräsentation positiv auf die Leistung der Probanden aus, sie reagierten schneller bei gleichzeitigen Tönen als in der unimodalen Bedingung. Somit stellten die Töne für die einfache temporale Detektionsaufgabe einen hilfreichen Hinweis dar: Durch das Absenken der Lautstärke während der Verdeckung konnte der lauter werdende Ton bei Wiedererscheinen für die Detektion genutzt werden. Dies spricht für die Annahme, dass auditorische Informationen eine hohe Relevanz bei zeitlichen Aufgaben durch die Überlegenheit der zeitlichen Auflösung des auditorischen Systems einnehmen (Welch et al., 1986). Die Richtung der Tonbewegung war hingegen nicht von Bedeutung. Beide Töne (passend/nicht passend zur visuellen Bewegungsrichtung) förderten gleichermaßen die Reaktion auf das Wiedererscheinen der Kugel. Da die Rauminformationen der Töne (auditorische Bewegungsrichtung) keinen weiteren Informationsvorteil für die Aufgabe bieten, wurden diese von den Probanden entsprechend nicht genutzt. Ob alternativ die Nichtnutzung der Tonrichtung darauf zurückzuführen ist, dass die Probanden die Bewegungsrichtung dieser nicht unterscheiden können, wurde in einem Kontrollexperiment getestet (siehe Abschnitt 5.3). Weitere Gründe für das Ausbleiben eines audiovisuellen Kongruenzeffekts werden in der Allgemeinen Diskussion (siehe Abschnitt 8.2) erörtert.

Des Weiteren reagierten die Probanden schneller, wenn sich die Kugel nach der Brücke passend zu ihrer Bewegungsbahn vor der Verdeckung weiterbewegte. Dies bestätigt den

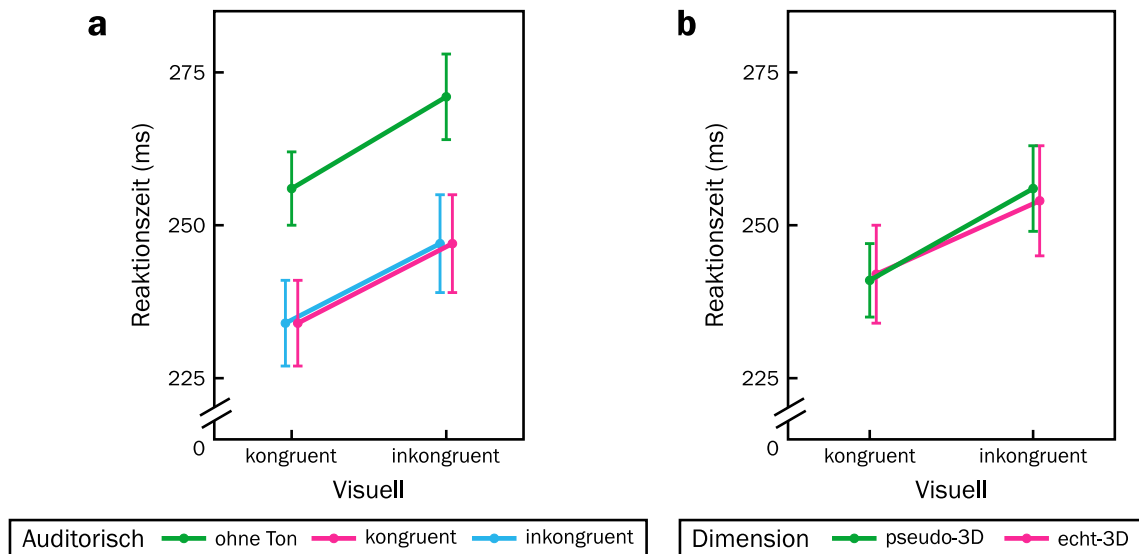


Abbildung 5. Reaktionszeit in Abhängigkeit vom visuellen Bahnverlauf nach der Verdeckung und von der auditorischen Stimulation (a) sowie von der visuellen Stimulation und der Präsentationsdimension (b). Die Fehlerbalken kennzeichnen den Standardfehler des Mittelwerts. Zur Unterscheidung der Fehlerbalken sind die Datenpunkte leicht versetzt dargestellt. Abbildung Teil (a) adaptiert nach: Dittrich und Noesselt (2018), S. 6.

visuellen Erleichterungseffekt für kongruente Bahnverläufe (Kawachi & Gyoba, 2006), wobei dieser mit der Präsentationsdimension interagierte und in der pseudo-3D-Bedingung größer ausfiel als in der echt-3D-Präsentation. Dies steht im Gegensatz zu bisherigen Studien. González et al. (2010) konnten zeigen, dass Disparität (das Unterscheidungsmerkmal zwischen pseudo- und echt-3D-Präsentation in diesem Experiment) einen entscheidenden Faktor für die Bewegungswahrnehmung in der Tiefe darstellt. Darüber hinaus war der Erleichterungseffekt in den Experimenten von Harrison et al. (2015) für eine echt-3D-Umgebung größer. Diese nutzten allerdings eine Diskriminationsaufgabe für die auditorische Bewegung, während hier eine visuelle Detektionsaufgabe verwendet wurde. Zudem lag der Fokus hier auf der Reaktionszeit-Auswertung, wohingegen Harrison et al. (2015) die korrekten Antworten analysierten. Auch hierfür werden mögliche Gründe für diese gegenläufigen Ergebnisse in der Allgemeinen Diskussion ausführlicher erörtert (siehe Abschnitt 8.6).

5.2 Experiment 2: Extrapolation

In Experiment 2 sollte der Einfluss auditorischer und visueller Informationen für eine zeitlich komplexere Aufgabe untersucht werden. Die Probanden schätzten, wann die Kugel eine bestimmte Position erreichen würde (Time-to-Arrival-/Contact-Schätzung). Da die Schätzung der Ankunftszeit mit zunehmender Zeit, in der das sich bewegende Objekt nicht sichtbar ist, schlechter wird (Lyon & Waag, 1995; Schiff & Oldak, 1990), sollten die Probanden in der fernen Schätzdistanz eine schlechtere Leistung erzielen. Außerdem sollten sich ähnlich wie in Experiment 1 Vorteile für Durchgänge mit auditorischer Stimulation zeigen. Hierbei stellt sich die Frage, ob im Gegensatz zu Experiment 1 ein Kongruenzeffekt gefunden werden kann und die Probanden während der komplexeren Extrapolationsaufgabe eher von räumlich kongruenten Tönen als inkongruenten profitieren oder ob sich hier erneut keine Unterschiede zwischen den Tonbedingungen finden lassen.

5.2.1 Methoden

An Experiment 2 nahmen insgesamt 32 Probanden (je 16 männlich/weiblich) teil. Die Versuchspersonen waren im Durchschnitt 23.34 Jahre alt ($SD = 3.70$). Es waren drei Teilnehmer links-, einer beid- und 28 rechtshändig.

Die Aufgabe bestand in diesem Experiment darin, anzugeben, zu welchem Zeitpunkt die Kugel einen bestimmten Ort erreichen würde. Daher tauchte die Kugel nach dem Verschwinden unter der Brücke nicht wieder auf. Für die zwei verschiedenen Schätzdistanzen (nah/fern) wurde der visuellen Szene entsprechend an zwei verschiedenen Positionen eine 0.34° breite rote Stange hinzugefügt. Diese befand sich entweder im Abstand von 0.46° (nahe Schätzung) oder 4.92° (ferne Schätzung) vor der Brücke aus Sicht der Probanden (siehe Abbildung 6). Die Probanden gaben mit der linken Pfeiltaste an, wann die Kugel die rote Stange erreichen würde, wenn diese sich nach der Verdeckung genauso fortbewegt hätte wie zuvor. Für die nahe Schätzung war die gesamte visuelle Szene insgesamt 4000 ms, für die ferne Schätzung 4667 ms zu sehen. So betrug das Antwortfenster, das nach der tatsächlichen Ankunftszeit zur Verfügung stand, für beide Distanzen insgesamt 1533 ms. Der Ton in bimodalen Durchgängen wurde bis zum Ende des Durchgangs präsentiert. Vor dem Übungsblock jedes Experimentalteils wurden den Probanden die vier Bewegungsbahnen der Kugel gezeigt, sodass sie einen Eindruck von der vollständigen Bewegung gewinnen konnten.

Da die Kugel nach der Brücke nicht wieder erschien, bestand der visuelle Faktor in diesem Experiment entsprechend nicht aus den Stufen visuell passend/nicht passend, sondern aus der nahen und fernen Schätzung. Als abhängige Variable wurde in diesem Experiment die absolute Abweichung der Reaktionszeit von der „tatsächlichen“ Ankunftszeit der Kugel genutzt, zu der die Kugel die Stange erreicht hätte, wenn sie sich kontinuierlich und sichtbar weiterbewegt hätte. Die Nutzung der relativen Abweichung weist zwar den Vorteil auf, dass diese eine Richtung der Abweichung angibt, jedoch unterschätzt sie bei der arithmetischen Mittlung die tatsächliche Abweichung durch das Eingehen positiver und negativer Werte. Um die Abweichung als Leistungsmaß der Probanden nicht systematisch zu unterschätzen, wurde die

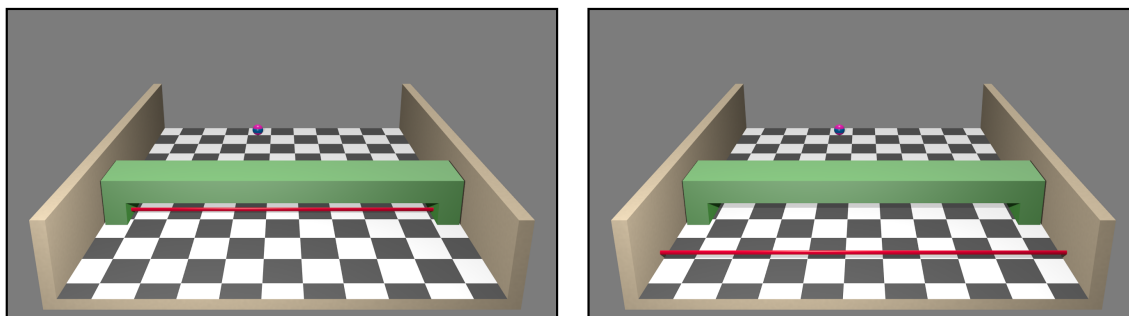


Abbildung 6. Beispielhaftes Bildmaterial aus Experiment 2. Die rote Stange kennzeichnet die Schätzdistanz: Links ist die nahe, rechts die ferne Schätzdistanz abgebildet. Die Kugel konnte sich auf vier verschiedenen Bahnen auf den Betrachter zubewegen, erschien jedoch nach der Brücke nicht wieder und die Probanden gaben per Tastendruck an, wann die Kugel die rote Stange erreichen würde. Abbildung adaptiert nach: Dittrich und Noesselt (2018), S. 4.

absolute Abweichung gewählt. Zusätzlich wurde zur absoluten Reaktionszeit-Abweichung als Maß für die Akkuratheit der Schätzung die Standardabweichung dieser als Variabilitätsmaß und damit Ausdruck der Präzision der Schätzung analysiert (Lugtigheid & Welchman, 2011).

5.2.2 Ergebnisse

Von der Datenauswertung wurden 4.61 % der Durchgänge ausgeschlossen, davon waren 0.15 % Trials mit nicht exakten Flipzeiten.

Bei der Analyse der Reaktionszeit-Abweichungen ergaben sich zwei signifikante Haupteffekte von Visuell, $F(1, 31) = 14.76$, $p = .001$, und Auditorisch, $F(1.06, 32.83) = 5.82$, $p = .020$. Die weiteren Effekte überschritten das Signifikanzniveau nicht: Dimension, $F(1, 31) = 0.03$, $p = .870$; Dimension \times Visuell, $F(1, 31) = 0.21$, $p = .648$; Dimension \times Auditorisch, $F(1.30, 40.17) = 1.06$, $p = .330$; Visuell \times Auditorisch, $F(1.04, 32.34) = 0.20$, $p = .667$, sowie Dimension \times Visuell \times Auditorisch $F(1.44, 44.73) = 0.68$, $p = .466$. Die Probanden wiesen hierbei eine geringere Abweichung für die nahe Distanz (226 ms) als für ferne Schätzungen (381 ms) auf. Zudem war die Abweichung unimodal (288 ms) kleiner als für auditorisch passende Bedingungen (313 ms), $p = .048$ sowie tendenziell ebenso verglichen zu auditorisch nicht passenden Bedingungen (310 ms), $p = .085$. Beide Tonbedingungen unterschieden sich nicht voneinander, $p = .629$. Die absoluten Reaktionszeit-Abweichungen in Abhängigkeit von der Schätzdistanz sowie der auditorischen Stimulation sind in Abbildung 7 zu sehen.

5.2.3 Post-hoc-Betrachtung interindividueller Unterschiede

Die Ergebnisse der Extrapolationsdaten scheinen einen negativen Einfluss zusätzlicher Tonstimulation auf die Schätzleistung herauszustellen. Dies steht im klaren Gegensatz zu Experiment 1, in dem ebenfalls eine zeitliche Aufgabe genutzt wurde und gleichzeitig präsentierte Töne die Leistung verbesserten. Daher wurden die Ergebnisse post-hoc probandenweise im Detail betrachtet. Augenscheinlich profitieren die meisten Probanden nur für eine Schätzdistanz von Tönen, bei der anderen Distanz wirken sich diese eher hinderlich aus. Um diese Vermutung zu bestätigen, wurde der Datensatz in zwei Hälften aufgeteilt: Jeweils eine wurde genutzt, um

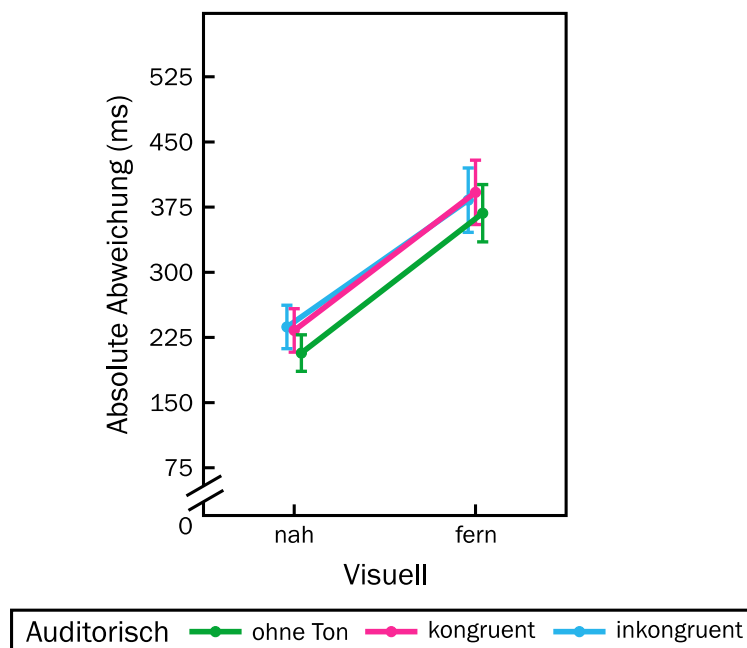


Abbildung 7. Absolute Reaktionszeit-Abweichung in Abhängigkeit von der Schätzdistanz und der auditorischen Stimulation. Die Fehlerbalken kennzeichnen den Standardfehler des Mittelwerts. Zur Unterscheidung der Fehlerbalken sind die Datenpunkte leicht versetzt dargestellt. Abbildung adaptiert nach: Dittrich und Noesselt (2018), S. 6.

die Probanden gemäß ihres Nutzungsverhaltens zu klassifizieren. Die andere Hälfte diente als Analysedatensatz, um das Schätzverhalten in Abhängigkeit von den experimentellen Faktoren zu untersuchen. Zur Crossvalidierung der Ergebnisse diente jede Datensatzhälfte einmal als Kategorisierungs- und einmal als Analysedatensatz¹. Die Aufteilung der Daten erfolgte anhand der Blocknummer in Daten aus geraden und ungeraden Blöcken. Im Analysedatensatz erfolgte die Auswertung per ANOVA mit den entsprechenden experimentellen Faktoren für die absolute Reaktionszeit-Abweichung sowie deren Standardabweichung (s. o.).

Nutzergruppeneinteilung Im Klassifizierungsdatensatz wurden die Probanden entsprechend ihres auditorischen Erleichterungseffekts eingeteilt. Dafür wurde die Differenz zwischen unimodaler und auditorisch passender Bedingung für jeweils beide Schätzdistanzen berechnet. Entsprechend des Vorzeichens wurde für jeden Probanden festgestellt, ob dieser für die nahe und die ferne Distanz Töne nutzte oder nicht. Beide Erleichterungseffekte für nah und fern ergaben vier verschiedene Kombinationen und dementsprechende Nutzergruppen für auditorische Informationen: 1) Beidtonnutzer (nutzen Töne für beide Distanzen), 2) Nahtonnutzer (nutzen Töne nur für die nahe Distanz), 3) Ferntonnutzer (nutzen Töne nur für die ferne Distanz) und 4) Nichtnutzer (nutzen für beide Distanzen keine Töne). Tabelle 1 stellt die Kategorisierungsergebnisse für beide Datensätze (gerade und ungerade) dar. Da in beiden Kategorisierungsdatensätzen die Beidton- und Nichttonnutzer jeweils eine Gruppe von $n < 10$ darstellten, wurden zur Auswertung der Analysedatensätze nur die Nah- und Ferntonnutzer

¹Für ein ähnliches Vorgehen zur Untersuchung individueller Unterschiede über die Kategorisierung und Analyse auf Basis verschiedener (Teil-)Datensätze siehe u. a. Albrecht et al. (2010); Albrecht und Mattler (2016); Hagmann und Russo (2016).

herangezogen. Die numerischen Werte der Erleichterungseffekte für beide Distanzen und die entsprechend zugeteilte Nutzergruppe sind für beide Kategorisierungsdatensätze (gerade vs. ungerade) den Tabellen A1 und A2 im Anhang zu entnehmen¹.

Analyse der Reaktionszeit-Abweichungen der Nahtonnutzer Zunächst erfolgte die Analyse der Daten aus ungeraden Blöcken nachdem die Kategorisierung auf Basis gerader Blöcke vorgenommen wurde. Für die Nahtonnutzer ($n = 11$) konnte für die absoluten Reaktionszeit-Abweichungen je ein signifikanter Haupteffekt für Visuell, $F(1, 10) = 30.71$, $p < .001$, und Auditorisch, $F(1.13, 11.26) = 9.66$, $p = .008$, gefunden werden. Ebenso wurde die Interaktion beider Faktoren signifikant, $F(2, 20) = 47.49$, $p < .001$. Die restlichen Effekte überschritten das Signifikanzniveau nicht: Dimension, $F(1, 10) = 0.03$, $p = .873$; Dimension \times Visuell, $F(1, 10) = 0.04$, $p = .843$; Dimension \times Auditorisch, $F(2, 20) = 1.02$, $p = .380$, sowie Dimension \times Visuell \times Auditorisch $F(1.21, 12.11) = 0.34$, $p = .613$. Aufgrund des disordinalen Interaktionseffekts (siehe Abbildung 8 oben links) wird auf eine Post-hoc-Interpretation der signifikanten Haupteffekte verzichtet. Die genauere Betrachtung der Interaktion offenbarte einen signifikanten Unterschied der unimodalen Vorhersage von beiden auditorischen Bedingungen für jeweils beide Schätzdistanzen, $p \leq .012$. Die Probanden wichen in der nahen Schätzdistanz weniger von der tatsächlichen Ankunftszeit ab, wenn zusätzlich ein Ton präsentiert wurde: Die Abweichung lag unimodal bei 155 ms verglichen zu 108 ms für die Schätzung mit passendem Ton und 112 ms mit nicht passendem Ton. Für die ferne Distanz verhielt es sich umgekehrt, die Abweichung der Probanden war hier mit Ton (477 ms für passenden, 469 ms für nicht passenden Ton) größer als ohne zusätzliche auditorische Stimulation (355 ms). Für beide Entfernungen ergab sich kein Unterschied in der Schätzleistung für die beiden Tonarten (passend/nicht passend), $p = 1.00$.

Tabelle 1

Anzahl der Probanden für die jeweilige Nutzergruppeneinteilung für gerade und ungerade Blöcke

Nutzergruppe	Anzahl Probanden (gerade)	Anzahl Probanden (ungerade)
Beidtonnutzer	5	4
Nahtonnutzer	11	9
Ferntonnutzer	10	11
Nichtnutzer	6	8

Anmerkungen. Anzahl der Probanden gemäß Kategorisierung für auditorische Erleichterungseffekte in Beidtonnutzer (nutzen Töne für beide Schätzdistanzen), Nahtonnutzer/Ferntonnutzer (nutzen Töne jeweils nur für die nahe/ferne Distanz) und Nichtnutzer (profitieren für beide Distanzen nicht von Tönen). Dargestellt für die Kategorisierung auf Basis gerader Blöcke (2. Spalte) bzw. ungerader Blöcke (3. Spalte). Details zur Nutzergruppeneinteilung siehe Text. Tabelle adaptiert nach: Dittrich und Noesselt (2018), S. 6.

¹Mit den Begriffen „(Nicht)Nutzer“ bzw. „(nicht) nutzen“ ist im Zusammenhang mit der Post-hoc-Analyse in diesem Experiment gemeint, dass die Probanden einen entsprechend ausgeprägten auditorischen Erleichterungseffekt aufweisen und somit in den dazugehörigen Bedingungen von zusätzlichen Tönen profitieren bzw. nicht von diesen profitieren. Es soll damit nicht zum Ausdruck gebracht werden, dass die Probanden bei entsprechender Kategorisierung und Bezeichnung als „Nichtnutzer“ auditorische Informationen vollständig nicht nutzen.

Die Auswertung der Daten gerader Blöcke nach erfolgter Kategorisierung aufgrund ungerader Blöcke zeigte ähnliche Ergebnisse. Für die Nahtonutzer ($n = 9$) ergaben sich auch hier (marginal) signifikante Einflüsse der beiden Hauptfaktoren Visuell, $F(1, 8) = 38.43$, $p < .001$, und Auditorisch, $F(1.08, 8.64) = 4.04$, $p = .075$. Durch den erneut disordinalen Interaktionseffekt zwischen Visuell und Auditorisch, $F(2, 16) = 59.79$, $p < .001$, werden die Haupteffekte nicht in die Interpretation einbezogen (siehe Abbildung 8 oben rechts). Die weiteren Effekte wurden nicht signifikant: Dimension, $F(1, 8) = 0.02$, $p = .888$; Dimension \times Visuell, $F(1, 8) = 0.01$, $p = .936$; Dimension \times Auditorisch, $F(2, 16) = 0.94$, $p = .412$, sowie Dimension \times Visuell \times Auditorisch $F(1.19, 9.51) = 1.16$, $p = .321$. Die Post-hoc-Betrachtung des Interaktionseffekts offenbarte eine bessere Schätzleistung für die nahe Distanz mit gleichzeitig präsentem Ton (116 ms für auditorisch passend, 118 ms für nicht passend) gegenüber der unimodalen Stimulation (176 ms), $p \leq .012$. Gleichzeitig war die Schätzung für die ferne Distanz schlechter bei gleichzeitiger Tonpräsentation (527 ms für auditorisch passend, 521 ms für nicht passend) gegenüber der unimodalen Bedingung (395 ms), $p \leq .012$. Die Tonbedingungen unterschieden sich für die nahe und ferne Schätzung jeweils nicht voneinander, $p = 1.00$.

Für die Analyse der Standardabweichung der Schätzleistung wurden im ungeraden Analysedatensatz nach gerader Kategorisierung die Haupteffekte Visuell, $F(1, 10) = 62.01$, $p < .001$, und Auditorisch, $F(2, 20) = 3.80$, $p = .040$, signifikant. Alle weiteren Effekte erwiesen sich als nicht signifikant: Dimension, $F(1, 10) = 0.02$, $p = .883$; Dimension \times Visuell, $F(1, 10) = 0.19$, $p = .677$; Dimension \times Auditorisch, $F(1.25, 12.51) = 2.50$, $p = .135$; Visuell \times Auditorisch, $F(2, 20) = 0.93$, $p = .410$, und Dimension \times Visuell \times Auditorisch $F(2, 20) = 0.72$, $p = .499$. Die Leistung der Probanden streute weniger bei naher (82 ms) als bei ferner (168 ms) Schätzung. Auditorisch ergab sich post-hoc nur ein marginal signifikanter Unterschied zwischen der unimodal größeren Streuung (137 ms) und der Schätzung bei auditorisch nicht passendem Ton (117 ms), $p = .086$. Der Unterschied zwischen unimodal und auditorisch passendem Ton (120 ms) wurde nicht signifikant, $p = .343$, ebenso der Unterschied zwischen beiden Tonarten, $p = 1.00$. Die Streuung der Reaktionszeit-Abweichung für die Nahtonutzer in Abhängigkeit von der Schätzdistanz und der gleichzeitigen auditorischen Stimulation ist in Abbildung 8 unten links gezeigt.

Für das Vorgehen mit entgegengesetzten Blockzuweisungen (ungerade Kategorisierung/gerade Analyse) ergaben sich ähnliche Effekte. Wieder waren die beiden Haupteffekte Visuell, $F(1, 8) = 68.90$, $p < .001$, und Auditorisch, $F(2, 16) = 6.29$, $p = .010$, signifikant. Der Effekt Dimension, $F(1, 8) = 0.01$, $p = .989$, sowie die Interaktionen wurden nicht signifikant: Dimension \times Visuell, $F(1, 8) = 0.02$, $p = .891$; Dimension \times Auditorisch, $F(2, 16) = 1.82$, $p = .194$; Visuell \times Auditorisch, $F(1.21, 9.69) = 0.83$, $p = .407$, und Dimension \times Visuell \times Auditorisch $F(2, 16) = 1.98$, $p = .170$. Die Streuung wies für die nahe Schätzung (87 ms) eine kleinere Ausprägung auf als für die ferne Distanz (181 ms). Für den auditorischen Effekt zeigte sich post-hoc ein marginal signifikanter Unterschied zwischen unimodaler Schätzung (151 ms) und auditorisch nicht passendem (125 ms), $p = .077$. Von auditorisch passenden Tönen (127 ms) war sie nicht verschieden, $p = .110$. Ebenso unterschieden sich die beiden Tonbedingungen nicht

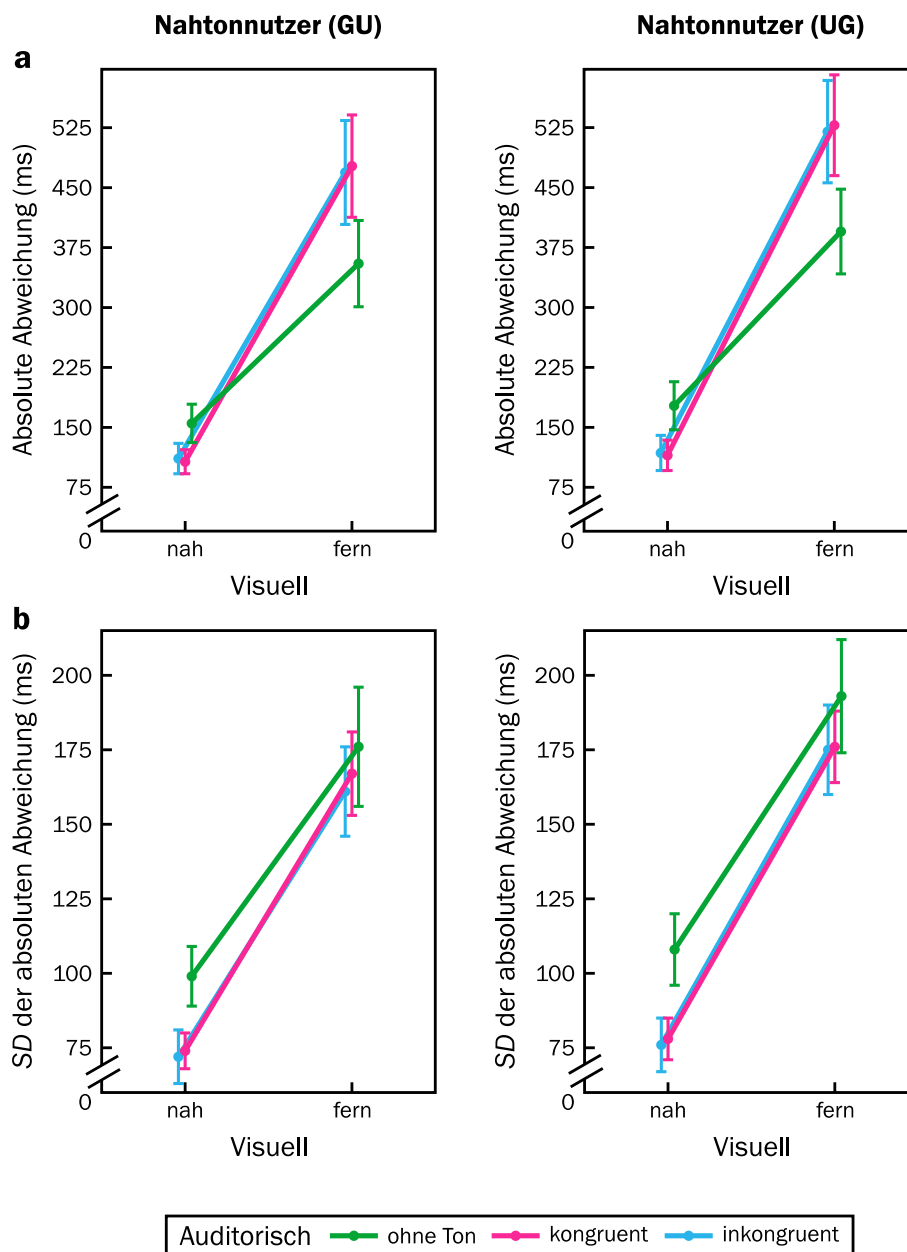


Abbildung 8. Reaktionszeit-Abweichung (oben, a) sowie die Streuung (SD) der Reaktionszeit-Abweichung (unten, b) in Abhängigkeit von der Schätzdistanz und der auditorischen Stimulation für die Nahtonutzer. Diese profitierten von zusätzlichen Tönen für die nahe Schätzdistanz, bei der fernen Distanz wirkten sich die Töne hinderlich aus (Details zur Nutzergruppeneinteilung siehe Text). Dargestellt sind die Ergebnisse auf Basis einer Split-Half-Analyse: Es wurden entweder gerade Blöcke zum Kategorisieren und ungerade Blöcke zum Analysieren genutzt (linke Seite, GU) oder ungerade Blöcke zum Kategorisieren und gerade Blöcke zum Analysieren (rechte Seite, UG). Die Fehlerbalken kennzeichnen den Standardfehler des Mittelwerts. Zur Unterscheidung der Fehlerbalken sind die Datenpunkte leicht versetzt dargestellt. Abbildung Teil (a) adaptiert nach: Dittrich und Noesselt (2018), S. 7.

voneinander, $p = 1.00$. Die Streuungen für diesen Analyseschritt in Abhängigkeit der visuellen und auditorischen Stimulation sind in Abbildung 8 unten rechts dargestellt.

Analyse der Reaktionszeit-Abweichungen der Ferntonnutzer Nach Kategorisierung der Probanden auf Grundlage der geraden Blöcke ergab sich für die Ferntonnutzer ($n = 10$) eine signifikante Interaktion für Schätzdistanz und auditorische Stimulation, $F(1.20, 10.82) = 40.46$, $p < .001$ (siehe Abbildung 9 oben links). Alle Haupteffekte wurden nicht signifikant: Dimension, $F(1, 9) = 0.89$, $p = .370$; Visuell, $F(1, 9) = 0.02$, $p = .897$; Auditorisch, $F(1.19, 10.73) = 0.47$, $p = .540$, ebenso wie die weiteren Interaktionen: Dimension \times Visuell, $F(1, 9) = 0.01$, $p = .953$; Dimension \times Auditorisch, $F(1.14, 10.25) = 0.19$, $p = .705$, und Dimension \times Visuell \times Auditorisch $F(2, 18) = 1.44$, $p = .264$. Post-hoc ergab sich für die Interaktion ein Unterschied der unimodalen Bedingung von beiden Tonarten jeweils für die nahe und ferne Schätzdistanz, $p \leq .012$. Die passenden und nicht passenden Töne unterschieden sich für beide Distanzen jeweils nicht voneinander, $p = 1.00$. Für die nahe Schätzung waren die Probanden ohne Ton besser (217 ms) als für passende (324 ms) und nicht passende Töne (326 ms). Bei der fernen Schätzdistanz verhielt es sich umgekehrt. Hier waren die Probanden unimodal (359 ms) schlechter als für beide Tonbedingungen (passend: 273 ms, nicht passend: 263 ms).

Das umgekehrte Vorgehen (ungerade Blöcke = Kategorisierung, gerade Blöcke = Analyse, $n = 11$) ergab ebenfalls einen signifikanten Interaktionseffekt zwischen Distanz und Tonpräsentation, $F(1.19, 11.88) = 29.43$, $p < .001$ (siehe Abbildung 9 oben rechts). Ebenso wurden erneut die weiteren Effekte nicht signifikant: Dimension, $F(1, 10) = 0.56$, $p = .472$; Visuell, $F(1, 10) = 0.91$, $p = .363$; Auditorisch, $F(1.26, 12.60) = 1.24$, $p = .299$; Dimension \times Visuell, $F(1, 10) = 0.30$, $p = .598$; Dimension \times Auditorisch, $F(1.34, 13.42) = 0.73$, $p = .448$, und Dimension \times Visuell \times Auditorisch $F(1.34, 13.43) = 0.18$, $p = .750$. Post-hoc waren die Probanden für die nahe Distanz unimodal (235 ms) genauer als in beiden Tonbedingungen (passend: 334 ms, nicht passend: 338 ms), $p \leq .012$. Für die ferne Schätzung waren sie für die auditorischen Bedingungen (passend: 235 ms, nicht passend: 232 ms) besser als in der nur visuellen Bedingung (307 ms), $p \leq .012$. In beiden Fällen unterschieden sich die Tonbedingungen jeweils nicht voneinander, $p = 1.00$.

Nach Kategorisierung auf Basis gerader Blöcke brachte eine ANOVA der ungeraden Blöcke für die Standardabweichungen der Reaktionszeit-Abweichungen signifikante Effekte für Dimension, $F(1, 9) = 8.07$, $p = .019$; Visuell, $F(1, 9) = 33.45$, $p < .001$, und Auditorisch, $F(2, 18) = 10.14$, $p = .001$ sowie eine signifikante Interaktion zwischen Visuell und Auditorisch hervor, $F(2, 18) = 21.03$, $p < .001$. Die weiteren Interaktionen wurden nicht signifikant: Dimension \times Visuell, $F(1, 9) = 1.20$, $p = .301$; Dimension \times Auditorisch, $F(1.11, 9.99) = 0.26$, $p = .649$, und Dimension \times Visuell \times Auditorisch $F(2, 18) = 0.40$, $p = .680$. Für die ferne Schätzdistanz war die Variabilität der Schätzungen größer (156 ms) als für die nahe Distanz (98 ms), ebenso während echt-3D-Stimulation (133 ms) verglichen zur pseudo-3D-Variante (121 ms). Auf die Betrachtung des auditorischen Haupteffekts wird aufgrund der Interaktion zwischen diesem und dem visuellen Faktor verzichtet (siehe Abbildung 9 unten links). Die Interaktion zwischen visueller und auditorischer Präsentation zeigte sich in keinen Variabilitätsunterschieden in

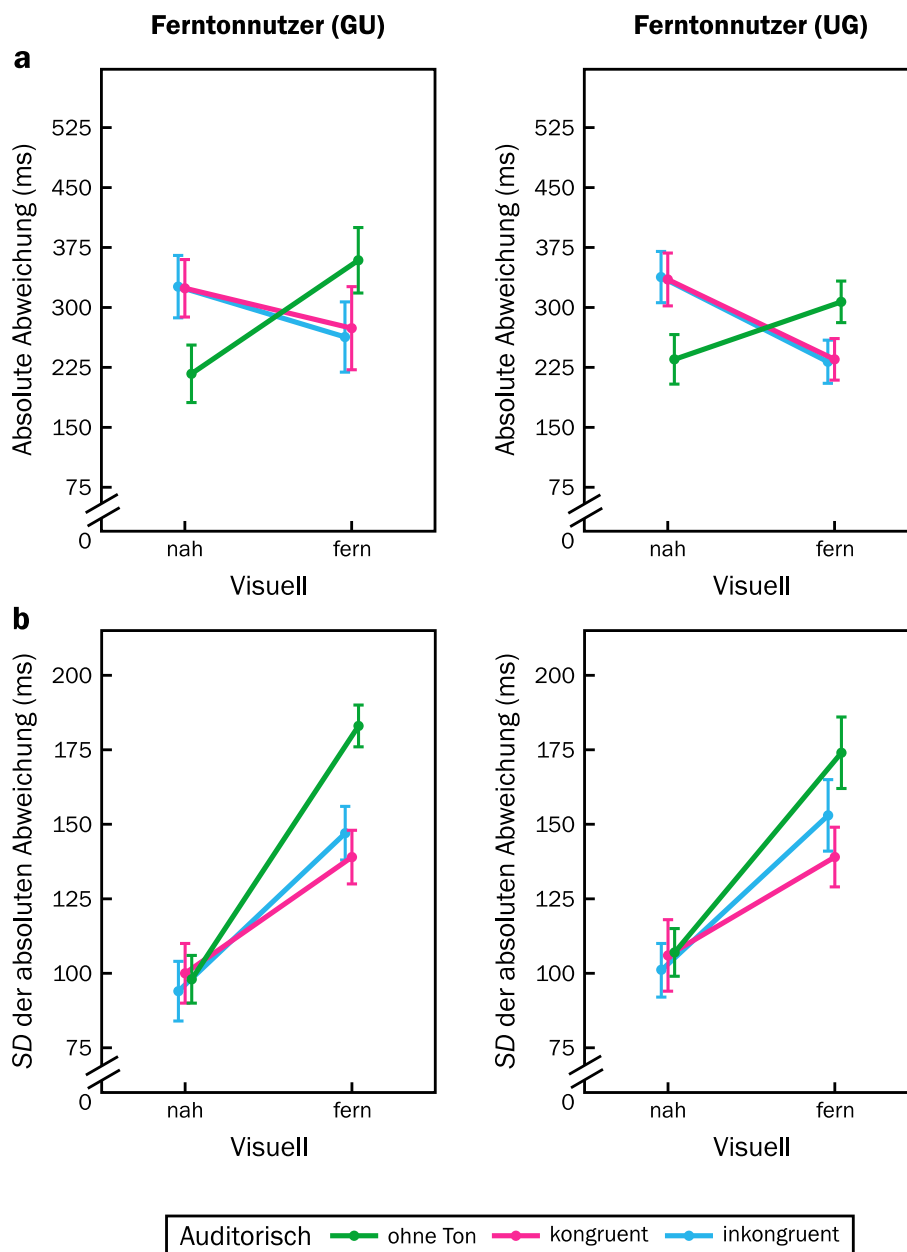


Abbildung 9. Reaktionszeit-Abweichung (oben, a) sowie die Streuung (SD) der Reaktionszeit-Abweichung (unten, b) in Abhängigkeit von der Schätzdistanz und der auditorischen Stimulation für die Ferntonnutzer. Diese profitierten von zusätzlichen Tönen für die ferne Schätzdistanz, bei der nahen Distanz wirkten sich die Töne hinderlich aus (Details zur Nutzergruppeneinteilung siehe Text). Dargestellt sind die Ergebnisse auf Basis einer Split-Half-Analyse: Es wurden entweder gerade Blöcke zum Kategorisieren und ungerade Blöcke zum Analysieren genutzt (linke Seite, GU) oder ungerade Blöcke zum Kategorisieren und gerade Blöcke zum Analysieren (rechte Seite, UG). Die Fehlerbalken kennzeichnen den Standardfehler des Mittelwerts. Zur Unterscheidung der Fehlerbalken sind die Datenpunkte leicht versetzt dargestellt. Abbildung Teil (a) adaptiert nach: Dittrich und Noesselt (2018), S. 7.

den Tonbedingungen für die nahe Schätzdistanz, $p = 1.00$ (unimodal: 100 ms, auditorisch passend: 103 ms, auditorisch nicht passend: 97 ms). Für die ferne Schätzdistanz waren die Standardabweichungen beider Töne ebenso nicht verschieden, $p = .726$, jedoch unterschieden sich beide Töne von der unimodalen Präsentation. Die Streuung war ohne Ton (186 ms) größer als für die Bedingungen mit auditorisch passendem Ton (144 ms), $p < .001$ und auditorisch nicht passendem Ton (154 ms), $p = .012$.

Bei Auswertung gerader Blöcke nach ungerader Kategorisierung ergab sich ein ähnliches Bild. Die Haupteffekte Visuell, $F(1, 10) = 51.99$, $p < .001$, und Auditorisch, $F(2, 20) = 4.98$, $p = .018$ wurden signifikant, ebenso die Interaktion beider Faktoren, $F(2, 20) = 8.04$, $p = .003$. Für die restlichen Effekte wurden keine signifikanten Ergebnisse gefunden: Dimension, $F(1, 10) = 1.20$, $p = .299$; Dimension \times Visuell, $F(1, 10) = 0.27$, $p = .617$; Dimension \times Auditorisch, $F(2, 20) = 1.18$, $p = .329$, und Dimension \times Visuell \times Auditorisch $F(2, 20) = 2.10$, $p = .149$. Die Streuung fiel für die nahe Distanz (105 ms) geringer aus als für die ferne Distanz (155 ms). Die Betrachtung des auditorischen Haupteffekts entfällt aufgrund der Interaktion mit dem visuellen Faktor (siehe Abbildung 9 unten rechts). Die Post-hoc-Betrachtung des Interaktionseffekts zwischen Visuell und Auditorisch zeichnete sich durch keine Unterschiede zwischen den Tonstufen für die nahe Distanz aus (unimodal: 111 ms, auditorisch passend: 112 ms, nicht passend: 108 ms), $p = 1.00$. Für die ferne Distanz hingegen ergaben sich Unterschiede. Die Streuung war für die unimodale Präsentation (180 ms) größer als für die auditorisch passende (147 ms), $p = .006$. Verglichen zum nicht passenden Ton (161 ms) ergab sich nach Korrektur kein Unterschied ($p = .276$, unkorrigiert: $p = .046$). Die beiden Töne unterschieden sich ebenfalls nicht, $p = .624$.

5.2.4 Diskussion

Bei der Bearbeitung einer anspruchsvolleren zeitlichen Extrapolationsaufgabe deuteten die Ergebnisse zunächst auf einen hinderlichen Einfluss der Tonpräsentation hin. Die Probanden waren laut Overall-Analyse in ihrer Schätzung, wann die Kugel die Stange erreichen würde, schlechter in den Tonbedingungen. Eine detaillierte Post-hoc-Betrachtung der Daten verdeutlichte jedoch, dass die meisten Probanden nur für eine der beiden Schätzdistanzen von zusätzlichen Tönen profitierten.

Bei der Einteilung der Probanden in die verschiedenen Nutzergruppen fällt auf, dass für diese Perspektive auf die Daten ein Großteil der Probanden bei der Bewegungsvorhersage auf Töne zurückgriff. Je nach Kategorisierungsdatensatz (gerade vs. ungerade) nutzten 26 bzw. 24 von 32 Probanden die Töne. Jedoch konnte mit der crossvalidierten Post-hoc-Auswertung gezeigt werden, dass stabile interindividuelle Unterschiede zwischen den Probanden vorliegen. Nur ein geringer Anteil (5 bzw. 4 Probanden) profitierte für beide Schätzdistanzen von den Tönen (Beidtonnutzer). Die übrigen konnten lediglich für jeweils eine der beiden Distanzen die Töne zu ihrem Vorteil nutzen (Nah- und Ferntonnutzer). Die Probanden sind in der Schätzung, wann die Kugel die Stange erreichen wird, wesentlich genauer bei gleichzeitiger Tonpräsentation für ihre präferierte Distanz (Abweichung um 47 bis 60 ms bzw. 75 bis 96 ms besser als ohne Töne). Auch die Streuung ist für Nah- und Ferntonnutzer bei dieser Distanz geringer und

somit ist ebenfalls die Präzision der Schätzung verbessert. Bei der jeweils anderen Distanz verhält es sich umgekehrt. Die Nahtonnutzer sind wesentlich schlechter (Abweichung 122 bis 132 ms größer) in Tonbedingungen für die ferne Distanz, ebenso erzielen die Ferntonnutzer eine im Mittel um 106 ms schlechtere Schätzung im nahen Bereich.

Dass viele Probanden die Töne nur für eine bestimmte Schätzdistanz zu ihrem Vorteil nutzten und im Gegenzug für eine andere Distanz gleichzeitig von diesen behindert wurden, spricht für einen Einfluss der Zeit auf die audiovisuelle Bewegungsvorhersage. Die Genauigkeit und Präzision der Schätzung kann nicht konstant über einen Zeitraum von mehreren (Milli-)Sekunden gewährleistet werden. Die Probanden unterscheiden sich im Zeitpunkt, wann auditorische Informationen genutzt werden können. Einige nutzen diese sehr früh nach der Verdeckung, sind aber im Anschluss nicht in der Lage, sie auch nach einem längeren Zeitraum gewinnbringend zu verwerten. Andere können zwar zu einem relativ späten Zeitpunkt nach der Verdeckung die Töne zu ihrem Vorteil nutzen, können diesen dafür jedoch nicht sehr früh nach der Verdeckung abrufen. Da es zusätzlich einige wenige Probanden gab, die für beide Distanzen die Töne nutzen konnten, sprechen diese Ergebnisse dafür, dass entweder verschiedene Strategien von den Probanden genutzt wurden oder unterschiedliche Lernvergangenheiten eine Rolle spielen. So könnten die Probanden eine Vorliebe für eine Schätzdistanz entwickelt und sich auf diese in der Aufgabenbearbeitung fokussiert haben. Oder sie konnten nur für eine Distanz ein Zeitgefühl entwickeln, das dem tatsächlichen Bahnverlauf nahekommt. Dies ließe sich durch das Hinzufügen von Feedback genauer untersuchen. Man könnte so überprüfen, ob die Probanden in der Lage sind, über die Zeit hinweg für beide Distanzen ihr Schätzverhalten zu verbessern und den Ton auch für die zweite Distanz zu ihrem Vorteil zu nutzen. Auf die Aspekte dieser interindividuellen Unterschiede im vorliegenden Experiment wird in der Allgemeinen Diskussion genauer eingegangen (siehe Abschnitt 8.2).

Außerdem zeigten sich im Gegensatz zu Experiment 1 keine Dimensionseffekte. Die zusätzlichen Disparitätsinformationen im echt-3D-Teil übten keinen Einfluss aus. Dies spricht dafür, dass für eine Interaktion mit der Präsentation im 3D-Raum (das heißt in den Experimenten hier mit Material, das Disparitätshinweise enthält) das für die Aufgabenstellung relevante Objekt zum Zeitpunkt der Antwortgabe sichtbar sein muss. Dies war in diesem Experiment nicht der Fall, da die Kugel nach der Brücke nicht wieder auftauchte und die Probanden somit ohne visuellen Bewegungs-Input angaben, wann die Stange erreichen werden würde. Die Rolle von 3D-Informationen für die Bewegungsvorhersage wird in Abschnitt 8.6 der Allgemeinen Diskussion genauer beleuchtet. Ob eine zusätzliche Relevanz der räumlichen Komponente bei der Aufgabenstellung den Einfluss von 3D-Material verstärkt, wird in den drei durchgeführten audiovisuellen Diskriminationsexperimenten mithilfe einer räumlichen Aufgabe untersucht (siehe Experimente 3 bis 5).

5.3 Kontrollexperiment: Bestimmbarkeit der Tonrichtung

In Experiment 1 und 2 zeigte sich während einer Detektions- und Extrapolationsaufgabe ein Vorteil für Durchgänge mit zusätzlichen Tönen. Allerdings spielte die Richtung des Tons keine Rolle. Es war unerheblich, ob sich dieser in die gleiche oder die entgegengesetzte Richtung wie der visuelle Stimulus bewegte. Beide Tonarten (passend/nicht passend) erleichterten das Verhalten der Probanden in ähnlicher Art und Weise.

