

Übersichtsartikel

Uta Noppeney*, Samuel A. Jones, Tim Rohe und Ambra Ferrari

Schauen Sie, was Sie hören – Wie das Gehirn mit allen seinen Sinnen die Welt wahrnimmt

<https://doi.org/10.1515/nf-2017-0066>

Zusammenfassung: Unsere Sinne werden fortwährend mit den unterschiedlichsten Signalen bombardiert. Um dieses Sinneschaos zu verstehen, muss das Gehirn Sinnesreize integrieren, wenn sie von einer Quelle kommen, aber separate verarbeiten, wenn sie von unterschiedlichen Quellen kommen. Somit beruht multisensorische Wahrnehmung entscheidend auf dem Erfassen der kausalen Struktur, die die Sinnesreize erzeugt hat. Verhaltensstudien legen nahe, dass das Gehirn zwischen Integration und Segregation wie von normativen Modellen der Bayesianischen kausalen Inferenz vorhergesagt abwägt. Neueste funktionelle Magnetresonanztomographie (fMRI) und Elektroenzephalographie (EEG) Studien zeigten, dass das Gehirn Bayesianische kausale Inferenz durchführt, indem es mehrere Wahrnehmungsschätzwerte dynamisch auf verschiedenen Ebenen der corticalen Hierarchie der Sinnesverarbeitung enkodiert. Erst an der Spitze der Hierarchie in anterioren parietalen Arealen formt das Gehirn Wahrnehmungsschätzwerte, die die Ungewissheit des Beobachters über die kausale Struktur der Umgebung berücksichtigt, wie von Modellen der Bayesianischen kausalen Inferenz vorhergesagt.

Schlüsselwörter: Wahrnehmung, multisensorische Wahrnehmung, multisensorische Integration, audiovisuell, Bayesianisch, Computermodell, Modellierung, perzeptuelle Inferenz, Bindungsproblem, Kausale Inferenz

*Corresponding authors **Uta Noppeney**, Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, UK, E-Mail: U.Noppeney@bham.ac.uk

Samuel A. Jones, Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, UK, E-Mail: saj409@student.bham.ac.uk

Tim Rohe, Department of Psychiatry and Psychotherapy, University of Tübingen, Germany, E-Mail: Tim.Rohe@med.uni-tuebingen.de

Ambra Ferrari, Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, UK, E-Mail: axf548@student.bham.ac.uk

Herausforderungen in der multi-sensorischen Wahrnehmung

Im alltäglichen Leben werden unsere Sinne fortwährend mit den unterschiedlichsten Signalen bombardiert: Das Motorengeräusch eines Lastwagens, ein funkelnendes Motorrad, das mit hoher Geschwindigkeit an uns vorbeibraust, der Geruch von Rauch und Abgasen und der Anblick von anderen Fußgängern. Wie überträgt das menschliche Gehirn diese Kakophonie der Sinne in eine wahrheitsgetreue Repräsentation der Welt? Den plötzlich auftauchenden Lastwagen als sprechend und glänzend und Ihren Begleiter als tosend und stinkend wahrzunehmen, könnte katastrophale Folgen haben! Dies verdeutlicht, dass multisensorische Integration und Segregation essenziell für unsere täglichen Interaktionen sind.

Die Integration von Informationen erhöht die Salienz sensorischer Signale und ermöglicht uns dadurch, auf wichtige Ereignisse schneller und gezielter zu reagieren. Beispielsweise können wir schneller auf den Bürgersteig zurückspringen, um nicht von einem herannahenden Lastwagen erfasst zu werden (Diederich und Colonius, 2004; Frassinetti et al., 2002; Gillmeister und Eimer, 2007; Noesselt et al., 2008). Ferner ermöglicht die Kombination von komplementärer (zum Beispiel die Erkennung eines Objektes, indem man es von vorne betrachtet und an der Hinterseite berührt) oder redundanter Information (zum Beispiel die Lokalisierung eines Objektes mittels Seh- und Hörvermögens) über die Sinne hinweg eine robustere und präzisere Wahrnehmung (Ernst und Bühlhoff, 2004).

Eine einflussreiche Theorie in den Neurowissenschaften besagt, dass das Gehirn ein probabilistisches generatives Modell der sensorischen Inputs bildet, welches in der Wahrnehmung als Inferenz invertiert (d. h. umgekehrt vollzogen) wird (Kersten et al., 2004; Kersten und Yuille, 2003; Knill und Pouget, 2004). Die Bayesianische Theorie formuliert, wie der Beobachter unsichere Informationen miteinander kombinieren muss, um eine möglichst präzise Repräsentation von der Welt zu erhalten.

In der multisensorischen Wahrnehmung muss das Gehirn zwei grundlegende Herausforderungen lösen. Zum

einen stellt sich das Problem der sogenannten 'kausalen Inferenz' oder das 'Bindungsproblem'. Das Gehirn muss entscheiden, ob Signale von einer gemeinsamen Ursache stammen und in der Wahrnehmung integriert werden sollen (Shams und Beierholm, 2010). Zum anderen, sollte das Gehirn Signale, die von einer Quelle stammen, entsprechend ihren Verlässlichkeiten oder Präzisionen kombinieren, um so die genaueste (d. h. die am geringsten variable) Repräsentation der Umgebung zu bilden (Alais und Burr, 2004; Ernst und Banks, 2002). Das Bayesianische kausale Inferenz-Modell löst diese beiden Herausforderungen, indem es explizit die potenziellen kausalen Strukturen (z. B. ob die Signale von einer oder von mehreren Quellen stammen) unserer Welt modelliert (Körding et al., 2007; siehe auch Deroy et al., 2016; Rohe und Noppeney, 2015a, 2015b; Shams und Beierholm, 2010; Wozny et al., 2010).

Nehmen wir ein einfaches Beispiel: Stellen Sie sich vor, Sie sind ein begeisterter Ornithologe, der in der Morgendämmerung im Wald umherstreift, um die besten Fotos und Gesangsaufzeichnungen von Vögeln zu machen. Plötzlich entdecken Sie auf einem Ast ein kleines Rotkehlchen und hören gleichzeitig im Gebüsch ein Rotkehlchen zwitschern. Wie richten Sie nun Ihre Kamera und Ihr Mikrofon aus? Sollen Sie die Information Ihres Seh- und Hörvermögens miteinander kombinieren, um eine genauere Abschätzung des Aufenthaltsortes des Vogels zu erhalten? Oder sollen Sie allein die Information Ihres Sehvermögens nutzen, um die Kamera zu platzieren und dementsprechend allein die Ihres Hörvermögens, um Ihr Mikrofon anzubringen? Die Antwort auf diese Frage ist abhängig von der zugrunde liegenden kausalen Struktur der Welt. Es existieren zwei Hypothesen oder potenzielle kausale Strukturen:

Im ersten Szenario handelt es sich bei dem Vogel, den Sie auf dem Ast sitzen sehen, um denselben Vogel, den Sie im Busch zwitschern hören. In diesem Falle einer 'gemeinsamen Signalquelle' sollten Sie die Signale des Seh- und Hörvermögens miteinander kombinieren. Dieses Szenario folgt demnach dem klassischen Modell der 'obligatorischen Integration', welches das multisensorische Forschungsfeld über die letzten zwei Jahrzehnte hinweg dominiert hat (Alais und Burr, 2004; Ernst und Banks, 2002; Hillis et al., 2004). Wie durch das Modell der *Maximum Likelihood Estimation* (MLE) beschrieben wird, erhält ein Beobachter in diesem Falle den präzisesten Schätzwert für den Ort des Rotkehlchens, indem er die Signale gewichtet nach ihrer relativen Verlässlichkeit (i. e. Inverse von Varianz des Signals) integriert. Beispielsweise würden Sie einem schwachen, unverlässlichen visuellen Signal in der Dämmerung eine geringere Bedeutung beimessen als

einem deutlichen visuellen Signal bei Tageslicht. Gemäß der Prinzipien der MLE-Integration sollte multisensorische Integration zu einer verlässlicheren oder präziseren Wahrnehmung führen. Diese multisensorische Varianzreduktion relativ zu der am geringsten variablen unisensorischen Wahrnehmung ist mit dem Faktor 2 am größten, wenn die Varianzen der beiden unisensorischen Signale gleich sind. Tatsächlich zeigten psychophysische Studien, dass der menschliche Beobachter Signale, welche von einer gemeinsamen Quelle ausgehen, ungefähr wie vom MLE Modell vorhergesagt integriert (Alais und Burr, 2004; Bresciani et al., 2006; Ernst und Banks, 2002; Hillis et al., 2004; Jacobs, 1999; Knill und Saunders, 2003). Dennoch ist die Evidenz nicht eindeutig. Zunehmend weisen Studien auch auf Situationen hin, in denen der menschliche Beobachter eine geringere multisensorische Varianzreduktion erreicht, als vom MLE-Modell vorhergesagt wird (Battaglia et al., 2011; Bentvelzen et al., 2009). Gleichfalls gewichten Beobachter eine Sinnesmodalität stärker als erwartet nach dem MLE-Modell (Battaglia et al., 2003; Burr et al., 2009; Butler et al., 2010; Rosas et al., 2005). Im letzteren Fall ist die Gewichtung der Sinnessignale nicht adäquat für spezielle experimentelle Situationen, sondern entspricht einer Gewichtung, die für die meisten Alltagssituationen dienlich ist (Battaglia et al., 2003).

Im zweiten Szenario gibt es zwei Vögel, einen, den Sie auf dem Ast sehen und einen, den Sie im Gebüsch zwitschern hören. In diesem Falle von 'mehreren unabhängigen Ursachen' oder 'vollständiger Segregation' wäre es nachteilig, die visuellen und auditorischen Signale zu integrieren. Stattdessen sollten Sie allein die auditorische Information nutzen, um Ihr Mikrofon auszurichten und die visuelle, um Ihre Kamera korrekt zu platzieren.

Das Gehirn muss nun aktiv von verschiedenen multisensorischen Korrespondenzen ableiten, wie wahrscheinlich die beiden Szenarien (d. h. eine oder zwei Signalquellen) als kausale Strukturen der Sinnesreize sind. Wichtige multisensorische Korrespondenzen sind beispielsweise die zeitliche Koinzidenz oder Korrelation der Sinnessignale (Lee & Noppeney, 2011a; Lewis und Noppeney, 2010; Magnotti et al., 2013; Maier et al., 2011; Munhall et al., 1996; Noesselt et al., 2007; Parise und Ernst, 2016; Parise et al., 2012; van Wassenhove et al., 2007), oder ihre räumliche Kollokation (Lewald und Guski, 2003; Slutsky und Recanzone, 2011; Spence, 2013). Aber auch andere höher geordnete statistische oder erlernte Korrespondenzen von Reizen wie ihre semantische (Adam und Noppeney, 2010; Bishop und Miller, 2011; Kanaya und Yokosawa, 2011; Lee und Noppeney, 2011b; Noppeney et al., 2010) oder metaphorische (Sadaghiani et al., 2009; Parise und Spence, 2009) Kongruenz sind hilfreich, um zu entscheiden, ob

