

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

**Test-Retest-Reliabilität von fMRT-Paradigmen:
Wortgenerierung und Subliminale Präsentation emotionaler
Stimuli**

Im Rahmen der Studie

*"Wirkungen von genetischen Risikovarianten für psychische Störungen auf Gehirnstruktur
und -funktion bei Gesunden und Erkrankten"*

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

Victoria Krohn aus Hofgeismar

Marburg, 2018

Angenommen vom Fachbereich Medizin der Philipps-Universität Marburg am:

Gedruckt mit Genehmigung des Fachbereichs.

Dekan: Herr Prof. Dr. H. Schäfer

Referent: Herr Prof. Dr. A. Jansen

1.Korreferent: Frau Prof. Dr. S. Knake

INHALTSVERZEICHNIS

1 ZUSAMMENFASSUNG	1
1 ABSTRACT.....	3
2 EINLEITUNG.....	6
2.1 RELIABILITÄT	6
2.2 WORTGENERIERUNG (WG).....	7
2.3 SUBLIMINALE PRÄSENTATION EMOTIONALER STIMULI (SPES).....	11
3 METHODIK.....	17
3.1 METHODISCHE GRUNDLAGEN	17
3.1.1 <i>Magnetresonanztomographie</i>	17
3.1.2 <i>Funktionelle Magnetresonanztomographie</i>	25
3.1.3 <i>Analyse von fMRT-Daten</i>	26
<i>Statistische Inferenz</i>	30
3.1.4 <i>Design eines fMRT-Experiments</i>	30
3.2 PROBAND_INNENDATEN	32
3.3 FMRT-PARAMETER	33
3.4 EXPERIMENTELLES DESIGN	34
3.4.1 <i>Wortgenerierungs- Aufgabe</i>	35
3.4.2 <i>Aufgabe der subliminalen Präsentation emotionaler Stimuli (SPES-Aufgabe)</i>	37
3.5 FUNKTIONELLE DATENAUSWERTUNG	40
3.5.1 <i>Vorverarbeitung</i>	40
3.5.2 <i>Statistische Analyse und Inferenz</i>	40
3.5.3 <i>Lokalisierung aktivierter Hirnregionen</i>	41
3.6 MESSUNG DER RELIABILITÄT	41
3.6.1 <i>Methoden der Reliabilitätsmessung</i>	41
3.6.2 <i>Vorgehensweise der Reliabilitätsmessung in dieser Arbeit</i>	46
4 ERGEBNISSE.....	49
4.1 ERGEBNISSE DER WG-AUFGABE.....	49
4.1.1 <i>Ergebnisse der Gruppenanalyse</i>	50
4.1.2 <i>Ergebnisse der Reliabilitätsmessung</i>	53
4.2 ERGEBNISSE DER SPES-AUFGABE	56
4.2.1 <i>Ergebnisse der Gruppenanalyse</i>	57

4.2.2 Ergebnisse der Reliabilitätsanalyse.....	64
5 DISKUSSION.....	72
5.1 ZUSAMMENFASSUNG	72
5.2 METHODENKRITIK.....	73
5.2.1 Experimentelles Design	73
5.2.2 Experimentelle Bedingungen	75
5.2.3 Technische Bedingungen.....	77
5.2.3.1 Datenaquisition	77
5.2.3.2 Datenanalyse	79
5.3 ERGEBNISDISKUSSION.....	82
5.3.1 WG-Aufgabe	82
5.3.1.1 Bewertung der Gruppenanalyse.....	82
5.3.1.2 Bewertung der Reliabilitätsanalyse	84
5.3.1.3 Bewertung des Paradigmen- designs	85
5.3.2 SPES-Aufgabe.....	87
5.3.2.1 Bewertung der Gruppenanalyse.....	87
5.3.2.2 Bewertung der Reliabilitätsanalyse	91
5.3.2.3 Bewertung des Paradigmen- designs	93
6 FAZIT	96
6 LITERATURVERZEICHNIS.....	98
7 ABBILDUNGSVERZEICHNIS	105
8 TABELLENVERZEICHNIS	107
9 ABKÜRZUNGSVERZEICHNIS.....	108
10 ANHANG.....	110
10.1 PROBAND_INNEN- INSTRUKTION.....	110
10.2 PROBAND_INNEN- AUFKLÄRUNGSBOGEN	113
11 AKADEMISCHE LEHRER	122

1 ZUSAMMENFASSUNG

Hintergrund

Die funktionelle Magnetresonanztomographie hat seit ihrem Einsatz in der medizinischen Diagnostik Anfang der 1990er Jahre die Forschungslandschaft der Psychiatrie und Neurologie grundlegend verändert. In rasanter Geschwindigkeit sind seither Annahmen über die neurologischen Korrelate der menschlichen kognitiven Fähigkeiten untersucht und die Übertragung der Erkenntnisse in therapeutische Bereiche in den Rahmen des Möglichen gerückt worden. Wissenschaftliche Standards für die Messung und Interpretation der Ergebnisse hinken dieser Entwicklung allerdings noch hinterher (Bennet und Miller, 2010). Das Ziel der vorliegenden Arbeit ist es demzufolge, die Reliabilität der fMRT-Messung unterschiedlicher kognitiver Aufgabenbereiche zu untersuchen.

Die Paradigmen dieser Arbeit werden im Kontext der Studie *„Erfassung von Wirkungen genetischer Risikovarianten für psychische Störungen auf Gehirnstruktur und –funktion bei Gesunden und Erkrankten“* verwendet. In dieser Studie wird unter anderem die neuronale Aktivität und funktionelle Konnektivität mittels fMRT-Paradigmen im zeitlichen Verlauf bei einem gesunden und erkrankten Probandenkollektiv untersucht und auf Risikovarianten überprüft. Da es sich hierbei um eine Longitudinalstudie handelt, ist es essentiell, dass die gewonnenen fMRT-Daten der jeweiligen Paradigmen keine zufälligen Messfehler darstellen und möglichst über einen längeren Zeitraum vergleichbar sind.

Methoden

In einem Test-Retest-Verfahren wurden mittels fMRT-Untersuchung 20 gesunde Proband_innen im Abstand von $28,4 \pm 6,0$ Tagen gemessen. Aus den gewonnenen Daten der beiden Messungen folgte zunächst eine qualitative Überprüfung der Überlappung der Aktivierungsareale. Im Anschluss wurde eine statistische Rechenanalyse mithilfe des sogenannten ICC-Wertes (*Intra-Class Correlation Coefficient*) durchgeführt. Darauf basierend fand eine Bewertung der Zuverlässigkeit der Aktivierungswerte hinsichtlich ihrer Reproduzierbarkeit statt.

Die Messungen überprüften insgesamt fünf verschiedene Domänen kognitiver Hirnfunktionen in sechs Paradigmen: das Ruhenetzwerk (*Resting State*), die semantische Wortgenerierung (*Verbal Fluency*, WG), das Einspeichern und Abrufen innerhalb des episodischen Gedächtnisses (*Encoding und Retrieval*), die Verarbeitung subliminal präsentierter emotionaler Stimuli (*Subprime*, SPES) und das Arbeitsgedächtnis (*n-back*). In dieser Arbeit wurde schwerpunktmäßig die Reliabilität des WG und des SPES- Paradigmas untersucht.

Ergebnisse

Sowohl die WG- als auch die SPES-Aufgabe zeigten überwiegend erwartungsgemäß aktivierte Hirnregionen, welche sich in der Auswertung der qualitativen Darstellung mittels Cluster-Überlappungsmethode und der Rechenanalyse jedoch erheblich unterschieden. Die WG-Aufgabe erzeugte deutlich mehr Aktivierungen während der Kontrollphase. In der Cluster-Überlappungsmethode überlappten sich hauptsächlich die Deaktivierungsareale und statistisch bestätigten die errechneten Reliabilitäten eine insgesamt schlechte Reliabilität des Paradigmas.

Hinsichtlich der SPES-Aufgabe zeigten sich in der Auswertung die charakteristischen Aktivierungsregionen der Gesichtsverarbeitung und auch in der Reliabilitätsanalyse für das aktivierte und das deaktivierte Netzwerk verhältnismäßig hohe mediane ICC-Werte im moderaten bis guten Reliabilitätsbereich. Die Überprüfung der subliminal stimulierten Emotionsverarbeitung am Beispiel der Emotion Angst wies demhingegen nur geringe Aktivitäten und für die charakteristischen Regionen eine niedrige Reliabilität auf. Zusammenfassend gilt die Reliabilität der SPES-Aufgabe somit als moderat bis gut zu beurteilen.

Diskussion

In der Beurteilung der Ergebnisse müssen viele Aspekte Beachtung finden, um die Frage der Übertragbarkeit auf Messungen der Rahmenstudie und auch andere Probandenkollektive beantworten zu können. Zunächst ist die Varianz der erhobenen fMRT-Daten wesentlich von experimentellen und technischen Bedingungen sowie dem experimentellen Design abhängig (Bennet und Miller, 2013). Trotz guter technischer Bedingungen zeigten sich in dieser Studie experimentelle Defizite hinsichtlich des Probandenkollektivs. Zudem wirkten sich

verschiedene Mängel im experimentellen Design beider Paradigmen auf die geringe Reliabilität aus. Als weiterer Aspekt muss die Reliabilitätsanalyse anhand der ICC-Werte hinterfragt werden. Sie entspricht zwar dem aktuellen Stand der Reliabilitätsüberprüfung, weist jedoch ebenfalls große Defizite in der Übertragung auf andere Probandenkollektive auf (Fliessbach et al., 2010).

Fazit

Aufgrund der geringen Reliabilität der in dieser Arbeit untersuchten Paradigmen müssen Ergebnisse im Kontext anderer Studien, wie beispielsweise der Rahmenstudie, mit großer Vorsicht interpretiert werden. Hinsichtlich zukünftiger Forschungserkenntnisse mittels fMRT-Daten wäre es notwendig, einen übertragbaren Standard in der Reliabilitätsmessung festzulegen, um vor allem für zukünftige diagnostische und therapeutische Anwendungen zuverlässige Grundlagen zur Verfügung zu stellen.

1 ABSTRACT

Background

Since the early 1990s the implementation of functional magnetic resonance imaging (fMRI) in the field of medical diagnostics has fundamentally changed the scientific landscape of psychiatry and neurology. Ever since, assumptions about the neurological correlates of human cognitive capabilities evolved quickly. Soon, the transfer of that knowledge to therapeutic means approached realistic dimensions. Yet, scientific standards for the measurement and interpretation of those results have not been developed accordingly (Bennet and Miller, 2010). It is against this background that this study aims to examine the reliability of the fMRI measurement in different cognitive fields.

The paradigms of the thesis at hand are situated within the broader context of the study *“Effects of genetic risk variants for psychiatric disorders on brain structure and function in healthy and affected subjects”*. Using fMRI paradigms in the temporal processes of healthy and affected individuals, this study evaluates, amongst others, neuronal activity and functional connectivity with a view to risk variants. Being longitudinal in nature, it is

essential that this study be based on reliable and comparable fMRI data, which is collected over a long period of time.

Methods

In a test-retest-procedure, applied across an interval of $28,4 \pm 6,0$ days, 20 healthy individuals were examined, using the fMRI measurement technique. The data of both measurements was then used to qualitatively review the overlap of the activated areas. Subsequently, with the aid of the so called ICC (*Intra-class Correlation Coefficient*), this study ran a statistic calculation analysis. Thereafter, based on the ICC values, the reliability of the activation coefficients was assessed in view of their reproducibility.

Overall, the measurements reviewed five different domains of cognitive brain functions by six paradigms: the activation network at rest (*Resting State*); semantic word generation (*Verbal Fluency, WG*); the episodic memory (*Encoding and Retrieval*); subliminal presentation of emotional stimuli (*Subprime, SPES*); and the working memory (*n-back*). The thesis focused on examining the reliability of the WG-paradigm as well as the SPES-paradigm.

Results

Both the WG-paradigm as well as the SPES-paradigm mainly showed activations in expected brain areas. However the evaluation of the qualitative presentation by the cluster-overlap method and the calculation analyses showed considerable differences. The WG-paradigm generated considerably more activation during the control period of the experiment. In the cluster overlap method mainly the deactivated areas interleaved. Statistically, the calculated reliabilities confirmed the overall deficient reliability of the paradigm.

As for the SPES-paradigm, the analysis showed characteristic activations in the processing of faces. The analysis of reliabilities of the activated and deactivated network showed comparatively high median ICC values in the range of moderate to good reliability. To the contrary, the examination of subliminal stimulated processes of emotions with the example of the emotion fear only showed low activities and in characteristic regions a low reliability. In summary, it can be stated that the reliability of the SPES paradigm can be judged moderate and good.

Discussion

When seeking to examine the transferability of the generated results in light of the overall contextual study mentioned above as well as other groups of subjects, various aspects need to be considered: At first, it needs to be clarified that the variance of the acquired fMRI data depends essentially on experimental and technical conditions as well as the experimental design (Bennet und Miller, 2013). Despite good technical conditions, experimental deficits regarding the group of subjects emerged in this study. Furthermore, various deficiencies in the experimental design of both paradigms affected the low reliability. At the same time, the aspect of analysing reliability by ICC values has to be questioned. Even though it represents the current state and latest practice in the evaluation of reliability, it still shows considerable deficits in the transfer to different groups of subjects (Fliessbach et al., 2010).

Conclusion

With a view to other studies, the results of the investigated paradigms have to be interpreted with caution. In regard of future scientific findings that are based on fMRI data it will be essential to standardize the measurement of reliability. The aim should be to develop and systematize a solid basis for prospective diagnostic and therapeutic implementation.

2 EINLEITUNG

Dieses Kapitel führt zu allererst in den Begriff der *Reliabilität* ein. Allgemein werden die Relevanz der *Reliabilität* bei wissenschaftlichen Testverfahren, seine unterschiedlichen Messmethoden und seine Bedeutung bei fMRT-Studien und speziell der Rahmenstudie dieser Arbeit vorgestellt. Im Anschluss daran werden die biopsychologischen, anatomischen und wissenschaftlichen Grundlagen der beiden kognitiven Domänen der Wortgenerierung und Verarbeitung emotionaler Stimuli erläutert, welche in dieser Arbeit schwerpunktmäßig auf ihre *Reliabilität* untersucht wurden. Schließlich wird auf die Zielsetzung dieser Arbeit eingegangen.

2.1 Reliabilität

Die Reliabilität ist bei der Datenerhebung neben der Objektivität und Validität eines der drei wichtigsten Gütekriterien eines Tests (Lienert und Raatz, 1998). Sie repräsentiert das Zuverlässigkeitskriterium und misst den Grad an formaler Genauigkeit einer wissenschaftlichen Messmethode. Würden lediglich die Testgütekriterien der Validität und Objektivität erfüllt und darüber hinaus signifikante Ergebnisse erhoben werden, wären diese bei mangelnder Reliabilität trotz allem bedeutungslos (Bennet und Miller, 2010). Reliable Testergebnisse sollten möglichst frei von zufällig erhobenen Fehlern und Störeinflüssen und unter den gleichen Rahmenbedingungen mit der gleichen Messung replizierbar sein. Die Reliabilitätsmessung leitet sich von der Klassischen Testtheorie ab, in welcher davon ausgegangen wird, dass jeder Messwert sich aus der tatsächlichen Datenausprägung und einem unsystematischen Mess- oder Zufallsfehler additiv zusammensetzt (Schweinitz, 2015).

Zur Abschätzung der Reliabilität einer wissenschaftlichen Messung sind drei Methoden gebräuchlich: die Testwiederholung (Test-Retest-Reliabilität), die Testhalbierung und die Interne Konsistenz (Caspar und Wirtz, 2002).

Im Folgenden wird nur auf die in dieser Arbeit angewandte Test-Retest-Reliabilität eingegangen. Bei dieser Methode wird an zwei auseinanderliegenden Zeitpunkten an denselben Proband_innen ein Test wiederholt. Die Korrelation der Messwerte zu den

unterschiedlichen Zeitpunkten gilt als Indikator für die Stabilität bzw. Gleichheit der Messergebnisse.

Die Bedeutung der Reliabilität in fMRT-Studien zeigt sich in vielfältiger Form bei der Anwendung von fMRT-Ergebnissen (Bennett und Miller, 2010).

Zu allererst sei hierbei die Reproduzierbarkeit und somit die Möglichkeit der Generalisierung von den Messergebnissen aufgeführt, die zu einer allgemeinen wissenschaftlichen Wahrheit beitragen. Allem voran ist hier zu bedenken, dass bei der Detektion der prinzipiell sehr schwachen Magnetresonanz-Signale eine hohe Beeinflussung durch Störfaktoren vorliegt, so dass es zu stark variierenden Messungen dieses Signales kommen kann. Von einer Gruppe Proband_innen zu einer anderen und von einem MRT zum anderen sollte genau erforscht sein, welche Variablen die Ergebnisse beeinflussen können und wie reliabel sie schlussendlich sein können. Eine hohe Reliabilität einer Messung begünstigt die Vergleichbarkeit der Ergebnisse und stellt somit die Grundlage für den wissenschaftlichen Austausch und die Möglichkeit der Integration der Ergebnisse in Metaanalysen dar. Des Weiteren werden fMRT-Daten in der klinischen Diagnostik und Therapie angewandt und führen u.a. zu chirurgischen Konsequenzen wie beispielsweise die Inzision epileptischer Areale nach zuvor mittels fMRT getesteten Sprachfunktionsdominanz (Fernandez et al., 2003) oder Anpassungen der Pharmakotherapie (Wise und Preston, 2010).

2.2 Wortgenerierung (WG)

Das neuroanatomische und -kognitive Konzept der Sprache hat sich seit seiner Entdeckung im 19. Jahrhundert wesentlich gewandelt. Während bei der klassischen anatomischen Einteilung noch die Artikulation der Sprache im Vordergrund stand, ist mittlerweile der Aspekt der Sprachverarbeitung Gegenstand der aktuellen Forschung. Der die Artikulation und Sprachverarbeitung umfassende kognitive Prozess der Sprache ist komplex und unterliegt einer besonderen mehrphasigen Dynamik verschiedener aktivierter Hirnregionen.

In dieser Arbeit wird der Prozess der semantischen Sprachproduktion untersucht und im folgenden Abschnitt daher ausschließlich auf diesen Teilbereich der Sprache eingegangen. Zunächst wird die Sprache in ihrer klassischen anatomischen Zuordnung zu den beteiligten Hirnregionen vorgestellt, im Anschluss daran folgt eine detaillierte Aufschlüsselung des Sprachgenerierungsprozesses basierend auf umfassenden aktuellen fMRT-Studien zu

diesem Thema. Schließlich wird auf die Aufgabe der semantischen Wortgenerierung im Kontext von bisherigen Forschungsergebnissen eingegangen.

Die beiden bekanntesten anatomischen Regionen, welche mit der Sprachverarbeitung im Gehirn in Verbindung gebracht werden, sind die Broca- und Wernicke-Region. Das klassische motorische Sprachzentrum stellt das nach seinem Erstbeschreiber Paul Broca im 19. Jahrhundert benannte Broca-Sprachzentrum dar. Es liegt im Bereich des Pars opercularis und zum Teil auch des Pars triangularis des Gyrus frontalis inferior und seinen umliegenden Regionen. In der vielzitierten zytoarchitektonischen Einteilung nach Korbinian Brodmann wurde es aufgrund mikroskopischer Analysen 1909 im Bereich der Areale 44 und 45 angesiedelt.

Das Wernicke-Sprachzentrum stellt das Zentrum des Sprachverständnisses in der sekundären Hörrinde und des Brodmann Areals 22 dar. Die Areale stehen über den Fasciculus arcuatus miteinander in Verbindung. Bei 95% der Bevölkerung liegt eine linkshemisphärische Sprachdominanz vor (Fernandez et al., 2009).

Bisher bekannte korrespondierende Hirnregionen sind die Wernicke Region, von wo aus afferente Fasern in das Broca-Zentrum reichen. Weiterhin bestehen Assoziationsfelder zum Gyrus Angularis, welche die zentralnervöse Komponente der Sprache integrieren. Efferent ziehen Fasern direkt und indirekt (via Basalganglien zwischen Kleinhirn und Thalamus) zum Gyrus Präzentralis, von wo aus über kortikonukleäre Bahnen die entsprechenden Hirnstammkerne für die Sprechmuskulatur aktiviert werden (Trepel, 2004).

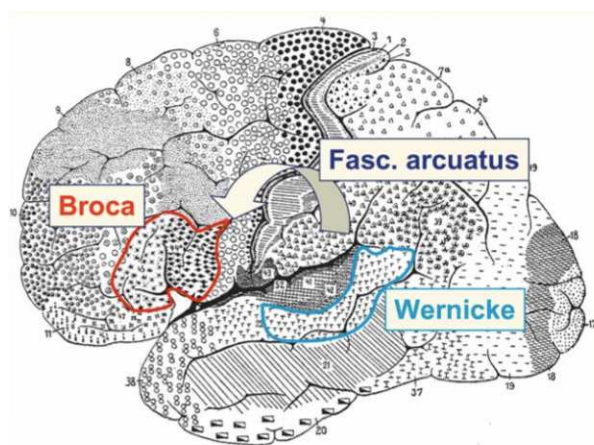


Abb. 2.1 Lage des motorischen und sensorischen Sprachzentrums in einer Brodmann-Karte. Die Broca-Region liegt in der Area 44, 45, rot und Wernicke-Region in der Area 22, blau. Beide Regionen haben nach heutigem Wissenstand größere Ausdehnungen. (entnommen aus Schneider und Fink, 2013).

Die Broca-Region ist an der Verarbeitung einer Vielzahl von linguistischen Prozessen und Informationen (Semantik, Syntax, Phonologie, Zweitsprachkontrolle) beteiligt und zeigt dabei eine funktionelle Spezialisierung innerhalb ihrer Region (Schneider und Fink, 2013).

In dem Übersichtsartikel ‚The Anatomy of Language‘ (*Die Anatomie der Sprache*, Price 2010) unterteilt Price den Sprachprozess der Broca-Region auf der Basis einer Metaanalyse von fMRT-Studien in verschiedene Level: Sprache entsteht demnach mit der Verarbeitung von Ideen (*Conceptual Processing*). Anhand der Metaanalyse von Binder et al. (Price, 2010) sind konsistent sieben links lateralisierte Hirnregionen an diesem Prozess beteiligt, die in Tabelle 2.1 aufgelistet sind.

Wortgenerierungsschritt	Aktivierte Hirnregionen
<i>Processing</i>	Inferiorer frontaler Gyrus Ventraler und dorsaler medialer präfrontaler Kortex Posteriorer inferiorer Parietallappen Mittlerer temporaler Gyrus Fusiformer Gyrus Parahippocampaler Gyrus Posteriorer Gyrus Cinguli
<i>Word Retrieval</i>	linker inferiorer und medialer frontaler Gyrus mit Pars opercularis, Pars triangularis und dem inferioren frontalen Sulcus
<i>Articulation</i>	bilateraler Motor- und Prämotorkortex Zerebellum SMA (supplementary motor area) superiorer temporaler Gyrus temporo-parietaler Kortex anteriore Insula mit links lateralisierter Aktivierung im Putamen
<i>Auditory Feedback</i>	Bilaterer auditiver Kortex SMA

Tab. 2.1 Übersicht der den jeweiligen Schritten der Wortgenerierung zugeordneten Hirnregionen nach Price (2010)

Auffällig ist hierbei, dass die Wörter je nach ihrer Kategorie andere Regionen mitaktivieren, beispielsweise bei somatosensorischen Kategorien den postzentralen und inferioren Parietallappen und bei motorischen Kategorien den intraparietalen Sulcus (Price, 2010).

Nach dem Schritt des *Conceptual Processing* folgt das Level des *Word Retrieval*: Wörter, die mit den oben beschriebenen Ideen assoziiert sind, werden aus dem Wortgedächtnis abgerufen und gleichzeitig die Wörter ähnlicher aber nicht adäquater Bedeutung

unterdrückt. Die Differenzierung zwischen linguistischen und nicht-linguistischen Lauten erfolgt im linken dorsalen Pars opercularis und die Funktion der Differenzierung zwischen Artikulation und Wahrnehmung eines Wortes übernimmt der linke ventrale Pars opercularis.

Es folgen die Wortgenerierungsschritte der Artikulation (*Articulation*) und des auditorischen Feedback (*Auditory Feedback*) über die Wernicke-Region. Auf diesen beiden Ebenen findet eine Sequenzierung des Sprachprozesses statt. So muss beispielsweise auf der Ebene der Artikulation eines Wortes die korrekte Kombination von Vokalen und Konsonanten zusammengestellt werden, dabei die Sprachmelodie mit Tonlage, Rhythmus und die Lautstärke gewählt werden, die natürlich auch von dem Kontext der Konversation abhängt. Den entworfenen phonetischen Plan setzt der motorische Kortex in einen Koordinationsplan der zu vibrierenden Stimmlippen, Zunge und weiteren beteiligten Muskeln um. Gleichzeitig muss die Atmung kontrolliert und das empfangende auditorische Signal überwacht werden, um die folgenden Prozesse zu adjustieren. Auf der Ebene des Auditorischen Feedback werden demnach die auditiven, artikulären und somatosensorischen Signale in das motorische Output integriert. In Anbetracht dieses umfangreichen Prozesses der Artikulation eines einzelnen Wortes wird deutlich, wie viele Hirnregionen an der Sprachproduktion beteiligt sind.

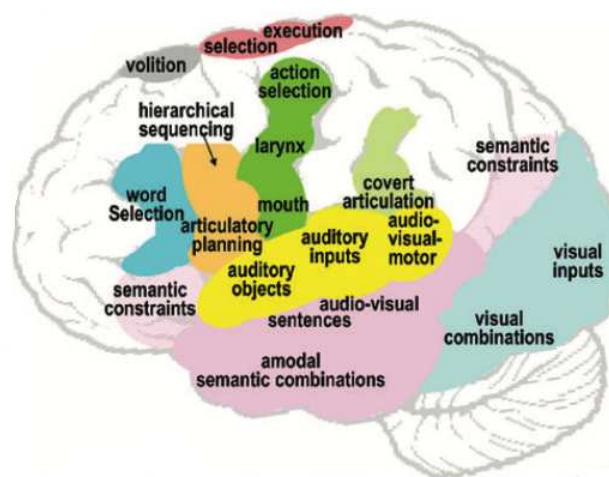


Abb. 2.2 Übersicht beteiligter Hirnregionen bei der Sprachproduktion (modifiziert nach Price, 2010)

Semantische Wortflüssigkeitsaufgaben untersuchen gezielt den Prozess der kontinuierlichen semantischen Sprachproduktion. Dabei sollen in einem vorgeschriebenen Zeitraum die Proband_innen möglichst viele Wörter, die zu bestimmten semantischen Kategorien

gehören (wie beispielsweise Wörter, die zum Oberbegriff ‚Kleidung‘ passen), aufzählen. Dies erfordert- abgesehen von den bisher genannten Aktivierungsbereichen der Sprachproduktion - weitere intakte Bereiche der Kognition: das verbale Gedächtnis, die Exekutivfunktion und die Verarbeitungsgeschwindigkeit. Für die Effizienz muss auch Zugang zum semantischen Lexikon, dem Speicher der Wortbedeutungen, sowie dessen Organisation intakt sein (Bouma, Mulder und Lindeboom, 1998). Dem semantischen Lexikon liegt als organische Basis eine Region des Temporallappens vor (Troyer et al., 2000).

Wortflüssigkeitsaufgaben prüfen wie o.g. unter anderem die Exekutivfunktion. Das ist jene Fähigkeit, die man benötigt, ein Wort in einem vorgegebenen Zusammenhang sinnvoll auszusprechen. Inhibiert werden müssen jedoch alle Brüche von festgelegten Regeln (Bouma et al., 1998). Diese Fähigkeit wird dem Frontallappen zugeordnet (Benton 1968). Um die Regeln einer Aufgabe beachten zu können, ist ein intaktes Arbeitsgedächtnis notwendig (Heinrich & Zakzanis 1998).

Die gefundenen Aktivierungsmuster korrelieren mit Alter und verbaler Intelligenz (Fu et al., 2005). Alter korreliert positiv mit einer bilateralen Aktivierung des Gyrus Frontalis medius und inferior, dem anterioren Gyrus Cinguli, dem linken Gyrus Präzentralis und der rechten Insula. Je geringer die verbale Intelligenz, desto höher ist die Aktivierung im rechten Gyrus Temporalis superior und im linken Gyrus Frontalis superior (Nagels et al., 2012).

2.3 Subliminale Präsentation emotionaler Stimuli (SPES)

Viel älter als die Entwicklung der Sprache und dementsprechend anatomisch in stammesgeschichtlich alten Hirnstrukturen lokalisiert ist der Prozess der Emotion. Nahezu alle psychologischen Forschungsaspekte wie beispielsweise Gedächtnis, Verhalten, Lernen und Wahrnehmung unterliegen dem Einfluss von Emotionen, wodurch die Erforschung der neuronalen Korrelate von emotionalen Verarbeitungsprozessen von großem Interesse ist. Demgegenüber steht die mangelnde Operationalisierbarkeit und Zugänglichkeit von Emotionen für Untersuchungsbedingungen. Bisher wurden durch bildgebende Verfahren vor allem die Bereiche des emotionalen Erlebens, der emotionalen Diskriminationsfähigkeit, des emotionalen Gedächtnisses und Lernprozesses charakterisiert (Kandel et al., 2013).

Diese Arbeit hat den Untersuchungsschwerpunkt auf die Wahrnehmung unterbewusster emotionaler Stimuli gelegt und wird im folgenden Einleitungsabschnitt primär darauf eingehen. Zu Beginn dieses Abschnittes wird der wissenschaftliche Begriff der Emotion erläutert und darauf aufbauend verschiedene Konzepte der anatomischen Zuordnung emotionaler Prozesse vorgestellt. Das in dieser Arbeit verwandte *Backward-masking* Konzept wird im Anschluss erläutert. Schließlich wird die Operationalisierbarkeit von Emotionen mittels Stimmungsinduktionsmethoden aufgeführt und daran anlehnend auf aktuelle Forschungserkenntnisse hingewiesen.

Für die sachliche Auseinandersetzung mit den kognitiven Prozessen der Emotionen wurde in der Wissenschaft eine Kategorisierung der Emotionen festgelegt, auf welche sich grundsätzlich geeinigt werden kann (Gazzaniga, Ivry und Mangun, 2015). Die erste Kategorie beinhaltet sechs Basisemotionen mit einzigartigen Eigenschaften: Angst, Freude, Trauer, Ekel, Überraschung und Wut. Die zweite Kategorie beschreibt auf den Basisemotionen aufbauende komplexe Emotionen, welche wie beispielsweise Liebe zum Teil sozial und kulturell erlernt werden. Die dritte Kategorie stellt die Dimensionen der Valenz und Stärke des Affektes dar. Diese Arbeit befasst sich ausschließlich mit den Basisemotionen Angst, Freude und Trauer der ersten Kategorie.

Das erste Konzept, welches die neuralen Strukturen der Emotionen zu erfassen suchte, stellt der von James Papez 1937 definierte *Papez Kreis* dar. Demnach besteht das Netzwerk emotionaler Antworten aus dem Hypothalamus, dem anterioren Thalamus, dem Gyrus Cinguli und dem Hippocampus (Trepel, 2004). 1949 erweiterte Paul MacLean dieses Konzept unter anderem um den orbitofrontalen Kortex und die Amygdala und nannte es das *Limbische System* (Gazzaniga et al., 2015). Die Idee eines geschlossenen Netzwerkes ist mittlerweile jedoch wiederlegt (Fusar-Poli et al., 2009). Vielmehr geht man heutzutage von vielen situations- und personenabhängigen Systemen mit individuellen Aktivierungen aus.

Im Mittelpunkt der Emotionsforschung der letzten drei Jahrzehnte steht die Amygdala. In ihrer anatomischen Nähe zum Hypothalamus, welcher das autonome Nervensystem reguliert, ist die Amygdala für die Detektion und Evaluation der Relevanz emotionaler Stimuli verantwortlich und bereitet somit hypothalamisch getriggerte Triebreaktionen vor (Schneider und Fink, 2013). Die darüber hinausgehende Differenzierung der Stimuli übernimmt der ventrale präfrontale Kortex (Dima, Klass und Roiser, 2011). Ihre zentrale

Funktion hat die Amygdala im impliziten emotionalen Lernen, wobei sie die Stärke der Emotionen durch Interaktionen mit dem Hippocampus modulieren kann und somit auch an expliziten Lernprozessen und dem Gedächtnis beteiligt ist. Darüber hinaus ist sie in der Wahrnehmung und Aufmerksamkeit bei sozialen Interaktionen involviert, was es beispielsweise ermöglicht, Emotionen aus den Gesichtsausdrücken und Augen unseres Gegenübers festzustellen (Gazzaniga et al., 2015).

Bei der Erfahrung emotionaler Stimuli beeinflusst die Amygdala, wie sensorische Informationen von Stimuli prozessiert werden (Rauch et al., 2010) und bei negativen Stimuli die Wahrnehmungssensibilität erhöhen (Barrett, Bliss-Moreau, Duncan, Rauch und Wright, 2007).

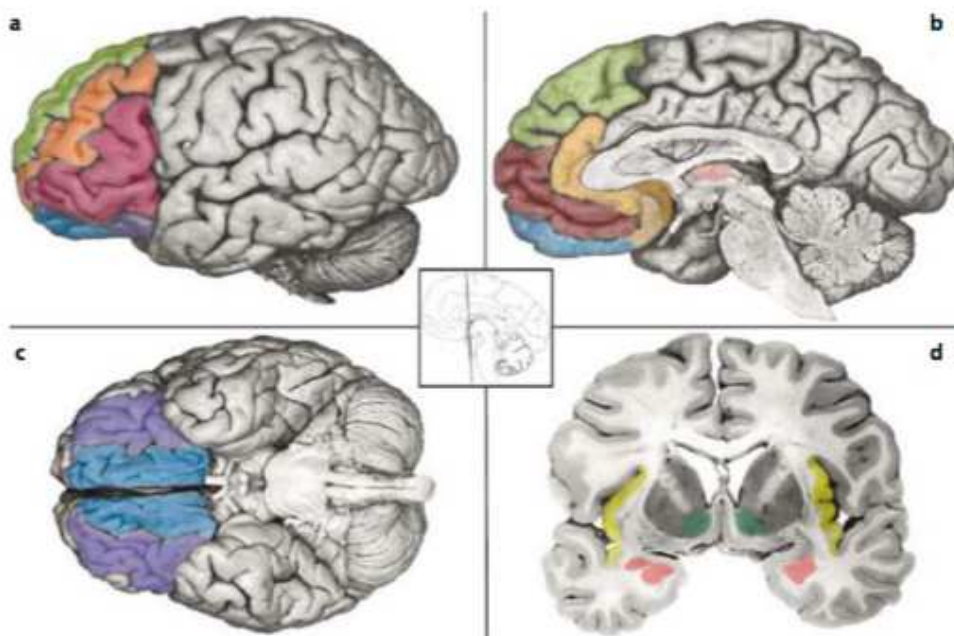


Abb 2.3: Spezifische Hirnregionen, die bei Emotionen eine Rolle spielen in der Ansicht von sagittal (a,b), axial (c) und coronar (d) nach Barrett et al. (2007): orbitofrontaler Kortex (violett), ventromedialer präfrontaler Kortex (blau), Insula (gelb), (subgenual) anteriorer zingulärer Kortex (beige und braun), dorsolateral präfrontaler Kortex (orange und pink), dorsomedial präfrontaler Kortex (hellgrün), ventrales Striatum (dunkelgrün), Thalamus (b, rosa) sowie die Amygdala (d, rosa) (modifiziert nach Barrett et al. 2007).

Die Amygdala stellt jedoch trotz ihrer oben aufgeführten Relevanz nur einen Teil der Emotionsverarbeitung dar, denn jede Basisemotion der oben genannten ersten Kategorie wird aus einem eigenen neuronalen Netzwerk generiert. Eine Metaanalyse von Fusar-Poli et al. (2009) identifiziert die folgenden emotionsspezifischen Muster: sowohl positive als auch

negative Emotionen zeigen eine verstärkte Aktivierung der Amygdala, Zerebellum, Putamen, temporoparietalen und präfrontalen Regionen und auch dem visuellen Kortex. Die Basalganglien haben durch die hohe Konzentration dopaminerger Neurone und ihrer Rolle im Rahmen des Belohnungssystems eine wesentliche Bedeutung für die Emotion Freude. Negative Affekte und insbesondere die Emotion Angst sind dagegen am häufigsten in der Amygdala lokalisiert.

