

Aus der Klinik für Psychiatrie und Psychotherapie
(*Direktor: Prof. Dr. med. Tilo Kircher*)
des Fachbereichs Medizin der Philipps-Universität Marburg

Die Bedeutung von experimentellem Paradigma und interindividuellen Unterschieden für die Interpretation von funktionellen MRT-Ergebnissen

Inaugural-Dissertation
zur Erlangung des Doktorgrades der gesamten Humanmedizin,
dem Fachbereich Medizin der Philipps-Universität Marburg
vorgelegt von

Kristin Marie Zimmermann aus **Wolfenbüttel**

Marburg, 2020

Angenommen vom Fachbereich Medizin der Philipps-Universität Marburg am:

23.10.2020.

Gedruckt mit Genehmigung des Fachbereichs

Dekan i.V. der Prodekan: Prof. Dr. R. Müller

Referent: Herr Prof. Dr. A. Jansen

1. Korreferentin: Frau Prof. Dr. I. Kamp-Becker

Inhaltsverzeichnis

<u>1 Einleitung.....</u>	<u>2</u>
<u>2 Aspekte der geschichtlichen Entwicklung neurowissenschaftlicher Methoden....</u>	<u>5</u>
<u>3 Einführung in die fMRT.....</u>	<u>9</u>
<u>4 Ergebnisbeeinflussende Faktoren bei fMRT-Untersuchungen.....</u>	<u>11</u>
<u>4.1 Der Einfluss des experimentellen Paradigmas.....</u>	<u>12</u>
<u>4.2 Der Einfluss interindividueller Unterschiede.....</u>	<u>13</u>
<u>5 Originalarbeiten.....</u>	<u>16</u>
<u>5.1 Studie 1. Comparison of fMRI paradigms assessing visuospatial processing: Robustness and reproducibility.....</u>	<u>16</u>
<u>5.1.1 Bibliographische Details.....</u>	<u>16</u>
<u>5.1.2 Zusammenfassung Studie 1.....</u>	<u>17</u>
<u>5.2 Studie 2. Illusory face detection in pure noise images: The role of interindividual variability in fMRI activation patterns.....</u>	<u>23</u>
<u>5.2.1 Bibliographische Details.....</u>	<u>23</u>
<u>5.2.2 Zusammenfassung Studie 2.....</u>	<u>23</u>
<u>5.3 Studie 3. Seeing things differently: Gaze shapes neural signal during mentalizing in alexithymia.....</u>	<u>28</u>
<u>5.3.1 Bibliographische Details.....</u>	<u>28</u>
<u>5.3.2 Zusammenfassung Studie 3.....</u>	<u>28</u>
<u>6 Diskussion und Ausblick.....</u>	<u>35</u>
<u>7 Zusammenfassung.....</u>	<u>41</u>
<u>8 Referenzen.....</u>	<u>43</u>

„In den Wissenschaften ist viel Gewisses, sobald man sich von den Ausnahmen nicht irre machen läßt und die Probleme zu ehren weiß.“

Johann Wolfgang von Goethe (1749 – 1832)

„Wilhelm Meisters Wanderjahre“

1 Einleitung

Die funktionelle Magnetresonanztomographie (fMRT) ist eine – nach bisherigen Erkenntnissen – nicht-invasive Methode, mit der stimulations- und aufgabenbezogene neurale Prozesse auch in tiefen Strukturen des menschlichen Gehirns abgebildet werden können. Die Qualität des in der fMRT genutzten Signals wird von vielen Faktoren beeinflusst, wie z.B. von physikalischen Messparametern, der Art der Datenvorverarbeitung, aber auch vom experimentellen Paradigma und interindividuellen Unterschieden. Trotz eines relativ geringen Signal-zu-Rausch-Verhältnisses von nur wenigen Prozent (Parrish et al., 2000) ist die fMRT eine der am häufigsten eingesetzten Methoden, um definierte Prozesse im menschlichen Gehirn zu untersuchen (Logothetis, 2008; Poldrack et al., 2017).

Die Zahl der fMRT-Studien wächst stetig und mit ihr auch die Palette an verwendeten experimentellen Aufgaben und statistischen Methoden. Nicht zuletzt aufgrund des geringen Signalanteils, im Vergleich zum Rauschen, führen diese Alternativen allerdings allzu häufig zu verschiedenen Ergebnissen. So steckt die Neurowissenschaft, und mit ihr die fMRT, in diesen Zeiten inmitten einer Replikationskrise (Barch and Yarkoni, 2013; Shrout and Rodgers, 2018). Ein aktuelles Beispiel aus der Neuropsychiatrie verdeutlicht die Schwere dieser Krise: im Rahmen einer Meta-Analyse fassten Müller et al. (2017) fMRT-Studien zusammen, die emotionsbezogene und kognitive Prozesse bei Depression untersuchten. Die Autoren fanden für keinen der untersuchten Prozesse signifikante Überschneidungen zwischen den Studien. Angesichts der mangelnden Konvergenz verschiedener Untersuchungen der vermeintlich gleichen neuronalen Prozesse stellt sich die Frage nach notwendigen Schritten, um eine Reform der Neurowissenschaften voranzutreiben (Barch and Yarkoni, 2013).

Verschiedene Probleme werden diskutiert. Einerseits wird allgemeine Kritik an den formalen Strukturen des Wissenschaftssystems laut. Diese System-

kritik lässt sich nicht nur auf die Neurowissenschaften anwenden, sondern betrifft auch weitere Wissenschaftsdisziplinen, welche sich ebenfalls in der Krise befinden. Kritisiert werden unter anderem fehlgeleitete Anreize, Interessenskonflikte und fragliche wissenschaftliche Praktiken, wie beispielsweise das sogenannte "p-hacking" (Bakker and Wicherts, 2011; Ioannidis, 2011; Simmons et al., 2011). Ein aktueller Ansatz, durch den Teile des Wissenschaftssystems reformiert werden sollen, ist die sogenannte "Open Science", die eine verstärkte Transparenz wissenschaftlicher Arbeitsweisen und Kommunikation einfordert (Alexander et al., 2012; Open Science Collobaration, 2015).

Ein anderes Problem im Rahmen der Replikationskrise bezieht sich direkt auf die Umsetzung wissenschaftlicher Fragestellungen in Experimente. Die Eignung experimenteller Aufgaben zur Untersuchung bestimmter Hypothesen wird nur in seltenen Fällen hinterfragt. Häufig werden Aufgaben ausgewählt, die augenscheinlich zur Fragestellung passen oder sich bereits in dem jeweiligen Forschungsbereich etabliert haben (Barch and Yarkoni, 2013; Logothetis, 2008). Eine Begutachtung der Qualität anhand formaler Kriterien findet in diesem Zusammenhang vergleichsweise selten statt, so dass hier ein dringender Optimierungsbedarf besteht. Um wissenschaftliche Designs zu verbessern, sollten vermehrt Eignungsprüfungen experimenteller Paradigmen und statistischer Analyseverfahren erfolgen. Für derartige Prüfungen stehen etablierte wissenschaftliche Gütekriterien, wie z.B. Reliabilität und Validität, zur Verfügung (vgl. Kapitel 4). In einer Spezialausgabe zum Thema *Reliabilität und Replikation in den kognitiven und affektiven Neurowissenschaften* fordern Barch und Yarkoni (2013) beispielsweise, "... that researchers should be examining the reliability of their particular task and analysis choices on a much more regular basis." Reliabilität, d.h. die Zuverlässigkeit der Ergebnisse, wird durch verschiedene Faktoren, wie etwa die statistische Schwelle oder das experimentelle Paradigma inklusive der gewählten Kontrollaufgabe und Kontraste, beeinflusst (Bennett and Miller, 2013). Es existieren verschiedene Maße zur Quantifizierung von Reliabilität, was sowohl Erhebung als auch Interpretation dieses Gütekriteriums erleichtert. Dass in der wissenschaftlichen Praxis dennoch vergleichsweise selten die Reliabilität experimenteller Aufgaben untersucht wird, hängt wiederum auch mit einem Mangel an Anreizen zusammen: viele renommierte Zeitschriften ziehen es vor, vermeintlich neue Fragestellungen statt Replikationsstudien zu berichten (Brembs et al., 2013; Open

Science Collobaration, 2015). So ist Studie 1 dieser kumulativen Dissertation, obwohl viele experimentelle Paradigmen zur Untersuchung räumlicher Prozesse existieren, die erste Studie, die einige dieser Paradigmen mittels Gütekriterien wie der Reliabilität hinsichtlich ihres Potenzials prüft, die Rechtslateralisation räumlicher Prozesse darzustellen (s. Kapitel 5.1).

Ein weiteres Problem, das im Rahmen der Replikationskrise diskutiert wird, ist die mangelnde *Power* durch zu kleine Stichproben (Button et al., 2013). Der Gedanke dahinter ist, dass aufgrund starker individueller Abweichungen vom Gruppenmittel wichtige funktionelle Areale nicht detektiert werden. Gruppenanalysen, deren Ergebnisse vor allem die Schnittmenge individueller Prozesse widerspiegeln, sind in der fMRT die Regel. Variabilität zwischen Probanden wird als Fehler klassifiziert. Dementsprechend wird im Rahmen der Replikationskrise zwar die Notwendigkeit größerer Stichproben zur Erhöhung der *Power* thematisiert, nicht jedoch eine verstärkte Untersuchung interindividueller Unterschiede. Dabei bilden Unterschiede zwischen Probanden oftmals bedeutsame Varianten ab, z.B. bezüglich der Lokalisation bestimmter Funktionen oder der Anwendung verschiedener Strategien (Lebreton et al., 2019; Seghier and Price, 2018; Zilles and Amunts, 2013). Studie 2 dieser Dissertation stellt als erste Untersuchung zum Thema *Face Pareidolia*, dem Eindruck ein Gesicht in einem Muster oder einer Struktur zu erkennen, dar, dass bedeutende interindividuelle Unterschiede zwischen den neuronalen Prozessen verschiedener Individuen vorliegen (Kapitel 5.2). In Studie 3 verwenden Kollegen und ich die bislang selten eingesetzte Methodenkombination aus fMRT und *Eyetracking*, um für verschiedene visuelle Aufmerksamkeitsfoki zwischen Probanden zu kontrollieren (Kapitel 5.3). Die Berücksichtigung des individuellen Blickverhaltens ermöglicht uns dabei, probandenspezifische Unterschiede in differenzierten Regionen des *Theory of Mind* Netzwerks abzubilden.

Zusammengefasst ist festzustellen, dass das fMRT-Signal nicht nur abhängig von der Einstellung physikalischer Messparameter oder der Präzision erforderlicher Vorverarbeitungsschritte ist. Auch die Eignung und Umsetzung der experimentellen Aufgabe, die Art der statistischen Auswertung sowie Unterschiede zwischen Personen spielen eine Rolle. In drei Studien untersuchten Kollegen

und ich den Einfluss verschiedener nicht-physikalischer Parameter auf das resultierende fMRT-Signal. In der ersten Studie verglichen wir Reliabilität und Robustheit spezifischer experimenteller Paradigmen, um die für sich anschließende Fragestellung vergleichsweise geeignetste Aufgabe zu identifizieren (Studie 1). Die beiden anderen Studien behandeln den Einfluss individueller Unterschiede auf das fMRT-Signal. Wir zeigen, dass zwischen Probanden trotz gleicher visueller Stimulation eine komplett unterschiedliche Rekrutierung früher visueller Areale erfolgen kann (Studie 2). In der letzten und aufwendigsten Untersuchung kombinieren wir die Methoden fMRT und Eyetracking (Studie 3). Auf diese Weise ermitteln wir die Effekte von individuellem Blickverhalten auf bestimmte neurale Prozesse.

In der hier vorliegenden kumulativen Dissertation gebe ich zunächst einen kurzen Einblick in bestimmte geschichtliche Aspekte der Neurowissenschaft, aus denen sich die Entwicklung der fMRT ergab (Kapitel 2). Anschließend führe ich in die Methode der fMRT ein (Kapitel 3) und verdeutliche spezifische, in den Studien dieser Dissertation behandelte, Einflussfaktoren auf das fMRT-Signal (Kapitel 4). Schließlich stelle ich die drei Studien zusammenfassend vor (Kapitel 5) und diskutiere ihre Ergebnisse im Hinblick auf die Thematik dieser Dissertation: die Bedeutung des experimentellen Paradigmas und individueller Unterschiede für die Interpretation von fMRT-Ergebnissen (Kapitel 6).

2 Aspekte der geschichtlichen Entwicklung neurowissenschaftlicher Methoden

Das menschliche Gehirn besteht aus circa 100 Milliarden Nervenzellen, die durch etwa 100 Billionen Schnittstellen miteinander verbunden sind. Dieses hochkomplexe physiologische System schafft im Menschen die Grundlage dafür, verschiedenste, zum Teil artspezifische, Funktionen ausbilden zu können (Tomasello et al., 2005). Ein Zusammenspiel aus Afferenzen (sensorische Eingänge) und Efferenzen (Signalausgänge an gezielte Effektoren) ermöglicht dabei Interaktionen zwischen Peripherie und zentralem Steuerorgan und bildet die Basis für Kommunikation zwischen Individuum und Umwelt. Ferner erlaubt der komplizierte Netzwerkcharakter dem Organ eine dynamische Anpassung an ebenso dynamische Externalitäten. Neben autonom ablaufenden Prozessen, wie beispielsweise der Regelung en-

dokriner, z.B. hypothalamisch-hypophysärer, Schaltkreise, stellen Wahrnehmung, Erwartung sowie Initiation bestimmter Verhaltensweisen besonders wichtige Funktionen des menschlichen Nervensystems dar.

Diese faszinierend komplexe Natur des menschlichen Gehirns motivierte Wissenschaftler bereits in frühen Zeiten dazu, sich der Erforschung zugrundeliegender Mechanismen und Funktionen des Organs zu widmen (Oeser, 2002). Mediziner und Philosophen der Renaissance und Post-Renaissance entwickelten eher allgemeine, grob gefasste Theorien. Im mittleren 17. Jahrhundert vermutete beispielsweise der Philosoph René Descartes (1596 - 1650), dass unsere Gedanken in der Epiphyse entstehen. Andere seiner Zeit machten das Corpus Callosum, welches über axonale Fasern die beiden Hirnhälften miteinander verbindet, für unsere Kognitionen verantwortlich. Mitte des 18. Jahrhunderts spekulierte der Naturforscher, Ingenieur und Theosoph Emanuel Swedenborg (1688 - 1772) bereits, dass sich die menschliche Seele in der Großhirnrinde befinden würde. Außerdem vermutete er, dass die verschiedenen Hirnfunktionen in unterschiedlichen Anteilen des Kortex lokalisiert sind. So nahm er beispielsweise an, dass die frontale Hirnrinde Vorstellungen und Willen formen würde. Seine Hypothesen formulierte er auf der Grundlage von Sektionen sowie Studien der menschlichen Physiologie. Die Präzision seiner Vermutungen waren Swedenborgs Zeit voraus und so fanden seine Theorien keine weitere Beachtung seitens seiner damaligen Kollegen. (Vgl. Huettel et al., 2004)

Zu Beginn des 19. Jahrhunderts wuchs ein allgemeines Interesse an der Beziehung zwischen individueller Physiognomie und Charakter. So hypothetisierten sogenannte Phrenologen (altgriechisch *phrenos* „Geist, Seele“, und *logos* „Lehre“), dass die Ausprägung eines bestimmten Verhaltens oder bestimmter Eigenschaften mit einem lokal vergrößerten Hirnvolumen einhergehen müsste. Dies setzte voraus, dass Hirnfunktionen an umschriebenen Orten angesiedelt sind und die Masse der Substanz an diesen Orten mit der Ausprägung der Funktion zunimmt. Bekannte Vertreter unter den Phrenologen waren der Arzt Franz Joseph Gall (1758 - 1828) sowie sein Mitarbeiter Johann Spurzheim (1776 - 1832). Da eine Messung des lokalen Hirnvolumens zu den Zeiten Galls und Spurzheims nicht möglich war, behelfen sie sich mit einer Begutachtung der Schädelphysiognomie. Sie beurteilten individuelle Schädelformen inklusive Dellen und Ausbuchtungen und beobachteten gleichzeitig die Verhaltensweisen ihrer Studienob-

jekte. Laut Gall saß das Intelligenzorgan beispielsweise hinter den Ohren: „Wer in Verlegenheit ist, kratzt sich gewöhnlich hinter den Ohren, um etwa das Organ der Schlaueheit oder jenes des Muthes in Thätigkeit zu setzen.“ (zitiert nach Bloede, 1807). Obwohl das Konzept der Phrenologie bereits in seinen Anfängen seitens der wissenschaftlichen Gemeinschaft angezweifelt wurde (Ackermann, 1806), hatte es in Laienkreisen noch viele Jahre Konjunktur. So behandelte beispielsweise auch die physische Anthropologie unterschiedliche Schädelformen und bereitete auf diese Art eine Grundlage für die wenig später aufkommende Rassenbiologie (Schott, 2002).

Durch zahlreiche Sektionsstudien war Ende des 18. Jahrhunderts das Wissen über makroanatomische Verhältnisse des zentralen Nervensystems bereits fortgeschritten (Bear et al., 2018). Anhand von Läsionsstudien wurden anatomische Topologien mit menschlichen Funktionen verknüpft (Huettel et al., 2004; Peiffer, 2004). So beschrieb Paul Broca (1824 – 1880) beispielsweise im Jahr 1861, dass eine Schädigung im linken Frontallappen mit der Unfähigkeit zur Sprachproduktion bei erhaltenem Sprachverständnis in Verbindung stehen würde (Broca, 1861). In der zweiten Hälfte des 19. Jahrhunderts gewannen Untersuchungen physiologischer Vorgänge des Gehirns vermehrt an Bedeutung.

Mittlerweile gibt es zahlreiche Bildgebungsmethoden (vgl. Abb. 1), von denen einige, wie auch die fMRT, eine – nach heutigem Erkenntnisstand – nicht-invasive Untersuchung physiologischer Zustandsveränderungen unseres Gehirns ermöglichen (Gaebler et al., 2016). Der besondere Vorteil der fMRT besteht darin, dass sie auch in tiefer gelegenen Hirnstrukturen physiologische Prozesse darstellen kann. Durch gezielte Stimulationen und Manipulationen können diese physiologischen Prozesse mit Phänomenen wie Verhalten, Sinneseindrücken oder Kognitionen verknüpft werden.