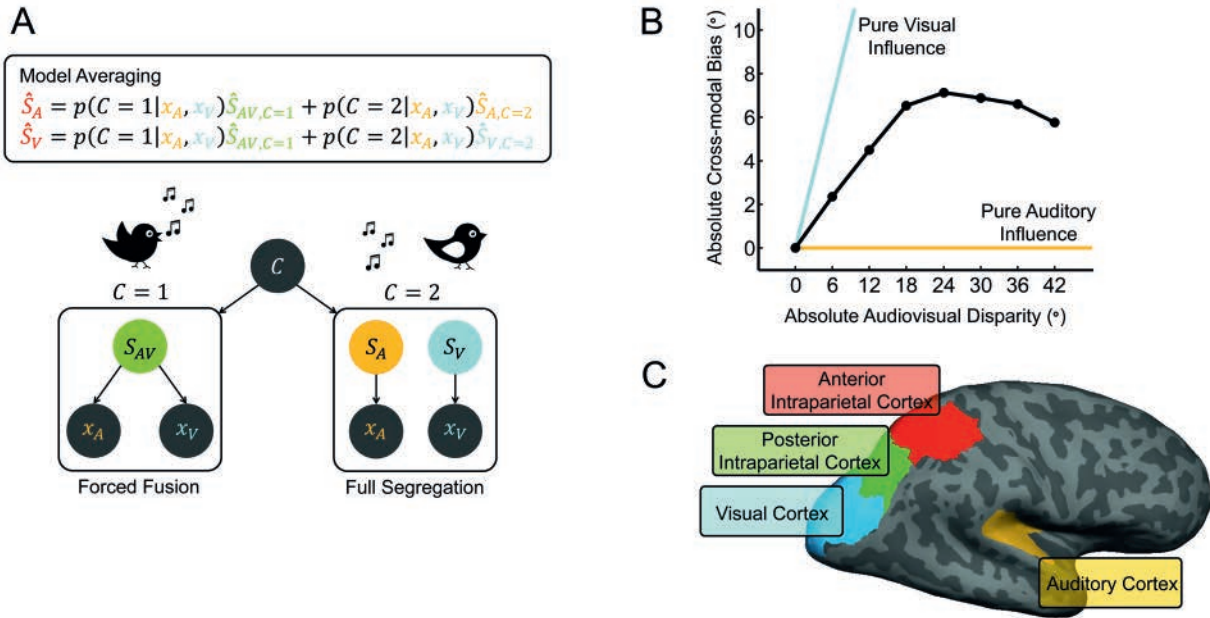


Abb. 1: A. Das Bayesianische kausale Inferenz Modell: Das generative Modell der Bayesianischen kausalen Inferenz modelliert explizit, ob die 'der gesehene Vogel' und das 'Zwitschern' auf einer gemeinsamen ($C=1$) oder zweier unabhängigen ($C=2$) Ursachen beruhen (Körding et al., 2007). Im Falle einer gemeinsamen Ursache wird ein einziger 'wahrer' audio-visueller Ort (S_{AV}) von einer 'prior' räumlichen Verteilung gewählt. Im Falle zweier unabhängiger Ursachen werden der 'wahre' auditorische (S_A) und der 'wahre' visuelle (S_V) Ort unabhängig voneinander von einer 'prior' räumlichen Verteilung gewählt. Schließlich werden unabhängige auditorische und visuelle Störsignale oder Rauschen eingeführt, um auditorische (x_A) und visuelle (x_V) Inputs zu generieren. **B.** Der visuelle 'bias' in der Wahrnehmung des Ortes eines Tones ist dargestellt als Funktion des räumlichen Abstands der auditorischen und visuellen Quellen voneinander. Wie von dem Bayesianischen kausalen Inferenz Modell vorhergesagt, integriert der Beobachter die geschätzten Orte der auditorischen und visuellen Quelle gewichtet proportional zu deren räumlichen Verlässlichkeiten. Dadurch, dass das visuelle Signal für räumliche Lokalisierung verlässlicher ist als das auditorische Signal, wird der Ort der auditorischen Quelle näher bei der visuellen Quelle wahrgenommen, i. e. eine sogenannter visuell induzierter Bias in der auditorischen Raumwahrnehmung. Dieser visuelle Bias nimmt zunächst zu, dann aber für größere räumliche Abstände zwischen der auditorischen und visuellen Quelle wieder ab. Mit anderen Worten audiovisuelle Interaktionen und Biases verringern sich, wenn es unwahrscheinlich ist, dass auditorische und visuelle Signale von einer einzigen Quelle stammen (Rohe und Noppeney, 2015b). **C.** Die Bayesianische kausale Inferenz innerhalb der corticalen Hierarchie. Primäre auditorische Areale encodieren den Ort des Tones, primär visuelle Areale den Ort des Blitzes. Entsprechend dem Prinzip der obligatorischen Integration integriert der posteriore intraparietale Cortex sensorische Signale proportional zu ihrer Verlässlichkeit. Der anteriore sulcus intraparietalis berechnet schließlich die auditorischen und visuellen Orte gemäß Bayesianischer kausaler Inferenz. Mit anderen Worten anteriore parietale Areale integrieren Signale, wenn sie wahrscheinlich von einer Quelle kommen, aber segregieren sie, wenn sie eher von zwei Quellen kommen (Rohe und Noppeney, 2015a). Nur an der Spitze der Verarbeitungshierarchie berücksichtigt das Gehirn somit seine Ungewissheit über die kausale Struktur, wie beispielsweise die Anzahl der Quellen, in seinen Berechnungen.

auditorische und visuelle Signale von derselben Quelle oder Ereignis stammen. Dennoch bleibt eine gewisse Ungewissheit über die kausale Struktur der Welt bestehen. Um dieser kausalen Ungewissheit gerecht zu werden, berechnet das Gehirn einen letztendlichen Schätzwert, indem es die Schätzwerte, die für die verschiedenen kausalen Strukturen (oder Szenarien) berechnet wurden, kombiniert (siehe im Detail: Wozny et al., 2010). Gemäß der Berechnungsstrategie namens 'Model Averaging' sollte das Gehirn beispielsweise die Ausrichtung des Mikrofons berechnen, indem es die räumlichen Schätzwerte, die für die Szenarien einer gemeinsamen (i. e. Modell der 'obligatorischen Integration') und zweier unabhängiger

Ursachen (i. e. Modell der 'vollständigen Segregation') ermittelt wurden, gewichtet nach ihren posterioren Wahrscheinlichkeiten kombiniert (Körding et al., 2007).

Zunehmend legen Studien nahe, dass der menschliche Beobachter zwischen sensorischer Integration und Segregation abwägt wie qualitativ von dem Bayesianischen Modell der kausalen Inferenz vorhergesagt (Beierholm et al., 2009; Bertelson und Radeau, 1981; Landy et al., 1995; Roach et al., 2006; Shams und Beierholm, 2010; Wallace et al., 2004). In Laborexperimenten wurde dies primär für räumliche Lokalisation (Körding et al., 2007; Rohe und Noppeney, 2015a, 2015b; Wozny et al., 2010) und Spracherkennungsaufgaben (Magnotti und Beauchamp, 2017;

Magnotti et al., 2013) gezeigt. Bei räumlichen Lokalisierungsexperimenten werden dem Beobachter auditorische (z. B. ein kurzer Ton) und visuelle Signale (z. B. ein Blitz) gleichzeitig entweder an demselben oder an unterschiedlichen Orten präsentiert. Bei jedem Durchlauf muss der Beobachter den Ort des Blitzes und/oder den des Tones bestimmen. Die Ergebnisse zeigen, dass der Beobachter einen Ton, der synchron aber an einem anderen Ort als der Blitz präsentiert wird, fehllokalisiert. In Abhängigkeit von den auditorischen und visuellen Verlässlichkeiten nimmt er den Ton näher beim Blitz und umgekehrt den Blitz näher beim Ton wahr. Wie vom Bayesianischen Modell der kausalen Inferenz vorhergesagt sind diese systematischen räumlichen Fehllokalisationen (sogenannte ‘crossmodale biases’) besonders ausgeprägt, wenn die auditorischen und visuellen Signale nahe beieinander sind. Sie vermindern sich mit zunehmendem audiovisuellen Abstand, bei dem es unwahrscheinlich ist, dass auditorische und visuelle Signale von einer gemeinsamen Quelle stammen (Körding et al., 2007; Rohe und Noppeney, 2015a, 2015b; Wozny et al., 2010).

Die Einflüsse eines visuellen Reizes auf die räumliche Wahrnehmung eines synchronen aber räumlich versetzten Geräusches sind auch die Grundlage für den Ventriquoist- oder auch Bauchredner-Effekt (Bertelson und Radeau, 1981; Bonath et al., 2007; Driver, 1996). Dieser beschreibt eine Wahrnehmungssillusion, die bereits in der Antike für religiöse Zwecke und später zu Unterhaltungszwecken auf Jahrmärkten genutzt wurde (Vox, 1981). Um die Illusion des Bauchredens zu erzeugen, spricht der Puppenspieler ohne seine Lippen zu bewegen. Gleichzeitig positioniert er die Puppe in der Nähe seines Gesichtes und bewegt die Lippen der Puppe synchron zu seinem eigenen Sprechen. Infolge der zeitlichen Korrelation zwischen dem auditorischen (i. e. Sprechsignal des Puppenspielers) und dem visuellen (i. e. Gesichtsbewegungen der Puppe) Signal schließt der Beobachter, dass die beiden Sinnessignale von einer gemeinsamen Ursache stammen, und integriert sie proportional zu ihren auditorischen und visuellen Verlässlichkeiten in eine kohärente Wahrnehmung. Da der Sehsinn präzisere räumliche Informationen liefert als der Hörsinn, fehllokalisiert der Betrachter das Sprechsignal des Puppenspielers und ordnet es der Puppe zu (Alais und Burr, 2004). Somit entsteht die Bauchredner-Illusion dadurch, dass das Gehirn Seh- und Hörsignale gemäß den Prinzipien der Bayesianischen kausalen Inferenz fälschlicherweise integriert, obwohl sie eigentlich von zwei verschiedenen Quellen kommen (siehe Exkurs 1 Box1 für Illusionen in der multisensorischen Wahrnehmung).

Multisensorische Interaktionen sind ubiquitär im Neocortex

Traditionell wurde angenommen, dass multisensorische Integration erst in höher geordneten Assoziationsarealen wie dem parietalen oder präfrontalen Cortex stattfindet (Avillac et al., 2007; Barraclough et al., 2005; Beauchamp et al., 2004; Calvert et al., 2000; Driver und Noesselt, 2008; Ghazanfar et al., 2008; Macaluso et al., 2003; Miller und D’Esposito, 2005; Sadaghiani et al., 2009; Schroeder und Foxe, 2002; Stevenson und James, 2009). Über die letzten beiden Jahrzehnte zeigte jedoch die Neuroimaging-Forschung am Menschen (Foxe et al., 2002; Lee und Noppeney, 2011a, 2014; Lehmann et al., 2006; Martuzzi et al., 2007; Molholm et al., 2002; Noesselt et al., 2007; Werner und Noppeney, 2010a), die Neurophysiologie an nicht-menschlichen Primaten oder Nagern (Atilgan et al., 2018; Bieler et al., 2017; Bizley et al., 2006; Bizley und King, 2009; Foxe und Schroeder, 2005; Ghazanfar et al., 2005; Ibrahim et al., 2016; Iurilli et al., 2012; Lakatos et al., 2007; Schroeder und Foxe, 2002; Kayser und Logothetis, 2007) und die Neuroanatomie (Falchier et al., 2002; Rockland und Ojima, 2003; Schroeder et al., 2003), dass multisensorische Integration bereits in frühen und sogar primär sensorischen Arealen stattfindet und dann progressiv entlang der corticalen Hierarchie zunimmt. Provokativ wurde sogar behauptet, dass ‘der gesamte Neocortex multisensorisch sei’ (Ghazanfar und Schroeder, 2006).