Wie in Tabelle 2.2 zusammengefasst, zeigen Wut und Angst erhöhte Aktivierungen des inferioren Gyrus Frontalis; eine Aktivierung der Insula findet selektiv bei Ekel und Wut statt, wobei die Aktivierung bei Ekel stärker ist; Trauer aktiviert den anterioren Gyrus Cinguli; Angst, Freude und Trauer aktivieren die Amygdala, Ekel und Wut allerdings nicht; Angst zeigt eine besonders starke Aktivierung der Amygdala (Fusar-Poli et al., 2007). Hirnregionen, die bei Freude bisher als aktiviert beobachtet werden konnten, sind der dorsolaterale präfrontale Kortex, der Gyrus Cinguli, der inferiore Gyrus Temporalis und das Zerebellum. Diese Aktivierungen zeigten sich auch bei der Emotion Trauer, wobei hier auch noch der parahippocampale Gyrus und der Precuneus aktiviert waren.

WUT	ANGST	TRAUER	FREUDE	EKEL
IFG	IFG	Anteriorer Gyrus Cinguli	Anteriorer Gyrus Cinguli	Insula
Insula	Amygdala	Amygdala	Amygdala	
	Hippocampus	Insula	Zerebellum	
		Precuneus	Inferiorer Gyrus Temporalis	
		Parahippocampaler Gyrus	Dorsolateraler präfrontaler Kortex	

Tab. 2.2 Übersicht der den Basisemotionen Wut, Trauer, Angst, Freude und Ekel zugeordneten Hirnregionen nach Fusar-Poli et al. (2009)

In der Zusammenschau der Aktivierungsnetzwerke der einzelnen Emotionen zeigt sich deutlich, dass sowohl kortikale Strukturen, wie beispielsweise parietale und präfrontale Regionen als auch subkortikal-limbische Hirnregionen in der Emotionsverarbeitung involviert sind. Vermutlich fördert die Interaktion zwischen Amygdala und kortikalen Regionen die Aufmerksamkeit und Wahrnehmungsverarbeitung emotionaler Reize

(Schneider und Fink, 2013). Kortikale Aktivierungen zeigen sich ebenfalls verstärkt bei für das Überleben direkt bedrohlichen Reizen, um schnelle körperliche Antworten auslösen zu können. Im Gegensatz dazu lösen traurige Gesichter als Stimulusmaterial nur minimale Aktivierungen aus, da hier keine direkte körperliche Reaktion folgen muss (LeDoux, 1998).

Neben den Aktivierungsunterschieden hinsichtlich der Bedrohlichkeit finden sich auch Unterschiede in der bewussten und unbewussten Wahrnehmung dieser Reize. Es wird von zwei verschiedenen neuronalen Verarbeitungswegen ausgegangen, welche verschiedene Ebenen der Emotionsanalyse und Reaktion ermöglichen. Bei bewusster Wahrnehmung emotionaler Reize wird die Information zunächst an die sensorischen Kortices und über den Thalamus direkt an die Amygdala weitergeleitet, welche eine grobe Bewertung des Stimulusmaterials vornimmt und simultan höher geordnete kortikale Prozesse eine genaue Analyse vornehmen, während unbewusste Reize subkortikal verarbeitet werden (Killgore und Yurgelun-Todd, 2004). Diese Vermutungen decken sich mit seit langem bekannten Theorien, dass das Gehirn viele Erfahrungen einzig auf unbewusster Ebene verarbeitet (Freud 1915).

In dieser Arbeit wurde die effektive Methode des *Backward-Masking* (Rückwärtsmaskierung) zur Untersuchung unbewusster emotionaler Prozesse angewandt. Hierbei erfolgt die Darbietung emotionalen Materials unterhalb der bewussten Wahrnehmungsschwelle. Diese unterbewusste, als subliminal bezeichnete Schwelle liegt bei gesunden Proband_innen bei 30-38ms (Murphey und Zajonc, 1993). Auf das emotionale Material folgend wird über einen längeren, bewussten Zeitraum ein neutraler Stimulus gezeigt, welcher den vorhergehenden Reiz somit maskiert.

Die Validität des *Backward-Masking* als Stimmungsinduktionsmethode wurde in vorhergehenden Experimenten bestätigt (Suslow, Roestel und Arolt, 2003; Murphey und Zajonc, 1993). Allerdings bleibt fraglich, wie genau sie tatsächlich diejenige Emotion auslöst, welche das experimentelle Design vorgibt. Grundlegend für die experimentelle Untersuchung der emotionalen Netzwerke ist die Verwendung standardisierter Methoden der Stimmungsinduktion. Hierzu werden unter anderem die freie Erinnerung an eigene Erlebnisse, das Hineinversetzen in emotionale Zustände anhand von Filmen, Musik, Geruch etc., die Rückmeldung von Erfolg und Misserfolg und experimentelle physiologische Veränderungen bspw. mittels Medikamentengabe, verwendet (Schneider und Fink, 2013).

Allerdings besteht beispielsweise noch kein wissenschaftlicher Konsens über die Induktion von Freude. Während einerseits diskutiert wird, dass Freude aus kognitiven, ästhetischen oder moralischen Erfolgen resultiert (Csiszentmihalyi, 1990) wird andererseits die Freude bei somatischer Zufriedenstellung gegenübergestellt.

Ein weiteres Problem der Induktion von Emotionen liegt in der Abhängigkeit von individuellen Einflussfaktoren. Hierzu gehören die jeweilige Tagesform, Persönlichkeitsvariationen, Geschlecht, hormonelle Schwankungen, Aufmerksamkeitslevel aber auch persönliche Erfahrungen und vieles mehr.

So gelang es vielen Forschungsgruppen bisher nicht, die Ergebnisse der Amygdalaaktivierungen während des emotionalen Erlebens zu reproduzieren. Anstelle davon beobachteten einige eine verstärkte Aktivierung in den Regionen des anterioren Gyrus Cinguli und des präfrontalen Kortex (Mayberg et al., 1999). Boubela et al. weisen sogar in einer fMRT Studie von 2015 auf die Möglichkeit einer bisherigen Verwechslung der Amygdala Aktivierung mit der sich in unmittelbarer Nähe befindenen Aktivierung im Bereich der Rosenthalvene hin (Boubela et al., 2015).

3 METHODIK

Im folgenden Kapitel werden die methodischen Grundlagen, die Probandendaten und fMRT-Parameter sowie das experimentelle Design dieser Studie vorgestellt. Im Anschluss wird zunächst die funktionelle Datenauswertung und schließlich das zentrale Thema dieser Arbeit, die Reliabilitätsanalyse, im Allgemeinen und in ihrer Anwendung in dieser Studie beschrieben.

3.1 Methodische Grundlagen

Dieser Abschnitt behandelt die Grundlagen der verwendeten Methode. Es wird zunächst auf die allgemeine Funktionsweise der MRT und der fMRT eingegangen. Es folgt die Darstellung der Analyse und Verarbeitung von fMRT-Daten. Der letzte Teil dieses Kapitels stellt das Design eines fMRT-Experiments vor.

3.1.1 Magnetresonanztomographie

Die Magnetresonanztomographie ist ein höchst differenziertes nicht invasives Bildgebungsverfahren, welches dafür genutzt wird, Strukturen des menschlichen Körpers in Form von Schnittbildern zu visualisieren. In der medizinischen Diagnostik wird die MRT seit den 80er Jahren eingesetzt und ist heute Routinebestandteil der klinischen Diagnostik.

Der physikalische Hintergrund der MRT ist die Darstellung der Kernspinresonanz, deren zugrunde liegende Funktionsweise erstmals durch die späteren Nobelpreisträger Edward Will Purcell (Purcell, Torrey und Pound, 1946) und Felix Bloch (Bloch, Hansen und Packard, 1946) beschrieben wurde (Jansen, 2004). Die Kernspinresonanz beruht auf der schwach magnetischen Eigenschaft von Wasserstoffatomen, welche hauptsächlich in Form von Wasser überall im menschlichen Körper vorhanden sind. Befindet sich der menschliche Körper in einem starken Magnetfeld, so richten sich die Kernspins der Wasserstoffatome entlang der Feldlinien aus, so dass eine Magnetisierung des Körpers selbst stattfindet. Werden auf die entstandene Magnetisierung hochfrequente elektromagnetische Felder eingestrahlt, so ändert sie sich für einen kurzen Moment, um danach wieder ihren ursprünglichen Zustand anzunehmen. Hierbei setzt der Körper Energie in Form von

Radiowellen ab, welche vom Tomographen gemessen werden können. Da der Zeitverlauf dieser Magnetresonanzsignale in den verschiedenen Geweben des Körpers unterschiedlich ist, kann eine Differenzierung der einzelnen Körperstrukturen erfolgen.

Im Folgenden werden der Begriff der Magnetisierung und die Entstehung von Magnetresonanzsignalen und Magnetresonanzbildern erläutert. Dabei werden wiederholt die Begriffe *z-Richtung*, *xy-Ebene* und *äußeres Magnetfeld B_0* verwendet. Die Bedeutung davon wird auf der folgenden Abbildung dargestellt.

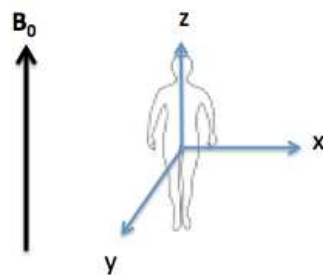


Abb. 3.1 Ausrichtung des äußeren Magnetfeldes. Das äußere Magnetfeld B_0 liegt parallel zur *z-Richtung* und verläuft von unten nach oben. Axialschichten liegen in der *xy-Ebene*, Sagittalschichten in der *zy-Ebene* und Koronarschichten in der *xz-Ebene*. Diese Konventionen sind angelehnt an den stereotaktischen Atlas von Talairach und Tournoux (Talairach, Tournoux 1988).

Kernspin und Lamorpräzession

Im Fokus der Erzeugung von Magnetresonanzsignalen stehen wie bereits erwähnt die Wasserstoffatome des menschlichen Körpers. Die im Kern dieser Wasserstoffatome vorhandenen Protonen führen eine ständige Eigenrotation aus, den sogenannten Kernspin, welcher ähnlich einem Kreisel ist (Weishaupt et al., 2014). Gleichzeitig hat das Proton eine rotierende elektrische Ladung mit einer magnetischen Wirkung in Richtung der Drehachse, welche hier als magnetischer Moment B beschrieben wird. Sie „(...) verhält sich wie ein kleiner Magnet, der von Magnetfeldern und elektromagnetischen Wellen beeinflusst wird und wenn er sich bewegt, in einer Empfangsspule eine Spannung induziert.“ (Weishaupt et al., 2014) Die Rotation selber ist jedoch nicht sichtbar, sondern nur die Lage der Rotationsachse am Magnetvektor B , welche in der Empfangsspule eines Tomographen das Signal erzeugt.

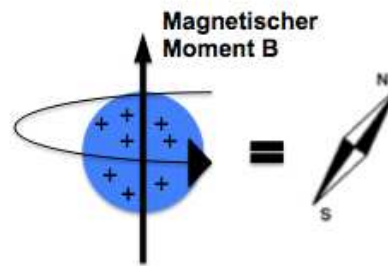


Abb. 3.2 Schematische Darstellung vom Magnetischen Moment B in einem Wasserstoffproton (modifiziert nach Weishaupt et al., 2014)

In einem Raum ohne äußeres Magnetfeld orientieren sich die Kernspins ohne Vorzugsrichtung an und weisen keine messbare Magnetisierung auf. Wird nun ein starkes äußeres Magnetfeld eingestrahlt, richten sich die Kernspins in parallelen und antiparallelen Feldlinien aus. Da die parallele Ausrichtung hierbei energetisch günstiger ist, existieren mehr parallele als antiparallele Spins. Dieser geringe Überschuss parallel orientierter Spins bewirkt eine Polarisierung der Spins. Aus der vektoriellen Summe der einzelnen magnetischen Momente B kann dann eine effektive Magnetisierung berechnet werden, die sogenannte makroskopische Magnetisierung M^{\rightarrow} , welche die nach außen messbare Magnetisierung der makroskopischen Probe darstellt.

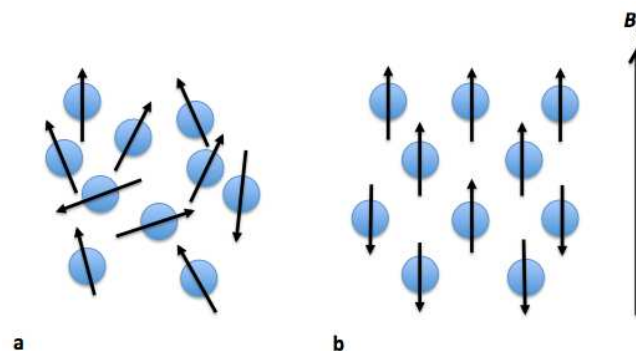


Abb. 3.3 Ausrichtung der Kernspins (a) ohne Einwirkung des Magnetfeldes B_0 ungeordnet und (b) im statischen Magnetfeld B_0 in bevorzugt paralleler und antiparalleler Orientierung

Wie bereits erwähnt, ähnelt der Eigendrehimpuls der Protonen einem magnetischen Kreisel. Er kann die Lage seiner Rotationsachse bei Einwirkung von beispielsweise Schwerkraft ändern und somit eine Ausweichbewegung durchführen, während das äußere Magnetfeld B_0 versucht, ihn wie eine Kompassnadel parallel zu sich auszurichten. Die

Ausweichbewegung ist als sogenannte *Präzessionsbewegung* bekannt, welche in einer charakteristischen Frequenz- der *Lamorfrequenz* ω_0 - um das äußere Magnetfeld B_0 präzediert und proportional zur Stärke des Magnetfeldes ist. Die *Lamorfrequenz* ω_0 wird in der folgenden Gleichung beschrieben:

$$\omega_0 = \gamma_0 \cdot B_0 \quad (3.1)$$

γ_0 bezeichnet in dieser Gleichung das gyromagnetische Verhältnis, welches für jeden Atomkern spezifisch ist. Auf der Lamorfrequenz beruht die gesamte MR-Bildgebung.

Die Richtung in welcher der magnetische Moment B des Atomkerns präzediert erfolgt in der z-Richtung des Magnetfeldes, entspricht demnach also einer Längsmagnetisierung. Die Wirkung der einzelnen magnetischen Momente der Wasserstoffatomkerne bleibt null, da sie außer Phase präzedieren, was bedeutet, dass sie, obwohl sie die gleiche Rotationsfrequenz besitzen, nur zufällig in der xy- Ebene orientiert sind. Die makroskopisch messbare Magnetisierung bleibt aufgrund dieser Phaseninkohärenz der einzelnen magnetischen Momente auf ihre z-Komponente beschränkt.

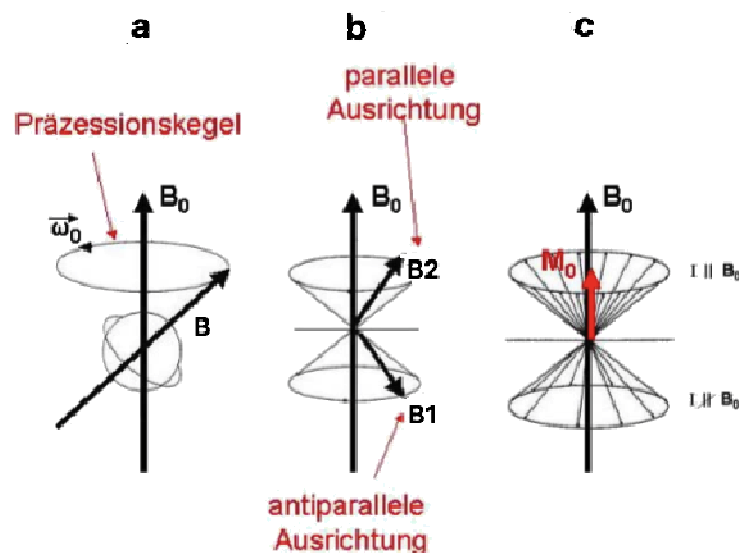


Abb. 3.4 Präzessionsbewegung. a) Präzessionsbewegung eines magnetischen Momentes B mit der Lamorfrequenz ω_0 in einem magnetischen Feld B_0 . b) Die beiden Präzessionskegel entsprechen den Energieniveaus, welche sich dadurch ergeben, dass die magnetischen Momente B_1 und B_2 sich parallel oder antiparallel zum magnetischen Feld B_0 ausrichten. c) Die parallele Ausrichtung ist energetisch bevorzugt und wird daher von mehr Wasserstoffkernen eingenommen. Es bildet sich eine makroskopisch messbare Magnetisierung heraus. (Abb. übernommen und modifiziert nach Jansen, 2004).

Dynamik der Magnetisierung: Resonanz und Relaxation

Wie im letzten Abschnitt erläutert, bewirkt der Energieunterschied der parallelen und antiparallelen Orientierung der magnetischen Momente die messbare makroskopische Längsmagnetisierung im stabilen Zustand. Diese Energie ist jedoch nicht direkt mit einem MR-Tomographen messbar (Schneider und Fink, 2013). Um sie messbar zu machen, muss in diesem stabilen System eine Wechselwirkung ausgelöst werden. Dies erreicht man mit einem elektromagnetischen Hochfrequenz-Impuls (*HF-Impuls*), dessen Frequenz exakt auf der Larmorfrequenz des präzedierenden Spins erfolgen muss, um der gleichen Energie zu entsprechen. Der HF-Impuls wird senkrecht zum bestehenden magnetischen Feld B_0 eingestrahlt. Mit der richtigen Energie kann der HF-Impuls eine Auslenkung um genau 90° bewirken und somit die gesamte Magnetisierung in die xy -Ebene umkippen. Die Bewegung der Magnetisierung aus der z - in die xy -Ebene induziert in der Empfangsspule eine Wechselspannung. Es ist eine *Resonanzbedingung* hergestellt worden und das MR-Signal entstanden (Weishaupt et al., 2014).

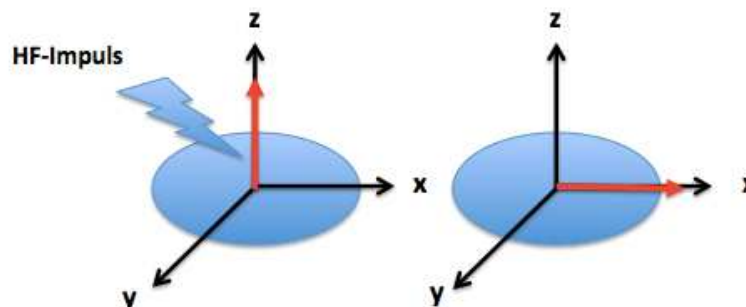


Abb.3.5 Auswirkung HF-Impuls. Durch einen geeigneten HF-Impuls kann eine Auslenkung des Vektors um 90° bewirkt werden. Die makroskopische Längsmagnetisierung M_z befindet sich nun in der xy -Ebene und ist somit als Quermagnetisierung M_{xy} ausgerichtet (Abb. modifiziert nach Jansen, 2004).

Wird der HF-Impuls wieder abgeschaltet, stellt sich der ursprüngliche Gleichgewichtszustand der Längsmagnetisierung M_z und Quermagnetisierung M_{xy} her und die eingestrahlte Energie wird wieder abgegeben. Man beschreibt diesen Vorgang als Relaxation, wobei man die beiden Komponenten Relaxation der Längsmagnetisierung und Relaxation der Quermagnetisierung als zwei voneinander unabhängige, gleichzeitig stattfindende Prozesse verstehen muss.

T1-Relaxation

Nach einer gewissen Zeit klappt die Quermagnetisierung von der transversalen Vektorprojektion zurück in die z-Richtung und baut dort die Längsmagnetisierung wieder auf, das MR-Signal nimmt hierbei ab. T1 gibt die Zeit nach einem HF-Impuls an, in der die ursprüngliche Längsmagnetisierung wieder auf 63% des Gleichgewichtswertes angestiegen ist. Bei diesem Prozess werden abhängig von unterschiedlichen Magnetfeldschwankungen durch Molekularbewegungen Energien der Spins an das umgebende Gewebe („Gitter“) abgegeben. T1 ist somit abhängig von der Magnetfeldstärke. Man versteht diesen Prozess als Spin-Gitter-Relaxation. In strukturellen Aufnahmen des Gehirns ist die T1-Wichtung von besonderer Bedeutung, da sich hierdurch eine deutliche Kontrastierung von weißer und grauer Substanz ergibt (Schneider und Fink, 2013).

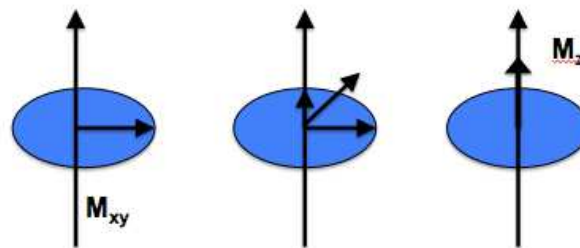


Abb. 3.6 Schema einer T1-Relaxation (modifiziert nach Weishaupt et al., 2014)

T2/T2-Relaxation*

Relaxationszeit T2 beschreibt die Zeit nach dem HF-Impuls, in welcher noch 37% der Quermagnetisierung vorhanden sind. Für den Zerfall spielen zwei Komponenten eine Rolle: einerseits regen sich verändernde Magnetfeldumgebungen benachbarte Spins zu einem Energieaustausch an. Die zu Beginn des HF-Impulses synchronen Präzessionsbewegungen der Spins werden dabei aufgehoben und das MR-Signal nimmt ab. Diesen Prozess bezeichnet man als Spin-Spin-Wechselwirkungen. Er ist jedoch nur in einem homogenen magnetischen Feld vorhanden, da der Körper allerdings vor allem an Gewebekontaktflächen seine Magnetfeldeinhomogenitäten hat, dephasieren die Spins dort schneller als bei T2.

Dies wird als T2*-Zerfall beschrieben. Er ist im Gegensatz zu der T1-Wichtung unabhängig von der Magnetfeldstärke.

Die genannten Relaxationszeiten T1 und T2/T2* werden noch durch einen dritten Parameter ergänzt, der Protonendichte. Sie beschreibt die Anzahl der erregbaren Kernspins, die das maximal mögliche Signal eines Gewebes erzeugen kann. Da jeder einzelne Parameter eigene Charakteristika hat, bilden sie Gewebe auf unterschiedliche Weisen ab (Jansen, 2004). Während sich beispielsweise Flüssigkeiten in der T2-gewichteten Aufnahme weiß darstellen, zeigen sie sich in der T1-Wichtung als dunkel. Dies stellt die Grundlage zur Bildkontrastgebung in der MRT dar.

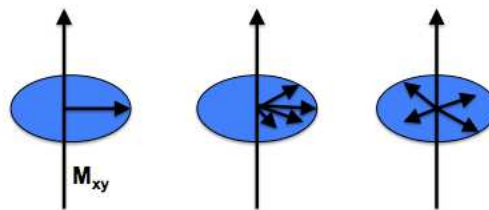


Abb. 3.7 Schema einer T2-Relaxation (modifiziert nach Weishaupt et al., 2014)

MR-Bildgebung

In den beiden vorangegangenen Abschnitten wurden die Grundlagen und die Dynamik der Magnetisierung beschrieben. Es stellt sich nun die Frage, wie anhand der entstandenen Signale ein Bild erstellt werden kann. Bei der Magnetresonanztomographie handelt es sich um ein Schnittbildverfahren, bei welchem Messkörper in Schichten dargestellt werden können. Der (beispielsweise menschliche) Körper wird für die Messung in Voxel eingeteilt, welche quadratische Volumenelemente von bestimmter Kantenlänge repräsentieren (Abb 3.8).

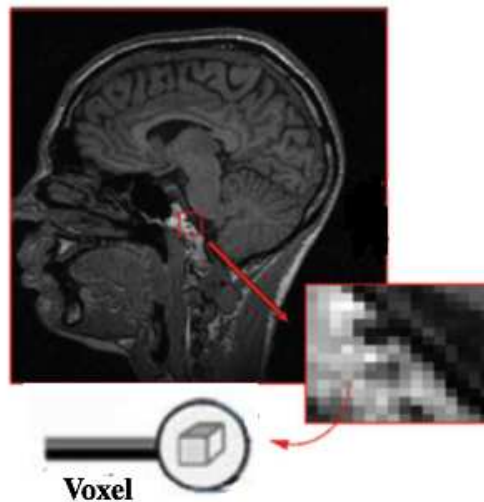


Abb. 3.8 Voxelschema. Jedem einzelnen Voxel wird eine Zahl zugeordnet, welche seiner Signalintensität entspricht. Die Auflösung eines Bildes wird durch die Größe dieser Voxel bestimmt (entnommen und modifiziert nach Schneider und Fink, 2013)

Der Mechanismus, welcher die einzelnen MR-Signale ihren jeweiligen anatomischen Strukturen zuordnet, beruht auf der Überlagerung eines variierenden schwachen magnetischen Feldgradienten auf das räumlich statische äußere magnetische Feld. Die Larmorfrequenzen der darin enthaltenen Protonen sind proportional zur Magnetfeldstärke, welche in der Raumrichtung proportional ansteigt. Indem eine bestimmte Frequenz des HF-Impulses gewählt wird, welche mit der Larmorfrequenz der gewünschten Schicht übereinstimmt, lässt sich selektiv eine bestimmte Schicht anregen.

Um nun die Raumposition zu analysieren, wird diese Schicht mehrmals hintereinander gemessen. Die Zeit zwischen den jeweiligen Messungen wird Repetitionszeit (*Time of repetition, TR*) genannt und beeinflusst den T1-Kontrast. Der T2-Kontrast wird durch die Zeit zwischen Anregung der Schicht und Messung des Signals, der sogenannten Echozeit (*Time to Echo, TE*) beeinflusst.

Es wird nun jedem einzelnen Voxel ein bestimmter Zahlenwert zugeordnet, der seine Signalintensität repräsentiert. Mittels einer sogenannten Fourier Transformation werden die jeweiligen Werte Graustufen zugeordnet und generieren ein strukturelles Bild. Diese Bildrekonstruktion ist mathematisch hoch komplex und weitere Ausführungen würden den Rahmen dieser Arbeit sprengen. Für weiterreichende Erläuterungen der Funktionsweise der

Magnetresonanztomographie wird an dieser Stelle auf die einschlägigen Lehrbücher verwiesen (u.a. Reiser et al. 2008, Weishaupt et al., 2014).

3.1.2 Funktionelle Magnetresonanztomographie

Wir haben nun das strukturelle Bildgebungsverfahren MRT kennengelernt. Die für diese Arbeit durchgeführte Studie befasst sich jedoch vorrangig mit der Darstellung von kognitiven Prozessen. Hierfür wurde die Methode der funktionellen Magnetresonanztomographie angewandt, welche erstmals durch Ogawa und Kwong Anfang 1990 beschrieben wurde (Kwong et al., 1992; Ogawa et al., 1990). Insbesondere der neurowissenschaftliche Bereich bietet seither diese Messmethode die Möglichkeit, sensorische, motorische und kognitive Prozesse in Form von Aktivierungen in den jeweiligen Hirnregionen abzubilden.

Allgemein besteht neben der Möglichkeit einer elektrophysiologischen Messung, welche die elektrischen bzw. magnetischen Aktivitäten der Nervenzellen detektiert, eine blutflusssensitive Messung, welche auf dem Prinzip der metabolischen Kopplung beruht (Jansen, 2004). Im Folgenden wird nur die blutflusssensitive Methode erläutert. Ihr liegt der Mechanismus zugrunde, dass regionale Stoffwechselveränderungen und somit Änderung der Durchblutung und der Sauerstoffsättigung des Blutes durch neuronale Aktivitäten ausgelöst wird. Oxygeniertes Hämoglobin (HbO_2) hat eine andere magnetische Eigenschaft als desoxygeniertes Hämoglobin (Hbr) und lässt sich dadurch in der fMRT differenziert darstellen (Pauling & Coryell, 1936). Dieser Mechanismus wird BOLD (*Blood Oxygen Level Dependent*)-Effekt genannt. Dem BOLD-Effekt liegt zugrunde, dass der Bedarf an HbO_2 lokal in genau den Hirnarealen erhöht ist, bei denen eine neuronale Aktivität, wie beispielsweise die Generierung von Worten, stattfindet. Hier verbrauchen die aktivierten Nervenzellen durch Generierung von Aktionspotenzialen Energie, welche in Form von Sauerstoffverbrauch bereitgestellt wird. Dabei kommt es zwar zunächst zu einer kurzfristigen Abnahme des HbO_2 im Blut, diese wird jedoch durch eine Erhöhung des lokalen zerebralen Blutflusses kompensiert und überwiegt schließlich sogar gegenüber dem Hbr (Klose, 1999).

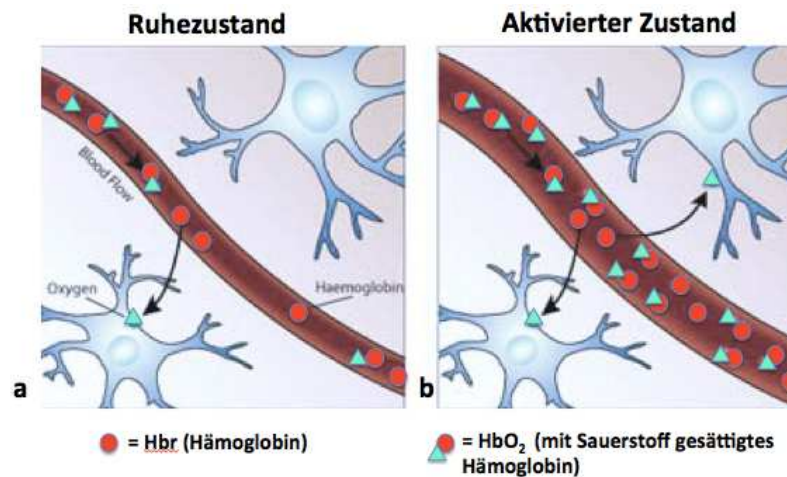


Abb. 3.9 BOLD-Effekt. Im Ruhezustand des zerebralen Blutflusses liegt unter Abgabe von Sauerstoff ein Gleichgewicht von Hbr und HbO₂ vor (a). Nach Stimulation der neuronalen Aktivität nehmen Blutfluss und Blutvolumen stark zu. Über metabolische Kopplung wird die Bereitstellung von HbO₂ erhöht, die Abgabe an das Nervengewebe jedoch nur gering erhöht, hierdurch entsteht das Signal im fMRT (b) (Abb. modifiziert nach Mette, 2014).

Um die hämodynamischen Antworten auf Stimmulationsereignisse in MR-Signale umzusetzen, macht man sich die verschiedenen magnetischen Eigenschaften des HbO₂ und des Hbr zunutze (Ogawa et al., 1990; Turner et al., 1991). Hbr ist paramagnetisch, was bedeutet, dass hier Magnetfeldinhomogenitäten vorliegen, die einen Signalzerfall beschleunigen. Im Gegensatz dazu hat das diamagnetische HbO₂ keinen Einfluss auf das Signal. Es liegt zudem noch in Folge der neuronalen Aktivität vermehrt vor und bewirkt eine relative Signalerhöhung. Eine T2*-gewichtete Aufnahme, beispielsweise die Echoplanarsequenz (Echo Planar Imaging, EPI) kann diese Verstärkung des Signals visualisieren. Erkennbar wird wie schon bei den strukturellen MRT-Aufnahmen eine Grauwertveränderung, welche auf dem BOLD-Kontrast basiert.

3.1.3 Analyse von fMRT-Daten

Bevor sich anhand der T2*-gewichteten EPI-Bilder eine zuverlässige Aussage über die Hirnaktivität machen lässt, müssen die Bilder verschiedenen Vorverarbeitungsschritten und einer statistischen Analyse unterzogen werden. Das Programm SPM 8 (*Statistical Parametric Mapping; Wellcome Department of Cognitive Neurology, London; <http://www.fil.ion.ucl.ac.uk/spm>*), welches die Programmiersprache MATLAB (*MathWorks, Natick, MA; <http://www.mathworks.com/>*) verwendet, ist auf die Auswertung der funktionellen Daten spezialisiert. Sie ist in drei aufeinanderfolgende Analyseschritte

eingeteilt: der räumlichen Vorverarbeitung, der statistischen Analyse und der statistischen Inferenz.

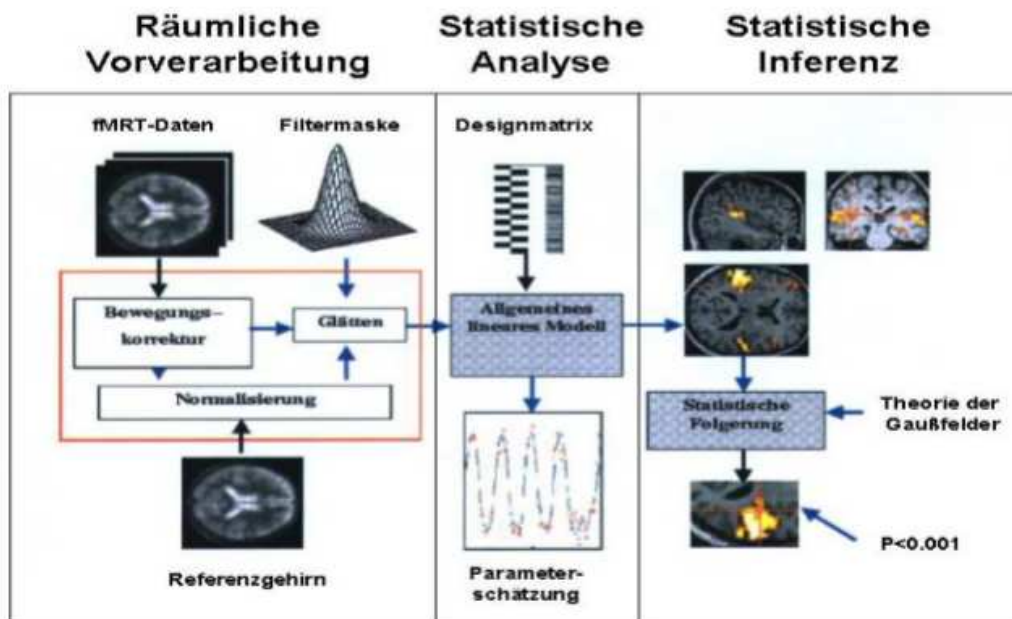


Abb. 3.10 Funktionelle Datenauswertung mit SPM 8 (Abb. entnommen aus Jansen, 2004)

Räumliche Vorverarbeitung

Bei der Bildaufnahme sind Bewegungsartefakte durch minimale Kopfbewegungen der gemessenen Personen selbst bei optimaler Compliance nicht vermeidbar. Diese werden in der räumlichen Vorverarbeitung korrigiert, was als *Realignment* bezeichnet wird. In Folge dessen wird die Aufnahme des Gehirns in dem Prozess der *Normalisierung* mit anderen Gehirnen vergleichbar gemacht. Abschließend wird noch eine räumliche Glättung, das *Smoothing*, durchgeführt.

Das Prinzip des *Realignment* basiert auf der Auswahl einer Referenzaufnahme aus den Bildern der Zeitserie. Alle übrigen Aufnahmen werden transformiert bzw. rotiert darauf abgebildet, um vorherige Translations- und Rotationsbewegungen rückgängig zu machen.

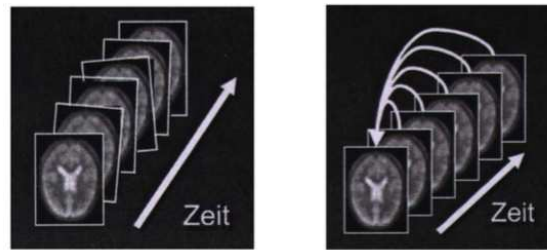


Abb. 3.11 Prinzip des Realignment (Abb. entnommen aus Jansen, 2004)

Der Schritt der Normalisierung hat zum Ziel, interindividuelle Variationen der Hirnstruktur und Hirngröße zu kompensieren. Hierzu wird die Probendenaufnahme auf ein standardisiertes Referenzgehirn angepasst. Dieser Vorgang der sogenannten Transformation besteht zunächst aus der Anpassung der Größe und Position und anschließend der Deformierungen des Gehirns (Jansen 2004). Das zu diesem Zweck von SPM verwandte Referenzgehirn ist das MNI-Gehirn (*Montreal Neurological Institute*), es wurde durch Mittelung mehrerer hundert Gehirnaufnahmen von gesunden Proband_innen entwickelt (Brett et al., 2002). Zusätzlich wird durch die Normalisierung bewirkt, dass die aktivierten Areale unterschiedlicher Studien verglichen werden können.