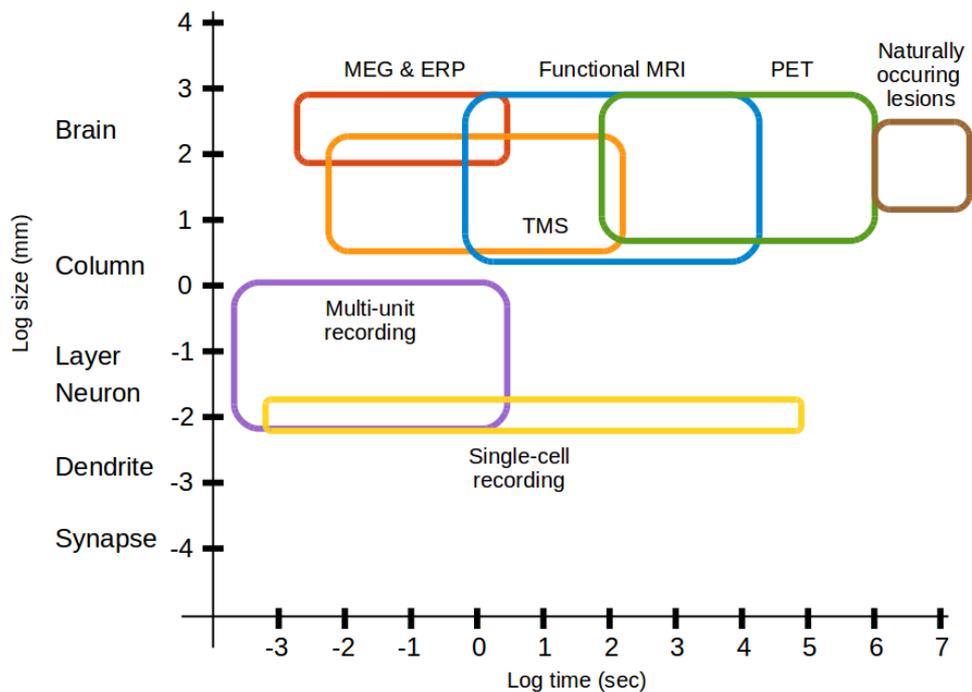


Abbildung 1 (modifiziert nach Grinvald und Hildesheim, 2004, und Sejnowski et al., 2014). Zeitliche und räumliche Auflösung verschiedener Methoden zur Untersuchung funktioneller Prozesse im Gehirn.

Die Entwicklung der MRT basiert auf der Erforschung physikalischer Grundlagen zu magnetischen und elektrischen Eigenschaften von Atomen aus der ersten Hälfte des 20. Jahrhunderts. 1952 bekamen Felix Bloch und Edward Purcell für ihre unabhängigen Untersuchungen zu magnetischer Resonanz von Masse den Nobelpreis für Physik. Die Nutzung ihrer Erkenntnisse beschränkte sich allerdings zunächst auf chemische und geologische Bereiche. Den Grundstein für erweiterte Nutzungsmöglichkeiten der MR-Technik, zum Beispiel im Rahmen biologischer Fragestellungen und schließlich für bildgebende Studien des menschlichen Gehirns, legten allen voran Raymond Damadian, Paul Lauterbur und Peter Mansfield, wobei allerdings lediglich die beiden letzteren 2003 dafür den Nobelpreis für Physiologie und Medizin erhielten. (Vgl. Huettel et al., 2004)

Der Genehmigung eines klinischen Einsatzes der MRT seitens der U.S. Food and Drug Administration (FDA) im Jahr 1985 folgte die Etablierung geeigneter Geräte an internationalen Krankenhäusern (Huettel et al., 2004). Fünf Jahre später wurden die ersten erfolgreichen funktionellen MRT-Experimente durchgeführt, so dass die Methode ab Beginn der 90er Jahre zunehmend auch für For-

schungszwecke eingesetzt wurde (Bandettini, 2012). Seitdem wächst die Anzahl an fMRT-Studien stetig. Dabei findet die fMRT Einsatz in verschiedensten Wissenschaftsdisziplinen, wie zum Beispiel Psychologie, Medizin, Ökonomie, experimentelle Physik oder Sprachwissenschaften. So etablierte sie sich in den vergangenen 20 Jahren als eines der gefragtesten Verfahren moderner Hirnbildgebung.

Trotz dieses rasanten Aufstiegs und einer kontinuierlichen technischen und methodischen Weiterentwicklung gibt es eine Reihe von Faktoren, die die Ergebnisse von fMRT-Untersuchungen verzerren. Nachdem im folgenden Kapitel zunächst eine kurze Einführung in die Funktionsweise der fMRT gegeben wird, werden in Kapitel 4 die Einflussfaktoren auf Ergebnisse von fMRT-Studien näher erläutert, welche in den drei Studien dieser Dissertation behandelt wurden (Kapitel 5).

3 Einführung in die fMRT

Die fMRT basiert auf der Detektion von Signalen, welche durch die elektromagnetischen Eigenschaften von Atomen entstehen. Legt man einen Körper in einen MR-Scanner, so richten sich die Eigenvektoren der Wasserstoffatome dieses Körpers in Magnetfeldrichtung aus. Wird anschließend ein Hochfrequenzpuls mit für die Atomeigenschaften spezifischer Frequenz („*Lamorfrequenz*“) appliziert, führt die auf das Atom übertragene Energie zu einer Richtungsänderung des Eigenvektors im Raum. Das Atom befindet sich nun in einem energiereichen Zustand. Nach kurzer Zeit klappen die Eigenvektoren der Atome wieder in ihre Ursprungsrichtung zurück und strahlen dabei Energie ab. Diese Energie beschreibt einen Teil des mittels MR gemessenen Signals. Ein weiterer Teil wird aufgrund von Wechselwirkungen zwischen den einzelnen Atomen emittiert. Nach dem Einstrahlen des Hochfrequenzpulses bewegen sich die Atome zunächst alle in Phase. Diese Einheit löst sich kurze Zeit später aufgrund heterogener Atomdynamiken auf, so dass es zu Phasenunterschieden und schließlich zu Atomkollisionen kommt. Bei diesen Kollisionen wird ebenfalls Energie nach außen abgegeben.

Mittels spezieller Messprotokolle werden Sequenzen aus anregenden Hochfrequenzpulsen mit jeweils gekoppelten Signalauslesephasen erstellt. Diese Sequenzen markieren die zeitliche Auflösung der Methode, welche durch Atomei-

genschaften, wie Relaxationszeiten, definiert wird. Die räumliche Auflösung wird durch den Einsatz zusätzlicher, durch Gradientenspulen erzeugter, Magnetfelder hergestellt. Durch ihre Schaltung wird auf eine kontrollierte Art Einfluss auf die Verteilung von Frequenz- bzw. Signalvarianten im zweidimensionalen Raum genommen, so dass ein räumlich enkodiertes Muster an Signalvarianten entsteht. Anhand der Kombination dieser sequentiell aufgenommenen, zweidimensionalen Schichtaufnahmen können dreidimensionale Signalkarten des gesamten Gehirns rekonstruiert werden.

Um eine Verknüpfung dieser Signalkarten mit Verarbeitungsprozessen im Gehirn herzustellen, werden die magnetischen Eigenschaften des eisenhaltigen Blutbestandteils Hämoglobin ausgenutzt. Hämoglobin übt unterschiedlichen Einfluss auf das umliegende Magnetfeld aus, je nachdem ob es in oxygenierter oder desoxygenierter Form vorliegt. So führen Änderungen der Oxygenierung im (kapillar-)venösen System (über Modulation der Magnetisierung des Hämoglobins) zu detektierbaren Magnetfeldänderungen. Im Gehirn ändert sich die Sauerstoffsättigung des Blutes durch den aeroben Energiestoffwechsel der Neurone. Die Aktivität der an einem zentralvenösen Prozess beteiligten Neuronenpopulationen führt so zu einem mittels MR-Technik lokalisierbaren Phänomen.

Typischerweise wird anhand dieser Zusammenhänge die neurale Reaktion auf bestimmte Reize und Aufgaben erfasst. Das bedeutet, dass die Signalkarten, die in der fMRT als „Aktivierungs-“karten bezeichnet werden, nicht direkt die Aktivierung von Nervenzellen widerspiegeln. Sie sind indirekte Darstellungen neuraler Aktivität, die anhand von Annahmen über Energiestoffwechsel, Blutströmungsdynamiken sowie vaskuläre Eigenschaften erstellt werden (*hämodynamische Antwortfunktion*). Das aus diesen Prozessen resultierende Signal wird als *BOLD-Antwort* (blood oxygenation level dependent) bezeichnet.

Die Herkunft der BOLD-Antwort verdeutlicht die Bedeutung physikalischer und physiologischer Faktoren für Qualität sowie Quantität des fMRT-Signals. Neben diesen Faktoren nehmen weitere, sowohl vom Versuchsleiter als auch vom Probanden abhängige, Variablen Einfluss auf das fMRT-Signal. Dazu gehören das experimentelle Paradigma inklusive Aufgabendesign und Stimulusmaterial sowie interindividuelle Unterschiede zwischen Individuen, wie zum Beispiel unterschiedliche Verarbeitungsstrategien oder unterschiedliches Verhalten.

Die folgenden Abschnitte beschäftigen sich mit den in den drei Studien dieser Dissertation behandelten Faktoren, die das fMRT-Signal und somit Ergebnisdarstellungen wissenschaftlicher Studien jenseits der beabsichtigten Effekte beeinflussen. (Vgl. z.B. Huettel et al., 2004)

4 Ergebnisbeeinflussende Faktoren bei fMRT-Untersuchungen

In der fMRT wird zunächst ein Rohsignal erfasst. Dieses Rohsignal wird durch zahlreiche Faktoren verzerrt, so dass ein Signal-zu-Rausch-Verhältnis von nur wenigen Prozent resultiert (Parrish et al., 2000). Das bedeutet, dass die Varianz des Rohsignals nur zu einem geringen Teil auf experimentelle Manipulationen zurückzuführen ist.

Die Qualität des Signals kann durch verschiedene Faktoren verbessert werden. Beispielsweise können durch die Wahl geeigneter physikalischer Messsequenzen Rauschanteile des Signals verringert werden. Auch die Art der Datenvorverarbeitung, bei der unter anderem für die Bewegungen des Probanden oder den zeitlichen Versatz der einzelnen Schichtaufnahmen korrigiert werden kann, ist von Bedeutung.

Neben diesen *physikalisch-technischen* Einflussfaktoren spielt die Übersetzung der theoretischen Fragestellung in ein wissenschaftliches Experiment eine entscheidende Rolle. Aus dem geringen Signal-zu-Rausch-Verhältnis im Rahmen der fMRT ergibt sich die Bedeutung einer umfangreichen Kontrolle der Parameter, die durch den Versuchsleiter beeinflusst werden können: je mehr Parameter kontrolliert werden, desto besser kann das erwartete Signal vorhergesagt und von Fluktuationen durch unkontrollierte (Stör-)faktoren abgegrenzt werden. Besonders für die *funktionelle* MRT, bei der die neuronalen Veränderungen erfasst werden sollen, die durch experimentelle Manipulationen hervorgerufen werden, ist die Eignung der experimentellen Aufgabe von zentraler Bedeutung. Über die Eignung eines Experiments für die Prüfung spezifischer Hypothesen geben zentrale Gütekriterien wissenschaftlicher Forschung Auskunft, darunter *Objektivität*, *Reliabilität* und *Validität* (Himme, 2009). Objektivität bezieht sich auf eine größtmögliche Unabhängigkeit der Studienergebnisse von den sogenannten Randbedingungen, wie beispielsweise den Umwelteinflüssen (Räumlichkeiten, Temperatur etc.) oder

dem Versuchsleiter (durch soziale Interaktionen, Meinungsbildung etc.). Reliabilität bezeichnet die Zuverlässigkeit der Ergebnisse, wie zum Beispiel, inwieweit die Wiederholung eines Experiments unter gleichen Messbedingungen zu den gleichen Ergebnissen führt (*Test-Retest-Reliabilität*). Validität beschreibt, ob durch das Experiment tatsächlich das gemessen wird, was gemessen werden soll. Hohe *interne Validität* eines Experiments besteht zum Beispiel dann, wenn die Varianz der abhängigen Variable durch Veränderungen der unabhängigen Variable bedingt ist, ohne dass eine alternative Ursache dafür in Frage kommt.

In den drei Studien dieser kumulativen Dissertation werden *nicht-physikalische* Einflussfaktoren auf des fMRT-Signal behandelt, die mit verschiedenen Aspekten der wissenschaftlichen Gütekriterien in Zusammenhang stehen. In *Studie 1* vergleichen wir verschiedene experimentelle Paradigmen, die in der Literatur für die Darstellung lateralisierter visuell-räumlicher Verarbeitungsprozesse genutzt werden. Das daraus hervorgehende geeignetste Paradigma zur Darstellung der Lateralisierung prüfen wir unter anderem hinsichtlich seiner Test-Retest-Reliabilität und leiten daraus Empfehlungen für die Untersuchung visuell-räumlicher Prozesse mittels fMRT ab. In *Studie 2* hinterfragen wir die interne Validität der experimentellen Aufgaben, die in der fMRT für die Darstellung des Pareidolienetzwerks genutzt werden. Dabei werden Schwierigkeiten bei der Ergebnisinterpretation verdeutlicht, die auftreten können, wenn das experimentelle Design individuell unterschiedliche Aufgabenausführungen zulässt. In *Studie 3* diskutieren wir die interne Validität von fMRT-Untersuchungen, in denen komplexe visuelle Stimuli genutzt werden, ohne dass für interindividuelle Unterschiede im Blickverhalten und deren Auswirkungen auf neurale Prozesse kontrolliert wird.

4.1 Der Einfluss des experimentellen Paradigmas

In vielen Fällen besteht das Ziel einer fMRT-Untersuchung darin, bestimmte funktionelle Strukturen, ein bestimmtes Netzwerk, oder andere funktionelle Eigenschaften wie die Lateralisation eines neuronalen Prozesses (Studie 1 dieser Dissertation) abzubilden. Dabei ist eine möglichst robuste Ergebnisdarstellung anzustreben. Aufgrund der zahlreichen Faktoren, die das fMRT-Signal beeinflussen, variieren die Ergebnisse verschiedener Untersuchungen zur Darstellung derselben funktionellen Strukturen oder Eigenschaften jedoch mitunter stark.

Dennoch gelten einige funktionelle Areale und Netzwerke angesichts der fMRT-Forschung der vergangenen 30 Jahre bereits als stabil darstellbar. Ein prominentes Beispiel hierfür ist die Beschreibung eines gesichtsverarbeitenden Moduls im Fusiformen Gyrus, der sogenannten *Fusiform Face Area (FFA)* (Kanwisher et al., 1997). Die FFA ist Teil eines anerkannten Netzwerks zur Verarbeitung von Gesichtern, welches sich in robuster und reliabler Weise im einzelnen Individuum abbilden lässt (Frässle et al., 2016).

Aufgrund der Stabilität, mit der verschiedene Regionen des Gesichternetzwerks im Individuum darstellbar sind, entwickelte sich der Einsatz sogenannter *Face Localizer*. Face Localizer beschreiben experimentelle Paradigmen, die sich besonders dafür eignen, am Gesichternetzwerk beteiligte Regionen zu identifizieren. Mittlerweile werden auch in vielen anderen Bereichen Localizer eingesetzt. Beispielsweise werden Aufgaben mit spezifischen Handbewegungen dafür genutzt, bedeutende motorische Regionen darzustellen (Kristo et al., 2014). Andere Localizer wiederum nutzen Bilder von emotionalen Gesichtern für die funktionelle Kartierung der Amygdala, einem Teil des für die Emotionsverarbeitung relevanten limbischen Systems (Dotterer et al., 2017; Hariri et al., 2002).

Funktionelle Localizer stellen ein Beispiel dafür dar, dass es mittels der Methode fMRT möglich ist, funktionelle Strukturen auf stabile Art und Weise zu detektieren und abzubilden. Genau wie funktionelle Localizer sollten andere in der fMRT verwendete experimentelle Paradigmen gewissen Anforderungen hinsichtlich der wissenschaftlichen Gütekriterien entsprechen (vgl. Kapitel 4). In Studie 1 dieser Dissertation verfolgten wir das Ziel, ein geeignetes Paradigma zur Darstellung der anerkannten Rechtsdominanz visuell-räumlicher Prozesse zu identifizieren. Dazu verglichen wir verschiedene, gängige Aufgaben und erhoben qualitative Maße wie *Robustheit*, *Test-Retest-Reliabilität* und *Konsistenz*. In Kapitel 5.1 werden die Ergebnisse zusammengefasst und im Hinblick auf ihre Implikationen für die Beantwortung wissenschaftlicher Fragestellungen diskutiert.

4.2 Der Einfluss interindividueller Unterschiede

In fMRT-Experimenten wird in der Regel auf Gruppenstatistiken zurückgegriffen, um Aussagen über Hirnfunktionen zu treffen. Treten unerwartet große oder breit gestreute Unterschiede zwischen Individuen auf, können diese die Ergebnisse von

fMRT-Experimenten dementsprechend beeinflussen. Interindividuelle Unterschiede finden sich in verschiedenen Bereichen, die für die fMRT eine Rolle spielen, wie beispielsweise in der Hirnanatomie oder bezüglich BOLD-Amplitude und zerebrovaskulärer Reaktivität (Ainslie and Duffin, 2009; Schuster et al., 2017).

Neben diesen anatomisch-physiologisch bedingten Faktoren, gibt es auch Unterschiede, die mit spezifischen funktionellen Prozessen in Verbindung stehen. In diesem Zusammenhang unterscheiden Seghier und Price (2018) *intrinsische*, *strategische* und *kontextuelle* Unterschiede. Intrinsische Unterschiede sind personeninhärent und zeitlich invariant. So ist beispielsweise die in der Regel rechtslateralisierte visuell-räumliche Verarbeitung in einigen Personen linkslateralisiert (vgl. Kapitel 5.1). Demgegenüber können strategische Unterschiede über die Zeit variieren. Sie sind in der Regel erlernt und beziehen sich auf bestimmte kognitive Muster, wie zum Beispiel verschiedene Techniken beim Kopfrechnen. Kontextuelle Unterschiede beschreiben wiederum veränderliche Merkmale wie Motivation, Habituation, Stress, oder auch die Vertrautheit mit MRT-Messungen. Dabei unterliegen kontextuelle Unterschiede schnelleren Veränderungen als strategische.

Im Rahmen der in der fMRT üblichen Gruppenstatistiken werden interindividuelle Unterschiede, die keine explizite Beachtung in den statistischen Modellen finden, als Rauschen klassifiziert. Um Unterschiede zwischen Probanden zu berücksichtigen können sie entweder kontrolliert oder ins statistische Modell integriert werden. Kontrolliert werden sie z.B. in Form von statistischem Matching oder durch standardisierte Instruktionen. Integriert werden können sie durch die Aufnahme individueller Ausprägungen als unabhängige Variable ins statistische Modell. Die Voraussetzung für eine Berücksichtigung ist jedoch zunächst, dass die Bedeutung der jeweiligen Unterschiede bekannt ist.