Zahlreiche fMRI – und EEG – Studien am Menschen zeigten, dass multisensorische Interaktionen in primär sensorischen Bereichen und in frühen Verarbeitungsstadien stattfinden und teilweise 100 ms nach dem Stimulus und sogar früher beobachtet werden können (Besle et al., 2008; Foxe et al., 2000; Molholm et al., 2002; Molholm et al., 2004). Gleichermaßen zeigten neurophysiologische Untersuchungen an nicht-humanen Primaten (Kayser et al., 2008, 2010; Lakatos et al., 2009; Schroeder und Foxe, 2005) oder Nagern (Atilgan et al., 2018; Bieler et al., 2017; Bizley und King, 2009; Bizley et al., 2007), dass die Antwort auf den präferierten (i. e. auditorischen) Stimulus im auditiven Cortex, durch einen synchronen Stimulus einer nicht-präferierten Sinnesmodalität (e. g. visuell) verstärkt oder auch unterdrückt werden kann. Diese frühen multisensorischen Interaktionen können auf Top-Down Einflüssen von höheren temporalen oder parietalen Cortices beruhen (Seltzer und Pandya, 1994) oder durch thalamo-corticale Mechanismen (zum Beispiel Pulvinar) oder direkte Verbindungen zwischen sensorischen Cortices (Mussacchia et al., 2014; Schroeder et al., 2003) vermittelt werden. In der Tat konnten neuroanatomische Tracer-Stu-

dien an nicht-menschlichen Primaten (Falchier et al., 2002; Rockland und Ojima, 2003) und Nagern (Bizley et al., 2007; Budinger et al., 2006; Campi et al., 2009; Ibrahim et al., 2016) direkte Verbindungen zwischen frühen oder sogar primären auditorischen und visuellen Arealen zeigen.

Die Allgegenwärtigkeit des multisensorischen Zusammenspiels auf allen Ebenen der corticalen Verarbeitung stellt das traditionelle, hierarchische Modell der späten Integration in Frage. Die vielseitigen multisensorischen Interaktionen in thalamo-corticalen, cortico-corticalen und höhergeordneten Assoziationscortices stattfindet (Musacchia und Schroeder, 2009; Schroeder et al., 2003) erfordern nun die genauere Charakterisierung ihrer funktionellen Eigenschaften und Verhaltensrelevanz.

Bisherige Studien haben in frühen sensorischen Cortices multisensorische Einflüsse beschrieben, in denen ein nicht-präferierter Reiz nicht nur die Antwort auf einen präferierten Reiz modulieren, sondern auch selbst eine Antwort auslösen kann (Atilgan et al., 2018; Bieler et al., 2017; Bizley und King, 2009; Bizley et al., 2007; Kayser et al., 2008; Lakatos et al., 2009; Meijer et al., 2017; Meredith und Allman, 2015). Sowohl Neuroimaging-Studien am Menschen (Leitão et al., 2012; Werner und Noppeney, 2011) als auch elektrophysiologische Untersuchungen an Nagern (Ibrahim et al., 2016; Iurilli et al., 2012) zeigten, dass nicht-präferierte Reize eine Deaktivierung oder synaptische Inhibition in sensorischen Cortices auslösen. Beispielsweise riefen visuelle Stimuli eine negative BOLD-Antwort in auditorischen Cortices hervor, während auditorische Stimuli synaptische Inhibition und fMRI-Deaktivierungen in visuellen Cortices erzeugten (Ibrahim et al., 2016; Iurilli et al., 2012; Leitão et al., 2012). Jedoch ruft ein Stimulus einer nicht-präferierten sensorischen Modalität nicht unbedingt eine verlässliche Antwort alleine hervor. In vielen Fällen moduliert er die Antwort auf einen Stimulus einer präferierten sensorischen Modalität. So moduliert ein synchroner visueller Input zum Beispiel die Antwort auf einen auditorischen Stimulus im primären und sekundären auditorischen Cortex (Kayser et al., 2010). Lakatos und Kollegen postulierten, dass diese modulatorischen Interaktionen auf Mechanismen des ‘Phase Resetting’ von Theta-Oszillationen beruhen (Lakatos et al., 2009; für relevante Forschung in Nagern siehe auch Sieben et al., 2012).

Da in unserer natürlichen Umgebung das visuelle Signal dem auditorischen oft vorangeht (z. B. artikulatorische Gesichtsbewegungen), kann dieses die auditorisch-induzierte Aktivität beeinflussen, indem es die Phase der neuronalen Oszillationen im auditorischen Cortex zurücksetzt (Lakatos et al., 2009; Schroeder et al., 2008). Dies könnte einen bedeutsamen Mechanismus darstellen,

durch welchen die Salienz von multisensorischen Ereignissen erhöht und deren Wahrnehmung erleichtert wird. Für zeitsensitive Mechanismen sprechen jüngste neurophysiologische (Kayser et al., 2010) und fMRI-Studien (Lewis und Noppeney, 2010; Werner und Noppeney, 2011), in denen audio-visuelle Interaktionen im Heschl’s Gyrus und dem Planum temporale von der temporalen Koinzidenz oder Korrelationen der auditorischen and visuellen Signale beeinflusst werden. Diese multisensorisch verstärkte Antwort in primär sensorischen Cortices wurde dann entsprechend ihrer Aufgabenrelevanz an höher geordnete Assoziationscortices weitergeleitet, beispielsweise an den ventralen temporalen Cortex zur Objekterkennung oder an hMT/V5 zur Bewegungswahrnehmung. Dies weist darauf hin, dass Integrationseffekte in frühen sensorischen Arealen höher geordnete Verarbeitungsprozesse und dadurch letztendlich das Verhalten beeinflussen (Lewis und Noppeney, 2010).

Höher geordnete Assoziationsareale wie der sulcus temporalis superior oder der parietale Cortex zeichnen sich durch größere zeitliche Integrationsfenster aus, so dass multisensorische Interaktionen geringer von der Synchronizität der sensorischen Inputs abhängen (Werner und Noppeney, 2011). Stattdessen integrieren diese Areale sensorische Signale in Repräsentationen (zum Beispiel Raum, Objekt, Sprache etc.), die wichtig für die jeweilige Aufgabe des Beobachters sind. In diesem Sinne präzisieren die Aktivierungsprofile in STS und IPS, ob der menschliche Beobachter durch multisensorische Integration seine Leistung verbessern kann: Je superadditiver die multisensorische Integration im sulcus temporalis superior, desto ausgeprägter verbesserte sich die Objektkategorisierungsleistung des Beobachters durch audiovisuelle Integration (Werner und Noppeney, 2010a, 2010b).

Selbst wenn sensorische Signale aufgrund ihrer Inkongruenz nicht in eine einheitliche Wahrnehmung integriert werden, so können sie doch auf der Ebene der Entscheidungsfindung interagieren. Zahlreiche Studien haben diese multisensorischen Interferenz-Effekte in Paradigmen der selektiven intersensorischen Aufmerksamkeit untersucht, in denen Versuchspersonen zwischen visuellen Reizen – beispielsweise den geschriebenen Wörtern ‘links’ oder ‘rechts’ unterscheiden müssen, während sie einen synchron präsentierten auditorischen Reiz (z. B. das gesprochene Wort ‘links’ oder ‘rechts’) ignorieren sollen. Die Versuchspersonen sind langsamer and machen häufiger Fehler, wenn der auditorische Reiz inkongruent zum visuellen Reiz ist (Noppeney et al., 2008) und vice versa (Krugliak und Noppeney, 2015; Marks, 1987). Wir konnten vor einigen Jahren in einer fMRI – Studie zeigen, dass der präfrontale Cortex – wie von einem Bayesianischen

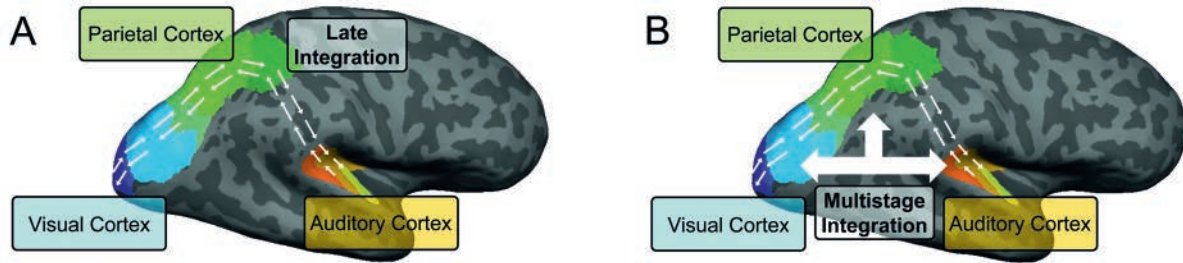


Abb. 2: Späte und mehrstufige Integrationsmodelle. Klassische Modelle nehmen an, dass multisensorische Integration in späten Phasen der Informationsverarbeitung in Assoziationsarealen stattfindet. Wir schlagen ein neues Modell vor, in dem unterschiedliche Arten von Informationen, die von den sensorischen Inputs abstrahiert wurden, auf verschiedenen Ebenen der corticalen Verarbeitungshierarchie integriert werden.

Kompatibilitätsmodell vorhergesagt – sensorische Informationen über die auditorischen und visuellen Stimuli akkumuliert bis ein Schwellenwert für eine Entscheidung erreicht und eine Antwort hervorgerufen wird (Noppeney et al., 2010). Ferner reagierte der präfrontale Cortex auf kongruente audiovisuelle Signale mit einer schwächeren Antwort als auf unisensorische Signale, was durch eine erleichterte und raschere Entscheidungsfindung erklärt werden kann (Sugihara et al., 2006; Werner und Noppeney, 2010a). Interessanterweise wurden die suppressiven Interaktionen vermehrt im Grenzbereich zwischen auditorisch und visuell dominanten Bereichen gefunden (Wallace et al., 2004). Dies konvergiert mit Studien an Nagern, die multisensorische Interaktionen überwiegend in den Transitionszonen zwischen sensorisch-dominanten Cortices zeigten.

Zusammengefasst zeigen diese Forschungsergebnisse, dass multisensorische Integration ein vielschichtiger Prozess ist, der sich auf mehreren Ebenen entlang der corticalen Hierarchie vollzieht. Während einige sensorische Interaktionen bereits früh und sogar in primären sensorischen Arealen beginnen, werden andere Informationen erst auf höheren corticalen Ebenen integriert. Möglicherweise dienen multisensorische Interaktionen in frühen sensorischen Arealen dazu, die Salienz von multisensorischen Ereignissen zu verstärken. Dies beeinflusst die repräsentationellen Integrationsprozesse und Prozesse der Entscheidungsfindung auf höheren Ebenen der corticalen Hierarchie (Werner und Noppeney, 2010a).

Wie das Gehirn kausale Inferenz und Verlässlichkeits-abhängige Integration vollzieht

Zu Beginn unseres Artikels erörterten wir, dass das menschliche Gehirn zwei entscheidende Herausforderungen in der multisensorischen Welt lösen muss. Zum einen muss es Signale, die von einer gemeinsamen Ursache ausgehen, identifizieren und integrieren, während es Signale von unterschiedlichen Ereignissen segregieren muss. Zum anderen müssen Signale, die von einer gemeinsamen Ursache ausgehen, proportional zu ihren Verlässlichkeiten oder Präzisionen integriert werden. Während der erste Teil des Artikels Evidenzen aus dem Bereich der Verhaltensforschung vorstellte, werden wir im Folgenden neueste Ergebnisse aus den Bereichen der Neurophysiologie und des Neuroimaging diskutieren, die wertvolle Einblicke in die zugrunde liegenden neuronalen Mechanismen geben.

Seit den bahnbrechenden Arbeiten von Stein und Kollegen, die multisensorische Integration in den Colliculi superiores untersuchten (Meredith und Stein, 1983; Wallace et al., 1996; Stein und Meredith, 1993), konnten viele neurophysiologische und Neuroimaging-Studien zeigen, dass multisensorische Interaktionen und kausale Inferenz von Faktoren wie räumlicher Kolokalisation, zeitlicher Synchronizität und Korrelationen abhängen (Stein und Stanford, 2008). Im Detail zeigten Stein und andere Gruppen, dass audio-visuelle Interaktionen superadditiv für räumlich kolokalisierte audio-visuelle Signale sind (d. h. die neuronale Antwort auf den audio-visuellen Stimulus ist größer als die Summe der unisensorischen Antworten), sich jedoch in additive oder sogar suppressive Interaktionen verwandeln, wenn die auditorischen und

visuellen Signale an unterschiedlichen Orten präsentiert werden und ein Signal außerhalb des rezeptiven Feldes eines anderen Reizes fällt (Stanford, 2005; Stanford und Stein, 2007; Wallace et al., 1996). Die neuronale Organisation in rezeptive Felder ermöglicht somit kausale Inferenz durch räumliche Integration and Segregation von Sinnesignalen.