Dieses Kontrollexperiment wurde durchgeführt, um auszuschließen, dass die Tonrichtung keinen Einfluss ausübt, weil die Probanden die Tonbewegungsrichtung nicht bestimmen können. Dafür sollten die Probanden für die vier in Experiment 1 und 2 verwendeten Töne (entsprechend den vier visuellen Bewegungsbahnen) sowie vier weitere in einer früheren Pilotversion dieser Experimentalreihe genutzten Töne die Bewegungsrichtung angeben.

5.3.1 Methoden

Probanden Für das Kontrollexperiment stellten sich 10 Probanden (2 männlich, 8 weiblich) zur Verfügung. Alle waren (korrigiert) normalsichtig. Sie waren im Durchschnitt 21.60 Jahre ($SD = 2.55$) alt, ein Teilnehmer war links-, die restlichen waren rechtshändig. Die Probanden berichteten keine neurologischen oder psychiatrischen Erkrankungen und kein eingeschränktes Hörvermögen. Sie nahmen freiwillig an der Untersuchung teil und erklärten ihre Teilnahme zuvor schriftlich. Die Untersuchung wurde in Einklang mit den Richtlinien des örtlichen Ethikkomitees durchgeführt. Die Teilnehmer wurden monetär oder mit Versuchspersonenstunden im Rahmen ihres Psychologie-Studiums vergütet.

Experimentelles Set-Up und Durchführung Insgesamt wurden den Probanden acht verschiedene Töne präsentiert, vier davon entsprachen den in Experiment 1 und 2 verwendeten Tönen. Die weiteren vier Töne unterschieden sich im Ausmaß ihrer Lautstärkenmodulation im Raum. Spannt man ein imaginäres Rechteck zwischen den vier Lautsprechern auf (siehe experimentelles Set-Up in Abbildung 3), kann man die Stärke, mit der sich der Ton in der Tiefe auf den Hörer zubewegt, variieren, indem man die Höhe des Rechtecks (entspricht der Seitenlänge orthogonal zum Hörer) verändert. Diese Höhe des Rechtecks unterschied sich zwischen den beiden Tongruppen um den Faktor 20. Dadurch entstand eine Tongruppe aus vier Tönen, bei denen die Bewegung in der Tiefendimension (vorn/hinten) im Vordergrund stand und eine Tongruppe, bei der die Bewegung in der Horizontalen (links/rechts) im Vordergrund stand. Die letztere Tongruppe entsprach den in Experiment 1 und 2 verwendeten Tönen.

Den Probanden wurden zunächst in Experimententeil 1 dieses Kontrollexperiments alle acht Töne in zufälliger Reihenfolge präsentiert. Im Anschluss wurden in Teil 2 und 3 nur die vier Töne jeweils einer Tongruppe präsentiert. Die Reihenfolge der Einzelpräsentation der beiden Tongruppen war zwischen den Probanden ausbalanciert. Während der Tonpräsentation waren die Probanden angehalten, die Bildschirmmitte zu fixieren und die Augen nicht zu schließen. Im Anschluss mussten die Probanden mit einer USB-Maus angeben, ob sich der zuvor gehörte Ton von links nach rechts oder rechts nach links bewegt hat (siehe Abbildung 10 a). Nach dieser Frage schätzten sie auf einer Ratingskala von 1 (überhaupt nicht sicher) bis 100 (vollkommen

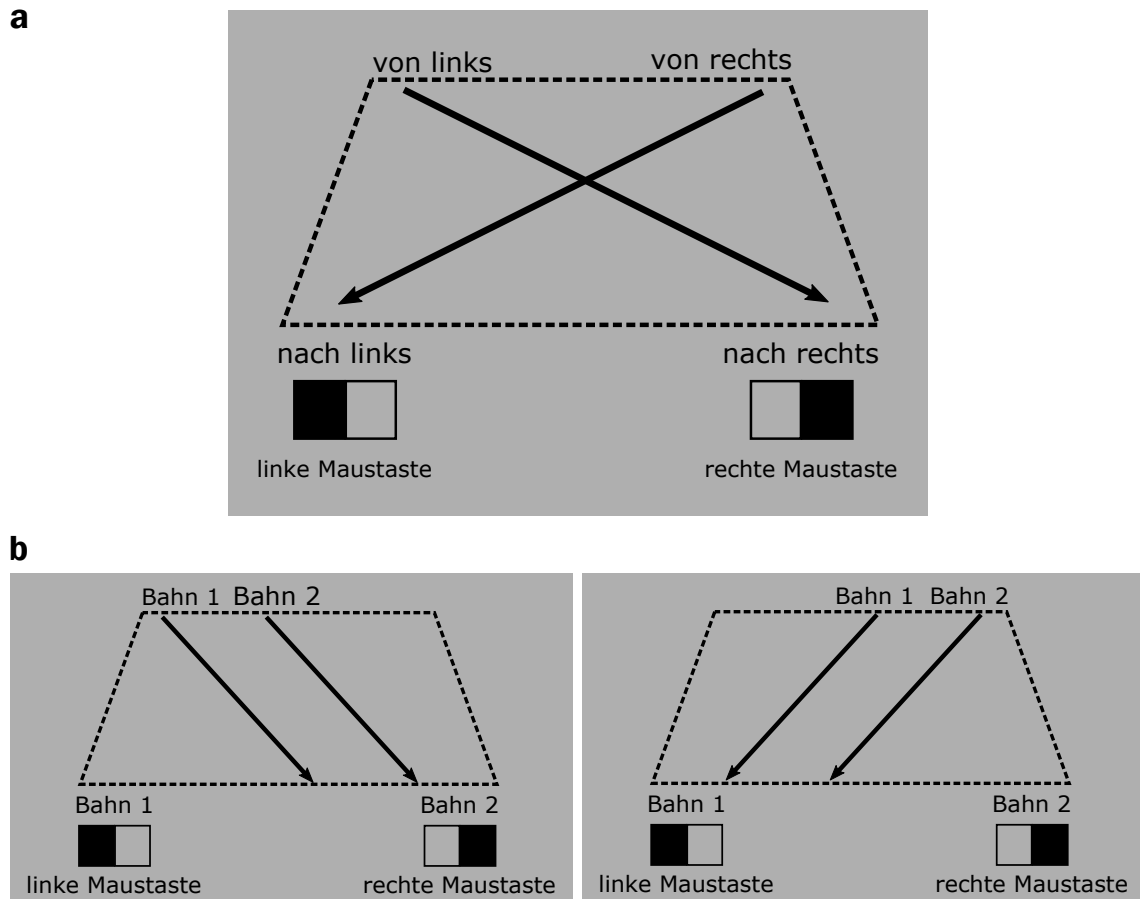


Abbildung 10. Abfragedisplays für die Richtung der Tonbewegung im Kontrollexperiment. In allen Experimentalteilen erfolgte zunächst die Abfrage, in welche Richtung sich der Ton bewegte (a). Die Probanden wählten per Maustaste zwischen „von links nach rechts“ und „von rechts nach links“. Im 2. und 3. Experimentalteil entschieden die Probanden zusätzlich, ob sich der Ton eher auf der äußeren oder inneren Bahn bewegte (b). Entsprechend der Antwort im ersten Abfragedisplay (a) wurde im Anschluss das passende Abfragedisplay aus (b) für links bzw. rechts angezeigt. Nähere Informationen siehe Text.

sicher) ein, wie sicher sie sich bei ihrer Entscheidung waren. Im 2. und 3. Experimentalteil gaben die Teilnehmer zusätzlich zur Richtung des Tons an, auf welcher Bahn der Ton sich gemäß ihrer Richtungsentscheidung (links oder mittig links bzw. rechts oder mittig rechts) bewegt hat (siehe Abbildung 10 b). In beiden Abfragearten wurden den Probanden skizzenhaft auf dem Bildschirm die zur Verfügung stehenden Bewegungsrichtungen bzw. -bahnen inklusive den jeweils damit assoziierten Maustasten angezeigt.

Auswertung Die Datenanalyse erfolgte mit SPSS 23.0 (IBM, USA). Es wurden die deskriptiven Werte der korrekten Antworten sowie der Konfidenz-Ratings betrachtet. Zudem wurde eine 2x4-ANOVA für die beiden Tongruppen mit den jeweils vier Tönen durchgeführt, um auf etwaige Unterschiede zwischen den Tongruppen und Tönen zu prüfen. Bei verletzter Sphärizitätsbedingung wurde gegebenenfalls nach Greenhouse-Geisser korrigiert.

5.3.2 Ergebnisse

Die mittleren korrekten Antworten für die Angabe der Bewegungsrichtung aller Bedingungen schwankten zwischen 98.0% und 99.7%. Die Werte für die Konfidenz-Ratings lagen zwischen 96.1 und 98.0. Die Haupteffekte Tongruppe, $F(1, 9) = 0.30$, $p = .599$, und Ton, $F(1.46, 13.18) = 0.67$, $p = .485$, wurden ebenso wie der Interaktionseffekt Ton \times Tongruppe, $F(1.89, 16.97) = 2.15$, $p = .149$, für die korrekten Antworten nicht signifikant. Auch für die Konfidenz-Ratings ergaben sich keine signifikanten Effekte: Tongruppe, $F(1, 9) = 3.49$, $p = .095$; Ton, $F(3, 27) = 1.74$, $p = .183$, und Ton \times Tongruppe, $F(3, 27) = 0.93$, $p = .439$. Darüber hinaus lag die mittlere Rate korrekter Antworten für die Angabe der Bewegungsbahn (äußere vs. innere) für die in Experiment 1 und 2 verwendete Tongruppe bei durchschnittlich 80.1% ($SD = 24.0\%$). Ein Proband erzielte für diese Abfrage eine extrem niedrige Accuracy von 13.9% und lag damit außerhalb des Intervalls von $\pm 2 SD$. Nach Ausschluss dieses Ausreißer-Probandens lag der mittlere Anteil korrekt bestimmter Bewegungsbahnen bei 87.5% ($SD = 6.3\%$).

5.3.3 Diskussion

Der Anteil korrekter Antworten für die in Experiment 1 und 2 verwendeten Töne ist mit annähernd 100% sehr hoch. Die Probanden haben demnach keine Schwierigkeiten, die Bewegungsrichtung der Töne zu bestimmen. Zudem sind sie sich bei einer Konfidenz-Angabe von beinahe 100 bei 100 möglichen „Punkten“ sehr sicher in ihrer Einschätzung, in welche Richtung sich der Ton bewegte. Außerdem ergaben sich keine Unterschiede bezüglich der Korrektheit und Sicherheit in den Antworten zwischen den vier in Experiment 1 und 2 verwendeten Tönen.

Die Ergebnisse des Kontroll-experiments sprechen somit dafür, dass die Teilnehmer in der Lage sind, die Bewegungsrichtung des Tons korrekt wahrzunehmen. Der nicht vorhandene Einfluss der Tonrichtung in den beiden ersten Experimenten kann nicht darauf zurückgeführt werden, dass die kongruenten und inkongruenten Töne von den Versuchspersonen nicht unterschieden werden können. Dies unterstreicht die Vermutung, dass die Bewegungsrichtung der Töne, und in diesem Sinne die räumliche Information, für die Detektion bzw. Extrapolation der Kugel nach der Verdeckung irrelevant ist und nur die zeitliche Information, die von den Tönen unabhängig der Bewegungsrichtung transportiert werden, eine Rolle zu spielen scheinen. Die Lautstärkenveränderung durch die Verdeckung stellt einen ausreichenden Hinweis für die Bearbeitung der Aufgabenstellung dar und die Probanden nutzen nicht die zusätzlichen Informationen über die Bewegungsrichtung, die vom Ton transportiert werden. Auf weitere mögliche Gründe für den fehlenden Kongruenzeffekt in Experiment 1 und 2 wird in Abschnitt 8.2 der Allgemeinen Diskussion genauer eingegangen.

6 Variation der visuellen Stimulation (Diskrimination)

6.1 Experiment 3: Diskrimination des Bewegungsverlaufs

Im Gegensatz zum ersten Experimententeil dieser Arbeit, wurde in diesem sowie den folgenden Experimenten die räumliche Komponente untersucht und eine Diskrimination über die Bahnbe-
wegung der Kugel durchgeführt. Für räumliche Aspekte sollten die visuellen Informationen bei
der Aufgabenbearbeitung entscheidender sein (Welch & Warren, 1980; Welch et al., 1986; Wit-
ten & Knudsen, 2005). Daher sollte in diesem Experiment untersucht werden, in welcher Form
sich zusätzliche Töne auf die Leistung in einer räumlichen Diskriminationsaufgabe auswirken.
Des Weiteren ergab sich in Experiment 1 und 2 kein Einfluss der Bewegungsrichtung des
Tons. Dass dies nicht darauf zurückzuführen ist, dass die Probanden diese nicht differenzieren
können, zeigte das durchgeführte Kontrollexperiment. Da in diesem Experiment räumliche
Aspekte in der Aufgabenstellung forciert wurden, könnte die Tonrichtung einen größeren Ein-
fluss ausüben und die Probanden entsprechend vor allem von multisensorisch kongruenter
Stimulation profitieren (Meyer et al., 2005).

6.1.1 Methoden

Am 3. Experiment nahmen insgesamt 41 Probanden teil, ein Proband wurde von der Analyse
ausgeschlossen, da dieser in einem Experimententeil die beiden Antworttasten vertauscht
hatte. Er wurde direkt entsprechend der Ausbalancierung ersetzt. Von den 40 verbleibenden
Probanden (je 20 männlich/weiblich) waren drei links-, einer beid- und 36 rechtshändig. Sie
hatten ein Durchschnittsalter von 23.55 Jahren ($SD = 3.43$).

Als Aufgabenstellung wurde in diesem Experiment die Diskrimination verwendet: Die
Probanden mussten nach dem Wiedererscheinen so schnell und korrekt wie möglich angeben,
ob sich die Kugel auf einer visuell passenden oder nicht passenden Bahn weiterbewegte.
Dafür betätigten sie die linke bzw. rechte Pfeiltaste. Die Tastenzuweisung war zwischen den
Probanden gemeinsam mit der Experimentalteilerihenfolge und dem Geschlecht ausbalanciert.
Das Stimulusmaterial war identisch mit dem aus Experiment 1, Beispielbilder für die visuellen
Stimuli sind daher in Abbildung 4 zu sehen.

6.1.2 Ergebnisse

Von der Datenauswertung wurden insgesamt 7.49 % der Durchgänge für die Reaktionszeit-
Analyse und 5.15 % für die Analyse der korrekten Antworten ausgeschlossen. Der Anteil
entfernter Durchgänge mit Flipungenaugigkeiten lag bei 0.13 % (für Reaktionszeiten und korrekte
Antworten).

Reaktionszeit-Analyse Es ergaben sich zwei signifikante Haupteffekte für die Faktoren Vi-
suell, $F(1, 39) = 6.39$, $p = .016$, und Auditorisch, $F(1.68, 65.58) = 4.92$, $p = .014$. Die Probanden
reagierten schneller bei visuell kongruenten Pfaden nach der Verdeckung (480 ms) als bei
nicht kongruenten Bahnen (495 ms). Sie waren zudem in Bedingungen ohne Ton (485 ms)

schneller als bei auditorisch passender (490 ms), $p = .059$, und nicht passender Stimulation (489 ms), $p = .047$. Die beiden Tonstufen unterschieden sich nicht voneinander, $p = 1.0$. Die Reaktionszeiten in Abhängigkeit von Bild- und Tonpräsentation sind in Abbildung 11 dargestellt. Eine signifikante Dreifachinteraktion aller Faktoren, $F(2, 78) = 3.27$, $p = .043$, schränkt jedoch die Interpretation der Haupteffekte ein. Post-hoc-ANOVAs für pseudo- und echt-3D-Stimulation getrennt zeigten keine signifikanten Einflüsse für die pseudo-3D-Bedingungen: Visuell, $F(1, 39) = 1.83$, $p = .184$; Auditorisch, $F(2, 78) = 1.77$, $p = .176$, und Visuell \times Auditorisch, $F(2, 78) = 0.08$, $p = .926$. Bei echt-3D-Präsentation wurden hingegen alle Effekte signifikant: Visuell, $F(1, 39) = 13.56$, $p = .001$; Auditorisch, $F(1.65, 64.25) = 4.01$, $p = .030$, und Visuell \times Auditorisch, $F(1.64, 63.91) = 3.84$, $p = .034$. Die Interaktion zwischen visueller und auditorischer Stimulation zeigte sich in einem größeren visuellen Erleichterungseffekt für auditorisch passende Bedingungen (28 ms) als für unimodale (13 ms), $p = .003$. Der Unterschied zwischen auditorisch passend und nicht passend (14 ms) zeigte unkorrigiert ebenfalls eine Tendenz zum Unterschied $p = .052$ (korrigiert: $p = .156$). Die Erleichterung ohne Ton und bei nicht passendem Ton, unterschieden sich nicht, $p = 1.00$. Alle weiteren Effekte der übergeordneten ANOVA erwiesen sich als nicht signifikant: Dimension, $F(1, 39) = 0.31$, $p = .580$; Dimension \times Visuell, $F(1, 39) = 1.69$, $p = .201$; Dimension \times Auditorisch, $F(2, 78) = 1.38$, $p = .258$; Visuell \times Auditorisch, $F(1.52, 59.44) = 1.60$, $p = .213$.

Eine genauere Post-hoc-Betrachtung der Reaktionszeit-Ergebnisse für alle drei in dieser Arbeit durchgeführten audiovisuellen Diskriminationsexperimente erfolgt in Abschnitt 6.4.

Analyse der korrekten Antworten Um zu überprüfen, ob ein Speed-Accuracy-Trade-Off vorliegt, wurden zusätzlich die korrekten Antworten mittels ANOVA analysiert. Es wurden hierbei keine der Effekte signifikant: Dimension, $F(1, 39) = 1.33$, $p = .255$; Visuell, $F(1, 39) = 1.73$, $p = .196$; Auditorisch, $F(2, 78) = 1.70$, $p = .190$; Dimension \times Visuell, $F(1, 39) = 0.31$, $p = .583$; Dimension \times Auditorisch, $F(2, 78) = 0.63$, $p = .536$; Visuell \times Auditorisch, $F(2, 78) = 1.88$, $p = .160$, und Dimension \times Visuell \times Auditorisch, $F(2, 78) = 0.34$, $p = .715$.

Die Analyse brachte keine signifikanten Einflüsse der Experimentalfaktoren auf die korrekten Antworten hervor. Zudem schwankten die Genauigkeiten für die Experimentalbedingungen nur wenig auf einem sehr hohen Niveau zwischen 96.9% und 98.5%. Daher kann davon ausgegangen werden, dass kein Speed-Accuracy-Trade-Off für die Daten vorliegt.

6.1.3 Diskussion

Für die Diskriminationsaufgabe ergab sich zunächst wie in Experiment 1 ein visueller Erleichterungseffekt. Die Probanden diskriminierten den visuellen Bahnverlauf bei passenden Bahnen schneller als bei nicht passenden Bahnen (Kawachi & Gyoba, 2006). Anders als bei der Detektion übten die Töne jedoch keinen positiven Einfluss aus. In unimodalen Bedingungen reagierten die Probanden schneller, die Töne wirkten sich hier eher hinderlich auf die Leistung aus, wobei die Art des Tons unerheblich war. Dies könnte daran liegen, dass in räumlichen Aufgaben die visuelle Modalität dominiert (Welch & Warren, 1980). Dass sich die Leistung bei zusätzlicher Tonstimulation verschlechtert, könnte auf mögliche Interferenz- oder auch Ablenkungsprozesse

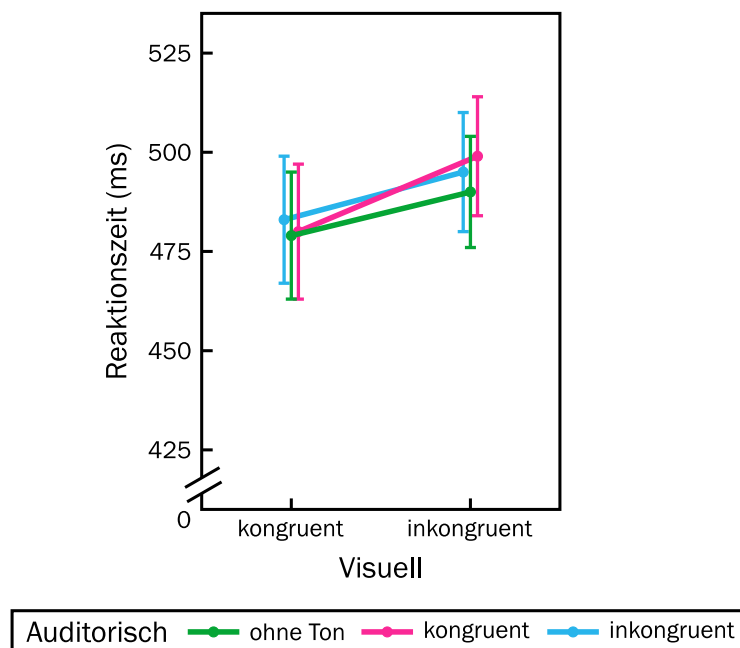


Abbildung 11. Reaktionszeit in Abhängigkeit vom visuellen Bahnverlauf nach der Verdeckung sowie von der auditorischen Stimulation. Die Fehlerbalken kennzeichnen den Standardfehler des Mittelwerts. Zur Unterscheidung der Fehlerbalken sind die Datenpunkte leicht versetzt dargestellt.

hindeuten (Banbury et al., 2001). Die Tonpräsentation zieht zu viel Aufmerksamkeit auf sich bzw. wird von den Probanden als störend empfunden, sodass sich dies in einer signifikanten Leistungseinbuße gegenüber unimodalen Durchgängen ohne Ton zeigt.

Allerdings scheinen die Einflüsse der visuellen und auditorischen Stimulation nur für die echt-3D-Präsentation zu gelten. Für die Analyse der pseudo-3D-Bedingungen wurden sie nicht signifikant. Dies spricht für einen größeren Einfluss der beiden Modalitäten unter realistischeren Bedingungen und ist im Einklang mit zahlreichen Studien, die einen Einfluss realitätsnaher Stimulation in Experimenten nahe legen (Conrad et al., 2013; Gaebler et al., 2014; Getzmann & Lewald, 2010; Harrison et al., 2015; Ogawa & Macaluso, 2013; Ogawa et al., 2013; Senkowski et al., 2007; Snow et al., 2011). Dies ist insofern von maßgeblicher Relevanz, da Studien mit vereinfachtem Labormaterial ggf. Effekte über- bzw. unterschätzen oder auch übersehen könnten.

Da die Datenanalyse für die drei durchgeführten audiovisuellen Experimente (3 bis 5) keine eindeutigen Effekte offenbarte, wurden diese in einer Post-hoc-Gesamtschau erneut im Detail betrachtet. Insbesondere eine Untersuchung auf interindividuelle Unterschiede in der Informationsnutzung (analog zu den gefundenen Effekten in Experiment 2) wurde durchgeführt. Die Ergebnisse dieser Post-hoc-Analysen sind in Abschnitt 6.4 zu finden.

6.2 Experiment 4: Erhöhung der visuellen Schwierigkeit

In diesem Experiment wurde erneut eine Diskriminationsaufgabe verwendet. Allerdings wurde die Sichtbarkeit des visuellen Stimulationsmaterials herabgesetzt, um zu prüfen, ob sich dadurch der Einfluss der Tonstimulation erhöht. Dafür wurde der visuelle Unterschied zwischen der Kugel und dem Boden der Box verringert. Durch eine Anpassung von Kugel- und Bodenmuster, hebt sich die Kugel weniger vom Hintergrund ab und ist schwieriger zu erkennen als dies in Experiment 3 der Fall war. Wenn die Reliabilität der einen sensorischen Modalität gesenkt wird, sollten die Probanden vermehrt auf die zweite zur Verfügung stehende Informationsquelle, in diesem Fall die Töne, zurückgreifen (Alais & Burr, 2004b; Bresciani & Ernst, 2007; Helbig & Ernst, 2007; Heron et al., 2004). Dies sollte sich darin zeigen, dass in diesem Experiment die Tonstimulation einen förderlichen Einfluss auf die Diskriminationsleistung ausübt.

6.2.1 Methoden

Insgesamt 40 Versuchspersonen (je 20 männlich/weiblich) nahmen an Experiment 4 teil. Diese waren im Durchschnitt 25.00 Jahre ($SD = 4.43$) alt. Es waren ein Proband links-, zwei beid- und 37 rechtshändig.

Der experimentelle Ablauf und die Stimulusgestaltung entsprachen weitgehend Experiment 3, jedoch mit folgenden Änderungen: Zur Erhöhung der visuellen Schwierigkeit wurde das Kugelmuster dem Boden der Box angepasst. Entsprechend war die Kugel statt mit farbigen Streifen in diesem Experiment ebenfalls durch dunkelgrau-weiße Karos gekennzeichnet. Das Kugelmuster wurde derart angepasst, dass beim Wiederscheinen nach der Brücke die Farbe der Kugelhälfte der Farbe des Schachfeldes auf dem Boden entsprach. Aus diesem Grund war das Kugelmuster zwischen den beiden linken und rechten Bewegungsbahnen gespiegelt, um die Schwierigkeit beim Wiedererkennen bei allen Bahnen gleichzuhalten. Zwei Beispielbilder mit der Start- und Endposition für die an den Untergrund angepasste Kugel sind in Abbildung 12 gezeigt.

Die Aufgabe der Probanden war auch hier wie in Experiment 3 die Unterscheidung, ob die Kugel nach dem Wiederscheinen sich auf einer visuell passenden oder nicht passenden Bahn weiterbewegte (räumliche Diskrimination). Dafür betätigten sie die linke oder rechte Pfeiltaste der Tastatur. Die Tastenzuweisung zu passend/nicht passend war erneut über die Reihenfolge der Experimentalteile und das Geschlecht der Probanden ausbalanciert.

6.2.2 Ergebnisse

Für die Analyse der Reaktionszeiten wurden 7.00 % der Durchgänge entfernt, für die der korrekten Antworten 5.14 %. Der Anteil von Trials mit Flipungenaugigkeiten lag für beide abhängigen Variablen insgesamt bei 0.14 %.

Reaktionszeit-Analyse Für die Reaktionszeiten wurden alle Haupteffekte und Interaktionen nicht signifikant: Dimension, $F(1, 39) = 0.08$, $p = .786$; Visuell, $F(1, 39) = 1.35$, $p = .253$; Auditorisch, $F(1.29, 50.14) = 2.18$, $p = .141$; Dimension \times Visuell, $F(1, 39) = 0.95$, $p = .336$;

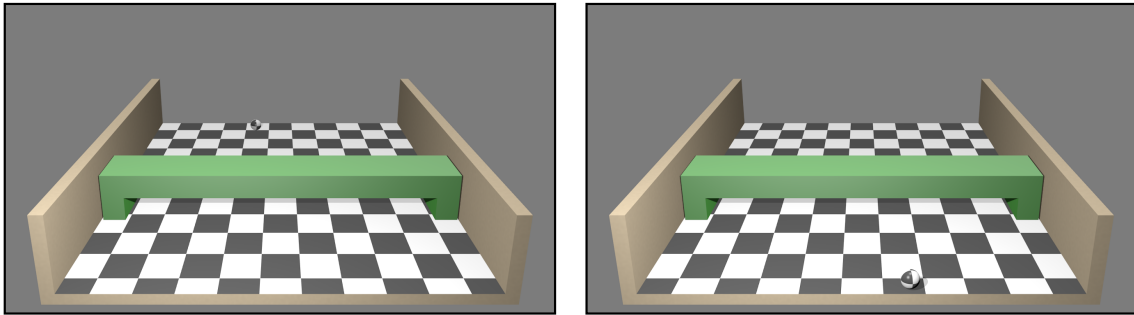


Abbildung 12. Beispielhaftes Bildmaterial aus Experiment 4. Links ist die Start-, rechts die Endposition der Kugel abgebildet. Die Kugel bewegte sich auf einer von vier möglichen Bahnen vom hinteren Rand der Box auf den Betrachter zu und wurde kurzzeitig durch die Brücke verdeckt. Sobald die Kugel auf der anderen Seite der Brücke wiedererschien, gaben die Probanden an, ob sich die Kugel auf einer passenden Bahn nach der Verdeckung weiterbewegte. Das Kugelmuster wurde an das Schachbrettmuster des Bodens angepasst, um die visuelle Schwierigkeit bei der Diskrimination zu erhöhen, da die Kugel so weniger vom Hintergrund unterscheidbar war.

Dimension \times Auditorisch, $F(2, 78) = 0.96$, $p = .386$; Visuell \times Auditorisch, $F(2, 78) = 0.09$, $p = .911$, sowie Dimension \times Visuell \times Auditorisch, $F(2, 78) = 0.99$, $p = .375$. Die Reaktionszeiten in Abhängigkeit der visuellen und auditorischen Stimulation sind in Abbildung 13 dargestellt.

Eine detaillierte Post-hoc-Betrachtung der Reaktionszeiten für Experiment 4 sowie die beiden weiteren audiovisuellen Diskriminationsexperimente findet sich in Abschnitt 6.4.

Analyse der korrekten Antworten Zusätzlich wurde ein Einfluss der Experimentalfaktoren auf die korrekten Antworten mittels ANOVA analysiert. Auch hier wurden alle Haupteffekte und Interaktionen nicht signifikant: Dimension, $F(1, 39) = 0.03$, $p = .855$; Visuell, $F(1, 39) = 1.85$, $p = .181$; Auditorisch, $F(2, 78) = 1.36$, $p = .263$; Dimension \times Visuell, $F(1, 39) = 1.65$, $p = .207$; Dimension \times Auditorisch, $F(2, 78) = 0.13$, $p = .822$; Visuell \times Auditorisch, $F(2, 78) = 0.48$, $p = .618$, sowie Dimension \times Visuell \times Auditorisch, $F(2, 78) = 0.01$, $p = .997$. Zudem schwankten die Bedingungsmittelwerte der korrekten Antworten geringfügig auf einem sehr hohen Niveau zwischen 97.4 % und 99.1 %. Die Daten sprechen daher weder für einen Einfluss der Experimentalfaktoren auf die korrekten Antworten noch für einen Speed-Accuracy-Trade-Off.

6.2.3 Diskussion

Für Experiment 4 konnten keine signifikanten Effekte gefunden werden. Das heißt, dass die experimentelle Stimulation keinen Einfluss auf die Reaktionszeiten auszuüben scheint. Der visuelle Faktor wurde im Gegensatz zu den bisherigen Experimenten nicht signifikant. Dies könnte darauf zurückzuführen sein, dass durch die Anpassung der Kugel an den Hintergrund die Schwierigkeit insoweit erhöht werden konnte, dass die Probanden nicht mehr von visuell passenden Bahnen profitieren konnten. Damit spricht dieses Ergebnis zunächst für eine erfolgreiche Erhöhung der visuellen Schwierigkeit.

Allerdings scheint sich im Gegenzug keine erhöhte Nutzung auditorischer Informationen einzustellen. Die gleichzeitige Tonpräsentation übte keinen Einfluss auf die Entscheidung

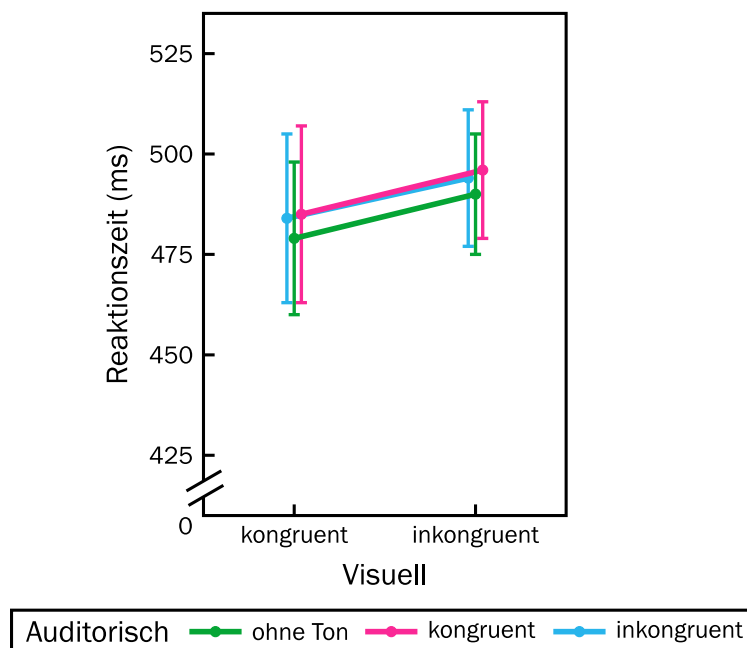


Abbildung 13. Reaktionszeit in Abhängigkeit vom visuellen Bahnverlauf nach der Verdeckung sowie von der auditorischen Stimulation. Die Fehlerbalken kennzeichnen den Standardfehler des Mittelwerts. Zur Unterscheidung der Fehlerbalken sind die Datenpunkte leicht versetzt dargestellt.

aus, ob die Kugel passend oder nicht passend wiedererschien. Dies deutet darauf hin, dass ggf. die Relevanz von Tönen für räumliche Aufgaben allgemein als niedrig zu betrachten ist. Bereits in Experiment 3 konnten die Probanden von den Tönen nicht profitieren, gleichwohl sie dort sogar hinderlich waren. Auch in diesem Experiment konnte keine Leistungsverbesserung durch Töne gefunden werden. Demnach scheinen Töne für räumliche Aufgaben von den Probanden nicht genutzt zu werden, da sie für diese Modalität nicht entscheidend sind (Welch & Warren, 1980; Welch et al., 1986). Alternativ ist denkbar, dass nicht die Modalität und die räumliche Ausrichtung der Aufgabe entscheidend für den fehlenden Toneinfluss ist, sondern die Aufgabenstellung per se. Dass in Diskriminationsaufgaben andere Effekte gefunden werden als für Detektionsaufgaben, konnte z. B. für das Temporal Orienting gezeigt werden (Correa et al., 2004). Auch in einer Studie von Gleiss und Kayser (2013) ergaben sich Toneinflüsse für eine Detektions-, nicht aber eine Diskriminationsaufgabe. Es könnten also Töne generell in (räumlichen) Diskriminationsaufgaben einen wenig hilfreichen Cue darstellen, der entsprechend von den Versuchspersonen nicht für die Bearbeitung der Aufgabe genutzt wird, selbst wenn weitere Informationsquellen wenig reliable Informationen liefern.

Ebenso konnte kein Einfluss der Dimensionsdarbietung gefunden werden. Es ergaben sich keine Unterschiede zwischen pseudo- und echt-3D-Stimulation. Im Gegensatz zu Experiment 1 und 3, in denen Einflüsse gefunden werden konnten, wurde hier die visuelle Erkennbarkeit reduziert. Es ist denkbar, dass ein Einfluss der Präsentationsdimension nur zum Tragen kommt, wenn der visuelle Input eindeutig gestaltet ist. Da hier durch die Anpassung der Kugel an den Schachbrettboden die visuelle Schwierigkeit erhöht war, könnte der Einfluss des visuellen Tiefenmerkmals Disparität an Bedeutung verloren haben, da der visuelle Input in

seiner Erkennbarkeit im Allgemeinen bereits reduziert war. Der Einfluss der Tiefendarbietung für Effekte alltagsnahen Materials wird daher im folgenden Experiment 5 genauer überprüft.

Aufgrund der nicht signifikanten Effekte in diesem Experiment und der wenig eindeutigen Ergebnisse aller drei audiovisuellen Diskriminationsexperimente, wurden diese gesamt post-hoc erneut betrachtet (siehe Abschnitt 6.4).

6.3 Experiment 5: Vereinfachtes zweidimensionales Stimulusmaterial

In diesem Experiment sollte der Einfluss von realitätsgetreuem Material genauer untersucht werden. Dafür wurde im Gegensatz zu den ersten vier Experimenten kein Vergleich zwischen Experimenten mit Material, das Disparitätsinformationen enthält und solchen ohne, durchgeführt. Stattdessen wurde das alltagsnahe echt-3D-Material mit einer vereinfachten 2D-Variante verglichen. Es wurde in diesem Experiment eine einfach-2D-Variante des Materials erstellt und als Vergleichsbedingung genutzt. Dafür wurden im Vergleichsmaterial neben den Disparitätsinformationen zusätzlich die perspektivischen Tiefenhinweise sowie der Schattenwurf der Objekte entfernt. Dadurch ergab sich ein visuelles Display mit einfachen geometrischen Formen – ähnlich dem häufig verwendeten planaren visuellen Material in der Mehrzahl der Studien zur Bewegungswahrnehmung (z. B. Chien et al., 2013; Freeman & Driver, 2008; Kirkham et al., 2012; Prime & Harris, 2010).

6.3.1 Methoden

An Experiment 5 nahmen 40 Probanden (je 20 männlich/weiblich) mit einem Durchschnittsalter von 23.80 Jahren ($SD = 3.36$) teil. Davon waren drei links- und 37 rechtshändig. Bis auf die folgend genannten Änderungen entsprachen alle Details zum experimentellen Ablauf und den verwendeten Stimuli denen in Experiment 3.

Für den Vergleich mit standardmäßig häufig verwendetem 2D-Stimulusmaterial wurde eine an die visuelle 3D-Szene angelehnte 2D-Szene erstellt. Dafür wurde die Box mit dem Schachbrettboden in einen Schachbretthintergrund mit Begrenzungen an der linken und rechten Seite in den gleichen Farben wie in der 3D-Variante umgewandelt (Höhe: 19.85° , Breite: 41.61°). Die Brücke wurde durch einen 38.99° breiten und 5.15° hohen grünen Balken ersetzt. Abbildung 14 zeigt die neu erstellte 2D-Variante der visuellen Stimulation. Der Balken verdeckte die Bewegung des nun 2D-Kreises (Durchmesser: 2.18° , ebenfalls magentafarben mit mittelblauem Streifen) für den gleichen Zeitraum wie die Brücke die Kugel in der 3D-Variante (567 ms). Der Flip, in dem die Verdeckung der Kugel begann und ebenso der Flip des Wiederauftauchens entsprachen der 3D-Variante des Experiments. Der Kreis bewegte sich ebenfalls diagonal über den Hintergrund, die Start- und Endpositionen der Bewegung wurden ebenfalls an die 3D-Variante angepasst. Entsprechend der einfachen 2D-Gestaltung veränderte sich der Faktor Dimension im Design in die Stufen einfach-2D und echt-3D.

Die Probanden entschieden erneut über eine Diskriminationsaufgabe analog den Experimenten 3 und 4, ob die Kugel bzw. der Kreis nach der Verdeckung einer passenden Bahn folgte oder nicht. Auch hier wurde die Tastenzuweisung wieder gemeinsam mit der Experimentreihenfolge und dem Geschlecht ausbalanciert.

6.3.2 Ergebnisse

Es wurden 7.13 % der Durchgänge für die Reaktionszeit- und 5.19 % für die Korrekte-Antworten-Analyse entfernt, davon waren für beide Analysevariablen 0.10 % Trials mit Flipungenauigkeiten.

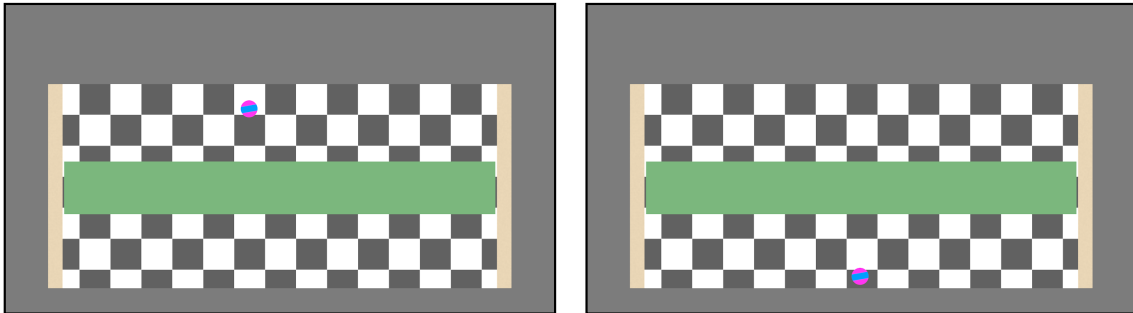


Abbildung 14. Beispielhaftes Bildmaterial aus Experiment 5 und 6. Es wurde einfaches 2D-Material als Vergleich zur echt-3D-Stimulation verwendet. Links ist die Start-, rechts die Endposition der Bewegung des Kreises abgebildet. Dieser bewegte sich auf vier möglichen Bahnen von oben nach unten über den Schachbretthintergrund. Die Probanden diskriminierten, ob der Kreis sich nach dem grünen Balken auf einer visuell passenden oder nicht passenden Bahn weiterbewegte. In der echt-3D-Version der Experimente wurde das echt-3D-Material aus Experiment 3 verwendet.

Reaktionszeit-Analyse Bei der Analyse der Reaktionszeiten ergab sich ein marginal signifikanter Effekt für die visuelle Stimulation, $F(1, 39) = 3.80$, $p = .059$. Die Probanden reagierten schneller für visuell passende Bahnen (474 ms) als für nicht passende (488 ms). Alle weiteren Haupteffekte und Interaktionen wurden nicht signifikant: Dimension, $F(1, 39) = 0.37$, $p = .544$; Auditorisch, $F(1.51, 59.05) = 1.28$, $p = .278$; Dimension \times Visuell, $F(1, 39) = 0.01$, $p = .934$; Dimension \times Auditorisch, $F(2, 78) = 0.24$, $p = .784$; Visuell \times Auditorisch, $F(2, 78) = 0.48$, $p = .623$, und Dimension \times Visuell \times Auditorisch, $F(2, 78) = 0.48$, $p = .620$. Die Reaktionszeiten bei visueller und auditorischer Stimulation sind in Abbildung 15 zu sehen.

Eine Post-hoc-Detailanalyse der Reaktionszeiten für Experiment 5 findet sich in der gemeinsamen Betrachtung aller drei audiovisuellen Diskriminationsexperimente in Abschnitt 6.4.

Analyse der korrekten Antworten Um einen Speed-Accuracy-Trade-Off ausschließen zu können, wurden neben den Reaktionszeiten die korrekten Antworten ausgewertet. Die Interaktion zwischen den Faktoren Dimension und Visuell wurde signifikant, $F(1, 39) = 12.06$, $p = .001$. Alle weitere Haupteffekte und Interaktionen wurde nicht signifikant: Dimension, $F(1, 39) = 0.59$, $p = .445$; Visuell, $F(1, 39) = 0.94$, $p = .338$; Auditorisch, $F(2, 78) = 2.17$, $p = .121$; Dimension \times Auditorisch, $F(2, 78) = 1.40$, $p = .252$; Visuell \times Auditorisch, $F(2, 78) = 1.18$, $p = .312$, sowie Dimension \times Visuell \times Auditorisch, $F(1.41, 54.81) = 0.97$, $p = .358$. Die Interaktion Dimension \times Visuell zeigte sich post-hoc in einem Unterschied während echt-3D-Stimulation zwischen visuell passender (98.8 %) und nicht passender Stimulation (97.8 %), $p = .010$, wohingegen sich die beiden visuellen Bedingungen bei der pseudo-3D-Stimulation nicht unterschieden (97.9 % vs. 98.4 %), $p = .470$.

Die Effekte auf die korrekten Antworten waren nicht gegenläufig zu denen der Reaktionszeiten, zudem zeigen die korrekten Antworten der verschiedenen Bedingungen nur geringe numerische Unterschiede (zwischen 97.2 % und 98.8 %). Die Daten deuten demnach nicht auf einen Speed-Accuracy-Trade-Off hin.

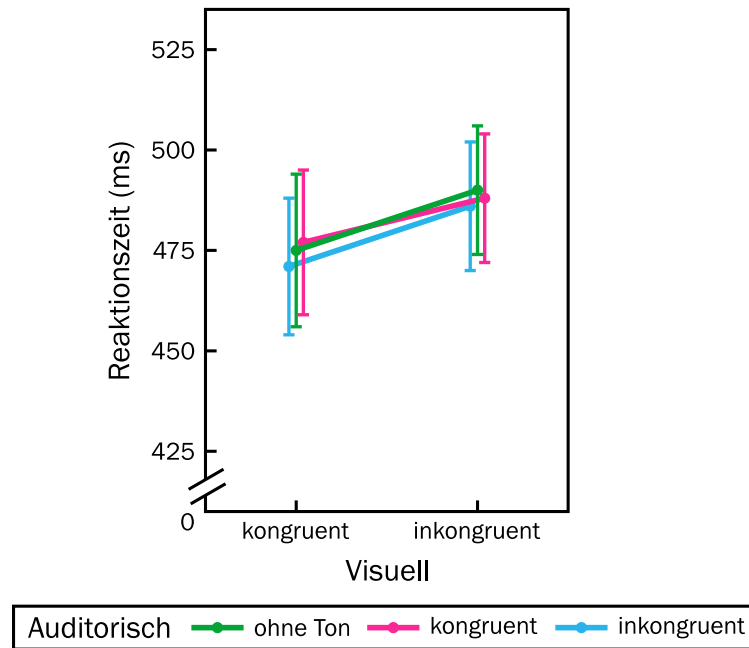


Abbildung 15. Reaktionszeit in Abhängigkeit vom visuellen Bahnverlauf nach der Verdeckung sowie von der auditorischen Stimulation. Die Fehlerbalken kennzeichnen den Standardfehler des Mittelwerts. Zur Unterscheidung der Fehlerbalken sind die Datenpunkte leicht versetzt dargestellt.

6.3.3 Diskussion

Auch für Experiment 5 zeigte sich ein Einfluss der visuellen Stimulation, wenn auch nur auf marginal signifikantem Niveau. Die Probanden profitierten von visuell passenden Bahnen und reagierten schneller als bei nicht passenden Bahnen. Die Töne wirkten sich jedoch auch hier nicht signifikant aus. Dies bestätigt die Ergebnisse von Experiment 4 und deutet darauf hin, dass auditorische Stimulation wahlweise für Aufgaben mit räumlichen Fokus und/oder Diskriminationsaufgaben keine Auswirkung auf die Reaktionszeiten ausübt. Auch der fehlende Einfluss der Präsentationsdimension unterstreicht den in Experiment 4 ebenfalls nicht gefundenen Effekt. Die Ergebnisse unterscheiden sich nicht zwischen einfach-2D und echt-3D-Präsentation, was für den Vergleich von pseudo- und echt-3D in Experiment 3 der Fall war. Somit scheint die Disparität und gleichzeitig ein gut sichtbarer visueller Input das entscheidende Kriterium für einen Einfluss der Präsentationsdimension zu sein. Allerdings soll an dieser Stelle darauf hingewiesen werden, dass die ausführliche Post-hoc-Analyse der drei Diskriminationsexperimente darauf hindeutet, dass die Präsentationsdimension in diesen Experimenten im Allgemeinen keinen Einfluss auszuüben scheint. Details dazu sind in Abschnitt 6.4 zu finden.

6.4 Post-hoc-Betrachtung der audiovisuellen Diskriminationsexperimente

Die Auswertung der Diskriminationsexperimente brachte nur wenige, instabile Effekte hervor. Diese waren entweder wenig intuitiv wie in Experiment 3, in dem für die pseudo-3D-Umgebung keine Effekte gefunden wurden, auch nicht auf Ebene der Haupteffekte oder aber nicht bzw. nur marginal signifikant – wie dies in Experiment 4 und 5 der Fall war. Daher wurden die Ergebnisse dieser drei Experimente post-hoc genauer betrachtet.

6.4.1 Bayes-Faktor-Analyse

Zunächst wurde mit den Daten eine Bayes'sche ANOVA (Rouder et al., 2012) durchgeführt, um herauszufinden, welches Erklärungsmodell die größte Evidenz für die Daten liefert. Dafür wurden die Daten mit dem Programm JASP (Version 0.8.5.1, <https://jasp-stats.org/>) jeweils mit Hilfe einer 2x2x3-Bayes-ANOVA mit Messwiederholung analysiert. Dafür wurden die Prior-Voreinstellungen sowie der Vergleich mit dem jeweils besten Modell verwendet. Die gesamten Bayes-Faktoren aller Modelle für die audiovisuellen Diskriminationsexperimente sind in Tabelle 2 aufgeführt. Für Experiment 3 und 5 ergibt sich eine eindeutige Präferenz für das Modell Visuell, die Bayes-Faktoren von 55.308 und 24.351 sprechen für eine (sehr) starke Evidenz. Zudem sind sie um einiges wahrscheinlicher als das jeweils nächst bessere Modell Dimension mit Bayes-Faktoren von 0.320 bzw. 0.293. Für Experiment 4 zeigt sich keine eindeutige Evidenz für ein Modell, zwar fällt auch hier der Bayes-Faktor für Visuell am höchsten aus, weist mit einem Wert von 0.954 jedoch nicht auf eine Evidenz für eine Alternativhypothese hin, sodass hier von einer größeren Evidenz für die Nullhypothese ausgegangen werden sollte.