In einigen Untersuchungskontexten ist die Identifikation von Faktoren die mit interindividuellen Aktivierungsmustern in Verbindung stehen könnten einfacher als in anderen. So kann Verhalten in der Regel beobachtet und quantifiziert werden, während kognitive Vorgänge weniger einfach zugänglich sind. Es ist zum Beispiel plausibel, dass neurale Aktivierungen die zeitlich kontingent mit einer Handbewegung, wie etwa einem Faustschluss, auftreten, mit dieser motorischen Handlung zusammenhängen. Im Vergleich dazu würde es weniger leicht fallen, die Zeitpunkte vorherzusagen, zu denen eine Person, die ein Bild betrachtet, ein

Gefühl von Traurigkeit empfindet, selbst dann, wenn es sich bei dem Bild um eine Darstellung eines mehrheitlich als traurig bewerteten Bilds handelt. Insgesamt scheinen phänomenologische, mit den Sinnen erfahrbare, Variablen leichter identifiziert und objektiviert werden zu können, als Gefühle, Kognitionen oder psychologische Konstrukte.

Studie 2 dieser Dissertation behandelt interindividuelle Unterschiede der neuralen Korrelate von *Face Pareidolia*, dem Eindruck, Gesichter in Strukturen oder Formen zu erkennen (Kapitel 5.2). Auch Face Pareidolia stellt kein beobachtbares Phänomen dar. Aus diesem Grund ist es schwer festzustellen, wann bei einer Person tatsächlich Face Pareidolia auftritt. Die Ergebnisse aus Studie 2 dieser Dissertation weisen darauf hin, dass Face Pareidolia im Rahmen konventioneller Studien zu dem Thema eventuell weniger häufig auftritt, als bisher vermutet: nur etwa die Hälfte aller Probanden zeigte die zu erwartende Aktivierung in Kernregionen des Gesichternetzwerks (vgl. Kapitel 4.1). Da in anderen Studien zu Face Pareidolia stets Gruppenergebnisse berichtet wurden, war das durch die Analyse der einzelnen Probanden sichtbar gewordene Ausmaß der interindividuellen Unterschiede überraschend. Um diesen Unterschieden auf den Grund zu gehen, sollten zukünftige Studien klären, ob sich die interindividuelle Varianz durch intrinsische, strategische, oder kontextuelle Faktoren ergibt. In Kapitel 5.2 diskutiere ich die Wahrscheinlichkeit dieser Möglichkeiten im Bezug auf Studie 2 und stelle potenzielle phänomenologische Variablen vor, die im Rahmen von Experimenten zu Face Pareidolia eine Rolle bei der Aufklärung zugrundeliegender Mechanismen interindividueller Unterschiede spielen könnten. In Kapitel 6 bespreche ich, wie mit intrinsischen, strategischen und kontextuellen Unterschieden im Allgemeinen umgegangen und zwischen ihnen differenziert werden könnte.

Das Beispiel Face Pareidolia verdeutlicht Schwierigkeiten, die auftreten wenn kognitive Vorgänge untersucht werden sollen, die nicht direkt zu beobachten sind. Demgegenüber sind interindividuelle Verhaltensstrategien aufgrund ihrer phänomenologischen Natur besser zu identifizieren und somit im Rahmen von Experimenten auch leichter zu kontrollieren. Trotz dieser Voraussetzung werden experimentelle Freiräume für unterschiedliches Verhalten innerhalb von fMRT-Studien häufig ignoriert.

Ein verbreitetes Beispiel dafür ist die Verwendung von komplexem visuellem Stimulusmaterial. Komplexe Stimuli können von verschiedenen Probanden auf unterschiedliche Art und Weise inspiziert werden. Diese verschiedenen Pfade visueller Aufmerksamkeit können mit der Rekrutierung unterschiedlicher neuraler Strukturen einhergehen (Hadjikhani et al., 2017b). Um für diese Unterschiede zu kontrollieren, wird in vielen Studien ein bestimmter visueller Fokus vorgegeben, der von allen Teilnehmern übernommen werden soll. Das hat wiederum den Nachteil, dass eine Untersuchung der neuralen Prozesse, die mit natürlichem Verhalten assoziiert sind, auf diese Weise nicht erfolgen kann. In Studie 3 dieser kumulativen Dissertation kombinierten wir Aufnahme und Analyse von fMRT-Signal und Eyetracking, um für Effekte individueller Unterschiede im Blickverhalten zu kontrollieren (Kapitel 5.3).

5 Originalarbeiten

In diesem Kapitel werden die drei Originalarbeiten der vorliegenden kumulativen Dissertation vorgestellt. Studie 1 und Studie 2 wurden nach einem Peer-Review-Verfahren in einer internationalen Fachzeitschrift veröffentlicht. Das Manuskript zu Studie 3 ist fertiggestellt und wurde ebenfalls bei einer internationalen Fachzeitschrift mit Peer-Review-Verfahren eingereicht. Zunächst werden jeweils bibliographische Details beschrieben. Im Anschluss erfolgt eine inhaltliche Zusammenfassung.

5.1 Studie 1. Comparison of fMRI paradigms assessing visuospatial processing: Robustness and reproducibility

5.1.1 Bibliographische Details

Diese Untersuchung wurde durch die Autoren Schuster, V., Herholz, P., Zimmermann, K. M., Westermann, S., Frässle, S. und Jansen, A. durchgeführt und veröffentlicht. Die Studie wurde 2017 bei PLoS One (ISSN 1932-6203, volume 12, issue 10, pages e0186344) unter Open-Access-Bedingungen publiziert. Der Artikel ist unter <https://doi.org/10.1371/journal.pone.0186344> erhältlich.

5.1.2 Zusammenfassung Studie 1

Das menschliche Gehirn besteht aus zwei Hälften. Makroanatomisch wirken die beiden Hirnhälften auf den ersten Blick wie gespiegelte Replikate ihrer selbst. Auch auf funktioneller Ebene scheinen viele Prozesse zunächst auf symmetrische Weise in beiden Hirnhälften repräsentiert zu sein. Bei näherer Betrachtung wird jedoch deutlich, dass bei vielen Funktionen eine Hirnhälfte stärker involviert ist, als die andere. Zum Beispiel treten im Rahmen einer unilateralen Läsion des Broca Areal im linken Frontallappen (vgl. Kapitel 3) typischerweise starke Einschränkungen der Sprachproduktion, bis hin zum Ausfall, auf. Traumata der gespiegelten Region im rechten Frontalhirn führen in der Regel nicht zu derartigen Beeinträchtigungen. Daraus folgt, dass bestimmte neurale Funktionen lateralisiert sind: eine der beiden Hirnhälften nimmt bezüglich der funktionellen Verarbeitung die dominante Rolle ein. Während im Bezug auf die Darstellung der Linksdominanz von Sprache bereits viele experimentelle Paradigmen auf ihre Eignung geprüft wurden, blieben systematische Untersuchungen zur Tauglichkeit der Aufgaben, die die Rechtslateralisation visuell-räumlicher Prozesse abbilden sollen, bislang aus.

In Studie 1 dieser Dissertation wurden verschiedene experimentelle Paradigmen bezüglich ihres Potenzials untersucht, auf robuste und reliable Weise die Rechtslateralisierung visuell-räumlicher Prozesse darzustellen. In einem ersten Experiment erfolgte zunächst ein Vergleich dreier Paradigmen (Abb. 2) im Bezug auf ihre *Robustheit*, die Rechtslateralisierung im einzelnen Probanden abzubilden. Robustheit definierten wir anhand bestimmter Kriterien auf Einzelprobanden- und Gruppenebene (z.B. ein für räumliche Prozesse typisches fronto-parietales Aktivierungsmuster bei einer unkorrigierten statistischen Schwelle von $p < 0.001$ und einer Clustergröße von $k = 20$).

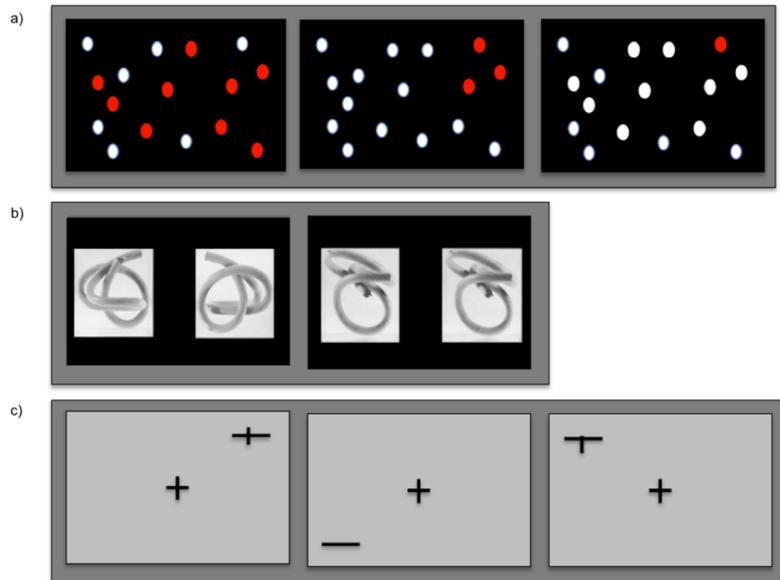


Abbildung 2 (aus Schuster et al., 2017). Drei in der fMRT gängige Paradigmen für die neuronale Darstellung visuell-räumlicher Prozesse. a) Aufgabe "Punkte-im-Raum" (Whitehouse et al., 2009): In der Experimentalbedingung (links und mittig) sollten Probanden angeben, ob sie das aktuelle Muster roter Punkte in der vorangegangenen Enkodierungsphase bereits präsentiert bekamen. In der Kontrollbedingung (rechts) sollte angegeben werden, ob ein roter Punkt zu sehen ist. b) Mentale Rotationsaufgabe (Dorst et al., 2008): In der Experimentalbedingung (links) sollten Probanden angeben, ob die rechte Figur die linke von hinten, unten, oben, rechts oder links darstellt. In der Kontrollbedingung (rechts) sollten die Probanden beurteilen, ob die beiden gezeigten Figuren identisch sind. c) Landmark Aufgabe (Fink et al., 2000): In der Experimentalbedingung (links) sollten Probanden entweder angeben, ob die vertikale Linie die horizontale Linie rechts oder links von deren Mitte kreuzt (Version A), oder ob die horizontale Linie exakt in der Mitte durch die vertikale Linie geteilt wird (Version B und C). In der Kontrollbedingung wurden die Probanden gefragt, ob eine kreuzende vertikale Linie abgebildet ist (mittig, Version A), oder ob die vertikale Linie die horizontale teilt oder nicht (rechts, Version B und C). In Version C waren die Abweichungen der vertikalen Linie von der Mitte der horizontalen Linie kleiner als in Version A und B.

Die *Landmark Aufgabe*, bei der die Probanden die räumliche Position einer vertikalen Linie bewerten müssen (vgl. Abb. 2), erwies sich als robustestes Paradigma. Zwei Versionen (Version B und C, s. Abb. 2) dieser Aufgabe wurden in einem zweiten Experiment auf *Test-Retest-Reliabilität* und *Konsistenz* geprüft. Konsistenz bezeichnet in diesem Zusammenhang, dass die abhängige Variable 'Rechtslateralisierung' in multiplen Messungen eine vordefinierte Mindestausprägung im Individuum erreicht.

Test-Retest-Reliabilität ist ein in der fMRT häufig eingesetztes Gütekriterium zur Überprüfung der Eignung eines experimentellen Paradigmas als Localizer (vgl. Kapitel 4.1). Gewöhnlich basieren Reliabilitätsbestimmungen auf der Quantifizierung des skalierten Messfehlers. Ein verbreitetes Instrument zur Quantifizierung und Skalierung des Messfehlers im Rahmen von Reliabilitätsuntersuchungen ist der *intra-class correlation coefficient (ICC)* (Shrout and Fleiss, 1979). In der fMRT wird dabei in der Regel auf den *ICC (3,1)* aus Shrout und Fleiss (1979) zurückgegriffen, der für zweifaktorielle Designs mit gemischten Effekten ("mixed-effects") formuliert ist¹:

$$ICC(3,1) = \frac{BV - RV}{BV + (k - 1)RV}$$

Der ICC (3,1) beschreibt den Anteil der Varianz innerhalb einer definierten Gruppe ("Between-Subject Varianz", BV, z.B. Varianz zwischen Probanden oder Voxel-Werten) zur Gesamtvarianz (Summe aus BV und Varianz innerhalb eines Probanden oder Voxels, "Within-Subject Varianz" oder residuelle Varianz, RV), wobei *k* für die Anzahl der "Rater", beziehungsweise bei fMRT-Reliabilitätsuntersuchungen typischerweise für die Anzahl der Messwiederholungen, steht. Der ICC (3,1) kann Werte zwischen -1 und 1 annehmen, wobei negative Werte zwar praktisch vorkommen, aber theoretisch nicht definiert sind (Giraudeau, 1996). Der ICC ist groß, wenn der Messfehler klein, oder der Anteil der BV an der Gesamtvarianz groß ist.

Mittels ICC bestimmten wir zunächst die Reliabilität der neuronalen Korrelate der Landmark Aufgabe. Diese Bestimmung erfolgte sowohl auf Voxel-ebene als auch innerhalb für visuell-räumliche Prozesse wichtiger Regions-of-Interest (ROIs) im frontalen und parietalen Kortex. Zur Beurteilung der Reliabilität richteten wir uns nach einer üblichen Einteilung, wobei ein $ICC \leq 0.4$ als "poor", $0.4 < ICC \leq 0.6$ als "fair", $0.6 < ICC \leq 0.8$ als "good" und $ICC > 0.8$ als "excellent" bewertet wird (Caceres et al., 2009a; Fliessbach et al., 2010). Die Landmark

¹ Für die Berechnung des ICC (3,1) werden "Rater" (in der fMRT typischerweise „Messwiederholungen“) als feste Effekte ("fixed effects") behandelt, während sie z.B. im Rahmen des ICC (2,1) als randomisierte Effekte ("random effects") modelliert werden. Daher werden beim ICC (3,1) Effekte, die auf den "Rater" bzw. die jeweilige Messung zurückzuführen sind, nicht berücksichtigt (siehe Shrout und Fleiß, 1979).

Aufgabe erreichte sowohl auf Voxel Ebene als auch innerhalb der ROIs jeweils nur schlechte ("poor") Reliabilitätswerte (Abb. 3).

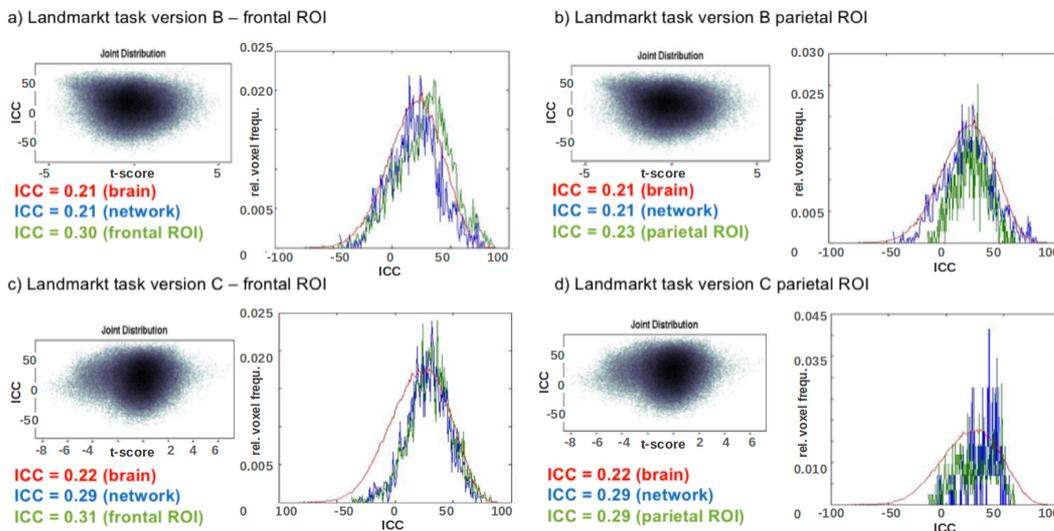


Abbildung 3 (modifiziert nach Schuster et al., 2017). ICC Ergebnisse, dargestellt für die jeweilige Version der Landmark Aufgabe (Version B, oben; Version C, unten) und die jeweilige ROI (frontal, links; parietal, rechts): "Joint probability distribution" der voxelweisen t-Werte und die jeweiligen ICC-Werte im Median (links) sowie ICC Häufigkeitsverteilungen für die Voxel des gesamten Gehirns ("brain", rot), für die Voxel des aktivierten Netzwerks ("network", blau) und für die frontalen und parietalen ROIs (grün). Das aktivierte Netzwerk wurde anhand der Ergebnisse der ersten Messung definiert (aktivierte Voxel bei $p < 0.001$, unkorrigiert, $k = 20$). k : Clustergröße.

Da das primäre Ziel der Studie darin bestand, eine experimentelle Aufgabe zu identifizieren, die eine stabile Rechtslateralisierung visuell-räumlicher Prozesse innerhalb der Probanden hervorruft, testeten wir den *Lateralisationsindex (LI)* auf seine Reliabilität. Der Lateralisationsindex ist eine in der fMRT häufig genutzte Methode, um das Ausmaß der Lateralisation neuraler Prozesse zu quantifizieren. Den LI beschreibt die folgende Formel,

$$LI = \frac{A_L - A_R}{A_L + A_R}$$

wobei A für ein Aktivitätsmaß der fMRT steht, hier bezogen auf die jeweils symmetrischen ROIs (frontal, parietal) der linken (L) und rechten (R) Hemisphäre. Zur Quantifizierung des Aktivitätsmaßes nutzten wir den derzeitigen Goldstandard in Form eines bootstrapping Ansatzes (Wilke and Schmithorst, 2006).

Für Version B der Landmark Aufgabe zeigten sich im Durchschnitt akzeptable ("fair") Reliabilitäten der LIs der frontalen (ICC = 0.57) und parietalen (ICC = 0.54) ROIs. Version C der Landmark Aufgabe erzielte im Mittel einen exzellenten ("excellent") ICC (ICC = 0.81) für die LIs der parietalen ROIs und einen schlechten ("poor") ICC (ICC = 0.33) für die LIs der frontalen ROIs (vgl. Abb. 4).