Gleichfalls müssen Signale in einem zeitlichen Fenster von ungefähr 500 ms eintreffen, um in den colliculi superiores integriert zu werden. Nach jüngsten Modellansätzen hängt das zeitliche Bindungsfenster für kontinuierliche Signale wie Sprache von den audiovisuellen Signalkorrelationen ab, die das Gehirn möglicherweise mittels eines Hassenstein-Reichardt-Detektors, einer relativ einfachen aber dennoch physiologisch plausiblen Modellkomponente, bestimmt (Parise und Ernst, 2016). Vor Kurzem wurde untersucht, ob dieses zeitliche Bindungsfenster durch die internen Rhythmen des Gehirns, d. h. die neuronalen Oszillationen, beeinflusst wird. Die Studien zeigten, dass ein Reiz in einer Sinnesmodalität die Verarbeitung eines Reizes in einer anderen Sinnesmodalität durch ‘Phase resetting’ von Oszillationen moduliert. Diese Modulation hängt von der oszillatorischen Zykluslänge und der audio-visuellen Asynchronizität ab (Lakatos et al., 2009). Darüberhinaus ließen jüngere Studien sogar vermuten, dass Änderungen in der Zykluslänge von Alpha-Oszillationen über die Zeit das sensorische Bindungsfenster dynamisch in einem Versuchsprobanden modulieren kann (Samaha und Postle, 2015). Gleichzeitig erklärte auch die Variabilität in der Alpha-Zykluslänge zwischen den Versuchsprobanden Unterschiede in der Weite ihres sensorischen Bindungsfenster. Schnellere Oszillationen waren mit kleineren zeitlichen Bindungsfenstern assoziiert (Cecere et al., 2015; Samaha und Postle, 2015). Die Vorstellung, dass Oszillationszyklen eine ähnliche Funktion für zeitliche Bindungsfenster erfüllen wie rezeptive Felder für räumliche Bindungsfenster, ist zwar faszinierend, jedoch sind weitere Studien und spezifischere Modelle nötig, um diese Hypothese zu festigen.

Jüngere neurophysiologische Studien an nicht-humanen Primaten untersuchten insbesondere, wie einzelne Neurone oder Neuronengruppen Signale proportional zu ihrer Verlässlichkeit oder Präzision integrieren. Fetsch und Kollegen zeigten, dass Makaken ähnlich wie menschliche Beobachter eine nahezu optimale multisensorische Varianzreduktion und Leistungsverbesserung zeigten, wenn sie die Richtung ihrer eigenen Bewegung auf Grund von visuellen und vestibulären Signalen beurteilten (Fetsch et al., 2012). Gleichzeitige neuronale Ableitungen im dorsalen Bewegungsareal MSTd konnten zeigen, dass kongruente Neurone visuelle und vestibuläre Inputs subadditiv

(Gu et al., 2008) und proportional zu ihrer jeweiligen Verlässlichkeit (Fetsch et al., 2012) kombinierten, wobei sie verlässlichere sensorische Signale stärker wichteten (Fetsch et al., 2002). Wie durch das Maximum Likelihood Estimation – Modell vorhergesagt reagierten Neurone unter visuo-vestibulärer Stimulierung präziser auf Eigenbewegung als unter unisensorischer Stimulierung (Gu et al., 2008), genauso wie die Makaken auf der Verhaltensebene besser in der Beurteilung ihrer Bewegungsrichtung waren (siehe auch Nikbakht et al., 2018 für analoge Ergebnisse in Nagern). Gleichfalls zeigten die Dekodierungsanalysen der neuronalen Gruppen in MSTd neuronalen Gruppen, dass die sensorischen Gewichtungen auf der neuronalen und Verhaltensebene sich entsprachen (Fetsch et al., 2012). Zusätzliche elektrische Mikrostimulierung und chemische Inaktivierung des Areals MSTd unterstützten ferner einen kausalen Zusammenhang zwischen der neuronalen Verarbeitung in MSTd und dem Verhalten der Makaken in einer Richtungsunterscheidungsaufgabe (Gu et al., 2012).

Zusammengefasst zeigt diese große Anzahl an exzellenten Studien, dass neuronale Populationen sowie Einzelneurone in MSTd visuelle und vestibuläre Signale proportional zu ihrer Verlässlichkeit in Repräsentationen der Eigenbewegung integrieren, welche relevant für die Wahrnehmung und Entscheidungen der nicht-humanen Primaten sind. Während diese Verarbeitung im Einklang mit dem Maximum Likelihood Estimation – Modell steht, können sie auch durch Mechanismen der divisiven Normalisierung erklärt werden (Ohshiro et al., 2011, 2017), eines kanonisch neuronalen Verarbeitungsmechanismus, der bereits für das visuelle System und Aufmerksamkeitsmodulierung vorgeschlagen wurde (Carandini und Heeger, 2012). Darüberhinaus kann die divisive Normalisierung auch eine Verstärkung der Antwort erklären, welche maximal ist, wenn die Stärke des individuellen Signals schwach ist – ein Prinzip, das seit den wegweisenden Studien von Stein und Kollegen als inverse Effektivität bekannt ist (Stein und Meredith, 1993). Wie vom Prinzip der inversen Effektivität beschrieben, zeigten Studien mittels vielfältiger Methoden und in unterschiedlichen Spezies, dass die Operationen (d. h. super- vs. subadditive) der multisensorischen Integration von der Signalstärke, sowie der Antwort eines Neurons oder Voxels (i. e. Volumenelement in Neuroimaging) auf unisensorische Stimuli abhängt (Kayser et al., 2008; Siemann et al., 2015; Stanford et al., 2005; Stein und Meredith, 1993; Stein et al., 2014; Werner und Noppeney, 2010b).

Auf neuronaler Systemebene demonstrierten funktionelle Bildgebungsstudien im Menschen, dass höhere Assoziationscortices wie der intraparietale sulcus oder der sulcus temporalis superior sensorische Signale pro-

portional zu ihrer Verlässlichkeit in der Spracherkennung (Nath und Beauchamp, 2011), bei räumlicher Lokalisierung (Rohe und Noppeney, 2018) sowie Formunterscheidungsaufgaben (Beauchamp et al., 2010; Helbig et al., 2012) integrieren. Vor Kurzem untersuchten zwei funktionelle Bildgebungsstudien, wie das menschliche Gehirn die Bayesianische kausale Inferenz in einem räumlichen Ventriloquist-Paradigma vollzieht (Rohe und Noppeney, 2015a, 2016). Im Magnetresonanztomografen wurden Probanden audio-visuelle Signale präsentiert, wobei der Abstand zwischen dem auditorischen und visuellen Signal und die visuelle Verlässlichkeit von Durchlauf zu Durchlauf variierten. Bei jedem Durchlauf mussten die Probanden entweder den Ton oder den visuellen Stimulus lokalisieren. Gemeinsam zeigten die Forschungsergebnisse mittels Psychophysik, fMRI, Bayesianischer Modellierung und multivariater Analysemethoden, dass das Gehirn Bayesianische kausale Inferenz berechnet, indem es multiple räumliche Schätzwerte auf mehreren Ebenen der corticalen Hierarchie encodiert. Am Fuße der Hierarchie kodierten auditorische Areale mehrheitlich Schätzwerte für den Ort des Tones und visuelle Areale überwiegend für den Ort des visuellen Stimulus (= Segregation, d. h. Signaltrennung). Im hinteren Sulcus intraparietalis wird der Stimulusort unter der Annahme geschätzt, dass die zwei Signale von einer gemeinsamen Quelle ausgehen (= obligatorische Integration). Nur an der Spitze der Hierarchie, im vorderen Sulcus intraparietalis, wird die Ungewissheit der kausalen Struktur der Welt mitberücksichtigt. Wie durch das Modell der Bayesianischen kausalen Inferenz vorhergesagt wird beispielsweise der Ort des Tones berechnet, indem die Schätzwerte, welche man mittels Signaltrennung und obligatorischer Integration erhalten hat, gemittelt in Abhängigkeit der posterioren Wahrscheinlichkeiten einer gemeinsamen oder zweier unabhängiger Quellen ermittelt werden. Somit berechnet der anteriore IPS einen Schätzwert des Stimulusortes, der elegant von Integration zu Segregation übergeht als Funktion der räumlichen Distanz zwischen den auditorischen und visuellen Signalen (Rohe und Noppeney, 2015a, 2016).

Zusammenfassung

In unserer natürlichen Umgebung werden unsere Sinne unaufhörlich mit den unterschiedlichsten Signalen bombardiert. Idealerweise sollte unser Gehirn die Signale proportional zu ihrer Verlässlichkeit integrieren, wenn sie von einer gemeinsamen Quelle stammen. Demgegenüber sollte es sie unabhängig voneinander verarbeiten, wenn

sie von unterschiedlichen Quellen kommen. Tatsächlich unterscheiden menschliche Beobachter zwischen Integration und Segregation wie vom Bayesianischen Modell der kausalen Inferenz vorhergesagt. Neurophysiologische Studien in nicht-humanen Primaten und anderen Spezies konnten auf neuronaler Ebene zeigen, dass das Gehirn Signale, die von gemeinsamen Quellen stammen, proportional zu ihrer relativen Verlässlichkeit in eine kohärente Repräsentation der Welt integriert. Erste Neuroimaging-Studien am Menschen zeigen, dass das Gehirn sensorische Signale, wie vom Bayesianischen Modell der kausalen Inferenz vorhergesagt, integriert, indem es mehrere perzeptuelle Schätzwerte für die Szenarien einer gemeinsamen oder zweier verschiedener Quellen entlang der corticalen Hierarchie bildet. Um unser Verständnis der neuronalen Mechanismen und Verarbeitungsprinzipien und ihrer Verhaltensrelevanz zu vertiefen, sind weitere Studien in unterschiedlichen Spezies erforderlich, die Psychophysik, Computermodelle, Neurophysiologie und Neuroimaging miteinander kombinieren.

Exkurs: Multisensorische Bindung als Mechanismus für perzeptuelle Illusion

Die Berechnungen unseres Wahrnehmungssystems sind optimiert, effektive Interaktionen mit unserer natürlichen Umgebung zu gewährleisten. Im Labor können wir die Wahrnehmung von Versuchspersonen täuschen, indem wir sie in Situationen versetzen, welche die natürliche Statistik, für welche ihre Wahrnehmungssysteme optimiert wurden, auf den Kopf stellt. So können wir multisensorische Wahrnehmungstäuschungen dadurch hervorrufen, dass wir sensorische Signale entlang einer bestimmten Dimension (zum Beispiel Raum, Zeit, Anzahl, Phonem) in Konflikt miteinander bringen und gleichzeitig genügend multisensorische Korrespondenzen und Kongruenzen entlang einer anderen Dimension liefern. Damit können wir das Gehirn dazu überlisten, Signale, die eigentlich im Widerspruch zueinander stehen, trotzdem in eine einheitliche illusorische Wahrnehmung zu verschmelzen. Multisensorische Integration liegt zahlreichen Wahrnehmungssillusionen zugrunde. Im Folgenden werden wir die bekanntesten Beispiele vorstellen:

In der Doppelblitz-Illusion (double flash illusion, Shams et al., 2000) werden Probanden mit einem einzelnen Lichtblitz konfrontiert, der zwischen zwei Piep-Tönen erscheint. In den meisten Fällen berichten die Versuchs-

personen, dass sie zwei Lichtblitze sahen. Ihre Wahrnehmung ist somit stärker durch die Anzahl der Tonsignale als den zeitlich unpräzisen Blitz beeinflusst. Während sich die Illusion des Bauchredners der räumlichen Ungewissheit des Gehörs bedient (in den meisten Fällen können wir ein Objekt besser durch unser Seh- als unser Hörvermögen lokalisieren), nutzt die Doppelblitz-Illusion die zeitliche Ungewissheit des Sehvermögens (die Ohren sind verlässlicher als die Augen, um zu entscheiden, wann etwas stattgefunden hat). Daher kann die Doppelblitz-Illusion als das zeitliche Äquivalent zu dem Bauchredner-Effekt angesehen werden. Aber selbstverständlich sind zeitliche und räumliche Dimensionen nicht völlig vergleichbar. Während der räumliche Bauchredner-Effekt zeigt, wie das Gehirn die räumliche Lokalisation eines Ereignisses abschätzt (d. h. Schätzaufgabe), verdeutlicht die Doppelblitz-Illusion, wie das Gehirn die Anzahl von Ereignissen bestimmt (d. h. Detektionsaufgabe).