Als letzter Schritt der räumlichen Vorverarbeitung wird das *Smoothing* durchgeführt. Hiermit soll erreicht werden, dass zufällige Aktivierungen –sogenanntes Rauschen - reduziert wird. Es trägt somit zu einer Verbesserung des Signal-zu-Rausch Verhältnisses (*SNR, Signal-to-noise-ratio*) bei. Als Filter zur Differenzierung zwischen tatsächlichen Signalen und zufälligen Aktivierungen verwendet man eine Gaußfunktion, welche Fehler einer Normalverteilung anpasst und somit die Validität der darauf folgenden statistischen Analyse verbessert.

Zudem bewirkt die räumliche Glättung eine Weichzeichnung der aktivierten Areale. Benachbarte Voxel von diesen aktivierten Arealen werden dabei in ihren Grauwerten verglichen und gemittelt. Die aktivierten Areale dehnen sich dabei aus und funktionell homologe Areale verschiedener Proband_innen können besser aufeinander abgebildet werden. Die Ausdehnung hat allerdings zum Nachteil, dass dabei vorher bestandene Aktivierungsgrenzen verbunden werden.

Statistische Analyse

Mit der statistischen Analyse soll eine den Messungen vorangestellte Hypothese durch den zeitlichen Verlauf der Aktivierungssignale in einem Voxel überprüft werden. Die Korrelation wird mit einem genau spezifizierten Modell verglichen, das von den verschiedenen Konditionen des experimentellen Studiendesigns abgeleitet wird (Schneider und Fink, 2013). Für jede einzelne Messung der jeweiligen Paradigmen wird dann abgeglichen, welche Bedingung zu ihrem Zeitpunkt gegeben war. So wird beispielsweise bei der Aufgabe der subliminalen Verarbeitung von Emotionen zeitlich genau spezifiziert, wann eine Emotion, ein neutrales Gesicht oder eine Ruhebedingung gezeigt wurde. Grundlage hierfür bildet das allgemeine lineare Modell (*GLM, General Linear Model*), das verschiedene parametrische Verfahren der Zeitreihenanalyse in einem Modell zusammenfasst (Friston, 2007). Die in jedem Voxel gemessene Zeitreihe der Signalintensitäten entspricht hierbei einer statistisch abhängigen Variable, die Hypothese wird durch unabhängige Variablen gekennzeichnet (Jansen, 2004). Es entsteht eine Variablenmatrix des Experimentes, welche als sogenannte Designmatrix die Parameter jedes Voxels schätzen kann (Parameterschätzung). Um Mittelwertschwankungen auszugleichen, kann als eine Variante des allgemeinen linearen Modells mittels einer Regressionsanalyse die Korrelation der unabhängigen Variablen mit den gemessenen Signalen berechnet werden.

In einer t-Karte werden die gemittelten Aktivierungsintensitäten der Voxel statistisch auf signifikante Unterschiede verglichen und mit einer Grauwertkodierung dargestellt.

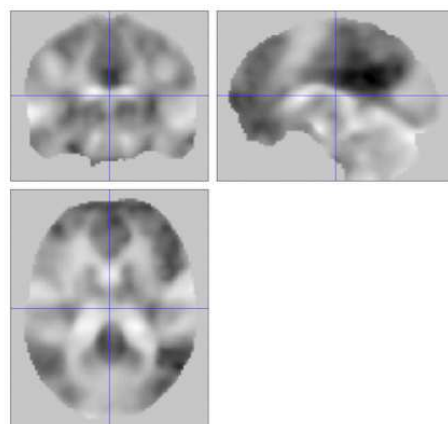


Abb. 3.12 T-Karte am Beispiel der Aktivierungskarte der Wortgenerierungsaufgabe (Abb. der Autorin der Dissertation)

Statistische Inferenz

Die statistische Analyse ermöglicht es nun zwar statistische Folgerungen zu treffen, es sollte dabei jedoch zunächst vermieden werden, Fehlinterpretationen bei einem erwarteten Signifikanzniveau von 5% zu treffen. SPM führt daher eine Überprüfung auf statistische Signifikanz eines jeden einzelnen Voxels durch. Da bei der fMRT vor allem nach Vorverarbeitungsschritten wie der räumlichen Glättung eine räumliche Korrelation der Voxel vorliegt, wird hier die Theorie der Gauß-Felder (Worsley, 1996) angewandt. Hier werden nur diejenigen Bildpunkte als aktiv berücksichtigt, welche einen bestimmten Schwellenwert (p-Wert) überschreiten. Der p-Wert kann die voxelweise Signifikanz (signifikante Intensitätsunterschiede einzelner Voxel) und die räumliche Ausdehnung zusammenhängender Aktivierungsfelder (signifikante Ausdehnung) mit einberechnen. Der Gedanke dahinter ist einerseits, dass ‚echte‘ Aktivierungen in der Regel mehrere benachbarte Voxel (Cluster) betreffen und andererseits, dass je größer ein solches Cluster ist, die Wahrscheinlichkeit geringer wird, dass es trotz Schwellenwertes nur zufällig einberechnet wird.

Die bisherigen Ergebnisse können auf die erhobenen Daten einzelner Proband_innen angewandt werden. Um wie in dieser Studie eine Gruppe von Proband_innen analysieren zu können, muss noch eine sogenannte Second-Level-Analyse durchgeführt werden. Sie erstellt ein gemitteltes Aktivierungsmuster der gesamten Gruppe und lässt Aussagen über das gesamte Kollektiv zu.

3.1.4 Design eines fMRT-Experiments

In der Modellierung eines fMRT Experimentes müssen viele Faktoren beachtet werden, welche die akquirierte Datenqualität entscheidend beeinflussen können. Während ganz allgemein die Qualität und Art des Stimulusmaterials einen wichtigen Teil des Designs bildet, gilt es bei der fMRT spezifisch auch auf die Zeitverzögerung der hämodynamischen Antwort zu achten. Ein BOLD-Signal benötigt etwa 4-6 Sekunden, um seine maximale Signalintensität zu erreichen und ist erst nach etwa 20 Sekunden wieder in seinen Ursprungszustand zurückgekehrt. Im Gegensatz dazu erreichen neuronale Antworten bereits nach wenigen Sekunden ihr Maximum. Daher muss eine optimale Aktivierungsintensität und- dauer bei Stimuluspräsentation vorliegen.

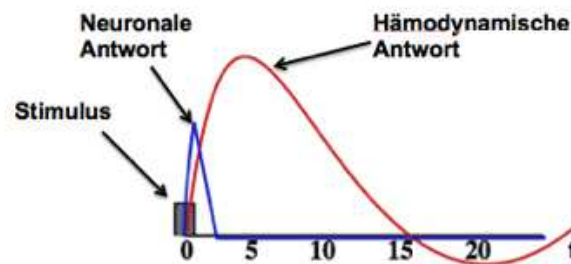


Abb. 3.13 Zeitverlauf einer Hämodynamischen Antwort und einer Neuronalen Antwort nach einer Stimulipräsentation (modifiziert nach Hermann, 2014)

In dieser Studie wurde daher ein Blockdesign (sogenanntes Boxcar-Design) verwendet, weil die Stimuli hiermit entweder länger andauernd oder mehrfach aufeinander folgend präsentiert werden können und sich die hämodynamischen Antworten zu einer ausgeprägten einzelnen Aktivierung ausdehnen.

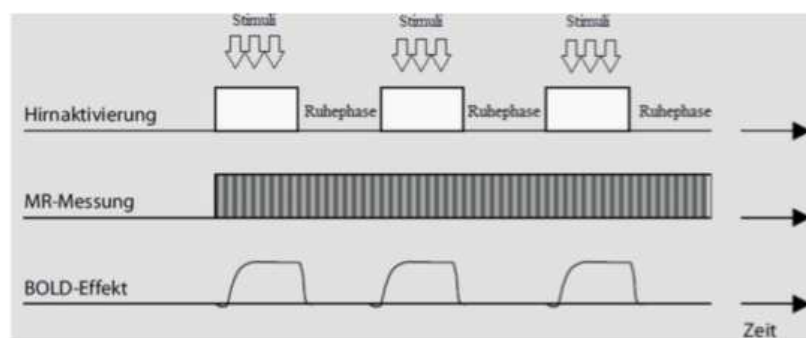


Abb. 3.14 Häemodynamische Antwort und Messung in einem Block-Design (modifiziert nach Weishaupt et al., 2014)

Prinzipiell beinhaltet das Konzept des Boxcar-Designs zwei zuvor festgelegte alternierende Konditionen, welche eine ON-Phase (Stimuli werden dargeboten) einer OFF-Phase (keine Stimulipräsentation) gegenüberstellen. Diese Phasen verlaufen in einem festen Zeitschema, welches unabhängig von den Proband_innen abläuft. Um bei der Erstellung des Aktivitätsmusters konfundierende Variablen bestmöglich zu reduzieren, wird der Aktivierungsdurchschnitt beider Phasen mehrfach miteinander verglichen und Hintergrundaktivitäten bzw. auditorisches Input subtrahiert. Dadurch wird ein besseres Signal-zu-Rausch-Verhältnis erreicht und nur die Aktivierungsunterschiede, welche auf die sich unterscheidende Komponente der Phasen zurückzuführen sind, bleiben übrig (Jansen, 2004).

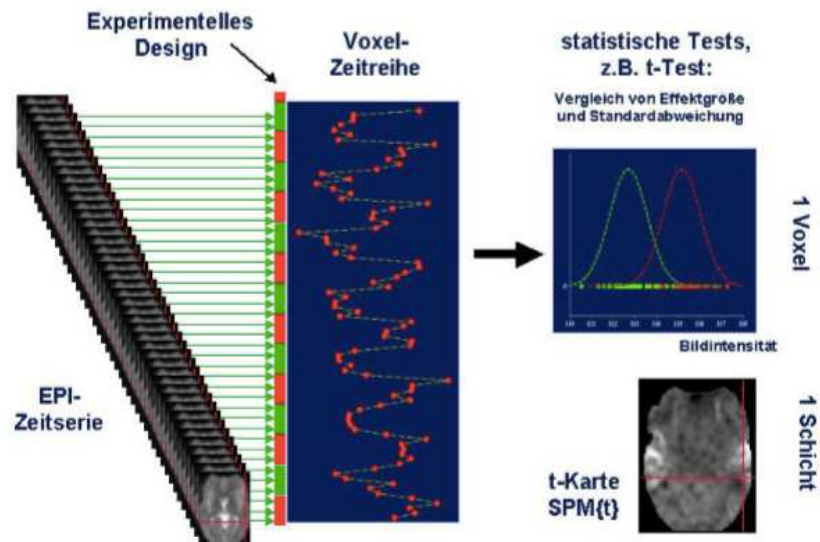


Abb. 3.15 Boxcar Design. Experimentelles Design der beiden Konditionen der Aktivierungs- (rot) und Kontrollphase (grün). Die Datenaufnahme erfolgt über die gesamte Messzeit hinweg. Für jeden Voxel werden die Mittelwerte der beiden Phasen statistisch miteinander verglichen und in ihrer Gesamtheit als T-Karte dargestellt (Abb. entnommen aus Jansen, 2004)

Im Vergleich zu anderen experimentellen Designs hat das Blockdesign den Vorteil, dass die Analyse und Interpretation der gewonnenen Daten simpel ist und von einer hohen statistischen Aussagekraft geprägt sind (Schneider und Fink, 2013). Allerdings kann sich bei den Proband_innen eine Art Erwartungshaltung auf die jeweilige Bedingung in den Darbietungsphasen herausbilden. Gleichzeitig ist durch die vorherige Festlegung des Ablaufs ein Randomisieren oder auch eine Variation in der Ereignishäufigkeit der Stimuli erschwert.

3.2 Proband_innendaten

10 weibliche und 10 männliche gesunde Proband_innen im Alter von 20 bis 28 Jahren (Durchschnittsalter: $23,7 \pm 2,3$ Jahre) wurden im Rahmen der Studie "Wirkungen von genetischen Risikovarianten für psychische Störungen auf Gehirnstruktur und -funktion bei Gesunden und Erkrankten" in zwei Sitzungen im Abstand von durchschnittlich $28,4 \pm 6,0$ Tagen zur Überprüfung der Reliabilität der Paradigmen gemessen. Als 'gesund' wurden solche Proband_innen bezeichnet, welche keine psychischen Erkrankungen gemäß der ICD 10 Klassifikation oder neurologische Störungen aufwiesen. Die Proband_innen waren deutsche Muttersprachler_innen und hatten mindestens die allgemeine Hochschulreife als vergleichbares Bildungsniveau erreicht. Um eventuelle zyklusabhängige Schwankungen der

Hirnaktivierung zu minimieren, wurden nur Frauen ausgewählt, welche hormonelle Kontrazeptiva zum Zeitpunkt der Messungen einnahmen. Bei der WG-Aufgabe wurden nur 19 Probanden bei der Auswertung miteinberechnet, weil bei einem Probanden die Daten aus technischen Gründen nicht vollständig erhoben werden konnten.

Im Voraus wurden die Proband_innen mithilfe eines Aufklärungsbogens über die Studie informiert und willigten in einer Probandeneinwilligungserklärung zu der Teilnahme daran ein (siehe Anhang). Desweiteren erhielt jede_r Proband_in einen Erklärungsbogen zu den einzelnen fMRT Paradigmen und ihrer Durchführung mit Beispielaufgaben. Vor der ersten Messung im Magnetresonanztomographen übte jede_r Proband_in die Ausführung der Aufgaben schließlich an einem Rechner, um das Verständnis zu garantieren.

Die Ethikkommission des Fachbereichs Humanmedizin der Philipps- Universität Marburg genehmigte diese Studie.

3.3 fMRT-Parameter

Die MRT-Daten wurden mit einem 3-Tesla Siemens Magnetom Tim Trio Scanner der Philipps- Universität Marburg aufgenommen. Zu Beginn der ersten Messung wurde eine hochaufgelöste anatomische T1-Aufnahme eines_r jeden Proband_in angefertigt und im Anschluss die funktionellen Bilder zu allen sechs Paradigmen in beiden Messungen erstellt. Die Aufnahme dieser funktionellen Bilder erfolgte mit einer T2*-gewichteten Echoplanarsequenz, die sensitiv für den BOLD-Kontrast war (64x 64 Matrix, Field of View (FoV) = 224 mm, 40 Schichten, 3,5 mm Schichtdicke, TR= 2500 ms, TE= 30 ms, Pulswinkel 90°, die Aufnahmesequenzen erfolgten interleaved und ohne gap). Die Schichten wurden parallel zur AC-PC Linie (einer imaginären Verbindungslinie zwischen anteriorer und posteriorer Kommissur) aufgenommen und transversal positioniert. Beim Paradigma der Wortgenerierung wurden 130 Bilder aufgenommen, beim Paradigma der subliminalen Präsentation emotionaler Stimuli 140 Bilder

3.4 Experimentelles Design

Für die Messung der Hirnaktivierung der gesunden Proband_innen wurden Stimuli in sechs unterschiedlichen Paradigmen mit der Software "Presentation" dargeboten (Neurobehavioral System Inc., San Francisco, CA; <http://www.neurobs.com/>). Diese Paradigmen wurden sowohl bei jeder_m Proband_in als auch bei der ersten und zweiten Messung in der gleichen Reihenfolge gezeigt. Die vor dem ersten Ablauf durchgeführte anatomische Messung dauerte 15 Minuten. Im Anschluss fand die funktionelle Datenaufnahme statt, welche sowohl bei erster als auch bei zweiter Messung (weitere) 33 Minuten dauerte. Insgesamt ergab sich daraus bei dem ersten Ablauf inklusive Pausen eine Gesamtzeit von 54 Minuten und bei der zweiten Messung 39 Minuten. Den Proband_innen wurden vor Beginn der Messungen Kopfhörer aufgesetzt, über welche die Instruktionen für die einzelnen Paradigmen eingespielt wurden. An diese Kopfhörer wurde ein 40 dB geräuschreduzierendes MRT-taugliches Mikrophonsystem mit einem anpassungsfähigen Filter für MRT-Geräusche (Optoacoustics Ltd.) angebracht, welches die generierten Worte bei der 'Verbal Fluency' Aufgabe aufnahm und als Kommunikationsmittel diente. Am rechten Oberschenkel wurde eine Konsole befestigt, an welcher die Proband_innen bei zwei der Paradigmen (Gedächtnisaufgaben) Tasten bedienen sollten.

Die sechs Paradigmen dieser Studie messen die kognitiven Domänen, welche bei Personen mit psychischen und neurologischen Erkrankungen als eingeschränkt beobachtet wurden. Hierzu gehören eine *Resting State* Aufgabe, Aufgaben zum Arbeitsgedächtnis und episodischem Gedächtnis, eine Wortgenerierungsaufgabe und eine Aufgabe zur Emotionsverarbeitung. Der Hauptgegenstand dieser Arbeit, das Wortgenerierungsparadigma (WG-Aufgabe, *Verbal Fluency*) und das Paradigma der Subliminalen Präsentation emotionaler Stimuli (SPES-Aufgabe, *Subliminal Priming*) werden im Anschluss im Detail erläutert. Die Zielsetzung der anderen Paradigmen soll jedoch an dieser Stelle zusammenfassend dargestellt werden.

Das erste Paradigma, der sogenannte Ruhezustand (*Resting State*), ermöglicht es, sogenannte Ruhenetzwerke der Proband_innen zu untersuchen. Hierzu werden die Proband_innen gebeten, für 5 Minuten mit geschlossenen Augen ruhig zu liegen und keine Aufgaben auszuführen.

Das dritte und vierte Paradigma dient der Untersuchung des episodischen Gedächtnisses, welches in zwei Paradigmen unterteilt ist, um sowohl die Einspeicherung (*Faces Encoding*) als auch das Abrufen (*Faces Retrieval*) von gelernten Inhalten messen zu können. Den Proband_innen wurden in dem *Faces Encoding* Paradigma 30 neutrale Gesichter gezeigt, deren Geschlecht sie mit Tastendruck auf der Konsole anzeigen sollten. In dem *Faces Retrieval* Paradigma wurden zwei Gesichter nebeneinander präsentiert, wovon eines bereits aus dem *Faces Encoding* Paradigma bekannt war. Dieses sollen die Proband_innen mittels linkem oder rechten Tastendruck identifizieren.

Das letzte Paradigma hatte das Ziel, das Arbeitsgedächtnis (*N-Back*) zu untersuchen. Dieses Paradigma bestand aus einem Blockdesign. Der zu variierende Faktor war die Arbeitsgedächtnisbelastung. In einem ersten Block (*0-Back*) wurden die Proband_innen gebeten, eine Taste bei dem Buchstaben 'X' zu drücken. In den folgenden Blöcken *1-Back* und *3-Back* sollten sie eine Taste drücken, sobald der aktuell präsentierte Buchstabe identisch mit dem zuletzt, zuvorletzt und zu drittletzt präsentierten Buchstaben war.

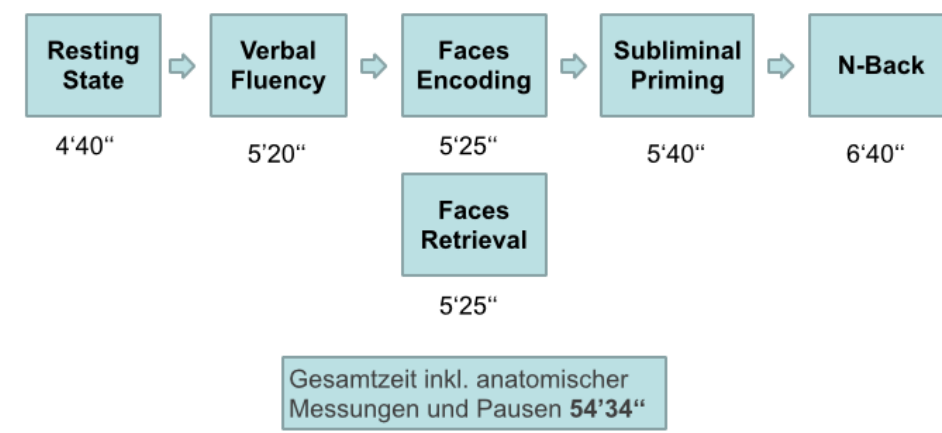


Abb. 3.16 Übersicht der Paradigmen

3.4.1 Wortgenerierungs- Aufgabe

Die WG-Aufgabe bestand aus einem Blockdesign mit zwei alternierenden Konditionen: Nach der Präsentation einer deutschen Kategorie (bspw. Tiere) wurden die Proband_innen dazu aufgefordert, so viele dazugehörige Unterbegriffe wie möglich zu bilden (bspw. Katze, Vogel, Käfer). Die einzelnen Wortgenerierungsblöcke wurden im Wechsel mit Ruheblöcken

angezeigt, bei welchen die Proband_innen auf die Instruktion 'Ruhe' folgend keine Wörter generieren sollten. (*Instruktion: ‚Zählen Sie während ‚+‘ Wörter zu den Kategorien auf. Bei ‚#‘ seien sie bitte ruhig.‘*)

Insgesamt wurden pro Messung je 10 Blöcke der beiden Konditionen Ruhe und Wortgenerierung präsentiert. Die Kategorien der Wortgenerierung sind in Tabelle 3.1 dargestellt.

Erste Messung	Zweite Messung
Tiere	Städte
Obst	Farbe
Fahrzeuge	Getränke
Blumen	Materialien
Hobbys	Gemüse
Lebensmittel	Verkehrsschilder
Kleidungsstücke	Musikinstrumente
Sportarten	Gewürze
Berufe	Spielzeug
Möbelstücke	Bauwerke

Tab. 3.1: Kategorien der Aufgabe Wortgenerierung der ersten und zweiten Messung

Zu Beginn eines jeden Wortgenerierungsblockes wurde für 3 Sekunden eine Kategorie eingeblendet, welcher für 12 Sekunden das '+' Zeichen folgte. Bei dem Ruheblock wurde zu Beginn 3 Sekunden lang das Wort 'RUHE' eingeblendet, hier folgte 12 Sekunden lang das Zeichen '#'. Wiederholungen von Wörtern oder das Aufzählen verschiedener grammatikalischer Formen der Wörter waren nicht erlaubt. Die Kategorien wurden in weißer Farbe auf einem schwarzen Hintergrund dargestellt. Die Wortgenerierung der Proband_innen wurde mit einem Scanner kompatiblen Mikrophon und der Audacity 1.2.6 Software aufgezeichnet.

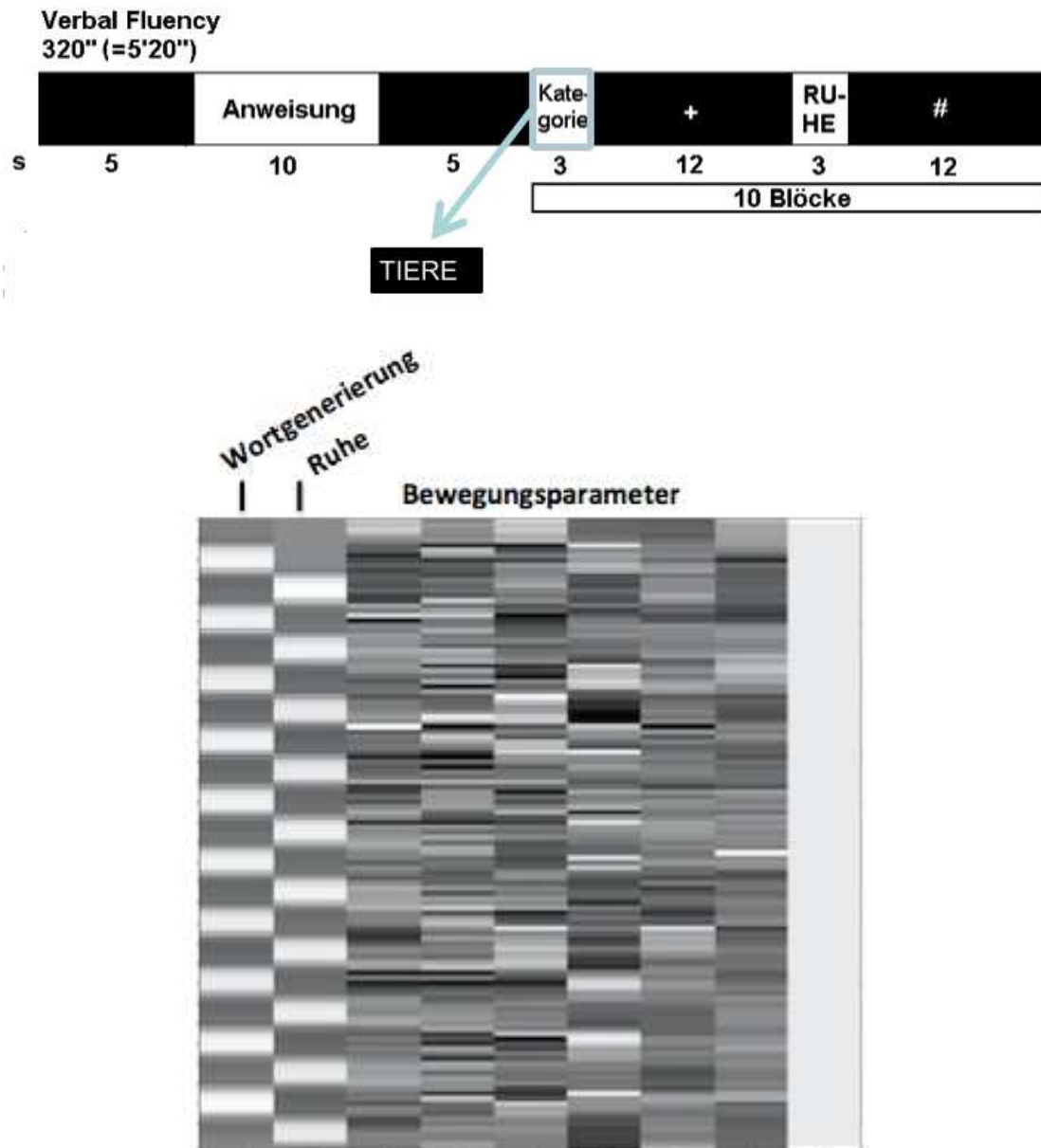


Abb. 3.17 Design und Designmatrix der WG- Aufgabe. Spalte 1 zeigt die Bedingung der Wortgenerierung, Spalte 2 die Ruhebedingung, die Spalten 3-8 zeigen die modellierten Bewegungsparameter, Spalte 9 die Modellierung der Mittelwertkorrektur.

3.4.2. Aufgabe der subliminalen Präsentation emotionaler Stimuli (SPES-Aufgabe)

Dieses Paradigma untersuchte die subliminale, also unterbewusste, Verarbeitung von emotionalen Stimuli bei gesunden Proband_innen.

Die Proband_innen wurden angewiesen, über den Zeitraum von 5 Minuten die gezeigten Gesichter zu betrachten. Darüber hinaus erhielten sie keine Informationen über die Aufgabe. Zu Beginn der Aufgabe wurde dem/r Proband_in für 10 Sekunden über den Kopfhörer die Instruktion „Bitte betrachten sie die folgenden Gesichter! (kein Tastendruck)“

eingespielt. Daraufhin folgte ein Blockdesign mit zwei Konditionen: eine Ruhekondition mit einem grauen Dia für 9 Sekunden gefolgt von einer Wahrnehmungskondition mit der subliminalen Präsentation affektiver Stimuli für weitere 31 Sekunden. Diese affektiven Stimuli bestanden aus monochromen Bildern mit neutralen, traurigen, ängstlichen und glücklichen Gesichtsausdrücken von 10 ausgewählten Schauspieler_innen verschiedener Altersgruppen, die bereits im Rahmen von anderen Studien bei gesunden Proband_innen und psychiatrischen Patient_innen auf ihre Validität untersucht worden sind (Suslow und Dannlowski, 2006).

Damit die Proband_innen die gezeigten Emotionen jedoch nicht direkt, sondern unterbewusst wahrnahmen, wurde pro Gesicht für 33 ms der jeweilige Gesichtsausdruck gezeigt und für weitere 467 ms mit dem neutralen Gesichtsausdruck derselben Person maskiert. Bei dem Block der neutralen Emotion wurde das Gesicht für die ersten 33 ms gespiegelt dargestellt, damit die Kontinuität im Bildwechsel erhalten blieb.

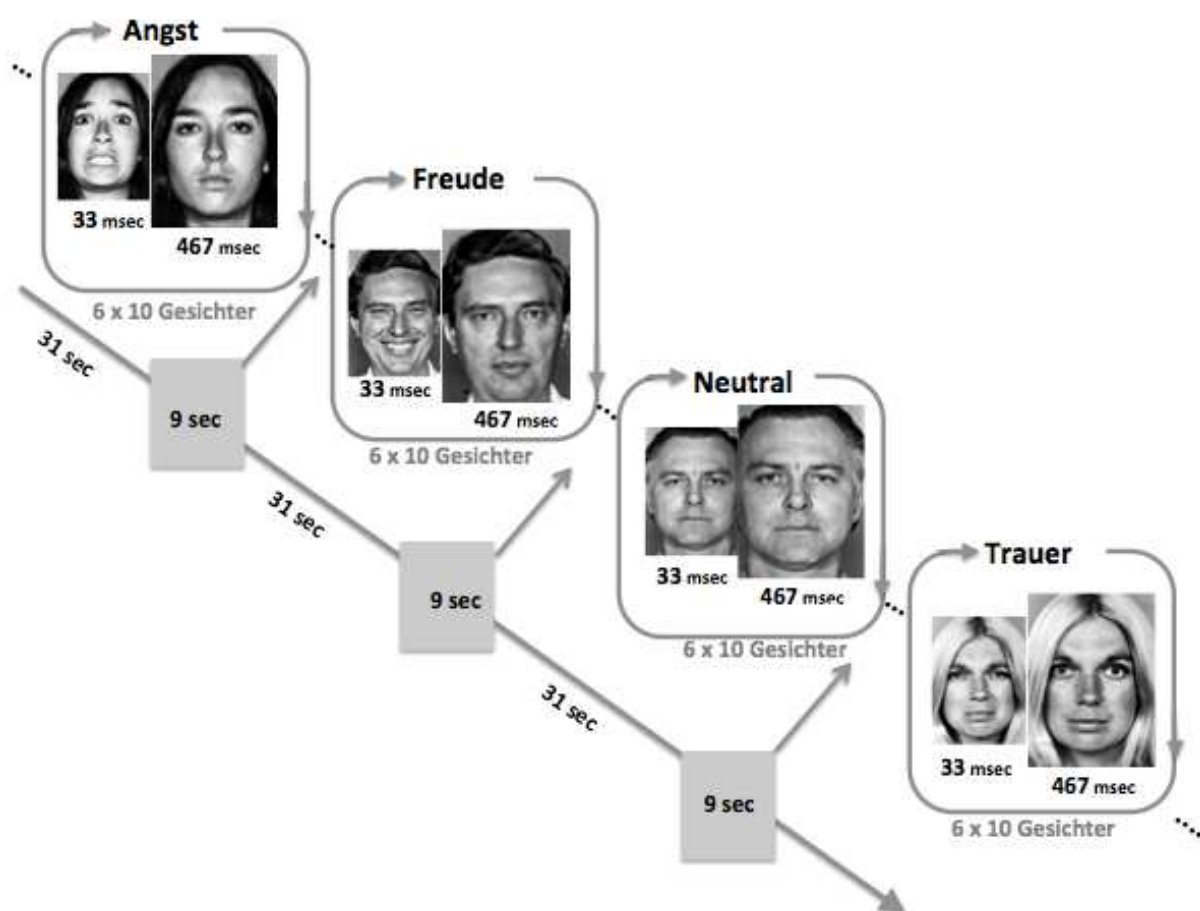


Abb. 3.18: Präsentation der Blöcke mit affektiven Stimuli

Jeder Block mit affektiven Stimuli bestand aus Gesichtern derselben 10 Personen, die 6 Mal randomisiert gezeigt wurden. Nach jedem Block wurde das graue Dia für 9 Sekunden gezeigt. Es gab insgesamt 8 Blöcke mit affektiven Stimuli. Die Präsentation der Blöcke ist in Abbildung 3.18 dargestellt. Am Ende wurde nochmals für 30 Sekunden das graue Dia gezeigt. Die zweite Messung war identisch mit der ersten. Alle Proband_innen negierten nach dem Ablauf der zweiten Messung, bewusst Gesichter mit affektiven Expressionen wahrgenommen zu haben.

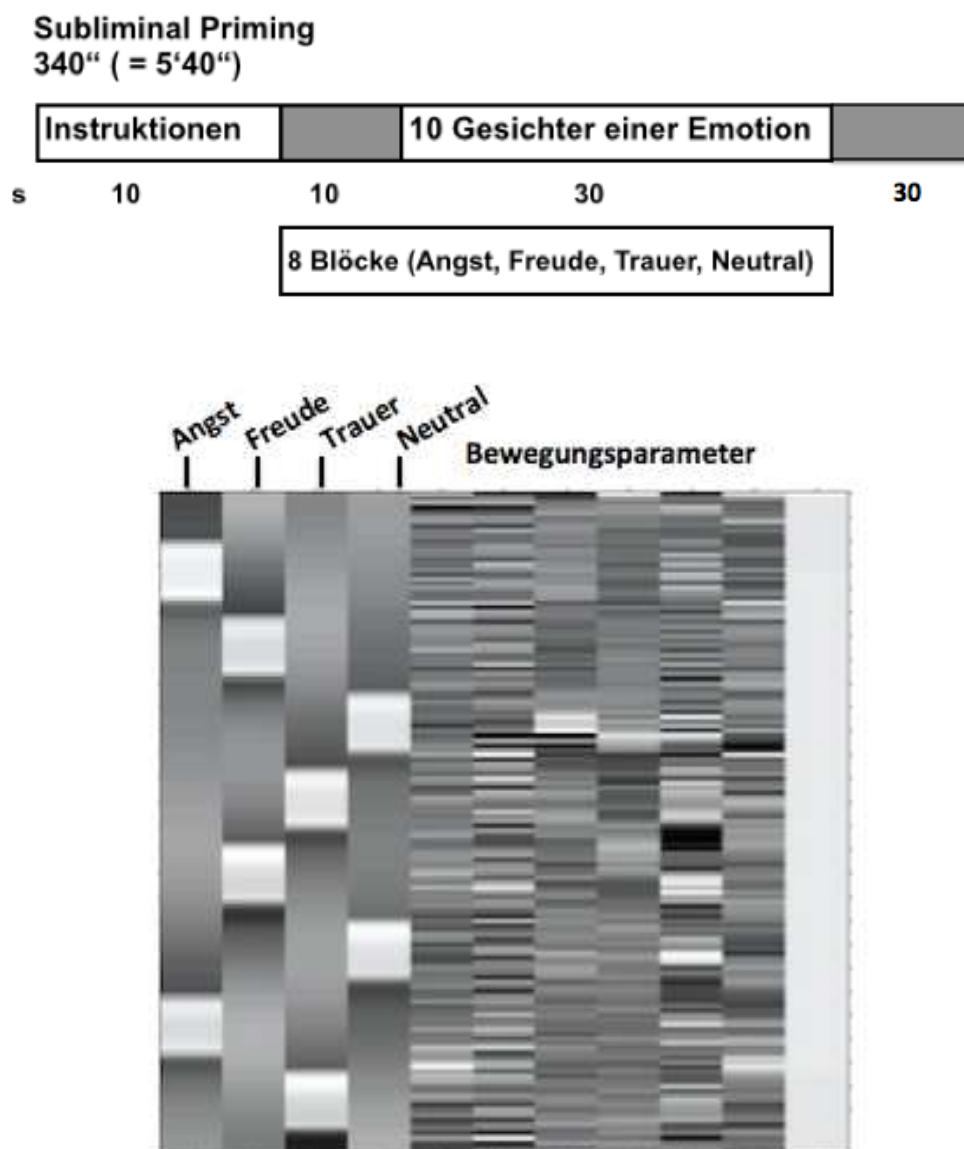


Abb. 3.19 Design und Designmatrix der SPES-Aufgabe. Spalte 1 zeigt die Bedingung der Emotion Angst, Spalte 2 die Bedingung der Emotion Freude, Spalte 3 die Bedingung der Emotion Trauer, Spalte 4 die Bedingung der neutralen Gesichtspräsentation. Die Spalten 5-10 zeigen die modellierten Bewegungssparameter, Spalte 11 die Modellierung der Mittelwertkorrektur.