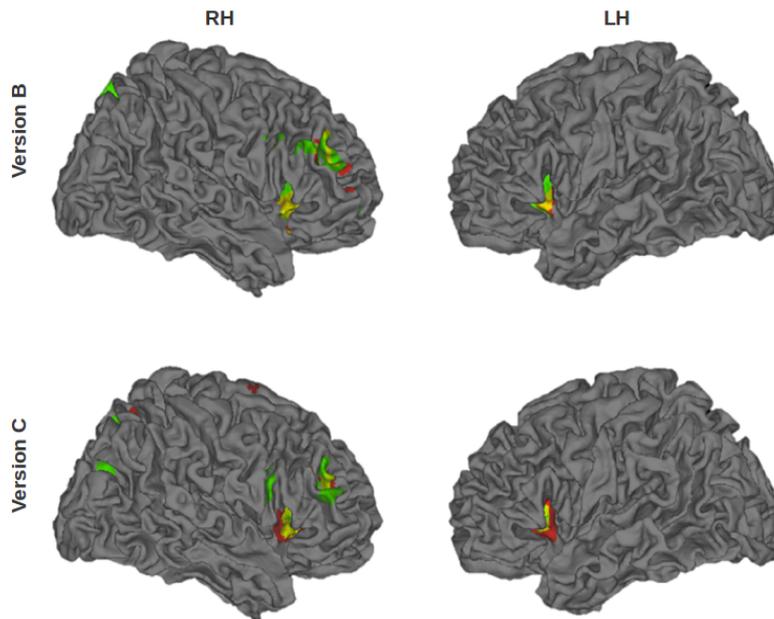


Abbildung 4 (modifiziert nach Schuster et al., 2017). Gruppenaktivierungen der wiederholten Messungen der Landmark Aufgabe (Versionen B, oben; Version C, unten). Aktivierungen der jeweils ersten Messung sind in grün, Aktivierungen der zweiten Messung in rot dargestellt ($p < 0.001$, unkorrigiert, $k = 20$). Überlappende Aktivierungen der beiden Messungen sind in gelb markiert. RH: rechte Hemisphäre, LH: linke Hemisphäre, k : Clustergröße.

In einem letzten Schritt betrachteten wir die *Konsistenz* der Lateralisierung, die die Landmark Aufgabe hervorruft. Dafür wurde der relative Anteil an Probanden identifiziert, die in beiden Messungen eine Rechtslateralisierung in den frontalen und parietalen ROIs aufwiesen. Bei Anwendung eines liberalen aber gängigen Kriteriums, eines $LI > 0.2$ (Deppe et al., 2000; Jansen et al., 2004), zeigten für Version B 80.0% der Probanden Rechtslateralisation in den frontalen ROIs und 75.0% in den parietalen ROIs sowie für Version C 70.0% der Probanden in den frontalen und 55.6% der Probanden in den parietalen ROIs.

In Studie 1 dieser Dissertation verglichen wir verschiedene experimentelle Paradigmen hinsichtlich ihrer Verlässlichkeit, rechtshemisphärische Dominanz visuell-räumlicher Prozesse abzubilden. Diesbezüglich erwies sich die Landmark Aufgabe im Vergleich mit anderen gängigen Aufgaben als robustes Paradigma. Nichtsdestotrotz war auch die Landmark Aufgabe nicht in allen Probanden mit rechtslateralisierten Aktivierungen assoziiert. Dieser Befund überrascht nicht sonderlich, denn obwohl es allseits akzeptiert ist, dass visuell-räumliche Verarbeitung auf rechtslateralisierte Prozesse zurückgreift, wurden bereits im Rahmen verschiedener Studien Abweichungen von dieser überwiegenden Rechtsdominanz dokumentiert. Regelmäßig zeigen sich in einer nicht zu vernachlässigenden Anzahl von Probanden auch linkslateralisierte Aktivierungsmuster. Diese interindividuellen Unterschiede treten nicht nur in Studien zu visuell-räumlicher Verarbeitung auf, sondern zeigen sich auch in Untersuchungen zu anderen Prozessen, die im Allgemeinen als lateralisiert gelten, wie z.B. Sprache (Jansen et al., 2007; Knecht et al., 2000b, 2000a; Springer et al., 1999).

Weiterführende Analysen der Landmark Aufgabe ergaben eine schlechte Test-Retest-Reliabilität auf Voxelenebene. Demgegenüber zeigte sich im Bezug auf die Hauptfragestellung der Studie, dass die Rechtslateralisation visuell-räumlicher Prozesse (ermittelt durch den LI) mit ausreichend hoher Reliabilität und Konsistenz durch die Landmark Aufgabe, Version B, abgebildet werden kann. In Anbetracht der Ergebnisse der Studie 1 dieser Dissertation ist eine Nutzung der Landmark Aufgabe, insbesondere der Version B, entsprechend dann zu empfehlen, wenn die Fragestellung der Studie die Lateralisation visuell-räumlicher Prozesse betrifft. Beispielsweise könnten verschiedene lateralisierte Funktionen hinsichtlich ihrer Interaktionseffekte untersucht werden. Sollte allerdings, wie im Rahmen von Konnektivitätsanalysen, ein Interesse am zeitlichen Signalverlauf auf Voxelenebene bestehen, ist von einer Nutzung der Landmark Aufgabe abzuraten. Insgesamt verdeutlicht Studie 1 dieser Dissertation die Bedeutung von Robustheit und Reliabilität eines Paradigmas für die Interpretation von fMRT-Ergebnissen. Zukünftige Untersuchungen, die zum Ziel haben, mittels eines Localizers weiterführende Fragestellungen zu bearbeiten, sollten sich zunächst analytisch mit der Tauglichkeit des Localizers für ihre Fragestellung auseinandersetzen.

5.2 Studie 2. Illusory face detection in pure noise images: The role of interindividual variability in fMRI activation patterns

5.2.1 Bibliographische Details

Diese Untersuchung wurde durch die Autoren Zimmermann, K. M., Stratil, A. S., Thome, I., Sommer, J. und Jansen, A. durchgeführt und verschriftlicht. Die Studie wurde 2019 bei PLoS One (ISSN 1932-6203, volume 14, issue 1, pages e0209310) unter Open-Access-Bedingungen publiziert. Der Artikel ist unter <https://doi.org/10.1371/journal.pone.0209310> erhältlich.

5.2.2 Zusammenfassung Studie 2

Als *Face Pareidolia* wird der Eindruck bezeichnet, in einem Muster oder einer Struktur ein Gesicht zu erkennen. Da die sensorische, sogenannte "bottom-up", Information bei Face Pareidolia unvollständig ist und somit alleinig nicht dafür verantwortlich sein kann, den Eindruck eines Gesichts zu erzeugen, wird davon ausgegangen, dass endogene, sogenannte "top-down" Prozesse, eine wichtige Rolle spielen.

Zur Untersuchung der neuronalen Korrelate von Face Pareidolia werden häufig verrauschte Bilder eingesetzt, die unsystematische, randomisierte Muster zeigen (Abb. 5). Die Probanden werden dabei aufgefordert anzugeben, bei welchen Bildern sie den Eindruck haben, ein Gesicht zu sehen.

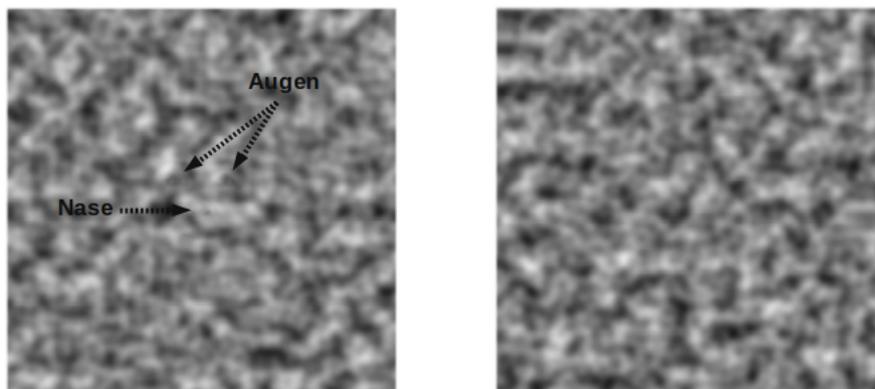


Abbildung 5 (modifiziert nach Zimmermann et al., 2019). Beispiel eines Trainingsstimulus (links) und eines experimentellen Stimulus (rechts) aus Zimmermann et al. (2019). Die Pfeile und die Beschriftungen des Gesichts im Trainingsstimulus wurden nachträglich eingefügt und waren nicht auf den originalen Stimuli abgebildet. Während einer Trainingsphase bestanden zunächst die Hälft-

te aller Stimuli aus Gesichtern, die mit randomisierten Mustern überlagert waren (links). In der Experimentalphase im MRT präsentierten wir nur noch die unsystematischen, randomisierten Muster ohne Gesichter ("pure noise images", rechts). Den Probanden wurde dennoch instruiert, dass sich in der Hälfte aller Stimuli der Experimentalphase ein Gesicht befinden würde. Sie wurden aufgefordert eine Taste zu drücken, sobald sie ein Gesicht in einem dargebotenen Muster erkennen.

Gruppenanalysen aus der bisherigen Literatur zu neuronalen Korrelaten von *Face Pareidolia* legten nahe, dass der im Rahmen von *Face Pareidolia* entstandene bloße Eindruck ein Gesicht zu sehen, frühe sensorische Gesichterareale involviert. Obwohl verrauschte Bilder also keine systematischen strukturellen Gesichtskonfigurationen enthalten, scheinen die Hinregionen rekrutiert zu werden, die eine Rolle für die visuell-sensorische Verarbeitung von Gesichtern spielen (Li et al., 2010; Liu et al., 2014; Zhang et al., 2008).

Das ursprüngliche Ziel der zweiten Studie dieser kumulativen Dissertation war die Untersuchung der "top-down" Prozesse, die für die Rekrutierung der sensorischen Gesichterareale bei *Face Pareidolia* verantwortlich gemacht werden. Mit Hilfe effektiver Konnektivitätsanalysen sollte der hypothetisierte directionale Einfluss frontaler Regionen auf die Gesichterareale überprüft werden.

Eine Voraussetzung von Konnektivitätsanalysen ist die robuste Aktivierung der Zielregionen im einzelnen Individuum. Die visuelle Wahrnehmung von Gesichtern ist typischerweise mit der robusten und reliablen Aktivierung eines rechtslateralisierten Netzwerks assoziiert, zu dem unter anderem die *Occipital Face Area (OFA)* und die *Fusiform Face Area (FFA)*, aber auch frontale Strukturen gehören (Duchaine and Yovel, 2015; Haxby et al., 2000) (vgl. auch Kapitel 4.1). Dementsprechend zeigten alle Probanden der Studie 2 Aktivierung in der rechten OFA und der rechten FFA, als ihnen im Rahmen eines Face Localizers (s. Kapitel 4.1) Bilder von Gesichtern präsentiert wurden (Abb. 6). Die Kernareale des Gesichternetzwerks wurden somit bei allen Probanden für die Verarbeitung tatsächlicher sensorischer Gesichtsinformationen rekrutiert.

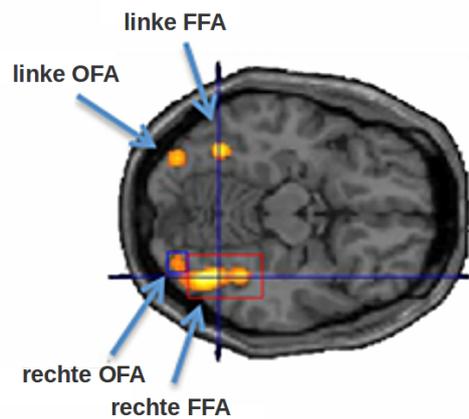


Abbildung 6 (modifiziert nach Zimmermann et al., 2019). Aktivierung eines Probanden aus Zimmermann et al. (2019) während des Face Localizers. Der Kontrast „2*Gesichter > (Objekte + verrauschte Bilder)“ ergab bei dem Probanden die dargestellte Aktivierung der bilateralen Occipital Face Area (OFA) und Fusiform Face Area (FFA) ($p < 0.05$, FWE-korrigiert, z-Koordinate = -14 mm).

In der Experimentalphase der Studie gaben die Probanden bei durchschnittlich 23.5% der verrauschten Bilder an, ein Gesicht in den Bildern erkannt zu haben (Standardabweichung 11.4%, Range 9.2% - 44.0%). Entgegen den Befunden bisheriger Untersuchungen zu Face Pareidolia zeigte sich in der Gruppenanalyse, trotz eines liberalen statistischen Schwellwerts, überraschenderweise jedoch keine Aktivierung in den Kernregionen der Gesichterverarbeitung (Abb. 7).

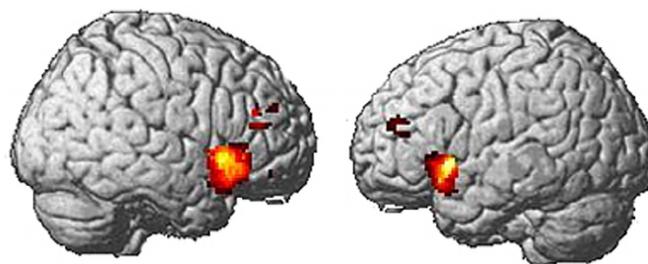


Abbildung 7 (aus Zimmermann et al., 2019). Gruppennaktivierungen der Experimentalphase für den Kontrast „Gesichter gesehen > keine Gesichter gesehen“ ($p < 0.001$, unkorrigiert), wobei den Probanden lediglich randomisiertes Rauschen präsentiert wurde (vgl. Abb. 5).

Eine Betrachtung der Aktivierungsmuster auf Einzelprobandenebene verdeutlichte, dass sich bezüglich der Rekrutierung der Kernareale des Gesichternetzwerks

eine nahezu dichotome Verteilung der Probanden ergab: nur etwa die Hälfte der Probanden zeigten die in der Literatur vorbeschriebene Aktivierung der sensorischen Gesichtsareale (Abb. 8).

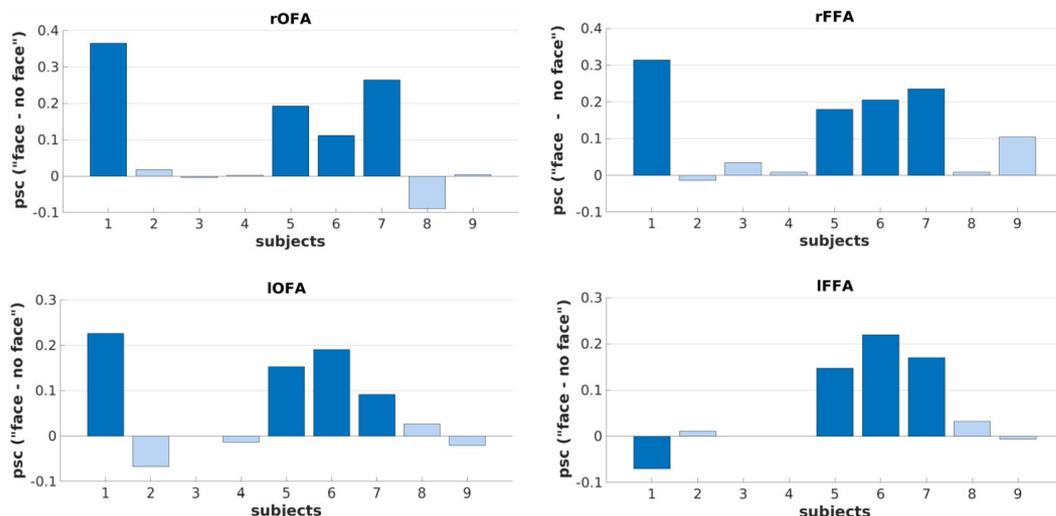


Abbildung 8 (modifiziert nach Zimmermann et al., 2019). Prozentuelle Signalveränderungen ("percent signal change", psc) in den Kernregionen der Gesichtsverarbeitung der einzelnen Probanden, bezogen auf den Kontrast „Gesichter gesehen > keine Gesichter gesehen“ der Experimentalphase.

In der fMRT führen multiple Faktoren dazu, dass eine gewisse Streubreite zwischen individuellen Aktivierungen durchaus typisch ist (vgl. Kapitel 4). Das Besondere an den Ergebnissen aus Studie 2 ist aber, dass sich aus den individuellen Aktivierungen der Gesichtsareale nicht erwartungsgemäß eine Gauß'sche, sondern eine dichotome Verteilung ergibt. Diese Extremform interindividueller Varianz kann verschiedene Ursachen haben. Einerseits könnten intrinsische Unterschiede zwischen den Probanden vorliegen (vgl. Kapitel 4.2), so dass grundlegend verschiedene neurale Netzwerke für Face Pareidolia rekrutiert werden (siehe z.B. Barch und Yarkoni, 2013; Seghier und Price, 2018). Auf der anderen Seite könnten die Probanden auf unterschiedliche Verhaltensstrategien zurückgegriffen haben, um die experimentelle Aufgabe zu bearbeiten. So könnten beispielsweise manche Probanden nach augenähnlichen Strukturen in den Rauschbildern gesucht haben, während andere sich auf den Umriss eines Kopfes oder die Nase konzentriert haben. In diesen Fällen wäre jedoch trotzdem die Erwartung, dass sensorische Gesichtsareale bei derartigen Eindrücken involviert sind.

Eine weitere Möglichkeit ist, dass kontextuelle Unterschiede zu der Ergebnisvarianz in Studie 2 geführt haben (Kapitel 4.2). In Studien zu Face Pareidolia wird die Zuordnung zu Experimental- oder Kontrollbedingung nicht durch vom Versuchsleiter etablierte Manipulationen definiert, sondern ergibt sich durch die individuellen Angaben der Probanden. Insbesondere da wir die Probanden instruierten, dass in 50 % der Fälle ein Gesicht in den verrauschten Bildern zu sehen sei (obwohl die Bilder nur randomisiertes Rauschen zeigten), liegt die Vermutung nahe, dass sie teilweise angaben, ein Gesicht erkannt zu haben, ohne dass tatsächlich Face Pareidolia aufgetreten ist. Auf diese Weise wären kontextuelle Faktoren im Rahmen eines Versuchsleitereffekts oder sozial erwünschten Verhaltens für die Unterschiede zwischen den Probanden verantwortlich.

Um die interne Validität von Studien zu Face Pareidolia zu verbessern, sollten sich zukünftige Untersuchungen nicht auf Gruppenergebnisse beschränken, sondern Aktivierungsmuster auf individueller Ebene analysieren. Außerdem sollten nicht nur die Antworten der Probanden als Indikator für *Face Pareidolia* verwendet werden. So wird beispielsweise bereits der Ansatz verfolgt, die Bilder in denen ein Proband angab, Gesichter gesehen zu haben, mit bestimmten Techniken zusammenzufassen (Smith et al., 2012). Durch Überlagerung der Stimuli können auf diese Weise gesichtsähnliche Strukturen sichtbar gemacht werden, auf die der Proband möglicherweise geachtet hat. Auch das Erheben von Maßen, die soziale Erwünschtheit erfassen, wäre zu empfehlen.