6.4.2 Interindividuelle Unterschiede in der Nutzung visueller und auditorischer Informationen

Unterschiede zwischen den Probanden könnten eine Ursache für die weniger eindeutigen Daten in den Diskriminationsexperimenten darstellen. Zum einen konnte in Experiment 2 gezeigt werden, dass ein Großteil der Versuchspersonen sich darin unterschied, für welche Schätzdi- stanz sie Töne gewinnbringend für die Vorhersage nutzen konnten und damit interindividuelle

Tabelle 2

Bayes-Faktoren (BF) für die audiovisuellen Diskriminationsexperimente

Effekt	BF Experiment 3	BF Experiment 4	BF Experiment 5
Dimension	0.320	0.125	0.293
Visuell	55.308	0.954	24.351
Auditorisch	0.040	0.040	0.031
Dimension × Visuell	0.200	0.210	0.141
Dimension × Auditorisch	0.049	0.047	0.041
Visuell × Auditorisch	0.058	0.044	0.048
Dimension × Visuell × Auditorisch	0.077	0.096	0.086

Anmerkungen. Berechnungen mit JASP unter Standard-Prior-Einstellungen und Vergleich mit dem besten Modell.

Unterschiede vorliegen. Zum anderen konnten zahlreiche Studien in der Vergangenheit interindividuelle Unterschiede bei multisensorischen Interaktionen belegen. Zunächst soll daher kurz die relevante Literatur für interindividuelle Unterschiede in Extrapolations- oder Multisensorikexperimenten vorgestellt werden. Im Anschluss wird das hier gewählte Vorgehen der Nutzergruppeneinteilung und Crossvalidierung erläutert. Es folgen die Darstellung der Ergebnisse der Post-hoc-Betrachtungen sowie ein zusammenfassendes Zwischenfazit.

Interindividuelle Unterschiede in der Bewegungsvorhersage und Multisensorik Nicht nur die visuelle (Halpern et al., 1999) und auditorische (Johnson et al., 1986; Kidd et al., 2007) Wahrnehmung im Allgemeinen, auch andere Prozesse, in denen visuelle Fähigkeiten involviert sind, weisen individuelle Unterschiede zwischen Probanden auf. So wurden z. B. für die visuelle Suche Unterschiede bei der Aufgabenbearbeitung gefunden (Becic et al., 2007; Boot et al., 2006, 2009). Die Probanden unterscheiden sich vor allem in ihren Augenbewegungen bei der visuellen Suche (offene vs. verdeckte Suche) sowie in der Nutzung verschiedener Strategien (seriell vs. non-seriell) und perzeptueller Gruppierungsmechanismen (lokal vs. global, Hogeboom & van Leeuwen, 1997). Auch für die attentive Verfolgung sich bewegender Objekte konnten Unterschiede in der Leistung zwischen den Probanden festgestellt werden (Drew & Vogel, 2008). Ebenso unterscheiden sich Versuchspersonen in ihrer Fähigkeit die Ankunftszeit sich bewegender Objekte vorherzusagen, wobei sie sich sowohl in der Leistung selbst als auch im Bias (Über- oder Unterschätzung) ihrer Antworten unterscheiden (Lugtigheid & Welchman, 2011). Die Experimente von Law et al. (1993) deuten darauf hin, dass dies mit der unterschiedlichen Ausprägung der Fähigkeit verschiedene Informationen (in deren Studie die relative Geschwindigkeit und die relative Distanz zweier Objekte zueinander) für die Vorhersage zu integrieren, zusammenhängt. Diese Unterschiede in der Bewegungsvorhersage zeigen sich auch in verschieden ausgeprägten Regressionsgeraden, die für jeden Probanden bei einer visuellen Time-to-Arrival-Schätzung aufgestellt wurden (Yakimoff et al., 1993).

Auch für multisensorische Phänomene konnte in zahlreichen Untersuchungen interindividuelle Variabilität in den Ergebnissen gefunden werden (Baum et al., 2016). Unterschiede zeigen sich zunächst für den McGurk-Effekt. Bei diesem tritt die Wahrnehmung einer anderen gesprochenen Silbe auf, wenn über ein Video und die Tonspur einer sprechenden Person unterschiedliche Silben präsentiert werden. Probanden unterscheiden sich in ihrer Empfänglichkeit und Frequenz den McGurk-Effekt wahrzunehmen (Ipser et al., 2017; Magnotti & Beauchamp, 2015; Mallick et al., 2015; Nath & Beauchamp, 2012). Auch bei Lokalisierung auditorischer Targets unter gleichzeitigem visuellen Einfluss (crossmodaler Bias) differieren die Leistungen der Versuchspersonen in hohem Maße untereinander (Hairston et al., 2003). Umgekehrt, beim Einfluss der auditorischen auf die visuelle Modalität, wird die schwelennahe Wahrnehmung visueller Reize durch Töne verschieden stark verbessert, wobei die Verbesserung durch Töne stärker bei Probanden ausfällt, die in den visuellen Bedingungen schlechter abschnitten (Caclin et al., 2011). Auch Wozny et al. (2010) ließen ihre Probanden eine audiovisuelle Lokalisationsaufgabe durchführen und diese unterschieden sich in ihrer Strategie während

der Aufgabenbearbeitung, wobei die Mehrheit der Probanden eine statistisch nicht optimale Strategie nutzte.

Am häufigsten wurden in Untersuchungen zu interindividuellen Unterschieden für multisensorische Interaktionen bisher verschiedene Formen der Gleichzeitigkeits- oder Reihenfolgebeurteilung verwendet. Es ergaben sich Unterschiede zwischen den Versuchspersonen im Temporal-Order-Judgement oder der Gleichzeitigkeits-/Synchronitätsbeurteilung für audiovisuelle Stimuli innerhalb einer homogenen Stichprobe (Grabot & van Wassenhove, 2017; Freeman et al., 2013; Kaganovich & Schumaker, 2016) und zwischen zwei verschiedenartigen Stichproben wie z. B. Videospiele vs. Nicht-Videospiele oder Musiker vs. Nicht-Musiker (Donohue et al., 2010; H. Lee & Noppeney, 2014). Bei der Überprüfung der Vorhersagen des Maximum-Likelihood-Estimation-Modells bei einer Aufgabe dieser Art fanden Maiworm und Röder (2011), dass sich die Probanden aufteilen ließen: Die eine Hälfte erzielte Ergebnisse, die anhand des Modells vorhergesagt wurden, die andere Hälfte teilte hingegen dem auditorischen Signal bei der Gleichzeitigkeitsbeurteilung systematisch ein zu hohes Gewicht zu. Zudem sind nicht alle Versuchspersonen in der Lage ihr multisensorisches Integrationsfenster, das in solchen Experimenten bestimmt werden kann, an sich ändernde Anforderungen zu adaptieren, wenn anstelle auf die Reihenfolge beider Stimuli nur auf den ersten präsentierten Stimulus reagiert werden soll. Nur die Hälfte der Probanden konnte sich durch eine Rekalibrierung ihres Fensters optimal an diese Aufgabe anpassen (Mégevand et al., 2013).

Auch in Experimenten mit komplexerem audiovisuellen Stimulusmaterial ergaben sich Varianzen in den Probandenleistungen (Eg & Behne, 2015): Der Point of Subjective Simultaneity (Punkt subjektiver Gleichzeitigkeit, an dem Probanden angeben, dass zwei Stimuli gleichzeitig präsentiert wurden) schwankt stark zwischen den Versuchspersonen für realitätsnahes audiovisuelles Videomaterial (Szenen eines Schachspiels, eines Schlagzeugspiels oder einer Nachrichtensendung). Außerdem konnten Stevenson et al. (2012) aufzeigen, dass neben den bereits vielfach gefundenen interindividuellen Differenzen in der Gleichzeitigkeitsbeurteilung zudem ein Zusammenhang zu weiteren multisensorischen Phänomenen besteht. Sie fanden eine Korrelation zwischen der Breite des Integrationsfensters (für den Fall, dass der auditorische Stimulus auf den visuellen folgt) mit der Stärke der Anfälligkeit für multisensorische Illusionen wie dem McGurk-Effekt und der Sound-Induced-Flash-Illusion (Wahrnehmung eines visuellen Blinkens, wenn während eines durchgängigen visuellen Stimulus gleichzeitig mehrere Töne präsentiert werden). Individuen unterscheiden sich ebenso in der Fähigkeit, trisensorische (audiovisuotaktile) Stimuli zu integrieren. In den Experimenten von Hagmann und Russo (2016) konnte die eine Hälfte der Probanden von trisensorischer Stimulation profitieren, während sich für die andere Hälfte kein Vorteil ergab.

Eine unterschiedliche Nutzung bimodaler Informationen zwischen den Probanden ergab sich in einer Untersuchung mit spielerischem Material von Sun et al. (2017). Hier wurde eine Art Angelspiel für die Probanden simuliert. Sie sahen Fisch-Stimuli, die sich durch einen Teich bewegten und währenddessen in ihrer Größe oszillierten. Die Versuchspersonen hatten die Aufgabe, Fische zu finden, die eine bestimmte Zieloszillationsrate aufwiesen. Die Fischbewegungen konnten zusätzlich von einem amplitudenmodulierten Ton begleitet werden,

wobei die Bewegung des Tons mit der Oszillationsrate der Fische übereinstimmen konnte oder nicht. Die Probanden unterschieden sich in ihrer Rate, mit der sie aus den Stimuli audiovisuelle Informationen ziehen konnten. Sie wurden in unterschiedlichem Ausmaß von gleichzeitigen auditorischen Informationen beeinflusst.

Zusammenfassend lassen sich in der Literatur diverse Hinweise für individuelle Unterschiede zwischen den Probanden für die Bewegungsvorhersage bzw. multisensorische Phänomene sowie insbesondere in der Nutzung visueller und auditorischer Informationen finden. Des Weiteren konnten diese auch für Experiment 2 durch eine nachträgliche crossvalidierte Nutzergruppeneinteilung und -analyse aufgezeigt werden. Daher soll in der Post-hoc-Betrachtung der Daten für die audiovisuellen Diskriminationsexperimente ebenfalls ein solches Vorgehen genutzt werden.

Nutzergruppeneinteilung und Crossvalidierung Um der Vermutung interindividueller Unterschiede in den hier durchgeführten audiovisuellen Diskriminationsexperimenten nachzugehen, wurden die Probanden ähnlich zu Experiment 2 in verschiedene Nutzergruppen aufgeteilt. Dafür wurden sie gemäß der für die Vorhersage zur Verfügung stehenden Informationen eingeteilt. Es wurden zwei Kriterien angewendet: 1) Die Nutzung visueller Informationen über den visuellen Erleichterungseffekt. Dafür wurde die Differenz zwischen den Bedingungen visuell passend/unimodal und visuell nicht passend/unimodal berechnet. Bei einer negativen Differenz wurde der Proband als Visuell-Nutzer eingestuft, bei positiver Differenz als Nicht-Visuell-Nutzer. 2) Die Nutzung auditorischer Informationen wurde analog mittels auditorischem Erleichterungseffekt zugeteilt. Es wurde die Differenz zwischen den Bedingungen visuell passend/auditorisch passend sowie visuell passend/unimodal berechnet und die Probanden entsprechend als Auditorisch-Nutzer oder Nicht-Auditorisch-Nutzer deklariert. Vereinigt man beide Kriterien, ergeben sich vier Kombinationsmöglichkeiten in der Informationsnutzung bei der Vorhersage: Nutzung visueller Informationen (Bildnutzer), Nutzung auditorischer Informationen (Tonnutzer), Nutzung audiovisueller Informationen (Beidnutzer) und Nutzung weder visueller noch auditorischer Informationen (Nichtnutzer)¹.

Wie in Experiment 2 wurde die Kategorisierung der Probanden jeweils mit einer Hälfte des Datensatzes durchgeführt, mit der zweiten Hälfte erfolgte die Analyse der Reaktionszeiten für die Nutzergruppen separat. Zur Crossvalidierung wurde jeder Auswertungsschritt für beide Datenhälften durchgeführt: Zunächst wurden gerade Blöcke zur Kategorisierung und ungerade zur Analyse, im Anschluss ungerade Blöcke zur Kategorisierung und gerade zur Analyse genutzt. Bei den hier vorgestellten Ergebnissen wurden für alle Analysen die Reaktionszeiten ausgewertet. Die Analyse der korrekten Antworten für die Nutzergruppen und entsprechenden Auswertreihenfolgen finden sich im Anhang C.

¹Mit den Begriffen „(Nicht)Nutzer“ bzw. „(nicht) nutzen“ ist im Zusammenhang mit der Post-hoc-Analyse in den audiovisuellen Diskriminationsexperimenten gemeint, dass die Probanden einen entsprechend ausgeprägten visuellen bzw. auditorischen Erleichterungseffekt aufweisen und somit in den dazugehörigen Bedingungen von visuell kongruenten/zusätzlichen auditorischen Informationen profitieren bzw. nicht von diesen profitieren. Es soll damit nicht zum Ausdruck gebracht werden, dass die Probanden bei entsprechender Kategorisierung und Bezeichnung als „Nichtnutzer“ visuelle bzw. auditorische Informationen vollständig nicht nutzen.

Ergebnisdarstellung In Tabelle 3 sind die Ergebnisse der Probanden-Kategorisierung für die vier Nutzergruppen auf Basis gerader bzw. ungerader Blöcke für die drei Experimente dargestellt. Die numerischen Werte der Reaktionszeit-Differenzen für jeden Probanden und die entsprechende Einteilung in die Nutzergruppen können den Tabellen A3 bis A8 im Anhang entnommen werden. Da die Ton- und Nichtnutzer für alle drei Experimente in beiden Kategorisierungen $n < 10$ war, wurden diese für die Reaktionszeit-Analyse im Folgeschritt nicht berücksichtigt und nur die Daten der Beid- und Bildnutzer analysiert. Post-hoc-Analysen und -Interpretationen etwaig signifikanter Effekte erfolgten nur, wenn sich diese stabil zeigten, d. h. sowohl bei der Gerade-Ungerade- als auch der Ungerade-Gerade-Auswertung auftraten (siehe Tabelle 4).

Experiment 3

Bei der Analyse der ungeraden Blöcke der Beidnutzer ($n = 6$) nach gerader Kategorisierung ergaben sich ein signifikanter Haupteffekt für den Faktor Visuell, $F(1, 5) = 28.74, p = .003$, sowie eine signifikante Interaktion von Dimension und Auditorisch, $F(2, 10) = 4.44, p = .042$. Alle weiteren Effekte wurden nicht signifikant, $F \leq 3.21, p \geq .084$. Für den geraden Datensatz nach ungerader Nutzergruppenzuweisung ($n = 12$ Beidnutzer) wurde ebenfalls der visuelle Faktor signifikant, $F(1, 11) = 26.24, p < .001$, alle anderen Faktoren und deren Interaktionen waren nicht signifikant, $F \leq 1.05, p \geq .365$. Die Reaktionszeiten in Abhängigkeit von der visuellen und auditorischen Stimulation für die Beidnutzer sind in Abbildung 16 oben dargestellt. Die weiteren statistischen Kennwerte der ANOVAs können Tabelle B1 im Anhang entnommen werden. Da nur der visuelle Faktor in beiden Analysen signifikant wurde (siehe auch Tabelle 4), wurde nur dieser post-hoc betrachtet. Die Probanden reagierten in beiden Analysen schneller für visuell passende (401 ms bzw. 427 ms) als nicht passende Pfade nach der Verdeckung (450 ms bzw. 465 ms).

Die Analyse des ungeraden Datensatzes ergab für die Bildnutzer ($n = 19$) einen signifikanten Haupteffekt der visuellen Stimulation, $F(1, 18) = 21.58, p < .001$. Alle weiteren Faktoren und Interaktionen wurden nicht signifikant, $F \leq 2.78, p \geq .095$. Bei der Betrachtung der geraden Blöcke nach Kategorisierung mit ungeraden Blöcken ($n = 14$) zeigte sich ebenfalls ein signifikanter Einfluss der visuellen Präsentation, $F(1, 13) = 15.85, p = .002$. Auch die Töne übten einen signifikanten Einfluss aus, $F(2, 26) = 5.23, p = .012$. Die weiteren Effekte wurden nicht signifikant, $F \leq 2.59, p \geq .132$. Die gesamten statistischen Ergebnisse beider ANOVAs für die Bildnutzer sind Tabelle B2 im Anhang zu entnehmen. Nur der visuelle Einfluss wurde für beide Analysen signifikant (siehe auch Tabelle 4), hier ergab sich post-hoc in beiden Fällen ein Vorteil für visuell passende Bahnen (461 ms bzw. 472 ms) gegenüber nicht passenden Bahnen (489 ms bzw. 501 ms). Die Reaktionszeiten sind für die visuelle und auditorische Stimulation bei den Bildnutzern in Abbildung 16 unten dargestellt.

Tabelle 3

Anzahl der Probanden für die jeweilige Nutzergruppeneinteilung für gerade und ungerade Blöcke in den drei audiovisuellen Diskriminationsexperimenten

Nutzergruppe	Experiment 3	Experiment 4	Experiment 5
Einteilung auf Basis gerader Blöcke			
Beidnutzer	6	12	11
Bildnutzer	19	15	16
Tonnutzer	6	5	8
Nichtnutzer	9	8	5
Einteilung auf Basis ungerader Blöcke			
Beidnutzer	12	11	11
Bildnutzer	14	17	17
Tonnutzer	8	8	8
Nichtnutzer	6	4	4

Anmerkungen. Anzahl der Probanden gemäß Kategorisierung für visuelle und auditorische Erleichterungseffekte in Beidnutzer (nutzen beide Arten von Informationen), Bildnutzer/Tonnutzer (nutzen jeweils nur visuelle/auditorische Informationen) und Nichtnutzer (nutzen keine der beiden Informationen). Dargestellt für die Kategorisierung auf Basis gerader Blöcke (oben) bzw. ungerader Blöcke (unten). Details zur Nutzergruppeneinteilung siehe Text.

Experiment 4

Im Anschluss an die Kategorisierung mit geraden Blöcken wurde für die Beidnutzer ($n = 12$) der Einfluss visueller Stimulation signifikant, $F(1, 11) = 39.61$, $p < .001$. Auch für die Analyse nach ungerader Kategorisierung ($n = 11$) war dieser Faktor signifikant, $F(1, 10) = 18.31$, $p = .002$. Alle weiteren Effekte und Interaktionen waren für beide Auswertreihenfolgen nicht signifikant, $F \leq 2.75$, $p \geq .088$. Die Probanden waren in beiden Fällen schneller, wenn die Kugel nach der Verdeckung visuell passend wiedererschien (425 ms bzw. 418 ms) im Vergleich zu nicht passenden Bahnen (449 ms bzw. 444 ms). Die gesamten statistischen Kennwerte beider Analysen sind in Tabelle B3 im Anhang zu finden, die Reaktionszeiten für die Beidnutzer sind in Abbildung 17 oben dargestellt.

Bei den Bildnutzern wurde ebenfalls der visuelle Faktor sowohl nach gerader Kategorisierung im ungeraden Datensatz ($n = 15$), $F(1, 14) = 17.68$, $p = .001$, als auch nach ungerader Einteilung für die geraden Blöcke ($n = 17$) signifikant, $F(1, 16) = 10.51$, $p = .005$. In der Analyse der geraden Blockdaten übte ebenso die Dimension einen signifikanten Einfluss aus, $F(1, 16) = 4.55$, $p = .049$. Da nur der visuelle Faktor in beiden Auswertungen signifikant wurde (siehe auch Tabelle 4), wurde nur dieser post-hoc in seiner Richtung betrachtet. Für beide Analysen reagierten die Probanden schneller für visuell passende Bahnen (472 ms bzw. 465 ms) als für nicht passende Bahnen (519 ms bzw. 502 ms). Alle weiteren Effekte wurden nicht signifikant $F \leq 2.33$, $p \geq .116$. Die gesamten statistischen Kennwerte für die Analysen der Bildnutzer sind in Tabelle B4 im Anhang aufgeführt, die Reaktionszeiten für die visuelle und auditorische Stimulation sind in Abbildung 17 unten gezeigt.

Tabelle 4

Vergleich der Ergebnisse für die drei audiovisuellen Diskriminationsexperimente für die Analyse gerader und ungerader Blöcke

Effekt	GU E3	UG E3	GU E4	UG E4	GU E5	UG E5
Beidnutzer						
Dimension	n. s.	n. s.	n. s.	n. s.	n. s.	n. s.
Visuell	s.	s.	s.	s.	s.	s.
Auditorisch	n. s.	n. s.	n. s.	n. s.	n. s.	m. s.
Dimension × Visuell	n. s.	n. s.	n. s.	n. s.	n. s.	n. s.
Dimension × Auditorisch	s.	n. s.	n. s.	n. s.	m. s.	n. s.
Visuell × Auditorisch	m. s.	n. s.	n. s.	n. s.	n. s.	n. s.
Dimension × Visuell × Auditorisch	n. s.	n. s.	n. s.	m. s.	n. s.	n. s.
Bildnutzer						
Dimension	n. s.	n. s.	n. s.	s.	n. s.	n. s.
Visuell	s.	s.	s.	s.	s.	s.
Auditorisch	n. s.	s.	n. s.	n. s.	m. s.	s.
Dimension × Visuell	n. s.	n. s.	n. s.	n. s.	n. s.	n. s.
Dimension × Auditorisch	m. s.	n. s.	n. s.	n. s.	n. s.	m. s.
Visuell × Auditorisch	n. s.	n. s.	n. s.	n. s.	n. s.	s.
Dimension × Visuell × Auditorisch	n. s.	n. s.	n. s.	n. s.	n. s.	n. s.

Anmerkungen. Effekte-Übersicht für Experiment 3 (E3), Experiment 4 (E4) und Experiment 5 (E5) für die Analyse ungerader Blöcke nach Kategorisierung auf Basis gerader Blöcke (GU) bzw. für die Analyse gerader Blöcke nach Kategorisierung auf Basis ungerader Blöcke (UG). Es wurde eine Varianzanalyse mit den Inner-subjektfaktoren Dimension, Visuell und Auditorisch durchgeführt. Effekte, die für die GU- und UG-Auswertung in beiden Fällen signifikant waren, sind fett hervorgehoben. s. = signifikant, m. s. = marginal signifikant, n. s. = nicht signifikant.

Experiment 5

Für die Auswertung in der Gerade-Ungerade-Reihenfolge ergab sich für die Beidnutzer ($n = 11$) ein signifikanter Einfluss der visuellen Bahn, $F(1, 10) = 12.07$, $p = .006$, welcher ebenfalls für die umgekehrte Reihenfolge Ungerade-Gerade ($n = 11$) signifikant wurde, $F(1, 10) = 14.76$, $p = .003$. Die Probanden reagierten in beiden Fällen für passendes Wiedererscheinen (412 ms bzw. 423 ms) schneller als für nicht passendes (450 ms bzw. 452 ms). Die weiteren Effekte zeigten kein signifikantes Ergebnis, $F \leq 3.04$, $p \geq .070$. Die numerischen Ergebnisse der Analysen sind in Tabelle B5 im Anhang zu finden. Die Ergebnisse der Beidnutzer sind in Abbildung 18 oben dargestellt.

Bei den Bildnutzern konnten erneut sowohl für die Gerade-Ungerade-Auswertung ($n = 16$) als auch für das umgekehrte Vorgehen ($n = 17$) signifikante visuelle Faktoren gefunden werden, $F(1, 15) = 28.69$, $p < .001$, bzw. $F(1, 16) = 28.21$, $p < .001$. Für die Ungerade-Gerade-Auswertung wurden zudem die auditorische Stimulation, $F(2, 32) = 4.10$, $p = .026$, sowie die Interaktion zwischen Visuell und Auditorisch signifikant, $F(2, 32) = 4.95$, $p = .013$. Die restlichen Effekte überschritten das Signifikanzniveau nicht, $F \leq 3.17$, $p \geq .070$. Die numerischen Werte der ANOVA-Kennwerte sind für beide Auswertreihenfolgen in Tabelle B6 im Anhang aufgeführt. Da nur der visuelle Faktor in beiden Auswertungen signifikant wurde

(siehe auch Tabelle 4) wurde nur für diesen eine Post-hoc-Betrachtung durchgeführt. In beiden Fällen reagierten die Probanden schneller, wenn die Kugel nach der Verdeckung passend (425 ms bzw. 430 ms) wiedererschien verglichen zu nicht passenden Bedingungen (466 ms bzw. 467 ms). Die Ergebnisse für die Bildnutzer in Abhängigkeit der visuellen und auditorischen Stimulation sind in Abbildung 18 unten zu sehen.

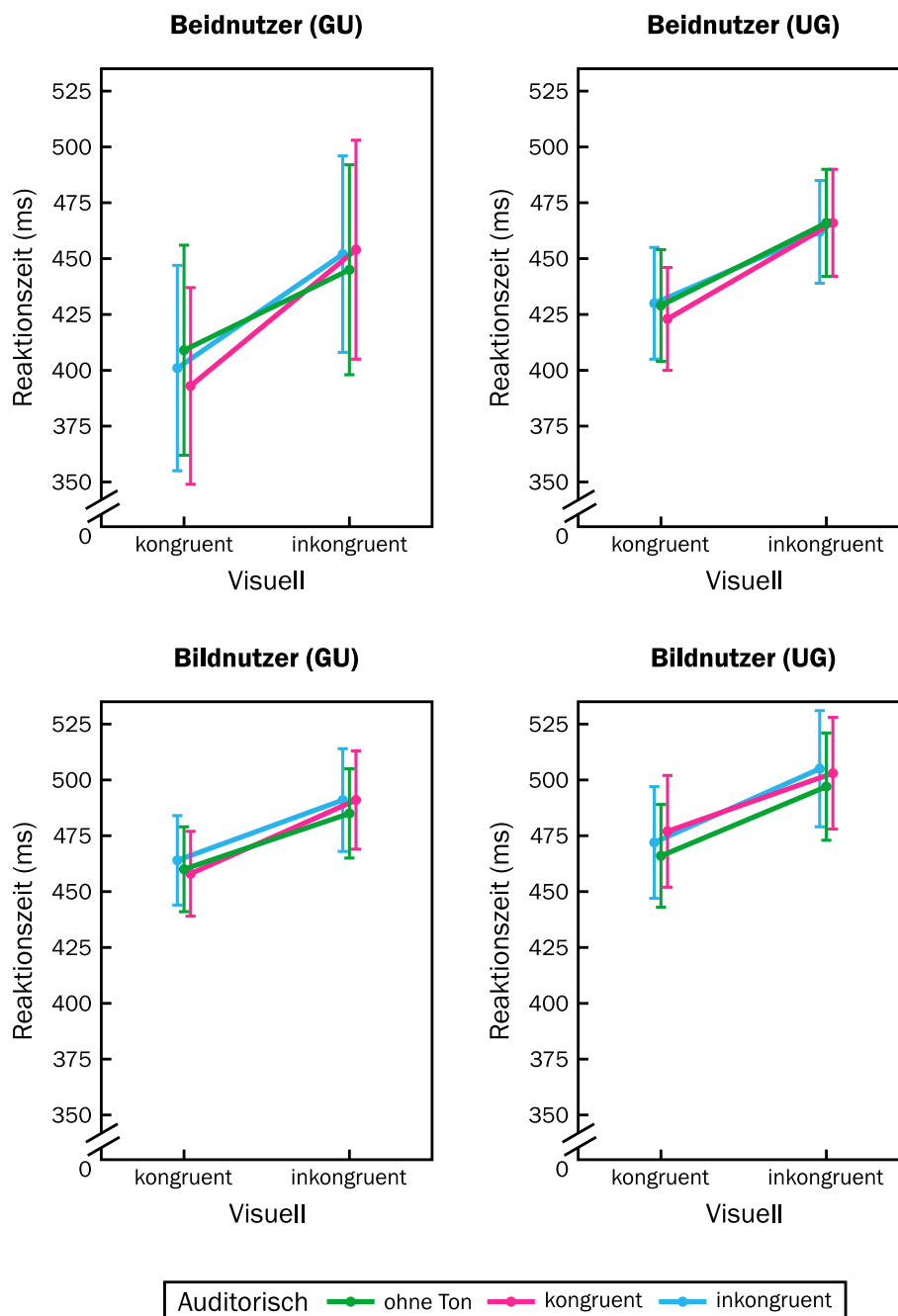


Abbildung 16. Reaktionszeit in Abhängigkeit vom visuellen Bahnverlauf nach der Verdeckung sowie von der auditorischen Stimulation für die Beidnutzer (oben) sowie die Bildnutzer (unten) in Experiment 3. Details zur Nutzergruppeneinteilung können dem Text entnommen werden. Dargestellt sind die Ergebnisse auf Basis einer Split-Half-Analyse: Es wurden entweder gerade Blöcke zum Kategorisieren und ungerade Blöcke zum Analysieren genutzt (linke Seite, GU) oder ungerade Blöcke zum Kategorisieren und gerade Blöcke zum Analysieren (rechte Seite, UG). Die Fehlerbalken kennzeichnen den Standardfehler des Mittelwerts. Zur Unterscheidung der Fehlerbalken sind die Datenpunkte leicht versetzt dargestellt.

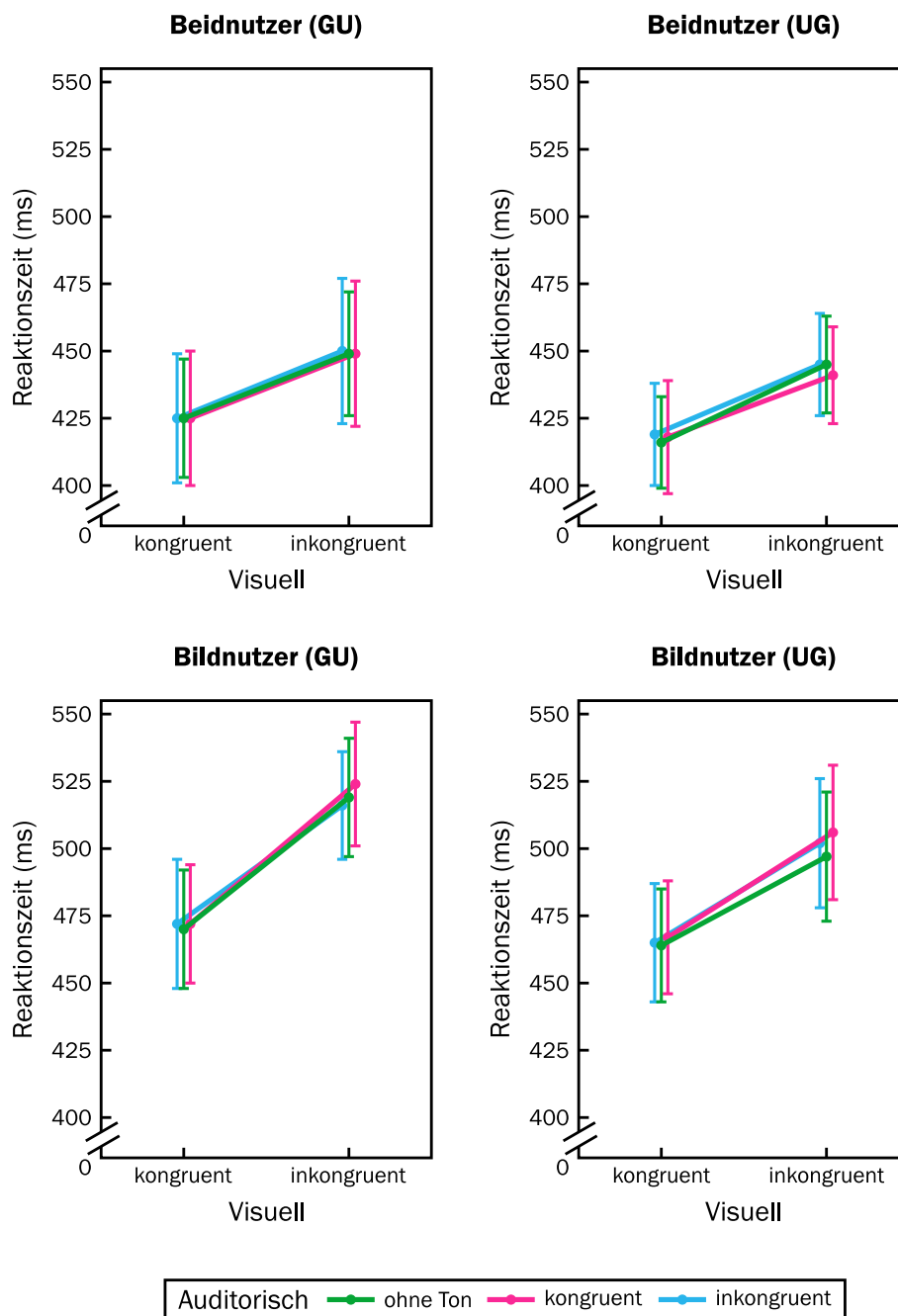


Abbildung 17. Reaktionszeit in Abhängigkeit vom visuellen Bahnverlauf nach der Verdeckung sowie von der auditorischen Stimulation für die Beidnutzer (oben) sowie die Bildnutzer (unten) in Experiment 4. Details zur Nutzergruppeneinteilung können dem Text entnommen werden. Dargestellt sind die Ergebnisse auf Basis einer Split-Half-Analyse: Es wurden entweder gerade Blöcke zum Kategorisieren und ungerade Blöcke zum Analysieren genutzt (linke Seite, GU) oder ungerade Blöcke zum Kategorisieren und gerade Blöcke zum Analysieren (rechte Seite, UG). Die Fehlerbalken kennzeichnen den Standardfehler des Mittelwerts. Zur Unterscheidung der Fehlerbalken sind die Datenpunkte leicht versetzt dargestellt.

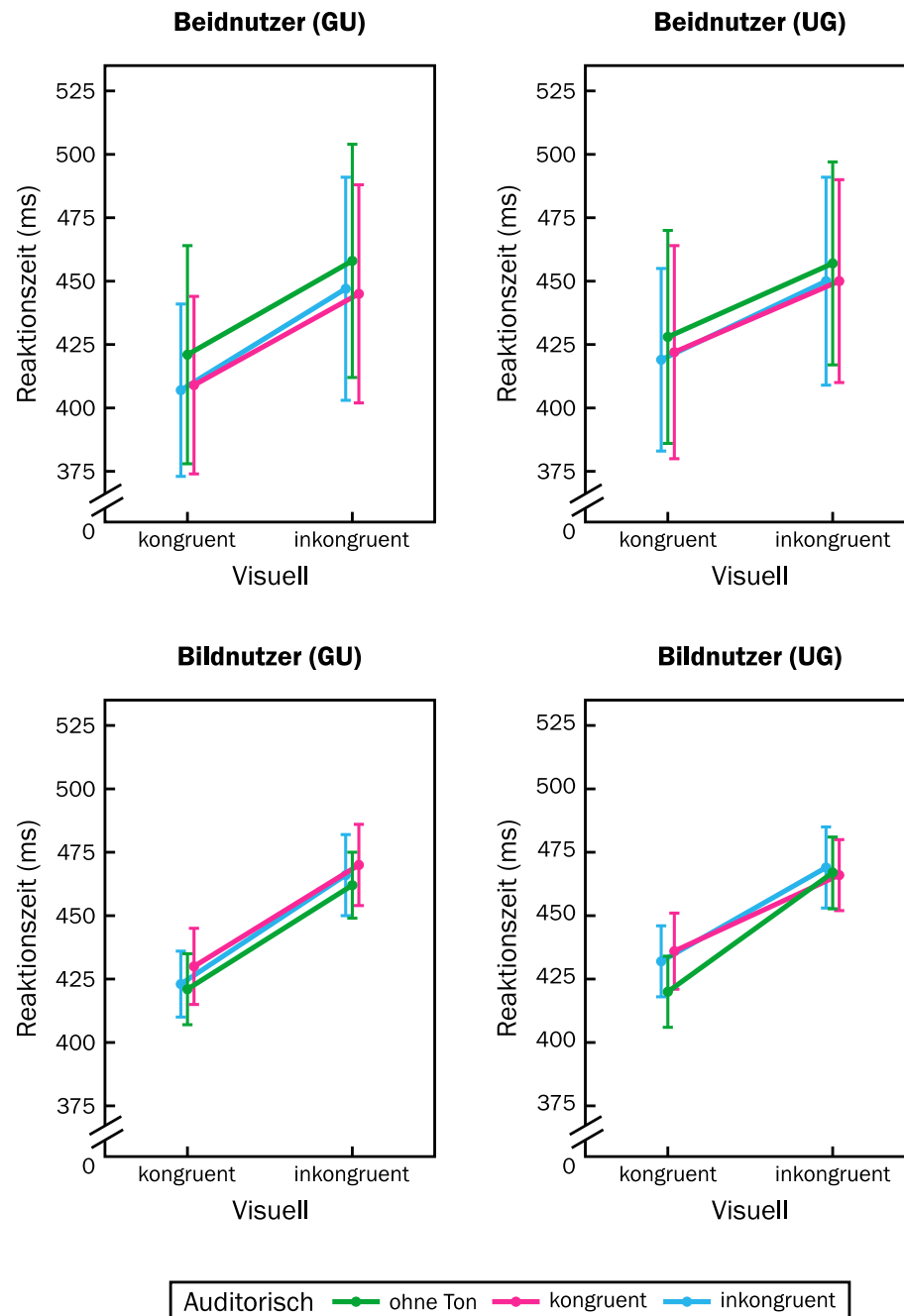


Abbildung 18. Reaktionszeit in Abhängigkeit vom visuellen Bahnverlauf nach der Verdeckung sowie von der auditorischen Stimulation für die Beidnutzer (oben) sowie die Bildnutzer (unten) in Experiment 5. Details zur Nutzergruppeneinteilung können dem Text entnommen werden. Dargestellt sind die Ergebnisse auf Basis einer Split-Half-Analyse: Es wurden entweder gerade Blöcke zum Kategorisieren und ungerade Blöcke zum Analysieren genutzt (linke Seite, GU) oder ungerade Blöcke zum Kategorisieren und gerade Blöcke zum Analysieren (rechte Seite, UG). Die Fehlerbalken kennzeichnen den Standardfehler des Mittelwerts. Zur Unterscheidung der Fehlerbalken sind die Datenpunkte leicht versetzt dargestellt.

6.4.3 Interindividuelle Unterschiede in der Stabilität der Informationsnutzung

Die Nutzergruppenanalyse brachte als einzigen stabilen Effekt für alle drei Experimente den Einfluss des visuellen Bahnverlaufs hervor. Der Einfluss auditorischer Stimulation zeigte sich hingegen als wenig kohärent. Dies erscheint vor allem vor dem Hintergrund erstaunlich, dass die Einteilung der Probanden in die Nutzergruppen einen auditorischen Einfluss hätte forcieren müssen (zumindest im Vergleich unimodal/auditorisch passend). Dies deutet darauf hin, dass die Probanden in der Mehrheit nicht konstant von auditorischen Informationen (nicht) profitieren. Daher wurde die Stabilität der Informationsnutzung in diesem Teil der Post-hoc-Analyse genauer betrachtet. Als Referenz wurden zudem die Ergebnisse für Experiment 2 (Extrapolation) herangezogen, da sich dort die Effekte als stabil erwiesen hatten.

Zunächst wurden die Einteilungen in die Nutzergruppen auf Basis gerader und ungerader Blöcke verglichen. In den Tabellen A9 bis A12 im Anhang ist für jeden Probanden in allen vier Experimenten aufgeführt, welche der beiden Erleichterungseffekte jeweils positiv bzw. negativ waren und welche Nutzergruppe sich daraus ergab. Im Anschluss wurde verglichen, ob für beide Kategorisierungsprozesse die Erleichterungseffekte das gleiche Vorzeichen aufwiesen und ob der Proband in die gleiche Nutzergruppe eingeteilt wurde oder nicht. Es ergaben sich für das Extrapolationsexperiment recht hohe Übereinstimmungsraten: 84.4 % der Probanden wurden in beiden Kategorisierungen gleich eingestuft. Auch die auditorischen Erleichterungseffekte wiesen für die nahe Distanz in 90.6 % und für die ferne Distanz in 93.8 % der Fälle für gerade und ungerade Blöcke das gleiche Vorzeichen und so auch die gleiche Richtung auf. Für die Diskriminationsexperimente waren die Übereinstimmungsraten wesentlich geringer: es wurden 57.5 % der Probanden in Experiment 3, 52.5 % in Experiment 4 und 70 % in Experiment 5 für gerade und ungerade Blöcke in die gleiche Gruppe eingeordnet. Dabei war die Übereinstimmung für die Ausprägung des visuellen Erleichterungseffekts in allen drei Experimenten größer als dies bei der auditorischen Reaktionszeit-Differenz der Fall war: Für die visuellen Differenzen ergaben sich 77.5 %, 87.5 % sowie 97.5 % Übereinstimmung im Vorzeichen für Experiment 3, 4 und 5. Beim auditorischen Erleichterungseffekt lagen diese hingegen bei 65 %, 60 % bzw. 70 %.

Die unterschiedlich hohen Übereinstimmungsraten für die visuellen und auditorischen Differenzen auf Basis gerader und ungerader Blockkategorisierung deuten darauf hin, dass der Einfluss der visuellen Stimulation stabiler zu sein scheint als der auditorische Einfluss. Um dies weiter explorativ zu begutachten, wurden im Folgenden die visuellen und auditorischen Erleichterungseffekte der drei Diskriminationsexperimente für jeweils alle Probanden blockweise auf Basis des Gesamtdatensatzes berechnet, um diese auf Stabilität zu untersuchen (siehe Tabellen D2 bis D4 im Anhang). Es fällt auf, dass die Probanden konstanter visuelle als auditorische Informationen (nicht) nutzen. Geht man von einer (Nicht-)Nutzungskonstanz aus, wenn $\frac{2}{3}$ bzw. $\frac{3}{4}$ der Blöcke eine visuelle bzw. auditorische Reaktionszeit-Differenz mit dem gleichen Vorzeichen aufweisen, ergibt sich folgendes Bild: Für die auditorischen Differenzen zeigten 9, 8 bzw. 9 Probanden eine $\frac{2}{3}$ -Konstanz für die Experimente 3, 4 bzw. 5. Bei der $\frac{3}{4}$ -Konstanz ergaben sich zusätzlich nur für Experiment 4 und 5 ein weiterer Proband bzw. 2 weitere Probanden. Für die visuellen Reaktionszeit-Differenzen fielen diese Werte hingegen

wesentlich höher aus: In den Experimenten 3, 4 und 5 ergaben sich 6, 14 bzw. 9 Probanden auf 2/3-Niveau und dazu 17, 13 und 17 weitere Probanden auf 3/4-Niveau Blöcke mit gleichem Vorzeichen. Nimmt man das stabile Extrapolationsexperiment 2 als Referenz (siehe Tabelle D1) ergeben sich für die auditorischen Differenzen beider Schätzdistanzen dort ebenso sehr hohe Probandenzahlen mit konstanten Differenzen. Hier zeigten sich für die nahe Distanz 7 bzw. 20 und für die ferne Distanz 6 bzw. 17 Probanden mit 2/3- bzw. 3/4-Konstanzen bei nur 32 insgesamt erhobenen Versuchspersonen.

6.4.4 Zwischenfazit

Die Bayes-Analyse der Gesamtdaten für die audiovisuellen Diskriminationsexperimente ergab für Experiment 3 und 5, dass das Modell mit dem Einfluss des visuellen Faktors favorisiert wird. In Experiment 4 wurde hingegen Evidenz für das Nullmodell gefunden. Dies deutet darauf hin, dass der visuelle Faktor in den beiden Experimenten mit der leicht erkennbaren Kugel die Diskrimination beeinflusste. Die Probanden konnten von visuell passenden Bahnen profitieren und schneller die Bewegungsbahn nach der Verdeckung angeben. In Experiment 4 mit der visuell schwierigeren Gestaltung ergab sich kein derartiger visueller Einfluss. Die Schwierigkeit in der visuellen Modalität konnte durch die Anpassung der Kugel an den Untergrund erhöht werden, die Probanden können sich weniger auf diese Information verlassen und daher übt der visuelle Faktor hier keinen Einfluss aus. Allerdings führte dies nicht dazu, dass die Versuchspersonen vermehrt auf die zweite Informationsquelle, die auditorischen Bewegungsinformationen, zurückgriffen. Für diese konnte kein Einfluss gefunden werden. Die Ergebnisse bestätigen daher die Dominanz visueller Informationen für die räumliche Domäne (Hidaka et al., 2015; Welch & Warren, 1980; Welch et al., 1986; Witten & Knudsen, 2005): Die Probanden verließen sich primär auf den visuellen Bahnverlauf als Informationsquelle, wenn sie ein räumliches Urteil treffen sollten. Auf zusätzliche Töne griffen sie nicht zurück, auch nicht, wenn die Reliabilität der visuellen Informationen gesenkt wurde. Dies steht nicht im Einklang mit vorherigen Studien, in denen die Probanden vermehrt Informationen einer Quelle nutzten, wenn eine zuvor mehr genutzte in ihrer Reliabilität herabgesetzt wurde (Bresciani & Ernst, 2007; Bresciani et al., 2008; Ernst & Banks, 2002; Helbig & Ernst, 2007; Helbig et al., 2012). Es ist vorstellbar, dass die Probanden die zusätzlich präsentierten Töne in der räumlichen Diskriminationsaufgabe nicht als relevante Quelle betrachteten, da diese keine Informationen über die Aufgabenstellung bereithielt. Die Töne transportierten zwar die Information, wann die Kugel nach der Verdeckung wiedererschien, enthielten aber keine Informationen über den visuellen Bahnverlauf, der in der Aufgabenstellung diskriminiert werden sollte.

Die Crossvalidierungsanalyse bestätigte die Ergebnisse der Bayes-Analyse größtenteils. Diese konnte zeigen, dass einzig der visuelle Faktor stabile Effekte für die Beid- und Bildnutzer hervorbrachte. Dies spricht dafür, dass der visuelle Erleichterungseffekt (Flombaum & Scholl, 2006; Kawachi & Gyoba, 2006) ein stabiles Phänomen darstellt, sofern Probanden visuelle Informationen nutzen. Dieser konnte mit der Einteilung der Probanden auch für Experiment 4 belegt werden. Das erscheint sinnvoll, da in Experiment 4 durch die herabgesetzte visuelle Reliabilität der visuelle Einfluss erst zum Vorschein tritt, wenn die Probanden entsprechend

ihres Nutzungsverhaltens eingeteilt werden und der visuelle Erleichterungseffekt in den Beid- und Bildnutzern dadurch forciert wird. Der Einfluss auditorischer Stimulation zeigte sich hingegen auch nach Nutzergruppeneinteilung als instabil. Weder für die Bildnutzer ergab sich ein stabiler negativer Einfluss der Töne, noch für die Beidnutzer ein stabiler positiver Effekt¹. Eine detaillierte blockweise Betrachtung der visuellen und auditorischen Erleichterungseffekte offenbarte, dass sich die Probanden nicht nur in ihrer Gruppenzugehörigkeit unterschieden, sondern vor allem in der Stabilität, mit der sie auf visuelle bzw. auditorische Informationen bei der Bewegungsvorhersage (nicht) zurückgriffen. Für den visuellen Erleichterungseffekt ergaben sich konstante (Nicht-)Nutzungsquoten über die Experimente hinweg, für den auditorischen Erleichterungseffekt waren diese wesentlich geringer. Erneut traten Unterschiede zwischen den Probanden auf: Während einige (wenige) recht konstant über das Experiment hinweg auditorische Informationen (nicht) nutzten, schwankte für andere die Nutzung stark, sodass sich am Ende weder ein eindeutiger Vor- noch Nachteil für die Toninformationen ergab. Bei den visuellen Informationen war dies nicht der Fall. In den meisten Fällen waren die Probanden hierbei konstant in ihrem Nutzungsverhalten über das Experiment hinweg, was sich in der Stabilität des visuellen Effekts für alle drei audiovisuellen Experimente in der Crossvalidierungsanalyse widerspiegelt.