3.5 Funktionelle Datenauswertung

3.5.1 Vorverarbeitung

Mit dem Ziel der optimalen Datenqualität wurden in der WG-Aufgabe 6 Bilder verworfen, so dass die ersten 15 Sekunden inklusive Instruktion nicht mit ausgewertet wurden und sich eine Gesamtanzahl von 124 Bildern ergab. Bei der SPES-Aufgabe wurden ebenso 6 Bilder verworfen, also insgesamt 134 Bilder erhalten. Auf diese Weise konnte zudem sichergestellt werden, dass das MR-Signal seinen Gleichgewichtszustand in den zur Analyse verwandten Bildern erreicht hatte (Jansen, 2004).

Zunächst wurde mittels einer Transformation der verbleibenden Bilderserie auf das erste Bild eine Bewegungskorrektur durchgeführt. Für jede_n Proband_in und Messung konnte dann aus diesen bewegungskorrigierten Bildern durch Mittelung ein funktionelles Bild eines jeden Paradigmas erstellt werden. Mit dem von SPM 8 verwendeten Standardgehirn (MNI-Template) wurden die funktionellen Bilder im Anschluss normalisiert, wodurch alle eine Voxelgröße von $2 \times 2 \times 2 \text{ mm}^3$ erreichten. Der letzte Schritt der räumlichen Vorverarbeitung bestand aus der Glättung der normalisierten, funktionellen Bilder mit einem isotropen Gauß'schen Filter (Halbwertsbreite 6 mm).

3.5.2 Statistische Analyse und Inferenz

Die Aktivierungs- und Kontrollbedingungen beider Paradigmen wurden mit dem zuvor erläuterten Boxcar-Design (siehe Kapitel 3.1.4) modelliert. Die Boxcar-Funktion wurde zur Abbildung des zu erwartenden Verlaufs des BOLD-Signals mit der HRF gefiltert (Friston et al., 1994). Es folgte für jeden einzelnen Voxel entsprechend einem T-Test ein statistischer Vergleich auf Unterschiede in den jeweiligen Bedingungen der beiden Messungen. Hierzu wurden *con*-Bilder (*contrast-images*) erstellt, welche einen Vergleich der Aktivierungsstärke der einzelnen Bedingungen lieferten. Um die Varianz der Messung berücksichtigen zu können, wurden diese *con*-Bilder durch die Varianz geteilt und somit T-Bilder erstellt.

3.5.3 Lokalisierung aktivierter Hirnregionen

Die Lokalisierung der in den Aktivierungskarten aktivierten Hirnregionen der jeweiligen Paradigmen und Messungen erfolgte mit der standardisierten, in SPM8 integrierten *Anatomy Toolbox*.

3.6 Messung der Reliabilität

In der Einführung in das Thema Reliabilität (s. dazu Kapitel 2.2) wird darauf hingewiesen, dass bei der Messung der MR-Signale grundsätzlich von unterschiedlichen Messergebnissen ausgegangen werden muss. Dies liegt vor allem an der nur bis zu einem gewissen Grad möglichen Minimierung der Geräuschkulisse, welche stets über auditorische Reize einen Einfluss auf die Gehirnaktivität und somit dem nur schwachen MR-Signal haben kann. Zudem sind Änderungen in der BOLD-Signalstärke zu beachten, welche von vielen weiteren, beispielsweise psychosomatischen Reaktionen, abhängig sind. Wie diese Tatsache bei der Evaluation der Reliabilität der Messergebnisse zum Tragen kommt, soll im Folgenden vorgestellt werden. Dabei wird zunächst allgemein auf die Methoden der Reliabilitätsmessung eingegangen und im Anschluss daran speziell auf die in dieser Arbeit gewählte Methode der Reliabilitätsmessung.

3.6.1 Methoden der Reliabilitätsmessung

In der Reliabilitätsmessung wurden viele unterschiedliche Methoden entwickelt, die je nach dem Aspekt der Ergebnisse, welcher über einen längeren Zeitraum Stabilität zeigen soll, ausgewählt werden können. Grundsätzlich können die Reliabilitätsberechnungen drei übergeordneten Ergebnisaspekten in der fMRT zugeschrieben werden: Die Stabilität signifikanter Voxel, die Stabilität der Aktivierungsstärke bestimmter Cluster und die Stabilität der Voxelaktivität über das gesamte Gehirn (Bennett und Miller, 2010).

Abgesehen von diesen Methoden kann die Reliabilität auch mit der Voxelzählung evaluiert werden. Berechnet wird hier die absolute Anzahl der aktivierten Voxel in der Test- und Retest-Messung. Diese Methode hat allerdings das große Defizit, dass nicht die Aktivierungsstärke der aktivierten Voxel beachtet wird, sie werden nur als „aktiv“ oder „nicht aktiv“ klassifiziert. Als eine Methode, welche diese Signifikanz beachtet, soll an dieser Stelle die sogenannte ‚*Cluster Overlap*‘ – Methode (*Cluster-Überlappungsmethode*)

vorgestellt werden. Als *Cluster* werden zusammenhängende, aktiv dargestellte Hirnregionen beschrieben. Im Zentrum dieser Messmethode steht hier ein festgelegter Schwellenwert, der angibt, ab wann *Cluster* für die Evaluation signifikant aktiviert werden. Die *Cluster Overlap*- Methode geht also der Frage nach, wie stabil die signifikanten *Cluster* bei demselben Schwellenwert in einer wiederholten Messung wieder aktiviert sind. Eine häufig verwendete Kenngröße zum Vergleich ist der *Dice*-Koeffizient (Rombouts et al., 1997).

$$R_{overlap} = \frac{2(V_{overlap})}{(V_1 + V_2)} \quad (3.2)$$

Der *Dice*-Koeffizient ist ein einfaches Maß für die Überlappung aktivierter Hirnregionen aus zwei Messungen. V_1 und V_2 stehen hier für die durchschnittliche Anzahl signifikant aktivierter *Cluster* der jeweiligen Messungen und $V_{overlap}$ für die Überlappung beider Cluster. Sind die Überlappungen mit den jeweiligen Messungen identisch, dann entspricht der Wert idealerweise 1. Je weniger Überlappungen allerdings vorhanden sind, desto geringer wird der Wert $R_{overlap}$ (Specht et al., 2003).

Ein dem *Dice*-Koeffizienten ähnliches Gütemaß zum Vergleich der Ergebnisse ist der *Jaccard*-Index.

$$R_{overlap} = \frac{V_{overlap}}{(V_1 + V_2 - V_{overlap})} \quad (3.3)$$

Der *Jaccard*-Index gibt wie der *Dice*-Koeffizient die Überlappungen der signifikant aktivierten *Cluster* an (Bennet und Miller, 2010). Hierbei wird die Anzahl der sich überlappenden *Cluster* durch die absolute Anzahl einzelner signifikant aktivierter Voxel in beiden Sitzungen geteilt.

Auch der *Jaccard*-Index zeigt hohe Übereinstimmungen der *Cluster* Überlappungen bei einem Wert von 1.0. Ein Wert von 0.0 würde bedeuten, dass keiner der signifikant aktivierten Voxel des Test- und Retest-Messungen übereinstimmen. Der Unterschied zwischen den beiden Kenngrößen ist die Art der Normierung, um absolut vergleichbare Messzahlen zu erhalten, da bei dem *Dice*-Koeffizienten die Überlappung der *Cluster* doppelt gewichtet werden (Bennet und Miller, 2010). Beide Kenngrößen können graphisch als farbig markierte Hirnregionen in Bildern dargestellt werden (Abb. 3.20).

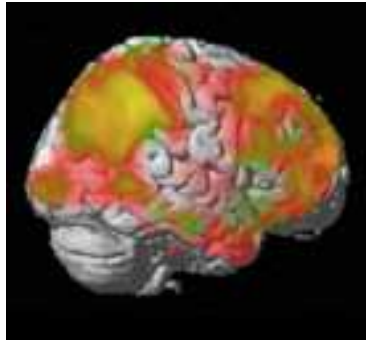


Abb. 3.20 Visualisierung der Cluster-Überlappungsmethode. Rote Regionen zeigen Aktivierungen der ersten Messung, grüne Regionen Aktivierungen der zweiten Messung und gelbe Regionen repräsentieren die sich überlappenden Aktivierungen beider Messungen der Proband_innen während der WG-Aufgabe dieser Studie (Abb. der Autorin der Dissertation).

Die *Jaccard* und die *Dice*-Methode sind maßgeblich von dem statistischen Schwellenwert abhängig, der definiert, welche Regionen als signifikant aktiv betrachtet werden und von der auswertenden Person festgelegt wird. Daraus ergibt sich das Problem, dass bei hohen Schwellenwerten weniger Bereiche aktiviert sind als bei niedrigen Schwellenwerten. Somit wird die Vergleichbarkeit steuerbar, da bei liberalen Schwellenwerten mehr zu vergleichende Regionen der Kenngrößenberechnung zur Verfügung stehen. Eine weitere Schwäche dieser Berechnung ist die Tatsache, dass selbst wenn beide aktivierten *Cluster* stabil wären, es keine Informationen darüber gäbe, ob jede_r Proband_in konsistent innerhalb der Gruppe Aktivierungen zeigt. Die Gruppenaktivierungen, welche zur *Cluster*-Überlappung herangezogen werden, könnten ebenso durch eine zufällige Kombination der individuellen Aktivierungen entstanden sein (Yoo et al. 2005 und Rombouts et al. 1997).

Als häufigste und äußerst stringente Methode umgeht dieses Problem die Berechnung der Intraklassenkorrelation (ICC), welche auch in dieser Arbeit angewandt wird. Hier werden erst die signifikanten Regionen der Gruppenaktivierung der beiden Messungen separat festgestellt. Im Anschluss können nicht nur die aktivierten Voxel verglichen werden, sondern es kann auch berücksichtigt werden, ob ihre Aktivierungsstärke in der Test- und Retest-Messung stabil bleibt (Caceres, 2009).

Grundlage der ICC Berechnungsweise ist eine Varianzanalyse, die solche Messfehler quantifiziert, welche die Ergebnisse maßgeblich beeinflussen können. Hierbei wird zwischen einer systematischen Varianz *zwischen* den Individuen (interindividuelle Varianz) und einer Fehlervarianz, basierend auf Aktivierungsunterschieden *innerhalb* eines Individuums

(intraindividuelle Varianz) differenziert. Die ICC stellt sich als Quotient aus der Differenz der interindividuellen Varianz und der Summe der Varianz *zwischen* und *innerhalb* der Individuen dar (Zandbelt et al., 2008).

$$ICC = \frac{\sigma_{\text{zwischen}}^2}{(\sigma_{\text{zwischen}}^2 + \sigma_{\text{innerhalb}}^2)} = \frac{\text{interindividuelle Varianz}}{\text{Gesamtvarianz}} \quad (3.4)$$

Gut reproduzierbare Aktivierungen sind solche, bei denen die interindividuelle Varianz kleiner als ein festgelegter Grenzwert ist. Da es in der fMRT Analyse bisher noch keinen definierten Standardwert dafür gibt, wird üblicherweise die Reliabilität und nicht die Reproduzierbarkeit bestimmt (Bland und Altman, 1986).

Die ICC ist ausschließlich von der Varianz abhängig und kann daher für jede Aktivierungsstärke kalkuliert werden. Somit sind auch Informationen über diejenigen Voxel möglich, welche einen festgelegten Schwellenwert zwar unterschreiten, trotzdem aber konsistent über die beiden Messungen sind (Caceres, Hall, Zelaya, Williams und Mehta, 2009).

Da es verschiedene Berechnungsweisen des ICC gibt, die jeweils unterschiedliche Maßgaben des experimentellen Designs auf Reliabilität untersuchen, haben Shrout und Fleiss 1979 eine Übersichtsarbeit zur Differenzierung der jeweiligen Bedingungen erstellt. Ein Schema dazu zeigt Abbildung 3.21.

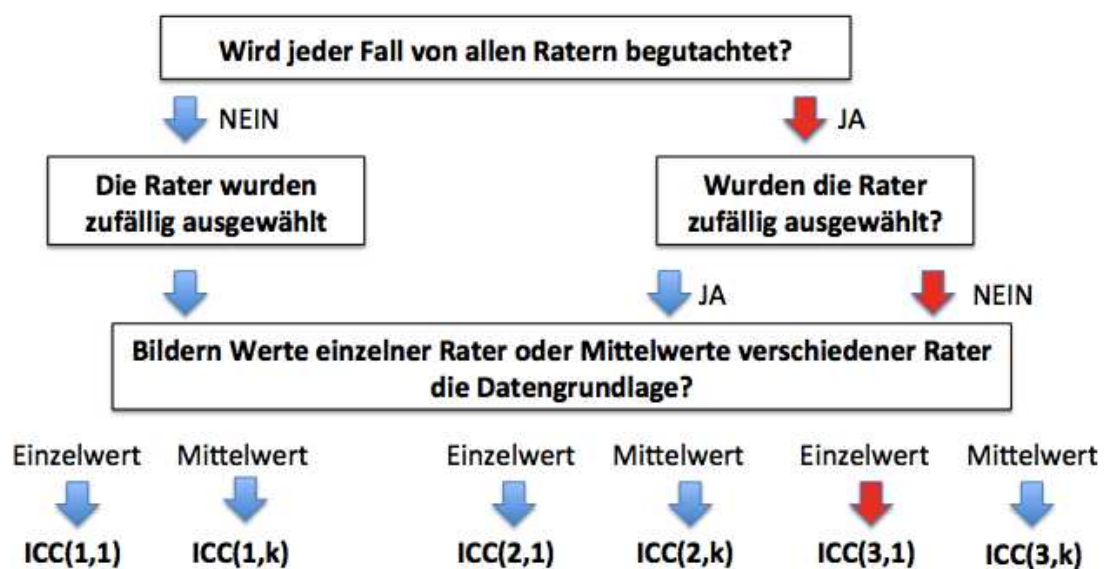


Abb. 3.21 Schema zur Differenzierung der ICC Berechnungsversion. *Beispielpfad der Bedingungen dieser Arbeit ist rot markiert (modifiziert nach Shrout und Fleiss, 1979).*

Der erste Schritt beachtet die Bedingung, ob jedes Messobjekt (Proband_in) von demselben ‚Rater‘ (Beurteiler, in diesem Fall Test und Retest Scanner) begutachtet wird. Am Beispiel des experimentellen Designs dieser Studie kann hier dem Pfad der Erfüllung dieser Bedingung gefolgt werden, da jede_r Proband_in in beiden Messungen von demselben Scanner gemessen wurde. Der nächste Schritt überprüft, ob die Rater zufällig ausgewählt wurden. Diese Bedingung wird in dieser Studie verneint, da der Rater nicht zufällig, sondern ein bestimmter, der 3-Tesla Siemens Magnetom Tim Trio Scanner der Philipps-Universität Marburg, ist. Die letzte Bedingung des Schemas differenziert zwischen den gemessenen Rohwerten, ob sie einzelne Werte oder Mittelwerte repräsentieren. Es wurden in den jeweiligen Messungen Einzelwerte aller Proband_innen erhoben. Den Bedingungen des experimentellen Designs dieser Arbeit folgend, ist der ICC (3,1) die Berechnungsversion, welche eine verlässliche Aussage zur Reliabilität treffen kann (Fließbach et al., 2010).

$$ICC(3,1) = \frac{\sigma_{zwischen}^2 - \sigma_{innerhalb}^2}{\sigma_{zwischen}^2 + \sigma_{innerhalb}^2} \quad (3.5)$$

Ein ICC- Wert von 1.0 entspricht wie schon bei der Cluster-Überlappungsmethode einer nahezu perfekten Übereinstimmung der Aktivierungswerte aus der ersten und der zweiten Messung. Hohe ICC- Werte resultieren nach der Gleichung 3.5 aus einer hohen interindividuellen Varianz ($\sigma_{zwischen}^2$), während niedrige ICC-Werte aus einer hohen intraindividuellen Varianz ($\sigma_{innerhalb}^2$) entstehen. Die intraindividuelle Varianz bezeichnet daher die Kenngröße für den relevanten Messfehler in der Reliabilitäsanalyse (Caceres et al., 2009).

Bei der Interpretation der ICC-Werte einigt sich ein großer Anteil der Autoren auf eine von Shrout und Fleiss beschriebene Skala von 0 = keine Reliabilität, <0,4 = schlechte, <0,6 = moderate, <0,8 = gute und >0,8 = exzellente Reliabilität. Allerdings sind in der fMRT jeweils geringere ICC Werte zur Bewertung heranzuziehen, da es sich um ein generell sehr stör anfälliges Messinstrument handelt.

Zur Berechnung des ICC bei fMRT Paradigmen werden sogenannte ROIs (*Regions of Interest*) festgelegt, die solche Hirnareale definieren, die bei bestimmten kognitiven Prozessen als signifikant aktiv erwartet werden. Hierbei ergeben sich zwei Hauptberechnungswege. Der

erste bestimmt den voxelweisen ICC für jede_n Proband_in, also vergleicht den Aktivierungswert (T-Wert) für jeden einzelnen Voxel mit dem aus dem Retest Scan innerhalb der ROI, wodurch sich ein ICC_{med} ergibt. Der andere Weg ist die ROI-weise Berechnung der ICCs. Durch Mittelung wird für jede ROI der einzelnen Proband_innen nur ein Aktivierungswert bestimmt, welcher der in der Gruppenanalyse stärkste (ICC_{max}), mittlere (ICC_{mean}) oder mediane (ICC_{med}) Aktivierungswert sein kann (Caceres et al., 2009).

3.6.2 Vorgehensweise der Reliabilitätsmessung in dieser Arbeit

Zu Beginn der Reliabilitätsanalyse der Messergebnisse dieser Arbeit wurden mit Hilfe der SPM-Toolbox ‚Anatomy‘ alle aktivierten Hirnareale identifiziert. Parallel dazu erfolgte die Erstellung von 3-D-Bildern mittels Cluster-Overlap Methode. Sie gab eine erste orientierende Visualisierung der Reliabilität der Paradigmen. Schließlich begann die eigentliche Reliabilitätsanalyse anhand der ICC-Berechnung. Nachfolgend wird die Anwendung der ICC-Berechnungsweise innerhalb dieser Studie ausführlich vorgestellt.

Zunächst wurden von jeder_m einzelnen Proband_in pro Paradigma *con*-Bilder (siehe Kapitel 3.5.2) erstellt. Hierzu wird die ICC-Toolbox „gui_icc“ in dem Programm MATLAB (MathWorks, Natick, MA; <http://www.mathworkds.com/>) benutzt. Es ergaben sich bei 20 Proband_innen und zwei Messungen insgesamt 40 *con*-Bilder pro Paradigma. Für jedes entstandene *con*-Bild überprüfte das Programm im Anschluss, ob jedem Voxel ein Aktivierungswert zugeordnet werden konnte. Durch aufnahmebedingte Artefakte könnten in einigen Hirnregionen keinerlei Aktivierungswerte zustande kommen, da beispielsweise Luftüberlagerungen deren Erhebung nicht zulassen. Am Ende dieser Überprüfung entstanden für die Test- und Retest-Messung entsprechende *m_con*-Bilder, die nun für jedes Paradigma in die richtige Reihenfolge gebracht wurden.

Aus den *m_con*-Bildern folgte die voxelweise Berechnung der ICCs, welche in der allgemeinen Einführung zur Reliabilitätsanalyse (Kapitel 3.6.1) erläutert wird. Es entstand ein Streudiagramm einer ICC-vs.-T-Karte, welches die Assoziation zwischen dem T-Wert (Aktivierungsstärke als unabhängige Variable) und der Reliabilitätsstärke (abhängige Variable) eines jeden Voxels darstellte. Höhere Schwellenwerte für den T-Wert führten hierbei zu anderen Aktivierungsmustern und höheren ICC-Werten (Caceres et al., 2009).

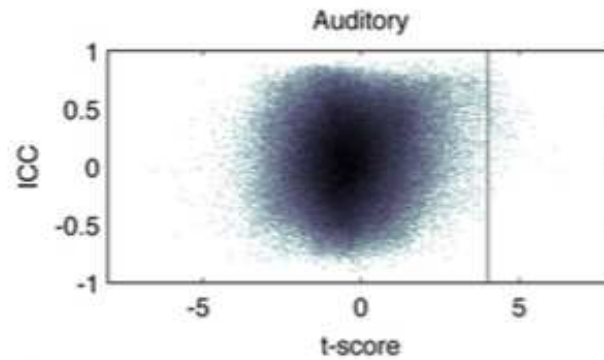


Abb. 3.22 Beispiel eines Streudiagramms der ICC und T-Werten Assoziationen (*modifiziert nach Caceres et al., 2009*)

Der nächste Verarbeitungsschritt mit der Toolbox erstellte eine ICC-Häufigkeitskarte, welche graphisch in drei Kurven zeigte, wie viele Voxel welchen ICC-Wert über das gesamte Gehirn, über das aktivierte Netzwerk und über die deaktivierten Areale relativ erreicht hatten.

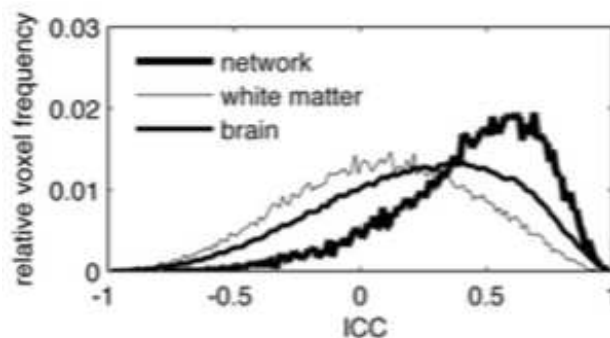


Abb. 3.23 Beispiel einer relativen Häufigkeitsverteilung des gesamten Gehirns (brain), der deaktivierten Areale (white matter) und der aktivierten Netzwerke (network) (*modifiziert nach Caceres et al., 2009*)

Zuletzt erfolgte die Bestimmung der ICC-Werte der für das jeweilige Paradigma definierten ROI. Hierfür mussten ROI-Masken verwendet werden, um die nicht interessierenden Regionen aus der Berechnung auszuschließen. Diese ROI-Maske bestand aus der Wahl eines p-Wertes und einer Mindestclustergröße, bei welchen möglichst überwiegend die gewünschte ROI Aktivierungen aufwies (Jansen et al., 2006). Aus den erhobenen ICC dieser Region wurde nun ein medianer ICC-Wert berechnet. Gegenüber einer vorgefertigten anatomisch definierten Maske mit dem WFU-Pickatlas (Maldjian et al., 2003) hat die funktionelle Referenzregion den Vorteil, dass sie analysiert, welche Voxel in der Region des Gehirns aktiv sind und zeigt nur diese an.

Zu den Möglichkeiten, die in die Reliabilitätsanalyse einfließenden Daten qualitativ in der Auswahl ihrer Aktivierungsstärke zu verbessern, gehörte die Wahl eines Schwellenwertes (p-Wert, T-Wert oder Clustergröße) und die Festlegung eines Dilatations-Faktors. Zuletzt Genannter ermöglichte die gleichmäßige Vergrößerung einer definierten Referenzregion. In der vorliegenden Studie wurde hierfür ein Dilatations Faktor von 2 gewählt. In einem letzten Schritt wurde noch für die gewählten Referenzregionen ein medianer (ICC_{med}) bestimmt.

4 ERGEBNISSE

Dieses Kapitel stellt zunächst die Gruppenanalysen der ersten Messungen der jeweiligen Paradigmen vor und zeigt im Anschluss daran die Ergebnisse der Reliabilitätsberechnung.

Die Darstellung der Ergebnisse der Gruppenanalyse erfolgt in drei Schritten: der erste Teil zeigt die Aktivierungsmuster der Gruppenaktivierung der ersten Messung. Der zweite Teil präsentiert eine tabellarische Übersicht der auf das MNI-Standardgehirn projizierten Koordinaten, den für diese Ergebnisse zugehörigen Parametern und den anatomischen Zuordnungen der aktivierten Hirnregionen. Im dritten Teil erfolgt die Darstellung der Aktivierungsüberlappungen. Die Gruppeneaktivierung wird aus dem Vergleich der aktivierten Areale in der Kontrollphase zur Aktivierungsphase in der ersten Messung berechnet.

Die Ergebnisse der Reliabilitätsmessung werden anhand der mittels Cluster Überlappungsmethode erstellten 3D- und Schnittbilder dargestellt und schließlich die berechneten ICCs im Detail erläutert.

4.1 Ergebnisse der WG-Aufgabe

Der Gruppenanalyse dieser Aufgabe konnten die Daten von nur 19 der 20 gemessenen Proband_innen zugrunde gelegt werden, da bei der sechsten Probandin während der ersten Messung aufgrund von technischen Schwierigkeiten die Daten nicht gespeichert werden konnten.

Bei der Gruppenanalyse wurde stets der Wert $k > 10$ Voxel als Mindestclustergröße für ein aktiv bewertetes Areal eingesetzt. In der Auswertung ergab der Schwellenwert $p(\text{fwe-unkorrigiert}) = 0,001$ lediglich ein aktiviertes Cluster im Bereich des linken Nucleus Caudatus. Die Ergebnisse, die im folgenden Abschnitt dargestellt werden, sind mit dem Schwellenwert $p(\text{fwe korrigiert}) = 0,05$ berechnet worden und zeigen daher weitaus mehr Aktivierungsareale.

4.1.1 Ergebnisse der Gruppenanalyse

Eine Übersicht der aktivierten Hirnregionen der Wortgenerierungsaufgabe mit anatomischen Zuordnungen zeigt das *Glassbrain*- Schnittbild in Abbildung 4.1.

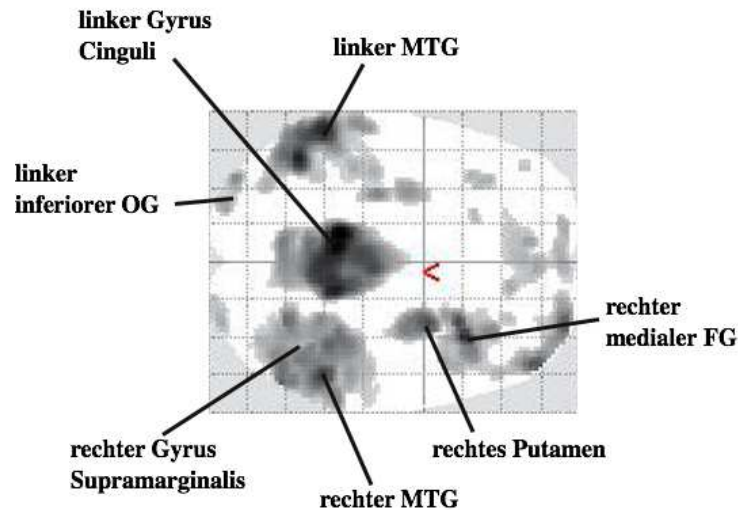


Abb. 4.1 Übersicht der Aktivierungen der WG-Aufgabe mit anatomischen Zuordnungen im Transversalschnitt. Aktivierungen der ersten Messung bei $p(\text{fwe-korrigiert})=0,05$. MTG= mediale Temporalgyrus, FG= Frontal Gyris, OG= Occipitalgyrus.

Das aktivierte Netzwerk der ersten Messung ist in Abbildung 4.2 weiterhin als 3D-Bild dargestellt. Die Hauptaktivierungen der WG-Aufgabe dieser Studie liegen in den folgenden Hirnregionen: Links dominierend der mediale Gyrus Cinguli, der mediale und inferiore Gyrus Temporalis und der Gyrus Angularis. Rechts dominierend waren der Gyrus Temporalis mediale, der superiore und mediale Gyrus Frontalis und Gyrus Supramarginalis sowie der visuelle Cortex mit dem Gyrus Occipitalis inferior hauptaktiviert. Das Putamen, der Gyrus Fusiformis und der Gyrus Temporalis inferior zeigten beidseits Aktivierungen.

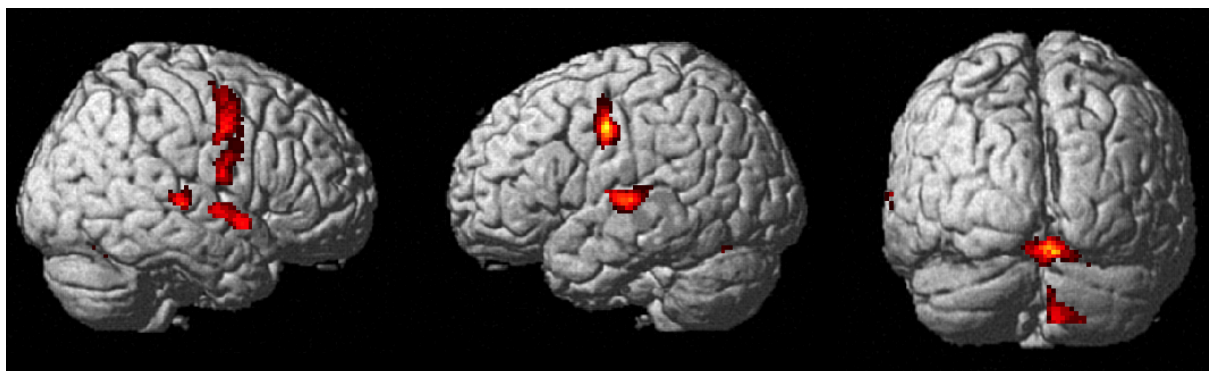


Abb. 4.2 Darstellung des aktivierten Netzwerks in der Gruppenanalyse der ersten Messung der WG-Aufgabe. $p(\text{fwe-korrigiert}) = 0,05$, $k = 10$. Die aktivierten Areale sind rot-gelb markiert.

MNI-Koordinaten			Clustergröße kE	T-Wert	Z	P	Anatomisches Korrelat
-10	-42	38	4307	13,89	6,58	0	Linker medialer Gyrus Cinguli
10	-46	30		11,53	6,12	0	Rechter posteriorer Gyrus Cinguli
-44	-60	28	2155	12,34	6,29	0	Linker medialer Gyrus Temporalis
-56	-48	40		12,08	6,24	0	Linker Gyrus Supramarginalis
-46	-66	42		9,83	5,7	0,001	Linker Gyrus Angularis
36	16	56	1500	11,62	6,14	0	Rechter medialer Gyrus Frontalis
30	64	12		9,99	5,75	0	Rechter superiorer Gyrus Frontalis
54	-48	32	3981	11	6	0	Rechter Gyrus Supramarginalis
40	-46	52		9,83	5,7	0,001	Rechter inferiorer Gyrus Parietalis
56	-68	30		9,81	5,7	0,001	Rechter medialer Gyrus Occipitalis
30	14	-32	193	10,75	5,94	0	Rechter medialer Polus Temporalis
28	6	-38		7,49	4,98	0,014	Rechter Gyrus Fusiformis
30	0	0	358	10,65	5,91	0	Rechtes Putamen
56	-52	0	315	8,99	5,47	0,002	Rechter medialer Gyrus Temporalis
52	-60	-4		8,51	5,32	0,003	Rechter inferiorer Gyrus Temporalis
-36	-90	-2	112	8,58	5,35	0,003	Linker inferiorer Gyrus Occipitalis
-58	-58	-2	169	8,53	5,33	0,003	Linker inferiorer Gyrus Temporalis
-60	-46	-4		7,52	4,99	0,013	Linker medialer Gyrus Temporalis
-30	-10	2	155	8,45	5,31	0,007	Linkes Putamen
6	50	14	497	7,97	5,15	0,007	Rechter superiorer Gyrus Medialis
-12	64	14		7,65	5,04	0,011	Linker superiorer Gyrus Frontalis

Tab. 4.1 Darstellung der anatomischen Korrelate der ersten Messung der WG-Aufgabe. $p(\text{fwe korrigiert}) = 0,05$, $k > 10$.

Tabelle 4.1 zeigt die anatomischen Zuordnungen der Koordinaten zu den Aktivierungen der 3D-Bilder bei $p = 0,05$ und einer Mindestclustergröße von $k > 10$ der ersten Messung.

Die zweite Messung zeigte bei $p(\text{fwe korrigiert}) = 0,05$ qualitativ dieselben aktivierten Hirnregionen mit etwas mehr Aktivierungsstärke, daher wird sie hier nicht separat dargestellt.

MNI-Koordinaten	Clustergröße kE	T-Wert	Z	P	Anatomisches Korrelat
-22 -64 -20	2291	4,96	3,89	0	Linkes Zerebellum
42 -70 -26		3,67	3,13	0,001	Rechtes Zerebellum
34 -48 8	365	4,22	3,47	0	Rechter lateraler Ventrikel
22 -28 20		3,82	3,23	0,001	Rechter lateraler Ventrikel
28 -36 12		2,95	2,63	0,004	Rechter lateraler Ventrikel
-2 2 20	78	3,91	3,28	0,001	Linker lateraler Ventrikel
64 -8 -2		2,71	2,45	0,007	Rechter STG
-56 -8 48	150	3,48	3	0,001	Linker präzentraler Gyrus
-46 -14 40		3,23	2,83	0,002	Linker präzentraler Gyrus
-8 -22 -16	43	3,45	2,98	0,001	Linker ventraler DC
2 50 -2	148	3,35	2,91	0,002	Rechter Lingualer Gyrus
20 -70 -50	146	3,3	2,88	0,002	Rechtes Zerebellum
14 -70 -44		2,96	2,63	0,004	Rechtes Zerebellum
-30 -52 8	367	3,26	2,85	0,002	Linker lateraler Ventrikel
-24 -38 12		3,16	2,78	0,003	Linker lateraler Ventrikel
60 -4 44	135	3,05	2,7	0,003	Rechter präzentraler Gyrus
-60 -14 2	48	2,89	2,58	0,005	Linker STG
58 4 -8	40	2,82	2,54	0,006	Rechter Temporalpol
-58 2 -4	15	2,77	2,49	0,006	Linker Temporalpol

Tab. 4.2 Darstellung der anatomischen Korrelate der zweiten Messung der WG-Aufgabe. $p(\text{unkorrigiert})=0,01$, $k > 10$.

In der Evaluation der zweiten Messung fanden sich darüber hinaus bei einem unkorrigierten p-Wert von 0,01 im Gegensatz zu der Analyse der Daten der ersten Messung mit $p(\text{fwe korrigiert})=0,05$ zusätzlich die für die motorischen Funktionen bekannten Aktivierungen. Hierzu zählen beispielsweise der postzentrale und präzentrale Gyrus und auch die Koordinaten für den Bereich des Larynx im Motorcortex (-46, -14, 40). Auch das Zerebellum war beidseits aktiviert (siehe Tabelle 4.2).

Die Messung der Kontrollphase zeigte unerwarteter Weise deutlich mehr aktivierte Hirnregionen an. Hier fand sich im Gegensatz zu der Aktivierungsphase ein großes frontoparietales Aktivierungsbild (siehe Abb. 4.3).

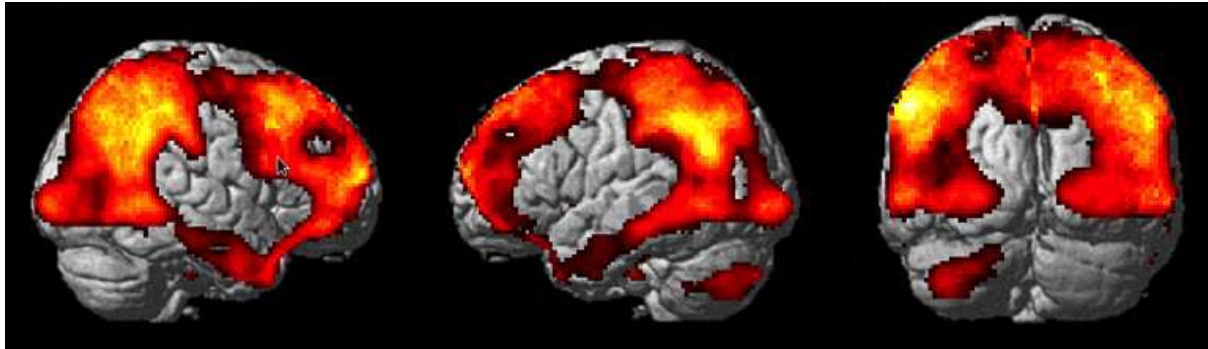


Abb. 4.3 Darstellung der Deaktivierungen der ersten Messung der WG-Aufgabe (Kontrolle > Aktivierungen). p (unkorrigiert) = 0,001 und $k > 10$ Voxel. Die deaktivierten Areale sind rot markiert.