Insgesamt wäre es wünschenswert, in Zukunft anhand von verdecktem statt offenem Verhalten auf Face Pareidolia schließen zu können. Verdeckte Antworten haben den Vorteil, dass sie unwillkürliche Reaktionen widerspiegeln, die wenig bis gar nicht aktiv beeinflusst werden können. Beispiele für typische Maße, die genutzt werden, um unbewusstes Verhalten zu quantifizieren, sind Reaktionszeiten und verschiedene Arten von Augenbewegungen. Eine kürzlich erschienene Studie von Wang et al. (2019), die unbewusste Muster von Augenbewegungen mit fMRT-Signalen in Verbindung brachte, könnte dabei von besonderem Interesse für die Untersuchung von Face Pareidolia sein. Die Autoren fanden, dass die Ausführung bestimmter Fixationsmuster mit differenzierten Signalen in der FFA zusammenhängt. Dabei führten Augenbewegungen, die während des Betrachtens eines Gesichtsstimulus aufgezeichnet wurden, auch in kompletter Abwesenheit einer Gesichtsstruktur zu gesichtsselektiven Aktivierungsmustern in der FFA. Für

die Untersuchung von Face Pareidolia könnte dies bedeuten, dass unbewusste Augenbewegungen eventuell zukünftig als verdecktes Maß herangezogen werden könnten, um zu detektieren, ob Face Pareidolia stattfindet. Zu diesem Zweck sollte der Zusammenhang zwischen Signalen in Kernregionen der Gesichtsverarbeitung und spezifischen Fixationsmustern im Kontext von Face Pareidolia näher untersucht werden.

5.3 Studie 3. Seeing things differently: Gaze shapes neural signal during mentalizing in alexithymia

5.3.1 Bibliographische Details

Die Studie wurde durch die Autoren Zimmermann, K. M., Schmidt, K., Gronow, F., Sommer, J., Leweke, F., und Jansen, A. durchgeführt und verschriftlicht. Das Manuskript ist fertiggestellt und wurde bei einer internationalen, wissenschaftlichen Zeitschrift mit Peer-Review Verfahren eingereicht.

5.3.2 Zusammenfassung Studie 3

In fMRT-Studien zu sozialen Kognitionen wird häufig komplexes visuelles Stimulusmaterial verwendet. Die Komplexität dieses Stimulusmaterials zeichnet sich dadurch aus, dass gleichzeitig verschiedene Aspekte visuell dargestellt werden, die der Proband fixieren und so in das Zentrum seiner visuellen Aufmerksamkeit rücken kann. Ist das Blickverhalten des Probanden nicht zentral für die eigentliche Fragestellung der Studie, wird die Kontrolle dieses Verhaltens in vielen Fällen vernachlässigt. Das ist problematisch, da sich visuelle Aufmerksamkeitsausrichtung auf verschiedene Stimulusaspekte in unterschiedlichen neuronalen Prozessen äußern kann.

Schon vor 15 Jahren beschrieben Dalton et al. (2005) Abhängigkeiten zwischen fMRT-Signal und natürlich auftretendem Blickverhalten. Dabei zeigten sich Signalveränderungen der FFA und der Amygdala, die mit der Fixationsdauer auf die Augen präsentierter Gesichter zusammenhingen. Seitdem bedienten sich nur wenige Studien zu sozialen Kognitionen ähnlicher Methoden, um den Einfluss natürlichen Blickverhaltens auf neurale Prozesse zu untersuchen (Dalton et al., 2008, 2007; Gamer et al., 2010; Gamer and Büchel, 2009; Kliemann et al., 2012). Die meisten fMRT-Studien aus dem Bereich kontrollierten Blickverhalten, indem sie

das zu fokussierende Ziel vorgaben, was mit der Unterbindung natürlicher Verhaltensweisen einhergeht (Hadjikhani et al., 2017a, 2017b; Lassalle et al., 2017; Morris et al., 2008, 2007; Perlman et al., 2011; Spilka et al., 2019; Zürcher et al., 2013). Andere wiederum zeichneten Augenbewegungen zu einem anderen Zeitpunkt auf als das fMRT-Signal (Tottenham et al., 2014), werteten simultan erhobene Augen- und fMRT-Daten getrennt voneinander aus (Kujala et al., 2012), oder ignorierten den potentiell verzerrenden Einfluss visueller Aufmerksamkeit. Das Ausmaß dieser Verzerrungen wurde erst kürzlich durch eine Untersuchung von Hadjikhani et al. (2017b) demonstriert, in der gesunde Probanden Gesichter unter zwei Bedingungen betrachten sollten: mit und ohne Aufmerksamkeitslenkung auf die Augen. Die Aufmerksamkeitslenkung auf die Augen war assoziiert mit stärkeren Aktivitäten nicht nur in frühen visuellen Arealen, sondern in multimodalen Assoziationskortex sowie hierarchisch höheren frontalen Bereichen. Dazu zählten auch die Regionen des *Theory of Mind (ToM)* Netzwerks.

In Studie 3 dieser Dissertation untersuchten wir neurale ToM Prozesse in Abhängigkeit von *Alexithymie*. Alexithymie bezeichnet ein Persönlichkeitsmerkmal, das bei hoher Ausprägung mit Schwierigkeiten einhergeht, sowohl die eigenen Gefühle als auch die des Gegenübers zu erkennen und zu beschreiben. Als experimentelle Aufgabe nutzten wir den *Reading the Mind in the Eyes Test (RMET)*, bei dem aus vier Wörtern das ausgewählt werden soll, welches am besten zu einem präsentierten Augenpaar passt (Abb. 9). Das Aufgabendesign des RMET bedient sich komplexer visueller Stimuli, die verschiedene Aufmerksamkeitsstrategien zulassen. Da in einem solchen Kontext unterschiedliches Blickverhalten das fMRT-Signal in ToM-assozierten Arealen bedeutend beeinflussen könnte (vgl. Hadjikhani et al., 2017b), kombinierten wir die Aufnahme und Analyse von fMRT-Signal mit *Eyetracking*.

Aus vorherigen Studien ging zudem hervor, dass Alexithymie mit einer atypisch geringen Aufmerksamkeit auf Augen assoziiert ist. Wir hypothesierten, dass hoch Alexithyme (HA) auch während einer ToM Aufgabe weniger auf die Augen fokussieren als niedrig Alexithyme (LA). Außerdem untersuchten wir, ob die Fokussierung verschiedener Stimulusaspekte bei HA und LA mit unterschiedlichen Prozessen im ToM Netzwerk einhergeht. Dazu fügten wir die kumulierte Dauer der Fixationen auf Augen und auf Wörter als parametrische Modulatoren in unser statistisches Modell zur Auswertung der fMRT-Daten ein (Abb. 9).

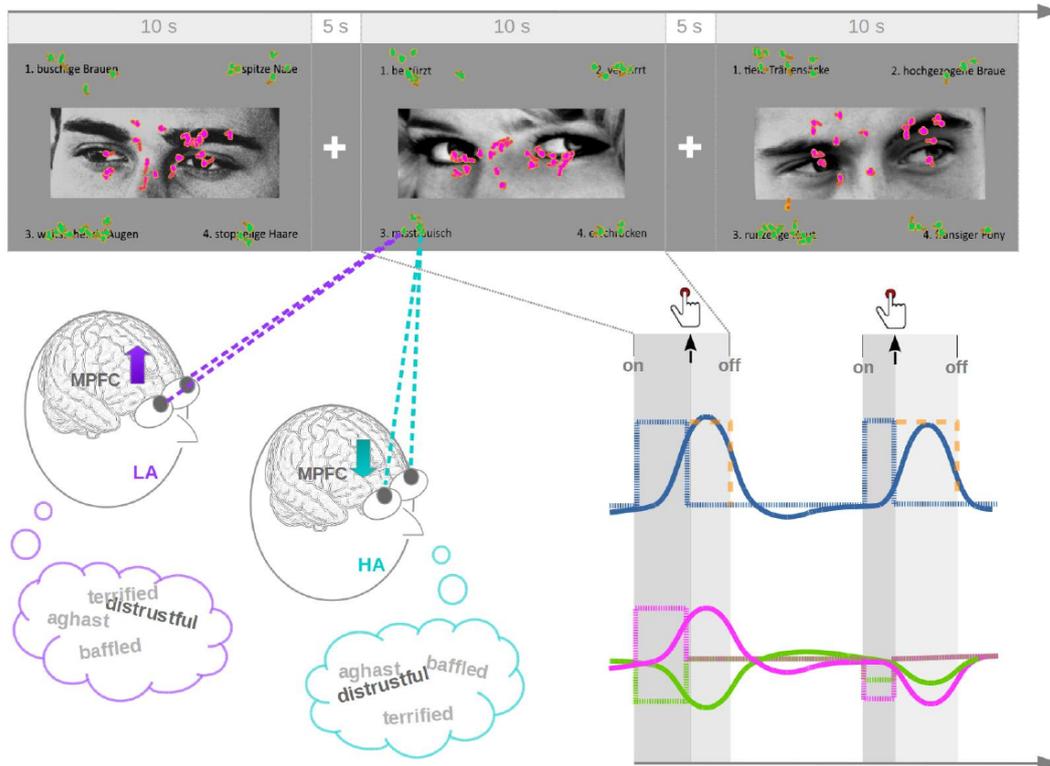


Abbildung 9 (aus Zimmermann et al., 2020). Überblick über das experimentelle Paradigma, die parametrische Modellierung und grundlegende Ergebnisse. Während der Präsentation des Stimulusmaterials (RMET, s. Bilder von Augenpaaren und Wörtern, oben) wurden sowohl Fixationen (magenta und grüne Kreise auf den Stimuli) als auch das fMRT-Signal der Probanden aufgezeichnet. Für die Analyse der ToM Bedingung (mittlerer Stimulus) wichtige Regressoren sind anhand der Grafik unten rechts beispielhaft veranschaulicht: die Zeit vom Beginn der Stimuluspräsentation ("on") bis zur Entscheidung für eines der Wörter (Tastendruck, Pfeil) modellierten wir durch eine Rechteckfunktion (fein gestrichelte blaue Linie), die mit der haemodynamischen Antwortfunktion (vgl. Kapitel 3) gefaltet wurde (durchgezogene blaue Linie). Dieser Regressor repräsentiert die mittlere BOLD-Antwort während der ToM Bedingung. Außerdem fügten wir die individuellen Fixationszeiten als orthogonalisierte parametrische Modulatoren in das statistische Modell ein. Pro Stimulus kumulierten wir dafür die Dauer aller Fixationen auf die Augen (magenta) und auf die Wörter (grün). Die gleichen Regressoren (mittlere BOLD-Antwort sowie parametrische Modulatoren für die Fixationsdauer auf Augen und auf Wörter) wurden auch für die Kontrollbedingung (rechter und linker Stimulus) erstellt und ins Modell aufgenommen. Zusätzlich modellierten wir pro Bedingung (ToM und Kontrollbedingung) die Zeit nach der Entscheidung des Probanden (Tastendruck, Pfeil) bis zum Präsentationsende ("off") des Stimulus (grob gestrichelte gelbe Rechteckfunktion für die ToM Bedingung). Sowohl das Ergebnis der Faltung der gelben Rechteckfunktion als auch die jeweils zwei zusätzlichen parametrischen Modulatoren (für Fixation auf die Augen und auf die Wörter pro Bedingung) sind aus Übersichtsgründen hier nicht abgebildet. Während der Fixation auf die Wörter der ToM Bedingung ergaben sich für LA und HA entgegengerichtete neurale Signale im medialen Präfrontalkortex (MPFC; ROI-Analysen). ToM: Theory of Mind; RMET: Reading the Mind in the Eyes Test; LA: niedrig Alexithyme; HA: hoch Alexithyme; MPFC: Medialer Präfrontalkortex; ROI: Region-of-Interest.

Es zeigte sich, dass Alexithymie auch während der Ausführung von ToM Aufgaben mit einer atypisch geringen Fokussierung auf Augen assoziiert ist. Interessanterweise zeigten die hoch alexithymen Probanden dieses Verhalten auch dann, wenn die ToM oder Kontrollaufgabe bereits erfüllt war und sie ohne weitere Instruktionen den aktuellen Stimulus betrachteten (Abb. 10).

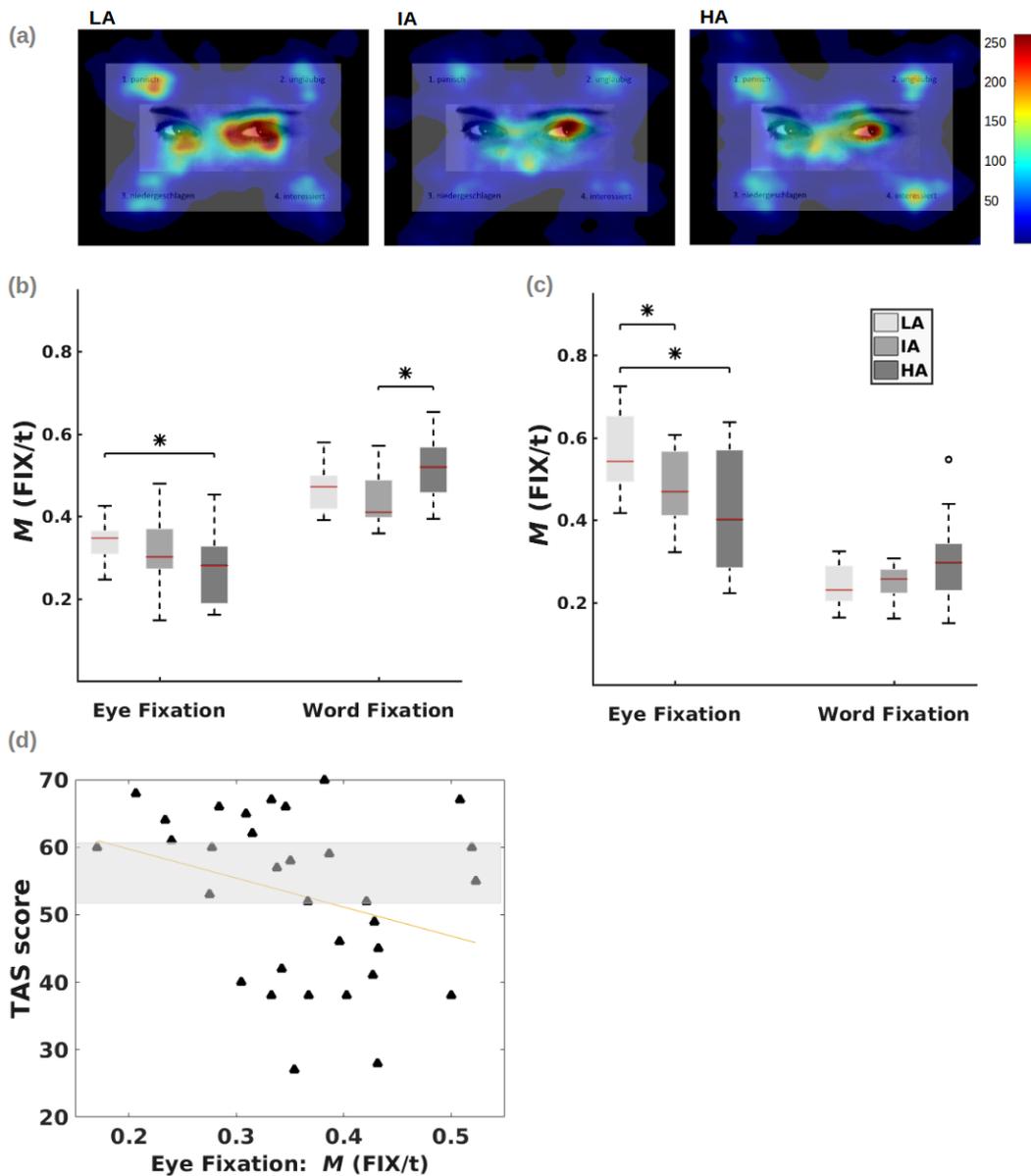


Abbildung 10 (aus Zimmermann et al., 2020). Darstellung der unterschiedlichen Verteilungen visueller Aufmerksamkeit in Abhängigkeit vom Alexithymielevel. (a) Heat maps der Fixationsdauer kumuliert über alle ToM Stimuli für LA (links), IA (mittig) und HA (rechts). (b, c) Darstellung der post-hoc Vergleiche (t -teste, $p < .05$) der mittleren relativen Fixationsdauer (" M FIX/t") auf Augen

oder Wörter, die sich aus der signifikanten Interaktion (ANOVA) von Alexithymielevel (LA, IA, HA) und visuellem Fokus (Fixationsdauer auf Augen und auf Wörter) ergaben. Es zeigte sich keine signifikante Interaktion mit dem Faktor ‚Experimentelle Aufgabe‘ (ToM, Kontrollbedingung), weshalb sich die Abbildungen auf Fixationen kumuliert über (b) beide experimentelle Aufgaben und (c) die Zeiten der Stimuluspräsentationen nachdem die jeweilige experimentelle Aufgabe durch den Tastendruck beendet war beziehen. Die Fixationszeiten wurden jeweils in Relation zur Dauer der Entscheidung (Beginn Stimuluspräsentation bis zum Tastendruck, (b)) oder zur Dauer von Entscheidung bis zum Ausblenden des aktuellen Bildes (Tastendruck bis zum Ende der Stimuluspräsentation, (c)) ermittelt ("M FIX/t"). (d) Signifikante Korrelation ($p < .05$) zwischen dem Alexithymielevel (individueller Wert auf der *Toronto Alexithymia Scale*, TAS) und der mittleren Fixationsdauer auf Augen oder Wörter während der ToM Bedingung. * $< .05$; ToM: Theory of Mind; LA: niedrig Alexithyme; IA: mittelgradig Alexithyme; HA: hoch Alexithyme; TAS: Toronto Alexithymia Scale.

Auf neuraler Ebene fanden wir, dass Alexithymie *per se*, anders als in einer vorherigen Studie vermutet (Moriguchi et al., 2006), nicht mit veränderten ToM Prozessen assoziiert war. Stattdessen zeigte sich in unserer Studie, dass die interindividuellen Unterschiede in neuralen Aktivierungen auf eine Interaktion zwischen Alexithymie und Blickverhalten zurückzuführen waren: während die Aktivität in Kernregionen des ToM Netzwerks bei Probanden mit niedrigerem Alexithymielevel zunahm, je länger sie auf kritische Stimulusaspekte blickten, nahm das Aktivierungsniveau bei Probanden mit hohem Alexithymielevel ab, je länger sie die gleichen Stimulusaspekte betrachteten (Abb. 11 und Abb. 12). Es ist zu betonen, dass sich diese Unterschiede nicht aufgrund von unterschiedlichem Blickverhalten ergaben, sondern im Rahmen des gleichen visuellen Fokus auftraten. Das bedeutet, dass sich die neuralen Aktivitäten von hoch und niedrig Alexithymen gegensätzlich verhielten, während die Probanden denselben Stimulusaspekt fixierten. Da sich diese Unterschiede als spezifisch für verschiedene neurale Regionen herausstellten, war es zudem möglich, die einzelnen Regionen mit verschiedenen funktionellen Aufgabenaspekten in Verbindung zu bringen. So zeigten sich für den medialen Präfrontalkortex (MPFC) nur Unterschiede zwischen den Probanden, wenn diese auf die Wörter der Stimuli blickten (Abb. 11). Für die anteriore temporo-parietale Junktion (TPJa) ergaben sich hingegen nur dann Unterschiede, wenn die Probanden dargestellte Augenpaare fixierten (Abb. 12).