Insgesamt deuten die Daten darauf hin, dass für die Bewegungsvorhersage in den durchgeführten Experimenten sowohl inter- als auch intraindividuelle Unterschiede bzw. Schwankungen vorliegen. Die Probanden unterscheiden sich zum einen darin, ob sie visuelle Informationen für die Bewegungsvorhersage nutzen, wobei dies für einen Großteil der Fall ist. Zum anderen liegen Unterschiede in der Stabilität der Nutzung auditorischer Informationen vor. Einige Probanden nutzen diese recht konstant (nicht), andere weisen starke intraindividuelle Fluktuationen in der Nutzung auf. Diese Unterschiede zwischen und innerhalb der Probanden werden in der Allgemeinen Diskussion (siehe Abschnitt 8.3) genauer betrachtet².

Außerdem ergaben sowohl die Bayes- als auch die Post-hoc-Crossvalidierungsanalyse erneut, wie auch die klassische Overall-Analyse, keinen stabilen Einfluss der Präsentations-

¹Die Stabilität des visuellen Einflusses wurde auch in einer ANOVA mit den drei Innersubjekt Faktoren und einem zusätzlichen Zwischensubjekt Faktor Experiment über alle drei Experimente hinweg bestätigt: In dieser wurde der visuelle Haupteffekt signifikant. Dass dieser in Experiment 4 und 5 in der Overall-Analyse nicht signifikant wurde, könnte mit einer größeren Variabilität begründet sein (vergleiche dazu auch die unterschiedlich großen Standardfehler in den Experimenten 3, 4 und 5 in den Abbildungen 11, 13 und 15). In die Instabilität des auditorischen Einflusses reiht sich dahingehend ein, dass auch der auditorische Faktor in der übergeordneten ANOVA für alle Experimente signifikant wurde, obwohl dieser Einfluss in den crossvalidierten Nutzergruppenanalysen nicht nachgewiesen werden konnte. Allerdings beachte man hierbei den Einfluss der Stichprobengröße, die in der übergeordneten ANOVA mit $n = 120$ sehr groß ausfiel und somit auch sehr kleine Unterschiede signifikant werden (in diesem Fall der post-hoc marginal signifikante Unterschied zwischen der unimodalen und der auditorisch passenden Bedingung von 3 ms). Die genauen statistischen Kennwerte der übergeordneten ANOVA sind in Tabelle B7 im Anhang zu finden.

²Da die Informationsnutzung über die Differenzen von Reaktionszeiten berechnet wurden, besteht die Möglichkeit, dass sich die Probanden verschiedener Nutzergruppen in ihren Antwortraten unterscheiden. So könnten z. B. die Nichtnutzer zwar keine schnellere Reaktion aufweisen, aber insgesamt eine höhere Accuracy erzielt haben, weil sie sich auf die Antwortgüte konzentrierten. Daher wurden die mittleren korrekten Antworten für alle Nutzergruppen in allen drei audiovisuellen Diskriminationsexperimenten für die beiden Analysedatensätze (Gerade vs. Ungerade, s. o.) verglichen. Es konnten keine stabilen signifikanten Unterschiede zwischen den Nutzergruppen für die drei Experimente gefunden werden. Zudem unterschied sich der Anteil korrekter Antworten jeweils nur wenig, die Raten schwankten für die drei Experimente zwischen den Nutzergruppen von 96.7% bis 99.5%. Die genauen Ergebnisse dieser Analysen sind in Anhang E zu finden.

dimension in den drei audiovisuellen Diskriminationsexperimenten. In Experiment 1 konnte ein solcher Zusammenhang gefunden werden, der visuelle Erleichterungseffekt fiel unter pseudo-3D-Stimulation größer aus als in echt-3D. Für die Extrapolationsaufgabe ergab sich ebenfalls kein Einfluss der Präsentationsdimension. Dies könnte darauf hindeuten, dass eine Aufgabenabhängigkeit für das Auftreten von Dimensionseffekten vorliegt. Die Ergebnisse der drei Diskriminationsexperimente deuten aber auch darauf hin, dass die auditorische Stimulation durch ihre instabilen Fluktuationseffekte ein Rauschen in den Daten erzeugt haben könnte und somit eine mögliche Störvariable für andere Effekte darstellt. Um den Einfluss der Tonstimulation auf die Bewegungsvorhersage zu eliminieren, wurde im Folgeexperiment auf auditorische Bewegungsinformationen verzichtet. Zudem wurde der Einfluss der Aufgabenstellung (Detektion vs. Diskrimination) innerhalb einer Stichprobe untersucht. Der Einfluss der Präsentationsdimension wird daher gemeinsam mit den Ergebnissen von Experiment 6 in der Allgemeinen Diskussion in einen Zusammenhang gesetzt (siehe Abschnitt 8.6).

7 Einfluss von Aufgabenstellung und visueller Gestaltung ohne Tonstimulation (Experiment 6)

Die Ergebnisse der audiovisuellen Diskriminationsexperimente zeigten, dass die Töne sowohl zwischen als auch innerhalb der Probanden nicht stabil als Informationsquelle genutzt werden. Daher sollte in diesem letzten Experiment überprüft werden, wie sich die visuelle Stimulation und die Präsentationsdimension auf das Verhalten der Probanden auswirken, wenn keine Töne als potentielle Störquelle gleichzeitig präsentiert werden. Dafür wurde in diesem Experiment eine Detektions- und Diskriminationsaufgabe für visuell passende und nicht passende Bahnen mit denselben Probanden durchgeführt. Beide Aufgaben wurden sowohl in echt-3D als auch einfach-2D präsentiert.

7.1 Methoden

An Experiment 6 nahmen 32 Versuchspersonen (je 16 männlich/weiblich) mit einem Durchschnittsalter von 22.97 Jahren ($SD = 2.94$) teil. Von diesen waren zwei links- und 30 rechtshändig. Alle Probanden absolvierten zwei Messtermine mit den beiden Aufgaben Detektion und Diskrimination. Zu beiden Terminen wurde die jeweilige Aufgabenstellung sowohl in einfach-2D als auch echt-3D präsentiert. Die Reihenfolge der beiden Präsentationsdimensionen wechselte zwischen beiden Terminen, sodass jeder Proband einmal mit einfach-2D und einmal mit echt-3D begann. Die Reihenfolge beider Aufgabenstellungen und die Reihenfolge der Präsentationsdimension sowie die Tastenzuweisung für die Diskriminationsaufgabe war gemeinsam mit dem Geschlecht der Probanden ausbalanciert.

Da in diesem Experiment keine Töne präsentiert wurden, reduzierte sich das Design auf die beiden Faktoren Visuell (passend vs. nicht passend) und Dimension (einfach-2D vs. echt-3D). Ergänzend kam der Faktor Aufgabe (Detektion vs. Diskrimination) hinzu, sodass sich insgesamt ein $2 \times 2 \times 2$ -Innersubjekt-Design ergab. Im Telexperiment Detektion reagierten die Probanden analog zu Experiment 1 mit Tastendruck sobald die Kugel nach der Brücke wiedererschien. Für die Diskrimination gaben sie wie in den Experimenten 3, 4 und 5 an, ob sich die Kugel auf einer visuell passenden oder nicht passenden Bahn weiterbewegte. Die weiteren Details der visuellen Stimulation entsprachen den bisherigen Experimenten und sind in der Allgemeinen Methode (siehe Abschnitt 4) bzw. für den einfach-2D-Teil im Methodenteil von Experiment 5 (siehe Abschnitt 6.3) zu finden.

7.2 Ergebnisse

Die Ergebnisse wurden mit Hilfe einer $2 \times 2 \times 2$ -ANOVA mit Messwiederholung mit den Faktoren Aufgabe, Dimension und Visuell ausgewertet. Es wurden alle drei Haupteffekte signifikant: Aufgabe, $F(1, 31) = 154.33$, $p < .001$; Dimension, $F(1, 31) = 18.87$, $p < .001$, und Visuell, $F(1, 31) = 35.41$, $p < .001$. Die Probanden reagierten schneller in der Detektionsaufgabe (267 ms) als bei der Diskrimination (483 ms). Sie waren ebenso schneller für visuell passende (365 ms) als nicht passende (384 ms) Bahnen. Außerdem fielen die Reaktionszeiten

bei echt-3D-Präsentation (362 ms) geringer als bei einfach-2D-Präsentation (387 ms). Des Weiteren ergab sich ein signifikanter Interaktionseffekt zwischen Aufgabe und visueller Stimulation, $F(1, 31) = 4.38, p = .045$. Die Interaktion zwischen visueller Stimulation und Präsentationsdimension wurde marginal signifikant, $F(1, 31) = 3.39, p = .075$. Die beiden weiteren Interaktionen wurden nicht signifikant, Aufgabe \times Dimension, $F(1, 31) = 1.97, p = .170$, sowie Aufgabe \times Dimension \times Visuell, $F(1, 31) = 1.01, p = .323$. Post-hoc zeigte sich für die Interaktion zwischen Aufgabe und Bahnverlauf ein größerer visueller Erleichterungseffekt für die Diskrimination (25 ms) als für die Detektion (13 ms), $t(31) = 2.07, p = .047$. Ebenso ergab sich eine Tendenz für eine größere Erleichterung während echt-3D-Präsentation (21 ms) als bei einfach-2D (15 ms), $t(31) = 1.91, p = .066$ (zweiseitig). Die Reaktionszeiten in Abhängigkeit von der visuellen Stimulation und der Präsentationsdimension sind für beiden Aufgabenstellungen in Abbildung 19 dargestellt.

Für die Diskrimination wurden zusätzlich die korrekten Antworten ausgewertet, um die Daten auf einen Speed-Accuracy-Trade-Off zu prüfen. Der Faktor Dimension sowie die Interaktion zwischen Dimension und Visuell wurden nicht signifikant, $F(1, 31) = 0.37, p = .545$ bzw. $F(1, 31) = 0.01, p = .962$. Für den visuellen Bahnverlauf ergab sich ein marginal signifikanter Einfluss, $F(1, 31) = 3.75, p = .062$. Die Versuchspersonen erzielten hierbei einen höheren Anteil korrekter Antworten für visuell passende Bahnverläufe (98.1 %) als für nicht passende (97.1%). Da die Probanden besser in Bedingungen abschnitten, in denen sie gleichzeitig schnellere Reaktionszeiten erzielten und zudem die korrekten Antworten nur geringfügig auf sehr hohem Niveau schwankten, ist nicht von einem Speed-Accuracy-Trade-Off auszugehen.

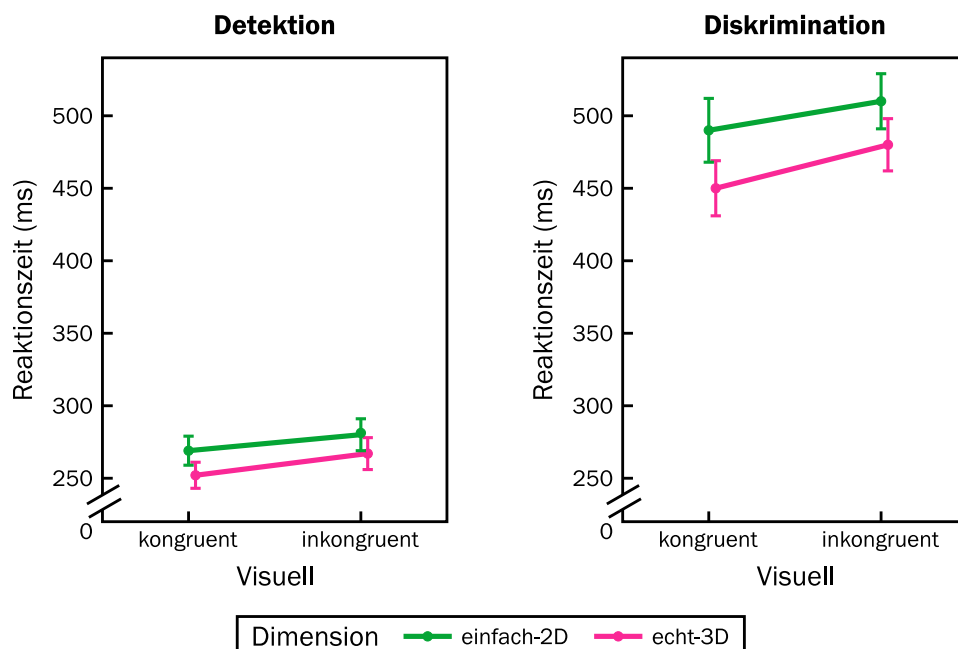


Abbildung 19. Reaktionszeit in Abhängigkeit vom visuellen Bahnverlauf nach der Verdeckung sowie von der Präsentationsdimension für Detektion (links) und Diskrimination (rechts). Die Fehlerbalken kennzeichnen den Standardfehler des Mittelwerts. Zur Unterscheidung der Fehlerbalken sind die Datenpunkte leicht versetzt dargestellt.

7.3 Diskussion

Die Probanden profitierten auch in diesem Experiment von visuell passenden Bahnen und reagierten entsprechend schneller als für inkongruente Bahnen. Dieser Erleichterungseffekt wurde außerdem sowohl von der Aufgabenstellung als auch der Präsentationsdimension moduliert. Er fiel während der Diskrimination sowie der echt-3D-Präsentation größer aus. Der Einfluss der Aufgabenstellung zeigte sich zudem in generell schnelleren Reaktionszeiten für die Detektion gegenüber der Diskrimination. Die Probanden reagierten außerdem insgesamt schneller während der echt-3D-Präsentation verglichen zur einfach-2D-Variante.

Die Ergebnisse bestätigen damit zunächst, dass die Detektion gegenüber der Diskrimination eine weniger anspruchsvolle Aufgabe darstellt, die entsprechend schneller absolviert werden kann (Bertelson & Boons, 1960). Dies könnte darauf zurückzuführen sein, dass die Diskrimination zusätzliche Aufmerksamkeitsprozesse gegenüber der Detektion benötigt (Luck & Hillyard, 1995). In diesem Kontext scheinen die Probanden zusätzlich unter erhöhten Ansprüchen in größerem Umfang von vorteilhaften Bedingungen zu profitieren. Dies äußert sich in einem stärker ausgeprägten Erleichterungseffekt für die Diskrimination, dort profitieren die Probanden stärker von visuell passenden Bahnen gegenüber nicht passenden Bahnen.

Dass die Probanden darüber hinaus schneller im echt-3D-Teil reagierten und der visuelle Erleichterungseffekt dort größer ausfiel, spricht für eine schnellere Verarbeitung in realitätsgetreuen Umgebungen und ist im Einklang mit Studien, die ebenfalls verbesserte Reaktionen bei 3D-Stimulation demonstrieren konnten (Conrad et al., 2013; Harrison et al., 2015; Senkowski et al., 2007). Die Ergebnisse widersprechen damit aber auch den Befunden aus Experiment 1, in welchem ein kleinerer visueller Erleichterungseffekt in der echt-3D-Umgebung gefunden wurde. Dies könnte auf die unterschiedliche Vergleichsbedingung zurückzuführen sein, da in Experiment 1 mit der pseudo-3D-Variante, in diesem Experiment mit der einfach-2D-Variante verglichen wurde. Zudem legen die gefundenen Dimensionseffekte in der Diskrimination hier nahe, dass die Töne in den audiovisuellen Diskriminationsexperimenten 3 bis 5 ein gewisses Rauschen in den Daten verursachten und somit etwaige Dimensionseffekte nicht gefunden werden konnten. Auf diese Zusammenhänge wird in einem übergreifenden Vergleich aller Befunde aus dem Bereich realitätsgetreues Stimulusmaterial in der Allgemeinen Diskussion (siehe Abschnitt 8.6) genauer eingegangen. Ebenso wird dort der Einfluss der Aufgabenstellung detaillierter betrachtet (siehe Abschnitt 8.4).

8 Allgemeine Diskussion

8.1 Zusammenfassung der wichtigsten Ergebnisse

Bei Verwendung von Aufgaben mit zeitlichem Fokus (Experiment 1 und 2, Detektion und Extrapolation) nutzten die Probanden mehrheitlich die zusätzlich präsentierten Töne für die Vorhersage, wobei die Bewegungsrichtung der Töne keine Rolle spielte. Zudem übte die Präsentationsdimension und damit zusätzliche Disparitätsinformationen einen Einfluss aus: Für die Detektion ergab sich ein größerer visueller Erleichterungseffekt für die pseudo-3D- als die echt-3D-Stimulation. Des Weiteren zeigten sich in Experiment 2 nach einer crossvalidierten Post-hoc-Betrachtung der Daten individuelle Unterschiede zwischen den Probanden in der Tonnutzung über die Zeit. Ein Großteil der Versuchspersonen konnte die Töne nur für jeweils eine Schätzdistanz zu ihrem Vorteil nutzen: Sie waren für diese besser in der Schätzung mit Ton, für die jeweils andere Distanz waren sie jedoch mit Ton schlechter in der Vorhersage, wann die Kugel die Zielposition erreichen würde.

In den drei räumlich orientierten Diskriminationsexperimenten (Experiment 3, 4 und 5) konnte die klassische Overall-Analyse für Experiment 3 und 5 Hinweise auf einen Vorteil passender Bahnen finden. Dieser wurde von einer nachträglichen Bayes-Analyse der Daten bestätigt. Da die weiteren Effekte wenig Stabilität aufwiesen, wurde ähnlich zu Experiment 2 post-hoc eine crossvalidierte Nutzergruppeneinteilung und -analyse durchgeführt. Diese offenbarte für alle drei Diskriminationsexperimente einen stabilen visuellen Erleichterungseffekt für Beid- und Bildnutzer, der einer Crossvalidierung standhielt. Für die weiteren Effekte ergaben sich keine kohärenten Einflüsse auf die Daten.

Für ein rein visuelles Experiment ohne Tonstimulation mit den Aufgaben Detektion und Diskrimination ergab sich erneut ein visueller Erleichterungseffekt für beide Aufgabenstellungen, wobei dieser in der Diskrimination größer ausfiel. Die Probanden waren zudem für die echt-3D-Präsentation schneller, ebenso war auch in dieser Präsentationsdimension der Erleichterungseffekt größer.

8.2 Einfluss audiovisueller Stimulation bei zeitlicher Bewegungsvorhersage

Informationsnutzung bei der Bewegungsvorhersage

Die Ergebnisse der durchgeführten zeitlichen Experimente bestätigen zunächst die Modalitäts-Angemessenheits-Hypothese (Welch & Warren, 1980; Welch et al., 1986) dahingehend, dass die Probanden für temporale Aufgaben verstärkt die zusätzlichen auditorischen Informationen in Form von Tönen nutzten. Während der Detektion (Experiment 1) und Extrapolation (Experiment 2) profitierten die Probanden von den Tönen während der Aufgabenbearbeitung (für Experiment 2 ersichtlich nach der Einteilung in Beidton-, Nahton- und Ferntonnutzer). Dies ist im Einklang mit zahlreichen Studien, in denen der Einfluss auditorischer Informationen in Aufgaben mit zeitlicher Relevanz belegt werden konnte (z. B. Aschersleben & Bertelson, 2003; Fendrich & Corballis, 2001; Guttman et al., 2005; Recanzone, 2003; Repp & Penel, 2002).

Töne stellen für Aufgaben zeitlicher Natur demnach hilfreiche Informationen dar, die von den Probanden entsprechend genutzt werden.

Allerdings fokussieren die Versuchspersonen nicht ausschließlich auditorische Informationen. Es ergab sich zusätzlich für die zeitliche Detektion ein visueller Erleichterungseffekt für räumlich passende Bahnen (Kawachi & Gyoba, 2006). Der Einfluss visueller Informationen ist somit nicht auf die räumliche Domäne begrenzt, sondern setzt sich auch in zeitlichen Aufgaben durch. Dies könnte auf die allgemeine Dominanz visueller Informationen, insbesondere bei der Bewegungswahrnehmung, zurückzuführen sein (Colavita, 1974; Hidaka et al., 2015; Koppen et al., 2009; Rummukainen & Mendonça, 2016; Soto-Faraco et al., 2003; Spence et al., 2012).

Fehlender Kongruenzeffekt zwischen visueller und auditorischer Stimulation

Auffällig ist, dass in beiden zeitlichen Experimenten kein Einfluss der Tonbewegungsrichtung auf die Leistung der Probanden gefunden wurde. Sowohl für die Detektion als auch die Extrapolation spielte es keine Rolle, ob sich der Ton in die gleiche oder die entgegengesetzte Richtung wie die visuelle Stimulation bewegte. Dies steht vermeintlich im Widerspruch mit Studien, in denen ein Kongruenzeffekt zwischen den Richtungen bewegter audiovisueller Stimuli mit einem Vorteil für gleichgerichtete Paarungen gefunden wurde (z. B. Conrad et al., 2010; Harrison et al., 2015; R. S. Kim et al., 2008; Meyer et al., 2005; Ogawa & Macaluso, 2013, aber siehe Alais & Burr, 2004a, die nur einen sehr kleinen Kongruenzeffekt fanden).

Eine Möglichkeit für den fehlenden Kongruenzeffekt liegt zunächst erneut in der Dominanz auditorischer Informationen für die zeitliche Domäne (z. B. Aschersleben & Bertelson, 2003; Fendrich & Corballis, 2001; Guttman et al., 2005; Recanzone, 2003; Repp & Penel, 2002). Diese könnte nicht nur für den allgemeinen Aufgabenfokus, sondern auch für spezifische Stimuluscharakteristika zutreffen. Die Probanden nutzen entsprechend für die Bearbeitung primär die zeitlichen Informationen der Tonstimuli (Spence, 2013). Die Information, wann die Kugel nach der Verdeckung wiederauftaucht, ist für beide Aufgabenstellungen ausreichend, genauere räumliche Toninformationen über den Bahnverlauf spielen eine untergeordnete Rolle. Für den Erleichterungseffekt hieße dies, dass die schnellere Reaktion bei raumzeitlicher Passung beim Wiederauftauchen primär über die räumlichen Informationen auf visueller Ebene moduliert wird. Zusätzliche räumliche Toninformationen werden hingegen anscheinend nicht genutzt.

Als Alternativ- bzw. Zusatzaspekt kommt das visuell dynamische Capture-Phänomen (Visual Dynamic Capture) in Frage. Bei diesem können Probanden in unisensorischen auditorischen Durchgängen die Bewegungsrichtung eines Tons korrekt bestimmen. Paart man den Ton hingegen mit einer visuellen Bewegung, dann nimmt die Angabe der Bewegungsrichtung für den auditorischen Stimulus in Richtung der visuellen Bewegungsrichtung zu (Soto-Faraco et al., 2002; Soto-Faraco, Spence & Kingstone, 2004; Soto-Faraco, Spence, Lloyd & Kingstone, 2004). Dieser Effekt tritt vor allem im audiovisuellen Binding-Fenster, in dem beide Modalitäten zeitlich eng beieinander präsentiert werden, vermehrt auf (Koppen & Spence, 2007). Übertragen auf die Experimente hier bedeutet dies, dass die Versuchspersonen in den bimodalen Durchgängen die Töne bei inkongruenten Bedingungen dennoch als kongruent wahrgenom-

men haben könnten, weil ein visueller Capture des Tons auftrat und dieser sich folglich in die gleiche Richtung wie die Kugel bewegte. Dies könnte durch die zeitlich synchrone Darbietung zusätzlich begünstigt worden sein. Dass der visuelle Capture für auditorische Bewegungen auch in der Tiefendimension auftritt (Alink et al., 2008; Kitagawa & Ichihara, 2002; Kitajima & Yamashita, 1999) und zudem für Looming-Stimuli größer ausfällt als für Receding-Stimuli (Harrison, 2012), unterstützt dies zusätzlich. Auch die Ergebnisse des durchgeführten Kontroll-experiments (siehe Abschnitt 5.3) legen diese Interpretation nahe. Dort wurden den Probanden die Töne unimodal präsentiert, die auch in der Detektion und Extrapolation verwendet wurden. Die Versuchspersonen waren in der Lage mit sehr hoher Genauigkeit nicht nur die Richtung der Töne anzugeben, sondern auch zu unterscheiden, ob es sich um einen Ton auf einer äußeren oder inneren Bewegungsbahn handelte.

Tonstimulation als potentielles Warnsignal (Alerting)

Der nicht vorhandene Kongruenzeffekt könnte jedoch auch darauf hindeuten, dass die Töne lediglich als Warnsignal wirkten. Die schnelleren Reaktionen in bimodalen Durchgängen kämen nicht durch multisensorische Erleichterung zustande, sondern durch eine einfache Cue-Wirkung der Tonstimuli. Die audiovisuelle Stimuluspaarung in den Experimenten war durch eine Passung zwischen Lautstärkenmodulation und Kugelverdeckung gekennzeichnet. Wenn die Probanden in Experiment 1 angaben, wann die Kugel nach der Brücke wiedererschien, dann fiel dieser Zeitpunkt mit der Erhöhung der Lautstärke zeitlich zusammen¹. In diesem Kontext könnte die plötzliche Änderung der Tonstimulation den Effekt durch eine Erhöhung der phasischen Aufmerksamkeit hervorrufen (Spence & Driver, 1997). In der Extrapolationsaufgabe bewegte sich die Kugel jedoch nach der Verdeckung für eine gewisse Distanz bis zur nahen bzw. fernen Stange weiter und auch unter diesen Bedingungen führten die Töne zu einer verbesserten Schätzleistung. Dies spricht dafür, dass ein simpler phasischer Alerting-Mechanismus ausgeschlossen werden und man eher von einer Erhöhung der tonischen Aufmerksamkeit durch die Tonstimulation ausgehen kann.

Interindividuelle Unterschiede in der Tonnutzung bei der Extrapolation

Die Probanden unterschieden sich im Einfluss der Töne auf die Schätzung, wann die Kugel die Stange erreicht. Nur wenige Versuchspersonen konnten die Töne vorteilhaft für beide Schätzdistanzen nutzen, ein Großteil profitierte nur für eine Entfernung von zusätzlichen Tönen. Für die zweite Distanz waren sie in bimodalen Durchgängen schlechter. Dies könnte einen Grund darstellen, warum in früheren Studien kein zusätzlicher Vorteil von Tönen auf die Bewegungsvorhersage gefunden werden konnte (z. B. DeLucia et al., 2016; Hassan, 2012; Hofbauer et al., 2004; Keshavarz et al., 2017; Schiff & Oldak, 1990; L. Zhou et al., 2007). In diesen Untersuchungen könnten die Effekte mit interindividuellen Unterschieden konfundiert

¹Die enge Kopplung von Ton- und Bildinformation wurde gewählt, um die Integrationswahrscheinlichkeit beider Modalitäten aufgrund zeitlicher Kongruenz zu erhöhen (z. B. Lewis & Noppeney, 2010; Shore et al., 2006; van Wassenhove et al., 2007).

sein und so potentielle positive Einflüsse von Tönen auf Gruppenniveau herausgemittelt werden.

Dass die Probanden nur für eine Entfernung von den Tönen profitieren konnten, deutet auf einen Einfluss des Zeitverlaufs auf die Schätzleistung hin. Die Nahtonnutzer konnten anscheinend die positive Tonwirkung nicht über einen längeren Zeitraum aufrecht erhalten, wohingegen die Ferntonnutzer erst zu einem späteren Zeitpunkt davon profitieren konnten. Eine Möglichkeit für diese Unterschiede könnte in verschiedenen Nutzungsraten audiovisueller Informationen liegen. In den Experimenten von Sun et al. (2017) unterschieden sich die Probanden in der Rate, mit der sie audiovisuelle Informationen in einer Angler-Aufgabe aus oszillierenden Fischen nutzen konnten. Bei der Schätzung im vorliegenden Experiment 2 ist ebenso vorstellbar, dass die Probanden sich darin unterscheiden, zu welchen Zeitpunkten sie audiovisuelle Informationen in welchem Umfang extrahieren und in Folge nutzen können. Es ist weiterhin möglich, dass die Unterschiede nicht in der Informationsextraktion, sondern deren Integration liegen. So unterschieden sich die Versuchspersonen bei Law et al. (1993) in ihrer Leistung für visuelle Time-to-Contact-Schätzungen, laut den Autoren aufgrund der verschieden ausgeprägten Fähigkeit unterschiedliche visuelle Informationsquellen zu integrieren. Für die audiovisuelle Extrapolation hier hieße dies, dass die Probanden durch unterschiedliche Integrationsfähigkeiten für visuelle und auditorische Informationen gekennzeichnet sind, welche zu verschiedenen Zeitpunkten der Vorhersage verschieden ausgeprägt sein könnten.

Als weitere Ursache kommen Unterschiede in der Strategienutzung in Frage. Bisherige Studien verdeutlichen, dass Probanden ihre Antwortstrategie innerhalb einer Temporal-Binding-Window-Aufgabe anpassen konnten, um ihre multisensorische Leistung zu optimieren (Mégevand et al., 2013). Außerdem nutzen Versuchspersonen nicht immer eine optimale Strategie, um eine Aufgabe, wie beispielsweise die Lokalisation eines audiovisuellen Stimulus, zu bearbeiten (Wozny et al., 2010). Bei der Schätzung, wann die Kugel die Stange erreichen würde, könnten sich die Probanden ebenfalls suboptimal verhalten haben, indem sie sich auf eine Schätzdistanz fokussierten und für diese die Töne nutzten, während sie bei der zweiten Distanz nicht mehr auf diese zurückgriffen. Durch die randomisierte Darbietung der Distanzbedingungen und folglich den stetigen Wechsel von naher und ferner Schätzung könnte dieses Verhalten zusätzlich provoziert worden sein. In einer Folgeuntersuchung mit blockweiser Variation der Schätzdistanz könnte dieser Zusammenhang genauer exploriert werden. Die Versuchspersonen könnten so stärker dazu tendieren, für ihre schlechtere bimodale Distanz ebenfalls Töne zu nutzen, da sie sonst für ein längeres Zeitintervall für ihre zunächst nicht priorisierte Schätzdistanz suboptimale Leistungen erzielen würden.

Des Weiteren ist denkbar, dass die Probanden unterschiedliche Vorstellungen entwickelten, wie sich die Kugel nach der Verdeckung weiterbewegen wird. Dies geht einher mit den Ergebnissen von Fulvio et al. (2015). Dort gab ein Teil der Probanden an, dass sich eine Kugel nach einer Bewegungsbahn ähnlich einer halben quadratischen Funktion linear weiterbewegen würde. Ein zweiter Teil hingegen urteilte, dass die Kugel im Anschluss ebenfalls einer quadratischen Bahn folgen würde. Beide Bewegungsarten waren in dem Fall theoretisch möglich, anscheinend war die Erfahrung der Probanden ausschlaggebend für ihr Urteil über den weite-

ren Bahnverlauf. Für die Schätzung der Ankunftszeit der Kugel hier ist vorstellbar, dass die Versuchspersonen z. B. aufgrund vorheriger (Lern-)Erfahrungen davon ausgingen, die Kugel würde in ihrer Bewegung abbremsen oder beschleunigen. Auf diese Weise konnten sie mit der zeitlichen Information der Töne nur für eine Schätzdistanz passende Ergebnisse erzielen. Im Einklang dazu konnten Makin et al. (2008, 2009) zeigen, dass sowohl die im Trial zuvor gesehene Geschwindigkeit als auch gelernte Eigenschafts-Geschwindigkeits-Assoziationen die Schätzung im aktuellen Durchgang beeinflussen können. Ebenso wirken sich Langzeitrepräsentationen wie z. B. Erwartungen aufgrund der Schwerkraft (Zago et al., 2010) oder die vertraute Größe von Objekten (Hosking & Crassini, 2010) auf Bewegungsvorhersagen aus.

8.3 Einfluss audiovisueller Stimulation bei räumlicher Diskrimination

Visuelle Informationen und interindividuelle Unterschiede

Vergleicht man die Ergebnisse der drei audiovisuellen Diskriminationsexperimente wird deutlich, dass die Nutzung visueller Informationen bei der räumlichen Bewegungsvorhersage im Vordergrund steht. Der Vorteil passender Bahnen zeigt sich in Experiment 3 und 5 bereits in der klassischen Overall-Analyse und wird auch von der Bayes-Analyse unterstützt. In Experiment 4, in dem die visuelle Reliabilität durch eine schlechtere Erkennbarkeit der Kugel reduziert wurde, konnte mithilfe dieser Analysen kein Hinweis auf einen visuellen Erleichterungseffekt gefunden werden. Damit deutet die visuelle Gestaltung in Experiment 4 auf eine erfolgreiche Reduzierung der Informationsgenauigkeit der visuellen Modalität hin. Die Probanden profitieren bei einer schlecht erkennbaren Kugel weniger von deren Bahnverlauf.

Jedoch ergaben sich bei der Post-hoc-Betrachtung der Daten in der Crossvalidierungsanalyse Unterschiede zwischen den Probanden in der Nutzung visueller Informationen. Dabei nutzte zwar ein Großteil der Probanden visuelle Informationen (25 bis 28 von 40 Probanden), allerdings ergab sich für die restlichen Probanden und so einem nicht geringen Anteil von 30 bis 37.5 % kein Vorteil für visuell kongruente Bahnverläufe. Die Verteilungen in den Nutzergruppen deuten auf Unterschiede zwischen den Probanden dahingehend hin, ob sie visuelle Informationen bei der Bewegungsvorhersage nutzen oder nicht. Die gefundenen Unterschiede weisen eine recht hohe Stabilität auf, die Übereinstimmung der Gruppeneinteilung aufgrund des visuellen Erleichterungseffekts lag durchschnittlich bei 87.5 %. Für die Probanden ergibt sich somit in den meisten Fällen ein recht konstanter visueller Erleichterungseffekt über das gesamte Experiment hinweg. Folglich scheinen die Daten auf recht stabile interindividuelle Unterschiede in der Nutzung visueller Informationen bei der räumlichen Bewegungsvorhersage hinzudeuten (siehe auch Abschnitt 8.5).

Auditorische Informationen und inter-/intraindividuelle Unterschiede

Für die Nutzung auditorischer Informationen in den audiovisuellen Diskriminationsexperimenten zeichnet sich ein komplexeres Bild ab. Zunächst ergab die klassische ANOVA-Auswertung für Experiment 3 einen hinderlichen Einfluss der Töne, für die weiteren beiden Experimente konnte kein signifikanter Einfluss der Töne gefunden werden. Dies wurde durch die Bayes-Analyse

bestätigt, die überdies für alle drei Experimente ein visuelles Modell bzw. das Nullmodell favorisierte. Erstaunlicherweise ergab sich auch in der Post-hoc-Auswertung bei Einteilung der Probanden gemäß ihres auditorischen Erleichterungseffekts kein stabiler Einfluss des auditorischen Faktors. Die Einteilung, ob ein Proband aufgrund eines positiven oder negativen auditorischen Erleichterungseffekts in die entsprechende Nutzergruppe eingeteilt wurde, gestaltete sich passend dazu sehr inkohärent. Die Übereinstimmungsrate (durchschnittlich 65 %) lag unter der für die Einteilung der visuellen Informationsnutzung. Insbesondere bei der Betrachtung der auditorischen Erleichterung über die Blöcke hinweg offenbarte sich eine zum Teil recht große Instabilität der Tonnutzung, auch innerhalb einzelner Probanden. Nur jeweils ungefähr ein Viertel der Probanden nutzte auditorische Informationen konstant innerhalb des gesamten Experiments (nicht).

Damit sprechen die Daten zum einen erneut wie bei der visuellen Informationsnutzung für interindividuelle Unterschiede in der Nutzung von zusätzlichen Tönen bei der Bewegungsvorhersage. Jedoch ergibt sich gleichzeitig durch die stark ausgeprägte Inkohärenz der Tonnutzung bei einem Großteil der Probanden klare Evidenz für intraindividuelle Unterschiede im Nutzungsverhalten. Auch bei Berger (2006) fanden sich Unterschiede zwischen Probanden in der Ausprägung des Inhibition-of-Return, welche für den einzelnen nicht über das gesamte Experiment hinweg stabil waren. Die Autorin folgerte daraus, dass die Varianz der Ergebnisse kein Indiz für konstante individuelle Unterschiede darstellt, sondern eher aufgrund von beispielsweise transienten Veränderungen in der Aufmerksamkeitsverteilung der Probanden zustande kommt.

Passend dazu konnten in anderen Studien Fluktuationen innerhalb der Probanden von Trial zu Trial für verschiedene Konstrukte gefunden werden, wie z. B. kognitive Flexibilität (Leber et al., 2008), Bewegungsdiskrimination (Sapir et al., 2005) und Aufmerksamkeit (Esterman et al., 2014; Leber, 2010; Rosenberg et al., 2015). Esterman et al. (2014) gehen von einem optimalen Aufmerksamkeitszustand aus, in dem eine effiziente Rekrutierung von Ressourcen ermöglicht wird und so genügend weitere Ressourcen für die Verarbeitung zusätzlicher, auch aufgabenirrelevanter, Merkmale zur Verfügung stehen. Es wäre denkbar, dass die Probanden in den audiovisuellen Diskriminationsexperimenten über das Experiment hinweg in ihrer Aufmerksamkeit schwankten und nicht in jedem Durchgang genügend Aufmerksamkeit für beide Informationsquellen zur Verfügung stand. Daher entschieden sie sich mehrheitlich dafür, sich primär auf die visuellen Informationen (gemäß der räumlichen Aufgabenstellung, s. o.) zu verlassen und konnten so nicht über die gesamte Experimentalzeit die Töne zusätzlich nutzen. Eine weitere Perspektive liefert die Untersuchung von Leber (2010). Bei ihm schwankten die Versuchspersonen ebenso in ihrem Aufmerksamkeitszustand, allerdings wurde hier der Fokus auf die Ablenkbarkeit innerhalb eines Experimentalteils gelegt. Die Probanden waren nicht in der Lage konstant irrelevante Reize zu ignorieren. Übertragen auf die Experimente hier ist vorstellbar, dass einige Probanden die Tonreize als nicht vorteilhaft ansahen und diese dementsprechend eher als Störreize empfanden, die sie eigentlich ignorieren wollten. Da die instabilen auditorischen Erleichterungseffekte sowohl positiv als auch negativ zwischen den

Probanden ausgeprägt waren, sind beide Erklärungsvarianten ferner gemeinsam möglich und könnten jeweils auf eine Subgruppe von Versuchspersonen zutreffen.

Auf einen anderen Ansatzpunkt weisen die Ergebnisse von Ribeiro et al. (2016) hin. Dort ging intraindividuelle Varianz mit Unterschieden in der Latenz der N1-Komponente im EEG während einer Wahlreaktionsaufgabe einher. Die Autoren folgerten, dass Schwankungen in der Leistung innerhalb eines Probanden bereits mit Fluktuationen in der sensorischen Verarbeitung der Stimuli einhergehen. Auch dies ist für die vorliegenden Ergebnisse denkbar: Die Probanden könnten sich bereits in der Wahrnehmung der auditorischen Stimuli von Trial zu Trial unterscheiden und der Einfluss der Töne sich entsprechend über das gesamte Experiment herausmitteln. Folgeuntersuchungen könnten zum einen durch eine Kombination mit dem EEG entsprechende ereigniskorrelierte Potentiale betrachten, um schwankende sensorische Verarbeitung genauer zu untersuchen. Zum anderen ist eine Kombination mit aufmerksamkeitsrelatierten Aufgabenmerkmalen denkbar, wie beispielsweise verschiedenen Distraktoren, um die genaue Verteilung der Aufmerksamkeit zwischen visuellen und auditorischen Informationen genauer zu untersuchen.

Auf konzeptueller Ebene lassen sich die gefundenen Unterschiede für die visuelle und auditorische Informationsnutzung zwischen und innerhalb der Probanden auf die in der differentiellen Psychologie verbreitete Latent-State-Trait-Theorie übertragen. Laut dieser lassen sich Merkmale einer Person in Traits und States unterscheiden. Traits sind hierbei Eigenschaften, die über die Zeit stabil und unabhängig von Zuständen oder Situationen sind. States hingegen bezeichnen Zustände, die zeitlich instabil und vom konkreten Zustand bzw. der Situation abhängig sind (Kelava & Schermelleh-Engel, 2012). Dies hieße für die vorliegenden Ergebnisse, dass die gefundenen Unterschiede im Nutzungsverhalten visueller Informationen eher Trait-Eigenschaften aufweisen. Die Nutzung auditorischer Informationen scheint hingegen eher State-Charakter zu besitzen. Auch für kognitive Phänomene wie z. B. das Arbeitsgedächtnis wurden die Annahmen der State-Trait-Theorie bereits zur Erklärung von Unterschieden in der Gedächtniskapazität herangezogen (Ilkowska & Engle, 2010). Ob auch die Informationsnutzung bei der Bewegungsvorhersage im Detail mit diesen Annahmen vereinbar ist, könnte in Folgeexperimenten untersucht werden.

Mögliche Ursachen für die instabilen auditorischen Effekte

In allen drei audiovisuellen Diskriminationsexperimenten traten unter Zusammenschau verschiedener Analysemethoden keine stabilen Toneinflüsse auf, die Probanden schwankten bei genauer Betrachtung zudem sehr häufig innerhalb eines Experiments in ihrer Nutzung auditorischer Informationen. Sogar eine Herabsetzung der visuellen Reliabilität in Experiment 4 führte nicht zu einem vermehrten Rückgriff auf auditorische Informationen. Daher stellt sich die Frage, warum die Probanden in diesen Experimenten die Töne zum größten Teil nicht (konstant) nutzten.

Zunächst kommt erneut die Domänenspezifität der visuellen und auditorischen Modalität in Frage (s. o.). Da in den ersten beiden zeitlichen Experimenten die Töne von den Probanden genutzt wurden, könnten die Effekte darauf zurückzuführen sein, dass den auditorischen

Informationen für die räumliche Diskrimination keine Relevanz eingeräumt wird. Hier böte sich eine Folgeuntersuchung mit einer zeitlichen Diskriminationsaufgabe an, bei der die Kugel zeitlich passend oder nicht passend nach der Verdeckung wiedererscheinen kann. Sollte sich hier ein Toneinfluss zeigen, spräche dies dafür, dass die Probanden in den durchgeführten Experimenten nur auf zeitliche Informationen der Töne zurückgreifen, wenn diese für die Aufgabenstellung von Bedeutung sind. Die Ergebnisse von Lukas et al. (2014) deuten auf einen solchen Zusammenhang hin. Diese konnten einen auditorischen Dominanzeffekt für eine multimodale Aufmerksamkeitsaufgabe finden, wenn die Probanden in einer zeitlichen Diskriminationsaufgabe die Stimulusdauer beurteilen mussten.

Dieser Aspekt lässt sich auf unterschiedlich ausgeprägte Informationsgewichtungen ausweiten. Informationen, die Probanden verstärkt nutzen, bekommen ein hohes internes Gewicht, (vermeintlich) weniger relevante Informationen ein niedrigeres Gewicht. Entsprechend entsteht ein Gesamtperzept, auf das die Versuchsperson reagiert. In den audiovisuellen Diskriminationsexperimenten könnte der Tonstimulus demgemäß eher niedrig gewichtet wurden sein, wohingegen die visuellen Informationen ein hohes Gewicht zugewiesen bekamen. Die Ergebnisse der vorliegenden Experimente sprechen dafür, dass keine allgemeingültigen Verteilungen für die Gewichtungen existieren, sondern die Probanden sich in dieser Zuweisung unterscheiden. Zudem ist denkbar, dass eine hohe Gewichtung mit einer stabilen Nutzung über die Zeit hinweg einhergeht, während ein niedrigeres Gewicht eher für eine inkohärente Nutzung sprechen könnte. Dass verschiedene Gewichtungen zugrundeliegender Prozesse Unterschiede zwischen den Probanden erklären können, konnte beispielsweise von Albrecht und Mattler (2016) bei der visuellen Metakontrast-Maskierung gezeigt werden. Auch für die audiovisuelle Integration unterscheiden sich Probanden in ihrer Gewichtung beider Modalitäten und entsprechend in ihrer Ausprägung audiovisueller Interaktionen (F. Zhou et al., 2007). In einer Untersuchung von Menciloglu et al. (2017) konnten Informationsgewichtungen zudem experimentell manipuliert werden. Durch zeitliche Erwartung erhöhte sich bei ihnen die Gewichtung visueller Informationen gegenüber auditorischen Informationen. Dies hieße, die Probanden in den vorliegenden Experimenten könnten beispielsweise durch das über das Experiment hinweg gleichbleibende Verdeckungsintervall eine gewisse zeitliche Erwartung an das Wiederauftauchen der Kugel entwickelt haben und entsprechend wurden die visuellen Informationen höher gewichtet, wohingegen Töne als weniger relevant eingeordnet wurden.

Des Weiteren ist ein größeres Gewicht visueller Informationen nicht nur in räumlichen Aufgaben, sondern generell bei unterschiedlichen Aspekten der Bewegungswahrnehmung denkbar. So kam es in einer Untersuchung zur Bewegungswahrnehmung bei Tischtennisspielern zu verschiedenen Informationseinflüssen je nach beurteilter Bewegungseigenschaft (S. H. Park et al., 2016). Die Probanden sahen Videos von Tischtennisspielern bei der Angabe ein Balles. Sollte die Bewegungsrichtung des Balles kategorisiert werden, profitierten die Probanden von einer größeren Menge visueller Informationen, zusätzliche auditorische Informationen führten zu keiner Verbesserung. Für die Vorhersage der Rotationsbewegung des Balls hingegen verhielt es sich umgekehrt: Toninformationen verbesserten die Leistung, weitere visuelle Informationen hingegen nicht. Entsprechend wäre ein Vorteil visueller Informationen

nicht per se für die räumlichen Aufgabenaspekte, sondern für das konkrete Urteil über den räumlichen Bahnverlauf denkbar und eine damit einhergehende Höhergewichtung visueller Informationen.

Darüber hinaus kann eine nicht stattgefundene Integration beider Modalitäten bei den Probanden eine Ursache darstellen. Zwar wurden Töne und Bildinformationen zeitlich synchron und im Falle kongruenter Durchgänge auch räumlich gekoppelt präsentiert, was die multisensorische Integration hätte begünstigen sollen (Meredith & Stein, 1986a; Meredith et al., 1987; Spence, 2011), allerdings spielen darüber hinaus weitere Faktoren für eine erfolgreiche multisensorische Integration eine Rolle. Einer davon ist die semantische Kongruenz zwischen den Modalitäten (Chen & Spence, 2017; Spence, 2007). Bereits für einfache Farbdiskriminationen zeigt sich eine audiovisuelle Erleichterung, wenn für den Farbreiz semantisch passend das entsprechende Wort eingespielt wird (Laurienti et al., 2004). Ähnliche Effekte finden sich bei Reaktionen auf semantisch passend präsentierte Tierstimuli bzw. Alltagsobjekte (z. B. Bild von einem Hund und Bellgeräusch oder Telefon und Klingelgeräusch, Molholm et al., 2004; Sued et al., 2009). Bei der Betrachtung von Stepptanzsequenzen ergeben sich audiovisuelle Interaktionen ebenfalls nur, wenn Ton- und Bildsignale synchron präsentiert werden und eine sinnhafte Integration beider Modalitäten gemäß dem natürlichen Perzept möglich ist (Arrighi et al., 2009). Ebenso werden Point-Light-Walker besser detektiert, wenn gleichzeitig Schrittgeräusche gegenüber einfachen Tonstimuli präsentiert werden. Auch bei Stream-and-Bounce-Experimenten erzeugt der Sound einer Billardkugel beim Zusammentreffen der beiden visuellen Stimuli mehr Bounce-Antworten als gleichzeitige Wassertropfen- oder Feuerwerkgeräusche (Grassi & Casco, 2010). Gleichfalls erleichtern bei der Identifikation von (maskierten) Bildern semantisch passende Töne die Leistung gegenüber einer White-Noise-Bedingung (Chen & Spence, 2010). Eine ähnliche Überlegenheit von Tönen gegenüber White-Noise-Bedingungen konnte auch Neuhoff (2001) für Looming- bzw. Receding-Stimuli finden. Da in den Experimenten hier eine visuell sehr realitätsnahe Kugelbewegung mit einer zwar räumlich modulierten, aber doch einfachen White-Noise-Bewegung kombiniert wurde, könnte dies die Integration beider Signale aufgrund mangelnder semantischer Kongruenz erschwert haben. In Folgeexperimenten könnte dies mit passenderen auditorischen Kugelbewegungen untersucht werden, indem den Probanden als Tonstimulus z. B. ein Rollgeräusch präsentiert wird.