4.1.2 Ergebnisse der Reliabilitätsmessung

Zur Präsentation der Ergebnisse der Reliabilität der WG-Aufgabe werden im Folgenden anhand von 3D-Bildern mittels Cluster-Überlappungsmethode jeweils die Aktivierungen und Deaktivierungen verglichen. Danach folgt eine detaillierte Vorstellung der Analyse der ICC-Werte.

Wie sehr sich die aktivierten Areale der ersten und zweiten Messung überlappen, wird in Abbildung 4.4 dargestellt. Die gleichen Netzwerke der beiden Messungen sind hier gelb markiert. Klar erkennbar ist hier, wie bereits beschrieben, dass die zweite Messung (grün) deutlich mehr Aktivierung zeigt.

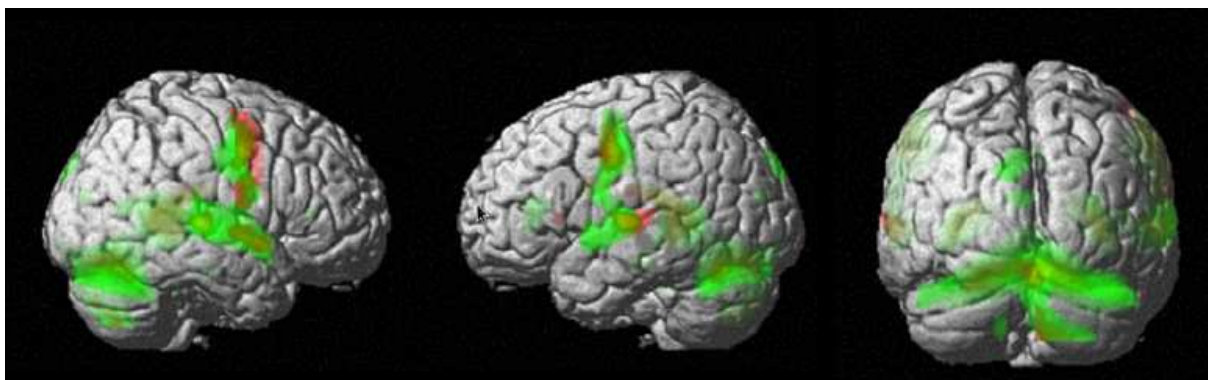


Abb. 4.4 Darstellung der Aktivierungen der beiden Messungen der WG-Aufgabe als Cluster-Überlappung. p (fwe korrigiert) = 0,05, $k > 10$, Messung 1 = rot, Messung 2 = grün und Überlappung = gelb.

Bei den Deaktivierungsergebnissen zeigen die 3D-Darstellungen deutliche Überlappungen der ersten und zweiten Messung. Dies ist in Abbildung 4.5. an den großflächigen gelben Markierungen gut erkennbar.

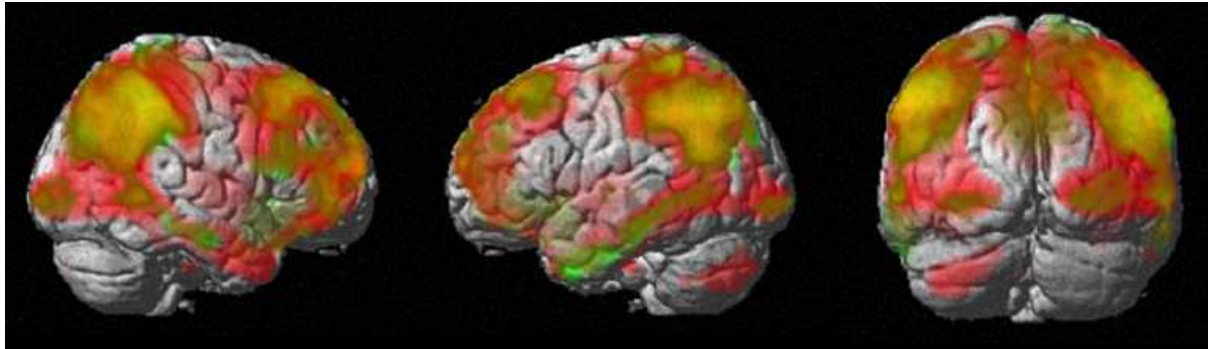


Abb. 4.5 Darstellung der Deaktivierungen (Kontrolle>Aktivierung) der beiden Messungen der WG-Aufgabe als Cluster-Überlappung. $p(\text{fwe unkorrigiert}) = 0,001$, $k > 10$, Messung 1 = rot, Messung 2 = grün und Überlappung = gelb.

Nach der Darstellung der Reliabilität der WG-Aufgabe mittels Cluster-Überlappungsmethode werden nun die Ergebnisse mittels ICC-Analyse vorgestellt. Hierzu wurde zunächst eine ICC-vs.-T-Karte erstellt (siehe Kapitel 3.6.2). Sie zeigt graphisch, wie die Aktivierungsstärken der einzelnen Voxel mit ihrer Reliabilitätsstärke assoziiert sind (Abb. 4.6). Der ICC-Wert wird in dieser und den folgenden ICC-vs.-T-Karten als ein Prozentwert zwischen -100 bis 100 angegeben. Ein Wert von 50% entspräche dementsprechend einem ICC von 0,5.

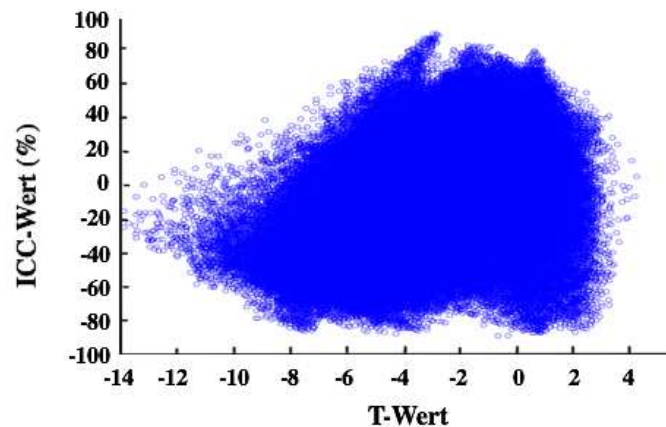


Abb. 4.6 ICC-vs.-T-Karte der WG-Aufgabe. Darstellung der T-Werte der ersten Messung und ihren assoziierten ICCs. Jeder blaue Kreis repräsentiert einen Voxel.

In der ICC-vs.-T-Karte lässt sich beobachten, dass der Großteil der Voxel im Bereich der negativen T-Werte zwischen $T \approx -8$ und $T \approx 2$ liegt. Es zeigen sich vor allem hohe ICC-Werte um den T-Wert-Bereich zwischen $T \approx -4$ und $T \approx 1$. Weiterhin liegen auch sehr viele Voxel im Bereich von negativen ICC-Werten zwischen $ICC \approx -0,8$ und $ICC \approx 0,3$. Insgesamt weisen sowohl Voxel mit sehr hohem T-Wert als auch Voxel mit stark negativen T-Werten keine

hohen ICC-Wert Korrelationen an, was die Vermutung nahe legt, dass stark (de)aktivierte Areale nicht gut reproduzierbar sind.

Abbildung 4.7 zeigt die relative Häufigkeitsverteilung der voxelweisen ICC-Werte. Hierbei wird insbesondere der mediane ICC für das aktivierte, das deaktivierte Netzwerk und für das Netzwerk des gesamten Gehirns der ersten Messung ersichtlich. Auch in dieser und den folgenden ICC-Häufigkeitskarten wird der ICC in Prozent angegeben.

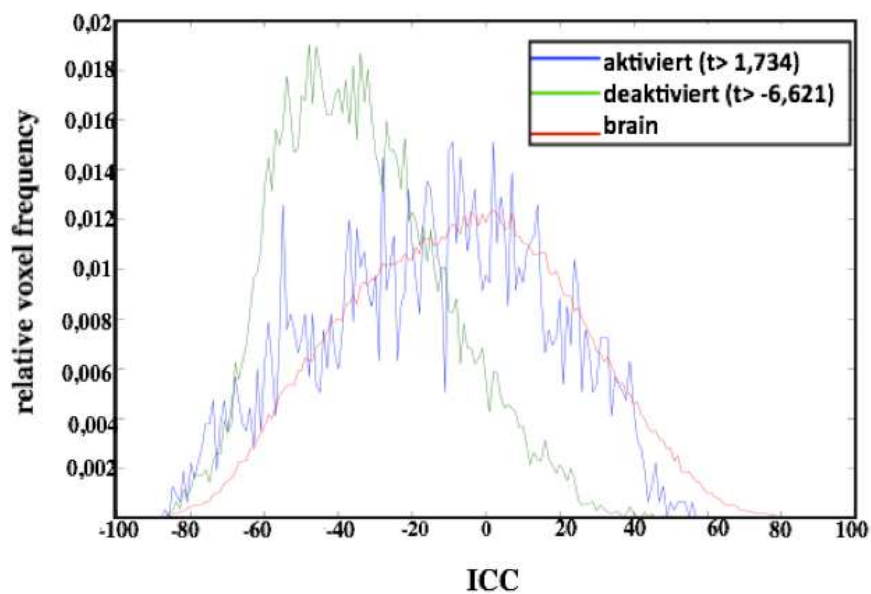


Abb. 4.7 ICC-Häufigkeitskarte der WG-Aufgabe. $p(\text{fwe korrigiert}) < 0,05$. Relative Häufigkeitsverteilung der ICCs der Voxel für das aktivierte Netzwerk (blau), das deaktivierte Netzwerk (grün) und das gesamte Gehirn (rot).

Für das gesamte Gehirn zeigt die Graphik mit voreingestelltem p-Wert von $p(\text{fwe korrigiert})=0,05$ die erwartete Gauß-förmige Kurve. Der mediane ICC des gesamten Gehirns liegt bei $\text{ICC}_{\text{med}} = -0,063$. Ein derart niedriger ICC_{med} für die Aktivierungen des gesamten Gehirns deutet auf eine generell geringe Reliabilität hin.

Das aktivierte Netzwerk zeigt bei einem mittleren Aktivierungswert von $t = 1,734$ ebenfalls einen Gauß-förmigen Verlauf. Der mediane ICC von $\text{ICC}_{\text{med}} = -0,128$ liegt hier sogar noch niedriger als der ICC_{med} des Gesamthirns. Stark aktivierte Hirnregionen sind demnach nicht reliabel aktiv. Auch das deaktivierte Netzwerk zeigt bei $t = -6,621$ niedrige Reliabilitätswerte mit dem $\text{ICC}_{\text{med}} = 0,358$.

Zusammenfassend lässt sich aus den vorangegangenen Analysen sagen, dass deutlich mehr Areale während der Kontrollphasen der einzelnen Messungen Aktivierungen zeigen. Sowohl die Aktivierungen der Aktivierungsphase als auch die Aktivierungen der Deaktivierungsphase scheinen in der Cluster-Überlappungsmethode vergleichbar aktiviert zu sein. Bei der Rechenanalyse der Reliabilitäten stellt sich jedoch heraus, dass die berechneten ICC-Werte insgesamt sehr niedrig sind und die aktivierten und deaktivierten Hirnregionen keine höhere Reliabilität zeigen als Bereiche, die nicht aktiv sind.

An dieser Stelle könnte eine ICC Berechnung für einzelne interessierende Regionen wie bspw. dem linken IFG als erwartetes Hauptaktivierungszentrum oder dem visuellen Cortex als typischerweise stark reliable Region durchgeführt werden. Erwartungsgemäß wären die berechneten ICC-Werte dieser ROI stärker als im gesamten aktivierten Netzwerk. In Abbildung 4.7 ist allerdings bereits deutlich erkennbar, dass es keine aktivierten Voxel gibt, die überhaupt in dem Bereich von $ICC > 0,6$ liegen, so dass diese Analyse nicht durchgeführt wurde.

4.2 Ergebnisse der SPES-Aufgabe

In der SPES-Aufgabe wurden insgesamt drei verschiedene Emotionen, ein neutrales Gesicht und ein Kontrolldia gezeigt, so dass sich viele verschiedene Möglichkeiten der Auswertung ergaben. Die Analyse der Daten auf Gruppenebene zeigte sehr unterschiedliche Aktivierungsmuster dieser möglichen Kontraste, welche in einem ersten Schritt untersucht wurden. Beispielsweise ergab sich in dem Kontrast der Aktivierungen in der Phase der subliminalen Präsentation trauriger Gesicht zu der Phase der neutralen Gesichter ein sehr geringes Aktivierungsmuster. Im Gegensatz dazu boten die Kontraste der freudigen und ängstlichen Gesichter zu neutralen Gesichtern weitaus mehr Aktivierungen zur Analyse an. Dieser Unterschied in dem Ausmaß der Aktivierungen von traurigen zu freudigen bzw. ängstlichen Stimuli wird im Kapitel 5.3.2 der Diskussion näher beleuchtet, da er vergleichbar zu anderen Studienergebnissen ist. Allerdings steht in dieser Arbeit die Frage der Reliabilität der Paradigmen im Vordergrund, so dass beispielhaft der Kontrast der subliminalen Präsentation ängstlicher Gesichter zu neutralen Gesichtern analysiert wird (ANGST>NEUTRAL). Zusätzlich dazu wird der Kontrast der subliminalen Präsentation aller

Gesichter zu dem grauen Kontrolldia analysiert (GESICHTER>KONTROLLE), da hiermit überprüft werden kann, ob das Paradigma abgesehen von der Verarbeitung von Emotionen auch generell den Prozess der Gesichterverarbeitung reliabel messen kann.

4.2.1 Ergebnisse der Gruppenanalyse

GESICHTER>KONTROLLE KONTRAST

Bei dieser Gruppenanalyse wurde wie bereits bei der WG-Aufgabe stets die Mindestclustergröße von $k > 10$ Voxel angegeben. In beiden Messungen wurde der unkorrigierte Schwellenwert von $p = 0,001$ für die Berechnung der Aktivierungsareale im *Glassbrain*-Schnittbild dargestellt. In dieser Ergebnisdarstellung werden im Gegensatz zu der WG-Aufgabe keine Deaktivierungsbilder dargestellt. Zwar wurden einige Deaktivierungswerte gemessen, jedoch nicht in einem so außergewöhnlich unverhältnismäßigen Anteil wie bei der WG-Aufgabe. Innerhalb der Aktivierungswerte kann daher genau das Netzwerk gezeigt werden, was interessiert. Die Übersicht der aktivierten Hirnregionen des Kontrastes aller Gesichter zu dem Kontrolldia wird mit den anatomischen Zuordnungen der Hauptaktivierungen in Abbildung 4.8 gezeigt.

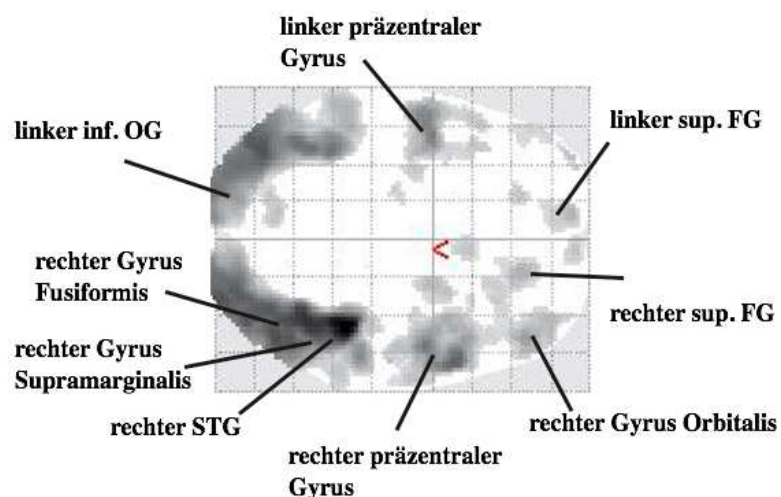


Abb. 4.8 Übersicht der Aktivierungen des Kontrastes Gesichter>Kontrolle der SPES-Aufgabe mit anatomischen Zuordnungen im Transversalschnitt. Aktivierungen der ersten Messung bei p (fwe unkorrigiert) = $0,001$. STG= superiorer Temporalgyrus, FG= Frontal Gyrus, OG= Occipitalgyrus.

Die Hauptaktivierungen der SPES-Aufgabe dieser Arbeit liegen mit bilateralen aber rechts dominanten Aktivierungen des Gyrus Fusiformis und des Gyrus Temporalis inferior in der sogenannten *Fusiform Face Area* (FFA) und mit Aktivierungen des bilateralen inferioren Gyrus Occipitalis im visuellen Cortex. Die bilateralen Aktivierungen des STG entsprechen dem superioren Sulcus Temporalis (STS), welcher ebenfalls für die Verarbeitung veränderbarer Aspekte von Gesichtern bekannt ist (Haxby, Hoffmann und Gobbini, 2000). In kleineren Clustern konnten zudem in einem Wahrscheinlichkeitsanteil von 16,2% die Amygdala und von 24% der Hippocampus sowie der IFG als aktiviert gemessen werden.

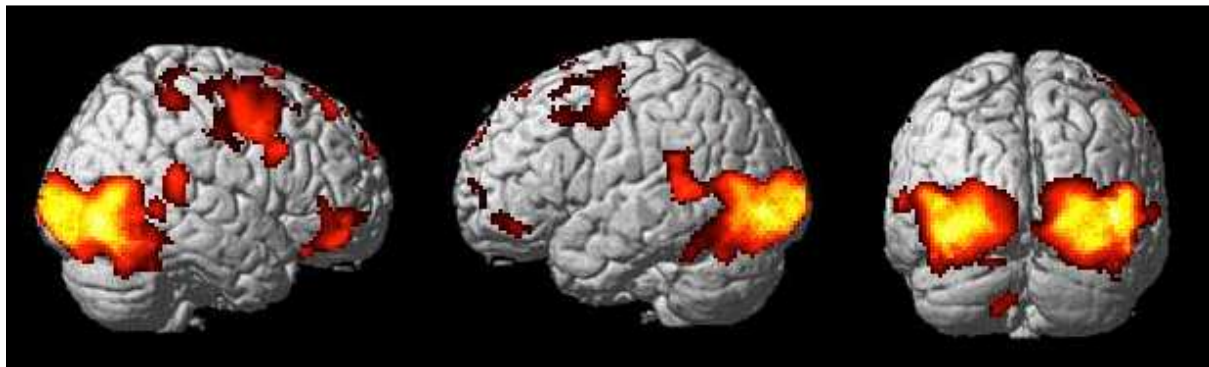


Abb. 4.9 Darstellung des aktivierten Netzwerks in der Gruppenanalyse der ersten Messung der SPES-Aufgabe. Gesichter > Kontrolle Kontrast. p (fwe unkorrigiert) = 0,001, $k = 10$. Die aktivierten Areale sind rot-gelb markiert.

In Abbildung 4.9 ist das aktivierte Netzwerk der ersten Messung als 3D-Bild dargestellt. Die Tabellen 4.3 und 4.4. zeigen die anatomischen Zuordnungen der Koordinaten zu den Aktivierungen der 3D-Bilder der beiden Messungen bei p (fwe unkorrigiert)= 0,001 und einer Mindestclustergröße von $k > 10$.

MNI-Koordinaten	Clustergröße kE	T-Wert	Z	P	Anatomisches Korrelat
42 -42 -22	5850	10,41	5,95		rechter Gyrus Fusiformis
46 -74 -2		7,83	5,17		rechter inferiorer Gyrus Occipitalis
-38 -82 -6	5059	7,69	5,12		linker inferiorer Gyrus Occipitalis
-32 -90 -4		6,93	4,84		linker inferiorer Gyrus Occipitalis
56 8 46	1392	7,25	4,96		rechter präzentraler Gyrus
50 -6 60		6,69	4,74		rechter präzentraler Gyrus
48 12 54		5,53	4,22		rechter mittlerer Gyrus Frontalis
-44 -2 60	445	6,52	4,67		linker präzentraler Gyrus
-50 -6 56		6,44	4,63		linker präzentraler Gyrus
-40 18 56		4,34	3,57		linker mittlerer Gyrus Frontalis
54 -36 16	358	6,03	4,46		rechter STG
46 42 -18	506	5,57	4,24		rechter lateraler Gyrus Orbitalis
50 42 -8		4,97	3,93		rechter Pars Triangularis des IFG
42 52 -14		4,74	3,8		rechter lateraler Gyrus Orbitalis
56 -38 58	213	5,07	3,85		rechter Gyrus Supramarginalis
38 -32 72		4,51	3,67		rechter postzentraler Gyrus
42 -40 68		3,98	3,35		rechter superiorer Parietallobe
16 38 56	160	4,96	3,92		rechter superiorer Gyrus Frontalis
16 42 48		4,83	3,85		rechter superiorer Gyrus Frontalis
-8 62 34	123	4,86	3,87		linker superiorer Gyrus Frontalis

Tab. 4.3 Darstellung der anatomischen Korrelate der ersten Messung der SPES-Aufgabe im Kontrast Gesichter>Kontrolle. $p(\text{unkorrigiert})=0,001$, $k>10$.

MNI-Koordinaten	Clustergröße kE	T-Wert	Z	P	Anatomisches Korrelat
50 -54 -18	14289	17,5	7,27	0	rechter inferiorer Gyrus Temporalis
44 -48 -20		13,32	6,59	0	rechter Gyrus Fusiformis
38 -62 -16		11,54	6,22	0	rechter occipitaler Gyrus Fusiformis
-34 -84 -18	5823	13,14	6,56	0	linker inferiorer Gyrus Occipitalis
-40 -48 -18		10,31	5,92	0	linker Gyrus Fusiformis
-42 -70 -10		9,87	5,78	0	linker inferiorer Gyrus Occipitalis
-46 -6 60	1871	7,37	5,01	0	Linker präzentraler Gyrus
-40 10 60		6,36	4,6	0	linker mittlerer Gyrus Frontalis
-34 -4 68		6,2	4,53	0	Linker präzentraler Gyrus
12 -78 -38	101	6,5	4,66	0	rechtes Zerebellum
36 34 -20	191	5,75	4,32	0	rechter posteriorer Gyrus Orbitalis
28 38 -16		4,89	3,88	0	rechter anteriorer Gyrus Orbitalis
12 20 52	36	4,56	4,7	0	rechte SMA
-38 -60 62	29	4,45	3,64	0	linker superiorer Parietallobe
-50 42 -4	25	4,08	3,41	0	linker Pars triangularis des IFG
-18 30 62	25	3,98	3,35	0	linker superiorer Gyrus Frontalis
-12 40 58		3,7	3,17	0,001	linker superiorer Gyrus Frontalis
30 -6 -22	32	3,9	3,3	0	rechte Amygdala (16,2%) und Hippocampus (24%)

Tab. 4.4 Darstellung der anatomischen Korrelate der zweiten Messung der SPES-Aufgabe im Kontrast Gesichter>Kontrolle. $p(\text{unkorrigiert})=0,001$, $k>10$.

ANGST>NEUTRAL KONTRAST

Repräsentativ für die Verarbeitung subliminaler Emotionen wird der Kontrast von ängstlichen zu neutralen Gesichtern hinsichtlich der Reliabilität ausgewertet. Auch hier wurde für die Gruppenanalyse die Mindestclustergröße von $k>10$ Voxel angegeben. In der ersten Messung wurde der unkorrigierte Schwellenwert von $p=0,01$ und in der zweiten Messung der unkorrigierte Schwellenwert von $p=0,001$ für die Berechnung der Aktivierungsareale eingesetzt. Die unterschiedlichen Schwellenwerte wurden gewählt, weil die zweite Messung stärkere Aktivierungen zeigte und somit ein restriktiverer Schwellenwert die Vergleichbarkeit der aktivierten Hirnregionen verbesserte.

Die Übersicht der aktivierten Hirnregionen des Kontrastes der subliminalen Präsentation von ängstlichen und neutralen Gesichtern wird mit den anatomischen Zuordnungen der Hauptaktivierungen in Abbildung 4.10 gezeigt.

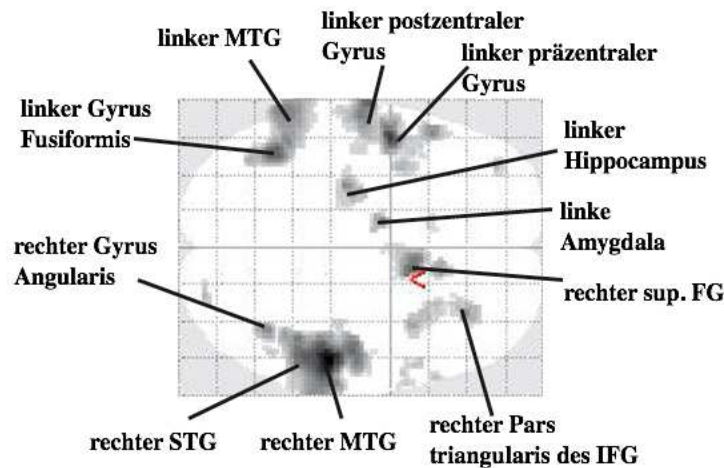


Abb. 4.10 Übersicht der Aktivierungen des Kontrastes Angst>Neutral der SPES-Aufgabe mit anatomischen Zuordnungen im Transversalschnitt. Aktivierungen der ersten Messung bei $p(\text{fwe unkorrigiert})=0,01$. MTG= medialer Temporalgyrus, STG= superiorer Temporalgyrus, FG= Frontal Gyrus, IFG= inferiorer FG

Die aktivierten Hirnregionen dieses Kontrastes liegen wie bereits in Abbildung 4.10 gut erkennbar ist, zu großen Teilen in subkortikalen Strukturen wie der Amygdala und dem Hippocampus. Zur Darstellung der aktivierten Netzwerke werden daher zusätzlich zu den 3D-Bildern Schnittbilder in drei Ebenen gezeigt, in welchen diese Strukturen erkennbar sind (Abb.4.11).

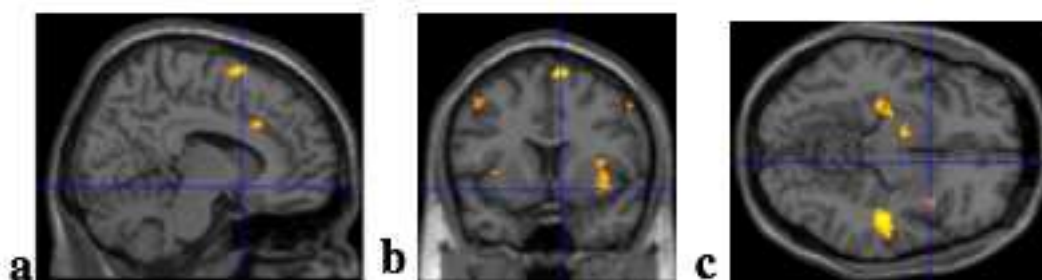


Abb. 4.11 Darstellung der Hauptaktivierungen der SPES-Aufgabe mit Angst>Neutral Kontrast der ersten Messung in Schnittbildern. $p(\text{fwe unkorrigiert})= 0,01$, $k>10$. a) Sagittalschnitt b) Koronarschnitt c) Transversalschnitt.

Die Hauptaktivierungen des ANGST>NEUTRAL Kontrastes der SPES-Aufgabe lagen in den folgenden Hirnregionen: bilateral im MTG und STG sowie dem Gyrus Fusiformis (siehe Abb.4.12). Als bekannte subkortikale Hirnareale waren bilateral die Amygdala und der Hippocampus aktiviert. Des Weiteren fanden sich Aktivierungen im Bereich des prä- und postzentralen Gyrus sowie der SMA.

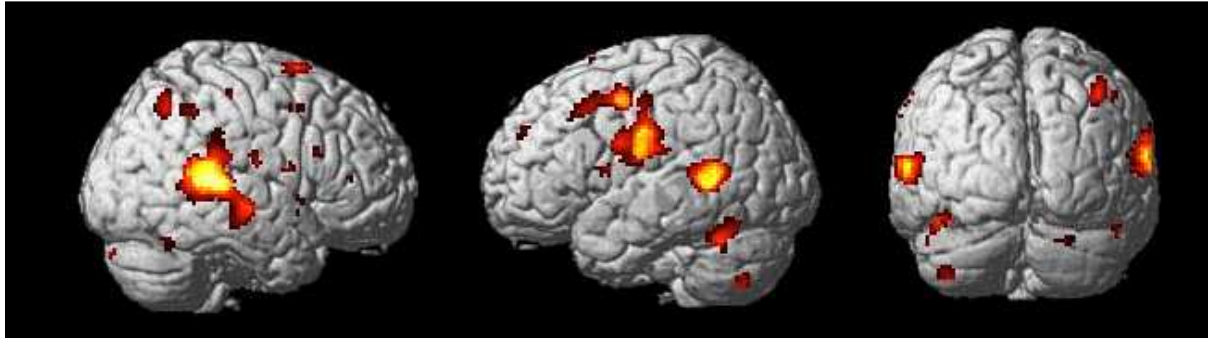


Abb. 4.12 Darstellung der Hauptaktivierungen der SPES-Aufgabe mit Angst>Neutral Kontrast der ersten Messung in 3D-Bildern. $p(\text{fwe unkorrigiert})= 0,01$, $k > 10$. Die Aktivierungen sind rot-gelb markiert.

Die Tabellen 4.5 und 4.6 zeigen die anatomischen Zuordnungen der Koordinaten der beiden Messungen bei $p(\text{fwe unkorrigiert})= 0,01$ der ersten Messung und $p(\text{fwe unkorrigiert})= 0,001$ der zweiten Messung und einer Mindestclustergröße von $k > 10$ Voxel. Da die Aktivierungen im Bereich des Hippocampus und der Amygdala nicht sicher anatomisch zugeordnet werden konnten, ist die Wahrscheinlichkeit für die angegebene Hirnregion in den folgenden Tabellen mit aufgeführt worden.

MNI-Koordinaten	Clustergröße kE	T-Wert	Z	P	Anatomisches Korrelat
54 -32 6	1816	5,67	4,28	0	rechter STG
52 -40 8		4,4	3,61	0	rechter MTG
50 -20 -8		4,14	3,45	0	rechter STG
-48 -2 48	258	4,64	3,75	0	linker präzentraler Gyrus
-50 18 44		3,64	3,13	0,001	linker mittlerer Gyrus Frontalis
-46 8 50		2,76	2,5	0,006	linker mittlerer Gyrus Frontalis
-42 -56 -22	229	4,18	3,48	0	linker Gyrus Fusiformis
-50 -56 -30		2,83	2,56	0,005	linkes Zerebellum
12 8 68	118	4,12	3,44	0	rechter superiorer Gyrus Frontalis
4 2 70		2,9	2,61	0,005	rechte SMA
-64 -12 24	700	4,11	3,43	0	linker postzentraler Gyrus
-58 -16 30		3,8	3,24	0,001	linker postzentraler Gyrus
-54 0 26		3,39	2,96	0,002	linker präzentraler Gyrus
-60 -50 4	596	3,75	3,2	0,001	linker MTG
-46 -50 8		3,69	3,17	0,001	linker MTG
-66 -42 12		3,12	2,77	0,003	linker STG
40 -60 48	80	3,68	3,16	0,001	rechter Gyrus Angularis
-10 -6 -10	27	3,65	3,14	0,001	linke Amygdala (3,3%) und Ventraler DC (45,3%)
-26 -24 -8	74	3,64	3,13	0,001	linker Hippocampus
-24 -16 -10		3,09	2,75	0,003	linke Amygdala (3,4%) und linker Hippocampus (23,5%)
32 34 10	203	3,55	3,07	0,001	Pars triangularis des IFG
38 34 10		3,32	2,91	0,002	rechte anteriore Insula
34 16 12		3,03	2,7	0,003	rechtes frontales Operculum

Tab. 4.5 Darstellung der anatomischen Korrelate der ersten Messung der SPES-Aufgabe im Kontrast Angst>Neutral. $p(fwe\ unkorrigiert)=0,01, k>10$.

Da die zweite Messung sich überwiegend mit der ersten Messung gleicht, wird sie hier nicht separat in Bildern dargestellt.

MNI-Koordinaten	Clustergröße kE	T-Wert	Z	P	Anatomisches Korrelat
44 -44 6	1227	7,39	5,01	0	rechter MTG
40 -60 6		5,61	4,26	0	rechter MTG
54 -38 16		5,43	4,17	0	rechter STG
-50 -52 10	1111	5,67	4,28	0	linker MTG
-42 -36 28		5,25	4,08	0	linkes parietales Operculum
-58 -48 12		5,22	4,06	0	linker STG
38 -24 -10	284	5,44	4,17	0	rechter Hippocampus
54 -18 -6		4,42	3,62	0	rechter STG
42 -24 -22		4,13	3,45	0	rechter Gyrus Fusiformis
-30 -6 -18	183	5,14	4,02	0	linke Amygdala
-32 -22 -10		4,95	3,92	0	linker Hippocampus
-38 -4 -24		4,46	3,64	0	linke Amygdala
30 -8 -18	86	5,1	4	0	rechte Amygdala (12,5%) und rechter Hippocampus (27,5%)
22 -6 -16		3,84	3,26	0,001	rechte Amygdala
20 12 -20	76	4,95	3,92	0	rechter medialer Orbitalgyrus
					rechtes Putamen(13%) und rechte Area
12 12 -12		4,44	3,63	0	Accumbens (12%)
20 18 -14		4,34	3,57	0	rechter medialer Orbitalgyrus
-38 -4 40	114	4,82	3,85	0	linker präzentraler Gyrus
-56 4 36		4,19	3,48	0	linker präzentraler Gyrus
16 14 40	81	4,81	3,84	0	rechte SMA
10 20 50		4,01	3,37	0	rechte SMA
-18 -48 48	86	4,65	3,75	0	linker Precuneus
46 -54 -24	192	4,65	3,75	0	rechter Gyrus Fusiformis

Tab. 4.6 Darstellung der anatomischen Korrelate der zweiten Messung der SPES-Aufgabe im Kontrast Angst>Neutral. $p(fwe\ unkorrigiert)=0,001$, $k>10$.

4.2.2 Ergebnisse der Reliabilitätsanalyse

Anhand von 3D-Bildern, welche mittels Cluster-Überlappungsmethode erstellt wurden, werden im Folgenden die Aktivierungen der beiden Messungen in den jeweiligen Kontrasten verglichen. Für die Ergebnisse des ANGST>NEUTRAL Kontrastes wird wie in der Ergebnisdarstellung der ersten Messung zusätzlich die Aktivierung in Schnittbildern dargestellt. Im Anschluss folgt eine detaillierte Vorstellung der Analyse der ICC-Werte.

GESICHTER>KONTROLLE KONTRAST

Im Gegensatz zu der WG-Aufgabe sind die Überlappungen der ersten und zweiten Messung deutlich als großflächige gelbe Markierung erkennbar (siehe Abb.4.13).

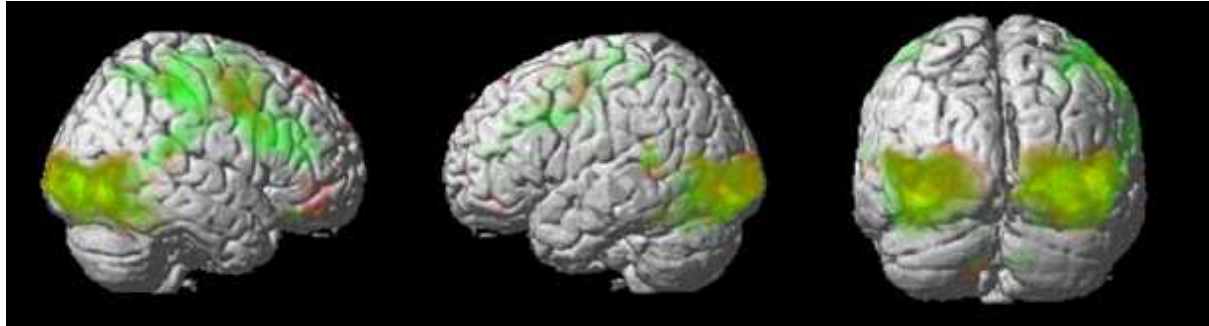


Abb. 4.13 Darstellung der Aktivierungen der beiden Messungen der SPES-Aufgabe im Gesichter>Kontrolle Kontrast als Cluster-Überlappung. $p(\text{unkorrigiert})= 0,001$, $k > 10$, Messung 1 = rot, Messung 2 = grün und Überlappung = gelb.