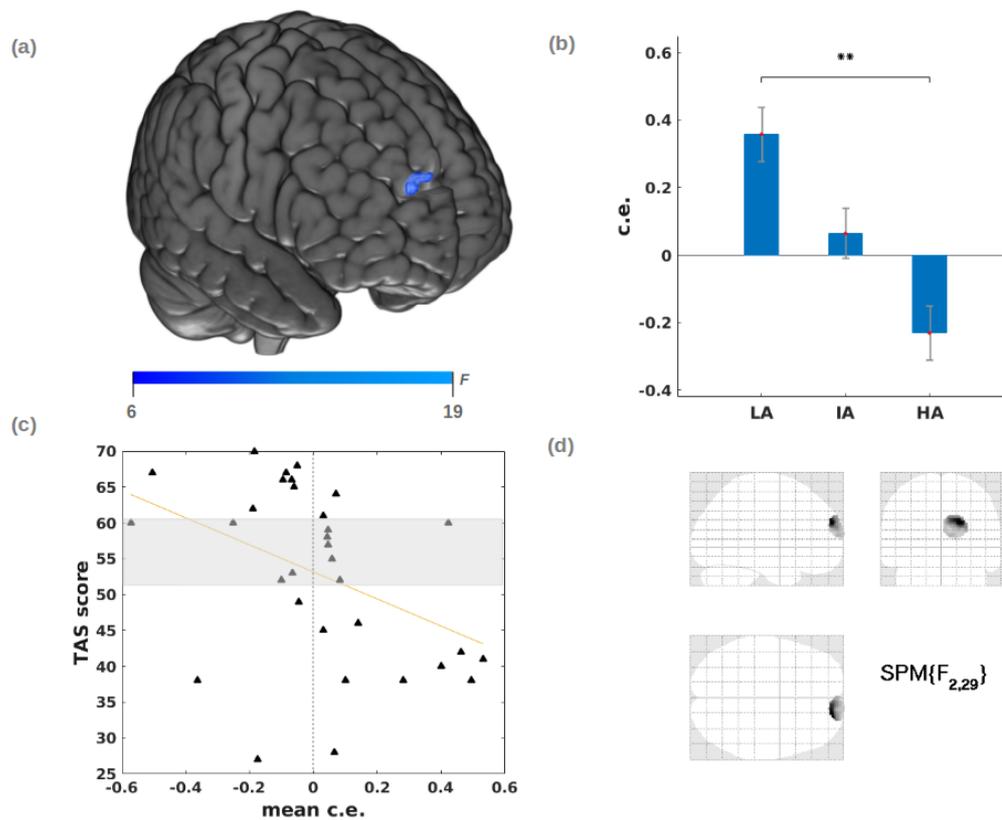


Abbildung 11 (aus Zimmermann et al., 2020). Effekte des Betrachtens von ToM Begriffen auf neurale Prozesse des MPFC in Abhängigkeit von Alexithymie. (a) Haupteffekt der einfaktoriellen ANOVA (Faktor Alexithymiegruppe): in Abhängigkeit der Alexithymiegruppe zeigten sich signifikant unterschiedliche Effekte von visueller Aufmerksamkeit auf Wörter (PMw) auf das Signal in der MPFC ROI ($p < .05$, FWE-korrigiert). Zum Zweck der Visualisierung definierten wir in der Abbildung eine statistische Schwelle von $p < .005$, unkorrigiert. (b) Darstellung der post-hoc Vergleiche (t -Teste): *contrast estimates* (c.e.) der am stärksten aktivierten Voxel ("peak voxel") pro Alexithymiegruppe. (c) *Scatter plot* der individuellen TAS Werte und des Kontrasts aus den c.e. der ToM Bedingung und den c.e. der Kontrollbedingung, gemittelt über die MPFC ROI ("mean c.e."). Die Abbildung stellt den signifikanten linearen Zusammenhang zwischen der individuellen Alexithymieausprägung und dem MPFC Signal während der Fixierung von Wörtern dar ($p < .05$). (d) *Glass brain* Visualisierung der Aktivierungsstärke innerhalb der MPFC ROI, in denen sich der in (a) abgebildete Haupteffekt zeigte. $**p < .01$; MPFC: medialer Präfrontalkortex; ToM: Theory of Mind; PMw: parametrischer Modulator bezogen auf die kumulierte Fixationsdauer von Wörtern; LA: niedrig Alexithymie; IA: mittelgradig Alexithymie; HA: hoch Alexithymie; TAS: Toronto Alexithymia Scale; c.e.: contrast estimate; ROI: Region-of-Interest.

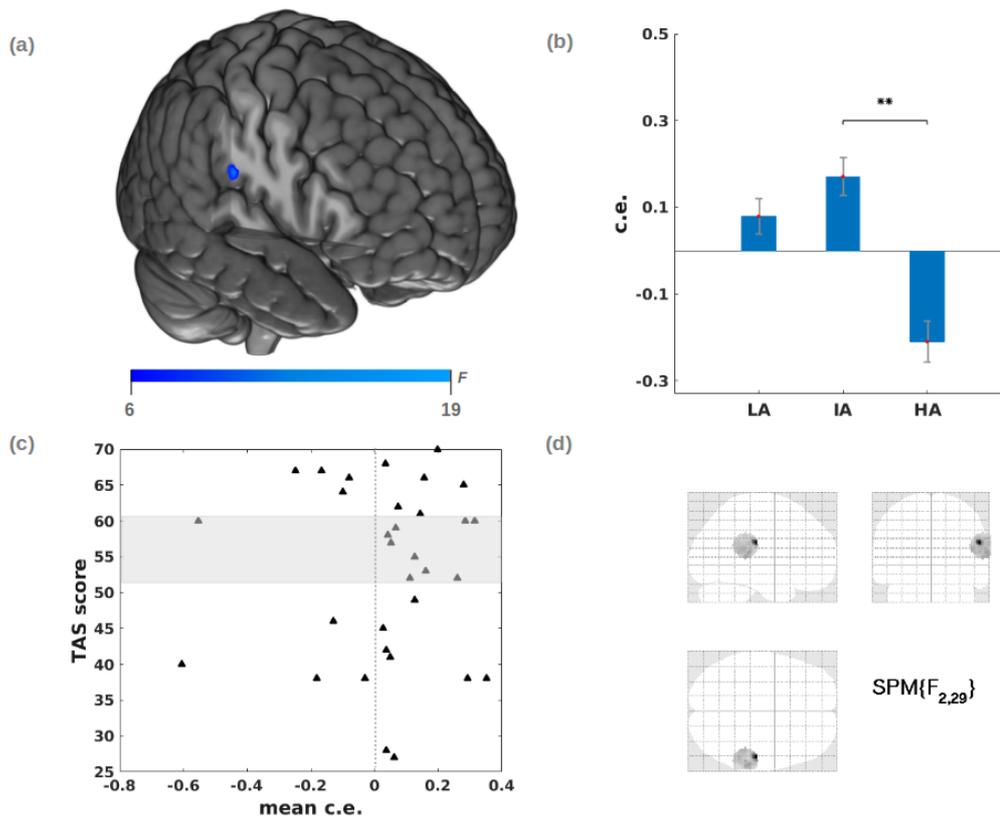


Abbildung 12 (nach Zimmermann et al., 2020). Effekte des Betrachtens der Augenpaare auf neurale Prozesse der TPJa während der ToM Bedingung in Abhängigkeit von Alexithymie. (a) Haupteffekt der einfaktoriellen ANOVA (Faktor Alexithymiegruppe): in Abhängigkeit der Alexithymiegruppe zeigten sich signifikant unterschiedliche Effekte des Betrachtens der Augenpaare (PMe) auf das Signal in der TPJa ROI ($p < .05$, FWE-korrigiert). Zum Zweck der Visualisierung definieren wir in der Abbildung eine statistische Schwelle von $p < .005$, unkorrigiert. (b) Darstellung der post-hoc Vergleiche (t -Teste): *contrast estimates* (c.e.) der am stärksten aktivierten Voxel ("peak voxel") pro Alexithymiegruppe. (c) *Scatter plot* der individuellen TAS Werte und des Kontrasts aus c.e. der ToM Bedingung und c.e. der Kontrollbedingung, gemittelt über die TPJa ROI ("mean c.e."). Bezogen auf die gesamte TPJa ROI ergab sich kein signifikanter linearer Zusammenhang zwischen der individuellen Alexithymieausprägung und dem TPJa Signal während der Fixierung der Augenpaare ($p < .05$). (d) *Glass brain* Visualisierung der Aktivierungsstärke innerhalb der TPJa ROI, in denen sich der in (a) abgebildete Haupteffekt zeigte. $**p < .01$; TPJa: anteriore temporo-parietale Junktion; ToM: Theory of Mind; PMe: parametrischer Modulator bezogen auf die kumulierte Fixationsdauer auf die Augen; LA: niedrig Alexithymie; IA: mittelgradig Alexithymie; HA: hoch Alexithymie; TAS: Toronto Alexithymia Scale; c.e.: contrast estimate; ROI: Region-of-Interest.

Mit unserer Studie verdeutlichen wir Effekte visueller Aufmerksamkeit auf das fMRT-Signal. Von besonderer Bedeutung ist die Kontrolle dieser Effekte, wenn Probanden untersucht werden, bei denen atypisches Blickverhalten bereits

bekannt ist, wie bei Personen mit Autismus-Spektrum-Störung, Schizophrenie, Depression oder sozialer Phobie (Toh et al., 2011). Aber auch Studien, in denen gesunde Probanden untersucht werden, können von einer Kontrolle des Blickverhaltens profitieren. So wurde bis vor einigen Jahren davon ausgegangen, dass das Blickverhalten von Gesunden vor allem durch den Kontext, d.h., durch Charakteristika des Stimulus, beeinflusst wird. Mittlerweile verdichten sich jedoch die Befunde dazu, dass Erfahrungen, Fertigkeiten, oder Eigenschaften eines Individuums, wie z.B. Neugier oder eben Alexithymie, die Verteilung der visuellen Aufmerksamkeit verändern (de Haas et al., 2019; Gottlieb and Oudeyer, 2018; Henderson, 2017; Risko et al., 2012; Schomaker et al., 2017).

Die Kombination von fMRT und Eyetracking könnte in Zukunft einen wichtigen Beitrag dazu leisten, die interne Validität von fMRT-Studien zu verbessern. Einerseits kann durch eine Kontrolle der Einflüsse von visueller Aufmerksamkeit auf das fMRT-Signal verhindert werden, dass Änderungen in der Aktivität, die eigentlich auf unterschiedliches Betrachten der Stimuli zurückzuführen sind, falsch attribuiert werden. Auf der anderen Seite ermöglicht eine kombinierte Analyse von Blickbewegungen und fMRT-Signal die Identifikation von neuronalen Subprozessen, die mit der Verarbeitung bestimmter Stimulusanteile zusammenhängen. Wie in Studie 3 dieser Dissertation können so Unterschiede zwischen Probanden mit einzelnen, spezifischen Funktionen in Verbindung gebracht werden, die einen isolierten Beitrag für die Bearbeitung der gesamten Aufgabe leisten. So könnte man sich verschiedenen kognitiven Strategien annähern, die unterschiedliche Probanden zur Bearbeitung derselben Aufgabe einsetzen. Insgesamt könnte die statistische Modellierung von Blickverhalten bei Einsatz eines Paradigmas mit komplexem Stimulusmaterial also signifikant zur Varianzaufklärung beitragen, nicht nur dann, wenn sich Probanden systematisch in ihrem Blickverhalten unterscheiden.

6 Diskussion und Ausblick

In dieser kumulativen Dissertation betrachte ich die Auswirkungen von experimentellem Paradigma und interindividuellen Unterschieden auf die Interpretation von fMRT-Ergebnissen. Nach einer kurzen geschichtlichen Zusammenfassung neurowissenschaftlicher Methodenentwicklungen folgte eine Einführung in die

funktionelle Magnetresonanztomographie (fMRT). Schließlich beschrieb ich Auswirkungen experimenteller Aufgaben und interindividueller Unterschiede auf das fMRT-Signal und stellte die in dieser Dissertation zusammengefassten Originalarbeiten vor.

In Studie 1 dieser Dissertation untersuchten wir verschiedene experimentelle Paradigmen hinsichtlich ihrer Robustheit, die Rechtslateralisierung visuell-räumlicher Verarbeitungsmechanismen abzubilden. Das robusteste Paradigma, die sogenannte Landmark Aufgabe, testeten wir auf seine Reliabilität und Konsistenz. Dabei ergaben sich nur schwache Reliabilitäten auf Voxel Ebene, aber gute bis exzellente Reliabilitäten und Konsistenzen bezüglich der Rechtslateralisation. Die Landmark Aufgabe ist demnach zur Untersuchung lateralisierter Prozesse, nicht aber zur Ermittlung konkreter Signalverläufe auf Voxel Ebene geeignet.

In Zeiten der Replikationskrise unterstreicht Studie 1 die Bedeutung von Eignungsprüfungen experimenteller Paradigmen in der fMRT, in der es einerseits üblich ist, komplexe Paradigmen einzusetzen, aber nur anhand der Augenscheinvalidität über die Eignung dieser zu urteilen. Jedoch hat auch die Quantifizierung der Reliabilität, wie sie in Studie 1 vorgenommen wurde, ihre Tücken. So hängt der ICC nicht nur vom intraindividuellen Messfehler ab, sondern auch von der Variabilität zwischen Probanden. Ist diese niedrig, können keine hohen ICC-Werte resultieren. Der ICC ist demnach auf eine heterogene Stichprobe angewiesen. Bei einem Paradigma, welches beispielsweise als Localizer eingesetzt werden soll, sind allerdings bei jedem Probanden starke Aktivierungen erwünscht. So könnte sich hinsichtlich der Aktivierungsstärken eine homogene Stichprobe ergeben. Typischerweise ist jedoch die Between-Subject Varianz größer als die Within-Subject Varianz (Wei et al., 2004), was die eventuelle Gruppenhomogenität im Regelfall relativieren sollte.

Neben diesen Schwierigkeiten, die sich aus der Konzeptualisierung des ICC ergeben, ist der Einsatz von Aktivierungswerten auf Voxel Ebene zur Beurteilung der Eignung eines Paradigmas an sich problematisch. Zum Beispiel zeigen Caceres et al. (2009), dass die Voxel, die die höchsten Aktivierungswerte auf Gruppenebene aufweisen, sich auch als reliabler bei Messwiederholung erweisen. Wie in Kapitel 3 dieser Arbeit dargestellt sind die Abhängigkeiten zwischen fMRT-Signal und neuraler Aktivierung aber indirekter Natur und werden von Dy-

namiken des oxygenierten und desoxygenierten Hämoglobins bestimmt. Bereits in den Anfängen der fMRT-Forschung wurde in diesem Zusammenhang gefunden, dass Voxel, die in großen drainierenden Venen liegen, nicht nur höhere Aktivierungswerte aufweisen als Voxel in kleineren, parenchymalen Venen, sondern auch mit höherer Wahrscheinlichkeit überhaupt signalgebend sind (Krings et al., 2001). Dabei sind kleinere, parenchymale Venen informativer was den eigentlichen Ursprung ihres Signals angeht, da das Drainagepotenzial kleinerer Venen auf ein kleineres Hirnvolumen beschränkt ist. Im Umkehrschluss ist es also möglich, dass sich höhere Reliabilitätswerte für Voxel aus großen Venen ergeben, die simultan mehrere verschiedene Hirnareale drainieren und in ihrer Lokalisation und ihrem Signalverlauf somit weniger spezifisch für den interessierenden Prozess sind (Boubela et al., 2015).

Aus der indirekten Beziehung zwischen fMRT-Signal und neuraler Aktivität ergibt sich also – trotz einer im Vergleich mit anderen bildgebenden Methoden hohen räumlichen Auflösung (vgl. Abb. 1) – eine Unsicherheit in der Lokalisation neuraler Prozesse (Menon, 2012). Neben Gütekriterien, wie diejenigen die in Studie 1 erhoben wurden (Robustheit, Reliabilität und Konsistenz), spielt dementsprechend auch die Spezifität der Aktivierungen eine Rolle. Die Spezifität des fMRT-Signals hängt nicht nur von physiologischen und physikalisch-technischen Faktoren, wie z.B. der Venenstruktur, Messequenzen oder Vorverarbeitungspipelines ab, sondern steht auch im direkten Zusammenhang mit dem experimentellen Paradigma, den gewählten Kontrasten und der statistischen Schwelle. Je präziser ein Paradigma die Voxel rekrutieren kann, welche die funktionspezifischen parenchymalen Anteile des Gehirns markieren, desto eher sollte es als Localizer für die bestimmte Funktion eingesetzt werden.

Insgesamt ist somit das bloße Ziel, durch ein experimentelles Paradigma Aktivierung in einem anatomisch definierten Bereich hervorzurufen, in Frage zu stellen. Localizer sollten eine präzise Lokalisation funktionell spezifischer Areale ermöglichen. Die Spezifität ist nicht nur im Rahmen der präoperativen Lokalisation funktioneller Domänen von Bedeutung (Brennan et al., 2016), sondern spielt auch für andere Einsatzbereiche eine Rolle, wie zum Beispiel für funktionelle Netzwerkanalysen (Stephan et al., 2010). Ein Mangel an Spezifität, Robustheit und Reliabilität von experimentellen Paradigmen ist mitverantwortlich für die Va-

riabilität der Ergebnisse verschiedener fMRT-Untersuchungen zu ein und demselben Thema, und letztendlich mitverantwortlich für die aktuelle Replikationskrise.

In Studie 2 beschäftigten wir uns mit den neuralen Korrelaten von Face Pareidolia. Die Studie stellte die erste Studie zu Face Pareidolia dar, die sich mit den Aktivierungsmustern auf individueller Ebene auseinandersetzte. Dabei verdeutlichen die Ergebnisse im Hinblick auf die hypothetisierte Rekrutierung der Kernareale des Gesichternetzwerks eine dichotome Verteilung der Probanden: nur etwa die Hälfte der Teilnehmer zeigte die erwarteten Aktivierungen, während diese bei der anderen Hälfte, auch bei liberaler statistischer Schwelle, nicht zu beobachten waren.

Die zukünftige Darstellung individueller Aktivierungskarten könnte dazu beitragen, Inkonsistenzen in Gruppenergebnissen bisheriger und folgender Studien zu Face Pareidolia auf den Grund zu gehen. Außerdem sollte dem Ursprung der interindividuellen Unterschiede bei Face Pareidolia nachgegangen werden.