Genauso können auch multisensorische Sprachsignale manipuliert werden, um illusorische Wahrnehmungen zu erzeugen. In der sogenannten McGurk-McDonald-Illusion (McGurk und MacDonald, 1976) wird dem Proband gleichzeitig eine Videoaufnahme der Lippenbewegungen eines Sprechers (e. g. /ga/) und die Tonaufnahme eines Phonems (e. g. /ba/) präsentiert. Da auditorische und visuelle Reize synchron präsentiert werden, integriert die Versuchsperson die widersprüchlichen audiovisuellen Signale in eine illusorische /da/ Wahrnehmung. Diese illusorische Wahrnehmung kann wiederum durch das Maximum Likelihood Estimation – Modell erklärt werden. Mithilfe eines Sprachgenerators lässt sich eine künstliche 'Phonem' Dimension schaffen, die ein /ba/ Phonem in ein /ga/ Phonem morphet. Die Wahrnehmung von 'ba' - 'da' - 'ga' Phonem-Kategorien entsteht als Folge menschlicher kategorischer Wahrnehmung (Liberman et al., 1957).

Die Gummihand-Illusion ('rubber hand illusion', Botvinick und Cohen, 1998) ist ein Beispiel dafür, wie unsere eigene Körperwahrnehmung getäuscht werden kann. Sie entsteht dadurch, dass der visuelle und der propriozeptive Sinn, der der Körperwahrnehmung im Raum zugrunde liegt, in Konflikt miteinander gebracht werden. Bei der Gummihand-Illusion legt der Teilnehmer beispielsweise die rechte Hand auf den Tisch. Die rechte Hand wird dann durch eine Trennwand optisch vor dem Blick der Versuchsperson verborgen. Statt auf seine rechte Hand blickt der Teilnehmer auf eine Ersatz-Gummihand, die vor ihm genau parallel zu seiner eigenen Hand gut sichtbar platziert wird. Der Versuchsleiter streicht dann sowohl die echte als auch die Gummihand gleichzeitig mit Pinseln. Dabei achtet er darauf, dass die Pinselstriche auf beiden Armen und Händen synchron so parallel wie möglich

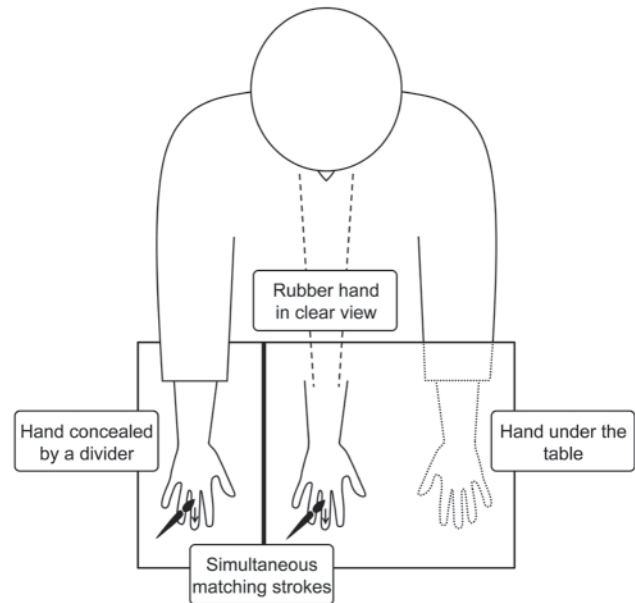


Abb. 3: Die Gummihand-Illusion. Der Teilnehmer wird an einen Tisch gesetzt. Während seine rechte Hand durch einen Schirm abgedeckt ist, blickt er auf eine Gummihand, die parallel zu seiner eigenen Hand liegt. Der Versuchsleiter streicht mit Pinseln gleichzeitig und parallel über die Hand des Teilnehmers und die Gummihand. Die meisten Versuchsteilnehmer nehmen dadurch die Gummihand als Teil ihres eigenen Körpers wahr. [Übersetzt von Dr. Joo-Hee Wälzlein]

durchgeführt werden. Dadurch, dass der Teilnehmer die Striche nun gleichzeitig auf der Gummihand sieht, sie jedoch auf der eigenen Hand fühlt, erhält er bald den Eindruck, dass die sichtbare Gummihand seine eigene Hand sei. Der visuelle Sinn ist also dominanter und überschreibt die propriozeptive Wahrnehmung des eigenen Armes im Raum, sodass ein Transfer des taktilen Gefühls in der rechten Hand auf die Gummihand erfolgt. Diese gefühlte Körperschema-Illusion ist bei vielen Menschen sehr ausgeprägt, auch wenn der Verstand ihnen sagt, dass die Gummihand kein Teil ihres Körpers ist. Um den maximalen Effekt dieses Phänomens zu erzielen, werden solche Demonstrationen oft damit beendet, dass der Versuchsleiter unerwartet mit einem Hammer auf die Gummihand schlägt.

Letztlich ist selbst unser Geschmackssinn nicht von multisensorischen Illusionen ausgenommen. Professor Charles Spence hat sich auf sensorische Wahrnehmung von Speisen spezialisiert. Seine Gruppe konnte zeigen, dass andere Sinne mitbestimmen, wie uns Speisen schmecken. Gewicht, die Beschaffenheit von Besteck (Harrar und Spence, 2013), Farbe und Form des Tellers (Piqueras-Fizman et al., 2012) und die Beschaffenheit der Verpackung (Piqueras-Fizman et al., 2012) beeinflussen, wie

wir Speisen wahrnehmen. Im Jahre 2008 wurde Professor Spence (zusammen mit seinem Kollegen Massimiliano Zampini) mit dem Ig Nobelpreis für die Entdeckung ausgezeichnet, dass Kartoffelchips mittels digitaler Tonmanipulation als knuspriger wahrgenommen werden.

Danksagung: Diese Studie wurde vom ERC-2012-StG_20111109 multisens finanziert.

Literatur

- Adam, R. and Noppeney, U. (2010). Prior auditory information shapes visual category-selectivity in ventral occipito-temporal cortex. *Neuroimage*, 52 (4), 1592–1602. <http://doi.org/10.1016/j.neuroimage.2010.05.002>
- Alais, D. and Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Curr. Biol.* 14 (3), 257–262. <http://doi.org/10.1016/j.cub.2004.01.029>
- Atilgan, H., Town, S., Wood, K., Jones, G., Maddox, R., Lee, A. and Bizley, J.K. (2018). Integration of visual information in auditory cortex promotes auditory scene analysis through multisensory binding. *Neuron*. 97 (3), 640–655. <http://doi.org/10.1016/j.neuron.2017.12.034>
- Avillac, M., Ben Hamed, S. and Duhamel, J.-R. (2007). Multisensory Integration in the Ventral Intraparietal Area of the Macaque Monkey. *J. Neurosci.* 27 (8), 1922–1932. <http://doi.org/10.1523/JNEUROSCI.2646-06.2007>
- Barracough, N.E., Xiao, D., Baker, C.I., Oram, M.W. and Perrett, D.I. (2005). Integration of Visual and Auditory Information by Superior Temporal Sulcus Neurons Responsive to the Sight of Actions. *J. Cogn. Neurosci.* 17 (3), 377–391. <http://doi.org/10.1162/0898929053279586>
- Battaglia, P.W., Jacobs, R.A. and Aslin, R.N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *J. Opt. Soc. Am. A*. 20 (7), 1391–1397. <http://doi.org/10.1364/JOSAA.20.001391>
- Battaglia, P.W., Kersten, D. and Schrater, P.R. (2011). How haptic size sensations improve distance perception. *PLoS Comput. Biol.* 7 (6). <http://doi.org/10.1371/journal.pcbi.1002080>
- Beauchamp, M.S., Lee, K., Argall, B. and Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*. 41, 809–823. [http://doi.org/10.1016/S0896-6273\(04\)00070-4](http://doi.org/10.1016/S0896-6273(04)00070-4)
- Beauchamp, M.S., Pasalar, S. and Ro, T. (2010). Neural substrates of reliability-weighted visual-tactile multisensory integration. *Front. Syst. Neurosci.* 4, 1–11. <http://doi.org/10.3389/fnsys.2010.00025>
- Beierholm, U.R., Quartz, S.R. and Shams, L. (2009). Bayesian priors are encoded independently from likelihoods in human multisensory perception. *J. Vis.* 9 (5), 23. <http://doi.org/10.1167/9.5.23>
- Bentvelzen, A., Leung, J. and Alais, D. (2009). Discriminating audiovisual speed: Optimal integration of speed defaults to probability summation when component reliabilities diverge. *Perception*. 38 (7), 966–987. <http://doi.org/10.1068/p6261>
- Bertelson, P. and Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Percept. Psychophys.* 29 (6), 578–584. <http://doi.org/10.3758/BF03214277>
- Besle, J., Fischer, C., Bidet-Caulet, A., Lecaigard, F., Bertrand, O. and Giard, M.-H. (2008). Visual Activation and Audiovisual Interactions in the Auditory Cortex during Speech Perception: Intracranial Recordings in Humans. *J. Neurosci.* 28 (52), 14301–14310. <http://doi.org/10.1523/JNEUROSCI.2875-08.2008>
- Bieler, M., Sieben, K., Cichon, N., Schildt, S., Röder, B. and Hanganu-Opatz, I.L. (2017). Rate and Temporal Coding Convey Multisensory Information in Primary Sensory Cortices. *eNeuro*. 4 (2), ENEURO-0037. <http://doi.org/10.1523/ENEURO.0037-17.2017>
- Bishop, C.W. and Miller, L.M. (2011). Speech cues contribute to audiovisual spatial integration. *PLOS ONE*. 6 (8). <http://doi.org/10.1371/journal.pone.0024016>
- Bizley, J.K. and King, A.J. (2009). Visual influences on ferret auditory cortex. *Hear. Res.* 258 (1–2), 55–63. <http://doi.org/10.1016/j.heares.2009.06.017>
- Bizley, J.K., Nodal, F.R., Bajo, V.M., Nelken, I. and King, A.J. (2006). Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cereb. Cortex*. 17 (9), 2172–2189.
- Bonath, B., Noesselt, T., Martinez, A., Mishra, J., Schwiecker, K., Heinze, H.-J. and Hillyard, S. a. (2007). Neural Basis of the Ventriloquist Illusion. *Curr. Biol.* 17 (19), 1697–1703. <http://doi.org/10.1016/j.cub.2007.08.050>
- Botvinick, M. and Cohen, J. (1998). Rubber hands “feel” touch that eyes see. *Nature*. 391 (6669), 756. <http://doi.org/10.1038/35784>
- Bresciani, J.-P., Dammeier, F. and Ernst, M.O. (2006). Vision and touch are automatically integrated for the perception of sequences of events. *J. Vis.* 6 (5), 2. <http://doi.org/10.1167/6.5.2>
- Budinger, E., Heil, P., Hess, A. and Scheich, H. (2006). Multisensory processing via early cortical stages: connections of the primary auditory cortical field with other sensory systems. *Neuroscience*. 143 (4), 1065–1083. <https://doi.org/10.1016/j.neuroscience.2006.08.035>
- Burr, D., Banks, M.S. and Morrone, M.C. (2009). Auditory dominance over vision in the perception of interval duration. *Exp. Brain Res.* 198 (1), 49–57. <http://doi.org/10.1007/s00221-009-1933-z>
- Butler, J.S., Smith, S.T., Campos, J.L. and Bühlhoff, H.H. (2010). Bayesian integration of visual and vestibular signals for heading. *J. Vis.* 10 (11), 23. <http://doi.org/10.1167/10.11.23>
- Calvert, G.A., Campbell, R. and Brammer, M.J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr. Biol.* 10 (11), 649–657. [http://doi.org/10.1016/S0969-9822\(00\)00513-3](http://doi.org/10.1016/S0969-9822(00)00513-3)
- Campi, K.L., Bales, K.L., Grunewald, R. and Krubitzer, L. (2009). Connections of auditory and visual cortex in the prairie vole (*Microtus ochrogaster*): evidence for multisensory processing in primary sensory areas. *Cereb. Cortex*. 20 (1), 89–108. <http://doi.org/10.1093/cercor/bhp082>
- Carandini, M. and Heeger, D. (2012). Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* 13 (1), 51–62. <http://doi.org/10.1038/nrn3136>. Normalization