Für die Reliabilitätsberechnung mittels ICC-Analyse wurde zunächst eine ICC-vs.-T-Karte erstellt (Abb. 4.14). In der Graphik werden die Korrelationen der Aktivierungsstärken einzelner Voxel mit ihrer Reliabilitätsstärke gezeigt. Generell stellt jeder der in Abb. 4.14 gezeigten blauen Kreise einen Voxel dar.

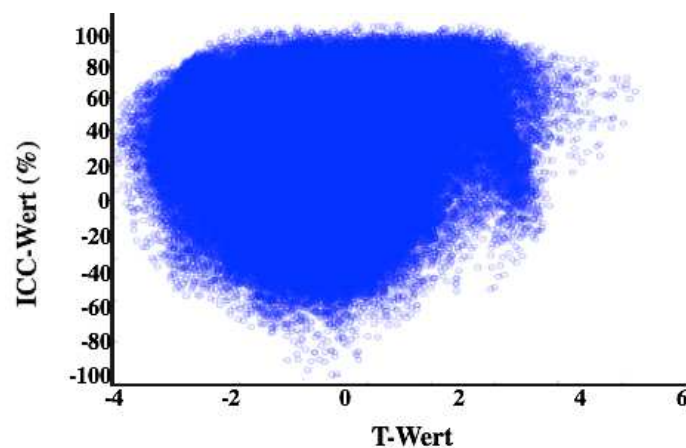


Abb. 4.14 ICC-vs.-T-Karte des Kontrastes Gesichter>Kontrolle der SPES-Aufgabe. Darstellung der T-Werte der ersten Messung und ihren assoziierten ICCs.

Ein Großteil der Voxel liegt hier im Bereich der negativen T-Werte zwischen -4 und 3. Zudem weisen die meisten Voxel höhere ICC-Werte von $ICC \approx 0,2$ bis $ICC \approx 1$. Graphisch wird deutlich, dass bei fast allen T-Werten die meisten Voxel im Bereich hoher ICC-Werte liegen. Somit liegt die Vermutung nahe, dass fast alle hohen T-Werte des beschriebenen Bereiches gut reproduzierbar sind.

Die relative Häufigkeitsverteilung der voxelweisen ICC-Werte wird anhand der ICC-Häufigkeitskarte in Abbildung 4.15 dargestellt. Wie schon bei der Reliabilitätsanalyse der WG-Aufgabe lässt sich anhand der relativen Häufigkeitsverteilung der mediane ICC für das aktivierte Netzwerk und das Netzwerk des gesamten Gehirns der ersten Messung aufzeigen.

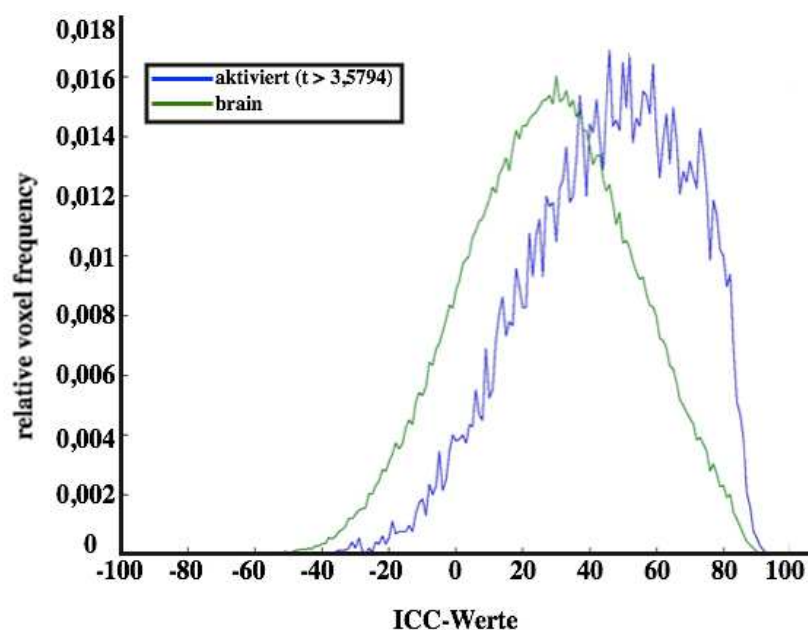


Abb. 4.15 ICC-Häufigkeitskarte des Kontrastes Gesichter>Kontrolle der SPES-Aufgabe. $p(fwe \text{ unkorrigiert}) < 0,001$. Relative Häufigkeitsverteilung der ICCs der Voxel für das aktivierte Netzwerk (blau) und das gesamte Gehirn (grün).

Die Graphik zeigt bei voreingestellten unkorrigierten p-Wert von $p = 0,001$ die erwartete Gauß-förmige Kurve sowohl für das gesamte Gehirn als auch für das aktivierte Netzwerk. Der mediane ICC des gesamten Gehirns liegt bei $ICC_{med} = 0,28$. Dieser Wert deutet darauf hin, dass generell eine gute Reliabilität vorliegen muss, denn in der Berechnung der Reliabilität des gesamten Gehirns fließen auch deaktivierte Voxel mit ein.

Bei einem mittleren Aktivierungswert von $T = 3,5794$ zeigt sich neben dem Gauß-förmigen Verlauf des aktivierten Netzwerkes der mediane ICC von $ICC_{med} = 0,47$. Folglich ist bei

Hirnregionen mit hohen Aktivierungswerten von einer moderaten Reliabilität für die beiden Messungen auszugehen.

Zusammenfassend lässt sich aus den vorangegangenen Analysen sagen, dass für den GESICHTER>KONTROLLE Kontrast sowohl in der Cluster-Überlappungsmethode als auch in den Rechenanalysen der Reliabilität deutlich vergleichbare Aktivierungen gezeigt werden können und sich insgesamt für das gesamte Gehirn, das aktivierte und auch das deaktivierte Netzwerk verhältnismäßig hohe ICC-Werte im gut moderaten Bereich ergeben.

Wie bereits zu Beginn des Kapitels 4.2.1 aufgeführt, umfassen die Hirnregionen, welche im Wesentlichen an der Gesichterverarbeitung beteiligt sind, die drei Bereiche FFA (Fusiform Face Area), den visuellen Kortex und den STS (superiorer Temporalsulcus). Innerhalb der Analyse des Kontrastes GESICHTER> KONTROLLE zeigen sich als am stärksten aktivierte Regionen rechts dominierend der Gyrus Fusiformis und der Gyrus Temporalis entsprechend der FFA. Diese aneinander liegenden Hirnareale werden als Referenzregion gewählt (siehe Abb. 4.15) und der mediane ICC-Wert hierfür mittels einer Maske berechnet, da eine höhere Reliabilität als für das gesamte aktivierte Netzwerk zu erwarten ist.

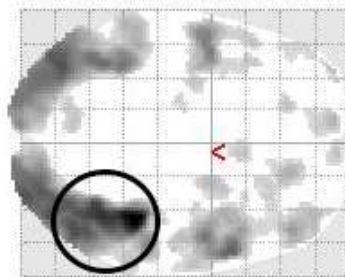


Abb. 4.15 Referenzregion des Kontrastes Gesichter>Kontrolle. p (fwe unkorrigiert) = 0,001. Der Kreis markiert die rechte FFA als Hauptaktivierungszentrum.

Das Ergebnis zeigt bei einem unkorrigierten p -Wert von 0,001 einen medianen ICC im guten Reliabilitätsbereich von $ICC_{med} = 0,62$ für die rechte Fusiform Face Area. Daraus kann geschlossen werden, dass das Paradigma und die Messung insgesamt reproduzierbare Ergebnisse liefern können.

ANGST>NEUTRAL KONTRAST

Die folgenden beiden Abbildungen 4.16 und 4.17 zeigen die Aktivierungsüberlappungen der beiden Messungen im ANGST>NEUTRAL Kontrast. Die erste Messung (rot) zeigt im Vergleich zu der zweiten Messung (grün) mehr Aktivierungen. Im Schnittbild sind zudem deutliche Überlappungen im Bereich der Amygdala und des Hippocampus zu sehen.

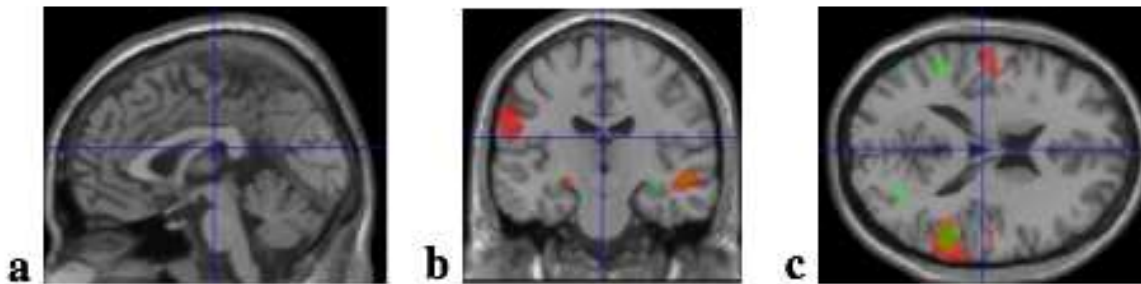


Abb. 4.16 Darstellung der Aktivierungen der beiden Messungen der SPES-Aufgabe im Angst>Neutral Kontrast in Schnittbildern. a) Sagittalschnitt b) Koronarschnitt c) Transversalschnitt $M1: p(\text{unkorrigiert})=0,01, k>10$ und $M2: p(\text{unkorrigiert})=0,001, k>10$. Messung 1 = rot, Messung 2 = grün und Überlappung = gelb.

Die Darstellung im 3D-Bild zeigt im Vergleich zu dem GESICHTER>KONTROLLE Kontrast insgesamt weniger Aktivierungen, welche sich darüber hinaus in geringerem Ausmaß zu überlappen scheinen.

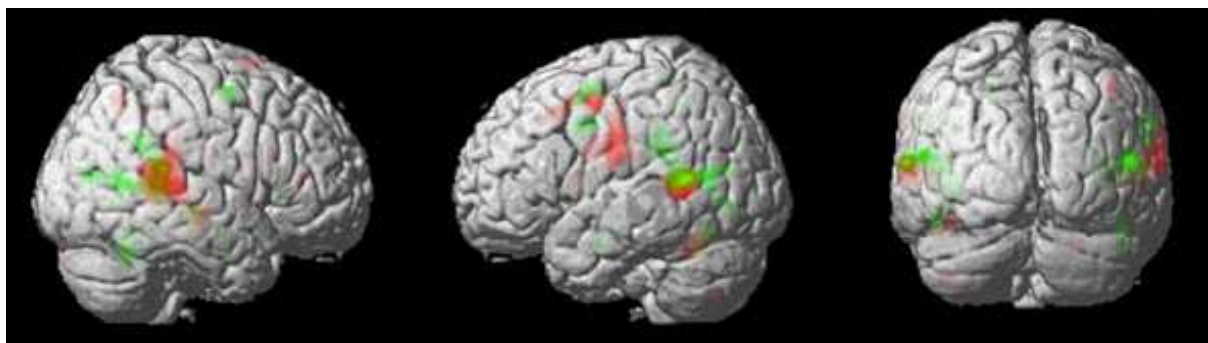


Abb. 4.17 Darstellung der Aktivierungen der beiden Messungen der SPES-Aufgabe im Angst>Neutral Kontrast als Cluster-Überlappung. $M1: p(\text{unkorrigiert})=0,01, k>10$ und $M2: p(\text{unkorrigiert})=0,001, k>10$. Messung 1 = rot, Messung 2 = grün und Überlappung = gelb.

Als stringenter Methode der Reliabilitätserschaffung wurde wie bereits in den Analysen zuvor eine ICC-vs.T-Karte erstellt (Abbildung 4.18).

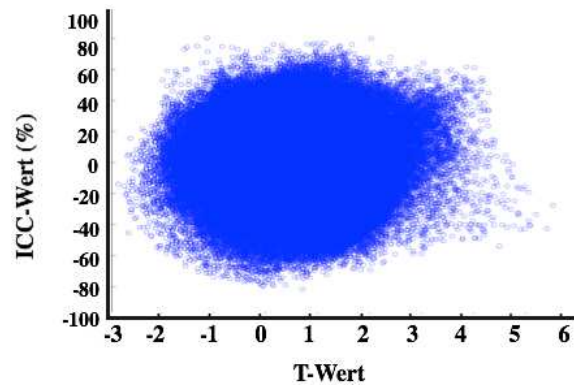


Abb. 4.18 ICC-vs.-T-Karte des Kontrastes Angst>Neutral der SPES-Aufgabe. Darstellung der T-Werte der ersten Messung und ihren assoziierten ICCs. Jeder blaue Kreis repräsentiert einen Voxel.

Die meisten Voxel liegen im Bereich von Aktivierungswerten zwischen -2 und 3 und ICC-Werten von -0,4 bis 0,7. In der Graphik wird in Ansätzen sichtbar, dass stärkere Aktivierungen mit höheren ICC-Werten einhergehen und andersherum. Allerdings liegen die höchsten ICC-Werte im Bereich von T-Werten zwischen 1 und 2. Im Vergleich zu der ICC-vs.-T-Karte des GESICHTER>KONTROLLE Kontrastes lässt sich beobachten, dass insgesamt geringere Aktivierungs- und ICC-Werte vorliegen. Dies legt die Vermutung nahe, dass die aktivierten Areale für die Wahrnehmung von Emotionen nicht so gut reproduzierbar wie die Gesichtserkennung sind.

Die nächste Abbildung 4.19 zeigt die relative Häufigkeitsverteilung der voxelweisen ICCs.

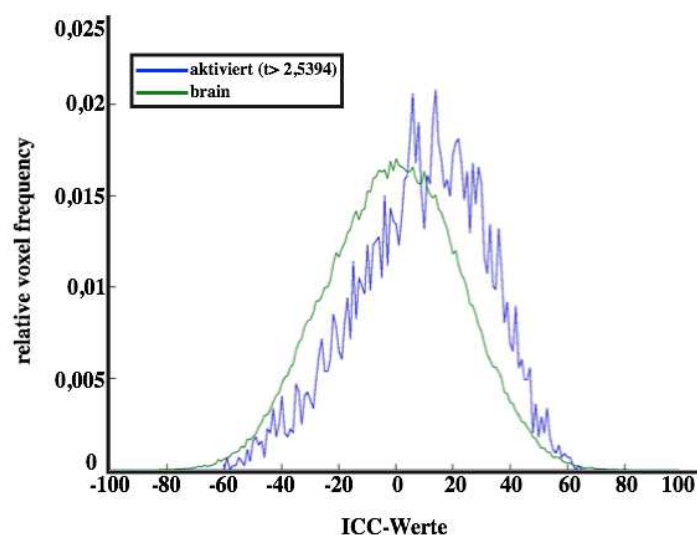


Abb. 4.19 ICC-Häufigkeitskarte des Kontrastes Angst>Neutral der SPES-Aufgabe. $p(\text{unkorrigiert}) < 0,001$. Relative Häufigkeitsverteilung der ICCs der Voxel für das aktivierte Netzwerk (blau) und das gesamte Gehirn (grün).

Bei einem unkorrigierten p-Wert von $p = 0,001$ zeigt sich wie schon bei dem GESICHTER>KONTROLLE Kontrast der erwartete Gauß-förmige Kurvenverlauf. Der mediane ICC für die Aktivierungen des gesamten Gehirns liegt bei $ICC_{med} = -0,004$, was auf eine generell geringe Reliabilität hinweist.

Das aktivierte Netzwerk zeigt ebenfalls eine Gauß-förmige Kurve und bei einem mittleren Aktivierungswert von $T = 2,5384$ den $ICC_{med} = 0,11$. Dieser Wert deutet zwar insgesamt eine schwache Reliabilität des Kontrastes an, allerdings können durchaus einzelne Aktivierungscluster eine moderate Reliabilität aufweisen und reproduzierbar sein.

Zusammenfassend lässt sich aus der Cluster-Überlappungsmethode und der Rechenanalyse schließen, dass die Aktivierungen des ANGST>NEUTRAL Kontrastes nicht so reliabel messbar sind wie in dem GESICHTER>KONTROLLE Kontrast. Die niedrige Reliabilität ist kongruent mit den geringen Überlappungen der beiden Messungen, welche mit der Cluster-Überlappungsmethode sichtbar sind. Es ist jedoch nicht auszuschließen, dass bei diesen Reliabilitätswerten einzelne Aktivierungsareale durchaus reproduzierbar sind, wie sich mit der nun folgenden, genaueren Rechenanalyse der Referenzregionen feststellen lässt.

Für die Verarbeitung von Emotionen sind hauptsächlich die Amygdala und temporoparietale sowie präfrontale Regionen bekannt (Fusar-Poli et al., 2009; Suslow et al., 2003; Schneider und Fink, 2013). Insbesondere die Emotion Angst löst bekanntermaßen eine starke Aktivierung im Bereich der Amygdala aus, wobei aktuell fraglich ist, inwiefern diese tatsächlich in der Amygdala und nicht der umgebenden Strukturen lokalisiert ist (Boubela et al., 2015).

Als Referenzregion wurden daher im ANGST>NEUTRAL Kontrast der SPES-Aufgabe drei Referenzregionen ausgewählt, deren mediane Reliabilitäten überprüft wurden, da sie hier erwartungsgemäß deutlich höher als im gesamten aktivierten Netzwerk liegen sollten. Diese Referenzregionen bilden der rechte MTG (siehe Abb.4.20) sowie die linke und rechte Amygdala (siehe Abb. 4.21).

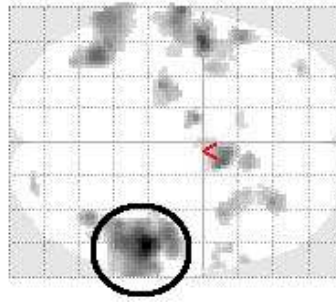


Abb. 4.20 Referenzregion des Kontrastes Angst>Neutral: rechter MTG. p (fwe unkorrigiert)= 0,001. Der Kreis markiert den rechten MTG als Hauptaktivierungszentrum.

Die Berechnung des medianen ICC für den rechten MTG ergab bei einem unkorrigierten p -Wert von 0,001 ein Ergebnis von $ICC_{med} = 0,41$. Die medianen ICC-Werte für die bilateralen Amygdalae ergab linksseitig den $ICC_{med} = -0,25$ und rechtsseitig den $ICC_{med} = -0,03$.

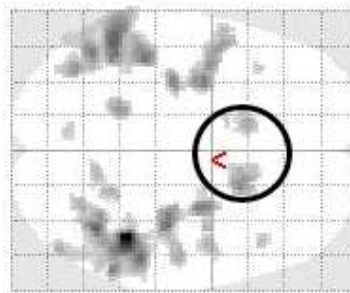


Abb. 4.21 Referenzregion des Kontrastes Angst>Neutral: bilaterale Amygdala. p (unkorrigiert)= 0,001. Der Kreis markiert die rechte und linke Amygdala als Hauptaktivierungszentrum.

Die genaue Analyse der Referenzregionen bestätigt, dass innerhalb der als insgesamt schlecht reliabel ausgewerteten aktivierten Netzwerke des ANGST>NEUTRAL Kontrastes (ICC_{med} von 0,11) mit dem rechten MTG sowohl Hirnregionen von moderater Reliabilität als auch beispielsweise mit den Amygdala solche mit schlechter Reliabilität liegen. Das Paradigma und die Messung können daher reproduzierbare Ergebnisse liefern, allerdings mit nur moderater Reliabilität und fraglicher Beteiligung der Amygdala am Prozess der Verarbeitung von emotionalen Stimuli.

5 DISKUSSION

Im folgenden Kapitel werden die Ergebnisse der Reliabilitätsprüfung der beiden Paradigmen SPES und WG diskutiert. Zunächst wird eine Zusammenfassung der Ergebnisse gegeben, der eine kritische Auseinandersetzung mit der Methodik dieser Arbeit folgt. Im Anschluss daran werden die Ergebnisse der jeweiligen Paradigmen anhand der Daten der Gruppen- und Reliabilitätsanalyse und des Designs diskutiert und mit aktuellen Forschungsergebnissen verglichen.

5.1 Zusammenfassung

In dieser Arbeit wurden die beiden Paradigmen Wortgenerierung und die subliminale Präsentation emotionaler Stimuli hinsichtlich ihrer Reliabilität im Rahmen der Longitudinalstudie „*Wirkungen von genetischen Risikovarianten für psychische Störungen auf Gehirnstruktur und –funktion bei Gesunden und Erkrankten*“ geprüft. Hierbei fanden sich sowohl Unterschiede zwischen den Reliabilitäten der einzelnen Paradigmen als auch innerhalb der jeweiligen Paradigmen.

Die WG-Aufgabe zeigte erwartungsgemäße Aktivierungen in den Hirnregionen des frontalen Gyrus, des medialen Temporallappens, des Gyrus Cinguli und des motorischen Kortex. Allerdings lagen die errechneten Reliabilitäten dieser Aktivierungen insgesamt im schlechten Bereich. Die aktivierten Hirnregionen zeigten erstaunlicherweise keine höhere Reliabilität als Bereiche, die nicht aktiv waren. Zudem waren in beiden Messungen deutlich mehr Aktivierungen während der Kontrollphase zu sehen, was auf Probleme mit dem Design hinweist. Zusammenfassend ist somit die Reliabilität der WG-Aufgabe als sehr schlecht zu beurteilen.

Hinsichtlich der SPES-Aufgabe muss zunächst zwischen den beiden analysierten Kontrasten unterschieden werden. Der erste untersuchte Kontrast GESICHTER > KONTROLLE sollte zum einen überprüfen, ob das Paradigma und die Messung insgesamt reproduzierbare Ergebnisse liefern können und zum anderen, ob die erwartungsgemäßen Aktivierungen in den für die Gesichterverarbeitung bekannten Hirnregionen liegen. Es zeigten sich in der Auswertung sowohl die charakteristischen Aktivierungen in den drei Bereichen FFA, visueller

Kortex und STS als auch in der Reliabilitätsanalyse für das aktivierte und das deaktivierte Netzwerk verhältnismäßig hohe mediane ICC-Werte im moderaten bis guten Reliabilitätsbereich.

Der Kontrast ANGST>NEUTRAL diente der Überprüfung der Messung von Emotionsverarbeitung am Beispiel der Emotion Angst. Wie bereits in dem ersten untersuchten Kontrast konnten in dieser Studie Aktivierungen in temporoparietalen und präfrontalen Hirnregionen sowie subkortikale Strukturen, insbesondere die Amygdala, reproduziert werden. Allerdings wiesen die aktivierten Netzwerke insgesamt eine nur geringe Reliabilität auf. Erst nach genauer Rechenanalyse einzelner Referenzregionen konnte eine moderate Reliabilität im rechten MTG und eine unzureichende Reliabilität in den Amygdala nachgewiesen werden. Zusammenfassend gilt die Reliabilität der SPES-Aufgabe somit als moderat bis gut zu beurteilen.

5.2 Methodenkritik

Die Zielsetzung dieser Arbeit war es, die Reliabilität der fMRT-Paradigmen Wortgenerierung und Subliminale Präsentation emotionaler Stimuli zu untersuchen. Unter Reliabilität ist hierbei die Stabilität der inter- und intraindividuellen Varianz in der BOLD-Signalstärke über einen definierten Zeitraum gemeint. Dies steht vor dem Hintergrund, dass die Stabilität von fMRT-Daten zu unterschiedlichen Zeitpunkten durch eine hohe Variabilität beeinträchtigt werden kann, die auf vielseitige Ursachen zurückgeführt werden kann. Die Varianz der Daten ist im Wesentlichen von drei Faktoren abhängig, welche im Folgenden diskutiert werden: dem experimentellen Design sowie den experimentellen und technischen Bedingungen.

5.2.1 Experimentelles Design

Die Wahl des experimentellen Designs hat auf die Ergebnisse und somit in dieser Studie auf die Reliabilität einen wesentlichen Einfluss (Fließbach et al., 2010, Bennett und Miller 2013). Im Vergleich mit ereigniskorrelierten Designs zeigen Blockdesigns eine stärkere statistische Sensitivität und damit einhergehend auch eine größere SNR (*Signal-to-noise ratio*). Zu beachten ist zusätzlich, dass motorische und sensorische Paradigmen eine höhere Reliabilität aufweisen können (Bennett und Miller, 2010). Für eine gute Einschätzung der

Parameter für die Datenanalyse sollte ein Design mindestens 100 Einzelereignisse beinhalten. Dies geht in dieser Studie allerdings auf Kosten der Länge des Experimentes. Inklusive der anatomischen Messungen dauerten die Aufnahmen der ersten Messung 54 Minuten und die der zweiten Messung noch 39 Minuten, was eine lange kognitive Beanspruchung erforderte und als zusätzlicher Stressfaktor für die Proband_innen gewertet werden kann.

Die Reihenfolge der Paradigmen wurde in diesem Experiment bewusst so gewählt, dass sich Aufgaben mit aktiver und bewusster kognitiver Beanspruchung mit denen von passiver Beanspruchung abwechselten. Dies sollte gewährleisten, dass die Proband_innen sich gefühlt erholen konnten. Diese Erholung sollte allerdings nicht auf Kosten der Aufmerksamkeit gehen. In der SPES-Aufgabe gibt es keinerlei Bestätigung über die Aufmerksamkeit der Proband_innen und es bleibt fraglich, inwiefern sich diese tatsächlich nach den zuvor bewältigten Paradigmen noch konzentrieren konnten. Vor allem während der kurzen subliminalen Exposition hätte die Einblendung eines Fixierungskreuzes die Aufmerksamkeit auf den Bildschirm bessern können (Dannowski und Suslow, 2006; Specht et al., 2003). Eine weitere Möglichkeit der späteren Überprüfung der Aufmerksamkeit wäre hier die Anwendung des sogenannten ‚*eyetracking*‘ (Blickerfassung) gewesen. In der WG-Aufgabe wurde die Aufmerksamkeit durch die Aufnahmen mit dem Programm Audacity verfolgt. Die Auswertung dieser Aufnahmen hätte zudem einen zusätzlichen Erkenntnisgewinn darüber geben können, wie viele Wörter die Proband_innen generieren konnten. Die Anzahl der generierten Worte hätte als parametrischer Regressor in der Datenanalyse der mit der Artikulation produzierten Artefakte herausgerechnet werden können (Wende et al., 2012). Allerdings hat die Auswertung den Rahmen dieser Studie in Anbetracht der insgesamt zwei Paradigmen mit Subanalysen überschritten und wurde nicht zusätzlich verfolgt.

Hinsichtlich des Stimulusmaterials wurden in der WG-Aufgabe sogenannte ‚Low Level Baselines‘ angewandt. Diese zeichnen sich durch eine minimale Aktivierung von nicht interessierenden Regionen aus. Beispielsweise wurden lediglich die Symbole ‚+‘ in der Aktivierungsphase und ‚#‘ in der Ruhephase angezeigt, um eine starke Aktivierung des visuellen Kortex zu vermeiden. Generell gehen Low Level Baselines mit höheren T-Werten einher und erzeugen dadurch bessere Reliabilitäten (Brandt et al., 2013). Es ist jedoch nicht

auszuschließen, dass während des Wortgenerierungsprozesses durch die Erinnerung an die Begriffe selber emotional kognitive Prozesse zusätzlich ausgelöst wurden.

In der SPES-Aufgabe wurden als Stimulusmaterial monochrome standardisierte Gesichter von Schauspieler_innen verwandt, welche bereits in einigen Studien zuvor auf ihre Fähigkeit, Emotionen auszulösen, überprüft wurden (Murphey und Zajonc, 1993; Dannlowski und Suslow, 2006).

Es ist wissenschaftlicher Standard, das experimentelle Design so zu entwerfen, dass durch die Unterschiede in der Aktivierungs- und Kontrollphase genau die Aktivierungsunterschiede erfasst werden können, die typisch für die Aufgabe sind, um darüber eine Aussage über die Reliabilität treffen zu können (Bennett und Miller, 2010).

Am Beispiel der WG-Aufgabe ist die Kalkulation der Reliabilität fragwürdig, weil hier untypischerweise deutlich mehr Aktivierungen in der Kontrollphase auftraten. Theoretisch könnte die zuvor abgelaufene Präsentation des Paradigmas ‚*Resting State*‘ einen Einfluss auf die Aktivierungen der WG-Aufgabe gehabt haben. Bennett und Miller zeigten in einer Studie 2013, dass beispielsweise Gedächtnisaufgaben einen verstärkenden Einfluss auf die kognitiven Leistungen haben können (Bennett und Miller, 2013) und legten die Vermutung nahe, dass ebenso eine Interaktion möglich ist, die in verminderten kognitiven Aktivierungen resultiert.

5.2.2 Experimentelle Bedingungen

Zunächst einmal soll das Probandenkollektiv an dieser Stelle beleuchtet werden. Die Gruppengröße von 20 Proband_innen wurde gewählt, da die probandenspezifische Fehlerentstehung ein großes Kollektiv benötigt, um statistisch einzelne Fehler berechnen zu können (Benett und Miller, 2010). Bisherige Reliabilitätsstudien schlossen in der Regel nicht mehr als 10 Proband_innen ein, so dass insgesamt von einer ausreichenden Aussagekraft ausgegangen werden kann. Hinsichtlich des Alters von 20 bis 28 Jahren und des Bildungsgrades ist ein sehr homogenes Probandenkollektiv gewählt worden. Dies hat einerseits den Vorteil, dass dadurch stark unterschiedliche Ergebnisse vermieden werden konnten, die Übertragbarkeit auf andere Probandenkollektive oder sogar Patient_innen jedoch massiv einschränkt ist. Schließlich wurde noch anamnestisch ausgeschlossen, dass die Proband_innen an psychiatrischen Erkrankungen litten. Diesbezüglich hätte ein

strukturiertes Interview für die DSM-IV (SKID) psychiatrische Störungen in einer standardisierten Weise ausschließen können.

Einen großen Einfluss auf die Reliabilität hat der kognitive Status der Proband_innen. Er ist über einen Zeitraum hinweg stets variabel und muss daher unabdinglich in der Test-Retest Diskussion eingebracht werden. Hierbei spielen der allgemeine Zustand wie die Tagesform, die emotionale Befindlichkeit, die Menge an Schlaf und Kaffeeconsum, die Anspannung und natürlich auch die physiologischen Einflüsse wie beispielsweise Harndrang während der Messung eine Rolle. Auch die allgemeine Kooperationsbereitschaft der Proband_innen kann zu der kognitiven Leistung beitragen bzw. sie minimieren. Hinzu kommen charakterliche Eigenschaften wie die generelle Fähigkeit zur Konzentration oder das Anspannungsniveau. In dieser Studie wurden die Proband_innen zusätzlich ausdrücklich im Vorfeld auf die Vermeidung von Medikamenten- und Drogenconsum hingewiesen, sie diesbezüglich aber nicht kontrolliert, weshalb die etwaige Beeinflussung solcher nicht ausgeschlossen werden kann.

Generell sollte bei jeder fMRT Studie vermieden werden, dass offensichtliche und vermeidbare Fehlerquellen zu Varianzen in den kognitiven Fähigkeiten der Proband_innen führen und daher möglichst gleiche Bedingungen für beide Messungen antizipiert werden. Der Tageszeitpunkt stellt einen beeinflussbaren Faktor dar, der die kognitiven Fähigkeiten durch hormonelle zirkadiane Schwankungen verändern kann (Carrier und Monk, 2000). Zwar wurde bei der Terminwahl mit den Proband_innen versucht, möglichst gleiche Tageszeiten auszuwählen, jedoch war dies aus organisatorischen Gründen nicht immer möglich. Durch die Wahl des Messabstandes von einem Monat wurde darüber hinaus bewusst versucht, hormonelle Zyklusschwankungen der Probandinnen zu minimieren. Geschlechtsspezifische, hormonelle und auch zyklusspezifische Aktivierungsunterschiede können durchaus in Aufgaben wie der semantischen Sprachproduktion auftreten (Konrad et al., 2008). Aus diesem Grunde wurden weiterhin nur Frauen in die Studie aufgenommen, welche die Einnahme von oralen Kontrazeptiva bestätigten. Die Wahl eines kurzen vierwöchigen Test-Retest Intervalles ist zudem eines der Kriterien, welche die Reliabilität verbessern können (Fließbach et al., 2010, Bennet und Miller, 2012). Längere Zeitintervalle gehen mit einer zunehmenden Reliabilitätsminderung einher.

Schließlich muss durch den Prozess der wiederholten Testung auch der sogenannte ‚Trainingseffekt‘ in der Auswertung der zweiten Messung Beachtung finden. Es ist davon auszugehen, dass die Proband_innen sich besser auf die Aufgabenstellungen in der wiederholten Messung konzentrieren können und Unsicherheiten darüber nach der ersten Messung klären konnten. Dies kann zu stärkeren kognitiven Aktivierungen der antizipierten Hirnfunktionen führen (Lohmann, Deppe, Jansen und Knecht, 2004).

Die bisher beschriebenen probandenabhängigen Varianzen beziehen sich vor allem auf die kognitiven und physiologischen Veränderungen zwischen zwei Messungen. Allerdings sind kognitive Veränderungen auch innerhalb einer Messung möglich, beispielsweise durch Aufmerksamkeitsminderung (Specht et al., 2003), Aufregung und Änderungen der kognitiven Strategien zur Bewältigung einer Aufgabe während der Aktivierungsphase. Ein relativ beeinflussbarer Faktor ist allerdings die Aufregung der Proband_innen. Erregung, Angst und Stress können sich durch die assoziierte Hyperventilation zu einer Verminderung des Kohlendioxidpartialdrucks auf die Signalhöhe auswirken (Foki, 2010). Es wurde daher sehr viel Wert auf die ausführliche Vorbereitung der Proband_innen für die Testung gelegt. Ein einige Tage vor der ersten Messung ausgehändigter Aufklärungsbogen sollte den Proband_innen eine erste Vorstellung über den Ablauf der Messungen geben. Schließlich wurde jede Aufgabe vor der Messung nochmals ausführlich mit den Proband_innen besprochen und anhand von einer Computertestsimulation eingeübt.

Trotz all dieser Beachtungen stellten die ungewohnte räumliche Einengung und der Geräuschpegel des Magnetresonanztomographen einen sicherlich weiterhin nicht unerheblichen Einflussfaktor auf die Aufregung der Proband_innen dar.

5.2.3 Technische Bedingungen

5.2.3.1 Datenaquisition

Bennett und Miller stellten 2010 in einer Übersichtsarbeit über Test-Retest Verfahren den Einfluss verschiedener Rauschquellen unabhängig von der Probandenwahrnehmung in den Vordergrund der Reliabilitätsanalyse. Ziel einer Reliabilitätsüberprüfung ist demnach die routinemäßige Detektion von Fehlerquellen, welche durch erhöhtes Rauschen, z.B. des Gerätes oder der Proband_innen entstehen, und somit nicht aufgabenbezogene kognitive

Prozesse herauszurechnen. Eine besonders wichtige Aussagekraft zur Qualität der fMRT-Daten besitzt diesbezüglich die SNR (*Signal-to-noise ratio*). Sie basiert auf dem Phänomen der in der funktionellen Bildgebung als ‚Noise‘ bekannten Rauschkulisse. Eine Anhebung der relevanten Signale unter gleichzeitiger Reduktion der Rauschkulisse resultiert in einer Maximierung der SNR und somit in einer Verbesserung der Datenqualität. Bennett und Miller unterscheiden zwei Subtypen der SNR. Die Bild-SNR bezieht sich auf die Datenqualität innerhalb eines fMRT-Volumens und die temporale SNR erweitert diese um die zeitliche Dimension, indem sie die Bild-SNR für jeden Voxel zu unterschiedlichen Zeitpunkten angibt (Gleichung 5.1). Hierzu wird der Quotient des mittleren Signalwertes aller Voxel in einer Zeitserie (μ timeseries) und der Standardabweichung aller Voxel in einer Zeitserie (σ timeseries) gebildet.

$$\text{SNR}_{\text{temporal}} = \frac{\mu \text{ timeseries}}{\sigma \text{ timeseries}} \quad (5.1)$$

Analog zu der temporalen SNR beschreiben Benett und Miller die Relevanz der temporalen CNR (*Contrast-to-noise ratio*), welche den maximalen Unterschied der Signalintensität der einzelnen Voxel zu verschiedenen Zeitpunkten beschreibt und somit den Unterschied der Aktivierung zweier Kontraste im Rahmen des Experiments (Bennett und Miller, 2010). Welchen Einfluss die technischen Bedingungen auf diese beiden Einflussgrößen haben können, wird im Folgenden diskutiert.