Seghier und Price differenzieren in diesem Zusammenhang (2018) intrinsische, kontextuelle und strategische Unterschiede (s. Kapitel 4.2). Je nach Art der individuellen Unterschiede können bestimmte Maßnahmen getroffen werden, um das experimentelle Design anzupassen. So können intrinsische Unterschiede zum Beispiel durch das Matching von Probanden anhand demographischer Variablen (Alter, Geschlecht, Händigkeit etc.) kontrolliert werden. Standardisierte Messprotokolle reduzieren die Wahrscheinlichkeit für kontextuelle Unterschiede. Eine Kontrolle für strategische Unterschiede findet in der Bildgebung hingegen vergleichsweise selten statt. Bezogen auf Studie 2 dieser Dissertation könnten strategische Unterschiede beispielsweise dann auftreten, wenn die Probanden in den Rauschbildern nach unterschiedlichen Mustern, etwa nach Augen oder einer Gesichtssilhouette, gesucht hätten. Andererseits könnten kontextuelle Unterschiede in Form eines Versuchsleitereffekts oder sozial erwünschten Handelns, hervorgerufen durch bestimmte Instruktionen, dazu führen, dass in mehr oder weniger Fällen fälschlicherweise angegeben wird, ein Gesicht zu sehen. Um derartigen Ursachen für interindividuelle Unterschiede auf den Grund zu gehen, könnte die Erhebung beobachtbarer unabhängiger Variablen, die im besten Fall unbewusstes Verhalten abbilden, von Nutzen sein. So gibt es beispielsweise psychologische Skalen, die sozial erwünschtes Antworten sowie verschiedene Gründe für derartiges

Verhalten messen (Larson, 2019). Von einer intrinsischen Natur der Unterschiede wäre bezüglich Studie 2 dann auszugehen, wenn die Auftretenswahrscheinlichkeit der Unterschiede über verschiedene Untersuchungen hinweg konstant bliebe und gleichzeitig andere Ursachen ausgeschlossen werden können.

Insgesamt wird durch Studie 2 betont, dass in der fMRT-Forschung der sorgfältigen Kontrolle und Validierung experimenteller Paradigmen mehr Aufmerksamkeit geschenkt werden sollte. Außerdem zeigen die Ergebnisse von Studie 2, dass es an der Zeit ist, Interpretationen nicht nur anhand von Gruppenstatistiken vorzunehmen, sondern sich vermehrt dem Vorkommen und den Ursprüngen individueller neuraler Prozesse zu widmen.

In Studie 3, der aufwändigsten Studie dieser Dissertation, beschäftigten wir uns mit dem in der fMRT weit verbreiteten Vorgehen, komplexes Stimulusmaterial zu verwenden, ohne potenzielle Einflüsse interindividueller Verteilungen visueller Aufmerksamkeit zu berücksichtigen. Um entgegen der gängigen Praxis für die Effekte natürlichen Blickverhaltens auf neurale Prozesse zu kontrollieren, kombinierten wir sowohl die Aufnahme als auch die Analyse von fMRT- und Eye-tracking-Daten. Unter dieser – vor allem in der Forschung zu sozialen Kognitionen – bislang selten eingesetzten Methodenkombination untersuchten wir ToM Mechanismen bei gesunden Probanden mit unterschiedlichen Ausprägungen im Merkmal Alexithymie. Eine hohe Ausprägung von Alexithymie wurde zuvor einerseits mit verändertem Blickverhalten und andererseits mit veränderten ToM Mechanismen in Verbindung gebracht. Außerdem zeigte eine erst kürzlich veröffentlichte Untersuchung, dass neben frühen visuellen Arealen auch hierarchisch höhere Regionen, inklusive des ToM Netzwerks, differenziert auf Unterschiede im Blickverhalten reagieren. Zunächst konnten wir die Assoziation zwischen atypischem Blickverhalten und Alexithymie bestätigen. Dabei zeigten wir einerseits, dass der Befund auf ToM Aufgaben generalisiert, und andererseits, dass eine atypisch geringe Fokussierung auf Augen bei Alexithymie auch spontan, d.h. in aufgabenfreiem Kontext, auftritt. Durch die Berücksichtigung der Effekte des individuellen Blickverhaltens auf das fMRT-Signal demonstrierten wir zudem, dass Alexithymie nicht per se mit veränderten ToM Prozessen assoziiert ist, wie in einer vorherigen Studie angenommen. Es war die Interaktion aus Alexithymie und Blickverhalten, die die Signale verschiedener Regionen des ToM Netzwerks erklärte. Während bei geringerer Alexithymie das Signal in diesen Regionen mit

längerer Fixationsdauer auf kritische Stimuluselemente zunahm, nahm es bei hoch Alexithymen mit Fixierung derselben Stimuluselemente ab. Diese Zusammenhänge stellten sich in Abhängigkeit von dem fixierten Stimulusaspekt in unterschiedlichen Hirnarealen dar, so dass sie jeweils mit spezifischen funktionellen Verarbeitungsmechanismen in Verbindung gebracht werden konnten.

Studie 3 dieser Dissertation macht darauf aufmerksam, dass es in der fMRT üblich ist, komplexe Paradigmen einzusetzen, den zahlreichen inhärenten Freiheitsgraden dabei aber weder durch Design noch Analyse gerecht zu werden. Die Interpretation von Hirnfunktionen anhand von Gruppenstatistiken setzt im Gegenteil voraus, dass alle Probanden die experimentelle Aufgabe auf die gleiche Art und Weise bearbeiten. So werden im Rahmen der üblichen Gruppenstatistiken interindividuelle Unterschiede, die keine explizite Beachtung in den statistischen Modellen finden, als Rauschen klassifiziert. Da die experimentellen Aufgaben in vielen Bereichen der fMRT-Forschung strategische Unterschiede sowohl auf kognitiver Ebene als auch auf Verhaltensebene erlauben, dürften strategische Unterschiede zwischen Probanden eher die Regel als die Ausnahme sein.

In diesen Fällen wäre es von Vorteil, wenn für die möglichen unterschiedlichen Verhaltensweisen oder kognitiven Strategien kontrolliert werden würde. Da die Ausrichtung visueller Aufmerksamkeit auf verschiedene Aspekte ein und desselben Stimulus mit unterschiedlichen neuralen Prozessen assoziiert ist, sollten interindividuelle Unterschiede im Blickverhalten in fMRT-Studien berücksichtigt werden. Von besonderer Bedeutung wäre dies im Rahmen von Studien, die sich mit Gruppen auseinandersetzen, bei denen atypisches Blickverhalten bereits bekannt ist, wie z.B. bei Untersuchungen zu Autismus-Spektrum-Störungen, Schizophrenie oder Depression (Pelphrey et al., 2002; Toh et al., 2011). Aber auch bei der Untersuchung gesunder Probanden häufen sich die Befunde dazu, dass sich Fixationsmuster zwischen Individuen mit verschiedenen Eigenschaften, wie z.B. Persönlichkeitsmerkmale, voneinander unterscheiden. Auch bei Gesunden lohnt sich demnach die Analyse von Augenbewegungen. Fixationsmuster könnten darüber Auskunft geben, welche Hirnregionen mit der Bearbeitung welcher spezifischen Stimulusaspekte in Verbindung stehen. Dass die Kombination aus fMRT und Eyetracking trotz ihres Potenzials bislang nur wenig Anwendung gefunden hat, liegt zum Teil an den unterschiedlichen zeitlichen Auflösungen beider Methoden (Peitek et al., 2018). In Studie 3 verwenden wir mit der parametrischen Mo-

dulation eine einfache Analysetechnik, die in allen gängigen Standard-Programmen zur Auswertung von fMRT-Daten implementiert ist. Die Studie könnte demnach als Wegbereiter für zukünftige Studien dienen, in denen die verschiedenen Einflüsse natürlichen Blickverhaltens auf das fMRT-Signal berücksichtigt werden sollen.

7 Zusammenfassung

Diese kumulative Dissertation umfasst drei Studien, die Einflüsse von experimentellem Paradigma und interindividuellen Unterschieden auf die Interpretation des resultierenden fMRT-Signals untersuchen. In der ersten Studie verglichen Kollegen und ich verschiedene, in fMRT-Studien eingesetzte, Paradigmen hinsichtlich ihrer Robustheit, ein bestimmtes Kriterium (Rechtslateralisation) zu erfüllen. Das robusteste Paradigma prüften wir weiterhin auf seine Test-Retest-Reliabilität. Die Untersuchungen verdeutlichen, dass sich gängige Paradigmen in der Erfüllung oftmals vorausgesetzter, aber nicht geprüfter, Kriterien (wie z.B. die Rekrutierung rechtslateralisierter Prozesse) unterscheiden. Außerdem kann sich die Reliabilität eines Paradigmas in Abhängigkeit von der jeweiligen Testgröße (z.B. ein einzelner Voxelwert vs. ein zusammengefasstes Maß wie der Lateralisationsindex) grundlegend ändern, was dazu führt, dass der Einsatz eines Paradigmas nur für getestete, umschriebene Zwecke empfohlen werden kann.

In der zweiten Studie betrachteten Kollegen und ich neurale Aktivierungsmuster (im Rahmen von Pareidolie), anders als übliche fMRT-Untersuchungen, auf individueller Ebene. Dabei stellten sich die Unterschiede zwischen Probanden extremer dar als unter Annahme der Gauß'schen Normalverteilung zu erwarten gewesen wäre. Bedeutende interindividuelle Unterschiede können im Rahmen von fMRT-Studien leicht unerkannt bleiben, da die Berichterstattung von Ergebnissen in der Regel lediglich auf der Basis von Gruppenstatistiken erfolgt. Gleichzeitig können diese Unterschiede die Interpretation der Studienergebnisse jedoch maßgeblich beeinflussen. In diesem Zusammenhang, sowie auch im Bezug auf die dritte Studie dieser Dissertation, diskutiere ich sowohl verschiedene Arten von interindividuellen Unterschieden als auch bestimmte Rahmenbedingungen, unter denen eine Berücksichtigung interindividueller Unterschiede in neuronalen Prozessen von besonderer Bedeutung ist.

In der dritten und aufwendigsten Studie dieser Dissertation beschäftigten Kollegen und ich uns mit den Einflüssen von in der fMRT besonders häufig vernachlässigten Verhaltensunterschieden: unter Einsatz von komplexem visuellen Stimulusmaterial ermittelten wir Effekte, die im Zusammenhang mit natürlichem Blickverhalten stehen. Zunächst zeigten wir, dass sich das Blickverhalten gesunder Probanden je nach Ausprägung einer bestimmten Persönlichkeitseigenschaft (Alexithymie) unterschied. Wir integrierten das Blickverhalten (in Form von individuellen Fixationszeiten) in unser Modell zur Analyse der fMRT-Daten und fanden – im Gegensatz zu einer vorherigen Studie – keine Unterschiede in den neuronalen Prozessen der Probanden, die sich durch die Merkmalsausprägung *per se* erklären ließen. Es zeigten sich jedoch Unterschiede, die in Verbindung mit dem Blickverhalten der Probanden standen: je länger Probanden mit geringer Merkmalsausprägung bestimmte Stimulusaspekte fixierten, desto stärker stellte sich das neurale Signal in für den Prozess bedeutenden Hirnarealen dar. Bei den Probanden mit extremer Merkmalsausprägung beobachteten wir einen entgegengesetzten Zusammenhang (längere Fixation, abnehmendes Signal). Die Ergebnisse unserer Studie verdeutlichen den Mehrwert einer Berücksichtigung des Blickverhaltens im Rahmen von fMRT-Studien. Demnach ist die Kontrolle des Blickverhaltens einerseits notwendig, um eine Fehlinterpretation der neuronalen Unterschiede zu vermeiden, die lediglich daraus resultieren, dass die Probanden auf verschiedene Stimulusaspekte achten. Andererseits können unterschiedliche kognitive Strategien identifiziert werden, die verschiedene Individuen bei der Verarbeitung derselben spezifischen Stimulusaspekte einsetzen.

Anhand der beschriebenen drei Studien verdeutlicht meine Dissertation, dass bei fMRT-Studien, insbesondere im Hinblick auf die zum gegenwärtigen Zeitpunkt bestehende Replikationskrise, sowohl die Eignungsprüfung des experimentellen Paradigmas als auch die Berücksichtigung interindividueller Unterschiede erstrebenswert sind. Die Prüfung des Paradigmas kann anhand verschiedener Gütekriterien erfolgen und Empfehlungen sollten unter Beachtung der speziellen Prüfgrößen ausgesprochen werden. Des Weiteren kann eine verstärkte Berücksichtigung der verschiedenen Arten interindividueller Unterschiede perspektivisch zu einer Reduktion nicht aufgeklärter Varianzen führen und somit nicht nur die Replikation von Gruppenergebnissen erleichtern, sondern auch zur Aufklärung individueller neuronaler Verarbeitungsmechanismen beitragen.

8 Referenzen

- Ackermann, J.F., 1806. Die Gall'sche Hirn- Schedel- und Organenlehre: vom Gesichtspunkte der Erfahrung aus beurtheilt und widerlegt. Heidelberg, Frankfurt a.M.
- Ainslie, P.N., Duffin, J., 2009. Integration of cerebrovascular CO₂ reactivity and chemoreflex control of breathing: mechanisms of regulation, measurement, and interpretation. *Am. J. Physiol. Integr. Comp. Physiol.* 296, R1473–R1495. <https://doi.org/10.1152/ajpregu.91008.2008>
- Alexander, A., Barnett-Cowan, M., Bartmess, E., Bosco, F.A., Brandt, M., Carp, J., Chandler, J.J., Clay, R., Cleary, H., Cohn, M., Costantini, G., Decoster, J., Dunn, E., Eggleston, C., Estel, V., Farach, F.J., Feather, J., Fiedler, S., Field, J.G., Foster, J.D., Frank, M., Frazier, R.S., Fuchs, H.M., Galak, J., Galliani, E.M., Garcia, S., Giammanco, E.M., Gilbert, E.A., Giner-Sorolla, R., Goellner, L., Goh, J.X., Justin Goss, R., Graham, J., Grange, J.A., Gray, J.R., Gripshover, S., Hartshorne, J., Hayes, T.B., Jahn, G., Johnson, K., Johnston, W., Joy-Gaba, J.A., Lai, C.K., Lakens, D., Lane, K., Lebel, E.P., Lee, M., Lemm, K., Mackinnon, S., May, M., Moore, K., Motyl, M., Muller, S.M., Munafo, M., Nosek, B.A., Olsson, C., Paunesku, D., Perugini, M., Pitts, M., Ratliff, K., Renkewitz, F., Rutchick, A.M., Sandstrom, G., Saxe, R., Selterman, D., Simpson, W., Smith, C.T., Spies, J.R., Strohminger, N., Talhelm, T., Van‘T Veer, A., Vianello, M., 2012. An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspect. Psychol. Sci.* <https://doi.org/10.1177/1745691612462588>
- Bakker, M., Wicherts, J.M., 2011. The (mis)reporting of statistical results in psychology journals. *Behav. Res. Methods* 43, 666–678. <https://doi.org/10.3758/s13428-011-0089-5>
- Bandettini, P.A., 2012. Twenty years of functional MRI: The science and the stories. *Neuroimage* 62, 575–588. <https://doi.org/10.1016/j.neuroimage.2012.04.026>
- Barch, D.M., Yarkoni, T., 2013. Introduction to the special issue on reliability and replication in cognitive and affective neuroscience research. *Cogn. Affect. Behav. Neurosci.* <https://doi.org/10.3758/s13415-013-0201-7>
- Bear, M.F., Connors, B.W., Paradiso, M.A., 2018. *Neurowissenschaften, Neurowissenschaften*. Springer Berlin Heidelberg, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-662-57263-4>

- Bennett, C.M., Miller, M.B., 2013. fMRI reliability: Influences of task and experimental design. *Cogn. Affect. Behav. Neurosci.* 13, 690–702. <https://doi.org/10.3758/s13415-013-0195-1>
- Bloede, K.A., 1807. Dr. F.J. Galls neue entdeckungen in der Gehirn-, Schedel- und Organenlehre. Müllersche Verlagshandlung, Karlsruhe.
- Boubela, R.N., Kalcher, K., Huf, W., Seidel, E.-M., Derntl, B., Pezawas, L., Našel, C., Moser, E., 2015. fMRI measurements of amygdala activation are confounded by stimulus correlated signal fluctuation in nearby veins draining distant brain regions. *Sci. Rep.* 5, 10499. <https://doi.org/10.1038/srep10499>
- Brembs, B., Button, K., Munafò, M., 2013. Deep impact: unintended consequences of journal rank. *Front. Hum. Neurosci.* 7. <https://doi.org/10.3389/fnhum.2013.00291>
- Brennan, N.P., Peck, K.K., Holodny, A., 2016. Language mapping using fMRI and direct cortical stimulation for brain tumor surgery the good, the bad, and the questionable. *Top. Magn. Reson. Imaging.* <https://doi.org/10.1097/RMR.0000000000000074>
- Broca, P., 1861. Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). *Bull. la Société Anat. Paris* 6, 330–357.
- Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J., Munafò, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* 14, 365–376. <https://doi.org/10.1038/nrn3475>
- Caceres, A., Hall, D.L., Zelaya, F.O., Williams, S.C.R., Mehta, M.A., 2009. Measuring fMRI reliability with the intra-class correlation coefficient. *Neuroimage.* <https://doi.org/10.1016/j.neuroimage.2008.12.035>
- Dalton, K.M., Holsen, L., Abbeduto, L., Davidson, R.J., 2008. Brain function and gaze fixation during facial-emotion processing in fragile X and autism. *Autism Res.* <https://doi.org/10.1002/aur.32>
- Dalton, K.M., Nacewicz, B.M., Alexander, A.L., Davidson, R.J., 2007. Gaze-Fixation, Brain Activation, and Amygdala Volume in Unaffected Siblings of Individuals with Autism. *Biol. Psychiatry* 61, 512–520. <https://doi.org/10.1016/j.biopsych.2006.05.019>
- Dalton, K.M., Nacewicz, B.M., Johnstone, T., Schaefer, H.S., Gernsbacher, M.A., Goldsmith, H.H., Alexander, A.L., Davidson, R.J., 2005. Gaze fixation and