- Cecere, R., Rees, G. and Romei, V. (2015). Individual differences in alpha frequency drive crossmodal illusory perception. *Curr. Biol.* 25 (2), 231–235. <http://doi.org/10.1016/j.cub.2014.11.034>
- Deroy, O., Spence, C. and Noppeney, U. (2016). Metacognition in Multisensory Perception. *Trends Cogn. Sci.* 20 (10), 736–747. <http://doi.org/10.1016/j.tics.2016.08.006>
- Diederich, A. and Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Percept. Psychophys.* 66 (8), 1388–1404. <http://doi.org/10.3758/BF03195006>
- Doehrmann, O. and Naumer, M.J. (2008). Semantics and the multisensory brain: How meaning modulates processes of audio-visual integration. *Brain Res.* 1242, 136–150. <http://doi.org/10.1016/j.brainres.2008.03.071>
- Driver, J. (1996). Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature.* 381 (6577), 66. <http://doi.org/10.1038/381066a0>
- Driver, J. and Noesselt, T. (2008). Multisensory Interplay Reveals Crossmodal Influences on “Sensory-Specific” Brain Regions, Neural Responses, and Judgments. *Neuron.* 57 (1), 11–23. <http://doi.org/10.1016/j.neuron.2007.12.013>
- Ernst, M.O. and Banks, M.S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature.* 415 (6870), 429–433. <http://doi.org/10.1038/415429a>
- Ernst, M.O. and Bühlhoff, H.H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci.* 8 (4), 162–169. <http://doi.org/10.1016/j.tics.2004.02.002>
- Falchier, A., Clavagnier, S., Barone, P. and Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.* 22 (13), 5749–5759. <http://doi.org/20026562>
- Fetsch, C.R., Pouget, A., Deangelis, G.C. and Angelaki, D.E. (2012). Neural correlates of reliability-based cue weighting during multisensory integration. *Nat. Neurosci.* 15 (1), 146–154. <http://doi.org/10.1038/nn.2983>
- Foxe, J.J., Morocz, I.A., Murray, M.M., Higgins, B.A., Javitt, D.C. and Schroeder, C.E. (2000). Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Cogn. Brain Res.* 10 (1), 77–83. [http://doi.org/10.1016/S0926-6410\(00\)00024-0](http://doi.org/10.1016/S0926-6410(00)00024-0)
- Foxe, J.J. and Schroeder, C.E. (2005). The case for feedforward multisensory convergence during early cortical processing. *NeuroReport.* 16 (5), 419–423. <http://doi.org/10.1097/00001756-200504040-00001>
- Foxe, J.J., Wylie, G.R., Martinez, A., Schroeder, C.E., Javitt, D.C., Guilfoyle, D., Ritter, W. and Murray, M.M. (2002). Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J. Neurophysiol.* 88 (1), 540–543. <http://doi.org/10.1152/jn.00694.2001>
- Frassinetti, F., Bolognini, N. and Ládavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Exp. Brain Res.* 147 (3), 332–343. <http://doi.org/10.1007/s00221-002-1262-y>
- Gepshtein, S. and Banks, M.S. (2003). Viewing Geometry Determines How Vision and Haptics Combine in Size Perception. *Curr. Biol.* 13 (6), 483–488. [http://doi.org/10.1016/S0960-9822\(03\)00133-7](http://doi.org/10.1016/S0960-9822(03)00133-7)
- Ghazanfar, A.A., Chandrasekaran, C. and Logothetis, N.K. (2008). Interactions between the Superior Temporal Sulcus and Auditory Cortex Mediate Dynamic Face/Voice Integration in Rhesus Monkeys. *J. Neurosci.* 28 (17), 4457–4469. <http://doi.org/10.1523/JNEUROSCI.0541-08.2008>
- Ghazanfar, A.A., Maier, J.X., Hoffman, K.L. and Logothetis, N.K. (2005). Multisensory Integration of Dynamic Faces and Voices in Rhesus Monkey Auditory Cortex. *J. Neurosci.* 25 (20), 5004–5012. <http://doi.org/10.1523/JNEUROSCI.0799-05.2005>
- Ghazanfar, A.A. and Schroeder, C.E. (2006). Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10 (6), 278–285. <http://doi.org/10.1016/j.tics.2006.04.008>
- Giard, M.H. and Peronnet, F. (1999). Auditory-Visual Integration during Multimodal Object Recognition in Humans: A Behavioral and Electrophysiological Study. *J. Cogn. Neurosci.* 11 (5), 473–490. <http://doi.org/10.1162/089892999563544>
- Gillmeister, H. and Eimer, M. (2007). Tactile enhancement of auditory detection and perceived loudness. *Brain Res.* 1160 (1), 58–68. <http://doi.org/10.1016/j.brainres.2007.03.041>
- Gu, Y., Angelaki, D.E. and DeAngelis, G.C. (2008). Neural correlates of multisensory cue integration in macaque MSTd. *Nat. Neurosci.* 11 (10), 1201–1210. <http://doi.org/10.1038/nn.2191>
- Gu, Y., DeAngelis, G.C. and Angelaki, D.E. (2012). Causal Links between Dorsal Medial Superior Temporal Area Neurons and Multisensory Heading Perception. *J. Neurosci.* 32 (7), 2299–2313. <http://doi.org/10.1523/JNEUROSCI.5154-11.2012>
- Harrar, V. and Spence, C. (2013). The taste of cutlery: how the taste of food is affected by the weight, size, shape, and colour of the cutlery used to eat it. *Flavour.* 2 (1), 21. <http://doi.org/10.1186/2044-7248-2-21>
- 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. <http://doi.org/10.1016/j.neuroimage.2011.09.072>
- Hillis, J.M., Watt, S.J., Landy, M.S. and Banks, M.S. (2004). Slant from texture and disparity cues: Optimal cue combination. *J. Vis.* 4 (12), 1. <http://doi.org/10.1167/4.12.1>
- Hollensteiner, K.J., Pieper, F., Engler, G., König, P. and Engel, A.K. (2015). Crossmodal integration improves sensory detection thresholds in the ferret. *PloS One.* 10 (5), e0124952. <https://doi.org/10.1371/journal.pone.0124952>
- Ibrahim, L.A., Mesik, L., Ji, X.Y., Fang, Q., Li, H.F., Li, Y.T., Zingg, B., Zhang, L.I. and Tao, H.W. (2016). Cross-modality sharpening of visual cortical processing through layer-1-mediated inhibition and disinhibition. *Neuron.* 89 (5), 1031–1045. <http://doi.org/10.1016/j.neuron.2016.01.027>
- Iurilli, G., Ghezzi, D., Olcese, U., Lassi, G., Nazzaro, C., Tonini, R., Tucci, V., Bonfenati, F. and Medini, P. (2012). Sound-driven synaptic inhibition in primary visual cortex. *Neuron.* 73 (4), 814–828. <http://doi.org/10.1016/j.neuron.2011.12.026>
- Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth. *Vis. Res.* 39 (21), 3621–3629. [http://doi.org/10.1016/S0042-6989\(99\)00088-7](http://doi.org/10.1016/S0042-6989(99)00088-7)
- Kanaya, S. and Yokosawa, K. (2011). Perceptual congruency of audio-visual speech affects ventriloquism with bilateral visual stimuli. *Psychon. Bull. Rev.* 18 (1), 123–128. <http://doi.org/10.3758/s13423-010-0027-z>
- Kayser, C. and Logothetis, N.K. (2007). Do early sensory cortices integrate cross-modal information? *Brain Struct. Funct.* 212 (2), 121–132. <http://doi.org/10.1007/s00429-007-0154-0>
- Kayser, C., Logothetis, N.K. and Panzeri, S. (2010). Visual Enhancement of the Information Representation in Auditory