Folgt man diesen Überlegungen stellt sich allerdings die Frage, warum in Experiment 5 für den einfach-2D-Teil keine Toneinflüsse gefunden werden konnten. In diesem war die visuelle Gestaltung durch den Verzicht auf Tiefenhinweise stark vereinfacht, wodurch eine stärkere Passung zwischen White-Noise-Tonbewegung und der Bewegung des Kreises über den Bildschirm entstehen könnte. Hierbei ist zu bedenken, dass die Kugelbewegung durch den Raum auf den Betrachter zu durch Wegfall der Tiefenhinweise in der einfach-2D-Präsentation zu einer Kreisbewegung wurde, die sich von oben nach unten über den Bildschirm bewegte. Die präsentierten Töne bewegten sich jedoch weiterhin auf den Betrachter zu. Daher könnte hier aufgrund der nicht vorhandenen räumlichen Kongruenz ebenfalls keine Integration beider Modalitäten stattgefunden haben (Meredith & Stein, 1986a; Spence, 2011).

Im Fall einer Nicht-Integration der visuellen Kugelbewegung und der auditorischen White-Noise-Bewegung könnte der auditorische Stimulus vermehrt als störend empfunden worden sein. Als Folge fühlten sich die Probanden von diesem in der Aufgabe abgelenkt, was sich ebenso in einer schlechteren visuellen Leistung zeigt und somit die kleineren visuellen Effekte bzw. die interindividuellen Unterschiede in der visuellen Informationsnutzung in der audiovisuellen Diskrimination erklären würde. Die Probanden könnten sich, wie in bisherigen Studien gezeigt, von auditorischen Distraktoren in der visuellen Aufgabe ablenken lassen (Berti & Schröger, 2003; Tellinghuisen & Nowak, 2003). Zudem unterscheiden sich Probanden in ihrer Ablenkbarkeit durch Tonreize (Fukuda & Vogel, 2011; Macken et al., 2009; Sörqvist & Rönneberg, 2014), was sich ebenso auf die Ergebnisse hier ausgewirkt haben und deren Inkohärenz zwischen und innerhalb der Probanden erklären könnte.

Die fehlende Aufmerksamkeit für die auditorische Modalität, sei es durch die oben erläuterten niedrigeren Gewichtungen oder die Wahrnehmung als Störquelle, könnte sich ebenfalls hinderlich auf die multisensorische Integration auswirken. In einigen Studien waren crossmodale Interaktionen davon abhängig, dass die Probanden ihre Aufmerksamkeit auf beide Modalitäten richteten (Mozolic et al., 2008; Talsma et al., 2007). Übertragen auf die Experimente in der vorliegenden Arbeit hieße dies, dass keine audiovisuellen Effekte auftreten konnten, weil die Probanden vermehrt eventuell nur eine Modalität, in diesem Fall die visuelle Kugelbewegung, fokussierten. Allerdings sprechen die Befunde anderer Untersuchungen gegen einen Einfluss der Aufmerksamkeitsausrichtung in dieser Form. So kam es in der Studie von Busse et al. (2005) auch bei nicht attendierten und aufgabenirrelevanten auditorischen Reizen bei gleichzeitiger Präsentation dennoch zu crossmodalen Interaktionen in Form von verbesserter visueller Detektion, die sich auch in den abgeleiteten neuronalen Korrelaten im EEG und fMRT widerspiegelte. Diese Effekte konnten von Starke et al. (2017) auch für den umgekehrten Zusammenhang, also den Einfluss irrelevanter und nicht attendierter visueller Stimuli auf die Detektion von auditorischen Reizen bestätigt werden. Auch in einer Studie zur Größendiskrimination mit visuellen und haptischen Reizen erfolgte eine automatische Integration beider Informationsquellen, auch wenn keine Aufmerksamkeit auf der visuellen Dimension lag (Helbig & Ernst, 2008). Modalitätsspezifische Aufmerksamkeit ist entsprechend diesen Ergebnissen für crossmodale Interaktionen nicht notwendig. Der genaue Einfluss von Aufmerksamkeit in crossmodalen Untersuchungen wird derzeit diskutiert (siehe dazu auch Macaluso et al., 2016) und sollte daher in zukünftigen Untersuchungen genauer exploriert werden.

Des Weiteren ist es möglich, dass die in den Experimenten verwendeten visuellen und auditorischen Bewegungen zu überschwellig präsentiert wurden. Gemäß dem Prinzip der inversen Effektivität (Meredith & Stein, 1986b; Stanford et al., 2005; B. E. Stein & Stanford, 2008) fällt der multisensorische Effekt am größten aus, je unerschwelliger die Einzelreize präsentiert werden, sodass der potentielle Nutzen einer weiteren Modalität zur korrekten Reaktion möglichst groß ausfällt. Entsprechend könnte die eindeutige visuelle Kugelbewegung hier ausreichen, um eine Entscheidung über den Bahnverlauf fällen zu können und zusätzliche Töne erbringen keinen weiteren Nutzen. Passend dazu ergab sich auch bei Hong und Shim

(2016) keine multisensorische Erleichterung bei Verwendung kontrastreicher Einzelstimuli, wohingegen die Detektion kontrastarmer visueller Stimuli durch gleichzeitige Töne verbessert werden konnte. Ebenso profitierten die Probanden bei von Saldern und Noppeney (2013) von audiovisuellen Stimuli vor allem für reduzierte, weniger reliable Einzelsignale.

Der Vollständigkeit halber sei als weitere mögliche Erklärung für die Nichtnutzung von Tönen der Präsentationsort angemerkt. Die Diskrimination des Bahnverlaufs erfolgte zentral im visuellen Gesichtsfeld. In den Experimenten von Schmiedchen et al. (2012) ergab sich für die zentrale Präsentation audiovisueller Bewegungsreize nur ein Einfluss von visuellen Reizen auf auditorische. Eine umgekehrte Beeinflussung von Tönen auf visuelle Reize fanden sie nur in der Peripherie. Passend dazu konnten Falchier et al. (2002) im Primaten eine exzentritätsabhängige Konnektivität zwischen dem primären visuellen Kortex und auditorischen Arealen finden, wobei diese in der Peripherie größer ausfiel. Gegebenenfalls ist eine stärkere Tonnutzung bei in der Peripherie präsentierten Bewegungsbahnen möglich.

8.4 Einfluss der Aufgabenstellung

Wie zu Beginn der Diskussion bereits ausgeführt, bestätigen die Ergebnisse die Modalitäts-Angemessenheits-Hypothese (Welch & Warren, 1980; Welch et al., 1986). Die Probanden nutzen für temporale Aufgaben die zusätzlichen auditorischen Informationen, die in der Detektion und Extrapolation hilfreiche Hinweise darstellten. Eine semantische Passung zwischen visuellen und auditorischen Bewegungsinformationen ist für diese Aufgaben anscheinend nicht notwendig, einfache Tonstimuli mit entsprechenden zeitlichen Hinweiscues sind ausreichend, damit die Probanden auf diese zurückgreifen.

Die vorliegenden Experimente konnten hingegen nicht bestätigen, dass in räumlichen Aufgaben die visuelle Modalität fokussiert wird. Zwar nutzten in allen Diskriminationsexperimenten ungefähr 2/3 der Probanden Bildinformationen, der relative Anteil ist damit aber geringer als der Tonnutzungsanteil für die temporalen Aufgaben. Zudem ergab sich auch in der zeitlichen Detektionsaufgabe ein deutlicher visueller Erleichterungseffekt für räumlich passende Bahnen, obwohl die räumliche Domäne für eine zeitliche Aufgabe keinen entscheidenden Einfluss ausüben sollte. Daher kann die Modalitäts-Angemessenheits-Hypothese in keinem Fall als absolut, sondern sollte vielmehr relativ betrachtet werden: Die meisten Probanden nutzen die für die Aufgabenstellung passendste Modalität, allerdings scheint dies verstärkt für auditorische Informationen der Fall zu sein. Diese werden in temporalen Aufgaben vermehrt genutzt, spielen aber für räumliche Aufgaben keine Rolle. Damit sprechen die Ergebnisse der hier durchgeführten Experimente eher für eine gewisse generelle visuelle Dominanz (Barrett & Newell, 2015); ggf. insbesondere bei der Bewegungswahrnehmung (Colavita, 1974; Rumukainen & Mendonça, 2016; Soto-Faraco et al., 2003), da sich diese Informationen auch in zeitlichen Aufgaben durchsetzen. Des Weiteren ergibt sich für räumliche Aufgaben zudem eine gewisse Abhängigkeit der Informationsnutzung vom jeweiligen Individuum. Die Probanden schwanken in ihrer Nutzung der beiden zur Verfügung stehenden Informationen. Insbesondere die semantische Passung zwischen Bild- und Tonbewegung könnte hier eine entscheiden-

de Voraussetzung für eine konstante Informationsnutzung darstellen (s. o.). Diese Aspekte schränken die Modalitäts-Angemessenheits-Hypothese zusätzlich in ihrer Gültigkeit ein.

Außerdem besteht die Möglichkeit, dass die Aufgabenstellung per se und nicht die dort relevante Modalität den maßgeblich beeinflussenden Faktor darstellt. Wie Gleiss und Kayser (2013) demonstrierten, könnte ein Einfluss von Tönen nur bei der Detektion, nicht aber bei der Diskrimination visueller Stimuli vorliegen. Auch in den vorliegenden Experimenten konnten (positive) Effekte zusätzlicher Töne für die Detektion, nicht aber die Diskrimination gefunden werden. Da die Töne sich auch in der Extrapolation auswirkten, müsste die Aufgabenabhängigkeit des Toneinflusses zumindest auf diese weitere Aufgabenstellung ausgeweitet werden. Auch das Temporal Orienting tritt primär in Detektionsexperimenten auf. Für einen Effekt während der Diskrimination sind bestimmte Bedingungen nötig wie z. B. eine geblockte Darbietung und die Verwendung intuitiv interpretierbarer Cues (Correa et al., 2004). Für das Auftreten eines Inhibition-of-Return in einer Diskriminationsaufgabe müssen ebenfalls bestimmte Voraussetzungen erfüllt sein, z. B. eine Cue-Target-Divergenz oder das Nutzen verdeckter Aufmerksamkeitsverlagerung (Pratt & Abrams, 1999). Die Knüpfung an bestimmte Bedingungen weist auf eine Relevanz bestimmter Aufgabenaspekte bzw. die Interaktion mit weiteren Merkmalen anstelle eines alleinigen Aufgabeneinflusses hin. Dies wurde von Beck et al. (2014) für räumliche und zeitliche Erwartungen in Detektions- und Diskriminationsaufgaben bestätigt. Es kam zu einem größeren Effekt der zeitlichen Erwartung während der Detektion sowie einer Interaktion beider Erwartungstypen. Für die Diskrimination hingegen wurde ein größerer Effekt der räumlichen Erwartung gefunden, räumliche und zeitliche Erwartung interagierten nicht miteinander. Im Einklang konnte in den hier durchgeführten Untersuchungen im ersten Experiment ein Einfluss zeitlicher Informationen (transportiert durch die Töne) bei der Detektion gefunden werden, wohingegen diese für die räumliche Diskrimination keine Rolle spielten.

Des Weiteren ist eine Abhängigkeit multisensorischer Einflüsse von der Aufgabenstellung denkbar. In einer Studie von Jaekl und Harris (2009) konnten zusätzliche Töne die Detektion von Gabor-Patches verbessern. Sollte deren Richtung diskriminiert werden, dann war die Leistung bei gleichzeitiger Tonestimulation schlechter. Dies geht einher mit den Ergebnissen von Sinnott et al. (2008). Sie zeigten, dass für das gleiche experimentelle Display je nach Aufgabenstellung multisensorische Erleichterungs- oder Inhibitionseffekte auftreten können. Auch in Experiment 1 und 3 dieser Arbeit wurde exakt das gleiche visuelle und auditorische Stimulationsmaterial genutzt, nur die Aufgabenstellung variierte. Bei der Detektion traten positive Effekte durch einen gleichzeitig präsentierten Ton auf, die Probanden reagierten mit diesem schneller. Bei der Diskrimination war dies hingegen nicht der Fall, der Ton wirkte sich nicht positiv auf die Leistung aus. Die Ergebnisse widersprechen jedoch den Befunden von D. Hecht et al. (2008), bei denen der multisensorische Gewinn in Diskriminationsaufgaben mit verschiedenen Kombinationen visueller, auditorischer und haptischer Stimuli größer ausfiel als in Detektionsaufgaben. Die Probanden sollten in den Experimenten dort entweder ein Signal detektieren oder eine Entscheidung über dessen Position treffen. Es besteht die Möglichkeit, dass bei konkreten räumlichen Aufgaben, wie einer Positionsbestimmung zusätzliche Töne die

Antwort erleichtern, wohingegen dies bei einer eher komplexen und dynamischen räumlichen Aufgabe, wie der Entscheidung über einen passenden Bahnverlauf, nicht der Fall ist.

Die vorgestellten Studien und die Ergebnisse der durchgeführten Experimente legen die Vermutung nahe, dass ein Toneinfluss bei der Bewegungsvorhersage ebenfalls nur für bestimmte Aufgabenstellungen, wie der zeitlichen Detektion bzw. Extrapolation auftritt. Für andere Aufgaben, wie die räumliche Diskrimination könnte er an bestimmte Bedingungen geknüpft sein, ggf. sind die semantische Passung zwischen beiden Modalitäten und eine konkretere räumliche Aufgabe (s. o.) entscheidende Faktoren, dass auch in diesen Aufgaben ein Toneinfluss auftritt. In Folgeuntersuchungen sollten diese Bedingungen genauer bestimmt werden.

Ein weiterer Einflussfaktor könnte in individuellen Unterschieden zwischen den Probanden liegen. Gerade die verschieden ausgeprägten visuellen Erleichterungseffekte und auch die wenigen Probanden, die in der Diskrimination stabil auditorische Informationen (nicht) nutzten, deuten darauf hin, dass hierfür Unterschiede vorliegen können. Bei Bridwell et al. (2013) unterschieden sich die Probanden in ihren Aufmerksamkeitsverteilungsstrategien je nachdem, ob sie eine Detektions- oder Diskriminationsaufgabe durchführen mussten. Nur wenige Probanden waren in der Lage ihre Aufmerksamkeit optimal für beide Aufgaben zu verteilen. Entsprechend könnte die Detektion hier eher eine von vielen Probanden verwendete Aufmerksamkeitsverteilung forciert haben, wohingegen die Diskrimination stärkere Unterschiede zwischen den Probanden zuließ, was sich in den verschieden ausgeprägten Erleichterungseffekten widerspiegelt.

Für Experiment 6 wurden keine Töne verwendet, entsprechend konnte in diesem Experiment das Verhalten der Probanden ohne die potentielle Störquelle Tonstimulation untersucht werden. Hier zeigte sich zunächst das erwartete Ergebnis, dass die Versuchspersonen in der Detektion schneller reagieren als in der Diskrimination (Bertelson & Boons, 1960). Darüber hinaus war der visuelle Erleichterungseffekt für die Diskrimination stärker ausgeprägt als in der Detektion. Die Probanden profitieren also in der anspruchsvolleren Aufgabe in größerem Ausmaß von visuell kongruenten Bahnen. Dies geht eher mit den Ergebnissen von D. Hecht et al. (2008) einher, die größere Effekte für die Diskrimination als die Detektion finden konnten. Auch Cueing-Effekte fallen für die Diskrimination größer aus als in der Detektion, auch bei identischem Stimulusmaterial (Brawn & Snowden, 2000). Diese Befunde können mit den Ergebnissen von Experiment 6 auf den visuellen Erleichterungseffekt ausgeweitet werden, der ebenfalls bei identischem Stimulusmaterial bei der Diskrimination größer ausfällt. Dass dieser Effekt bei gleichzeitiger Tonstimulation nicht gefunden werden konnte (Vergleich von Experiment 1 und 3), könnte daran liegen, dass die visuelle und auditorische Modalität auf gemeinsame Aufmerksamkeitsressourcen zurückgreifen, insbesondere, wenn räumliche Aufgabenaspekte im Vordergrund stehen (Wahn & König, 2017). Zusätzlich konnte gezeigt werden, dass die Detektion auditorischer Stimuli verringert ist, wenn gleichzeitig eine visuelle Aufgabe mit hoher Belastung durchgeführt wird (Macdonald & Lavie, 2011; Raveh & Lavie, 2015), wobei dieser Prozess von geringer ausgeprägten auditorisch evozierten Potentialen für die gleichzeitig präsentierten Töne begleitet wird (Molloy et al., 2015; Otten et al., 2000).

Demnach könnten in den audiovisuellen Diskriminationsexperimenten keine ausreichenden Aufmerksamkeitsressourcen für die Verarbeitung der Töne vorhanden sein und diese folglich keinen oder nur einen instabilen Einfluss ausüben.

8.5 Individuelle Unterschiede

Für eine anspruchsvollere zeitliche Aufgabe (Extrapolation der Kugelbewegung nach der Verdeckung) konnten individuelle Unterschiede in der Toninformationsnutzung über die Zeit gefunden werden. Nur sehr wenige Probanden konnten die Töne über den gesamten Zeitraum der Schätzung für beide Schätzdistanzen nutzen. Für die audiovisuellen Diskriminationsexperimente ergaben sich recht konstante Unterschiede in der Nutzung visueller Informationen, wobei ein Großteil der Probanden diese nutzte. Bei der auditorischen Nutzung kam es ebenso zu Unterschieden zwischen den Probanden, wobei diese weniger kohärent waren und sich vermehrt auch intraindividuelle Fluktuationen für den Einfluss der Töne zeigten.

Insgesamt wurden damit in den Experimenten 2 bis 5 Hinweise auf individuelle Unterschiede in der Informationsnutzung bei der audiovisuellen Bewegungsvorhersage gefunden. Dies ist im Einklang mit Studien, in denen ebenfalls für zahlreiche andere multisensorische Phänomene individuelle Unterschiede auftraten, wie beispielsweise dem McGurk-Effekt (z. B. Nath & Beauchamp, 2012), der Ausprägung des temporalen Binding-Fensters für zeitliche Reihenfolgeaufgaben (z. B. Freeman et al., 2013), der Binding-Tendenz (Odegaard & Shams, 2016; Odegaard et al., 2017) und der Integration trimodaler Stimuli (z. B. Haggmann & Russo, 2016). Die Ergebnisse der vorliegenden Untersuchungen können diese Erkenntnisse auf die audiovisuelle Bewegungsvorhersage ausweiten, für die sich ebenfalls individuelle Unterschiede finden lassen konnten. Dies unterstützt die Ansicht, dass individuelle Unterschiede bei Untersuchung multisensorischer Effekte bedacht werden müssen, um valide Schlussfolgerungen aus den Ergebnissen ziehen zu können (Baum et al., 2016). Die Ergebnisse anderer Studien sollten daher mit Vorsicht betrachtet werden, da sich die Hinweise darauf mehren, dass klassische Effekte wie der Redundante-Target-Effekt bzw. die multisensorische kongruente Erleichterung nicht allgemeingültiger Natur zu sein scheinen. Deren Auftreten ist vielmehr an bestimmte Bedingungen wie dem Aufgabenfokus und der Tongestaltung gebunden und könnte zusätzlich von individuellen Begebenheiten der Probanden abhängig sein.

Noch nicht geklärt sind in diesem Zusammenhang zwei entscheidende Aspekte: Zum einen, inwieweit diese Unterschiede flexibel sind und zum anderen, was die Ursachen dieser Unterschiede zwischen den Probanden darstellen. Bezüglich der Veränderbarkeit steht vor allem die Frage nach Strategien und Lernerfahrungen im Vordergrund (siehe dazu auch Abschnitt 8.2). Denkbar wäre eine Strategie sich auf bestimmte Aufgabenaspekte zu fokussieren, in der Extrapolation z. B. auf eine Schätzdistanz, in der Diskrimination auf nur visuelle Informationen oder (einige wenige Probanden) auf beide Modalitäten. Die genauere Rolle von Strategienutzung ließe sich mithilfe von Feedback oder Instruktionshinweisen untersuchen. Würden die Probanden damit ihr Verhalten anpassen und entsprechend ihr Verhalten ändern, spräche dies für eine Formbarkeit der gefundenen Unterschiede. Hinweise für eine Anpassung der Aufgabenstrategie finden sich z. B. für die visuelle Suche (Boot et al., 2009; Irons & Leber,

2016; Proulx, 2011). Bei Boot et al. (2009) nutzten die Probanden für eine visuelle Suchaufgabe zunächst ihre eigene interne Suchstrategie, konnten diese aber nach explizitem Feedback ändern. Auch in den Untersuchungen von Mégevand et al. (2013) waren die Versuchspersonen in der Lage ihr temporales Binding-Fenster an die Aufgabenanforderungen anzupassen und somit ihr multisensorisches Integrationsvermögen entsprechend zu verändern. Auffällig ist jedoch, dass dies nicht allen Probanden gleichermaßen gelang, sondern nur ungefähr der Hälfte der Versuchspersonen. Dies weist darauf hin, dass auch für eine eventuelle Flexibilität der Unterschiede ebenfalls Unterschiede vorliegen können.

Für die Ursachen individueller Unterschiede bei multisensorischen Phänomenen wurden zahlreiche Vorschläge angebracht. So könnten die Fähigkeiten in den Einzelmodalitäten eine Rolle spielen. Caclin et al. (2011) konnten zeigen, dass das Ausmaß einer verbesserten visuellen Detektion durch gleichzeitige Tonpräsentation davon abhängt, inwieweit die visuellen Sehfähigkeiten ausgeprägt sind. Die Verbesserung ist umso stärker, je schwächer die visuellen Fähigkeiten der Probanden sind. Da alle Probanden in den durchgeführten Experimenten hier normalsichtig waren, ist diese Erklärung für die gefundenen Unterschiede allerdings als unwahrscheinlich zu betrachten. Außerdem ist denkbar, dass die Aufmerksamkeit der Probanden eine mögliche Ursache darstellt. Probanden mit Tendenz zur Nutzung auditorischer Informationen könnten wahlweise insgesamt größere Aufmerksamkeitskapazitäten aufweisen oder einen größeren Anteil an Aufmerksamkeitsressourcen für die Aufgabe aufgewandt haben, sodass sie neben den visuellen Informationen zusätzlich auf die nicht primär aufgabenrelevanten Töne zurückgreifen konnten (Marois & Ivanoff, 2005). Probanden mit Fokus auf visuelle Informationen könnten im Gegenzug die Töne eher als störend empfunden haben und konnten daher die einhergehenden auditorischen Informationen nicht gewinnbringend nutzen. Die als irrelevant wahrgenommenen Töne hätten in diesem Fall eine distraktive Wirkung und lenken die Aufmerksamkeit weg von der eigentlichen Aufgabenbearbeitung, womit sich die Leistung der Probanden in der Diskrimination bei gleichzeitiger Tonpräsentation verschlechtert (Banbury et al., 2001; Dalton & Spence, 2007). Tremblay et al. (2007) weisen aufgrund von Korrelationen zwischen verschiedenen multisensorischen Phänomenen auf einen gemeinsamen Mechanismus als Ursache für interindividuelle Unterschiede zwischen den Probanden hin, wobei die Autoren auch eine Aufmerksamkeitskomponente als entscheidenden Faktor vermuten. Dieser Erklärungsansatz ließe sich in Folgeuntersuchungen überprüfen, indem zusätzliche Experimente zur Ausprägung der Aufmerksamkeitsressourcen z. B. in Form des Multiple-Object-Trackings (Drew & Vogel, 2008; Drew et al., 2009) durchgeführt werden, um etwaige Zusammenhänge zur Informationsnutzung bei der Bewegungsvorhersage aufzudecken (siehe dazu auch Abschnitt 8.3).

Eine weitere potentielle Ursache liegt in unterschiedlichen Erfahrungen bzw. Ausprägungen in Fertigkeiten, die mit dem jeweiligen multisensorischen Phänomen zusammenhängen. So fanden Donohue et al. (2010) Unterschiede in Gleichzeitigkeits- und Reihenfolgebeurteilungen für audiovisuelle Stimuli in Abhängigkeit davon, ob die Probanden regelmäßig Action-Videospiele spielen oder nicht. Da Videospiele zumeist multimodal gestaltet sind und man oftmals besser abschneidet, wenn man Hinweise aus verschiedenen Modalitäten nutzt, steigert dies die

multisensorische Integrationsfähigkeit. Ähnliche Effekte lassen sich bei Musikern und Nicht-Musikern finden, wobei die Musiker eine bessere multisensorische Integrationsfähigkeit aufweisen (Bidelman, 2016). Ebenso konnten für Unterschiede beim McGurk-Effekt Zusammenhänge mit den Augenbewegungen (Gurler et al., 2015) sowie der Lippenlesefertigkeit der Probanden (Strand et al., 2014) gefunden werden. Speziell für die Bewegungsvorhersage schlugen Schiff und Oldak (1990) eine Abhängigkeit von der Erfahrung in koordinativen Sportarten wie Tennis und Skifahren oder der Ausprägung räumlicher Orientierungs- und Visualisierungsfähigkeiten vor. Dies wird von Zmigrod und Zmigrod (2016) bestätigt, die einen Zusammenhang zwischen der Breite des audiovisuellen Binding-Fensters und der Leistung in Problemlöseaufgaben fanden. Ein Hinweis für biologische Ursachen findet sich bei Takeuchi et al. (2017), die auf verschieden ausgeprägte Neurotransmitterkonzentrationen (v. a. für Glutamat) bei individuell verschiedenen Bewegungswahrnehmungsphänomenen hinweisen.

Damit könnten die Unterschiede zwischen den Probanden für die audiovisuelle Informationsnutzung bei der Bewegungsvorhersage in anderen Erfahrungen, Fähigkeiten, Fertigkeiten oder auch neuronalen Unterschieden begründet sein. Im Vordergrund der vorliegenden Studie stand, zu untersuchen, wie sich audiovisuelle Stimulation auf die Bewegungsvorhersage auswirkt. Die Hinweise auf Unterschiede zwischen und innerhalb der Probanden ergaben sich erst bei Post-hoc-Betrachtung der Ergebnisse. In zukünftigen Untersuchungen sollte daher zunächst überprüft werden, ob sich Unterschiede zwischen den Probanden bei der Bewegungsvorhersage replizieren lassen. In einem weiteren Schritt sollten vermehrt potentielle Erklärungsfaktoren für diese Unterschiede erhoben werden, um systematische Zusammenhänge mit der Bewegungsvorhersage aufdecken zu können.

8.6 Quasi-natürliches Stimulusmaterial und ökologische Validität

In den durchgeführten Untersuchungen ließen sich verschiedene Einflüsse des Stimulusmaterials finden. In der zeitlichen Detektion in Experiment 1 war der visuelle Erleichterungseffekt in der pseudo-3D-Präsentation stärker ausgeprägt als in der echt-3D-Variante. Für die unimodal visuelle (kombinierte) Detektion und Diskrimination (Experiment 6) war die visuelle Erleichterung hingegen in der echt-3D-Variante größer als in der einfach-2D-Version. Für die zeitliche Extrapolation (Experiment 2) sowie die audiovisuelle Diskrimination (Experiment 3 bis 5) ergaben sich keine Dimensionseffekte.

In Experiment 1 unterschied sich der Faktor Dimension nur in der Disparität, die in der echt-3D-Präsentation vorhanden war. Der größere Einfluss visueller Stimulation in der pseudo-3D-Bedingung in Experiment 1 erscheint daher zu nächst kontraintuitiv, da zum einen binokulare Cues einen relevanten Beitrag zur Bewegungswahrnehmung in der Tiefe leisten (K. R. Brooks & Stone, 2004; Cumming & Parker, 1994; González et al., 2010; Rokers et al., 2009). Zum anderen konnten Harrison et al. (2015) einen größeren Erleichterungseffekt in einer 3D-Umgebung mit Disparität zeigen. Allerdings verglichen diese die korrekten Antworten der Versuchspersonen, welche die Richtung einer Tonbewegung diskriminieren sollten. Möglicherweise hängt ein stärkerer Erleichterungseffekt von der Modalität ab, auf die reagiert werden soll bzw. ist nur für die Antwortgüte, nicht aber die Reaktionsgeschwindigkeit zu finden (siehe dazu auch

Godwin et al., 2017, die einen Einfluss einer zusätzlichen Tiefendimension in der visuellen Suche für die korrekten Antworten, nicht aber die Reaktionszeiten fanden). Dass der Effekt hingegen primär für räumliche Diskriminationen auftritt, ist ebenso denkbar. Dies geht einher mit dem größeren Erleichterungseffekt in der echt-3D-Präsentation für den Diskriminationsteil in Experiment 6. Allerdings trat in den Experimenten 3 bis 5 bei einer solchen Aufgabenstellung kein Dimensionseffekt auf.

Alternativ ist ein verschlechternder Einfluss durch die 3D-Stimulation denkbar. In den Untersuchungen von Cooper et al. (2016) und Katz et al. (2015) war die Sensitivität der Probanden für 2D-Bewegungen größer als für 3D-Bewegungen. Dafür kam es zu Änderungen im Kriterium, welches sich verschob in Richtung zu Bewegungen auf den Betrachter zu (Cooper et al., 2016). Katz et al. (2015) schlugen Unterschiede in der zeitlichen Verarbeitung der Bewegungssignale sowie ein verbessertes Signal-zu-Rauschen-Verhältnis für 2D- gegenüber 3D-Bewegungen als Ursachen für diese Befunde vor. Eine weitere Idee kommt von Bonnen et al. (2017), in deren Untersuchung verfolgten die Probanden Bewegungen in einer 2D- oder 3D-Umgebung. Auch dort war die Sensitivität für die 2D-Variante größer, wobei die Autoren dies auf eine langsamere Disparitätsverarbeitung zurückführten. Da sich die pseudo- und echt-3D-Bedingungen in den Experimenten hier ebenfalls nur durch das Vorhandensein von Disparität unterschieden, sind die erläuterten Einschränkungen in der Verarbeitung als Erklärung denkbar.

Auffällig ist, dass sich die Ausprägung des visuellen Erleichterungseffekts in Experiment 6 umgekehrt verhält, hier war er in der echt-3D-Präsentation stärker ausgeprägt. Allerdings erfolgte in diesem Experiment der Vergleich mit dem einfach-2D-Material, das keine Tiefenhinweise enthielt, im Gegensatz zur pseudo-3D-Variante in Experiment 1, das durch perspektivische und Schatten-Cues gekennzeichnet war. Daraus lässt sich folgern, dass letztere Tiefenhinweise bereits ausreichend scheinen, um eine Umgebung für die Tiefenwahrnehmung zu erzeugen, die zu einer Reaktionserleichterung führt. Beim Vergleich dieses Materials mit zusätzlicher Disparitätskomponente greifen dann die verlangsamenenden Verarbeitungsmechanismen (s. o.). Ob vor allem die perspektivische Gestaltung oder die Schatteneffekte den entscheidenden Unterschied zum einfach-2D-Material ausmachen, kann mit den hier durchgeführten Experimenten nicht eindeutig abgeleitet werden. Jedoch sprechen Studien, in denen die herausragende Rolle des Schlagschattens als Tiefencue herausgestellt wird, eher für einen gewichtigeren Einfluss der Schatteneffekte. So spielen Schatteninformationen bei der Wahrnehmung der Tiefenstruktur einer Szene eine wichtige Rolle (Khuu et al., 2014). Auch für bewegte Stimuli ist der Schlagschatten ein entscheidender Einflussfaktor (Mamassian et al., 1998). In den ursprünglichen Ball-in-a-Box-Experimenten von Kersten et al. (1997) führte das Hinzufügen verschiedener Schlagschatten zu einer qualitativen Wahrnehmungsveränderung der Bewegungsbahn bei den Probanden und unterstützt damit erneut die Wichtigkeit des Schattens für ein sich bewegendes Objekt. Zudem konnten Imura et al. (2008) zeigen, dass der Vorteil von Looming- gegenüber Receding-Stimuli vor allem von einem zur Bewegung des Stimulus passenden Schatten moduliert wird. Die Autoren folgerten aus ihren Experimenten, dass der Schatten eines sich bewegendes Objekts ein entscheidender Faktor dafür ist, eine Näherung

zum Betrachter hin wahrzunehmen. Die Untersuchungen von N.-G. Kim und Grocki (2006) bestätigen zusätzlich, dass die Probanden bei der Bewegungsvorhersage zwar verschiedene Informationsquellen integrieren, aber der Schlagschatten hierbei eine sehr reliable Quelle darstellt. Auch in einer Representational-Momentum-Aufgabe nutzten Versuchspersonen für die Lokalisierung eines Objekts vor allem dessen Schlagschatten (Taya & Miura, 2010). Außerdem konnte in weiteren Studien demonstriert werden, dass die Disparität bei der Wahrnehmung von Bewegungen in der Tiefe eher eine weniger wichtige Rolle einnimmt (Sakano et al., 2012) und stattdessen andere Hinweise wie die interokulare Geschwindigkeitsdifferenz essentiell sind (Sakano & Allison, 2014). Die angeführten Studien legen die Vermutung nahe, dass die Schatteninformationen somit den Unterschied zwischen dem vereinfachtem 2D-Material und dem alltagsnäheren 3D-Material ausmachen könnten. Letztlich müsste dies aber in einem Folgeexperiment, in dem bei der pseudo-3D-Variante der hier durchgeführten Experimente zwischen Bildern mit nur perspektivischen Tiefeninformationen und solchen mit nur Schatteneffekten unterschieden wird, untersucht werden.

Da sich insbesondere die Effekte der Disparität recht inhomogen darstellen, stellt sich die Frage, unter welchen Umständen Disparität eine Wirkung ausübt. Gegebenenfalls könnten erneut Unterschiede zwischen den Probanden eine Rolle spielen. Es konnte gezeigt werden, dass sich Versuchspersonen darin unterscheiden, ob sie für die Tiefenschätzung eher perspektivische Cues oder Disparität als Hinweis nutzen (Zalevski et al., 2007). Ebenso präferieren Probanden mit unterschiedlicher Vorliebe bei konfligierenden Informationen eher Disparitäts- oder Perspektivenhinweise (Sato & Howard, 2001). Auch für die Nutzung von sich verändernder Disparität oder der interokularen Geschwindigkeit bei der 3D-Bewegungswahrnehmung konnten individuelle Unterschiede gefunden werden (Allen et al., 2015; Nefs et al., 2010). Demnach wäre denkbar, dass einige Versuchspersonen tendenziell Disparität als Tiefenhinweis nutzen, andere hingegen nicht. Die inhomogenen Ergebnisse könnten in diesem Fall zumindest teilweise auf verschieden zusammengesetzte Stichproben bezüglich der Nutzungswahrscheinlichkeit von Tiefencues zurückzuführen sein. Ob dem tatsächlich so ist, müsste in Folgeexperimenten genauer untersucht werden.

Darüber hinaus stellt sich die Frage, warum in den Experimenten 2 bis 5 keine Dimensionseffekte gefunden werden konnten. Der fehlende Einfluss in der Extrapolation könnte durch die reduzierte räumliche Komponente zu erklären sein. Die Kugel tauchte nach dem Verdecken nicht wieder auf und der räumliche Bahnverlauf spielte für die Aufgabe keine Rolle. Damit könnten für Effekte von realitätsgetreuem Material die Verwendung von Aufgaben entscheidend sein, in denen räumliche Aspekte in irgendeiner Form eine Rolle spielen (Conrad et al., 2013; Getzmann & Lewald, 2010; Harrison et al., 2015), wie dies z. B. auch für die zeitliche Detektionsaufgabe mit dem variierten Bahnverlauf der Fall war. Der Nicht-Effekt könnte zudem in der Extrapolationsaufgabe dadurch verstärkt bzw. forciert werden, dass die Kugel zum Zeitpunkt der Reaktion nicht mehr zu sehen war. Es ist denkbar, dass für einen Einfluss des Stimulusmaterials ein Objekt, auf das reagiert werden soll, visuell sichtbar sein muss.

Auch für die drei audiovisuellen Diskriminationsexperimente fanden sich keine Effekte der Präsentationsdimension. Dies ist vor allem deshalb verwunderlich, weil mit exakt demselben

Stimulusmaterial in Experiment 1 Effekte gefunden wurden (vergleiche Experiment 3, jedoch wurde dort eine andere Aufgabe verwendet) und zudem Einflüsse in Experiment 6 ebenso auftraten (vergleiche Experiment 5, jedoch wurden dort zusätzlich Töne präsentiert). Dies könnte auf die Tonstimulation in den Experimenten 3 bis 5 zurückzuführen sein. Gegebenenfalls haben die Töne, die in der audiovisuellen Diskrimination verstärkt intraindividuelle Fluktuationen erzeugten (s. o.) die Daten verrauscht, womit etwaige Effekte der Präsentationsdimension in diesen Experimenten nicht sichtbar werden konnten. Zukünftige Untersuchungen sollten ebenfalls mit natürlicherem Tonmaterial arbeiten, indem beispielsweise statt des weißen Rauschens eine realistischere Rollbewegung als auditorischer Stimulus verwendet wird. Käme es in diesen Experimenten erneut zu Einflüssen der Präsentationsdimension, dann stellte die semantische Passung von visueller und auditorischer Bewegungsstimulation eine wichtige Voraussetzung dar, damit in Experimenten mit 3D-Umgebungen valide Ergebnisse gefunden werden können.

Der inhomogene Einfluss der Präsentationsdimension könnte zudem auf die eher niedrige Komplexität des hier verwendeten Stimulusmaterials sowie der verwendeten Aufgaben zurückzuführen sein. Viele Untersuchungen, in denen ein Unterschied zwischen einer 2D- und 3D-Experimentalumgebung gefunden werden konnte, sind durch einen höheren Komplexitätsgrad und/oder anspruchsvollere Aufgaben gekennzeichnet, wie z. B. die räumliche Navigation in einem virtuellen Labyrinth bzw. Gebäude (Kober et al., 2012; Slobounov et al., 2015), das Falten von Origamiobjekten (Dan & Reiner, 2017, 2018) oder das Lernen biologischer Sachverhalte (Richards & Taylor, 2015). Diese Annahme wird von McIntire et al. (2012) aus dem Bereich der Human-Factor-Forschung bestätigt. Diese führten zahlreiche Studien, die einen Einfluss von 3D-Präsentation bzw. -Umgebungen untersucht haben, zusammen und kamen zu dem Schluss, dass Effekte vor allem für komplexe Aufgaben, wie bei der Orientierung im Raum, auftraten oder wenn von den Versuchspersonen eine Handlung im Raum durchgeführt werden sollte.

Insgesamt bestätigen die gefundenen Unterschiede für die verschiedenen genutzten Materialgestaltungen dennoch zahlreiche Studien, die einen Einfluss des Stimulusmaterials, insbesondere für 2D- vs. 3D-Darbietung bzw. artifiziell/vereinfacht vs. realitätsnah, belegen konnten (Conrad et al., 2013; Gaebler et al., 2014; Getzmann & Lewald, 2010; Harrison et al., 2015; Ogawa & Macaluso, 2013; Ogawa et al., 2013; Senkowski et al., 2007; Snow et al., 2011). Sowohl Disparität allein als auch die Kombination von perspektivischen Tiefenhinweisen und Schatteneffekten üben einen Einfluss auf die (audiovisuelle) Bewegungsvorhersage im Raum aus. Semantisch nicht passende Töne könnten zudem Effekte des visuellen Materials in multisensorischen Umgebungen abschwächen. Die Nutzung realitätsnaher Stimuli aller verwendeten Modalitäten sowie eine realitätsgetreue Gestaltung des gesamten Experimental-Set-Ups ist daher angezeigt, da bei der Verwendung von vereinfachtem Stimulusmaterial in Untersuchungen zur audiovisuellen Bewegungsvorhersage gegebenenfalls Effekte nicht nur über- bzw. unterschätzt, sondern auch vollständig übersehen werden könnten. Da zahlreiche Studien zudem belegen, dass sich diese Materialunterschiede zusätzlich auf die Untersuchung neuronaler Korrelate und zeitlicher Dynamiken auswirken, sollte dies für zukünftige Studien

entschieden bedacht werden (Gaebler et al., 2014; Getzmann & Lewald, 2010; Harrison et al., 2015; Ogawa & Macaluso, 2013; Ogawa et al., 2013; Senkowski et al., 2007; Snow et al., 2011). Nur so ist es möglich, aus den Ergebnissen in Experimenten valide Aussagen über das Verhalten und das damit einhergehende neuronale Geschehen beim Menschen im Alltag treffen zu können. Allerdings scheint der Einfluss von realitätsgetreuem Material an viele Einflussfaktoren gebunden: Aufgabendomäne (zeitlich vs. räumlich), Aufgabenanforderung (Detektion vs. Extrapolation vs. Diskrimination) und Aufgabenmodalität (visuell vs. auditorisch sowie uni vs. multimodal) scheinen einige relevante davon zu sein. Ziel von Folgeuntersuchungen sollte es sein, diese genauer sowohl in ihrer Einzelwirkung als auch ihren Interaktionen zu untersuchen.

9 Ausblick

Für zukünftige Untersuchungen multisensorischer Phänomene sollten interindividuelle Unterschiede nicht nur bedacht, sondern genauer spezifiziert werden. Insbesondere eine genauere Charakterisierung, welche weiteren multisensorischen Interaktionen ebenfalls betroffen sind und was zugrunde liegende Ursachen für diese Unterschiede darstellen, sollten dabei im Vordergrund stehen. Auch für einhergehende neuronale Korrelate müssen individuelle Unterschiede sowohl in der Lokalisierung von Aktivierungen, aber auch für deren Signalstärken sowie zeitlichen Dynamiken in Betracht gezogen werden (Baum et al., 2016). Darüber hinaus stellt sich die Frage, ob das Vorhandensein individueller Unterschiede im multisensorischen Bereich an bestimmte Bedingungen geknüpft ist. Es sollte daher genauer exploriert werden, unter welchen Umständen in multisensorischen Experimenten allgemeingültige Ergebnisse und wann Unterschiede zwischen den Probanden zu erwarten sind. Darüber hinaus ist eine Untersuchung bzw. Beachtung individueller Unterschiede auch für allgemeinspsychologische Phänomene generell wünschenswert (Albrecht & Mattler, 2010; Vogel & Awh, 2008).

Außerdem ist in der zukünftigen Forschung verstärkt die Verwendung realistischer und damit alltagsnaher Stimulationsmaterials angezeigt. Die Interaktionen mit der Materialgestaltung in den Experimenten der vorliegenden Arbeit belegen, dass vereinfachtes Stimulusmaterial andere Effekte hervorruft und die Ergebnisse aus solchen Experimenten nur bedingt auf das alltägliche Verhalten übertragbar sind. Um aus Experimenten ökologisch valide Schlussfolgerungen ziehen zu können, ist eine realistische Experimentalumgebung demnach unabdingbar. Zudem sollte zukünftig auf die Nutzung konfligierender Stimuli verzichtet werden, da diese in unserem Alltag nur extrem selten vorkommen und daher die Ergebnisse durch Adaptionsprozesse verfälscht werden könnten (de Gelder & Bertelson, 2003). Insbesondere die Entwicklungen im Bereich der Virtuellen Realität sind bei der Umsetzung alltagsnaher Forschung als vielversprechend anzusehen. Diese bieten nicht nur die Möglichkeit, für Experimente eine realistische Umgebung zu schaffen, sondern eröffnen zudem die Chance durch Handlungen und Konsequenzen in diesem Raum eine noch alltagsnähere Gestaltung zu erzielen (Bernhard et al., 2011; Bonfiglioli et al., 2004; Gröhn et al., 2003; Riecke et al., 2009; Rose et al., 2018; Sella et al., 2014). Auch eine intensivere Untersuchung von Handlungskonsequenzen, z. B. in Form von Abfangbewegungen für sich bewegende Objekte (Bosco et al., 2012, 2015; Cesqui et al., 2012), bietet das Potential für Untersuchungen mit hohem ökologischen Validitätsgehalt.

Für den Anwendungskontext bieten Untersuchungen zu multisensorischen Vorhersageprozessen ebenso ein breites Spektrum: Vor allem Lernumgebungen, in denen Bewegungen eine Rolle spielen, könnten dadurch nutzerfreundlicher gestaltet werden, z. B. im Bereich der Fahrschule oder des medizinischen Trainings (Heng et al., 2004; Toet et al., 2013). Auch Navigationssysteme (Asif & Boll, 2010; Baldwin, 2009; E. Park et al., 2017) oder die Gestaltung von Warnsignalen in Fahrzeugen (Ho & Spence, 2005; Ho et al., 2007; Meng et al., 2016) können von den Erkenntnissen aus diesen Experimenten profitieren und in ihrer Ausführung angepasst werden, je nachdem, ob eine multisensorische Gestaltung Nutzen mit sich bringt oder nicht.

10 Schlussfolgerungen

Mit den Ergebnissen der Experimente in dieser Arbeit konnte gezeigt werden, dass die Aufgabenstellung einen Einfluss auf das Informationsnutzungsverhalten der Probanden bei der audiovisuellen Bewegungsvorhersage nimmt. In zeitlichen Aufgaben greifen Probanden vermehrt auf zusätzlich präsentierte Töne zurück, in Aufgaben mit räumlichem Fokus ist dies weniger der Fall. Zudem konnten mehrfach Hinweise auf individuelle Unterschiede zwischen den Probanden gefunden werden: Die Probanden unterscheiden sich in ihrer Fähigkeit, für welche Distanzen sie Töne bei der Extrapolation von Bewegungen nutzen können. Zudem ergaben sich Unterschiede im Vorteil visuell passender Bahnen für die audiovisuellen Diskriminationsexperimente. In diesen verursachten Töne zudem sehr inhomogene Effekte. Dies ist vermutlich auf eine geringe semantische Übereinstimmung und einer einhergehenden fehlenden multisensorischen Integration beider Modalitäten bei den meisten Probanden zurückzuführen. Stattdessen werden die Töne eher als störend wahrgenommen und die Aufmerksamkeit dafür fluktuiert innerhalb der Probanden über die Experimentaldauer hinweg. Damit müssen in zukünftigen Untersuchungen zwingend individuelle Unterschiede als potentieller Einflussfaktor bedacht werden. Diese sollten dabei nicht als Störquelle angesehen werden, die man mit mehr Versuchspersonen und einer höheren Anzahl an Durchgängen umgehen kann. Vielmehr sollte man sie als Chance begreifen und in weiteren Experimenten im Detail charakterisieren. Außerdem ergaben sich Interaktionen mit der Gestaltung des visuellen Stimulusmaterials. Das Hinzufügen von Disparität führte zu einem geringer ausgeprägten visuellen Erleichterungseffekt, wohingegen beim Vergleich von vereinfachtem Stimulusmaterial mit 3D-Material mit diversen Tiefenhinweisen, die Reaktionen in der realitätsgetreueren 3D-Darstellung stärker erleichtert wurden. Daher sollte zukünftig auf vereinfachtes Stimulusmaterial in Untersuchungen zur multisensorischen Bewegungsvorhersage verzichtet werden, insbesondere eine realitätsnahe Darstellung in der Tiefendimension ist ebenso wie die Komplexität des Untersuchungssettings von maßgeblicher Relevanz um valide Folgerungen für das Verhalten der Menschen im Alltag treffen zu können.