Grundsätzlich kann eine bessere SNR durch höhere Magnetfeldstärken des Scanners erreicht werden (Fließbach et al., 2010). Ein wie in dieser Arbeit verwendeter 3-Tesla Scanner kann die SNR im Vergleich zu einem 1,5-Tesla Gerät sogar verdoppeln. Hierdurch entstehen allerdings auch zusätzliche Artefakte in Regionen, die bspw. von stark lufthaltigen Strukturen wie dem orbitofrontalen Kortex umgeben sind. Durch eine anders gewählte Orientierung der Schichtaufnahmen als die in dieser Arbeit gewählte AC/PC Schichtpositionierung könnte wiederum auch die SNR in solchen Artefakt-anfälligen Regionen deutlich verbessert werden (Deichmann, Hutton und Turner, 2003). Vorteilhaft für die Reliabilität kann ebenso die Auswahl bestimmter Parameter für die Datenaquisition sein. Hierzu gehört die Auslassung von sogenannten ‚gaps‘, also Lücken zwischen den Schichtaufnahmen, wie bei der Parameterbestimmung dieser Arbeit berücksichtigt. Um darüber hinaus den Austausch von Signalen an Schichtgrenzen und somit die Entstehung

von Artefakten zu vermeiden, wurden die Aufnahmen überlappend (*interleaved*) aufgenommen. Ein weiterer, die Reliabilität beeinflussender Parameter, stellt die Voxelgröße dar. Die Auswahl der Voxelgröße mit $3,0 \text{ mm}^3$ von dieser Studie wurde getroffen, um die SNR zu vergrößern, gleichzeitig wird hierbei allerdings die räumliche Auflösung etwas verschlechtert (Fließbach et al., 2010).

Schließlich sind in Anbetracht der geringen BOLD-Signalstärke, die kaum über dem Rauschpegel des MRT-Gerätes liegt, routinemäßige Qualitätskontrollen empfehlenswert. Hierdurch könnten Inhomogenitäten im Magnetfeld, welche die SNR stark beeinflussen, verhindert werden (Schneider und Fink, 2013). Im MR-Tomograph, an dem die Messungen dieser Studie durchgeführt wurden, werden daher dreimal wöchentlich Kontrollmessungen mit einem MR-Phantom durchgeführt.

Insgesamt lässt sich konstatieren, dass der Einfluss der fMRT-Technik auf die Reliabilität schwerwiegend sein kann, sich jedoch beeinflussen lässt. Eine optimale Bildgewinnung ist mit den bisherigen technischen Möglichkeiten allerdings noch nicht möglich. Es kann in technischer Hinsicht jedoch davon ausgegangen werden, dass die bestmöglichen Bedingungen für die Datenaquisition dieser Studie bestanden.

5.2.3.2 Datenanalyse

Es gibt einige unterschiedliche Methoden der Reliabilitätsberechnung, die bereits in Kapitel 3.6.1 vorgestellt worden sind. In dieser Arbeit wurden die Methoden der *Cluster*-Überlappung und der ICC-Stabilitätsprüfung angewandt.

Bei der Methode der *Cluster*-Überlappung wird überprüft, welche *Cluster* ab einem gewissen Schwellenwert in beiden Messungen Aktivierungen zeigen. Mathematisch ließe sich dies mittels *Dice*- oder *Jaccard*- Koeffizient ausdrücken und optisch als 3D- Bilder darstellen. Hier wurde die Darstellung mittels 3D- Bildern gewählt. Der wesentliche Nachteil dieser Methode ist, dass der Schwellenwert von der auswertenden Person frei gewählt werden kann und sich das Ergebnis daher mit einer liberalen Wahl an Schwellenwerten stark beeinflusst werden kann.

Während die Methode der *Cluster*-Überlappung demzufolge mehr als orientierendes optisches Maß der Reliabilität gewertet werden kann, bietet die Berechnungsanalyse der

ICCs eine sehr stringente Stabilitätsüberprüfung der Ergebnisse. Zur Berechnung der ICCs wird die inter- und intraindividuelle Varianz der Proband_innen verwandt. Reproduzierbare Aktivierungen sind hierbei solche, deren intraindividuelle Varianz kleiner als ein bestimmter Grenzwert ist (Bland und Altman, 1986). Somit stellt die intraindividuelle Varianz den eigentlichen Marker der Fehlerbestimmung dar. Die ICC-Stabilitätsprüfung kann zum einen für das gesamte Gehirn oder für einzelne ROIs berechnet werden.

Der größte Vorteil der ICC-Berechnungsanalyse ist sicherlich die Tatsache, dass sie einfach anzuwenden und zu interpretieren ist. So stellt sie mittlerweile eine Art Standard in der Reliabilitätsforschung dar, da sie das gängigste Maß zur Analyse von fMRT-Reliabilitäten ist (Specht et al., 2003). Die Beachtung der Aktivierungsstärken der signifikanten *Cluster* ist das wesentliche Merkmal der ICC-Berechnungsweise, welches die Aussagekraft der Reliabilitätsmethoden verfeinert hat. Eine hohe Reliabilität würde demnach aus einer möglichst gleichen Aktivierungsstärke innerhalb der Messungen einer einzelnen gemessenen Person bei einer unterschiedlichen Aktivierungsstärke zwischen allen gemessenen Personen resultieren (Fließbach et al., 2010). Hierbei zeigt sich jedoch bereits eine wesentliche Schwäche der ICC-Berechnungsweise. Sie ist zugleich von inter- als auch von intraindividuelle Varianz abhängig. Würde also der exakt gleiche Grad an intraindividuelle Varianz tatsächlich vorliegen, wäre theoretisch eine Stabilität der Test- und Retest-Werte erreicht. Allerdings gilt dies nur für eine gewisse Studienpopulation. In Experimenten mit größeren interindividuellen Varianzen, wie es in klinischen Populationen häufig der Fall ist, wäre die berechnete Reliabilität kaum mehr übertragbar, da hier andere ICC-Werte berechnet würden als bei einem Experiment mit geringer interindividuelle Varianz (Bennet und Miller, 2013). Dies ist sehr wichtig für die Ergebnisse dieser Studie, da hier die ICC-Werte für eine sehr homogene akademische Studienpopulation berechnet wurden. Die Übertragung dieser ICC-Werte auf ein Probandenkollektiv aus dem psychiatrischen Erkrankungsfeld ist daher als problematisch zu bewerten.

Ein weiterer umstrittener Punkt bei der Berechnung der ICC-Werte ist die Möglichkeit, die Aktivierungsstärken mit den Reliabilitäten bestimmter Hirnareale zu korrelieren (Caceres et al., 2009). Einerseits verfügt diese Methode über die Aussagekraft, wie stark gut reproduzierbare Areale tatsächlich aktiv waren. Andererseits ist dabei zu beachten, wieso die geringen Aktivierungswerte erfolgt sind. Beispielsweise könnte die Aktivierung in der

Kontrollphase ähnlich hoch gewesen sein oder aus einer geringeren BOLD-Reaktion als der Erwarteten resultieren.

Insgesamt ist die ICC-Berechnung sehr rauschanfällig, was sich in generell geringeren Reliabilitätswerten zeigen kann. Zudem ist die voxelweise Berechnung der ICC_{med} nicht zwangsläufig bedeutungsvoll, da aktivierte Netzwerke oft aus inhomogenen funktionellen Gebieten bestehen. Um anatomische und funktionelle Rahmenbedingungen zu kombinieren, eignet sich die ROI-Analyse. Sie prüft, wie zuverlässig Aktivierungen in interessierenden aufgabenspezifischen Hirnarealen vorliegen. Da diese Analyse einen größeren Voxelbereich einberechnet, ist sie weniger rauschanfällig. Nachteilig ist hier jedoch, dass nur Bereiche, die mit der Aufgabe assoziiert werden, in die Kalkulation eingehen und teilweise aufgabenspezifische Aktivierungen außerhalb der anatomischen Region unbeachtet bleiben und vice versa. In dieser Arbeit wurden daher nur aktivierte Voxel über einem festgelegten Schwellenwert in den interessierenden Regionen berücksichtigt.

Weiterhin muss die Auswahl an ICC-Berechnungsweisen, wie sie in Kapitel 3.6 beschrieben wurden, als kritisch betrachtet werden. Der in dieser Arbeit verwendete Subtyp ICC 3,1 wurde mit Hilfe des ICC-Algorithmus nach Shrout und Fleiss gewählt. Allerdings ist bisher nicht konkret definiert worden, wann welcher ICC zur Anwendung kommen sollte. Dies kann die Wahl unangemessener Varianten begünstigen und somit Unterschiede in den Reliabilitätsanalysen bewirken (Bennett und Miller, 2010).

Als letzter Kritikpunkt soll an dieser Stelle sowohl bei der *Cluster*-Überlappungsmethode als auch an der ICC-Berechnungsanalyse die Abhängigkeit des Vergleiches von Voxeln angeführt werden. Gemäß dem aktuellen neurowissenschaftlichen Forschungsstand gilt grundlegend anzunehmen, dass Aktivitäten in Form von korrelierenden Neuronengruppen stattfinden, welche wiederum räumlich zusammenhängend mit anatomischen Formen übereinstimmen (Cecchi, Garg und Rao, 2009). Die fMRT Datenanalyse mittels voxelweiser Aktivitätsangaben hat den Nachteil, dass sie kleine Unterschiede in der Position der Voxel nicht auf die interindividuelle Ebene übertragen kann und nur auf Kosten von räumlicher Auflösung im Vorverarbeitungsprozess zu minimieren versucht. Dabei gehen Aktivierungsdaten, die an Voxelgrenzen liegen, teilweise verloren. Als ein neuer Algorithmus, um funktionelle *Cluster* zu identifizieren, könnte die sogenannte '*weighted cluster coverage*' (gewichtete *Cluster*-

Erfassungs-) Methode nach Cecchi et al. bessere Möglichkeiten bieten, um die Information räumlich zusammenhängender *Cluster* zu nutzen.

Zusammenfassend gilt es mehrere Schwächen in der Anwendung der ICC basierten Reliabilitätsbestimmung zu beachten. Die Übertragbarkeit der Ergebnisse ist sehr begrenzt und sollte daher stets in seiner Limitation betrachtet werden.

5.3 Ergebnisdiskussion

Dieses Kapitel widmet sich der Bewertung und Interpretation der Ergebnisse der beiden Paradigmen. Zunächst werden die Ergebnisse der Gruppenanalyse und der Reliabilitätsanalyse getrennt beleuchtet und anschließend zusammenhängend bewertet. Im Anschluss werden die Ergebnisse mit aktuellen Forschungsergebnissen verglichen und die Schwächen und Stärken der Designs der beiden Paradigmen analysiert.

Grundlegend für die Bewertung der Reliabilität werden hierzu folgende ICC-Wertigkeiten als Referenz gewählt: $ICC \leq 0,4$ – schlecht reliabel, $0,4 < ICC \leq 0,6$ – moderat reliabel, $0,6 < ICC \leq 0,8$ – gut reliabel und $ICC > 0,8$ sehr gut reliabel (Caceres et al., 2009, Fließbach et al., 2010, Shrout und Fleiss, 1979).

5.3.1 WG-Aufgabe

5.3.1.1 Bewertung der Gruppenanalyse

Erwartungsgemäß zeigt eine Wortgenerierungsaufgabe hauptsächlich eine Aktivierung im Bereich des als Broca-Region bekannten Pars triangularis und Pars Opercularis des inferioren Gyrus Frontalis (IFG) und im Bereich des motorischen Sprachzentrums (siehe Kapitel 2.2). Zusätzlich dazu gehören unter anderem auch der mediale Temporallappen, der Gyrus Fusiformis, der Gyrus Cinguli, der ventrale und dorsale präfrontale Kortex sowie der posteriore und inferiore Parietallappen zu wichtigen Hirnregionen für die Sprachgenerierung (Fu et al., 2006; Price 2010; Wende et al., 2012; Nagels et al., 2012, Schlosser et al. 1998). Die in dieser Arbeit erhaltenen Ergebnisse sind größtenteils kongruent mit diesen Aktivierungsbereichen. Dagegen fehlt sowohl in der ersten als auch in der zweiten Messung

eine Aktivierung der Broca Region. Dies könnte an der Wahl des Schwellenwertes liegen, wodurch bestimmte Aktivierungen erst ab einem Schwellenwert über $p(\text{fwe unkorrigiert})=0,05$ in die Auswertung übernommen wurden. Der Schwellenwert ist jedoch bereits durchaus liberal gewählt und eine noch geringere Anwendung des Schwellenwertes bei der Gruppenanalyse hätte ein wesentlich undifferenzierteres Aktivierungsbild zur Folge gehabt. Zudem fehlt ausschließlich der IFG in der Auswertung der aktivierten Hirnregionen. Sowohl der superiore als auch der mediale Gyrus Frontalis zeigten rechtsseitig Aktivierungen großer Clustergrößen und auch linksseitig war der superiore Gyrus Frontalis aktiviert. Der mediale Gyrus Frontalis wird für den Wortgenerierungsschritt des *Word Retrieval* als aktiviert erwartet (Price, 2010). Darüber hinaus wird ihm in vielen Studien über die semantische Sprachproduktion die Vermittlerfunktion (Jeon und Polich, 2003), die Kontrollfunktion (Whitney et al., 2009) und die Beteiligung an der semantischen Ordnung bei sogenannten ‚lead-in‘-Prozessen (Indefrey und Levelt, 2000) der Sprache zugewiesen.

Des Weiteren ist erwähnenswert, dass starke Aktivierungen in den Hirnregionen des Gyrus Cinguli und des medialen Gyrus Temporalis lagen. Diese sind erwartungsgemäß im Wortgenerierungsschritt des *Processing* aktiv, zu welchem auch die in dieser Studie aufgetretene Aktivierung des Gyrus Fusiformis typischerweise gehört. Die für den Schritt der *Articulation* erwarteten motorischen Aktivierungen traten nur bei der zweiten Messung und einem p-Wert von 0,01 auf. Vor allem die Koordinaten für den Bereich des Larynx und der Zunge von Price stimmten mit den Messungen dieser Arbeit nahezu überein (-40 -12 30 und 59 -5 17, Price, 2010).

Die starke Aktivierung des Gyrus Supramarginalis spricht für eine Aktivierung der Hörrinde, wobei diese Hirnregion typischerweise für das Sprachverständnis im Wernicke Areal bekannt ist. Auch der für den auditiven Kortex bekannte superiore Gyrus Temporalis zeigte einige Aktivierungen in den Auswertungen der zweiten Messung, was insgesamt mit dem Schritt des *auditiven Feedbacks* übereinstimmt (Price, 2010).

Zuletzt spricht auch die Aktivierung des rechten Temporalpols klassisch für die Sprachproduktion, er stellt eine zentrale Verbindungsstelle für die Sprache dar, da er bei semantischer Demenz geschädigt ist (Schneider und Fink, 2013.)

Insgesamt kann die Kongruenz der Aktivierungen des WG-Paradigmas mit den bisher gewonnenen fMRT-Forschungserkenntnissen über die für die semantische Wortgenerierung zuständigen Hirnregionen bestätigt werden.

In der Auswertung der Deaktivierungen fand sich jedoch unerwarteterweise ein großes frontoparietales Aktivierungsbild und es waren insbesondere deutlich mehr Hirnregionen als in der eigentlichen Aktivierungsphase aktiv. Das im Design antizipierte Hauptaktivierungsbild während der Wortgenerierungsphase stellte sich in der Auswertung nun als umgekehrter Fall mit Hauptaktivierung in der Ruhephase des Paradigmas dar. Dies ist ein erster entscheidender Hinweis für ein Designproblem des Paradigmas.

5.3.1.2 Bewertung der Reliabilitätsanalyse

In der Reliabilitätsanalyse waren deskriptiv mittels *Cluster*-Überlappungsmethode in beiden Messungen die gleichen Netzwerke in der Aktivierungsphase aktiv. Die zweite Messung zeigte darüber hinaus deutlich mehr Aktivierungen als die erste. Möglicherweise waren die Proband_innen durch die Wiederholung der Messung in vielerlei Hinsicht der eigentlichen Untersuchung der Wortgenerierung zugänglicher. Die Rahmenbedingungen mit dem ungewohnten Lärm eines Magnetresonanztomographen und dem engen Raum bei der Messung sowie die Aufregung und Unsicherheiten, ob die Aufgaben richtig verstanden wurden, spielen hier sicherlich eine nicht unwesentliche Rolle. Zudem ist zu beachten, dass die Wortgenerierungsaufgabe die erste Aufgabe für die Proband_innen war, bei welcher sie sich aktiv beteiligten. Es ist also davon auszugehen, dass bei den ersten Messungen mehr Confounder in die Aktivierungen eingewirkt haben und erst zum Ende der Messungen hin die jeweiligen Phasen des Paradigmas, die hauptsächlich die mit der Aufgabe assoziierten Aktivierungen hervorrufen konnten (Lohmann et al., 2004).

Auch in der *Cluster*-Überlappung der Deaktivierungsphase bestätigte sich der Verdacht auf ein Designproblem des Paradigmas, da nicht nur in der ersten Messung eine deutlich stärkere Aktivierung während der Kontrollphase im Gegensatz zur Aktivierungsphase aufgetreten war, sondern auch in der zweiten Messung. Zudem zeigten die Überlappungsbilder, dass überwiegend die gleichen Areale in beiden Messungen aktiviert waren.

Es folgt nun die Bewertung der Reliabilitätsanalyse anhand der berechneten ICC-Werte. Bei der ICC-vs.-T-Karte wird üblicherweise ein U-förmiger Verlauf um den T-Wert von 0 erwartet. Dies bedeutet, dass hohe Aktivierungen und hohe Deaktivierungen typischerweise auch mit größeren ICC-Werten und somit einer erhöhten Reliabilität einhergehen (Fliessbach et al., 2010, Caceres et al. 2009). In der Analyse des WG-Paradigmas ergab sich allerdings eine überraschende Graphik. Ein Großteil der Voxel ordnete sich im Bereich von negativen T-Werten an. Das ist für das Design des WG-Paradigmas ungewöhnlich, jedoch kongruent mit dem stärkeren Aktivierungsbild der Deaktivierungsphase in der 3D-Darstellung. Eine gute Kontrollaufgabe sollte bestenfalls eine möglichst ähnliche Aktivierung zeigen wie die Aktivierungsaufgabe und nur die Regionen, die interessieren, sollten in der Aktivierungsaufgabe mehr Aktivierungsstärke zeigen (Jansen et al., 2006). Im Falle der WG-Aufgabe dieser Studie ist allerdings das Gegenteil der Fall, was erneut Probleme im Design des Paradigmas bestätigt. Weiterhin liegen sehr viele Voxel im Bereich von negativen ICC-Werten, was auf eine insgesamt niedrige Reliabilität hinweist. Auch mit dem sehr niedrigen berechneten medianen ICC über das gesamte Gehirn von $ICC_{med} = -0,063$ findet sich ein weiterer Hinweis auf eine niedrige Reliabilität, denn typische Gesamthirnverteilungen haben in der Regel eine höhere Reliabilität (Caceres et al., 2009 und Shrout und Fleiss, 1979).

Zusammenfassend kann konstatiert werden, dass die errechneten Reliabilitäten der WG-Aufgabe insgesamt sehr niedrig sind und die aktivierten und deaktivierten Hirnregionen erstaunlicherweise keine höhere Reliabilität zeigen als Bereiche, die nicht aktiv sind. Die Reliabilität dieses Paradigmas ist demnach sehr schlecht.

5.3.1.3 Bewertung des Paradigmen Designs

In der Bewertung der Ergebnisse des WG-Paradigmas verfestigte sich der Verdacht, dass das Design nicht den gewünschten kognitiven Prozess der semantischen Wortgenerierung ausreichend reliabel aktivieren kann. Zunächst scheint dieses Ergebnis überraschend, da das Design einige Stärken aufweist. Bennet und Miller arbeiteten in einer Studie von 2013 mehrere Bedingungen für gute Paradigmen Designs heraus. Eine davon ist die Anwendung eines Block-Designs, welches viel häufiger als andere Designs mit einer guten Reliabilität einhergeht. Auch die Wahl des Kontrastes von Kondition vs. Ruhe im WG-Design erzeugt in

der Regel höhere Reliabilitäten als beispielsweise Kontraste von Kondition vs Kondition. Zuletzt ist auch der Abstand der beiden Messungen in einem günstigen Zeitraum von nur einem Monat gewählt worden, denn Bennet und Miller beobachteten abnehmende Reliabilitäten bei längeren Test-Retest Intervallen.

Das Design selbst ist mit der Bedingung, in einem Zeitraum von 12 Sekunden so viele Wörter einer Kategorie zu benennen, wie bekannt sind, ein häufig verwandtes Design in vielen Studien zur semantischen Wortflüssigkeit (Lezak, Howieson und Bigler, 2011). Zudem ist als Bedingung ein visuell minimal stimulierendes Bild gewählt worden, indem lediglich für kurze Zeit die Kategorie vor schwarzem Bildschirm eingeblendet wurde und von einem '+' Zeichen während des Wortgenerierungsprozesses gefolgt wurde.

Im Vergleich zu anderen Studien der semantischen Sprachproduktion fällt jedoch auf, dass überwiegend eine stille Wortgenerierung als Bedingung angewandt wurde. Somit wurden Bewegungsartefakte minimiert und erhöhte Voxelaktivierungen in den interessierenden Regionen erreicht (Yetkin et al., 1995, Friedman et al., 1998). In anderen Studien sollten die Proband_innen auch in den Kontrollphasen weiter sprechen, indem sie ein Wort wiederholten oder zählten, so dass der Wortgenerierungsprozess der zuvor gegebenen Kategorie beendet werden musste (Costafreda, Everitt, Fu, Lee, Brammer und David, 2006). In der Studie dieser Arbeit konnte durch die Ruhe in der Kontrollphase nicht sichergestellt werden, dass die Proband_innen nicht während der Kontrollphase weiter über den Kategorien zugehörige Wörter grübelten. Möglicherweise lösten Emotionen wie Ärger über nicht erwähnte Wörter noch zusätzliche Aktivierungen aus (Price, 2010).

Die Tatsache, dass das Paradigma an zweiter Stelle in der Abfolge der Paradigmen stand, sollte ebenfalls nicht vernachlässigt werden. Denn somit befanden sich die Proband_innen noch in einer potenziell sehr *Confounder*-anfälligen Phase der Messung. Zwar wurde bereits ein *Resting State* Paradigma bewusst vor das der Wortgenerierung gestellt, damit sich die Proband_innen bereits an die Situation der MRT-Messung gewöhnen konnten, trotzdem war es die erste Aufgabe, bei welcher sie mitarbeiten sollten.

5.3.2 SPES-Aufgabe

5.3.2.1 Bewertung der Gruppenanalyse

GESICHTER>KONTROLLE

Mit dem Kontrast GESICHTER>KONTROLLE sollte primär untersucht werden, ob das Paradigma grundsätzlich zur Messung der Gesichtererkennung und –verarbeitung geeignet ist und die Messung an sich reliable Ergebnisse liefern kann. Erste Studienergebnisse über die Erkennung von Gesichtern präsentierten bereits Kanwisher, McDermott und Chun 1997, welche zunächst die Rolle der FFA in der Gesichterverarbeitung erkannten. In vielfältigen darauffolgenden Studien wurden die Ergebnisse von Kanwisher et al. bestätigt, weiter differenziert und mit anderen Hirnregionen zur Gesichterverarbeitung ergänzt (McCarthy et al., 1997; Woiciulik et al. 1998; Haxby et al., 2000; Kanwisher und Yovel 2006; Gobbini und Haxby 2007; Ishai, 2008 u.v.m.).

Die Hauptaktivierungen des ersten ausgewerteten Kontrastes aller präsentierten Gesichter gegenüber den grauen Kontrolldias dieser Arbeit liegen in den drei Kernzentren des extrastriären visuellen Kortex (Schneider und Fink, 2013). Hierzu gehören die bilateralen aber rechts dominanten Aktivierungen des Gyrus Fusiformis und des Gyrus Temporalis inferior in dem ersten Kernzentrum, der sogenannten *Fusiform Face Area* (FFA), welche für die Repräsentation der Gesichtsidentität zuständig ist. Die bilateralen Aktivierungen des inferioren Gyrus Occipitalis gehören dem zweiten Kernzentrum, den visuellen Cortex an, welcher für die frühe Phase der Gesichtererkennung zuständig ist. Die bilateralen Aktivierungen des STG entsprechen dem superioren Sulcus Temporalis (STS), welcher als drittes Kernzentrum für die Verarbeitung veränderbarer Aspekte von Gesichtern bekannt ist (Haxby et al., 2000). In kleineren Clustern konnten zudem in einem Wahrscheinlichkeitsanteil von 16,2% die Amygdala und von 24% der Hippocampus sowie der IFG als aktiviert gemessen werden.

Die beschriebenen Hauptaktivierungen sind auch bei sehr restriktiven Schwellenwerten von $p(\text{unkorrigiert}) = 0,001$ vollständig kongruent mit den erwarteten Aktivierungslokalisationen für die Gesichtererkennung. Es kann somit insgesamt festgestellt werden, dass das Paradigma hinsichtlich der Gesichterverarbeitung valide ist. Zusätzlich fällt auf, dass auch

Aktivierungen in den für die Verarbeitung emotionaler Stimuli bekannten Hirnregionen der Amygdala, des Hippocampus und des IFG gemessen werden konnten. Dies ist höchstwahrscheinlich darauf zurückzuführen, dass im überwiegenden Anteil des Paradigmas die maskierten emotionalen Stimuli gleichzeitig mit der Gesichterpräsentation gezeigt wurden und somit stets parallel zur Gesichtsverarbeitung auch kognitive Prozesse der Emotionsverarbeitung stattfanden und gemessen wurden.

ANGST>NETURAL

Die Messung der subliminalen Präsentation emotionaler Stimuli sollte nun der Frage nachgehen, ob das Paradigma neben der Verarbeitung der Gesichtswahrnehmung auch zuverlässig unbewusst präsentierte Emotionen untersuchen kann. Hierzu wurden drei der sechs Basisemotionen- Angst, Freude und Trauer- zur Stimmungsinduktion verwendet. In der Analyse der Messdaten sind verschiedene Konstellationen von Kontrasten ausgewertet jedoch nicht in die Ergebnisdarstellung aufgenommen wurden. Da diese Arbeit sich der Frage der Reliabilität des Paradigmas widmet und nicht der unterschiedlichen Verarbeitung von Emotionen, wird repräsentativ nur die Analyse der Messdaten der unterbewussten Verarbeitung von Angst Stimuli ausführlich diskutiert werden. Zudem soll in diesem Abschnitt auch der besondere Aspekt der subliminalen Wahrnehmung und die damit einhergehenden kognitiven Verarbeitungsmodifikationen beleuchtet werden.

Die Hauptaktivierungen des ANGST>NEUTRAL Kontrastes der SPES-Aufgabe liegen in den folgenden Hirnregionen: bilateral im MTG und STG sowie dem Gyrus Fusiformis. Als bekannte subkortikale Hirnareale fanden sich bilaterale Aktivierungen in der Amygdala und dem Hippocampus. Des Weiteren waren motorische Regionen wie der prä- und postzentrale Gyrus sowie die SMA aktiviert.

In der zweiten Messung konnten deutlich stärkere Aktivierungen vor allem im Bereich der Amygdala und des Hippocampus erhoben werden, so dass in der Datenanalyse der restriktivere Schwellenwert von $p(\text{unkorrigiert})= 0,001$ im Vergleich zu $p(\text{unkorrigiert})= 0,01$ der ersten Messung verwendet wurde. Die höhere Aktivierungsstärke ist möglicherweise darauf zurückzuführen, dass die Proband_innen nach dem ersten Messvorgang an die Messbedingungen gewöhnt und weniger aufgeregt und ablenkbar waren. Da vor allem die Wahrnehmung der unbewussten Stimuli stark von der Aufmerksamkeit der Proband_innen

abhängt, kann dies sich in der zweiten Messung vorteilhaft auf die Aktivierungsstärke ausgewirkt haben.

In einer fMRT-Übersichtsarbeit über die kognitive Verarbeitung von Emotionen beschrieben Fusar-Poli et al. 2009 sowohl allgemeine Aktivierungen für alle Basisemotionen als auch differenziert auf die jeweiligen Qualitäten zuordnenbare Aktivierungszentren (siehe Kapitel 2.3). In der Messung dieser Arbeit konnten sowohl die allgemeinen Aktivierungen im Bereich des visuellen Kortex, temporoparietaler und präfrontaler Regionen als auch der subkortikalen Strukturen reproduziert werden.

Die subkortikalen Aktivierungen des Kontrastes ANGST>NEUTRAL im Bereich der Amygdala und des Hippocampus sind kongruent mit den bisherigen Forschungsergebnissen (Yang-Menon et al., 2002; Killgore und Yurgelun-Todd, 2004; Costafreda et al. 2009; Fusar-Poli et al. 2009; Dima et al., 2011). Auch die in dieser Arbeit gezeigte Aktivierung des IFG stimmt mit aktuellen Forschungsergebnissen überein (Suslow und Dannlowski, 2006; Fusar-Poli et al., 2009).

Eine besondere Rolle nehmen die hohen Aktivierungswerte im Bereich des STG und MTG ein, da sie weniger für die Emotionsverarbeitung als vielmehr für die Einspeicherung von Gedächtnisinhalten bekannt sind (Jansen et al. 2006, Leube et al. 2001, Tulving et al. 1994). Zwischen dem Gedächtnis und Emotionen besteht jedoch ein wichtiger Zusammenhang, denn wie in Studien belegt werden konnte, ist die Intensität der Emotionen während des Einspeicherns von Inhalten ausschlaggebend für die Gedächtnisleistung (Schneider und Fink, 2013). Auch während des Abrufens emotionaler Inhalte ist eine hohe Korrelation mit der Amygdalaaktivität feststellbar (Jackson et al. 2008). Insofern wird von einem, MTL-Gedächtnissystem' gesprochen, welches mit dem Hippocampus, dem entorhinalen Kortex, dem perirhinalen Kortex sowie der Amygdala die Region des medialen Temporallappens betrifft (Murty, Ritchey, Adcock und LaBar, 2010). Die starken Aktivierungen dieser Arbeit im Bereich des MTG sprechen demnach dafür, dass die Proband_innen die Gesichter nicht nur betrachteten, sondern gleichzeitig versuchten, das Gesehene einzuspeichern. Dieses Verhalten kann in Zusammenhang mit den beiden unmittelbar vor der SPES-Aufgabe gezeigten Paradigmen des Einspeichern (*Faces Encoding*) und Abrufen (*Faces Decoding*) von Gesichtern stehen.

Von den subliminal gezeigten Emotionen erzeugte Angst die verhältnismäßig stärksten Aktivierungen. Dieser Aspekt ist insbesondere deshalb interessant, als dass Angst eine Emotion von Überlebenswert ist. Im Vergleich hierzu löste beispielsweise die Emotion Trauer im Kontrast TRAUER>NEUTRAL kaum Aktivierungen aus, was im Kontext des Überlebenswertes nachvollziehbar ist. Trauer verlangsamt das Verhalten und Gedankenprozesse und aktiviert nur minimal Bereiche wie den anterioren Gyrus Cinguli und die Insula (Mayberg et al. 2000; Killgore und Yurgelun-Todd, 2004). Da keine direkte Überlebensreaktion ausgelöst werden muss, erfolgt die Verarbeitung hier über den Thalamus und dauert länger.

Angst hingegen kann als biologisch relevanter Stimulus gewertet werden, welcher primär die sensorischen Kortices aktiviert und über eine grobe Bewertung der Signifikanz der Stimuli in den Amygdala schnelle körperliche Antworten auslöst (LeDoux, 1998).

Trotz der in den erwarteten Hirnregionen lokalisierten Aktivierungen ist eine nur sehr schwache Aktivität im Bereich der Amygdala und des Hippocampus zu verzeichnen, welche in der ersten Messung auch nur in geringen Wahrscheinlichkeitsanteilen tatsächlich diesen Hirnregionen zugewiesen werden konnte (linke Amygdala 3,4% und linker Hippocampus 23,5%). Die Reproduktion der Aktivierungen im Bereich der Amygdala stellte bisher die größte Herausforderung in der Emotionserforschung dar (Lipp et al., 2014; Plichta et al., 2012). Eine Ursache hierfür kann an hohen Magnetfeldinhomogenitäten im Bereich der Amygdala Region liegen. Hierdurch können wesentliche Aktivitäten bei geringen Signal zu Rausch- Verhältnissen nicht gemessen werden. Aus diesem Grunde führten Boubela et al. eine Studie mit beschleunigten Magnetresonanz Sequenzen und hoher zeitlicher Bildauflösung bei einer TR= 333ms durch (Boubela et al., 2015). Hierdurch konnte das Signal zu Rausch- Verhältnis deutlich verbessert werden. Es stellte sich in den Messungen von Boubela et al. jedoch heraus, dass die Aktivierungen keinesfalls in den Amygdala, sondern in der daneben liegenden Rosenthalvene zu vermessen waren.

Die geringe Aktivierung kann auch auf die maskierte Präsentation der Stimuli zurückzuführen sein. Während manche Autoren beispielsweise Aktivitäten in der Amygdala und dem anterioren Cingulären Gyrus als charakteristisch für die unbewusste, subkortikale Verarbeitung beschreiben (Killgore und Yurgelun-Todd, 2004), stellen andere Autoren neben der Amygdala Aktivierungen im inferioren Gyrus Frontalis als typische Hirnregion bei

subliminal präsentierter Angst vor (Suslow et al., 2003 und Fusar-Poli et al., 2009). Wieder andere konnten eine Amygdalabeteiligung nur bei offener- nicht bei maskierter- Präsentation von Angst Stimuli messen (Costafreda et al. 2008; Phillips et al., 2004).

Die Verarbeitung von Emotionen scheint demnach zwar an die Beteiligung der Amygdala gebunden zu sein, jedoch bleiben die konkreten Verarbeitungswege für bewusst und unterbewusst präsentierte Stimuli an unterschiedliche Netzwerke in subkortikalen und kortikalen Strukturen gebunden. Darüber hinaus müssen auch die verschiedenen Verarbeitungswege der einzelnen Basisemotionen beachtet werden.

5.3.2.2 Bewertung der Reliabilitätsanalyse

Mit der *Cluster*-Überlappungsdarstellung werden mit großflächigen gelben Arealen deutliche Überlappungen innerhalb des GESICHTER>KONTROLLE Kontrastes sichtbar. Die Stärke der Reliabilität dieser Überlappungen bestätigt sich allerdings nur bedingt in den ICC-vs.T-Karten. Der erwartete U-förmige Verlauf wird sichtbar, was bedeutet, dass hohe (De-)aktivierungswerte mit hohen ICC-Werten korrelieren. Allerdings liegen die gemessenen Aktivierungswerte überwiegend im negativen Bereich um den Wert von $T = -7$ und nicht um den T-Wert von 0. Dementsprechend zeigen nur die hohen Deaktivierungswerte starke ICC-Werte an, während hohe Aktivierungswerte nicht gemessen werden konnten. Prinzipiell sollten sich die Deaktivierungs- und Aktivierungswerte gleichen, da eine gute Kontrollaufgabe sich lediglich in der Aktivierung der interessierenden Areale unterscheidet. Zur genaueren Berechnung der aktivierten Bereiche ergibt die Analyse mit der relativen Häufigkeitsverteilung einen medianen ICC der aktivierten Netzwerke von 0,47, was einer moderaten Reliabilität entspricht. Auch die Berechnung des medianen ICC des gesamten Gehirns liegt verhältnismäßig hoch bei $ICC_{med} = 0,28$. Hier zeigt sich nun, dass in den aktivierten Hirnregionen stärkere Reliabilitäten vorliegen, bestätigt durch die in der Reliabilitätsanalyse der Referenzregion FFA mit einem ICC_{med} von 0,62 entsprechend guten Reliabilität. Die Gesichterverarbeitung ist auch in vergleichbaren Studien mit hohen Reliabilitäten messbar gewesen (Kanwisher et al., 1997; Haxby et al., 2000). Mit der Reliabilitätsanalyse des GESICHTER>KONTROLLE Kontrastes lässt sich daher für diese Arbeit zeigen, dass sowohl die Messung als auch das Paradigma gute Reliabilitäten erzeugen können.