- Cortex. *Curr. Biol.* 20 (1), 19–24. <http://doi.org/10.1016/j.cub.2009.10.068>
- Kayser, C., Petkov, C.I. and Logothetis, N.K. (2008). Visual modulation of neurons in auditory cortex. *Cereb. Cortex.* 18 (7), 1560–1574. <http://doi.org/10.1093/cercor/bhm187>
- Kersten, D., Mamassian, P. and Yuille, A. (2004). Object Perception as Bayesian Inference. *Annu. Rev. Psychol.* 55, 271–304. <http://doi.org/10.1146/psych.2004.55.issue-1>
- Kersten, D. and Yuille, A. (2003). Bayesian models of object perception. *Curr. Opin. Neurobiol.* 13 (2), 150–158. [http://doi.org/10.1016/S0959-4388\(03\)00042-4](http://doi.org/10.1016/S0959-4388(03)00042-4)
- Knill, D.C. and Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends Neurosci.* 27 (12), 712–719. <http://doi.org/10.1016/j.tins.2004.10.007>
- Knill, D.C. and Saunders, J.A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vis. Res.* 43 (24), 2539–2558. [http://doi.org/10.1016/S0042-6989\(03\)00458-9](http://doi.org/10.1016/S0042-6989(03)00458-9)
- Körding, K.P., Beierholm, U.R., Ma, W.J., Quartz, S.R., Tenenbaum, J.B. and Shams, L. (2007). Causal inference in multisensory perception. *PLoS One.* 2 (9), e943. <http://doi.org/10.1371/journal.pone.0000943>
- Krugliak, A. and Noppeney, U. (2015). Synesthetic interactions across vision and audition. *Neuropsychologia.* 88, 65–73. <http://doi.org/10.1016/j.neuropsychologia.2015.09.027>
- Lakatos, P., Chen, C.M., O’Connell, M.N., Mills, A. and Schroeder, C.E. (2007). Neuronal Oscillations and Multisensory Interaction in Primary Auditory Cortex. *Neuron.* 53 (2), 279–292. <http://doi.org/10.1016/j.neuron.2006.12.011>
- Lakatos, P., O’Connell, M.N., Barczak, A., Mills, A., Javitt, D.C. and Schroeder, C.E. (2009). The Leading Sense: Supramodal Control of Neurophysiological Context by Attention. *Neuron.* 64 (3), 419–430. <http://doi.org/10.1016/j.neuron.2009.10.014>
- Landy, M.S., Maloney, L.T., Johnston, E.B. and Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vis. Res.* 35 (3), 389–412. [http://doi.org/10.1016/0042-6989\(94\)00176-M](http://doi.org/10.1016/0042-6989(94)00176-M)
- Lee, H. and Noppeney, U. (2011a). Long-term music training tunes how the brain temporally binds signals from multiple senses. *Proc. Nat. Acad. Sci. U.S.A.* 108 (51), E1441–50. <http://doi.org/10.1073/pnas.1115267108>
- Lee, H. and Noppeney, U. (2011b). Physical and Perceptual Factors Shape the Neural Mechanisms That Integrate Audiovisual Signals in Speech Comprehension. *J. Neurosci.* 31 (31), 11338–11350. <http://doi.org/10.1523/JNEUROSCI.6510-10.2011>
- Lee, H. and Noppeney, U. (2014). Temporal prediction errors in visual and auditory cortices. *Curr. Biol.* 24 (8), R309–R310. <http://doi.org/10.1016/j.cub.2014.02.007>
- Lehmann, C., Herdener, M., Esposito, F., Hubl, D., di Salle, F., Scheffler, K., ... Seifritz, E. (2006). Differential patterns of multisensory interactions in core and belt areas of human auditory cortex. *NeuroImage.* 31 (1), 294–300. <http://doi.org/10.1016/j.neuroimage.2005.12.038>
- Leitão, J., Thielscher, A., Werner, S., Pohmann, R. and Noppeney, U. (2012). Effects of parietal TMS on visual and auditory processing at the primary cortical level—a concurrent TMS-fMRI study. *Cereb. Cortex.* 23 (4), 873–884. <http://doi.org/10.1093/cercor/bhs078>
- Lewald, J. and Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Cogn. Brain Res.* 16 (3), 468–478. [http://doi.org/10.1016/S0926-6410\(03\)00074-0](http://doi.org/10.1016/S0926-6410(03)00074-0)
- Lewis, R. and Noppeney, U. (2010). Audiovisual Synchrony Improves Motion Discrimination via Enhanced Connectivity between Early Visual and Auditory Areas. *J. Neurosci.* 30 (37), 12329–12339. <http://doi.org/10.1523/JNEUROSCI.5745-09.2010>
- Lieberman, A.M., Harris, K.S., Hoffman, H.S. and Griffith, B.C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *J. Exp. Psychol.* 54 (5), 358–368. <http://doi.org/10.1037/h0044417>
- Macaluso, E., Driver, J. and Frith, C.D. (2003). Multimodal Spatial Representations Engaged in Human Parietal Cortex during Both Saccadic and Manual Spatial Orienting. *Curr. Biol.* 13 (12), 990–999. [http://doi.org/10.1016/S0960-9822\(03\)00377-4](http://doi.org/10.1016/S0960-9822(03)00377-4)
- Macaluso, E., Frith, C.D. and Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science.* 289 (5482), 1206–1208. <http://doi.org/10.1126/science.289.5482.1206>
- Magnotti, J.F. and Beauchamp, M.S. (2017). A Causal Inference Model Explains Perception of the McGurk Effect and Other Incongruent Audiovisual Speech. *PLoS Comput. Biol.* 13 (2), e1005229. <http://doi.org/10.1371/journal.pcbi.1005229>
- Magnotti, J.F., Ma, W.J. and Beauchamp, M.S. (2013). Causal inference of asynchronous audiovisual speech. *Front. Psychol.* 4, 1–10. <http://doi.org/10.3389/fpsyg.2013.00854>
- Maier, J.X., Di Luca, M. and Noppeney, U. (2011). Audiovisual Asynchrony Detection in Human Speech. *J. Exp. Psychol. Hum. Percept. Perform.* 37 (1), 245–256. <http://doi.org/10.1037/a0019952>
- Marks, L.E. (1987). On Cross-Modal Similarity: Auditory-Visual Interactions in Speeded Discrimination. *J. Exp. Psychol. Hum. Percept. Perform.* 13 (3), 384–394. <http://doi.org/10.1037/0096-1523.13.3.384>
- Martuzzi, R., Murray, M.M., Michel, C.M., Thiran, J.P., Maeder, P.P., Clarke, S. and Meuli, R.A. (2007). Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cereb. Cortex.* 17 (7), 1672–1679. <http://doi.org/10.1093/cercor/bhl077>
- Mcgurk, H. and Macdonald, J. (1976). Hearing lips and seeing voices. *Nature.* 264 (5588), 746–748. <http://doi.org/10.1038/264746a0>
- Meijer, G.T., Montijn, J.S., Pennartz, C.M. and Lansink, C.S. (2017). Audiovisual Modulation in Mouse Primary Visual Cortex Depends on Cross-Modal Stimulus Configuration and Congruency. *J. Neurosci.* 37 (36), 8783–8796. <http://doi.org/10.1523/JNEUROSCI.0468-17.2017>
- Meredith, M.A. and Allman, B.L. (2015). Single-unit analysis of somatosensory processing in the core auditory cortex of hearing ferrets. *Eur. J. Neurosci.* 41 (5), 686–698. <http://doi.org/10.1111/ejn.12828>
- Meredith, M. and Stein, B. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science.* 221 (4608), 389–391. <http://doi.org/10.1126/science.6867718>
- Miller, L.M. and D’Esposito, M. (2005). Perceptual Fusion and Stimulus Coincidence in the Cross-Modal Integration of Speech. *J. Neurosci.* 25 (25), 5884–5893. <http://doi.org/10.1523/JNEUROSCI.0896-05.2005>
- Molholm, S., Ritter, W., Javitt, D.C. and Foxe, J.J. (2004). Multisensory Visual-Auditory Object Recognition in Humans: A High-density Electrical Mapping Study. *Cereb. Cortex.* 14 (4), 452–465. <http://doi.org/10.1093/cercor/bhh007>

- Molholm, S., Ritter, W., Murray, M.M., Javitt, D.C., Schroeder, C.E. and Foxe, J.J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cogn. Brain Res.* 14 (1), 115–128. [http://doi.org/10.1016/S0926-6410\(02\)00066-6](http://doi.org/10.1016/S0926-6410(02)00066-6)
- Munhall, K.G., Gribble, P., Sacco, L. and Ward, M. (1996). Temporal constraints on the McGurk effect. *Percept. Psychophys.* 58 (3), 351–362. <http://doi.org/10.3758/BF03206811>
- Musacchia, G., Large, E.W. and Schroeder, C.E. (2014). Thalamo-cortical mechanisms for integrating musical tone and rhythm. *Hear. Res.* 308, 50–59. <https://doi.org/10.1016/j.heares.2013.09.017>
- Musacchia, G. and Schroeder, C.E. (2009). Neuronal mechanisms, response dynamics and perceptual functions of multisensory interactions in auditory cortex. *Hear. Res.* 258 (1–2), 72–79. <http://doi.org/10.1016/j.heares.2009.06.018>
- Nath, A.R. and Beauchamp, M.S. (2011). Dynamic Changes in Superior Temporal Sulcus Connectivity during Perception of Noisy Audiovisual Speech. *J. Neurosci.* 31 (5), 1704–1714. <http://doi.org/10.1523/JNEUROSCI.4853-10.2011>
- Noesselt, T., Bergmann, D., Hake, M., Heinze, H.J. and Fendrich, R. (2008). Sound increases the saliency of visual events. *Brain Res.* 1220, 157–163. <http://doi.org/10.1016/j.brainres.2007.12.060>
- Noesselt, T., Rieger, J.W., Schoenfeld, M.A., Kanowski, M., Hinrichs, H., Heinze, H.-J. and Driver, J. (2007). Audiovisual Temporal Correspondence Modulates Human Multisensory Superior Temporal Sulcus Plus Primary Sensory Cortices. *J. Neurosci.* 27 (42), 11431–11441. <http://doi.org/10.1523/JNEUROSCI.2252-07.2007>
- Noppeney, U., Josephs, O., Hocking, J., Price, C.J. and Friston, K.J. (2008). The effect of prior visual information on recognition of speech and sounds. *Cereb. Cortex.* 18 (3), 598–609. <http://doi.org/10.1093/cercor/bhm091>
- Noppeney, U., Ostwald, D. and Werner, S. (2010). Perceptual Decisions Formed by Accumulation of Audiovisual Evidence in Prefrontal Cortex. *J. Neurosci.* 30 (21), 7434–7446. <http://doi.org/10.1523/JNEUROSCI.0455-10.2010>
- Nikbakht, N., Tafreshi, A., Zoccolan, D. and Diamond, M.E. (2018). Supralinear and supramodal integration of visual and tactile signals in rats: psychophysics and neuronal mechanisms. *Neuron.* 97 (3), 626–639.e8. <http://doi.org/10.1016/j.neuron.2018.01.003>
- Ohshiro, T., Angelaki, D.E. and DeAngelis, G.C. (2011). A normalization model of multisensory integration. *Nat. Neurosci.* 14 (6), 775–782. <http://doi.org/10.1038/nn.2815>
- Ohshiro, T., Angelaki, D.E. and DeAngelis, G.C. (2017). A Neural Signature of Divisive Normalization at the Level of Multisensory Integration in Primate Cortex. *Neuron.* 95 (2), 399–411. <http://doi.org/10.1016/j.neuron.2017.06.043>
- Parise, C.V. and Spence, C. (2009). “When birds of a feather flock together”: Synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS One.* 4 (5), e5664. <http://doi.org/10.1371/journal.pone.0005664>
- Parise, C.V. and Ernst, M.O. (2016). Correlation detection as a general mechanism for multisensory integration. *Nature Commun.* 7 (12), 11543. <http://doi.org/10.1038/ncomms11543>
- Parise, C.V., Spence, C. and Ernst, M.O. (2012). When correlation implies causation in multisensory integration. *Curr. Biol.* 22 (1), 46–49. <http://doi.org/10.1016/j.cub.2011.11.039>
- Piqueras-Fiszman, B., Alcaide, J., Roura, E. and Spence, C. (2012). Is it the plate or is it the food? Assessing the influence of the color (black or white) and shape of the plate on the perception of the food placed on it. *Food Qual. Pref.* 24 (1), 205–208. <http://doi.org/10.1016/j.foodqual.2011.08.011>
- Roach, N.W., Heron, J. and McGraw, P.V. (2006). Resolving multisensory conflict: a strategy for balancing the costs and benefits of audio-visual integration. *Proc. Biol. Sci.* 273 (1598), 2159–68. <http://doi.org/10.1098/rspb.2006.3578>
- Rockland, K.S. and Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *Int. J. Psychophysiol.* 50 (1–2), 19–26. [http://doi.org/10.1016/S0167-8760\(03\)00121-1](http://doi.org/10.1016/S0167-8760(03)00121-1)
- Rohe, T. and Noppeney, U. (2015a). Cortical Hierarchies Perform Bayesian Causal Inference in Multisensory Perception. *PLoS Biol.* 13 (2), e1002073. <http://doi.org/10.1371/journal.pbio.1002073>
- Rohe, T. and Noppeney, U. (2015b). Sensory reliability shapes Bayesian Causal Inference in perception via two mechanisms. *J. Vis.* 15, 1–38. <http://doi.org/10.1167/15.5.22>
- Rohe, T. and Noppeney, U. (2016). Distinct computational principles govern multisensory integration in primary sensory and association cortices. *Curr. Biol.* 1 (4), 509–514. <http://doi.org/10.1016/j.cub.2015.12.056>
- Rohe, T. and Noppeney, U. (2018). Reliability-Weighted integration of audiovisual signals can be modulated by top-down control. *eNeuro.* <https://doi.org/10.1523/ENEURO.0315-17.2018>
- Rosas, P., Wagemans, J., Ernst, M.O. and Wichmann, F.A. (2005). Texture and haptic cues in slant discrimination: reliability-based cue weighting without statistically optimal cue combination. *J. Opt. Soc. Am. A.* 22 (5), 801–809. <https://doi.org/10.1364/JOSAA.22.000801>
- Ross, L.A., Saint-Amour, D., Leavitt, V.M., Javitt, D.C. and Foxe, J.J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cereb. Cortex.* 17 (5), 1147–1153. <http://doi.org/10.1093/cercor/bhl024>
- Sadaghiani, S., Maier, J.X. and Noppeney, U. (2009). Natural, Metaphoric, and Linguistic Auditory Direction Signals Have Distinct Influences on Visual Motion Processing. *J. Neurosci.* 29 (20), 6490–6499. <http://doi.org/10.1523/JNEUROSCI.5437-08.2009>
- Samaha, J. and Postle, B.R. (2015). The Speed of Alpha-Band Oscillations Predicts the Temporal Resolution of Visual Perception. *Curr. Biol.* 25 (22), 2985–2990. <http://doi.org/10.1016/j.cub.2015.10.007>
- Schroeder, C.E. and Foxe, J.J. (2002). The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Cogn. Brain Res.* 14 (1), 187–198. [http://doi.org/10.1016/S0926-6410\(02\)00073-3](http://doi.org/10.1016/S0926-6410(02)00073-3)
- Schroeder, C.E. and Foxe, J.J. (2005). Multisensory contributions to low-level, “unisensory” processing. *Curr. Opin. Neurobiol.* 15 (4), 454–458. <http://doi.org/10.1016/j.conb.2005.06.008>
- Schroeder, C.E., Lakatos, P., Kajikawa, Y., Partan, S. and Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends Cogn. Sci.* 12 (3), 106–113. <http://doi.org/10.1016/j.tics.2008.01.002>
- Schroeder, C.E., Smiley, J., Fu, K.G., McGinnis, T., O’Connell, M.N. and Hackett, T.A. (2003). Anatomical mechanisms and functional implications of multisensory convergence in early