Literatur

- Alais, D. & Burr, D. C. (2004a). No direction-specific bimodal facilitation for audiovisual motion detection. *Brain Research. Cognitive Brain Research*, *19* (2), 185–194. doi: 10.1016/j.cogbrainres.2003.11.011
- Alais, D. & Burr, D. C. (2004b). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, *14* (3), 257–262. doi: 10.1016/j.cub.2004.01.029
- Albrecht, T., Klapötke, S. & Mattler, U. (2010). Individual differences in metacontrast masking are enhanced by perceptual learning. *Consciousness and Cognition*, *19* (2), 656–666. doi: 10.1016/j.concog.2009.12.002
- Albrecht, T. & Mattler, U. (2010). Individual differences in metacontrast masking: A call for caution when interpreting group data. *Consciousness and Cognition*, *19* (2), 672–673. doi: 10.1016/j.concog.2010.03.010
- Albrecht, T. & Mattler, U. (2016). Individually different weighting of multiple processes underlies effects of metacontrast masking. *Consciousness and Cognition*, *42*, 162–180. doi: 10.1016/j.concog.2016.03.006
- Alink, A., Euler, F., Galeano, E., Krugliak, A., Singer, W. & Kohler, A. (2012). Auditory motion capturing ambiguous visual motion. *Frontiers in Psychology*, *2*, 391. doi: 10.3389/fpsyg.2011.00391
- Alink, A., Singer, W. & Muckli, L. (2008). Capture of auditory motion by vision is represented by an activation shift from auditory to visual motion cortex. *The Journal of Neuroscience*, *28* (11), 2690–2697. doi: 10.1523/JNEUROSCI.2980-07.200
- Allen, B., Haun, A. M., Hanley, T., Green, C. S. & Rokers, B. (2015). Optimal Combination of the Binocular Cues to 3D Motion. *Investigative Ophthalmology & Visual Science*, *56* (12), 7589–7596. doi: 10.1167/iovs.15-17696
- Arrighi, R., Marini, F. & Burr, D. C. (2009). Meaningful auditory information enhances perception of visual biological motion. *Journal of Vision*, *9* (4), 25. doi: 10.1167/9.4.25
- Aschersleben, G. & Bertelson, P. (2003). Temporal ventriloquism: Crossmodal interaction on the time dimension. *International Journal of Psychophysiology*, *50* (1-2), 157–163. doi: 10.1016/S0167-8760(03)00131-4
- Asif, A. & Boll, S. (2010). Where to turn my car? Comparison of a Tactile Display and a Conventional Car Navigation System under High Load Condition. In A. K. Dey (Hrsg.), *Proceedings of 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2010, Pittsburgh, PA, USA, November 11-12, 2010* (S. 64–71). New York: ACM. doi: 10.1145/1969773.1969786
- Assad, J. A. & Maunsell, J. H. R. (1995). Neuronal correlates of inferred motion in primate posterior parietal cortex. *Nature*, *373* (6514), 518–521. doi: 10.1038/373518a0
- Assmus, A., Marshall, J. C., Noth, J., Zilles, K. & Fink, G. R. (2005). Difficulty of perceptual spatiotemporal integration modulates the neural activity of left inferior parietal cortex. *Neuroscience*, *132* (4), 923–927. doi: 10.1016/j.neuroscience.2005.01.047

- Assmus, A., Marshall, J. C., Ritzl, A., Noth, J., Zilles, K. & Fink, G. R. (2003). Left inferior parietal cortex integrates time and space during collision judgments. *NeuroImage*, *20*, S82-S88. doi: 10.1016/j.neuroimage.2003.09.025
- Baker, C. I., Keyser, C., Jellema, T., Wicker, B. & Perrett, D. I. (2001). Neuronal representation of disappearing and hidden objects in temporal cortex of the macaque. *Experimental Brain Research*, *140* (3), 375–381.
- Baldwin, C. L. (2009). Individual differences in navigational strategy: Implications for display design. *Theoretical Issues in Ergonomics Science*, *10* (5), 443–458. doi: 10.1080/14639220903106379
- Banbury, S. P., Macken, W. J., Tremblay, S. & Jones, D. M. (2001). Auditory Distraction and Short-Term Memory: Phenomena and Practical Implications. *Human Factors*, *43* (1), 12–29. doi: 10.1518/001872001775992462
- Barrett, M. M. & Newell, F. N. (2015). Task-Specific, Age Related Effects in the Cross-Modal Identification and Localisation of Objects. *Multisensory Research*, *28* (1-2), 111–151. doi: 10.1163/22134808-00002479
- Baum, S. H., Colonus, H., Thelen, A., Micheli, C. & Wallace, M. T. (2016). Above the Mean: Examining Variability in Behavioral and Neural Responses to Multisensory Stimuli. *Multisensory Research*, *29* (6-7), 663–678. doi: 10.1163/22134808-00002536
- Baumann, O. & Greenlee, M. W. (2007). Neural Correlates of Coherent Audiovisual Motion Perception. *Cerebral Cortex*, *17* (6), 1433–1443. doi: 10.1093/cercor/bhl055
- Becic, E., Kramer, A. F. & Boot, W. R. (2007). Age-related differences in visual search in dynamic displays. *Psychology and Aging*, *22* (1), 67–74. doi: 10.1037/0882-7974.22.1.67
- Beck, M. R., Hong, S. L., van Lamsweerde, A. E. & Ericson, J. M. (2014). The Effects of Incidentally Learned Temporal and Spatial Predictability on Response Times and Visual Fixations during Target Detection and Discrimination. *PLoS ONE*, *9* (4), e94539. doi: 10.1371/journal.pone.0094539
- Bennett, S. J., Baurès, R., Hecht, H. & Benguigui, N. (2010). Eye movements influence estimation of time-to-contact in prediction motion. *Experimental Brain Research*, *206* (4), 399–407. doi: 10.1007/s00221-010-2416-y
- Berger, A. (2006). Individual performance based on cognitive experimental measurements? The case of inhibition of return. *Experimental Psychology*, *53* (3), 209–217. doi: 10.1027/1618-3169.53.3.209
- Bernhard, M., Grosse, K. & Wimmer, M. (2011). Bimodal task-facilitation in a virtual traffic scenario through spatialized sound rendering. *ACM Transactions on Applied Perception*, *8* (4), 1–22. doi: 10.1145/2043603.2043606
- Bertelson, P. & Boons, J.-P. (1960). Time Uncertainty and Choice Reaction Time. *Nature*, *187* (4736), 531–532. doi: 10.1038/187531a0
- Berti, S. & Schröger, E. (2003). Working memory controls involuntary attention switching: Evidence from an auditory distraction paradigm. *European Journal of Neuroscience*, *17* (5), 1119–1122. doi: 10.1046/j.1460-9568.2003.02527.x

- Besle, J., Fort, A., Delpuech, C. & Giard, M.-H. (2004). Bimodal speech: Early suppressive visual effects in human auditory cortex. *The European Journal of Neuroscience*, *20* (8), 2225–2234. doi: 10.1111/j.1460-9568.2004.03670.x
- Beudel, M., Renken, R., Leenders, K. L. & de Jong, B. M. (2009). Cerebral representations of space and time. *NeuroImage*, *44* (3), 1032–1040. doi: 10.1016/j.neuroimage.2008.09.028
- Bidelman, G. M. (2016). Musicians have enhanced audiovisual multisensory binding: experience-dependent effects in the double-flash illusion. *Experimental Brain Research*, *234* (10), 3037–3047. doi: 10.1007/s00221-016-4705-6
- Bonfiglioli, C., Pavani, F. & Castiello, U. (2004). Differential Effects of Cast Shadows on Perception and Action. *Perception*, *33* (11), 1291–1304. doi: 10.1068/p5325
- Bonnen, K., Huk, A. C. & Cormack, L. K. (2017). Dynamic mechanisms of visually guided 3D motion tracking. *Journal of Neurophysiology*, *118* (3), 1515–1531. doi: 10.1152/jn.00831.2016
- Boot, W. R., Becic, E. & Kramer, A. F. (2009). Stable individual differences in search strategy? The effect of task demands and motivational factors on scanning strategy in visual search. *Journal of Vision*, *9* (3), 7. doi: 10.1167/9.3.7
- Boot, W. R., Kramer, A. F., Becic, E., Wiegmann, D. A. & Kubose, T. (2006). Detecting Transient Changes in Dynamic Displays: The More You Look, the Less You See. *Human Factors*, *48* (4), 759–773. doi: 10.1518/001872006779166424
- Bosco, G., Monache, S. D., Gravano, S., Indovina, I., La Scaleia, B., Maffei, V., . . . Lacquaniti, F. (2015). Filling gaps in visual motion for target capture. *Frontiers in Integrative Neuroscience*, *9*, 13. doi: 10.3389/fnint.2015.00013
- Bosco, G., Monache, S. D. & Lacquaniti, F. (2012). Catching What We Can't See: Manual Interception of Occluded Fly-Ball Trajectories. *PLoS ONE*, *7* (11), e49381. doi: 10.1371/journal.pone.0049381
- Boutet, I., Lemieux, C. L., Goulet, M.-A. & Collin, C. A. (2017). Faces elicit different scanning patterns depending on task demands. *Attention, Perception, & Psychophysics*, *79* (4), 1050–1063. doi: 10.3758/s13414-017-1284-y
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10* (4), 433–436. doi: 10.1163/156856897X00357
- Brawn, P. T. & Snowden, R. J. (2000). Attention to Overlapping Objects: Detection and Discrimination of Luminance Changes. *Journal of Experimental Psychology: Human Perception and Performance*, *1* (342–358). doi: 10.1037//0096-1523.26.1.342
- Bremner, J. G., Slater, A., Johnson, S. P., Mason, U. & Spring, J. (2012). The Effects of Auditory Information on 4-Month-Old Infants' Perception of Trajectory Continuity. *Child Development*, *83* (3), 954–964. doi: 10.1111/j.1467-8624.2012.01739.x
- Bresciani, J.-P., Dammeier, F. & Ernst, M. O. (2008). Tri-modal integration of visual, tactile and auditory signals for the perception of sequences of events. *Brain Research Bulletin*, *75* (6), 753–760. doi: 10.1016/j.brainresbull.2008.01.009

- Bresciani, J.-P. & Ernst, M. O. (2007). Signal reliability modulates auditory-tactile integration for event counting. *NeuroReport*, *18* (11), 1157–1161. doi: 10.1097/WNR.0b013e3281ace0ca
- Bridwell, D. A., Hecker, E. A., Serences, J. T. & Srinivasan, R. (2013). Individual differences in attention strategies during detection, fine discrimination, and coarse discrimination. *Journal of Neurophysiology*, *110* (3), 784–794. doi: 10.1152/jn.00520.2012
- Brooks, A., van der Zwan, R., Billard, A., Petreska, B., Clarke, S. & Blanke, O. (2007). Auditory motion affects visual biological motion processing. *Neuropsychologia*, *45* (3), 523–530. doi: 10.1016/j.neuropsychologia.2005.12.012
- Brooks, K. R. & Stone, L. S. (2004). Stereomotion speed perception: Contributions from both changing disparity and interocular velocity difference over a range of relative disparities. *Journal of Vision*, *4* (12), 1061–1079. doi: 10.1167/4.12.6
- Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H. & Woldorff, M. G. (2005). The spread of attention across modalities and space in a multisensory object. *Proceedings of the National Academy of Sciences*, *102* (51), 18751–18756. doi: 10.1073/pnas.0507704102
- Caclin, A., Bouchet, P., Djoulah, F., Pirat, E., Pernier, J. & Giard, M.-H. (2011). Auditory enhancement of visual perception at threshold depends on visual abilities. *Brain Research*, *1396*, 35–44. doi: 10.1016/j.brainres.2011.04.016
- Caird, J. K. & Hancock, P. A. (1994). The Perception of Arrival Time for Different Oncoming Vehicles at an Intersection. *Ecological Psychology*, *6* (2), 83–109. doi: 10.1207/s15326969eco0602_1
- Cappe, C., Thut, G., Romei, V. & Murray, M. M. (2009). Selective integration of auditory-visual looming cues by humans. *Neuropsychologia*, *47* (4), 1045–1052. doi: 10.1016/j.neuropsychologia.2008.11.003
- Cesqui, B., d'Avella, A., Portone, A. & Lacquaniti, F. (2012). Catching a Ball at the Right Time and Place: Individual Factors Matter. *PLoS ONE*, *7* (2), e31770. doi: 10.1371/journal.pone.0031770
- Chen, Y.-C. & Spence, C. (2010). When hearing the bark helps to identify the dog: Semantically-congruent sounds modulate the identification of masked pictures. *Cognition*, *114* (3), 389–404. doi: 10.1016/j.cognition.2009.10.012
- Chen, Y.-C. & Spence, C. (2017). Assessing the Role of the 'Unity Assumption' on Multisensory Integration: A Review. *Frontiers in Psychology*, *8*, 445. doi: 10.3389/fpsyg.2017.00445
- Chica, A. B., Lupianez, J. & Bartolomeo, P. (2006). Dissociating inhibition of return from endogenous orienting of spatial attention: Evidence from detection and discrimination tasks. *Cognitive Neuropsychology*, *23* (7), 1015–1034. doi: 10.1080/02643290600588277
- Chien, S.-E., Ono, F. & Watanabe, K. (2013). A Transient Auditory Signal Shifts the Perceived Offset Position of a Moving Visual Object. *Frontiers in Psychology*, *4*, 70. doi: 10.3389/fpsyg.2013.00070
- Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, *16* (2), 409–412. doi: 10.3758/BF03203962

- Conrad, V., Bartels, A., Kleiner, M. & Noppeney, U. (2010). Audiovisual interactions in binocular rivalry. *Journal of Vision*, *10* (10), 27. doi: 10.1167/10.10.27
- Conrad, V., Kleiner, M., Bartels, A., O'Brien, J. H., Bühlhoff, H. H. & Noppeney, U. (2013). Naturalistic Stimulus Structure Determines the Integration of Audiovisual Looming Signals in Binocular Rivalry. *PLoS ONE*, *8* (8), e70710. doi: 10.1371/journal.pone.0070710
- Cooper, E. A., van Ginkel, M. & Rokers, B. (2016). Sensitivity and bias in the discrimination of two-dimensional and three-dimensional motion direction. *Journal of Vision*, *16* (10), 5. doi: 10.1167/16.10.5
- Correa, Á., Lupiáñez, J., Milliken, B. & Tudela, P. (2004). Endogenous temporal orienting of attention in detection and discrimination tasks. *Perception & Psychophysics*, *66* (2), 264–278. doi: 10.3758/BF03194878
- Cumming, B. G. & Parker, A. J. (1994). Binocular Mechanisms for Detecting Motion-in-Depth. *Vision Research*, *34* (4), 483–495. doi: 10.1016/0042-6989(94)90162-7
- Dalton, P. & Spence, C. (2007). Attentional capture in serial audiovisual search tasks. *Perception & Psychophysics*, *69* (3), 422–438. doi: 10.3758/BF03193763
- Damian, M. F. (2010). Does variability in human performance outweigh imprecision in response devices such as computer keyboards? *Behavior Research Methods*, *42* (1), 205–211. doi: 10.3758/BRM.42.1.205
- Dan, A. & Reiner, M. (2017). EEG-based cognitive load of processing events in 3D virtual worlds is lower than processing events in 2D displays. *International Journal of Psychophysiology*, *122*, 75–84. doi: 10.1016/j.ijpsycho.2016.08.013
- Dan, A. & Reiner, M. (2018). Reduced mental load in learning a motor visual task with virtual 3D method. *Journal of Computer Assisted Learning*, *34* (1), 84–93. doi: 10.1111/jcal.12216
- de Gelder, B. & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences*, *7* (10), 460–467. doi: 10.1016/j.tics.2003.08.014
- Deluca, C., Golzar, A., Santandrea, E., Lo Gerfo, E., Eštočinová, J., Moretto, G., . . . Chelazzi, L. (2014). The cerebellum and visual perceptual learning: Evidence from a motion extrapolation task. *Cortex*, *58*, 52–71. doi: 10.1016/j.cortex.2014.04.017
- DeLucia, P. R. (2004a). Chapter 11. Multiple sources of information influence time-to-contact judgments: Do heuristics accommodate limits in sensory and cognitive processes? *Advances in Psychology*, *135*, 243–285. doi: 10.1016/S0166-4115(04)80013-X
- DeLucia, P. R. (2004b). Time-to-Contact Judgments of an Approaching Object That Is Partially Concealed by an Occluder. *Journal of Experimental Psychology: Human Perception and Performance*, *30* (2), 287–304. doi: 10.1037/0096-1523.30.2.287
- DeLucia, P. R. & Liddell, G. W. (1998). Cognitive motion extrapolation and cognitive clocking in prediction motion tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *24* (3), 901–914. doi: 10.1037/0096-1523.24.3.901
- DeLucia, P. R., Preddy, D. & Oberfeld, D. (2016). Audiovisual Integration of Time-to-Contact Information for Approaching Objects. *Multisensory Research*, *29* (4-5), 365–395. doi: 10.1163/22134808-00002520

- de'Sperati, C. & Deubel, H. (2006). Mental extrapolation of motion modulates responsiveness to visual stimuli. *Vision Research*, *46* (16), 2593–2601. doi: 10.1016/j.visres.2005.12.019
- Dittrich, S. & Noesselt, T. (2018). Temporal Audiovisual Motion Prediction in 2D- vs. 3D-Environments. *Frontiers in Psychology*, *9*, 368. doi: 10.3389/fpsyg.2018.00368
- Donohue, S. E., Green, J. J. & Woldorff, M. G. (2015). The effects of attention on the temporal integration of multisensory stimuli. *Frontiers in Integrative Neuroscience*, *9*, 32. doi: 10.3389/fnint.2015.00032
- Donohue, S. E., Woldorff, M. G. & Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. *Attention, Perception, & Psychophysics*, *72* (4), 1120–1129. doi: 10.3758/APP.72.4.1120
- Drew, T., McCollough, A. W., Horowitz, T. S. & Vogel, E. K. (2009). Attentional enhancement during multiple-object tracking. *Psychonomic Bulletin & Review*, *16* (2), 411–417. doi: 10.3758/PBR.16.2.411
- Drew, T. & Vogel, E. K. (2008). Neural Measures of Individual Differences in Selecting and Tracking Multiple Moving Objects. *The Journal of Neuroscience*, *28* (16), 4183–4191. doi: 10.1523/JNEUROSCI.0556-08.2008
- Ecker, A. J. & Heller, L. M. (2005). Auditory-visual interactions in the perception of a ball's path. *Perception*, *34*, 59–75. doi: 10.1068/p5368
- Eg, R. & Behne, D. M. (2015). Perceived synchrony for realistic and dynamic audiovisual events. *Frontiers in Psychology*, *6*, 736. doi: 10.3389/fpsyg.2015.00736
- Ernst, M. O. & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415* (6870), 429–433. doi: 10.1038/415429a
- Esterman, M., Rosenberg, M. D. & Noonan, S. K. (2014). Intrinsic Fluctuations in Sustained Attention and Distractor Processing. *The Journal of Neuroscience*, *34* (5), 1724–1730. doi: 10.1523/JNEUROSCI.2658-13.2014
- Falchier, A., Clavagnier, S., Barone, P. & Kennedy, H. (2002). Anatomical Evidence of Multimodal Integration in Primate Striate Cortex. *The Journal of Neuroscience*, *22* (13), 5749–5759. doi: 10.1523/JNEUROSCI.22-13-05749.2002
- Fendrich, R. & Corballis, P. M. (2001). The temporal cross-capture of audition and vision. *Perception & Psychophysics*, *63* (4), 719–725. doi: 10.3758/BF03194432
- Flombaum, J. I. & Scholl, B. J. (2006). A temporal same-object advantage in the tunnel effect: Facilitated change detection for persisting objects. *Journal of Experimental Psychology: Human Perception and Performance*, *32* (4), 840–853. doi: 10.1037/0096-1523.32.4.840
- Flombaum, J. I., Scholl, B. J. & Santos, L. R. (2009). Spatiotemporal priority as a fundamental principle of object persistence. In B. M. Hood & L. R. Santos (Hrsg.), *The Origins of Object Knowledge* (S. 135–164). Oxford: Oxford University Press.
- Freeman, E. D. & Driver, J. (2008). Direction of Visual Apparent Motion Driven Solely by Timing of a Static Sound. *Current Biology*, *18* (16), 1262–1266. doi: 10.1016/j.cub.2008.07.066

- Freeman, E. D., Ipser, A., Palmbaha, A., Paunoiu, D., Brown, P., Lambert, C., ... Driver, J. (2013). Sight and sound out of synch: Fragmentation and renormalisation of audiovisual integration and subjective timing. *Cortex*, *49* (10), 2875–2887. doi: 10.1016/j.cortex.2013.03.006
- Fukuda, K. & Vogel, E. K. (2011). Individual Differences in Recovery Time From Attentional Capture. *Psychological Science*, *22* (3), 361–368. doi: 10.1177/0956797611398493
- Fulvio, J. M., Maloney, L. T. & Schrater, P. R. (2015). Revealing individual differences in strategy selection through visual motion extrapolation. *Cognitive Neuroscience*, *6* (4), 169–179. doi: 10.1080/17588928.2014.1003181
- Gaebler, M., Biessmann, F., Lamke, J.-P., Müller, K.-R., Walter, H. & Hetzer, S. (2014). Stereoscopic depth increases intersubject correlations of brain networks. *NeuroImage*, *100*, 427–434. doi: 10.1016/j.neuroimage.2014.06.008
- Getzmann, S. (2007). The Effect of Brief Auditory Stimuli on Visual Apparent Motion. *Perception*, *36* (7), 1089–1103. doi: 10.1068/p5741
- Getzmann, S. & Lewald, J. (2010). Effects of natural versus artificial spatial cues on electrophysiological correlates of auditory motion. *Hearing Research*, *259* (1–2), 44–54. doi: 10.1016/j.heares.2009.09.021
- Girard, S., Collignon, O. & Lepore, F. (2011). Multisensory gain within and across hemispaces in simple and choice reaction time paradigms. *Experimental Brain Research*, *214* (1), 1–8. doi: 10.1007/s00221-010-2515-9
- Gleiss, S. & Kayser, C. (2013). Eccentricity dependent auditory enhancement of visual stimulus detection but not discrimination. *Frontiers in Integrative Neuroscience*, *7*, 52. doi: 10.3389/fnint.2013.00052
- Gleiss, S. & Kayser, C. (2014). Oscillatory mechanisms underlying the enhancement of visual motion perception by multisensory congruency. *Neuropsychologia*, *53*, 84–93. doi: 10.1016/j.neuropsychologia.2013.11.005
- Godwin, H. J., Menneer, T., Liversedge, S. P., Cave, K. R., Holliman, N. S. & Donnelly, N. (2017). Adding depth to overlapping displays can improve visual search performance. *Journal of Experimental Psychology: Human Perception and Performance*, *43* (8), 1532–1549. doi: 10.1037/xhp0000353
- González, E. G., Allison, R. S., Ono, H. & Vinnikov, M. (2010). Cue conflict between disparity change and looming in the perception of motion in depth. *Vision Research*, *50* (2), 136–143. doi: 10.1016/j.visres.2009.11.005
- Gordon, M. S. & Rosenblum, L. D. (2005). Effects of intrastimulus modality change on audiovisual time-to-arrival judgments. *Perception & Psychophysics*, *67* (4), 580–594. doi: 10.3758/BF03193516
- Grabot, L. & van Wassenhove, V. (2017). Time Order as Psychological Bias. *Psychological Science*, *28* (5), 670–678. doi: 10.1177/0956797616689369
- Grassi, M. & Casco, C. (2010). Audiovisual bounce-inducing effect: When sound congruence affects grouping in vision. *Attention, Perception, & Psychophysics*, *72* (2), 378–386. doi: 10.3758/APP.72.2.378

- Gröhn, M., Lokki, T. & Takala, T. (2003). Comparison of auditory, visual, and audio-visual navigation in a 3D space. *Proceedings of the 2003 International Conference on Auditory Display, ICAD03*, 200–203. doi: 10.1145/1101530.1101558
- Gurler, D., Doyle, N., Walker, E., Magnotti, J. & Beauchamp, M. (2015). A link between individual differences in multisensory speech perception and eye movements. *Attention, Perception, & Psychophysics*, 77 (4), 1333–1341. doi: 10.3758/s13414-014-0821-1
- Guttman, S. E., Gilroy, L. A. & Blake, R. (2005). Hearing what the eyes see: Auditory encoding of visual temporal sequences. *Psychological Science*, 16 (3), 228–235. doi: 10.1111/j.0956-7976.2005.00808.x
- Hagmann, C. E. & Russo, N. (2016). Multisensory integration of redundant trisensory stimulation. *Attention, Perception, & Psychophysics*, 78 (8), 2558–2568. doi: 10.3758/s13414-016-1192-6
- Hairston, W. D., Wallace, M. T., Vaughan, J. W., Stein, B. E., Norris, J. L. & Schirillo, J. A. (2003). Visual Localization Ability Influences Cross-Modal Bias. *Journal of Cognitive Neuroscience*, 15 (1), 20–29. doi: 10.1162/089892903321107792
- Halpern, S. D., Andrews, T. J. & Purves, D. (1999). Interindividual Variation in Human Visual Performance. *Journal of Cognitive Neuroscience*, 11 (5), 521–534. doi: 10.1162/089892999563580
- Harrison, N. R. (2012). Auditory Motion in Depth is Preferentially ‘Captured’ by Visual Looming Signals. *Seeing and Perceiving*, 25 (1), 71–85. doi: 10.1163/187847611X620928
- Harrison, N. R., Witheridge, S., Makin, A., Wuerger, S. M., Pegna, A. J. & Meyer, G. F. (2015). The effects of stereo disparity on the behavioural and electrophysiological correlates of perception of audio-visual motion in depth. *Neuropsychologia*, 78, 51–62. doi: 10.1016/j.neuropsychologia.2015.09.023
- Harrison, N. R., Wuerger, S. M. & Meyer, G. F. (2010). Reaction time facilitation for horizontally moving auditory–visual stimuli. *Journal of Vision*, 10 (14), 16. doi: 10.1167/10.14.16
- Hassan, S. E. (2012). Are normally sighted, visually impaired, and blind pedestrians accurate and reliable at making street crossing decisions? *Investigative Ophthalmology & Visual Science*, 53 (6), 2593–2600. doi: 10.1167/iovs.11-9340
- Hecht, D., Reiner, M. & Karni, A. (2008). Multisensory enhancement: gains in choice and in simple response times. *Experimental Brain Research*, 189 (2), 133–143. doi: 10.1007/s00221-008-1410-0
- Hecht, H. & Savelsbergh, G. J. P. (2004). Chapter 1. Theories of Time-to-Contact Judgment. *Advances in Psychology*, 135, 1–11. doi: 10.1016/S0166-4115(04)80003-7
- Helbig, H. B. & Ernst, M. O. (2007). Optimal integration of shape information from vision and touch. *Experimental Brain Research*, 179 (4), 595–606. doi: 10.1007/s00221-006-0814-y
- Helbig, H. B. & Ernst, M. O. (2008). Visual-haptic cue weighting is independent of modality-specific attention. *Journal of Vision*, 8 (1), 21. doi: 10.1167/8.1.21

- Helbig, H. B., Ernst, M. O., Ricciardi, E., Pietrini, P., Thielscher, A., Mayer, K. M., . . . Noppeney, U. (2012). The neural mechanisms of reliability weighted integration of shape information from vision and touch. *NeuroImage*, *60* (2), 1063–1072. doi: 10.1016/j.neuroimage.2011.09.072
- Heng, P.-A., Cheng, C.-Y., Wong, T.-T., Xu, Y., Chui, Y.-P., Chan, K.-M. & Tso, S.-K. (2004). A Virtual-Reality Training System for Knee Arthroscopic Surgery. *IEEE Transactions on Information Technology in Biomedicine*, *8* (2), 217–227. doi: 10.1109/TITB.2004.826720
- Heron, J., Whitaker, D. & McGraw, P. V. (2004). Sensory uncertainty governs the extent of audio-visual interaction. *Vision Research*, *44* (25), 2875–2884. doi: 10.1016/j.visres.2004.07.001
- Hidaka, S., Manaka, Y., Teramoto, W., Sugita, Y., Miyauchi, R., Gyoba, J., . . . Iwaya, Y. (2009). Alternation of Sound Location Induces Visual Motion Perception of a Static Object. *PLoS ONE*, *4* (12), e8188. doi: 10.1371/journal.pone.0008188
- Hidaka, S., Teramoto, W. & Sugita, Y. (2015). Spatiotemporal Processing in Crossmodal Interactions for Perception of the External World: A Review. *Frontiers in Integrative Neuroscience*, *9*, 62. doi: 10.3389/fnint.2015.00062
- Hidaka, S., Teramoto, W., Sugita, Y., Manaka, Y., Sakamoto, S. & Suzuki, Y. (2011). Auditory Motion Information Drives Visual Motion Perception. *PLoS ONE*, *6* (3), e17499. doi: 10.1371/journal.pone.0017499
- Ho, C., Reed, N. & Spence, C. (2007). Multisensory in-car warning signals for collision avoidance. *Human Factors*, *49* (6), 1107–1114. doi: 10.1518/001872007X249965
- Ho, C. & Spence, C. (2005). Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. *Journal of Experimental Psychology: Applied*, *11* (3), 157–174. doi: 10.1037/1076-898X.11.3.157
- Hofbauer, M., Wuerger, S. M., Meyer, G. F., Roehrbein, F., Schill, K. & Zetsche, C. (2004). Catching audiovisual mice: Predicting the arrival time of auditory-visual motion signals. *Cognitive, Affective, & Behavioral Neuroscience*, *4* (2), 241–250. doi: 10.3758/CABN.4.2.241
- Hogeboom, M. & van Leeuwen, C. (1997). Visual search strategy and perceptual organization covary with individual preference and structural complexity. *Acta Psychologica*, *95* (2), 141–164. doi: 10.1016/S0001-6918(96)00049-2
- Hong, S. W. & Shim, W. M. (2016). When audiovisual correspondence disturbs visual processing. *Experimental Brain Research*, *234* (5), 1325–1332. doi: 10.1007/s00221-016-4591-y
- Hosking, S. G. & Crassini, B. (2010). The effects of familiar size and object trajectories on time-to-contact judgements. *Experimental Brain Research*, *203* (3), 541–552. doi: 10.1007/s00221-010-2258-7
- Huber, S. & Krist, H. (2004). When is the ball going to hit the ground? Duration estimates, eye movements, and mental imagery of object motion. *Journal of Experimental Psychology: Human Perception and Performance*, *30* (3), 431–444. doi: 10.1037/0096-1523.30.3.431

- Hulme, O. J. & Zeki, S. (2006). The Sightless View: Neural Correlates of Occluded Objects. *Cerebral Cortex*, *17* (5), 1197–1205. doi: 10.1093/cercor/bhl031
- Ilkowska, M. & Engle, R. W. (2010). Trait and State Differences in Working Memory Capacity. In A. Gruszka, G. Matthews & B. Szymura (Hrsg.), *Handbook of Individual Differences in Cognition* (S. 295–320). New York: Springer. doi: 10.1007/978-1-4419-1210-7_18
- Imura, T., Shirai, N., Tomonaga, M., Yamaguchi, M. K. & Yagi, A. (2008). Asymmetry in the perception of motion in depth induced by moving cast shadows. *Journal of Vision*, *8* (13), 10. doi: 10.1167/8.13.10
- Ipser, A., Agolli, V., Bajraktari, A., Al-Alawi, F., Djaafara, N. & Freeman, E. D. (2017). Sight and sound persistently out of synch: stable individual differences in audiovisual synchronisation revealed by implicit measures of lip-voice integration. *Scientific Reports*, *7*, 46413. doi: 10.1038/srep46413
- Irons, J. L. & Leber, A. B. (2016). Choosing attentional control settings in a dynamically changing environment. *Attention, Perception, & Psychophysics*, *78* (7), 2031–2048. doi: 10.3758/s13414-016-1125-4
- Jaekl, P. M. & Harris, L. R. (2009). Sounds can affect visual perception mediated primarily by the parvocellular pathway. *Visual Neuroscience*, *26* (5-6), 477–486. doi: 10.1017/S0952523809990289
- Jain, A., Sally, S. L. & Papathomas, T. V. (2008). Audiovisual short-term influences and aftereffects in motion: Examination across three sets of directional pairings. *Journal of Vision*, *8* (15), 7. doi: 10.1167/8.15.7
- Johnson, D. M., Watson, C. S. & Jensen, J. K. (1986). Individual differences in auditory capabilities. I. *The Journal of the Acoustical Society of America*, *81* (2), 427–438. doi: 10.1121/1.394907
- Jonikaitis, D., Deubel, H. & de'Sperati, C. (2009). Time gaps in mental imagery introduced by competing saccadic tasks. *Vision Research*, *49* (17), 2164–2175. doi: 10.1016/j.visres.2009.05.021
- Kaas, A., Weigelt, S., Roebroek, A., Kohler, A. & Muckli, L. (2010). Imagery of a moving object: The role of occipital cortex and human MT/V5+. *NeuroImage*, *49* (1), 794–804. doi: 10.1016/j.neuroimage.2009.07.055
- Kaganovich, N. & Schumaker, J. (2016). Electrophysiological correlates of individual differences in perception of audiovisual temporal asynchrony. *Neuropsychologia*, *86*, 119–130. doi: 10.1016/j.neuropsychologia.2016.04.015
- Katz, L. N., Hennig, J. A., Cormack, L. K. & Huk, A. C. (2015). A Distinct Mechanism of Temporal Integration for Motion through Depth. *The Journal of Neuroscience*, *35* (28), 10212–10216. doi: 10.1523/JNEUROSCI.0032-15.2015
- Kawachi, Y. & Gyoba, J. (2006). A new response-time measure of object persistence in the tunnel effect. *Acta Psychologica*, *123* (1-2), 73–90. doi: 10.1016/j.actpsy.2006.04.003
- Kelava, A. & Schermelleh-Engel, K. (2012). Latent-State-Trait-Theorie (LST-Theorie). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 363–381). Berlin, Heidelberg: Springer.

- Kersten, D., Mamassian, P. & Knill, D. C. (1997). Moving cast shadows induce apparent motion in depth. *Perception*, *26* (2), 171–192. doi: 10.1068/p260171
- Keshavarz, B., Campos, J. L., DeLucia, P. R. & Oberfeld, D. (2017). Estimating the relative weights of visual and auditory tau versus heuristic-based cues for time-to-contact judgments in realistic, familiar scenes by older and younger adults. *Attention, Perception, & Psychophysics*, *79* (3), 929–944. doi: 10.3758/s13414-016-1270-9
- Khuu, S. K., Gordon, J., Balcomb, K. & Kim, J. (2014). The perception of three-dimensional cast-shadow structure is dependent on visual awareness. *Journal of Vision*, *14* (3), 25. doi: 10.1167/14.3.25
- Kidd, G. R., Watson, C. S. & Gygi, B. (2007). Individual differences in auditory abilities. *The Journal of the Acoustical Society of America*, *122* (1), 418–435. doi: 10.1121/1.2743154
- Kim, N.-G. & Grocki, M. J. (2006). Multiple sources of information and time-to-contact judgments. *Vision Research*, *46* (12), 1946–1958. doi: 10.1016/j.visres.2005.12.011
- Kim, R. S., Seitz, A. R. & Shams, L. (2008). Benefits of Stimulus Congruency for Multisensory Facilitation of Visual Learning. *PLoS ONE*, *3* (1), e1532. doi: 10.1371/journal.pone.0001532
- Kirkham, N. Z., Wagner, J. B., Swan, K. A. & Johnson, S. P. (2012). Sound support: Intermodal information facilitates infants' perception of an occluded trajectory. *Infant Behavior and Development*, *35* (1), 174–178. doi: 10.1016/j.infbeh.2011.09.001
- Kitagawa, N. & Ichihara, S. (2002). Hearing visual motion in depth. *Nature*, *416* (6877), 172–174. doi: 10.1038/416172a
- Kitajima, N. & Yamashita, Y. (1999). Dynamic capture of sound motion by light stimuli moving in three-dimensional space. *Perceptual and Motor Skills*, *89* (3), 1139–1158. doi: 10.2466/pms.1999.89.3f.1139
- Kleiner, M., Brainard, D. H. & Pelli, D. G. (2007). What's new in Psychtoolbox-3? *Perception*, *36*, ECVF Abstract Supplement.
- Kober, S. E., Kurzmann, J. & Neuper, C. (2012). Cortical correlate of spatial presence in 2D and 3D interactive virtual reality: An EEG study. *International Journal of Psychophysiology*, *83* (3), 365–374. doi: 10.1016/j.ijpsycho.2011.12.003
- Koppen, C., Levitan, C. A. & Spence, C. (2009). A signal detection study of the Colavita visual dominance effect. *Experimental Brain Research*, *196* (3), 353–360. doi: 10.1007/s00221-009-1853-y
- Koppen, C. & Spence, C. (2007). Audiovisual asynchrony modulates the Colavita visual dominance effect. *Brain Research*, *1186*, 224–232. doi: 10.1016/j.brainres.2007.09.076
- Landwehr, K., Baurès, R., Oberfeld, D. & Hecht, H. (2013). Visual discrimination thresholds for time to arrival. *Attention, Perception, & Psychophysics*, *75* (7), 1465–1472. doi: 10.3758/s13414-013-0497-y
- Laurienti, P. J., Kraft, R. A., Maldjian, J. A., Burdette, J. H. & Wallace, M. T. (2004). Semantic congruence is a critical factor in multisensory behavioral performance. *Experimental Brain Research*, *158* (4), 405–414. doi: 10.1007/s00221-004-1913-2

- Law, D. J., Pellegrino, J. W., Mitchell, S. R., Fischer, S. C., McDonald, T. P. & Hunt, E. B. (1993). Perceptual and Cognitive Factors Governing Performance in Comparative Arrival-Time Judgments. *Journal of Experimental Psychology: Human Perception and Performance*, *19* (6), 1183–1199. doi: 10.1037//0096-1523.19.6.1183
- Leber, A. B. (2010). Neural Predictors of Within-Subject Fluctuations in Attentional Control. *The Journal of Neuroscience*, *30* (34), 11458–11465. doi: 10.1523/JNEUROSCI.0809-10.2010
- Leber, A. B., Turk-Browne, N. B. & Chun, M. M. (2008). Neural predictors of moment-to-moment fluctuations in cognitive flexibility. *Proceedings of the National Academy of Sciences*, *105* (36), 13592–13597. doi: 10.1073/pnas.0805423105
- Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, *5* (4), 437–459. doi: 10.1068/p050437
- Lee, H. & Noppeney, U. (2014). Music expertise shapes audiovisual temporal integration windows for speech, sinewave speech, and music. *Frontiers in Psychology*, *5*, 868. doi: 10.3389/fpsyg.2014.00868
- Leone, L. M. & McCourt, M. E. (2013). The roles of physical and physiological simultaneity in audiovisual multisensory facilitation. *i-Perception*, *4* (4), 213–228. doi: 10.1068/i0532
- Lewis, R. & Noppeney, U. (2010). Audiovisual Synchrony Improves Motion Discrimination via Enhanced Connectivity between Early Visual and Auditory Areas. *The Journal of Neuroscience*, *30* (37), 12329–12339. doi: 10.1523/JNEUROSCI.5745-09.2010
- Luck, S. J. & Hillyard, S. A. (1995). The role of attention in feature detection and conjunction discrimination: An electrophysiological analysis. *International Journal of Neuroscience*, *80* (1-4), 281–297. doi: 10.3109/00207459508986105
- Lutgheid, A. J. & Welchman, A. E. (2011). Evaluating methods to measure time-to-contact. *Vision Research*, *51* (20), 2234–2241. doi: 10.1016/j.visres.2011.08.019
- Lukas, S., Philipp, A. M. & Koch, I. (2014). Crossmodal attention switching: Auditory dominance in temporal discrimination tasks. *Acta Psychologica*, *153*, 139–146. doi: 10.1016/j.actpsy.2014.10.003
- Lupiáñez, J., Milliken, B., Solano, C., Weaver, B. & Tipper, S. P. (2001). On the strategic modulation of the time course of facilitation and inhibition of return. *The Quarterly Journal of Experimental Psychology A*, *54* (3), 753–773. doi: 10.1080/02724980042000453
- Lyon, D. R. & Waag, W. L. (1995). Time course of visual extrapolation accuracy. *Acta Psychologica*, *89* (3), 239–260. doi: 10.1016/0001-6918(95)98945-Z
- Macaluso, E., Noppeney, U., Talsma, D., Vercillo, T., Hartcher-O'Brien, J. & Adam, R. (2016). The Curious Incident of Attention in Multisensory Integration: Bottom-up vs. Top-down. *Multisensory Research*, *29* (6-7), 557–583. doi: 10.1163/22134808-00002528
- Macdonald, J. S. P. & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics*, *73* (6), 1780–1789. doi: 10.3758/s13414-011-0144-4
- Macken, W. J., Phelps, F. G. & Jones, D. M. (2009). What causes auditory distraction? *Psychonomic Bulletin & Review*, *16* (1), 139–144. doi: 10.3758/PBR.16.1.139

- Maeda, F., Kanai, R. & Shimojo, S. (2004). Changing pitch induced visual motion illusion. *Current Biology*, 14 (23), R990-R991. doi: 10.1016/j.cub.2004.11.018
- Magnotti, J. F. & Beauchamp, M. S. (2015). The noisy encoding of disparity model of the McGurk effect. *Psychonomic Bulletin & Review*, 22 (3), 701–709. doi: 10.3758/s13423-014-0722-2
- Maiworm, M. & Röder, B. (2011). Suboptimal auditory dominance in audiovisual integration of temporal cues. *Tsinghua Science and Technology*, 16 (2), 121–132. doi: 10.1016/S1007-0214(11)70019-0
- Makin, A. D. J. & Bertamini, M. (2014). Do different types of dynamic extrapolation rely on the same mechanism? *Journal of Experimental Psychology: Human Perception and Performance*, 40 (4), 1566–1579. doi: 10.1037/a0036680
- Makin, A. D. J. & Poliakoff, E. (2011). Do common systems control eye movements and motion extrapolation? *Quarterly Journal of Experimental Psychology*, 64 (7), 1327–1343. doi: 10.1080/17470218.2010.548562
- Makin, A. D. J., Poliakoff, E., Chen, J. & Stewart, A. J. (2008). The effect of previously viewed velocities on motion extrapolation. *Vision Research*, 48 (18), 1884–1893. doi: 10.1016/j.visres.2008.05.023
- Makin, A. D. J., Stewart, A. J. & Poliakoff, E. (2009). Typical object velocity influences motion extrapolation. *Experimental Brain Research*, 193 (1), 137–142. doi: 10.1007/s00221-008-1678-0
- Mallick, D. B., Magnotti, J. F. & Beauchamp, M. S. (2015). Variability and stability in the McGurk effect: contributions of participants, stimuli, time, and response type. *Psychonomic Bulletin & Review*, 22 (5), 1299–1307. doi: 10.3758/s13423-015-0817-4
- Mamassian, P., Knill, D. C. & Kersten, D. (1998). The perception of cast shadows. *Trends in Cognitive Sciences*, 2 (8), 288–295. doi: 10.1016/S1364-6613(98)01204-2
- Marois, R. & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences*, 9 (6), 296–305. doi: 10.1016/j.tics.2005.04.010
- Mather, G. (2009). Depth Perception. In G. Mather (Hrsg.), *Foundations of Sensation and Perception* (S. 296–325). Hove: Psychology Press.
- Mays, A. & Schirillo, J. (2005). Lights can reverse illusory directional hearing. *Neuroscience Letters*, 384 (3), 336–338. doi: 10.1016/j.neulet.2005.05.021
- McIntire, J. P., Havig, P. R. & Geiselman, E. E. (2012). What is 3D good for? A review of human performance on stereoscopic 3D displays. In *Head- and Helmet-Mounted Displays XVII; and Display Technologies and Applications for Defense, Security, and Avionics VI* (Bd. 2012, S. 83830X). SPIE. doi: 10.1117/12.920017
- Mégevand, P., Molholm, S., Nayak, A. & Foxe, J. J. (2013). Recalibration of the Multisensory Temporal Window of Integration Results from Changing Task Demands. *PLoS ONE*, 8 (8), e71608. doi: 10.1371/journal.pone.0071608
- Menceloglu, M., Grabowecky, M. & Suzuki, S. (2017). Temporal expectation weights visual signals over auditory signals. *Psychonomic Bulletin & Review*, 24 (2), 416–422. doi: 10.3758/s13423-016-1109-3

- Meng, F., Li, S., Cao, L., Peng, Q., Li, M., Wang, C. & Zhang, W. (2016). Designing Fatigue Warning Systems: The perspective of professional drivers. *Applied Ergonomics*, *53*, 122–130. doi: 10.1016/j.apergo.2015.08.003
- Meredith, M. A., Nemitz, J. W. & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *The Journal of Neuroscience*, *7* (10), 3215–3229.
- Meredith, M. A. & Stein, B. E. (1986a). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research*, *365* (2), 350–354. doi: 10.1016/0006-8993(86)91648-3
- Meredith, M. A. & Stein, B. E. (1986b). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, *56* (3), 640–662. doi: 10.1152/jn.1986.56.3.640
- Meyer, G. F. & Wuerger, S. M. (2001). Cross-modal integration of auditory and visual motion signals. *NeuroReport*, *12* (11), 2557–2560. doi: 10.1097/00001756-200108080-00053
- Meyer, G. F., Wuerger, S. M., Röhrbein, F. & Zetsche, C. (2005). Low-level integration of auditory and visual motion signals requires spatial co-localisation. *Experimental Brain Research*, *166* (3-4), 538–547. doi: 10.1007/s00221-005-2394-7
- Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14* (2), 247–279. doi: 10.1016/0010-0285(82)90010-X
- Molholm, S., Ritter, W., Javitt, D. C. & Foxe, J. J. (2004). Multisensory Visual–Auditory Object Recognition in Humans: a High-density Electrical Mapping Study. *Cerebral Cortex*, *14* (4), 452–465. doi: 10.1093/cercor/bhh007
- Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E. & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cognitive Brain Research*, *14* (1), 115–128. doi: 10.1016/S0926-6410(02)00066-6
- Molloy, K., Griffiths, T. D., Chait, M. & Lavie, N. (2015). Inattentive Deafness: Visual Load Leads to Time-Specific Suppression of Auditory Evoked Responses. *The Journal of Neuroscience*, *35* (49), 16046–16054. doi: 10.1523/JNEUROSCI.2931-15.2015
- Mozolic, J. L., Hugenschmidt, C. E., Peiffer, A. M. & Laurienti, P. J. (2008). Modality-specific selective attention attenuates multisensory integration. *Experimental Brain Research*, *184* (1), 39–52. doi: 10.1007/s00221-007-1080-3
- Nath, A. R. & Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *NeuroImage*, *59* (1), 781–787. doi: 10.1016/j.neuroimage.2011.07.024
- Nefs, H. T., O'Hare, L. & Harris, J. M. (2010). Two independent mechanisms for motion-in-depth perception: Evidence from individual differences. *Frontiers in Psychology*, *1*, 155. doi: 10.3389/fpsyg.2010.00155
- Neuhoff, J. G. (2001). An Adaptive Bias in the Perception of Looming Auditory Motion. *Ecological Psychology*, *13* (2), 87–110. doi: 10.1207/S15326969ECO1302_2

- Odegaard, B. & Shams, L. (2016). The Brain's Tendency to Bind Audiovisual Signals Is Stable but Not General. *Psychological Science*, 27 (4), 583–591. doi: 10.1177/0956797616628860
- Odegaard, B., Wozny, D. R. & Shams, L. (2017). A simple and efficient method to enhance audiovisual binding tendencies. *PeerJ*, 5, e3143. doi: 10.7717/peerj.3143
- Ogawa, A., Bordier, C. & Macaluso, E. (2013). Audio-Visual Perception of 3D Cinematography: An fMRI Study Using Condition-Based and Computation-Based Analyses. *PLoS ONE*, 8 (10), e76003. doi: 10.1371/journal.pone.0076003
- Ogawa, A. & Macaluso, E. (2013). Audio–visual interactions for motion perception in depth modulate activity in visual area V3A. *NeuroImage*, 71, 158–167. doi: 10.1016/j.neuroimage.2013.01.012
- Oldfield, R. C. (1971). The Assessment and Analysis of Handedness: The Edinburgh Inventory. *Neuropsychologia* (9), 97–113. doi: 10.1016/0028-3932(71)90067-4
- Olson, I. R., Gatenby, J. C., Leung, H.-C., Skudlarski, P. & Gore, J. C. (2003). Neuronal representation of occluded objects in the human brain. *Neuropsychologia*, 42 (1), 95–104. doi: 10.1016/S0028-3932(03)00151-9
- O'Reilly, J., Mesulam, M. M. & Nobre, A. C. (2008). The Cerebellum Predicts the Timing of Perceptual Events. *The Journal of Neuroscience*, 28 (9), 2252–2260. doi: 10.1523/JNEUROSCI.2742-07.2008
- Otten, L. J., Alain, C. & Picton, T. W. (2000). Effects of visual attentional load on auditory processing. *NeuroReport*, 11 (4), 875–880. doi: 10.1097/00001756-200003200-00043
- Park, E., Kim, K. J. & Kwon, S. J. (2017). Evaluation of automobile navigation systems with multisensory information channels. *Behaviour & Information Technology*, 36 (10), 1014–1019. doi: 10.1080/0144929X.2017.1323233
- Park, S. H., Kim, S., Kwon, M. & Christou, E. A. (2016). Differential contribution of visual and auditory information to accurately predict the direction and rotational motion of a visual stimulus. *Applied Physiology, Nutrition, and Metabolism*, 41 (3), 244–248. doi: 10.1139/apnm-2015-0390
- Pratt, J. & Abrams, R. A. (1999). Inhibition of return in discrimination tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 25 (1), 229–242. doi: 10.1037//0096-1523.25.1.229
- Prime, S. L. & Harris, L. R. (2010). Predicting the position of moving audiovisual stimuli. *Experimental Brain Research*, 203 (2), 249–260. doi: 10.1007/s00221-010-2224-4
- Proulx, M. J. (2011). Individual Differences and Metacognitive Knowledge of Visual Search Strategy. *PLoS ONE*, 6 (10), e27043. doi: 10.1371/journal.pone.0027043
- Raveh, D. & Lavie, N. (2015). Load-induced inattentional deafness. *Attention, Perception, & Psychophysics*, 77 (2), 483–492. doi: 10.3758/s13414-014-0776-2
- Recanzone, G. H. (2003). Auditory Influences on Visual Temporal Rate Perception. *Journal of Neurophysiology*, 89 (2), 1078–1093. doi: 10.1152/jn.00706.2002