- the neural circuitry of face processing in autism. *Nat. Neurosci.* 8, 519–526. <https://doi.org/10.1038/nm1421>
- de Haas, B., Iakovidis, A.L., Schwarzkopf, D.S., Gegenfurtner, K.R., 2019. Individual differences in visual salience vary along semantic dimensions. *Proc. Natl. Acad. Sci.* 201820553. <https://doi.org/10.1073/pnas.1820553116>
- Deppe, M., Knecht, S., Papke, K., Lohmann, H., Fleischer, H., Heindel, W., Ringelstein, E.B., Henningsen, H., 2000. Assessment of hemispheric language lateralization: A comparison between fMRI and fTCD. *J. Cereb. Blood Flow Metab.* <https://doi.org/10.1097/00004647-200002000-00006>
- Dorst, J., Haag, A., Knake, S., Oertel, W.H., Hamer, H.M., Rosenow, F., 2008. Functional transcranial Doppler sonography and a spatial orientation paradigm identify the non-dominant hemisphere. *Brain Cogn.* <https://doi.org/10.1016/j.bandc.2008.02.123>
- Dotterer, H.L., Hyde, L.W., Swartz, J.R., Hariri, A.R., Williamson, D.E., 2017. Amygdala reactivity predicts adolescent antisocial behavior but not callous-unemotional traits. *Dev. Cogn. Neurosci.* 24, 84–92. <https://doi.org/10.1016/j.dcn.2017.02.008>
- Duchaine, B., Yovel, G., 2015. A Revised Neural Framework for Face Processing. *Annu. Rev. Vis. Sci.* 1, 393–416. <https://doi.org/10.1146/annurev-vision-082114-035518>
- Fink, G.R., Marshall, J.C., Shah, N.J., Weiss, P.H., Halligan, P.W., Grosse-Ruyken, M., Ziemons, K., Zilles, K., Freund, H.J., 2000. Line bisection judgments implicate right parietal cortex and cerebellum as assessed by fMRI. *Neurology.* <https://doi.org/10.1212/WNL.54.6.1324>
- Fliessbach, K., Rohe, T., Linder, N.S., Trautner, P., Elger, C.E., Weber, B., 2010. Retest reliability of reward-related BOLD signals. *Neuroimage.* <https://doi.org/10.1016/j.neuroimage.2010.01.036>
- Frässle, S., Paulus, F.M., Krach, S., Jansen, A., 2016. Test-retest reliability of effective connectivity in the face perception network. *Hum. Brain Mapp.* 37, 730–744. <https://doi.org/10.1002/hbm.23061>
- Gaebler, M., Paschke, L.M., Magrabi, A., 2016. Neurowissenschaft, in: *Handbuch Handlungstheorie.* J.B. Metzler, Stuttgart, pp. 414–421. https://doi.org/10.1007/978-3-476-05359-6_47
- Gamer, M., Büchel, C., 2009. Amygdala activation predicts gaze toward fearful eyes. *J. Neurosci.* 29, 9123–9126. <https://doi.org/10.1523/JNEUROSCI.1883-09.2009>

- Gamer, M., Zurowski, B., Buchel, C., 2010. Different amygdala subregions mediate valence-related and attentional effects of oxytocin in humans. *Proc. Natl. Acad. Sci.* 107, 9400–9405. <https://doi.org/10.1073/pnas.1000985107>
- Giraudeau, B., 1996. Negative values of the intraclass correlation coefficient are not theoretically possible. *J. Clin. Epidemiol.* 49, 1205. [https://doi.org/10.1016/0895-4356\(96\)00053-4](https://doi.org/10.1016/0895-4356(96)00053-4)
- Gottlieb, J., Oudeyer, P.Y., 2018. Towards a neuroscience of active sampling and curiosity. *Nat. Rev. Neurosci.* <https://doi.org/10.1038/s41583-018-0078-0>
- Grinvald, A., Hildesheim, R., 2004. VSDI: A new era in functional imaging of cortical dynamics. *Nat. Rev. Neurosci.* <https://doi.org/10.1038/nrn1536>
- Hadjikhani, N., Åsberg Johnels, J., Zürcher, N.R., Lassalle, A., Guillon, Q., Hippolyte, L., Billstedt, E., Ward, N., Lemonnier, E., Gillberg, C., 2017a. Look me in the eyes: constraining gaze in the eye-region provokes abnormally high subcortical activation in autism. *Sci. Rep.* 7, 3163. <https://doi.org/10.1038/s41598-017-03378-5>
- Hadjikhani, N., Zurcher, N.R., Lassalle, A., Hippolyte, L., Ward, N., Johnels, J.Å., 2017b. The effect of constraining eye-contact during dynamic emotional face perception—an fMRI study. *Soc. Cogn. Affect. Neurosci.* 12, 1197–1207. <https://doi.org/10.1093/scan/nsx046>
- Hariri, A.R., Tessitore, A., Mattay, V.S., Fera, F., Weinberger, D.R., 2002. The amygdala response to emotional stimuli: A comparison of faces and scenes. *Neuroimage.* <https://doi.org/10.1006/nimg.2002.1179>
- Haxby, J. V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. *Trends Cogn. Sci.* [https://doi.org/10.1016/S1364-6613\(00\)01482-0](https://doi.org/10.1016/S1364-6613(00)01482-0)
- Henderson, J.M., 2017. Gaze Control as Prediction. *Trends Cogn. Sci.* 21, 15–23. <https://doi.org/10.1016/j.tics.2016.11.003>
- Himme, A., 2009. Gütekriterien der Messung: Reliabilität, Validität und Generalisierbarkeit, in: *Methodik Der Empirischen Forschung.* https://doi.org/10.1007/978-3-322-96406-9_31
- Huettel, a. S., Song, A.W., McCarthy, G., 2004. *Functional Magnetic Resonance Imaging.* Sunderland, MA: Sinauer Associates.
- Ioannidis, J.P.A., 2011. An epidemic of false claims. Competition and conflicts of interest distort too many medical findings. *Sci. Am.* <https://doi.org/10.1038/scientificamerican0611-16>

- Jansen, A., Flöel, A., Deppe, M., van Randenborgh, J., Dräger, B., Kanowski, M., Knecht, S., 2004. Determining the hemispheric dominance of spatial attention: A comparison between fTCD and fMRI. *Hum. Brain Mapp.* 23, 168–180. <https://doi.org/10.1002/hbm.20055>
- Jansen, A., Lohmann, H., Scharfe, S., Sehlmeier, C., Deppe, M., Knecht, S., 2007. The association between scalp hair-whorl direction, handedness and hemispheric language dominance: Is there a common genetic basis of lateralization? *Neuroimage*. <https://doi.org/10.1016/j.neuroimage.2006.12.025>
- Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J. Neurosci.* <https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997>
- Kliemann, D., Dziobek, I., Hatri, A., Baudewig, J., Heekeren, H.R., 2012. The Role of the Amygdala in Atypical Gaze on Emotional Faces in Autism Spectrum Disorders. *J. Neurosci.* 32, 9469–9476. <https://doi.org/10.1523/JNEUROSCI.5294-11.2012>
- Knecht, S., Deppe, M., Dräger, B., Bobe, L., Lohmann, H., Ringelstein, E.-B., Henningsen, H., 2000a. Language lateralization in healthy right-handers. *Brain* 123, 74–81. <https://doi.org/10.1093/brain/123.1.74>
- Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., Ringelstein, E.B., Henningsen, H., 2000b. Handedness and hemispheric language dominance in healthy humans. *Brain*. <https://doi.org/10.1093/brain/123.12.2512>
- Krings, T., M.H.T., R., S., E., S., K., V., R., U., S., M., K., K., W., J.M., G., 2001. Functional MRI for presurgical planning: problems, artefacts, and solution strategies. *J. Neurol. Neurosurg. Psychiatry* 70, 749–760. <https://doi.org/10.1136/jnnp.70.6.749>
- Kristo, G., Rutten, G.J., Raemaekers, M., de Gelder, B., Rombouts, S.A.R.B., Ramsey, N.F., 2014. Task and task-free fMRI reproducibility comparison for motor network identification. *Hum. Brain Mapp.* <https://doi.org/10.1002/hbm.22180>
- Kujala, M. V., Carlson, S., Hari, R., 2012. Engagement of amygdala in third-person view of face-to-face interaction. *Hum. Brain Mapp.* 33, 1753–1762. <https://doi.org/10.1002/hbm.21317>
- Larson, R.B., 2019. Controlling social desirability bias. *Int. J. Mark. Res.* 61, 534–547. <https://doi.org/10.1177/1470785318805305>
- Lassalle, A., Åsberg Johnels, J., Zürcher, N.R., Hippolyte, L., Billstedt, E., Ward, N., Lemonnier, E., Gillberg, C., Hadjikhani, N., 2017. Hypersensitivity to

- low intensity fearful faces in autism when fixation is constrained to the eyes. *Hum. Brain Mapp.* 38, 5943–5957. <https://doi.org/10.1002/hbm.23800>
- Lebreton, M., Bavard, S., Daunizeau, J., Palminteri, S., 2019. Assessing inter-individual differences with task-related functional neuroimaging. *Nat. Hum. Behav.* <https://doi.org/10.1038/s41562-019-0681-8>
- Li, J., Liu, J., Liang, J., Zhang, H., Zhao, J., Rieth, C.A., Huber, D.E., Li, W., Shi, G., Ai, L., Tian, J., Lee, K., 2010. Effective connectivities of cortical regions for top-down face processing: A Dynamic Causal Modeling study. *Brain Res.* 1340, 40–51. <https://doi.org/10.1016/j.brainres.2010.04.044>
- Liu, J., Li, J., Feng, L., Li, L., Tian, J., Lee, K., 2014. Seeing Jesus in toast: Neural and behavioral correlates of face pareidolia. *Cortex.* <https://doi.org/10.1016/j.cortex.2014.01.013>
- Logothetis, N.K., 2008. What we can do and what we cannot do with fMRI. *Nature.* <https://doi.org/10.1038/nature06976>
- Menon, R.S., 2012. The great brain versus vein debate. *Neuroimage* 62, 970–974. <https://doi.org/10.1016/j.neuroimage.2011.09.005>
- Moriguchi, Y., Ohnishi, T., Lane, R.D., Maeda, M., Mori, T., Nemoto, K., Matsuda, H., Komaki, G., 2006. Impaired self-awareness and theory of mind: An fMRI study of mentalizing in alexithymia. *Neuroimage* 32, 1472–1482. <https://doi.org/10.1016/j.neuroimage.2006.04.186>
- Morris, J.P., Green, S.R., Marion, B., McCarthy, G., 2008. Guided saccades modulate face- and body-sensitive activation in the occipitotemporal cortex during social perception. *Brain Cogn.* 67, 254–263. <https://doi.org/10.1016/j.bandc.2008.01.011>
- Morris, J.P., Pelphey, K.A., McCarthy, G., 2007. Controlled scanpath variation alters fusiform face activation. *Soc. Cogn. Affect. Neurosci.* 2, 31–38. <https://doi.org/10.1093/scan/nsl023>
- Oeser, E., 2002. *Die Geschichte der Hirnforschung. Von der Antike bis zur Gegenwart.* Wissenschaftliche Buchgesellschaft, Darmstadt.
- Open Science Collobaration, 2015. Estimating the reproducibility of psychological science. *Science*, 349, aac4716–aac4716. <https://doi.org/10.1126/science.aac4716>
- Parrish, T.B., Gitelman, D.R., LaBar, K.S., Mesulam, M.M., 2000. Impact of signal-to-noise on functional MRI. *Magn. Reson. Med.* [https://doi.org/10.1002/1522-2594\(200012\)44:6<925::AID-MRM14>3.0.CO;2-M](https://doi.org/10.1002/1522-2594(200012)44:6<925::AID-MRM14>3.0.CO;2-M)

- Peiffer, J., 2004. Hirnforschung in Deutschland 1849–1974: Briefe zur Entwicklung von Psychiatrie und Neurowissenschaften sowie zum Einfluss des politischen Umfeldes auf Wissenschaftler, in: *Schriften Der Mathematisch-Naturwissenschaftlichen Klasse Der Heidelberger Akademie Der Wissenschaften*. Springer, Berlin.
- Peitek, N., Siegmund, J., Parnin, C., Apel, S., Brechmann, A., 2018. Toward joint analysis of simultaneous eye-tracking and fMRI data for program-comprehension studies, in: *Proceedings of the Workshop on Eye Movements in Programming - EMIP '18*. ACM Press, New York, New York, USA, pp. 1–5. <https://doi.org/10.1145/3216723.3216725>
- Pelphrey, K.A., Sasson, N.J., Reznick, J.S., Paul, G., Goldman, B.D., Piven, J., 2002. Visual Scanning of Faces in Autism. *J. Autism Dev. Disord.* <https://doi.org/10.1023/A:1016374617369>
- Perlman, S.B., Hudac, C.M., Pegors, T., Minshew, N.J., Pelphrey, K.A., 2011. Experimental manipulation of face-evoked activity in the fusiform gyrus of individuals with autism. *Soc. Neurosci.* 6, 22–30. <https://doi.org/10.1080/17470911003683185>
- Poldrack, R.A., Baker, C.I., Durnez, J., Gorgolewski, K.J., Matthews, P.M., Munafò, M.R., Nichols, T.E., Poline, J.-B., Vul, E., Yarkoni, T., 2017. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat. Rev. Neurosci.* 18, 115–126. <https://doi.org/10.1038/nrn.2016.167>
- Risko, E.F., Anderson, N.C., Lanthier, S., Kingstone, A., 2012. Curious eyes: Individual differences in personality predict eye movement behavior in scene-viewing. *Cognition* 122, 86–90. <https://doi.org/10.1016/j.cognition.2011.08.014>
- Schomaker, J., Walper, D., Wittmann, B.C., Einhäuser, W., 2017. Attention in natural scenes: Affective-motivational factors guide gaze independently of visual salience. *Vision Res.* <https://doi.org/10.1016/j.visres.2017.02.003>
- Schott, H., 2002. Schädel, Hirn und Seele – Ursprung der modernen Neurowissenschaft. *Dtsch. Arztebl.* 99, A 1420–1422.
- Schuster, V., Herholz, P., Zimmermann, K.M., Westermann, S., Frässle, S., Jansen, A., 2017. Comparison of fMRI paradigms assessing visuospatial processing: Robustness and reproducibility. *PLoS One* 12, e0186344. <https://doi.org/10.1371/journal.pone.0186344>

- Seghier, M.L., Price, C.J., 2018. Interpreting and Utilising Intersubject Variability in Brain Function. *Trends Cogn. Sci.* 22, 517–530. <https://doi.org/10.1016/j.tics.2018.03.003>
- Sejnowski, T.J., Churchland, P.S., Movshon, J.A., 2014. Putting big data to good use in neuroscience. *Nat. Neurosci.* 17, 1440–1441. <https://doi.org/10.1038/nn.3839>
- Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: Uses in assessing rater reliability. *Psychol. Bull.* 86, 420–428. <https://doi.org/10.1037/0033-2909.86.2.420>
- Shrout, P.E., Rodgers, J.L., 2018. Psychology, Science, and Knowledge Construction: Broadening Perspectives from the Replication Crisis. *Annu. Rev. Psychol.* <https://doi.org/10.1146/annurev-psych-122216-011845>
- Simmons, J.P., Nelson, L.D., Simonsohn, U., 2011. False-Positive Psychology. *Psychol. Sci.* 22, 1359–1366. <https://doi.org/10.1177/0956797611417632>
- Smith, M.L., Gosselin, F., Schyns, P.G., 2012. Measuring Internal Representations from Behavioral and Brain Data. *Curr. Biol.* 22, 191–196. <https://doi.org/10.1016/j.cub.2011.11.061>
- Spilka, M.J., Pittman, D.J., Bray, S.L., Goghari, V.M., 2019. Manipulating visual scanpaths during facial emotion perception modulates functional brain activation in schizophrenia patients and controls. *J. Abnorm. Psychol.* 128, 855–866. <https://doi.org/10.1037/abn0000468>
- Springer, J.A., Binder, J.R., Hammeke, T.A., Swanson, S.J., Frost, J.A., Bellgowan, P.S.F., Brewer, C.C., Perry, H.M., Morris, G.L., Mueller, W.M., 1999. Language dominance in neurologically normal and epilepsy subjects. *Brain* 122, 2033–2046. <https://doi.org/10.1093/brain/122.11.2033>
- Stephan, K.E., Penny, W.D., Moran, R.J., den Ouden, H.E.M., Daunizeau, J., Friston, K.J., 2010. Ten simple rules for dynamic causal modeling. *Neuroimage.* <https://doi.org/10.1016/j.neuroimage.2009.11.015>
- Toh, W.L., Rossell, S.L., Castle, D.J., 2011. Current visual scanpath research: a review of investigations into the psychotic, anxiety, and mood disorders. *Compr. Psychiatry* 52, 567–579. <https://doi.org/10.1016/j.comppsy.2010.12.005>
- Tomasello, M., Carpenter, M., Call, J., Behne, T., Moll, H., 2005. In search of the uniquely human. *Behav. Brain Sci.* <https://doi.org/10.1017/S0140525X05540123>
- Tottenham, N., Hertzog, M.E., Gillespie-Lynch, K., Gilhooly, T., Millner, A.J., Casey, B.J., 2014. Elevated amygdala response to faces and gaze aversion in au-

- tism spectrum disorder. *Soc. Cogn. Affect. Neurosci.* 9, 106–117. <https://doi.org/10.1093/scan/nst050>
- Wang, L., Baumgartner, F., Kaule, F.R., Hanke, M., Pollmann, S., 2019. Individual face- and house-related eye movement patterns distinctively activate FFA and PPA. *Nat. Commun.* <https://doi.org/10.1038/s41467-019-13541-3>
- Wei, X., Yoo, S.-S., Dickey, C.C., Zou, K.H., Guttman, C.R.G., Panych, L.P., 2004. Functional MRI of auditory verbal working memory: long-term reproducibility analysis. *Neuroimage* 21, 1000–1008. <https://doi.org/10.1016/j.neuroimage.2003.10.039>
- Whitehouse, A.J.O., Badcock, N., Groen, M.A., Bishop, D.V.M., 2009. Reliability of a novel paradigm for determining hemispheric lateralization of visuospatial function. *J. Int. Neuropsychol. Soc.* 15, 1028–1032. <https://doi.org/10.1017/S1355617709990555>
- Wilke, M., Schmithorst, V.J., 2006. A combined bootstrap/histogram analysis approach for computing a lateralization index from neuroimaging data. *Neuroimage.* <https://doi.org/10.1016/j.neuroimage.2006.07.010>
- Zhang, H., Liu, J., Huber, D.E., Rieth, C.A., Tian, J., Lee, K., 2008. Detecting faces in pure noise images: A functional MRI study on top-down perception. *Neuroreport.* <https://doi.org/10.1097/WNR.0b013e3282f49083>
- Zilles, K., Amunts, K., 2013. Individual variability is not noise. *Trends Cogn. Sci.* 17, 153–155. <https://doi.org/10.1016/j.tics.2013.02.003>
- Zimmermann, K.M., Stratil, A.-S., Thome, I., Sommer, J., Jansen, A., 2019. Illusory face detection in pure noise images: The role of interindividual variability in fMRI activation patterns. *PLoS One* 14, e0209310. <https://doi.org/10.1371/journal.pone.0209310>
- Zürcher, N.R., Donnelly, N., Rogier, O., Russo, B., Hippolyte, L., Hadwin, J., Lemonnier, E., Hadjikhani, N., 2013. It's All in the Eyes: Subcortical and Cortical Activation during Grotesqueness Perception in Autism. *PLoS One* 8, e54313. <https://doi.org/10.1371/journal.pone.0054313>