- cortical processing. *Int. J. Psychophysiol.* 50 (1–2), 5–17. [http://doi.org/10.1016/S0167-8760\(03\)00120-X](http://doi.org/10.1016/S0167-8760(03)00120-X)
- Schwartz, J.L., Berthommier, F. and Savariaux, C. (2004). Seeing to hear better: Evidence for early audio-visual interactions in speech identification. *Cognition.* 93 (2), 69–78. <http://doi.org/10.1016/j.cognition.2004.01.006>
- Seltzer, B. and Pandya, D.N. (1994). Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: A retrograde tracer study. *J. Comp. Neurol.* 343 (3), <http://doi.org/10.1002/cne.903430308>
- Shams, L. and Beierholm, U.R. (2010). Causal inference in perception. *Trends Cogn. Sci.* 14 (9), 425–432. <http://doi.org/10.1016/j.tics.2010.07.001>
- Shams, L., Kamitani, Y. and Shimojo, S. (2000). What you see is what you hear. *Nature.* 408, 2000. <http://doi.org/10.1038/35048669>
- Sieben, K., Röder, B. and Hanganu-Opatz, I.L. (2013). Oscillatory entrainment of primary somatosensory cortex encodes visual control of tactile processing. *J. Neurosci.* 33 (13), 5736–5749. <http://doi.org/10.1523/JNEUROSCI.4432-12.2013>
- Siemann, J.K., Muller, C.L., Bamberger, G., Allison, J.D., Veenstra-VanderWeele, J. and Wallace, M.T. (2015). A novel behavioral paradigm to assess multisensory processing in mice. *Front. Behav. Neurosci.* 8, 456. <http://doi.org/10.3389/fnbeh.2014.00456>
- Slutsky, D. and Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *NeuroReport.* 12 (1), 7–10. <http://doi.org/10.1097/00001756-200101220-00009>
- Soto-Faraco, S. and Alsius, A. (2009). Deconstructing the McGurk-MacDonald Illusion. *J. Exp. Psychol. Hum. Percept. Perform.* 35 (2), 580–587. <http://doi.org/10.1037/a0013483>
- Soto-Faraco, S., Kingstone, A. and Spence, C. (2006). Integrating motion information across sensory modalities: The role of top-down factors. *Progress Brain Res.* 155, 273–286. [http://doi.org/10.1016/S0079-6123\(06\)55016-2](http://doi.org/10.1016/S0079-6123(06)55016-2)
- Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Ann. NY Acad. Sci.* 1296 (1), 31–49. <http://doi.org/10.1111/nyas.12121>
- Stanford, T.R. (2005). Evaluating the Operations Underlying Multisensory Integration in the Cat Superior Colliculus. *J. Neurosci.* 25 (28), 6499–6508. <http://doi.org/10.1523/JNEUROSCI.5095-04.2005>
- Stanford, T.R. and Stein, B.E. (2007). Superadditivity in multisensory integration: Putting the computation in context. *NeuroReport.* 18 (8), 787–792. <http://doi.org/10.1097/WNR.0b013e3280c1e315>
- Stein, B.E. and Meredith, M.A. (1993). *The merging of the senses.* The MIT Press.
- Stein, B.E. and Stanford, T.R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* 9 (4), 255–266. <http://doi.org/10.1038/nrn2331>
- Stein, B.E., Stanford, T.R. and Rowland, B.A. (2014). Development of multisensory integration from the perspective of the individual neuron. *Nat. Rev. Neurosci.* 15 (8), 520.
- Stevenson, R.A. and James, T.W. (2009). Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. *NeuroImage.* 44 (3), 1210–1223. <http://doi.org/10.1016/j.neuroimage.2008.09.034>
- Sugihara, T., Diltz, M.D., Averbeck, B.B. and Romanski, L.M. (2006). Integration of Auditory and Visual Communication Information in the Primate Ventrolateral Prefrontal Cortex. *J. Neurosci.* 26 (43), 11138–11147. <http://doi.org/10.1523/JNEUROSCI.3550-06.2006>
- Van Atteveldt, N., Formisano, E., Goebel, R. and Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron.* 43 (2), 271–282. <http://doi.org/10.1016/j.neuron.2004.06.025>
- van Wassenhove, V., Grant, K.W. and Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia.* 45 (3), 598–607. <http://doi.org/10.1016/j.neuropsychologia.2006.01.001>
- Vox, V. (1981). *I can see your lips moving: The history and art of ventriloquism.* Kaye & Ward.
- Vroomen, J. and Gelder, B. de. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *J. Exp. Psychol. Hum. Percept. Perform.* 26 (5), 1583–1590. <http://doi.org/10.1037/0096-1523.26.5.1583>
- Wallace, M.T., Roberson, G.E., Hairston, W.D., Stein, B.E., Vaughan, J.W. and Schirillo, J. (2004). Unifying multisensory signals across time and space. *Exp. Brain Res.* 158 (2), 252–258. <http://doi.org/10.1007/s00221-004-1899-9>
- Wallace, M.T., Ramachandran, R. and Stein, B.E. (2004). A revised view of sensory cortical parcellation. *Proc. Nat. Acad. Sci.,* 101 (7), 2167–2172. <http://doi.org/10.1073/pnas.0305697101>
- Wallace, M.T., Wilkinson, L.K. and Stein, B.E. (1996). Representation and integration of multiple sensory inputs in primate superior colliculus. *J. Neurophysiol.* 76 (2), 1246–1266. <http://doi.org/10.1152/jn.1996.76.2.1246>
- Werner, S. and Noppeney, U. (2010a). Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. *J. Neurosci.* 30 (7), 2662–2675. <http://doi.org/10.1523/JNEUROSCI.5091-09.2010>
- Werner, S. and Noppeney, U. (2010b). Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. *Cereb. Cortex.* 20 (8), 1829–1842. <http://doi.org/10.1093/cercor/bhp248>
- Werner, S. and Noppeney, U. (2011). The contributions of transient and sustained response codes to audiovisual integration. *Cereb. Cortex.* 21 (4), 920–931. <http://doi.org/10.1093/cercor/bhq161>
- Wozny, D.R., Beierholm, U.R. and Shams, L. (2010). Probability Matching as a Computational Strategy Used in Perception. *PLoS Comput. Biol.* 6 (8), e1000871. <http://doi.org/10.1371/journal.pcbi.1000871>
- Zampini, M. and Spence, C. (2004). The role of auditory cues in modulating the perceived crispness and staleness of potato chips. *J. Sens. Stud.* 19 (5), 347–363. <http://doi.org/10.1111/j.1745-459x.2004.080403.x>

Anmerkung: Die englische Version des Artikels ist online verfügbar unter <https://doi.org/10.1515/nf-2017-A066>

Autoreninformation



Uta Noppeney
 Computational Neuroscience and Cognitive
 Robotics Centre
 School of Psychology
 University of Birmingham
 Edgbaston
 Birmingham
 B15 2TT
 Großbritannien
 Tel. +44 121 414 7211

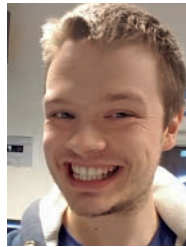
E-Mail: u.noppeney@bham.ac.uk

Uta Noppeney ist Professorin für Computational Neuroscience und Direktorin des Zentrums für Computational Neuroscience and Cognitive Robotics an der Universität Birmingham, Großbritannien. Sie studierte Medizin (1997, Universität Freiburg), promovierte in Medizin (1998, Universität Freiburg) und Neurowissenschaften (2004, University College London, UK). Nach einer Ausbildung in Neurologie am Universitätsklinikum Aachen forschte sie neurowissenschaftlich an der Universität Magdeburg und anschließend am Wellcome Trust Center for Neuroimaging, London. 2005 wurde sie Forschungsgruppenleiterin am Max-Planck-Institut für biologische Kybernetik in Tübingen. Sie kombiniert Psychophysik, funktionelle Bildgebung (M / EEG, fMRI, TMS) und Computermodellierung, um zu untersuchen, wie das menschliche Gehirn Informationen über die Sinne zu einer kohärenten Wahrnehmung der Umgebung integriert.



Tim Rohe
 Department of Psychiatry and
 Psychotherapy
 University of Tübingen
 Germany
E-Mail: Tim.Rohe@med.uni-tuebingen.de

Tim Rohe ist Postdoktorand an der Klinik für Psychiatrie und Psychotherapie des Universitätsklinikums Tübingen. Er promovierte zwischen 2010 und 2014 am Max-Planck-Institut für biologische Kybernetik in Tübingen. Zuvor studierte er Psychologie an der Universität Freiburg (2004–2010).



Samuel A. Jones
 Computational Neuroscience and Cognitive
 Robotics Centre
 School of Psychology
 University of Birmingham
 Edgbaston
 Birmingham
 B15 2TT
 Großbritannien

E-Mail: saj409@student.bham.ac.uk

Samuel A. Jones erhielt einen BSc in Psychologie und einen MSc in psychologischer Forschung von der Bangor University, Wales. Er ist jetzt im letzten Jahr seines PhD an der Universität von Birmingham. Er nutzt fMRT, Psychophysik und Computermodellierung, um den Einfluss von Gehirnalterung auf die multisensorische Integration zu untersuchen.



Ambra Ferrari
 Computational Neuroscience and Cognitive
 Robotics Centre
 School of Psychology
 University of Birmingham
 Edgbaston
 Birmingham
 B15 2TT
 Großbritannien

E-Mail: axf548@student.bham.ac.uk

Ambra Ferrari studierte Kognitive Neurowissenschaften und Neuropsychologie an der Universität von Trient, Italien. Sie promoviert derzeit an der University of Birmingham in Neurowissenschaften. Sie kombiniert Psychophysik, Computermodellierung und MRT-Techniken, um die Beziehung zwischen multisensorischer Integration und höheren kognitiven Prozessen wie Aufmerksamkeit und Belohnungslernen zu untersuchen.