- Repp, B. H. & Penel, A. (2002). Auditory dominance in temporal processing: New evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 28 (5), 1085–1099. doi: 10.1037//0096-1523.28.5.1085
- Ribeiro, M. J., Paiva, J. S. & Castelo-Branco, M. (2016). Spontaneous Fluctuations in Sensory Processing Predict Within-Subject Reaction Time Variability. *Frontiers in Human Neuroscience*, 10, 200. doi: 10.3389/fnhum.2016.00200
- Richards, D. & Taylor, M. (2015). A Comparison of learning gains when using a 2D simulation tool versus a 3D virtual world: An experiment to find the right representation involving the Marginal Value Theorem. *Computers & Education*, 86, 157–171. doi: 10.1016/j.compedu.2015.03.009
- Riecke, B. E., Våljamäe, A. & Schulte-Pelkum, J. (2009). Moving sounds enhance the visually-induced self-motion illusion (circular vection) in virtual reality. *ACM Transactions on Applied Perception*, 6 (2), 7:1–7:27. doi: 10.1145/1498700.1498701
- Rokers, B., Cormack, L. K. & Huk, A. C. (2009). Disparity- and velocity-based signals for three-dimensional motion perception in human MT+. *Nature Neuroscience*, 12 (8), 1050–1055. doi: 10.1038/nn.2343
- Roper, J. M. & Hassan, S. E. (2014). How Do Vision and Hearing Impact Pedestrian Time-to-Arrival Judgments? *Optometry and Vision Science*, 91 (3), 303–311. doi: 10.1097/OPX.0000000000000161
- Rose, T., Nam, C. S. & Chen, K. B. (2018). Immersion of virtual reality for rehabilitation - Review. *Applied Ergonomics*, 69, 153–161. doi: 10.1016/j.apergo.2018.01.009
- Rosenberg, M. D., Finn, E. S., Constable, R. T. & Chun, M. M. (2015). Predicting moment-to-moment attentional state. *NeuroImage*, 114, 249–256. doi: 10.1016/j.neuroimage.2015.03.032
- Rouder, J. N., Morey, R. D., Speckman, P. L. & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56 (5), 356–374. doi: 10.1016/j.jmp.2012.08.001
- Rummukainen, O. & Mendonça, C. (2016). Reproducing Reality: Multimodal Contributions in Natural Scene Discrimination. *ACM Transactions on Applied Perception*, 14 (1), 1–19. doi: 10.1145/2915917
- Sakano, Y. & Allison, R. S. (2014). Aftereffect of motion-in-depth based on binocular cues: Effects of adaptation duration, interocular correlation, and temporal correlation. *Journal of Vision*, 14 (8), 21. doi: 10.1167/14.8.21
- Sakano, Y., Allison, R. S. & Howard, I. P. (2012). Motion aftereffect in depth based on binocular information. *Journal of Vision*, 12 (1), 11. doi: 10.1167/12.1.11
- Sanabria, D., Correa, Á., Lupiáñez, J. & Spence, C. (2004). Bouncing or streaming? Exploring the influence of auditory cues on the interpretation of ambiguous visual motion. *Experimental Brain Research*, 157 (4), 537–541. doi: 10.1007/s00221-004-1993-z

- Sanabria, D., Lupiáñez, J. & Spence, C. (2007). Auditory motion affects visual motion perception in a speeded discrimination task. *Experimental Brain Research*, 178 (3), 415–421. doi: 10.1007/s00221-007-0919-y
- Sapir, A., d'Avossa, G., McAvoy, M., Shulman, G. L. & Corbetta, M. (2005). Brain signals for spatial attention predict performance in a motion discrimination task. *Proceedings of the National Academy of Sciences*, 102 (49), 17810–17815. doi: 10.1073/pnas.0504678102
- Sato, M. & Howard, I. P. (2001). Effects of disparity–perspective cue conflict on depth contrast. *Vision Research*, 41 (4), 415–426. doi: 10.1016/S0042-6989(00)00272-8
- Schiff, W. & Oldak, R. (1990). Accuracy of Judging Time to Arrival: Effects of Modality, Trajectory, and Gender. *Journal of Experimental Psychology: Human Perception and Performance*, 16 (2), 303–316. doi: 10.1037/0096-1523.16.2.303
- Schmiedchen, K., Freigang, C., Nitsche, I. & Rübsemann, R. (2012). Crossmodal interactions and multisensory integration in the perception of audio-visual motion — A free-field study. *Brain Research*, 1466, 99–111. doi: 10.1016/j.brainres.2012.05.015
- Scholl, B. J. (2007). Object Persistence in Philosophy and Psychology. *Mind & Language*, 22 (5), 563–591. doi: 10.1111/j.1468-0017.2007.00321.x
- Scholl, B. J. & Pylyshyn, Z. W. (1999). Tracking Multiple Items Through Occlusion: Clues to Visual Objecthood. *Cognitive Psychology*, 38, 259–290. doi: 10.1006/cogp.1998.0698
- Sella, I., Reiner, M. & Pratt, H. (2014). Natural stimuli from three coherent modalities enhance behavioral responses and electrophysiological cortical activity in humans. *International Journal of Psychophysiology*, 93 (1), 45–55. doi: 10.1016/j.ijpsycho.2013.11.003
- Senkowski, D., Saint-Amour, D., Kelly, S. P. & Foxe, J. J. (2007). Multisensory processing of naturalistic objects in motion: A high-density electrical mapping and source estimation study. *NeuroImage*, 36 (3), 877–888. doi: 10.1016/j.neuroimage.2007.01.053
- Shore, D. I., Barnes, M. E. & Spence, C. (2006). Temporal aspects of the visuotactile congruency effect. *Neuroscience Letters*, 392 (1-2), 96–100. doi: 10.1016/j.neulet.2005.09.001
- Shuwairi, S. M., Curtis, C. E. & Johnson, S. P. (2007). Neural substrates of dynamic object occlusion. *Journal of Cognitive Neuroscience*, 19 (8), 1275–1285. doi: 10.1162/jocn.2007.19.8.1275
- Sinnett, S., Soto-Faraco, S. & Spence, C. (2008). The co-occurrence of multisensory competition and facilitation. *Acta Psychologica*, 128 (1), 153–161. doi: 10.1016/j.actpsy.2007.12.002
- Slobounov, S. M., Ray, W., Johnson, B., Slobounov, E. & Newell, K. M. (2015). Modulation of cortical activity in 2D versus 3D virtual reality environments: An EEG study. *International Journal of Psychophysiology*, 95 (3), 254–260. doi: 10.1016/j.ijpsycho.2014.11.003
- Snow, J. C., Pettypiece, C. E., McAdam, T. D., McLean, A. D., Stroman, P. W., Goodale, M. A. & Culham, J. C. (2011). Bringing the real world into the fMRI scanner: Repetition effects for pictures versus real objects. *Scientific Reports*, 1, 130. doi: 10.1038/srep00130
- Sörqvist, P. & Rönnerberg, J. (2014). Individual differences in distractibility: An update and a model. *PsyCh Journal*, 3 (1), 42–57. doi: 10.1002/pchj.47

- Soto-Faraco, S., Kingstone, A. & Spence, C. (2003). Multisensory contributions to the perception of motion. *Neuropsychologia*, *41* (13), 1847–1862. doi: 10.1016/S0028-3932(03)00185-4
- Soto-Faraco, S., Lyons, J., Gazzaniga, M., Spence, C. & Kingstone, A. (2002). The ventriloquist in motion: Illusory capture of dynamic information across sensory modalities. *Brain Research. Cognitive Brain Research*, *14* (1), 139–146. doi: 10.1016/S0926-6410(02)00068-X
- Soto-Faraco, S., Spence, C. & Kingstone, A. (2004). Cross-Modal Dynamic Capture: Congruency Effects in the Perception of Motion Across Sensory Modalities. *Journal of Experimental Psychology: Human Perception and Performance*, *30* (2), 330–345. doi: 10.1037/0096-1523.30.2.330
- Soto-Faraco, S., Spence, C., Lloyd, D. & Kingstone, A. (2004). Moving Multisensory Research Along Motion Perception Across Sensory Modalities. *Current Directions in Psychological Science*, *13* (1), 29–32.
- Spence, C. (2007). Audiovisual multisensory integration. *Acoustical Science and Technology*, *28* (2), 61–70. doi: 10.1250/ast.28.61
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73* (4), 971–995. doi: 10.3758/s13414-010-0073-7
- Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Annals of the New York Academy of Sciences*, *1296*, 31–49. doi: 10.1111/nyas.12121
- Spence, C. & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, *59* (1), 1–22. doi: 10.3758/BF03206843
- Spence, C., Parise, C. & Chen, Y.-C. (2012). The Colavita Visual Dominance Effect. In M. M. Murray & M. T. Wallace (Hrsg.), *The Neural Bases of Multisensory Processes*. Boca Raton (FL): CRC Press/Taylor & Francis.
- 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
- Starke, J., Ball, F., Heinze, H.-J. & Noesselt, T. (2017). The spatio-temporal profile of multisensory integration. *The European Journal of Neuroscience*, Elektronische Veröffentlichung vor dem Druck. Zugriff am 24.05.2018 auf <https://onlinelibrary.wiley.com/doi/abs/10.1111/ejn.13753> doi: 10.1111/ejn.13753
- 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. doi: 10.1038/nrn2331
- Stein, T., Zwicker, J., Ritter, J., Kitzmantel, M. & Schneider, W. X. (2009). The effect of fearful faces on the attentional blink is task dependent. *Psychonomic Bulletin & Review*, *16* (1), 104–109. doi: 10.3758/PBR.16.1.104

- Stevenson, R. A. & Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Experimental Brain Research*, *227* (2), 249–261. doi: 10.1007/s00221-013-3507-3
- Stevenson, R. A., Zemtsov, R. K. & Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental Psychology: Human Perception and Performance*, *38* (6), 1517–1529. doi: 10.1037/a0027339
- Strand, J., Cooperman, A., Rowe, J. & Simenstad, A. (2014). Individual differences in susceptibility to the McGurk effect: Links with lipreading and detecting audiovisual incongruity. *Journal of Speech, Language, and Hearing Research*, *57* (6), 2322–2331. doi: 10.1044/2014_JSLHR-H-14-0059
- Suied, C., Bonneel, N. & Viaud-Delmon, I. (2009). Integration of auditory and visual information in the recognition of realistic objects. *Experimental Brain Research*, *194* (1), 91–102. doi: 10.1007/s00221-008-1672-6
- Sun, Y., Hickey, T. J., Shinn-Cunningham, B. & Sekuler, R. (2017). Catching Audiovisual Interactions With a First-Person Fisherman Video Game. *Perception*, *46* (7), 793–814. doi: 10.1177/0301006616682755
- Takeuchi, T., Yoshimoto, S., Shimada, Y., Kochiyama, T. & Kondo, H. M. (2017). Individual differences in visual motion perception and neurotransmitter concentrations in the human brain. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372* (1714), 20160111. doi: 10.1098/rstb.2016.0111
- 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
- Taya, S. & Miura, K. (2010). Cast shadow can modulate the judged final position of a moving target. *Attention, Perception, & Psychophysics*, *72* (7), 1930–1937. doi: 10.3758/APP.72.7.1930
- Teder-Sälejärvi, W. A., McDonald, J. J., Di Russo, F. & Hillyard, S. A. (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
- Tellinghuisen, D. J. & Nowak, E. J. (2003). The inability to ignore auditory distractors as a function of visual task perceptual load. *Perception & Psychophysics*, *65* (5), 817–828. doi: 10.3758/BF03194817
- Teramoto, W., Hidaka, S., Gyoba, J. & Suzuki, Y. (2010). Auditory temporal cues can modulate visual representational momentum. *Attention, Perception, & Psychophysics*, *72* (8), 2215–2226. doi: 10.3758/APP.72.8.2215
- Teramoto, W., Manaka, Y., Hidaka, S., Sugita, Y., Miyauchi, R., Sakamoto, S., . . . Suzuki, Y. (2010). Visual motion perception induced by sounds in vertical plane. *Neuroscience Letters*, *479* (3), 221–225. doi: 10.1016/j.neulet.2010.05.065

- Toet, A., Houtkamp, J. M. & van der Meulen, R. (2013). Visual and Auditory Cue Effects on Risk Assessment in a Highway Training Simulation. *Simulation & Gaming, 44* (5), 732–753. doi: 10.1177/1046878113495349
- Tremblay, C., Champoux, F., Bacon, B. A. & Théoret, H. (2007). Evidence for a Generic Process Underlying Multisensory Integration. *The Open Behavioral Science Journal, 1* (1), 1–4. doi: 10.2174/187423000701011000
- Tresilian, J. R. (1995). Perceptual and cognitive processes in time-to-contact estimation: Analysis of prediction-motion and relative judgment tasks. *Perception & Psychophysics, 57* (2), 231–245. doi: 10.3758/BF03206510
- Tresilian, J. R. (1999). Visually timed action: Time-out for 'tau'? *Trends in Cognitive Sciences, 3* (8), 301–310. doi: 10.1016/S1364-6613(99)01352-2
- Triesch, J., Ballard, D. H., Hayhoe, M. M. & Sullivan, B. T. (2003). What you see is what you need. *Journal of Vision, 3* (1), 86–94. doi: 10.1167/3.1.9
- Tyll, S., Bonath, B., Schoenfeld, M. A., Heinze, H.-J., Ohl, F. W. & Noesselt, T. (2013). Neural basis of multisensory looming signals. *NeuroImage, 65*, 13–22. doi: 10.1016/j.neuroimage.2012.09.056
- van Wassenhove, V., Grant, K. W. & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia, 45* (3), 598–607. doi: 10.1016/j.neuropsychologia.2006.01.001
- Vogel, E. K. & Awh, E. (2008). How to Exploit Diversity for Scientific Gain. *Current Directions in Psychological Science, 17* (2), 171–176. doi: 10.1111/j.1467-8721.2008.00569.x
- von Saldern, S. & Noppeney, U. (2013). Sensory and Striatal Areas Integrate Auditory and Visual Signals into Behavioral Benefits during Motion Discrimination. *The Journal of Neuroscience, 33* (20), 8841–8849. doi: 10.1523/JNEUROSCI.3020-12.2013
- Wahn, B. & König, P. (2017). Is Attentional Resource Allocation Across Sensory Modalities Task-Dependent? *Advances in Cognitive Psychology, 13* (1), 83–96. doi: 10.5709/acp-0209-2
- Welch, R. B., DuttonHurt, L. D. & Warren, D. H. (1986). Contributions of audition and vision to temporal rate perception. *Perception & Psychophysics, 39* (4), 294–300. doi: 10.3758/BF03204939
- Welch, R. B. & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88* (3), 638–667. doi: 10.1037//0033-2909.88.3.638
- Witten, I. B. & Knudsen, E. I. (2005). Why Seeing Is Believing: Merging Auditory and Visual Worlds. *Neuron, 48* (3), 489–496. doi: 10.1016/j.neuron.2005.10.020
- Wozny, D. R., Beierholm, U. R. & Shams, L. (2010). Probability Matching as a Computational Strategy Used in Perception. *PLoS Computational Biology, 6* (8), e1000871. doi: 10.1371/journal.pcbi.1000871
- Wuerger, S. M., Hofbauer, M. & Meyer, G. F. (2003). The integration of auditory and visual motion signals at threshold. *Perception & Psychophysics, 65* (8), 1188–1196. doi: 10.3758/BF03194844

-
- Wuerger, S. M., Meyer, G. F., Hofbauer, M., Zetsche, C. & Schill, K. (2010). Motion extrapolation of auditory–visual targets. *Information Fusion*, *11* (1), 45–50. doi: 10.1016/j.inffus.2009.04.005
- Yakimoff, N., Mateeff, S., Ehrenstein, W. H. & Hohnsbein, J. (1993). Motion Extrapolation Performance: A Linear Model Approach. *Human Factors*, *35* (3), 501–510. doi: 10.1177/001872089303500307
- Zago, M., Iosa, M., Maffei, V. & Lacquaniti, F. (2010). Extrapolation of vertical target motion through a brief visual occlusion. *Experimental Brain Research*, *201* (3), 365–384. doi: 10.1007/s00221-009-2041-9
- Zalevski, A. M., Henning, G. B. & Hill, N. J. (2007). Cue combination and the effect of horizontal disparity and perspective on stereoacuity. *Spatial Vision*, *20* (1–2), 107–138. doi: 10.1163/156856807779369706
- Zannoli, M., Cass, J., Mamassian, P. & Alais, D. (2012). Synchronized Audio-Visual Transients Drive Efficient Visual Search for Motion-in-Depth. *PLoS ONE*, *7* (5), e37190. doi: 10.1371/journal.pone.0037190
- Zhou, F., Wong, V. & Sekuler, R. (2007). Multi-sensory integration of spatio-temporal segmentation cues: one plus one does not always equal two. *Experimental Brain Research*, *180* (4), 641–654. doi: 10.1007/s00221-007-0897-0
- Zhou, L., Yan, J., Liu, Q., Li, H., Xie, C., Wang, Y., ... Sun, H.-J. (2007). Visual and Auditory Information Specifying an Impending Collision of an Approaching Object. In J. A. Jacko (Hrsg.), *Human-Computer Interaction: Interaction Platforms and Techniques (HCI International 2007, Beijing, China)* (S. 720–729). Berlin: Springer. doi: 10.1007/978-3-540-73107-8_80
- Zmigrod, L. & Zmigrod, S. (2016). On the Temporal Precision of Thought: Individual Differences in the Multisensory Temporal Binding Window Predict Performance on Verbal and Nonverbal Problem Solving Tasks. *Multisensory Research*, *29* (8), 679–701. doi: 10.1163/22134808-00002532

A Reaktionszeit-Differenzen für die Nutzergruppeneinteilung

Tabelle A1

Auditorische Reaktionszeit-Differenzen für beide Schätzdistanzen in Experiment 2 auf Grundlage gerader Blöcke

Proband	RT-Differenz auditorisch nah ^a	RT-Differenz auditorisch fern ^b	Nutzergruppe ^c
1	+0.0920	+0.0369	Nichtnutzer
2	+0.2747	-0.0878	Ferntonnutzer
3	-0.0861	-0.0752	Beidtonnutzer
4	-0.0878	+0.0340	Nahtonnutzer
5	+0.0462	-0.0286	Ferntonnutzer
6	+0.0982	-0.0423	Ferntonnutzer
7	+0.0572	-0.0494	Ferntonnutzer
8	-0.0368	+0.2020	Nahtonnutzer
9	+0.1528	+0.0016	Nichtnutzer
10	-0.0069	+0.0189	Nahtonnutzer
11	-0.0564	-0.0451	Beidtonnutzer
12	-0.1313	+0.0763	Nahtonnutzer
13	-0.0453	+0.0412	Nahtonnutzer
14	-0.0506	+0.1559	Nahtonnutzer
15	+0.0889	+0.0661	Nichtnutzer
16	+0.1349	-0.1082	Ferntonnutzer
17	-0.0192	+0.2323	Nahtonnutzer
18	+0.0381	-0.0566	Ferntonnutzer
19	-0.0496	+0.0722	Nahtonnutzer
20	+0.0846	-0.0644	Ferntonnutzer
21	-0.1104	+0.1518	Nahtonnutzer
22	-0.0010	-0.0507	Beidtonnutzer
23	+0.1322	+0.1675	Nichtnutzer
24	+0.1471	-0.1811	Ferntonnutzer
25	-0.0167	+0.2227	Nahtonnutzer
26	-0.0343	-0.0748	Beidtonnutzer
27	+0.0647	+0.1638	Nichtnutzer
28	+0.0598	-0.1502	Ferntonnutzer
29	-0.0405	+0.0537	Nahtonnutzer
30	+0.0612	-0.0258	Ferntonnutzer
31	+0.0223	+0.1397	Nichtnutzer
32	-0.0445	-0.0203	Beidtonnutzer

Anmerkungen. Reaktionszeit (RT) in Sekunden (Ausgangswert = Absolute Abweichung von der tatsächlichen Ankunftszeit). Berechnung und Kategorisierung auf Grundlage gerader Blöcke, im Anschluss Analyse der RT-Abweichungen mit Daten der ungeraden Blöcke.

^a Differenz der Bedingungen auditorisch passend und unimodal für die nahe Schätzdistanz. Ein negativer Wert steht für eine Nutzung auditorischer Informationen, ein positiver Wert für eine Nichtnutzung.

^b Differenz der Bedingungen auditorisch passend und unimodal für die ferne Schätzdistanz. Ein negativer Wert steht für eine Nutzung auditorischer Informationen, ein positiver Wert für eine Nichtnutzung.

^c Die Zuweisung erfolgte anhand der Nutzung auditorischer Informationen für beide Distanzen. Auditorische Informationen für beide Distanzen = Beidtonnutzer. Auditorische Informationen nur für die nahe Distanz = Nahtonnutzer. Auditorische Informationen nur für die ferne Distanz = Ferntonnutzer. Keine auditorischen Informationen = Nichtnutzer.

Tabelle A2

Auditorische Reaktionszeit-Differenzen für beide Schätzdistanzen in Experiment 2 auf Grundlage ungerader Blöcke

Proband	RT-Differenz auditorisch nah ^a	RT-Differenz auditorisch fern ^b	Nutzergruppe ^c
1	+0.0902	+0.0311	Nichtnutzer
2	+0.2834	-0.1064	Ferntonnutzer
3	-0.1027	-0.0079	Beidtonnutzer
4	-0.0819	+0.0976	Nahtonnutzer
5	+0.0708	-0.0643	Ferntonnutzer
6	+0.1507	-0.0353	Ferntonnutzer
7	+0.0365	+0.0018	Nichtnutzer
8	-0.0511	+0.1084	Nahtonnutzer
9	+0.1152	-0.0514	Ferntonnutzer
10	+0.0005	+0.0021	Nichtnutzer
11	-0.0500	-0.1150	Beidtonnutzer
12	-0.0592	+0.1494	Nahtonnutzer
13	-0.0756	+0.0875	Nahtonnutzer
14	-0.0653	+0.1351	Nahtonnutzer
15	+0.0816	+0.0767	Nichtnutzer
16	+0.1478	-0.0793	Ferntonnutzer
17	-0.0070	+0.2413	Nahtonnutzer
18	+0.0293	-0.0888	Ferntonnutzer
19	-0.0451	+0.1316	Nahtonnutzer
20	+0.1079	-0.1585	Ferntonnutzer
21	-0.1180	+0.1286	Nahtonnutzer
22	+0.0576	-0.0276	Ferntonnutzer
23	+0.1080	+0.1956	Nichtnutzer
24	+0.1375	-0.1158	Ferntonnutzer
25	-0.0151	+0.1944	Nahtonnutzer
26	-0.0449	-0.0946	Beidtonnutzer
27	+0.1223	+0.1566	Nichtnutzer
28	+0.0190	-0.1494	Ferntonnutzer
29	+0.0001	+0.0754	Nichtnutzer
30	+0.0842	-0.0596	Ferntonnutzer
31	+0.0262	+0.0984	Nichtnutzer
32	-0.0386	-0.0203	Beidtonnutzer

Anmerkungen. Reaktionszeit (RT) in Sekunden (Ausgangswert = Absolute Abweichung von der tatsächlichen Ankunftszeit). Berechnung und Kategorisierung auf Grundlage ungerader Blöcke, im Anschluss Analyse der RT-Abweichungen mit Daten der geraden Blöcke.

^a Differenz der Bedingungen auditorisch passend und unimodal für die nahe Schätzdistanz. Ein negativer Wert steht für eine Nutzung auditorischer Informationen, ein positiver Wert für eine Nichtnutzung.

^b Differenz der Bedingungen auditorisch passend und unimodal für die ferne Schätzdistanz. Ein negativer Wert steht für eine Nutzung auditorischer Informationen, ein positiver Wert für eine Nichtnutzung.

^c Die Zuweisung erfolgte anhand der Nutzung auditorischer Informationen für beide Distanzen. Auditorische Informationen für beide Distanzen = Beidtonnutzer. Auditorische Informationen nur für die nahe Distanz = Nahtonnutzer. Auditorische Informationen nur für die ferne Distanz = Ferntonnutzer. Keine auditorischen Informationen = Nichtnutzer.

Tabelle A3

Auditorische und visuelle Reaktionszeit-Differenzen in Experiment 3 auf Grundlage gerader Blöcke

Proband	RT-Differenz auditorisch ^a	RT-Differenz visuell ^b	Nutzergruppe ^c
1	+0.0313	+0.0018	Nichtnutzer
2	-0.0010	+0.0056	Tonnutzer
3	+0.0066	-0.0421	Bildnutzer
4	+0.0069	+0.0637	Nichtnutzer
5	+0.0004	+0.0520	Nichtnutzer
6	+0.0127	-0.0679	Bildnutzer
7	+0.0216	-0.0243	Bildnutzer
8	+0.0386	-0.0968	Bildnutzer
9	-0.0160	-0.0370	Beidnutzer
10	-0.0092	-0.0101	Beidnutzer
11	+0.0098	+0.0113	Nichtnutzer
12	+0.0440	-0.0199	Bildnutzer
13	+0.0127	-0.0235	Bildnutzer
14	+0.0106	-0.0233	Bildnutzer
15	+0.0113	-0.0203	Bildnutzer
16	+0.0244	-0.0742	Bildnutzer
17	-0.0580	+0.0047	Tonnutzer
18	+0.0100	+0.0027	Nichtnutzer
19	+0.0175	+0.0815	Nichtnutzer
20	+0.0110	-0.0768	Bildnutzer
21	+0.0020	+0.0793	Nichtnutzer
22	+0.0419	+0.0640	Nichtnutzer
23	-0.0055	-0.0842	Beidnutzer
24	+0.0196	-0.0577	Bildnutzer
25	+0.0096	+0.0015	Nichtnutzer
26	+0.0033	-0.0165	Bildnutzer
27	-0.0026	+0.0321	Tonnutzer
28	+0.0186	-0.0411	Bildnutzer
29	+0.0049	-0.0578	Bildnutzer
30	+0.0126	-0.0384	Bildnutzer
31	+0.0190	-0.0423	Bildnutzer
32	+0.0081	-0.0121	Bildnutzer
33	-0.0350	-0.0086	Beidnutzer
34	-0.0282	-0.0108	Beidnutzer
35	+0.0166	-0.0099	Bildnutzer
36	-0.0170	+0.0163	Tonnutzer
37	+0.0241	-0.0459	Bildnutzer
38	-0.0113	+0.0067	Tonnutzer
39	-0.0082	-0.0197	Beidnutzer
40	-0.0189	+0.0017	Tonnutzer

Anmerkungen. Reaktionszeit (RT) in Sekunden. Berechnung und Kategorisierung auf Grundlage gerader Blöcke, im Anschluss Analyse der RTs mit Daten der ungeraden Blöcke.

^a Differenz der Bedingungen visuell passend/auditorisch passend und visuell passend/unimodal. Ein negativer Wert steht für eine Nutzung auditorischer Informationen, ein positiver Wert für eine Nichtnutzung.

^b Differenz der Bedingungen visuell passend/unimodal und visuell nicht passend/unimodal. Ein negativer Wert steht für eine Nutzung visueller Informationen, ein positiver Wert für eine Nichtnutzung.

^c Die Zuweisung erfolgte anhand der Nutzung auditorischer und visueller Informationen. Beide Informationen = Beidnutzer. Nur visuelle Informationen = Bildnutzer. Nur auditorische Informationen = Tonnutzer. Keine Informationen = Nichtnutzer.

Tabelle A4

Auditorische und visuelle Reaktionszeit-Differenzen in Experiment 3 auf Grundlage ungerader Blöcke

Proband	RT-Differenz auditorisch ^a	RT-Differenz visuell ^b	Nutzergruppe ^c
1	+0.0166	+0.0045	Nichtnutzer
2	+0.0187	-0.0038	Bildnutzer
3	+0.0067	-0.0242	Bildnutzer
4	+0.0324	+0.0297	Nichtnutzer
5	-0.0135	+0.0575	Tonnutzer
6	+0.0023	-0.0641	Bildnutzer
7	+0.0250	-0.0264	Bildnutzer
8	+0.0005	-0.0665	Bildnutzer
9	-0.0120	-0.0381	Beidnutzer
10	-0.0111	-0.0063	Beidnutzer
11	+0.0186	-0.0123	Bildnutzer
12	+0.0011	-0.0063	Bildnutzer
13	-0.0146	+0.0046	Tonnutzer
14	+0.0016	-0.0168	Bildnutzer
15	-0.0358	-0.0027	Tonnutzer
16	-0.0234	-0.0536	Beidnutzer
17	-0.0261	-0.0629	Beidnutzer
18	-0.0263	+0.0153	Tonnutzer
19	-0.0603	+0.1047	Tonnutzer
20	-0.0023	-0.0224	Beidnutzer
21	+0.0412	+0.0227	Nichtnutzer
22	+0.0178	+0.0816	Nichtnutzer
23	-0.0125	-0.0919	Beidnutzer
24	-0.0191	-0.0462	Beidnutzer
25	+0.0094	+0.0060	Nichtnutzer
26	-0.0325	-0.0015	Beidnutzer
27	-0.0112	+0.0547	Tonnutzer
28	+0.0064	-0.0301	Bildnutzer
29	+0.0123	-0.0663	Bildnutzer
30	+0.0101	-0.0363	Bildnutzer
31	+0.0416	-0.0340	Bildnutzer
32	-0.0125	+0.0220	Tonnutzer
33	-0.0131	-0.0376	Beidnutzer
34	-0.0385	-0.0171	Beidnutzer
35	+0.0007	+0.0208	Nichtnutzer
36	-0.0088	+0.0224	Tonnutzer
37	-0.0129	-0.0377	Beidnutzer
38	+0.0160	-0.0179	Bildnutzer
39	-0.0130	-0.0206	Beidnutzer
40	+0.0120	-0.0192	Bildnutzer

Anmerkungen. Reaktionszeit (RT) in Sekunden. Berechnung und Kategorisierung auf Grundlage ungerader Blöcke, im Anschluss Analyse der RTs mit Daten der geraden Blöcke.

^a Differenz der Bedingungen visuell passend/auditorisch passend und visuell passend/unimodal. Ein negativer Wert steht für eine Nutzung auditorischer Informationen, ein positiver Wert für eine Nichtnutzung.

^b Differenz der Bedingungen visuell passend/unimodal und visuell nicht passend/unimodal. Ein negativer Wert steht für eine Nutzung visueller Informationen, ein positiver Wert für eine Nichtnutzung.

^c Die Zuweisung erfolgte anhand der Nutzung auditorischer und visueller Informationen. Beide Informationen = Beidnutzer. Nur visuelle Informationen = Bildnutzer. Nur auditorische Informationen = Tonnutzer. Keine Informationen = Nichtnutzer.

Tabelle A5

Auditorische und visuelle Reaktionszeit-Differenzen in Experiment 4 auf Grundlage gerader Blöcke

Proband	RT-Differenz auditorisch ^a	RT-Differenz visuell ^b	Nutzergruppe ^c
1	+0.0308	+0.0147	Nichtnutzer
2	+0.0110	-0.0537	Bildnutzer
3	+0.0079	-0.0342	Bildnutzer
4	-0.0083	+0.0446	Tonnutzer
5	-0.0088	-0.0596	Beidnutzer
6	-0.0231	-0.0002	Beidnutzer
7	+0.0176	-0.0248	Bildnutzer
8	-0.0194	-0.0264	Beidnutzer
9	-0.0236	-0.0123	Beidnutzer
10	+0.0131	-0.0510	Bildnutzer
11	-0.0359	-0.0227	Beidnutzer
12	-0.0117	-0.0118	Beidnutzer
13	+0.0242	-0.1747	Bildnutzer
14	+0.0223	+0.1597	Nichtnutzer
15	+0.1004	+0.1901	Nichtnutzer
16	-0.0223	-0.0400	Beidnutzer
17	-0.0166	-0.0111	Beidnutzer
18	-0.0094	-0.0116	Beidnutzer
19	+0.0012	-0.0358	Bildnutzer
20	+0.0063	-0.0414	Bildnutzer
21	-0.0056	-0.0130	Beidnutzer
22	+0.0294	-0.0143	Bildnutzer
23	+0.0030	-0.1053	Bildnutzer
24	-0.0273	+0.0505	Tonnutzer
25	+0.0253	+0.0036	Nichtnutzer
26	+0.0230	-0.0907	Bildnutzer
27	+0.0086	+0.0181	Nichtnutzer
28	+0.0045	+0.0378	Nichtnutzer
29	+0.0130	-0.0034	Bildnutzer
30	-0.0430	+0.0056	Tonnutzer
31	+0.0131	-0.0070	Bildnutzer
32	-0.0028	+0.0478	Tonnutzer
33	+0.0071	-0.0149	Bildnutzer
34	+0.0323	-0.0317	Bildnutzer
35	-0.0082	-0.0022	Beidnutzer
36	-0.0035	+0.0234	Tonnutzer
37	+0.0237	+0.0159	Nichtnutzer
38	-0.0074	-0.0226	Beidnutzer
39	+0.0430	+0.0640	Nichtnutzer
40	+0.0009	-0.0644	Bildnutzer

Anmerkungen. Reaktionszeit (RT) in Sekunden. Berechnung und Kategorisierung auf Grundlage gerader Blöcke, im Anschluss Analyse der RTs mit Daten der ungeraden Blöcke.

^a Differenz der Bedingungen visuell passend/auditorisch passend und visuell passend/unimodal. Ein negativer Wert steht für eine Nutzung auditorischer Informationen, ein positiver Wert für eine Nichtnutzung.

^b Differenz der Bedingungen visuell passend/unimodal und visuell nicht passend/unimodal. Ein negativer Wert steht für eine Nutzung visueller Informationen, ein positiver Wert für eine Nichtnutzung.

^c Die Zuweisung erfolgte anhand der Nutzung auditorischer und visueller Informationen. Beide Informationen = Beidnutzer. Nur visuelle Informationen = Bildnutzer. Nur auditorische Informationen = Tonnutzer. Keine Informationen = Nichtnutzer.

Tabelle A6

Auditorische und visuelle Reaktionszeit-Differenzen in Experiment 4 auf Grundlage ungerader Blöcke

Proband	RT-Differenz auditorisch ^a	RT-Differenz visuell ^b	Nutzergruppe ^c
1	+0.0239	+0.0148	Nichtnutzer
2	+0.0009	-0.0666	Bildnutzer
3	+0.0241	-0.0539	Bildnutzer
4	-0.0081	+0.0307	Tonnutzer
5	+0.0043	-0.0365	Bildnutzer
6	+0.0045	-0.0230	Bildnutzer
7	-0.0101	-0.0217	Beidnutzer
8	-0.0104	+0.0008	Tonnutzer
9	+0.0294	-0.0304	Bildnutzer
10	-0.0139	-0.0458	Beidnutzer
11	+0.0043	-0.0180	Bildnutzer
12	-0.0066	-0.0215	Beidnutzer
13	+0.0183	-0.1684	Bildnutzer
14	+0.0276	+0.1009	Nichtnutzer
15	+0.1662	+0.1593	Nichtnutzer
16	-0.0108	-0.0297	Beidnutzer
17	-0.0147	-0.0264	Beidnutzer
18	-0.0093	-0.0290	Beidnutzer
19	-0.0337	+0.0318	Tonnutzer
20	-0.0032	-0.0567	Beidnutzer
21	+0.0124	-0.0183	Bildnutzer
22	+0.0237	-0.0408	Bildnutzer
23	+0.0038	-0.0918	Bildnutzer
24	-0.0214	+0.0345	Tonnutzer
25	-0.0032	+0.0095	Tonnutzer
26	+0.0234	-0.0963	Bildnutzer
27	+0.0094	-0.0169	Bildnutzer
28	-0.0047	+0.0108	Tonnutzer
29	+0.0294	-0.0315	Bildnutzer
30	-0.0035	+0.0138	Tonnutzer
31	-0.0199	-0.0108	Beidnutzer
32	-0.0174	+0.0555	Tonnutzer
33	+0.0112	-0.0230	Bildnutzer
34	-0.0186	-0.0123	Beidnutzer
35	+0.0088	-0.0439	Bildnutzer
36	+0.0022	-0.0001	Bildnutzer
37	+0.0644	-0.0051	Bildnutzer
38	-0.0126	-0.0113	Beidnutzer
39	+0.0133	+0.0743	Nichtnutzer
40	-0.0029	-0.0487	Beidnutzer

Anmerkungen. Reaktionszeit (RT) in Sekunden. Berechnung und Kategorisierung auf Grundlage ungerader Blöcke, im Anschluss Analyse der RTs mit Daten der geraden Blöcke.

^a Differenz der Bedingungen visuell passend/auditorisch passend und visuell passend/unimodal. Ein negativer Wert steht für eine Nutzung auditorischer Informationen, ein positiver Wert für eine Nichtnutzung.

^b Differenz der Bedingungen visuell passend/unimodal und visuell nicht passend/unimodal. Ein negativer Wert steht für eine Nutzung visueller Informationen, ein positiver Wert für eine Nichtnutzung.

^c Die Zuweisung erfolgte anhand der Nutzung auditorischer und visueller Informationen. Beide Informationen = Beidnutzer. Nur visuelle Informationen = Bildnutzer. Nur auditorische Informationen = Tonnutzer. Keine Informationen = Nichtnutzer.

Tabelle A7

Auditorische und visuelle Reaktionszeit-Differenzen in Experiment 5 auf Grundlage gerader Blöcke

Proband	RT-Differenz auditorisch ^a	RT-Differenz visuell ^b	Nutzergruppe ^c
1	+0.0486	-0.0226	Bildnutzer
2	+0.0254	-0.0530	Bildnutzer
3	-0.0566	+0.0596	Tonnutzer
4	+0.0601	-0.1347	Bildnutzer
5	+0.0190	-0.0478	Bildnutzer
6	-0.0061	+0.0147	Tonnutzer
7	+0.0236	-0.0908	Bildnutzer
8	+0.0052	-0.0641	Bildnutzer
9	-0.0335	+0.0233	Tonnutzer
10	-0.0543	+0.0123	Tonnutzer
11	+0.0076	-0.0200	Bildnutzer
12	-0.0014	-0.0348	Beidnutzer
13	-0.0028	+0.0857	Tonnutzer
14	+0.0049	+0.0173	Nichtnutzer
15	+0.0506	+0.0259	Nichtnutzer
16	+0.0139	-0.0444	Bildnutzer
17	+0.0299	-0.0662	Bildnutzer
18	-0.0004	-0.0102	Beidnutzer
19	-0.0295	-0.0064	Beidnutzer
20	+0.0102	-0.0232	Bildnutzer
21	-0.0084	-0.0295	Beidnutzer
22	-0.0104	-0.0058	Beidnutzer
23	-0.0139	-0.0327	Beidnutzer
24	+0.0215	-0.0828	Bildnutzer
25	+0.0073	-0.0429	Bildnutzer
26	-0.0054	+0.0062	Tonnutzer
27	-0.0105	-0.0230	Beidnutzer
28	+0.0307	-0.0430	Bildnutzer
29	-0.0072	-0.0153	Beidnutzer
30	-0.0095	+0.1049	Tonnutzer
31	-0.0099	-0.0664	Beidnutzer
32	-0.0162	+0.0322	Tonnutzer
33	-0.0045	-0.0279	Beidnutzer
34	-0.0090	-0.0202	Beidnutzer
35	+0.0021	-0.0657	Bildnutzer
36	+0.0426	+0.0443	Nichtnutzer
37	+0.0100	-0.0209	Bildnutzer
38	+0.0271	+0.0178	Nichtnutzer
39	+0.0274	+0.0427	Nichtnutzer
40	+0.0097	-0.0308	Bildnutzer

Anmerkungen. Reaktionszeit (RT) in Sekunden. Berechnung und Kategorisierung auf Grundlage gerader Blöcke, im Anschluss Analyse der RTs mit Daten der ungeraden Blöcke.

^a Differenz der Bedingungen visuell passend/auditorisch passend und visuell passend/unimodal. Ein negativer Wert steht für eine Nutzung auditorischer Informationen, ein positiver Wert für eine Nichtnutzung.

^b Differenz der Bedingungen visuell passend/unimodal und visuell nicht passend/unimodal. Ein negativer Wert steht für eine Nutzung visueller Informationen, ein positiver Wert für eine Nichtnutzung.

^c Die Zuweisung erfolgte anhand der Nutzung auditorischer und visueller Informationen. Beide Informationen = Beidnutzer. Nur visuelle Informationen = Bildnutzer. Nur auditorische Informationen = Tonnutzer. Keine Informationen = Nichtnutzer.

Tabelle A8

Auditorische und visuelle Reaktionszeit-Differenzen in Experiment 5 auf Grundlage ungerader Blöcke

Proband	RT-Differenz auditorisch ^a	RT-Differenz visuell ^b	Nutzergruppe ^c
1	+0.0273	-0.0018	Bildnutzer
2	+0.0286	-0.0295	Bildnutzer
3	-0.0014	+0.0026	Tonnutzer
4	+0.0271	-0.0745	Bildnutzer
5	+0.0184	-0.0646	Bildnutzer
6	+0.0276	-0.0056	Bildnutzer
7	+0.0224	-0.0892	Bildnutzer
8	+0.0071	-0.0447	Bildnutzer
9	-0.0437	+0.0439	Tonnutzer
10	-0.0546	+0.0293	Tonnutzer
11	+0.0051	-0.0007	Bildnutzer
12	-0.0014	-0.0294	Beidnutzer
13	-0.0162	+0.0707	Tonnutzer
14	-0.0068	+0.0190	Tonnutzer
15	+0.0641	+0.0039	Nichtnutzer
16	+0.0021	-0.0384	Bildnutzer
17	+0.0209	-0.0357	Bildnutzer
18	-0.0104	-0.0346	Beidnutzer
19	-0.0057	-0.0210	Beidnutzer
20	-0.0056	-0.0080	Beidnutzer
21	-0.0102	-0.0252	Beidnutzer
22	-0.0070	-0.0021	Beidnutzer
23	-0.0041	-0.0422	Beidnutzer
24	+0.0155	-0.0862	Bildnutzer
25	+0.0066	-0.0471	Bildnutzer
26	+0.0076	+0.0039	Nichtnutzer
27	-0.1158	-0.0631	Beidnutzer
28	+0.0109	-0.0179	Bildnutzer
29	+0.0065	-0.0190	Bildnutzer
30	+0.0124	+0.0943	Nichtnutzer
31	-0.0024	-0.0754	Beidnutzer
32	-0.0118	+0.0957	Tonnutzer
33	+0.0074	-0.0315	Bildnutzer
34	+0.0035	-0.0640	Bildnutzer
35	-0.0174	-0.0495	Beidnutzer
36	-0.0124	+0.0400	Tonnutzer
37	+0.0165	-0.0401	Bildnutzer
38	-0.0358	+0.0551	Tonnutzer
39	+0.0144	+0.0308	Nichtnutzer
40	-0.0406	-0.0221	Beidnutzer

Anmerkungen. Reaktionszeit (RT) in Sekunden. Berechnung und Kategorisierung auf Grundlage ungerader Blöcke, im Anschluss Analyse der RTs mit Daten der geraden Blöcke.

^a Differenz der Bedingungen visuell passend/auditorisch passend und visuell passend/unimodal. Ein negativer Wert steht für eine Nutzung auditorischer Informationen, ein positiver Wert für eine Nichtnutzung.

^b Differenz der Bedingungen visuell passend/unimodal und visuell nicht passend/unimodal. Ein negativer Wert steht für eine Nutzung visueller Informationen, ein positiver Wert für eine Nichtnutzung.

^c Die Zuweisung erfolgte anhand der Nutzung auditorischer und visueller Informationen. Beide Informationen = Beidnutzer. Nur visuelle Informationen = Bildnutzer. Nur auditorische Informationen = Tonnutzer. Keine Informationen = Nichtnutzer.

Tabelle A9

Vergleich der Einteilungskriterien für die Nutzergruppen in Experiment 2 für die Kategorisierung auf Grundlage gerader vs. ungerader Blöcke

Pb	Diff GN	Diff UN		Diff GF	Diff UF		Gruppe ^a G	Gruppe ^a U	
1	+	+	✓	+	+	✓	Nicht	Nicht	✓
2	+	+	✓	-	-	✓	Fern	Fern	✓
3	-	-	✓	-	-	✓	Beid	Beid	✓
4	-	-	✓	+	+	✓	Nah	Nah	✓
5	+	+	✓	-	-	✓	Fern	Fern	✓
6	+	+	✓	-	-	✓	Fern	Fern	✓
7	+	+	✓	-	+	x	Fern	Nicht	x
8	-	-	✓	+	+	✓	Nah	Nah	✓
9	+	+	✓	+	-	x	Nicht	Fern	x
10	-	+	x	+	+	✓	Nah	Nicht	x
11	-	-	✓	-	-	✓	Beid	Beid	✓
12	-	-	✓	+	+	✓	Nah	Nah	✓
13	-	-	✓	+	+	✓	Nah	Nah	✓
14	-	-	✓	+	+	✓	Nah	Nah	✓
15	+	+	✓	+	+	✓	Nicht	Nicht	✓
16	+	+	✓	-	-	✓	Fern	Fern	✓
17	-	-	✓	+	+	✓	Nah	Nah	✓
18	+	+	✓	-	-	✓	Fern	Fern	✓
19	-	-	✓	+	+	✓	Nah	Nah	✓
20	+	+	✓	-	-	✓	Fern	Fern	✓
21	-	-	✓	+	+	✓	Nah	Nah	✓
22	-	+	x	-	-	✓	Beid	Fern	x
23	+	+	✓	+	+	✓	Nicht	Nicht	✓
24	+	+	✓	-	-	✓	Fern	Fern	✓
25	-	-	✓	+	+	✓	Nah	Nah	✓
26	-	-	✓	-	-	✓	Beid	Beid	✓
27	+	+	✓	+	+	✓	Nicht	Nicht	✓
28	+	+	✓	-	-	✓	Fern	Fern	✓
29	-	+	x	+	+	✓	Nah	Nicht	x
30	+	+	✓	-	-	✓	Fern	Fern	✓
31	+	+	✓	+	+	✓	Nicht	Nicht	✓
32	-	-	✓	-	-	✓	Beid	Beid	✓

Anmerkungen. Vergleich der Differenzwerte (Diff), die als Grundlage für die Einteilung in die Nutzergruppen dienten, für die Kategorisierung auf Basis gerader Blöcke (G) mit der Kategorisierung auf Basis ungerader Blöcke (U). Positive Differenzen sind mit einem +, negative mit einem - gekennzeichnet. Es wurden die Differenzen für die nahe Distanz (N, Spalten 2 & 3), die Differenzen für die ferne Distanz (F, Spalten 5 & 6) sowie die resultierende Nutzergruppe (Gruppe, Spalten 8 & 9) verglichen. Eine Übereinstimmung ist mit einem Haken, eine Nicht-Übereinstimmung mit einem x gekennzeichnet.

^a Beid = Beidtonnutzer, Nah = Nahtonnutzer, Fern = Ferntonnutzer, Nicht = Nichtnutzer. Details zu den Nutzergruppen siehe Text.

Tabelle A10

Vergleich der Einteilungskriterien für die Nutzergruppen in Experiment 3 für die Kategorisierung auf Grundlage gerader vs. ungerader Blöcke

Pb	Diff GA	Diff UA		Diff GV	Diff UV		Gruppe ^a G	Gruppe ^a U	
1	+	+	✓	+	+	✓	Nicht	Nicht	✓
2	-	+	x	+	-	x	Ton	Bild	x
3	+	+	✓	-	-	✓	Bild	Bild	✓
4	+	+	✓	+	+	✓	Nicht	Nicht	✓
5	+	-	x	+	+	✓	Nicht	Ton	x
6	+	+	✓	-	-	✓	Bild	Bild	✓
7	+	+	✓	-	-	✓	Bild	Bild	✓
8	+	+	✓	-	-	✓	Bild	Bild	✓
9	-	-	✓	-	-	✓	Beid	Beid	✓
10	-	-	✓	-	-	✓	Beid	Beid	✓
11	+	+	✓	+	-	x	Nicht	Bild	x
12	+	+	✓	-	-	✓	Bild	Bild	✓
13	+	-	x	-	+	x	Bild	Ton	x
14	+	+	✓	-	-	✓	Bild	Bild	✓
15	+	-	x	-	+	x	Bild	Ton	x
16	+	-	x	-	-	✓	Bild	Beid	x
17	-	-	✓	+	-	x	Ton	Beid	x
18	+	-	x	+	+	✓	Nicht	Ton	x
19	+	-	x	+	+	✓	Nicht	Ton	x
20	+	-	x	-	-	✓	Bild	Beid	x
21	+	+	✓	+	+	✓	Nicht	Nicht	✓
22	+	+	✓	+	+	✓	Nicht	Nicht	✓
23	-	-	✓	-	-	✓	Beid	Beid	✓
24	+	-	x	-	-	✓	Bild	Beid	x
25	+	+	✓	+	+	✓	Nicht	Nicht	✓
26	+	-	x	-	-	✓	Bild	Beid	x
27	-	-	✓	+	+	✓	Ton	Ton	✓
28	+	+	✓	-	-	✓	Bild	Bild	✓
29	+	+	✓	-	-	✓	Bild	Bild	✓
30	+	+	✓	-	-	✓	Bild	Bild	✓
31	+	+	✓	-	-	✓	Bild	Bild	✓
32	+	-	x	-	+	x	Bild	Ton	x
33	-	-	✓	-	-	✓	Beid	Beid	✓
34	-	-	✓	-	-	✓	Beid	Beid	✓
35	+	+	✓	-	+	x	Bild	Nicht	x
36	-	-	✓	+	+	✓	Ton	Ton	✓
37	+	-	x	-	-	✓	Bild	Beid	x
38	-	+	x	+	-	x	Ton	Bild	x
39	-	-	✓	-	-	✓	Beid	Beid	✓
40	-	+	x	+	-	x	Ton	Bild	x

Anmerkungen. Vergleich der Differenzwerte (Diff), die als Grundlage für die Einteilung in die Nutzergruppen dienen, für die Kategorisierung auf Basis gerader Blöcke (G) mit der Kategorisierung auf Basis ungerader Blöcke (U). Positive Differenzen sind mit einem +, negative mit einem - gekennzeichnet. Es wurden auditorische Differenzen (A, Spalten 2 & 3), visuelle Differenzen (V, Spalten 5 & 6) sowie die resultierende Nutzergruppe (Gruppe, Spalten 8 & 9) verglichen. Eine Übereinstimmung ist mit einem Haken, eine Nicht-Übereinstimmung mit einem x gekennzeichnet.

^a Beid = Beidnutzer, Bild = Bildnutzer, Ton = Tonnutzer, Nicht = Nichtnutzer. Details zu den Nutzergruppen siehe Text.

Tabelle A11

Vergleich der Einteilungskriterien für die Nutzergruppen in Experiment 4 für die Kategorisierung auf Grundlage gerader vs. ungerader Blöcke

Pb	Diff GA	Diff UA		Diff GV	Diff UV		Gruppe ^a G	Gruppe ^a U	
1	+	+	✓	+	+	✓	Nicht	Nicht	✓
2	+	+	✓	-	-	✓	Bild	Bild	✓
3	+	+	✓	-	-	✓	Bild	Bild	✓
4	-	-	✓	+	+	✓	Ton	Ton	✓
5	-	+	x	-	-	✓	Beid	Bild	x
6	-	+	x	-	-	✓	Beid	Bild	x
7	+	-	x	-	-	✓	Bild	Beid	x
8	-	-	✓	-	+	x	Beid	Ton	x
9	-	+	x	-	-	✓	Beid	Bild	x
10	+	-	x	-	-	✓	Bild	Beid	x
11	-	+	x	-	-	✓	Beid	Bild	x
12	-	-	✓	-	-	✓	Beid	Beid	✓
13	+	+	✓	-	-	✓	Bild	Bild	✓
14	+	+	✓	+	+	✓	Nicht	Nicht	✓
15	+	+	✓	+	+	✓	Nicht	Nicht	✓
16	-	-	✓	-	-	✓	Beid	Beid	✓
17	-	-	✓	-	-	✓	Beid	Beid	✓
18	-	-	✓	-	-	✓	Beid	Beid	✓
19	+	-	x	-	+	x	Bild	Ton	x
20	+	-	x	-	-	✓	Bild	Beid	x
21	-	+	x	-	-	✓	Beid	Bild	x
22	+	+	✓	-	-	✓	Bild	Bild	✓
23	+	+	✓	-	-	✓	Bild	Bild	✓
24	-	-	✓	+	+	✓	Ton	Ton	✓
25	+	-	x	+	+	✓	Nicht	Ton	x
26	+	+	✓	-	-	✓	Bild	Bild	✓
27	+	+	✓	+	-	x	Nicht	Bild	x
28	+	-	x	+	+	✓	Nicht	Ton	x
29	+	+	✓	-	-	✓	Bild	Bild	✓
30	-	-	✓	+	+	✓	Ton	Ton	✓
31	+	-	x	-	-	✓	Bild	Beid	x
32	-	-	✓	+	+	✓	Ton	Ton	✓
33	+	+	✓	-	-	✓	Bild	Bild	✓
34	+	-	x	-	-	✓	Bild	Beid	x
35	-	+	x	-	-	✓	Beid	Bild	x
36	-	+	x	+	-	x	Ton	Bild	x
37	+	+	✓	+	-	x	Nicht	Bild	x
38	-	-	✓	-	-	✓	Beid	Beid	✓
39	+	+	✓	+	+	✓	Nicht	Nicht	✓
40	+	-	x	-	-	✓	Bild	Beid	x

Anmerkungen. Vergleich der Differenzwerte (Diff), die als Grundlage für die Einteilung in die Nutzergruppen dienen, für die Kategorisierung auf Basis gerader Blöcke (G) mit der Kategorisierung auf Basis ungerader Blöcke (U). Positive Differenzen sind mit einem +, negative mit einem - gekennzeichnet. Es wurden auditorische Differenzen (A, Spalten 2 & 3), visuelle Differenzen (V, Spalten 5 & 6) sowie die resultierende Nutzergruppe (Gruppe, Spalten 8 & 9) verglichen. Eine Übereinstimmung ist mit einem Haken, eine Nicht-Übereinstimmung mit einem x gekennzeichnet.

^a Beid = Beidnutzer, Bild = Bildnutzer, Ton = Tonnutzer, Nicht = Nichtnutzer. Details zu den Nutzergruppen siehe Text.

Tabelle A12

Vergleich der Einteilungskriterien für die Nutzergruppen in Experiment 5 für die Kategorisierung auf Grundlage gerader vs. ungerader Blöcke

Pb	Diff GA	Diff UA	Diff GV	Diff UV	Gruppe ^a G	Gruppe ^a U			
1	+	+	✓	-	✓	Bild	Bild	✓	
2	+	+	✓	-	✓	Bild	Bild	✓	
3	-	-	✓	+	✓	Ton	Ton	✓	
4	+	+	✓	-	✓	Bild	Bild	✓	
5	+	+	✓	-	✓	Bild	Bild	✓	
6	-	+	x	+	-	x	Ton	Bild	x
7	+	+	✓	-	✓	Bild	Bild	✓	
8	+	+	✓	-	✓	Bild	Bild	✓	
9	-	-	✓	+	✓	Ton	Ton	✓	
10	-	-	✓	+	✓	Ton	Ton	✓	
11	+	+	✓	-	✓	Bild	Bild	✓	
12	-	-	✓	-	✓	Beid	Beid	✓	
13	-	-	✓	+	✓	Ton	Ton	✓	
14	+	-	x	+	+	✓	Nicht	Ton	x
15	+	+	✓	+	+	✓	Nicht	Nicht	✓
16	+	+	✓	-	-	✓	Bild	Bild	✓
17	+	+	✓	-	-	✓	Bild	Bild	✓
18	-	-	✓	-	-	✓	Beid	Beid	✓
19	-	-	✓	-	-	✓	Beid	Beid	✓
20	+	-	x	-	-	✓	Bild	Beid	x
21	-	-	✓	-	-	✓	Beid	Beid	✓
22	-	-	✓	-	-	✓	Beid	Beid	✓
23	-	-	✓	-	-	✓	Beid	Beid	✓
24	+	+	✓	-	-	✓	Bild	Bild	✓
25	+	+	✓	-	-	✓	Bild	Bild	✓
26	-	+	x	+	+	✓	Ton	Nicht	x
27	-	-	✓	-	-	✓	Beid	Beid	✓
28	+	+	✓	-	-	✓	Bild	Bild	✓
29	-	+	x	-	-	✓	Beid	Bild	x
30	-	+	x	+	+	✓	Ton	Nicht	x
31	-	-	✓	-	-	✓	Beid	Beid	✓
32	-	-	✓	+	+	✓	Ton	Ton	✓
33	-	+	x	-	-	✓	Beid	Bild	x
34	-	+	x	-	-	✓	Beid	Bild	x
35	+	-	x	-	-	✓	Bild	Beid	x
36	+	-	x	+	+	✓	Nicht	Ton	x
37	+	+	✓	-	-	✓	Bild	Bild	✓
38	+	-	x	+	+	✓	Nicht	Ton	x
39	+	+	✓	+	+	✓	Nicht	Nicht	✓
40	+	-	x	-	-	✓	Bild	Beid	x

Anmerkungen. Vergleich der Differenzwerte (Diff), die als Grundlage für die Einteilung in die Nutzergruppen dienen, für die Kategorisierung auf Basis gerader Blöcke (G) mit der Kategorisierung auf Basis ungerader Blöcke (U). Positive Differenzen sind mit einem +, negative mit einem - gekennzeichnet. Es wurden auditorische Differenzen (A, Spalten 2 & 3), visuelle Differenzen (V, Spalten 5 & 6) sowie die resultierende Nutzergruppe (Gruppe, Spalten 8 & 9) verglichen. Eine Übereinstimmung ist mit einem Haken, eine Nicht-Übereinstimmung mit einem x gekennzeichnet.

^a Beid = Beidnutzer, Bild = Bildnutzer, Ton = Tonnutzer, Nicht = Nichtnutzer. Details zu den Nutzergruppen siehe Text.

B Statistische Kennwerte für die Varianzanalysen der audiovisuellen Diskriminationsexperimente

Tabelle B1

Statistische Kennwerte der Varianzanalyse Dimension \times Visuell \times Auditorisch für die Beidnutzer in Experiment 3

Effekt	df	F	p ^a
Gerade = Kategorisierung, Ungerade = Analyse (n = 6)			
Dimension	1, 5	0.40	.556
Visuell	1, 5	28.74	.003
Auditorisch	2, 10	0.48	.635
Dimension \times Visuell	1, 5	0.49	.516
Dimension \times Auditorisch	2, 10	4.44	.042
Visuell \times Auditorisch	2, 10	3.21	.084
Dimension \times Visuell \times Auditorisch	2, 10	0.49	.626
Ungerade = Kategorisierung, Gerade = Analyse (n = 12)			
Dimension	1, 11	0.09	.772
Visuell	1, 11	26.24	< .001
Auditorisch	2, 22	0.38	.688
Dimension \times Visuell	1, 11	0.49	.498
Dimension \times Auditorisch	2, 22	1.05	.365
Visuell \times Auditorisch	2, 22	0.85	.441
Dimension \times Visuell \times Auditorisch	2, 22	0.04	.963

Anmerkungen. Varianzanalyse für die Reaktionszeiten mit den Innersubjektfaktoren Dimension (pseudo-3D/echt-3D), Visuell (passend/nicht passend), Auditorisch (ohne Ton/passend/nicht passend) für die Beidnutzer. Analysiert wurden ungerade Blöcke nach Kategorisierung auf Basis gerader Blöcke (oben) bzw. gerade Blöcke nach Kategorisierung auf Basis ungerader Blöcke (unten).

^a Signifikante *p*-Werte auf einem Niveau von <.05 sind fett hervorgehoben.

Tabelle B2

Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Bildnutzer in Experiment 3

Effekt	<i>df</i> ^a	<i>F</i>	<i>p</i> ^b
Gerade = Kategorisierung, Ungerade = Analyse (<i>n</i> = 19)			
Dimension	1, 18	1.71	.207
Visuell	1, 18	21.58	<.001
Auditorisch	2, 36	1.17	.323
Dimension × Visuell	1, 18	0.09	.771
Dimension × Auditorisch	1.45, 26.17	2.78	.095
Visuell × Auditorisch	2, 36	0.62	.544
Dimension × Visuell × Auditorisch	2, 36	2.01	.149
Ungerade = Kategorisierung, Gerade = Analyse (<i>n</i> = 14)			
Dimension	1, 13	0.06	.816
Visuell	1, 13	15.85	.002
Auditorisch	2, 26	5.23	.012
Dimension × Visuell	1, 13	2.59	.132
Dimension × Auditorisch	2, 26	1.82	.181
Visuell × Auditorisch	2, 26	0.65	.529
Dimension × Visuell × Auditorisch	2, 26	0.02	.983

Anmerkungen. Varianzanalyse für die Reaktionszeiten mit den Innersubjektfaktoren Dimension (pseudo-3D/echt-3D), Visuell (passend/nicht passend), Auditorisch (ohne Ton/passend/nicht passend) für die Bildnutzer. Analysiert wurden ungerade Blöcke nach Kategorisierung auf Basis gerader Blöcke (oben) bzw. gerade Blöcke nach Kategorisierung auf Basis ungerader Blöcke (unten).

^a Ggf. korrigiert nach Greenhouse-Geisser.

^b Signifikante *p*-Werte auf einem Niveau von <.05 sind fett hervorgehoben.

Tabelle B3

Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Beidnutzer in Experiment 4

Effekt	<i>df</i>	<i>F</i>	<i>p</i> ^a
Gerade = Kategorisierung, Ungerade = Analyse (<i>n</i> = 12)			
Dimension	1, 11	0.92	.358
Visuell	1, 11	39.61	<.001
Auditorisch	2, 22	0.01	.988
Dimension × Visuell	1, 11	0.75	.405
Dimension × Auditorisch	2, 22	0.33	.725
Visuell × Auditorisch	2, 22	0.02	.985
Dimension × Visuell × Auditorisch	2, 22	1.03	.374
Ungerade = Kategorisierung, Gerade = Analyse (<i>n</i> = 11)			
Dimension	1, 10	0.39	.549
Visuell	1, 10	18.31	.002
Auditorisch	2, 20	0.47	.632
Dimension × Visuell	1, 10	0.86	.376
Dimension × Auditorisch	2, 20	0.04	.959
Visuell × Auditorisch	2, 20	0.26	.772
Dimension × Visuell × Auditorisch	2, 20	2.75	.088

Anmerkungen. Varianzanalyse für die Reaktionszeiten mit den Innersubjektfaktoren Dimension (pseudo-3D/echt-3D), Visuell (passend/nicht passend), Auditorisch (ohne Ton/passend/nicht passend) für die Beidnutzer. Analysiert wurden ungerade Blöcke nach Kategorisierung auf Basis gerader Blöcke (oben) bzw. gerade Blöcke nach Kategorisierung auf Basis ungerader Blöcke (unten).

^a Signifikante *p*-Werte auf einem Niveau von <.05 sind fett hervorgehoben.

Tabelle B4

Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Bildnutzer in Experiment 4

Effekt	<i>df</i> ^a	<i>F</i>	<i>p</i> ^b
Gerade = Kategorisierung, Ungerade = Analyse (<i>n</i> = 15)			
Dimension	1, 14	1.21	.290
Visuell	1, 14	17.68	.001
Auditorisch	2, 28	0.69	.511
Dimension × Visuell	1, 14	0.83	.378
Dimension × Auditorisch	2, 28	0.84	.443
Visuell × Auditorisch	2, 28	0.38	.688
Dimension × Visuell × Auditorisch	2, 28	2.33	.116
Ungerade = Kategorisierung, Gerade = Analyse (<i>n</i> = 17)			
Dimension	1, 16	4.55	.049
Visuell	1, 16	10.51	.005
Auditorisch	1.24, 19.82	1.03	.341
Dimension × Visuell	1, 16	0.01	.942
Dimension × Auditorisch	2, 32	0.37	.691
Visuell × Auditorisch	2, 32	0.74	.485
Dimension × Visuell × Auditorisch	2, 32	2.14	.134

Anmerkungen. Varianzanalyse für die Reaktionszeiten mit den Innersubjektfaktoren Dimension (pseudo-3D/echt-3D), Visuell (passend/nicht passend), Auditorisch (ohne Ton/passend/nicht passend) für die Bildnutzer. Analysiert wurden ungerade Blöcke nach Kategorisierung auf Basis gerader Blöcke (oben) bzw. gerade Blöcke nach Kategorisierung auf Basis ungerader Blöcke (unten).

^a Ggf. korrigiert nach Greenhouse-Geisser.

^b Signifikante *p*-Werte auf einem Niveau von <.05 sind fett hervorgehoben.

Tabelle B5

Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Beidnutzer in Experiment 5

Effekt	<i>df</i> ^a	<i>F</i>	<i>p</i> ^b
Gerade = Kategorisierung, Ungerade = Analyse (<i>n</i> = 11)			
Dimension	1, 10	0.97	.348
Visuell	1, 10	12.07	.006
Auditorisch	1.18, 11.82	2.52	.136
Dimension × Visuell	1, 10	0.76	.404
Dimension × Auditorisch	2, 20	2.95	.076
Visuell × Auditorisch	1.30, 12.99	0.06	.874
Dimension × Visuell × Auditorisch	1.32, 13.22	0.32	.643
Ungerade = Kategorisierung, Gerade = Analyse (<i>n</i> = 11)			
Dimension	1, 10	0.34	.572
Visuell	1, 10	14.76	.003
Auditorisch	2, 20	3.04	.070
Dimension × Visuell	1, 10	0.81	.389
Dimension × Auditorisch	2, 20	1.05	.370
Visuell × Auditorisch	1.21, 12.07	0.02	.924
Dimension × Visuell × Auditorisch	1.33, 13.25	0.98	.367

Anmerkungen. Varianzanalyse für die Reaktionszeiten mit den Innersubjektfaktoren Dimension (einfach-2D/echt-3D), Visuell (passend/nicht passend), Auditorisch (ohne Ton/passend/nicht passend) für die Beidnutzer. Analysiert wurden ungerade Blöcke nach Kategorisierung auf Basis gerader Blöcke (oben) bzw. gerade Blöcke nach Kategorisierung auf Basis ungerader Blöcke (unten).

^a Ggf. korrigiert nach Greenhouse-Geisser.

^b Signifikante *p*-Werte auf einem Niveau von <.05 sind fett hervorgehoben.

Tabelle B6

Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Bildnutzer in Experiment 5

Effekt	<i>df</i> ^a	<i>F</i>	<i>p</i> ^b
Gerade = Kategorisierung, Ungerade = Analyse (<i>n</i> = 16)			
Dimension	1, 15	1.12	.307
Visuell	1, 15	28.69	<.001
Auditorisch	1.26, 18.86	3.17	.083
Dimension × Visuell	1, 15	0.24	.635
Dimension × Auditorisch	2, 30	1.19	.318
Visuell × Auditorisch	2, 30	0.18	.837
Dimension × Visuell × Auditorisch	2, 30	0.69	.509
Ungerade = Kategorisierung, Gerade = Analyse (<i>n</i> = 17)			
Dimension	1, 16	2.30	.149
Visuell	1, 16	28.21	<.001
Auditorisch	2, 32	4.10	.026
Dimension × Visuell	1, 16	1.66	.216
Dimension × Auditorisch	2, 32	2.89	.070
Visuell × Auditorisch	2, 32	4.95	.013
Dimension × Visuell × Auditorisch	1.46, 23.38	0.27	.696

Anmerkungen. Varianzanalyse für die Reaktionszeiten mit den Innersubjektfaktoren Dimension (einfach-2D/echt-3D), Visuell (passend/nicht passend), Auditorisch (ohne Ton/passend/nicht passend) für die Bildnutzer. Analysiert wurden ungerade Blöcke nach Kategorisierung auf Basis gerader Blöcke (oben) bzw. gerade Blöcke nach Kategorisierung auf Basis ungerader Blöcke (unten).

^a Ggf. korrigiert nach Greenhouse-Geisser.

^b Signifikante *p*-Werte auf einem Niveau von <.05 sind fett hervorgehoben.

Tabelle B7

Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch × Experiment für die drei audiovisuellen Diskriminationsexperimente

Effekt	<i>df</i> ^a	<i>F</i>	<i>p</i> ^b
Dimension	1, 117	0.23	.636
Dimension × Experiment	2, 117	0.24	.785
Visuell	1, 117	9.14	.003
Visuell × Experiment	2, 117	0.09	.918
Auditorisch	1.43, 167.35	3.48	.048
Auditorisch × Experiment	2.86, 167.35	1.79	.153
Dimension × Visuell	1, 117	1.55	.216
Dimension × Visuell × Experiment	2, 117	0.48	.620
Dimension × Auditorisch	2, 234	0.91	.403
Dimension × Auditorisch × Experiment	4, 234	0.84	.498
Visuell × Auditorisch	1.77, 207.14	0.19	.830
Visuell × Auditorisch × Experiment	3.54, 207.14	1.06	.374
Dimension × Visuell × Auditorisch	2, 234	1.41	.246
Dimension × Visuell × Auditorisch × Experiment	4, 234	1.65	.163
Experiment	2, 117	0.06	.945

Anmerkungen. Varianzanalyse für die Reaktionszeiten mit den Innersubjektfaktoren Dimension (pseudo-3D bzw. einfach-2D/echt-3D), Visuell (passend/nicht passend), Auditorisch (ohne Ton/passend/nicht passend) und dem Zwischensubjektfaktor Experiment (Experiment 3/4/5).

^a Ggf. korrigiert nach Greenhouse-Geisser.

^b Signifikante *p*-Werte auf einem Niveau von <.05 sind fett hervorgehoben.

C Analyse korrekter Antworten der audiovisuellen Diskriminationsexperimente

Für die Beid- und Bildnutzer wurden für die drei audiovisuellen Diskriminationsexperimente (Experiment 3, 4 und 5) zusätzlich die korrekten Antworten für beide Auswertreihenfolgen (Gerade-Ungerade vs. Ungerade-Gerade) analysiert. Die Ergebnisse in Hinblick auf Signifikanzen sind in Tabelle C1 zusammengefasst. Keiner der etwaigen Effekte wurde für beide Reihenfolgen signifikant und ist somit als stabil zu betrachten. Daher wurde auf eine ausführliche Post-hoc-Betrachtung verzichtet. Da die Werte der korrekten Antworten für alle Experimente, beide Nutzergruppen und Auswertreihenfolgen zudem in nur sehr geringen Umfang auf hohem Niveau für alle Bedingungen schwankten ($\geq 95.1\%$ in Experiment 3, $\geq 95.0\%$ in Experiment 4 und $\geq 96.1\%$ in Experiment 5), ist davon auszugehen, dass der Einfluss auf die korrekten Antworten zum einen als geringfügig anzusehen ist und zum anderen die Daten nicht auf einen Speed-Accuracy-Trade-Off hindeuten. Die statistischen Kennwerte aller durchgeführten Varianzanalysen für die Analyse der korrekten Antworten sind in den Tabellen C2 bis C7 aufgeführt.

Tabelle C1

Vergleich der Ergebnisse für die audiovisuellen Diskriminationsexperimente für die Analyse gerader und ungerader Blöcke (Auswertung korrekter Antworten)

Effekt	GU E3	UG E3	GU E4	UG E4	GU E5	UG E5
Beidnutzer						
Dimension	n. s.	n. s.	n. s.	n. s.	n. s.	n. s.
Visuell	n. s.	s.	n. s.	s.	s.	n. s.
Auditorisch	n. s.	n. s.	n. s.	s.	m. s.	n. s.
Dimension \times Visuell	n. s.	n. s.	n. s.	n. s.	n. s.	m. s.
Dimension \times Auditorisch	n. s.	n. s.	n. s.	n. s.	n. s.	n. s.
Visuell \times Auditorisch	m. s.	m. s.	n. s.	n. s.	n. s.	n. s.
Dimension \times Visuell \times Auditorisch	n. s.	n. s.	n. s.	n. s.	n. s.	n. s.
Bildnutzer						
Dimension	n. s.	n. s.	s.	n. s.	n. s.	n. s.
Visuell	m. s.	n. s.	n. s.	n. s.	n. s.	n. s.
Auditorisch	n. s.	n. s.	n. s.	n. s.	n. s.	n. s.
Dimension \times Visuell	n. s.	n. s.	n. s.	n. s.	n. s.	n. s.
Dimension \times Auditorisch	n. s.	m. s.	n. s.	n. s.	n. s.	n. s.
Visuell \times Auditorisch	n. s.	n. s.	n. s.	n. s.	n. s.	n. s.
Dimension \times Visuell \times Auditorisch	n. s.	n. s.	n. s.	n. s.	n. s.	n. s.

Anmerkungen. Effekte-Übersicht für Experiment 3 (E3), Experiment 4 (E4) und Experiment 5 (E5) für die Analyse ungerader Blöcke nach Kategorisierung auf Basis gerader Blöcke (GU) bzw. für die Analyse gerader Blöcke nach Kategorisierung auf Basis ungerader Blöcke (UG). Es wurde eine Varianzanalyse mit den Innersubjektfaktoren Dimension, Visuell und Auditorisch für die korrekten Antworten durchgeführt. Effekte, die für die GU- und UG-Auswertung in beiden Fällen signifikant waren, sind fett hervorgehoben (hier nicht vorhanden).

Tabelle C2

Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Beidnutzer in Experiment 3 (Korrekte Antworten)

Effekt	<i>df</i>	<i>F</i>	<i>p</i> ^a
Gerade = Kategorisierung, Ungerade = Analyse (<i>n</i> = 6)			
Dimension	1, 5	0.01	.939
Visuell	1, 5	0.63	.464
Auditorisch	2, 10	1.28	.319
Dimension × Visuell	1, 5	0.02	.901
Dimension × Auditorisch	2, 10	0.05	.952
Visuell × Auditorisch	2, 10	3.98	.054
Dimension × Visuell × Auditorisch	2, 10	0.41	.677
Ungerade = Kategorisierung, Gerade = Analyse (<i>n</i> = 12)			
Dimension	1, 11	0.06	.816
Visuell	1, 11	4.88	.049
Auditorisch	2, 22	0.95	.401
Dimension × Visuell	1, 11	0.59	.459
Dimension × Auditorisch	2, 22	0.01	.991
Visuell × Auditorisch	2, 22	3.26	.057
Dimension × Visuell × Auditorisch	2, 22	1.58	.229

Anmerkungen. Varianzanalyse für die korrekten Antworten mit den Innersubjektfaktoren Dimension (pseudo-3D/echt-3D), Visuell (passend/nicht passend), Auditorisch (ohne Ton/passend/nicht passend) für die Beidnutzer. Analysiert wurden ungerade Blöcke nach Kategorisierung auf Basis gerader Blöcke (oben) bzw. gerade Blöcke nach Kategorisierung auf Basis ungerader Blöcke (unten).

^a Signifikante *p*-Werte auf einem Niveau von < .05 sind fett hervorgehoben.

Tabelle C3

Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Bildnutzer in Experiment 3 (Korrekte Antworten)

Effekt	<i>df</i> ^a	<i>F</i>	<i>p</i>
Gerade = Kategorisierung, Ungerade = Analyse (<i>n</i> = 19)			
Dimension	1, 18	0.19	.670
Visuell	1, 18	3.20	.090
Auditorisch	2, 36	0.91	.410
Dimension × Visuell	1, 18	2.47	.134
Dimension × Auditorisch	2, 36	1.05	.362
Visuell × Auditorisch	1.48, 26.64	0.53	.543
Dimension × Visuell × Auditorisch	2, 36	0.08	.921
Ungerade = Kategorisierung, Gerade = Analyse (<i>n</i> = 14)			
Dimension	1, 13	0.11	.746
Visuell	1, 13	0.11	.750
Auditorisch	2, 26	1.67	.207
Dimension × Visuell	1, 13	0.37	.554
Dimension × Auditorisch	2, 26	3.36	.050
Visuell × Auditorisch	2, 26	0.54	.589
Dimension × Visuell × Auditorisch	2, 26	1.05	.365

Anmerkungen. Varianzanalyse für die korrekten Antworten mit den Innersubjektfaktoren Dimension (pseudo-3D/echt-3D), Visuell (passend/nicht passend), Auditorisch (ohne Ton/passend/nicht passend) für die Bildnutzer. Analysiert wurden ungerade Blöcke nach Kategorisierung auf Basis gerader Blöcke (oben) bzw. gerade Blöcke nach Kategorisierung auf Basis ungerader Blöcke (unten).

^a Ggf. korrigiert nach Greenhouse-Geisser.

Tabelle C4

Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Beidnutzer in Experiment 4 (Korrekte Antworten)

Effekt	<i>df</i> ^a	<i>F</i>	<i>p</i> ^b
Gerade = Kategorisierung, Ungerade = Analyse (<i>n</i> = 12)			
Dimension	1, 11	2.38	.151
Visuell	1, 11	1.49	.247
Auditorisch	2, 22	0.18	.837
Dimension × Visuell	1, 11	0.15	.710
Dimension × Auditorisch	2, 22	0.90	.420
Visuell × Auditorisch	1.31, 14.37	2.51	.129
Dimension × Visuell × Auditorisch	2, 22	2.02	.157
Ungerade = Kategorisierung, Gerade = Analyse (<i>n</i> = 11)			
Dimension	1, 10	0.04	.851
Visuell	1, 10	8.21	.017
Auditorisch	2, 20	3.55	.048
Dimension × Visuell	1, 10	0.33	.580
Dimension × Auditorisch	2, 20	0.04	.964
Visuell × Auditorisch	2, 20	0.71	.503
Dimension × Visuell × Auditorisch	1.32, 13.22	0.16	.853

Anmerkungen. Varianzanalyse für die korrekten Antworten mit den Innersubjektfaktoren Dimension (pseudo-3D/echt-3D), Visuell (passend/nicht passend), Auditorisch (ohne Ton/passend/nicht passend) für die Beidnutzer. Analysiert wurden ungerade Blöcke nach Kategorisierung auf Basis gerader Blöcke (oben) bzw. gerade Blöcke nach Kategorisierung auf Basis ungerader Blöcke (unten).

^a Ggf. korrigiert nach Greenhouse-Geisser.

^b Signifikante *p*-Werte auf einem Niveau von <.05 sind fett hervorgehoben.

Tabelle C5

Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Bildnutzer in Experiment 4 (Korrekte Antworten)

Effekt	<i>df</i> ^a	<i>F</i>	<i>p</i> ^b
Gerade = Kategorisierung, Ungerade = Analyse (<i>n</i> = 15)			
Dimension	1, 14	5.66	.032
Visuell	1, 14	0.56	.468
Auditorisch	2, 28	0.24	.786
Dimension × Visuell	1, 14	1.04	.326
Dimension × Auditorisch	2, 28	1.45	.251
Visuell × Auditorisch	2, 28	0.45	.645
Dimension × Visuell × Auditorisch	2, 28	2.27	.123
Ungerade = Kategorisierung, Gerade = Analyse (<i>n</i> = 17)			
Dimension	1, 16	0.07	.791
Visuell	1, 16	2.54	.131
Auditorisch	2, 32	0.39	.678
Dimension × Visuell	1, 16	1.13	.303
Dimension × Auditorisch	1.49, 23.89	0.05	.904
Visuell × Auditorisch	2, 32	0.64	.535
Dimension × Visuell × Auditorisch	2, 32	0.84	.441

Anmerkungen. Varianzanalyse für die korrekten Antworten mit den Innersubjektfaktoren Dimension (pseudo-3D/echt-3D), Visuell (passend/nicht passend), Auditorisch (ohne Ton/passend/nicht passend) für die Bildnutzer. Analysiert wurden ungerade Blöcke nach Kategorisierung auf Basis gerader Blöcke (oben) bzw. gerade Blöcke nach Kategorisierung auf Basis ungerader Blöcke (unten).

^a Ggf. korrigiert nach Greenhouse-Geisser.

^b Signifikante *p*-Werte auf einem Niveau von <.05 sind fett hervorgehoben.

Tabelle C6

Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Beidnutzer in Experiment 5 (Korrekte Antworten)

Effekt	<i>df</i>	<i>F</i>	<i>p</i> ^a
Gerade = Kategorisierung, Ungerade = Analyse (<i>n</i> = 11)			
Dimension	1, 10	1.01	.339
Visuell	1, 10	7.23	.023
Auditorisch	2, 20	3.43	.052
Dimension × Visuell	1, 10	0.27	.615
Dimension × Auditorisch	2, 20	0.25	.780
Visuell × Auditorisch	2, 20	0.98	.392
Dimension × Visuell × Auditorisch	2, 20	0.43	.657
Ungerade = Kategorisierung, Gerade = Analyse (<i>n</i> = 11)			
Dimension	1, 10	0.49	.500
Visuell	1, 10	2.12	.176
Auditorisch	2, 20	0.51	.611
Dimension × Visuell	1, 10	4.33	.064
Dimension × Auditorisch	2, 20	0.42	.664
Visuell × Auditorisch	2, 20	1.48	.251
Dimension × Visuell × Auditorisch	2, 20	0.33	.726

Anmerkungen. Varianzanalyse für die korrekten Antworten mit den Innersubjektfaktoren Dimension (einfach-2D/echt-3D), Visuell (passend/nicht passend), Auditorisch (ohne Ton/passend/nicht passend) für die Beidnutzer. Analysiert wurden ungerade Blöcke nach Kategorisierung auf Basis gerader Blöcke (oben) bzw. gerade Blöcke nach Kategorisierung auf Basis ungerader Blöcke (unten).

^a Signifikante *p*-Werte auf einem Niveau von <.05 sind fett hervorgehoben.

Tabelle C7

Statistische Kennwerte der Varianzanalyse Dimension × Visuell × Auditorisch für die Bildnutzer in Experiment 5 (Korrekte Antworten)

Effekt	<i>df</i>	<i>F</i>	<i>p</i>
Gerade = Kategorisierung, Ungerade = Analyse (<i>n</i> = 16)			
Dimension	1, 15	0.03	.874
Visuell	1, 15	0.03	.858
Auditorisch	2, 30	0.78	.469
Dimension × Visuell	1, 15	1.26	.279
Dimension × Auditorisch	2, 30	0.60	.555
Visuell × Auditorisch	2, 30	0.08	.924
Dimension × Visuell × Auditorisch	2, 30	1.38	.267
Ungerade = Kategorisierung, Gerade = Analyse (<i>n</i> = 17)			
Dimension	1, 16	0.02	.898
Visuell	1, 16	2.51	.132
Auditorisch	2, 32	2.25	.122
Dimension × Visuell	1, 16	2.40	.141
Dimension × Auditorisch	2, 32	0.08	.924
Visuell × Auditorisch	2, 32	0.47	.631
Dimension × Visuell × Auditorisch	2, 32	0.29	.747

Anmerkungen. Varianzanalyse für die korrekten Antworten mit den Innersubjektfaktoren Dimension (einfach-2D/echt-3D), Visuell (passend/nicht passend), Auditorisch (ohne Ton/passend/nicht passend) für die Bildnutzer. Analysiert wurden ungerade Blöcke nach Kategorisierung auf Basis gerader Blöcke (oben) bzw. gerade Blöcke nach Kategorisierung auf Basis ungerader Blöcke (unten).

D Blockweise Betrachtung der Reaktionszeit-Differenzen

Tabelle D1

Anzahl Blöcke mit positiver bzw. negativer auditorischer Reaktionszeit-Differenz für nahe und ferne Distanzen in Experiment 2

Proband	Differenz (-) nah	Differenz (+) nah	Differenz (-) fern	Differenz (+) fern
1	2	30	13	19
2	0	32	24	8
3	28	4	17	15
4	26	6	9	23
5	10	22	23	9
6	1	31	20	12
7	9	23	18	14
8	26	6	4	28
9	0	32	19	13
10	18	14	14	18
11	29	3	26	6
12	28	4	4	28
13	20	12	12	20
14	28	4	6	26
15	8	24	10	22
16	1	31	23	9
17	22	10	0	32
18	6	26	24	8
19	26	6	10	22
20	4	28	24	8
21	26	6	7	25
22	12	20	19	13
23	2	30	0	32
24	2	30	28	4
25	21	11	0	32
26	23	9	28	4
27	11	21	0	32
28	7	25	29	3
29	20	12	8	24
30	7	25	21	11
31	13	19	4	28
32	21	11	18	14

Anmerkungen. Anzahl der Blöcke je Proband, in denen Töne für die nahe Schätzdistanz genutzt wurden (negative Differenz nah, Spalte 2) bzw. nicht genutzt wurden (positive Differenz nah, Spalte 3). Spalte 4 und 5 enthalten analog die Angaben für die ferne Distanz. Insgesamt wurden pro Proband 32 Blöcke erhoben.

Tabelle D2

Anzahl Blöcke mit positiver bzw. negativer auditorischer und visueller Reaktionszeit-Differenz in Experiment 3

Proband	Differenz (-) A	Differenz (+) A	Differenz (-) V	Differenz (+) V
1	12	20	14	18
2	15	17	14	18
3	13	19	24	8
4	13	19	13	19
5	17	15	7	25
6	14	18	29	3
7	15	17	19	13
8	13	19	27	5
9	21	11	31	1
10	20	12	21	11
11	12	20	17	15
12	15	17	15	17
13	17	15	16	16
14	13	19	24	8
15	18	14	15	17
16	18	14	24	8
17	22	10	20	12
18	21	11	15	17
19	21	11	8	24
20	14	18	25	7
21	11	21	9	23
22	15	17	8	24
23	18	14	31	1
24	18	14	28	4
25	13	19	15	17
26	19	13	17	15
27	19	13	8	24
28	12	20	26	6
29	12	20	31	1
30	15	17	26	6
31	10	22	22	10
32	18	14	16	16
33	19	13	20	12
34	24	8	19	13
35	10	22	11	21
36	22	10	10	22
37	17	15	24	8
38	15	17	18	14
39	18	14	23	9
40	16	16	17	15

Anmerkungen. Anzahl der Blöcke je Proband, in denen auditorische Informationen (A) genutzt wurden (negative Differenz, Spalte 2) bzw. nicht genutzt wurden (positive Differenz, Spalte 3). Spalte 4 und 5 enthalten analog die Angaben für die Nutzung visueller Informationen (V). Insgesamt wurden pro Proband 32 Blöcke erhoben.

Tabelle D3

Anzahl Blöcke mit positiver bzw. negativer auditorischer und visueller Reaktionszeit-Differenz in Experiment 4

Proband	Differenz (-) A	Differenz (+) A	Differenz (-) V	Differenz (+) V
1	10	22	14	18
2	12	20	30	2
3	12	20	22	10
4	17	15	7	25
5	15	17	23	9
6	17	15	21	11
7	18	14	21	11
8	20	12	17	15
9	16	16	24	8
10	18	14	28	4
11	19	13	22	10
12	18	14	19	13
13	10	22	31	1
14	13	19	9	23
15	5	27	7	25
16	19	13	28	4
17	19	13	20	12
18	21	11	23	9
19	17	15	16	16
20	15	17	25	7
21	14	18	20	12
22	9	23	23	9
23	17	15	32	0
24	23	8	4	27
25	13	19	14	18
26	15	17	23	9
27	11	21	15	17
28	14	18	15	17
29	12	20	21	11
30	18	14	11	21
31	16	16	17	15
32	20	12	9	23
33	13	19	23	9
34	15	17	20	12
35	19	13	20	12
36	16	16	7	25
37	10	22	16	16
38	23	9	21	11
39	13	19	3	29
40	16	16	29	3

Anmerkungen. Anzahl der Blöcke je Proband, in denen auditorische Informationen (A) genutzt wurden (negative Differenz, Spalte 2) bzw. nicht genutzt wurden (positive Differenz, Spalte 3). Spalte 4 und 5 enthalten analog die Angaben für die Nutzung visueller Informationen (V). Insgesamt wurden pro Proband 32 Blöcke erhoben. Für Proband 24 blieben in Block 2 im echt-3D-Teil zu wenige Trials nach dem Datenausschluss übrig, sodass keine Differenzwerte berechnet werden konnten. Deswegen wurden für Proband 24 nur 31 Blöcke ausgewertet.

Tabelle D4

Anzahl Blöcke mit positiver bzw. negativer auditorischer und visueller Reaktionszeit-Differenz in Experiment 5

Proband	Differenz (-) A	Differenz (+) A	Differenz (-) V	Differenz (+) V
1	9	23	16	16
2	10	22	27	5
3	19	13	15	17
4	9	23	29	3
5	11	21	23	9
6	14	18	14	18
7	11	21	32	0
8	14	18	30	2
9	21	11	14	18
10	27	5	15	17
11	18	14	18	14
12	14	18	26	6
13	14	18	4	28
14	17	15	11	21
15	8	24	18	14
16	16	16	28	4
17	12	20	26	6
18	16	16	18	14
19	20	12	19	13
20	16	16	23	9
21	17	15	26	6
22	18	14	17	15
23	18	14	25	7
24	14	18	28	4
25	12	20	30	2
26	17	15	18	14
27	18	14	19	13
28	12	20	24	8
29	17	15	21	11
30	18	14	2	30
31	21	11	32	0
32	15	17	9	23
33	12	20	24	8
34	18	14	21	11
35	20	12	24	8
36	11	21	10	22
37	14	18	23	9
38	17	15	16	16
39	14	18	12	20
40	22	10	22	10

Anmerkungen. Anzahl der Blöcke je Proband, in denen auditorische Informationen (A) genutzt wurden (negative Differenz, Spalte 2) bzw. nicht genutzt wurden (positive Differenz, Spalte 3). Spalte 4 und 5 enthalten analog die Angaben für die Nutzung visueller Informationen (V). Insgesamt wurden pro Proband 32 Blöcke erhoben.

E Vergleich korrekter Antworten der Nutzergruppen der audiovisuellen Diskriminationsexperimente

Es wurde überprüft, ob für die Nutzergruppen in den drei audiovisuellen Diskriminationsexperimenten Unterschiede für die korrekten Antworten vorliegen. Dafür wurden die mittleren korrekten Antworten für jeden Probanden im jeweiligen Analysedatensatz berechnet und diese zwischen den Nutzergruppen, die gemäß Kategorisierungsdatensatz zugeteilt wurden, verglichen. Dieses Vorgehen wurde für jedes Experiment crossvalidiert, das heißt, eine Hälfte des Datensatzes diente als Kategorisierungs- und die andere Hälfte als Analysedatensatz. Im Anschluss wurde dieses Vorgehen mit getauschten Zuweisungen zur Art des Datensatzes (Analyse vs. Kategorisierung) wiederholt. Details zu diesem Vorgehen können Abschnitt 6.4 entnommen werden. Da die zu vergleichenden Nutzergruppen zum Teil sehr kleine Stichprobengrößen aufweisen (Ton- und Nichtnutzer immer $n < 10$, kleinste Stichprobe $n = 4$), erfolgte der Gruppenvergleich jeweils mithilfe des nichtparametrischen Kruskal-Wallis-Tests.

Die mittleren Antwortraten für die vier Nutzergruppen für alle drei Experimente für beide Auswertreihenfolgen sind in Tabelle E1 dargestellt. Die entsprechenden statistischen Kennwerte der Gruppenvergleiche sind in Tabelle E2 zu finden. Zwar wurde ein signifikanter Unterschied zwischen den Gruppen für Experiment 4 im geraden Datensatz nach ungerader Kategorisierung gefunden, jedoch bestätigte sich dieser Effekt nicht in der umgekehrten Analyse des ungeraden Datensatzes, womit kein stabiler Effekt vorliegt. Es kommt hinzu, dass die mittleren Antwortraten nur sehr wenig auf einem sehr hohen Niveau zwischen 96.9% und 99.5% schwanken. Außerdem sind die zu vergleichenden Stichproben zum Teil sehr klein und unterschiedlich groß. Daher wird nicht davon ausgegangen, dass in den vorliegenden drei audiovisuellen Diskriminationsexperimenten Unterschiede in den korrekten Antwortraten zwischen den Nutzergruppen vorliegen.

Dennoch ist es möglich, dass sich die Probanden darin unterscheiden, ob sie ihren Fokus in der Aufgabenbearbeitung auf die Schnelligkeit oder die Antwortgüte legen, auch weil in der Instruktion dazu aufgefordert wurde so schnell *und* korrekt wie möglich zu antworten. Ob insbesondere die Nichtnutzer dazu tendieren eher korrekt zu antworten, ließe sich in

Tabelle E1

Antwortraten für die vier Nutzergruppen in den drei audiovisuellen Diskriminationsexperimenten

Nutzergruppe	GU E3	UG E3	GU E4	UG E4	GU E5	UG E5
Beidnutzer	97.5	97.4	98.7	99.0	98.7	97.5
Bildnutzer	97.4	96.9	97.4	97.4	98.3	98.2
Tonnutzer	98.3	98.2	99.3	98.2	98.2	97.5
Nichtnutzer	98.7	99.4	99.0	99.5	98.9	99.0

Anmerkungen. Mittlere Antwortraten der vier Nutzergruppen für Experiment 3 (E3), Experiment 4 (E4) und Experiment 5 (E5) für die Analyse ungerader Blöcke nach Kategorisierung auf Basis gerader Blöcke (GU) bzw. für die Analyse gerader Blöcke nach Kategorisierung auf Basis ungerader Blöcke (UG). Beidnutzer: nutzen beide Arten von Informationen. Bild-/Tonnutzer: nutzen jeweils nur visuelle/auditorische Informationen. Nichtnutzer: nutzen keine der beiden Informationen. Alle Angaben in Prozent.

einem Folgeexperiment untersuchen. In diesem sollte jedoch die Schwierigkeit erhöht werden (z. B. durch eine geringere räumliche Abweichung zwischen passenden und nicht passenden Bahnverläufen), sodass etwaige Unterschiede in den korrekten Antworten zwischen den Gruppen genauer untersucht werden können. Da der Fokus der vorliegenden Experimente auf der Reaktionszeit-Analyse lag, wurde auf eine zu schwierige Diskriminationsaufgabe verzichtet.

Tabelle E2

Statistische Kennwerte des Kruskal-Wallis-Tests für den Vergleich der Nutzergruppen in den audiovisuellen Diskriminationsexperimenten

Experiment/Datensatz	<i>df</i>	χ^2	<i>p</i> ^a
Experiment 3 (GU)	3	1.27	.736
Experiment 3 (UG)	3	6.30	.098
Experiment 4 (GU)	3	4.29	.232
Experiment 4 (UG)	3	7.98	.047
Experiment 5 (GU)	3	0.78	.855
Experiment 5 (UG)	3	4.08	.253

Anmerkungen. Kruskal-Wallis-Test für den Vergleich auf Unterschiede zwischen den Nutzergruppen für die Experimente 3, 4 und 5. Analysiert wurden ungerade Blöcke nach Kategorisierung auf Basis gerader Blöcke (GU) bzw. gerade Blöcke nach Kategorisierung auf Basis ungerader Blöcke (UG).

^a Signifikante *p*-Werte auf einem Niveau von $<.05$ sind fett hervorgehoben.

Eigenständigkeitserklärung

Hiermit erkläre ich, Sandra Dittrich, dass ich die eingereichte Dissertation mit dem Titel *Audiovisuelle Bewegungsvorhersage im dreidimensionalen Raum* selbständig verfasst, nicht bereits als Dissertation verwendet sowie die benutzten Quellen und Hilfsmittel vollständig angegeben habe. Außerdem habe ich weder diese noch eine andere Arbeit zur Erlangung des akademischen Grades Dr. rer. nat. an anderen Einrichtungen eingereicht.

Magdeburg, den 14. 06. 2018

Sandra Dittrich