Auf diese Feststellung folgend werden die Ergebnisse des ANGST>NEUTRAL Kontrastes zusammenfassend beleuchtet. Wie sich bereits in der Gruppenanalyse angedeutet hat, sind die Aktivierungsareale nicht eindeutig anatomisch lokalisierbar, was sich auch in der Reliabilitätsanalyse widerspiegelt. Zunächst überlappen sich die aktivierten Hirnregionen- wie in der *Cluster*-Überlappungsmethode gut sichtbar- nur in geringem Ausmaße. Anhand der ICC-vs.-T-Karte lässt sich weiterhin feststellen, dass nicht wie typischerweise erwartet die hohen Aktivierungs- bzw. Deaktivierungswerte mit hohen ICC-Werten assoziiert sind. Zwar liegen die Voxel um den T-Wert von 0 beinahe U-förmig verteilt, so dass durchaus ein Anstieg der ICC-Werte bei höheren (De-)aktivierungswerten zu vermerken ist, allerdings liegen die stärksten ICC-Werte um den Bereich von T= 0 bis 1 und sind selbst dann in ICC-Bereichen von etwa 0,6. Neben der fraglichen Lokalisation der aktivierten Hirnregionen weisen diese demnach eine nur schlechte Reproduzierbarkeit innerhalb der beiden Messungen auf. Dies bestätigt sich mit den medianen ICC-Werten der relativen Häufigkeitsverteilung mit $ICC_{med}=-0,004$ für die gesamte Gehirnaktivität und $ICC_{med}= 0,11$ für die aktivierten Netzwerke.

Das vor der Messung erwartete Hauptaktivierungszentrum der Amygdala bei der subliminalen Präsentation von Angst zeigt schließlich in der Reliabilitätsanalyse der Referenzregionen äußerst schlechte mediane ICC-Werte von -0,25 links und -0,03 rechtsseitig. Die geringen Werte sind maßgeblich darauf zurückzuführen, dass nicht wie sonst üblich die stärksten aktivierten Regionen als Referenzmaske ausgewählt wurden, sondern die durch andere Studien am häufigsten beschriebenen Regionen.

In der am stärksten aktivierten Region der Messungen dieser Arbeit, dem rechten medialen Temporalgyrus, zeigt sich demhingegen ein moderater ICC_{med} von 0,41.

Zusammenfassend deutet die Reliabilitätsanalyse der SPES-Aufgabe dieser Arbeit darauf hin, dass mit diesem Paradigma und dieser Messung durchaus gut reproduzierbare Ergebnisse erzielt werden können. Die Aktivierungsareale müssen jedoch entweder bei der Verarbeitung von subliminal präsentierten Emotionen in anderen Hirnregionen als bisher angenommen liegen oder die Messungen waren durch Designdefizite des Paradigmas an sich beeinträchtigt. Letztere sollen im folgenden Abschnitt diskutiert werden.

5.3.2.3 Bewertung des Paradigmen-Designs

Das besondere Design der SPES-Aufgabe wurde bewusst ausgewählt, um die komplexen Vorgänge der Emotionsverarbeitung möglichst gut untersuchen zu können. Zunächst soll daher das Blockdesign der Aufgabe diskutiert werden. Wie schon bereits bei dem Paradigma der Wortgenerierung ausgeführt, haben zahlreiche Reliabilitätsstudien nachweisen können, dass sich das Blockdesign positiv auf die Reliabilität auswirkt (Bennet und Miller, 2013). Dadurch wurde die Auswahl an Stimmungsinduktionsmethoden jedoch erheblich eingeschränkt. Event-related Designs hätten alternativ hierzu wirkungsstärkere Stimuli präsentieren und möglicherweise auch individuell und spontan ausgelöste Affekte besser untersuchen können. Die Präsentation von Gesichtern hat jedoch bisher die signifikantesten Ergebnisse geliefert und bereits gute Möglichkeiten der Reproduzierbarkeit dargeboten (Van den Bussche, Van den Noortgate und Reynvoet, 2009; Costafreda et al., 2008), weshalb sie in dieser Studie zum Einsatz kam. Es bleibt dennoch fraglich, wie neutral die maskierenden Gesichter tatsächlich auf das jeweilige Individuum wirken (Johnstone et al., 2005) und ob die grauen Dias als Kontrollstimuli geeignet sind. Die grauen Dias wurden gewählt, um einen Abbau der BOLD-Antworten zu gewähren, damit die Reaktion auf die folgende Emotionsdarbietung so wenig wie möglich durch die zuvor gezeigten Emotionen beeinflusst werden kann. Möglicherweise reagierten die Proband_innen dessen ungeachtet mehr auf das Wiedererkennen von Gesichtern in den Aktivierungsphasen nach Präsentation der grauen Dias in der Kontrollphase. Dies könnte auch die gut reliabel messbaren Aktivitäten in der Gesichtsverarbeitung im Gegensatz zu den schlecht bis moderat reliabel messbaren Aktivitäten in der Verarbeitung der Angststimuli erklären. Als Lösungsansatz zur zusätzlichen Verifizierung der unterbewussten Bewertung verwendeten Suslow und Dannlowski Maskierungsstimuli in Form von Chinesischen Ideogrammen, welche die Proband_innen mittels einer Tastatur unmittelbar bewerten sollten (Suslow und Dannlowski, 2006). Diese Methode hat die Vorteile, dass sie zum einen das Problem der fraglichen Neutralität der Gesichter umgehen kann und gleichzeitig die individuellen Bewertungen der maskierten Stimuli feststellt. Nicht zuletzt kann diese von persönlichen Assoziationen mit den dargestellten Gesichtern und auch der Tagesform der einzelnen Proband_innen beeinflusst werden.

Eine weitere Besonderheit in dem Paradigmen Design der SPES-Aufgabe liegt in der Programmierung der Gesichterdarstellung. Diese erfolgte stets durch einen Trigger nach einer festgelegten Anzahl von Bildaufnahmen. Dies kann allerdings mögliche zeitliche Verschiebungen in der Präsentation der maskierten Stimuli ausgelöst haben, so dass diese später nicht mehr exakt 33ms lang waren. Optimal sind zur Rückwärtsmaskierung Präsentationszeiten der maskierten Stimuli von 31- 38ms (Murphey und Zajonc, 1993). Der Vorteil dieser Programmierung lag darin, dass die schnellen zeitlichen Abfolgen der maskierten und offenen Stimuli so besser ablaufen und in der Reihenfolge der Gesichter randomisiert werden konnten.

Die Methode des *Backward Masking* soll als ein weiterer wichtiger Bestandteil der SPES-Aufgabe diskutiert werden. Zunächst muss dabei beachtet werden, dass unbewusst wahrgenommene Stimuli andere, vornehmlich subkortikale Netzwerke aktivieren und somit viele der in der Literatur angegebenen typischen Aktivierungsareale- wie in der Bewertung der Gruppenanalyse ausführlich beschrieben- hier nicht zu erwarten sind. Die unbewusste Verarbeitung von Stimuli weist dennoch eine hohe Abhängigkeit von der Aufmerksamkeit der Proband_innen auf (Wojciulik, Kanwisher und Driver, 1998). Sollten diese beispielsweise in dem nur kurzen Moment der 33ms andauernden Präsentation maskierter Stimuli nicht ihre Aufmerksamkeit auf die Gesichter richten, hätte dies zur Folge, dass in dem unbewussten Verarbeitungsprozess die Amygdala Aktivierung unterdrückt wird (Brassen et al. 2010, Morawetz et al. 2011; Williams et al. 2005).

Wie bereits in dem Diskussionsunterpunkt der Methoden dargelegt, kann es als Defizit dieser Studie betrachtet werden, dass die Aufmerksamkeit der Proband_innen während des Paradigmas nicht sichergestellt wurde. Beispielsweise hätte ein Fixationskreuz im Zentrum des Bildschirms 500ms vor der Präsentation des subliminalen Stimulus die Aufmerksamkeit erhöhen können (Suslow und Dannlowski, 2006) oder mittels ‚eyetracking‘ die Aufmerksamkeit nachgewiesen werden können. Erschwerend kommt an dieser Stelle die Länge der einzelnen Blöcke hinzu. Ausreichend starke BOLD-Signale hätten auch in kürzerer Zeit als 30 Sekunden erzeugt werden können, was die Aufmerksamkeitskapazitäten der Proband_innen zusätzlich verbessert hätte. Übersteigen Wahrnehmungsaufgaben beispielsweise aufgrund der Länge die Aufmerksamkeitskapazität der Proband_innen, so

werden als irrelevant bewertete Reize sehr früh herausgefiltert und nicht weiter verarbeitet (Lavie, 1995).

Als letzter Punkt sei an dieser Stelle erwähnt, dass die Proband_innen zwar vor Aufnahme in die Studiengruppe hinsichtlich psychiatrischer Erkrankungen befragt wurden, allerdings dies nicht anhand objektiver Methoden bestätigt wurde. Es ist daher nicht vollständig auszuschließen, dass die Proband_innen sich selbst fälschlicherweise als psychisch gesund eingeschätzt und beschrieben haben. Suslow und Dannlowski evaluierten beispielsweise vor Studienbeginn anhand des ‚*Becks Depressions Inventory*‘ depressive Symptome, da diese mit generell reduzierten Aktivierungsstärken bei emotionalen Stimuli einhergehen können (Mayberg et al. 2000).

Es bleibt letztendlich unklar, ob die schlechten Reliabilitätsergebnisse für die subliminale Verarbeitung der Angstreize in Defiziten des Paradigmen Designs begründet liegen oder insgesamt durch die aktuell sehr begrenzten Messoptionen erklärt werden können. Allerdings wäre es für zukünftige Auswertungen von hoher Relevanz, neben der Verbesserung von Design Defiziten grundsätzlich die Möglichkeit eines anderen neuronalen Netzwerkes zur Verarbeitung von unbewusst präsentierten Emotionen in Betracht zu ziehen.

6 FAZIT

Im Hinblick auf die Verwendbarkeit der Paradigmen für das Gesamtprojekt ‚*Erfassung von Wirkungen genetischer Risikovarianten für psychische Störungen auf Gehirnstruktur und –funktion bei Gesunden und Erkrankten*‘ sind einige Verbesserungen im experimentellen Design sowie den experimentellen und technischen Rahmenbedingungen beider Paradigmen notwendig.

Die Berechnung der Reliabilität für ein sehr homogenes Probandenkollektiv stellte sich als fehleranfällig heraus. Starke Schwankungen zwischen den einzelnen Proband_innen resultierten zudem aus dem Anspannungs- und Geräuschpegel, der Aufmerksamkeit bei langanhaltender kognitiver Beanspruchung und der Abhängigkeit von physiologischen Einflüssen.

Die beiden untersuchten Paradigmen zeigten insgesamt eine schlechte bis moderate Reliabilität:

Bei der WG-Aufgabe ergaben sich in der Auswertung Probleme mit dem Design, welche beispielsweise durch mehr Aktivierungen in der Kontrollphase deutlich wurden. Erwartete Hirnregionen wie die Broca-Region waren zum Teil nicht aktiviert. Aktivierungsunterschiede bei psychisch Erkrankten, welche mit diesem Paradigma gemessen werden, beruhen bei diesen Reliabilitätsergebnissen höchstwahrscheinlich auf Zufälligkeiten und lassen daher keinerlei Rückschlüsse auf den Verlauf einer psychischen Erkrankung zu.

Die SPES-Aufgabe lieferte bestenfalls moderate Reliabilitäten bei den Aktivierungsnetzwerken der Emotion Angst. Allerdings ließen sich diese nur für bestimmte ausgewählte Referenzregionen wie den rechten MTG feststellen, nicht jedoch für die Amygdala. Auch hier können demnach aus Aktivierungsunterschieden bei psychisch Erkrankten wie in der Rahmenstudie keine zuverlässigen Folgerungen aus gemessenen Aktivierungsänderungen im Krankheitsprozess gezogen werden.

Die Tatsache, dass die in dieser Arbeit untersuchten Paradigmen Teil einer Messung von sechs verschiedenen Paradigmen waren, ist ebenfalls hervorzuheben. Drei weitere Paradigmen, welche sich mit der Gedächtnisleistung befassten und nach bzw. vor dem WG-

und dem SPES-Paradigma gezeigt wurden, ergaben ebenfalls nur moderate bis mäßig moderate Reliabilitäten (Hermann, 2013). Möglicherweise beeinflussten sich die Paradigmen gegenseitig, so dass sie die Aktivitäten untereinander verstärkten oder hemmten (Bennet und Miller, 2013). Eine kognitive Erschöpfung nach dem Einspeichern und Abrufen des episodischen Gedächtnisses könnte beispielsweise eine Aktivitätsminderung in der SPES-Aufgabe bedingt haben.

In dieser Studie wurde bei allen Paradigmen das Block Design angewandt. Unterschiedliche Paradigmen-Designs innerhalb einer Testreihe könnten zusätzlich die Aktivitätsdynamik beeinflussen. Eine genaue Vergleichsanalyse der kognitiven Dynamik bei der Messung der unterschiedlichen kognitiven Funktionen und unterschiedlichen Paradigmen-Designs könnte daher ein Impuls für zukünftige Studien sein.

Der ICC-Wert ist trotz vieler Vorteile defizitär (Caceres et al. 2009, Bennet und Miller 2010, Fliessbach et al., 2010). Dadurch, dass er aktuell als häufigster Marker in der Reliabilitätsanalyse verwendet wird, hat er voraussichtlich weiterhin eine große Relevanz. Zur Verbesserung könnten standardisierte Reliabilitätsmessungen mit Test-Retest Verfahren beitragen, insbesondere bei einem so rauschanfälligen Messverfahren wie der fMRT. In der Auswertung sollte stets bedacht werden, dass der ICC-Wert aufgrund seiner Abhängigkeit von der interindividuellen Varianz insbesondere in der Übertragung der Ergebnisse auf andere Kollektive in der Aussagekraft massiv eingeschränkt ist. Für die Übertragbarkeit der Ergebnisse sollte daher ein möglichst heterogenes, einem Patientenkollektiv angenähertes Probandenkollektiv ausgewählt werden.

Wünschenswert wäre als Standard der Reliabilitätsmessung ein von der interindividuellen Varianz unabhängiger Marker, um dem wissenschaftlichen Wert innerhalb der vielversprechenden fMRT-Forschung die nötige zuverlässige Grundlage zu bieten.

6 LITERATURVERZEICHNIS

BARRETT LF, BLISS-MOREAU E, DUNCAN SL, RAUCH SL, WRIGHT CI. *The amygdala and the experience of affect.* Soc Cogn Affect Neurosci. 2007 **Jun;2(2):**73-83.

BENNETT CM & MILLER MB (2010). *How reliable are the results from functional magnetic resonance imaging?* Ann N Y Acad Sci, **1191**, 133-155.

BENNETT CM & MILLER MB (2013). *fMRI reliability: Influences of task and experimental design.* Cogn Affect Behav Neurosci, **13(4)**, 690-702.

Benton AL (1968) *Differential behavioral effects in frontal lobe disease.* Neuropsychologia, **6**, 53-60.

Bland JM, & Alman DG (1986). *Statistical methods for assessing agreement between two methods of clinical measurement.* Lancet, **1(8476)**, 307-310.

BLOCH F, HANSEN W & PACKARD M (1946). *Nuclear Induction.* Phys Rev, **69**: 127.

BOUBELA RN, KALCHER K, HUF W, SEIDEL EM, DERNTL B, PEZAWAS L, NASEL C, MOSERA E *fMRI measurements of amygdala activation are confounded by stimulus correlated signal fluctuation in nearby veins draining distant brain regions.* Sci Rep. 2015; **5**: 10499.

BOUMA A, MULDER J, LINDEBOOM J (1998) *Neuropsychologische Diagnostik.* Handboek Swets & Zeitlinger.

BRANDT DJ, SOMMER J, KRACH S, BEDENBENDER J, KIRCHER T, PAULUS FM, & JANSEN A (2013). *Test-Retest Reliability of fMRI Brain Activity during Memory Encoding.* Front Psychiatry, **4**, 163.

BRASSEN S, GAMER M, ROSE M, BÜCHEL C (2010). *The influence of directed covert attention on emotional face processing.* Neuroimage. **50(2):** 545-51. doi: 10.1016/j.neuroimage.2009.12.073. Epub 2009 Dec 24.

CACERES A, HALL DL, ZELAYA FO, WILLIAMS C, & MEHTA MA (2009). *Measuring fMRI reliability with the intra-class correlation coefficient.* Neuroimage, **45(3)**, 758-768.

CARRIER J, & MONK TH (2000). *Circadian rhythms of performance: new trends.* Chronobiol Int, **17(6)**, 719-732.

CASPAR F, WIRTZ M: *Beurteilerübereinstimmung und Beurteilerreliabilität,* Hogrefe Verlag, Göttingen 2002

CECCHI GA, GARG R, RAO AR (2009). *A cluster overlap measure for comparison of activations in fMRI studies.* Med Image Comput Comput Assist Interv. 2009; **12(Pt 1)**:1018-25.

COSTAFREDA SG, FU CHY, LEE L, EVERITT B, BRAMMER M, DAVID AS (2006): *A systematic review and quantitative appraisal of MRI studies of verbal fluency: Role of the left inferior frontal gyrus.* Human Brain Mapping, **Vol. 27, Issue 10**: 799-810.

COSTAFREDA SG, BRAMMER MJ, DAVID AS & FU CHY (2008). *Predictors of amygdala activation during the processing of emotional stimuli: a meta-analysis of 385 PET and fMRI studies.* Brain Res Rev **58**, 57–70

DANNLOWSKI U, SUSLOW T (2006). *Test-Retest Reliability of subliminal facial affective priming.* Psychological Reports 2006, **98**, 153-158

DEICHMANN R, GOTTFRIED JA, HUTTON C, TURNER R 2003. *Optimized EPI for fMRI studies of the orbitofrontal cortex.* NeuroImage **19**, 430–441.

DIMA D, KLAAS ES, ROISER JP (2011) *Effective Connectivity during Processing of Facial Affect: Evidence for Multiple Parallel Pathways.* J Neurosci **31**: 14378–14385

FERNANDEZ G, SPECHT K, WEIS S, TENDOLKAR I, REUBER M, FELL J, KLAVER P, RUHLMANN J, REUL J, ELGER CE (2003). *Intrasubject reproducibility of presurgical language lateralization and mapping using fMRI.* Neurology **60**: 969–975.

FLIESSBACH K, ROHE T, LINDER NS, TRAUTNER P, ELGER CE, & WEBER B (2010). *Retest reliability of reward-related BOLD signals.* Neuroimage, **50(3)**, 1168-1176.

FOKI T (2010). *Methodische Probleme klinischer funktioneller MRT-Untersuchungen.* In R. Beisteiner (Ed.), (pp. 104-109): Radiologe, Springer-Verlag.

FREUD S (1915). *The unconscious.* In: Strachey, J. (Ed.), The Standard Edition, **vol. 14**. Hogarth Press, London.

FRIEDMAN L, KENNY JT, WISE AL, WU D, STUVE TA, MILLER DA, JESBERGER JA, LEWIN JS: *Brain activation during silent word generation evaluated with functional MRI.* Brain Lang **1998**; **64**:231–256 CrossRef, Medline

FUSAR-POLI P, PLACENTINO A, CARLETTI F, LANDI P, ALLEN P, SURGULADZE S, BENEDETTI F, ABBAMONTE M, GASPAROTTI R, BARALE F, PEREZ J, MC GUIRE P, POLITI P *Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies.* J Psychiatry Neurosci. 2009 Nov; **34(6)**:418-32.

GAZZANIGA MS, IVRY RB, MANGUN GR (2015) Cognitive Neuroscience: The Biology of the Mind **4e**, pp. 425-467

GOBBINI MI, HAXBY JV (2007) *Neural systems for recognition of familiar faces. Neuropsychologia* **45**: 32–41

HAXBY JV, HOFFMANN EA, GOBBINI MI: *The distributed human neural system for face perception*, Elsevier Science Ltd. PII: Trends in Cognitive Sciences – **Vol. 4, No. 6, June 2000**

HEINRICHS RW & ZAKZANIS KK (1998) *Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. Neuropsychology*, **12**, 426-445.

HERMANN B: *Test-Retest-Reliabilität von fMRT-Paradigmen: Episodisches Gedächtnis und Arbeitsgedächtnis*, Marburg 2014

INDEFREY P & LEVELR W (2000). *The neural correlates of language production. M. Gazzaniga (Ed.), The new cognitive neurosciences. Cambridge, MA:MIT Press.*

ISHAI A (2008) *Let's face it: It's a cortical network. NeuroImage* **40**: 415–419

JACKSON MC, WOLF C, JOHNSTON SJ, RAYMOND JE, LINDEN DE (2008) *Neural correlates of enhanced visual short-term memory for angry faces: an FMRI study. PLoS One* **3**: e3536

JANSEN A (2004). *Atypische Lateralisation kognitiver Hirnfunktionen*

JANSEN A, MENKE R, SOMMER J, FÖRSTER AF, BRUCHMANN S, HEMPLEMAN J, KNECHT S (2006). *The assessment of hemispheric lateralization in functional MRI--robustness and reproducibility. Neuroimage*, **33(1)**, 204-217.

JEON YW, POLICH J. (2003). *Meta-analysis of P300 and schizophrenia: patients, paradigms, and practical implication. Psychophysiology*; **40 (5)**:684-701

JOHNSTONE T, SOMERVILLE LH, ALEXANDER AL, OAKES TR, DAVIDSON RJ, KALIN NH, WHALEN PJ (2005). *Stability of amygdala BOLD response to fearful faces over multiple scan sessions. Neuroimage*. **25(4)**: 1112-23.

KANDEL ER, SCHWARTZ JH, JESSEL TM, SIEGELBAUM SA, HUDSPETH AJ (2013) Principles of Neural Science, **Fifth Edition**, pp 1079-1089

KANWISHER N, MC DERMOTT J & CHUN, MM (1997). *The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci*, **17(11)**, 4302-4311.

KANWISHER N, YOVEL G (2006) *The fusiform face area: a cortical region specialized for the perception of faces. Phil Trans R Soc B* **361**: 2109–2128

KILLGORE WDS, YURGELUN-TODD DA (2004) *Activation of the amygdala and anterior cingulate during nonconscious processing of sad versus happy faces, NeuroImage* **21**, pp 1215- 1223

KONRAD C, ENGELIEN A, SCHÖNING S, ZWITSERLOOD P, JANSEN A, PLETZIGER E, KUGEL H (2008). *The functional anatomy of semantic retrieval is influenced by gender, menstrual cycle, and sex hormones.* J Neural Transm, **115(9)**, 1327-1337.

LAVIE N (1995) *Perceptual load as a necessary condition for selective attention.* J Exp Psychol Hum Percept Perform **21**: 451–468

LEDOUX JE, 1998. *Cognitive-emotional interactions in the brain.* Cognition and Emotion **3 (4)**, 267–289.

LEZAK D, HOWIESON D, BIGLER E (2012). *Neuropsychological Assessment.* Oxford University Press.

LIENERT GA & U RAATZ: *Testaufbau und Testanalyse, Weinheim 1998,* Psychologie Verlags Union

LIPP I, MURPHEY K, WISE RG & CASERAS X (2014). *Understanding the contribution of neural and physiological signal variation to the low repeatability of emotion-induced bold responses.* Neuroimage **86**, 335–342

LOHMANN H, DEPPE M, JANSEN A, SCHWINDT W, KNECHT S (2004). *Task repetition can affect functional magnetic resonance imaging-based measures of language lateralization and lead to pseudoincreases in bilaterality.* J Cereb Blood Flow Metab. 2004 Feb; **24(2)**:179-87.

MALDIJAN JA, LAURIENTI PJ, KRAFT RA, & BURDETTE JH (2003). *An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets.* Neuroimage, **19(3)**, 1233-1239.

MAYBERG HS, LIOTTI M, BRANNAN SK, MCGINNIS S, MAHURIN RK, JERABEK PA, SILVA JA, TEKELL JL, MARTIN CC, LANCASTER JL, FOX PT (1999). *Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness.* Am J Psychiatry. 1999 May; **156(5)**:675-82.

MAYBERG HS, BRANNAN SK, TEKELL JL et al. (2000) *Regional metabolic effects of fluoxetine in major depression: serial changes and relationship to clinical response.* Biol Psychiatry **48**: 830–843

MCCARTHY G1, PUCE A, GORE JC, ALLISON T (1997) *Face-specific processing in the human fusiform gyrus.* J Cogn Neurosci ; **9(5)**: 605-10. doi: 10.1162/jocn.1997.9.5.605.

METTE M(2014). *Der Einfluss des Einzelnukleotidpolymorphismus rs1006737 des Gen CACNA1C auf neuronale Korrelate des Arbeitsgedächtnis. Eine Studie mit funktioneller Magnetresonanztomographie*

MORAWETZ C, BAUDEWIG J, TREUE S, DECHENT P (2011) *Effects of spatial frequency and location of fearful faces on human amygdala activity.* Brain Res **31**: 87–99

MURPHEY ST & ZAJONC RB (1993). *Affect, cognition and awareness: affective priming with optimal and suboptimal stimulus exposures.* Journal of Personality and Social Psychology **64**, 723-739.

MURTY VP, RITCHEY M, ADCOCK RA, LABAR KS (2010) *fMRI studies of successful emotional memory encoding: A quantitative metaanalysis.* Neuropsychologia **48**: 3459–3469

NAGELS A, KIRCHER T, DIETSCH B, BACKES H, MARQUETAND J, KRUG A (2012) *Neural processing of overt word generation in healthy individuals: The effect of age and word knowledge.* Neuroimage. **16; 61 (4)**: 832-840.

PHILIPPS ML, WILLIAMS LM, HEINING M, HERBA CM, RUSSEL T, ANDREW C, BULLMORE ET, BRAMMER MJ, WILLIAMS SCR, MORGAN M, YOUNG AW (2004). *Differential neural responses to overt and covert presentations of facial expressions of fear and disgust,* NeuroImage **21**, 1484–1496

PLICHTA MM, SCHWARZ AJ, GRIMM O, MORGEN K, MIER D, HADDAD L, GERDES AB, SAUER C, TOST H, ESSLINGER C, COLMAN P, WILSON F, KIRSCH P, MEYER-LINDENBERG A. (2012). *Test-retest reliability of evoked bold signals from a cognitive-emotive fMRI test battery.* Neuroimage **60**, 1746–1758

PRICE, CJ (2010): *The anatomy of language: a review of 100 fMRI studies published in 2009.* Annals of the New York Academy of Sciences, **Issue: The Year in Cognitive Neuroscience**

PURCELL, EM, H TORREY & R POUND (1946). *Resonance Absorption by Nuclear Magnetic Moments in a Solid.* Phys. Rev., **69**:37 38.

RAUCH AV, REKER M, OHRMANN P, PEDERSON A, BAUER J, DANNLOWSKI U, HARDING L, KOELKEBECK K, KONRAD C, KUGEL H, AROLT V, HEINDEL W, SUSLOW T (2010). *Increased amygdala activation during automatic processing of facial emotion in schizophrenia.* Psychiatry Res. 2010 Jun 30; **182(3)**:200-6. doi: 10.1016/j.psychres.2010.03.005. Epub 2010 May 20.

ROMBOUTS SA, BARKHOF F, HOOGENRAAD FG, SPRENGER M, VALK J & SCHELTENS P (1997). *Test-retest analysis with functional MR of the activated area in the human visual cortex.* AJNR Am J Neuroradiol, **18(7)**, 1317-1322.

SCHLOSSER R, HUTCHINSON M, JOSEFFER S, RUSINEK H, SAARIMIKI A, STEVENSON J, DEWEY SL, BRODIE JD (1998): *Functional magnetic resonance imaging of human brain activity in a verbal fluency task.* J Neurol Neurosurg Psychiatry 1998; **64**:492–498 CrossRef, Medline

SCHNEIDER F, FINK GR (2013). *Funktionelle MRT in Psychiatrie und Neurologie (2., überarb. und aktualisierte Aufl. ed.)*. Berlin [u.a.]: Springer Medizin Berlin Heidelberg.

SCHWEINITZ HT: *Erhebung der Test-Retest-Reliabilität der deutschen Adaptation der Eltern-, Lehrer-, Selbstbeurteilungsfragebögen der Conners 3 Langversion*, Diplomarbeit im Fachbereich Psychologie, Marburg 2015

SHROUT PE, & FLEISS JL (1979). *Intraclass correlations: uses in assessing rater reliability*. Psychol Bull, **86(2)**, 420-428.☒

SPECHT K, WILLMES K, SHAH NJ, JÄNCKE L, 2003. *Assessment of reliability in functional imaging studies*. J. Magn. Reson. Imaging **17 (4)**, 463–471.☒

SUSLOW T, ROESTEL C, AROLT V (2003). *Affective priming in schizophrenia with and without affective negative symptoms*, Eur Arch Psychiatry Clin Neurosci (2003) **253** : 292 – 300

TREPEL, M (2004): *Neuroanatomie- Struktur und Funktion*. **3. Auflage**, 217-218

TROYER AK. (2000): *Normative Data for Clustering and Switching on Verbal Fluency Tasks*. J Clin Exp Neuropsychology. **22:3**, 370-378

VAN DEN BUSSCHE E, VAN DEN NOORTGATE W, REYNFOET B. (2009). *Mechanisms of masked priming: a meta-analysis*. Psychol Bull. **135(3)**:452-77. doi: 10.1037/a0015329.

WEISHAUPT, D., KÖCHLI, V. D., & MARINCEK, B. (2014). *Wie funktioniert MRI?* **7.Auflage**, Springer Verlag

WENDE K, STRAUBE B, STRATMANN M, SOMMER J, KIRCHER T, NAGELS A (2012). *Neural correlates of continuous causal word generation*, NeuroImage **62**, 1399–1407

WHITNEY C, WEIS S, KRINGS T, HUBER W, GROSSMANN M, KIRCHER T (2009). *Task dependent modulations of prefrontal and hippocampal activity during intrinsic word production*. J Cogn Neurosci.;**21(4)**:697-712.

WILLIAMS MA, MC GLONE F, ABBOTT DF, MATTINGLEY JB (2005) *Differential amygdale responses to happy and fearful facial expressions depend on selective attention*. Neuroimage **24**: 417–425

WISE, RG & PRESTON, C (2010). *What is the value of human FMRI in CNS drug development?* Drug Discov Today, **15(21-22)**, 973-980.

WOJCIULIK E, KANWISHER N, DRIVER J. (1998). *Covert visual attention modulates face-specific activity in the human fusiform gyrus: fMRI study*. J Neurophysiol. **79(3)**: 1574-8.

YANG TT, MENON V, ELIEZ S, BLASEY C, WHITE CD, REID AJ, GOTLIB IH, REISS AL (2002). *Amygdalar activation associated with positive and negative facial expressions*, NeuroReport **Vol 13 No 14**, 1737-1741

YETKIN FZ, HAMMEKE TA, SWANSON SJ, MORRIS GL, MUELLER WM, MC AUCLIFFE TL, HAUGHTON VM (1995): *A comparison of functional MR activation patterns during silent and audible language tasks.* AJNR Am J Neuroradiol 1995; **16**:1087–1092

YOO, SS, WIE, X, DICKEY, CC, GUTTMANN, CRG, PANYCH, LP, 2005. *Long-term reproducibility analyses of fMRI using hand motor task.* Int. J. Neurosci. **115 (1)**, 55-77.

ZANBELT BB, GLADWIN TE, RAEMAKERS M, van BUUREN M, NEGGERS SF, KAHN RS, RAMSEY NF, VINK M. (2008): *Within-subject variation in BOLD-fMRI signal changes across repeated measurements: quantification and implications for sample size.* Neuroimage. 2008Aug 1;**42(1)**:196-206.

7 ABBILDUNGSVERZEICHNIS

Abb. 2.1 Lage des motorischen u. sensorischen Sprachzentrums in einer Brodmann-Karte...8	8
Abb. 2.2 Übersicht beteiligter Hirnregionen bei der Sprachproduktion.....10	10
Abb. 2.3 Spezifische Hirnregionen, die bei Emotionen eine Rolle spielen.....13	13
Abb. 3.4 Präzessionsbewegung..20	20
Abb. 3.5 Auswirkung HF-Impuls21	21
Abb. 3.6 Schema einer T1-Relaxation.....22	22
Abb. 3.7 Schema einer T2-Relaxation.....23	23
Abb. 3.8 Voxelschema24	24
Abb. 3.9 BOLD-Effekt26	26
Abb. 3.10 Funktionelle Datenauswertung mit SPM 827	27
Abb. 3.11 Prinzip des Realignment.....28	28
Abb. 3.12 T-Karte am Beispiel der Aktivierungskarte der Wortgenerierungs-aufgabe29	29
Abb. 3.13 Zeitverlauf einer Hämodynamischen Antwort und einer Neuronalen Antwort nach einer Stimulipräsentation31	31
Abb. 3.14 Hämodynamische Antwort und Messung in einem Block-Design.....31	31
Abb. 3.15 Boxcar Design32	32
Abb. 3.16 Übersicht der Paradigmen35	35
Abb. 3.17 Design und Designmatrix der WG- Aufgabe37	37
Abb. 3.18: Präsentation der Blöcke mit affektiven Stimuli38	38
Abb. 3.19 Design und Designmatrix der SPES-Aufgabe.....43	43
Abb. 3.21 Schema zur Differenzierung der ICC Berechnungsversion44	44
Abb. 3.22 Beispiel eines Streudiagramms der ICC und T-Werten Assoziatione.....47	47
Abb. 3.23 Beispiel einer relativen Häufigkeitsverteilung des gesamten Gehirns (brain), der deaktivierten Areale (white matter) und der aktivierten Netzwerke (network)47	47
Abb. 4.1 Übersicht der Aktivierungen der WG-Aufgabe mit anatomischen Zuordnungen im Transversalschnitt.....50	50
Abb. 4.2 Darstellung des aktivierten Netzwerks in der Gruppenanalyse der ersten Messung der WG-Aufgabe50	50

Abb. 4.3 Darstellung der Deaktivierungen der ersten Messung der WG-Aufgabe (Kontrolle > Aktivierungen).....	53
Abb. 4.4 Darstellung der Aktivierungen der beiden Messungen der WG-Aufgabe als Cluster-Überlappung	53
Abb. 4.5 Darstellung der Deaktivierungen (Kontrolle>Aktivierung) der beiden Messungen der WG-Aufgabe als Cluster-Überlappung	54
Abb. 4.6 ICC-vs.-T-Karte der WG-Aufgabe.....	54
Abb. 4.7 ICC-Häufigkeitskarte der WG-Aufgabe.....	55
Abb. 4.8 Übersicht der Aktivierungen des Kontrastes Gesichter>Kontrolle der SPES-Aufgabe mit anatomischen Zuordnungen im Transversalschnitt	57
Abb. 4.9 Darstellung des aktivierten Netzwerks in der Gruppenanalyse der ersten Messung der SPES-Aufgabe.....	58
Abb. 4.10 Übersicht der Aktivierungen des Kontrastes Angst>Neutral der SPES-Aufgabe mit anatomischen Zuordnungen im Transversalschnitt	61
Abb. 4.11 Darstellung der Hauptaktivierungen der SPES-Aufgabe mit Angst>Neutral Kontrast der ersten Messung in Schnittbildern.....	61
Abb. 4.12 Darstellung der Hauptaktivierungen der SPES-Aufgabe mit Angst>Neutral Kontrast der ersten Messung in 3D-Bildern	62
Abb. 4.13 Darstellung der Aktivierungen der beiden Messungen der SPES-Aufgabe im Gesichter>Kontrolle Kontrast als Cluster-Überlappung	65
Abb. 4.14 ICC-vs.-T-Karte des Kontrastes Gesichter>Kontrolle der SPES-Aufgabe.....	65
Abb. 4.15 ICC-Häufigkeitskarte des Kontrastes Gesichter>Kontrolle der SPES-Aufgabe.....	66
Abb. 4.15 Referenzregion des Kontrastes Gesichter>Kontrolle.....	67
Abb. 4.16 Darstellung der Aktivierungen der beiden Messungen der SPES-Aufgabe im Angst>Neutral Kontrast in Schnittbildern.....	68
Abb. 4.17 Darstellung der Aktivierungen der beiden Messungen der SPES-Aufgabe im Angst>Neutral Kontrast als Cluster-Überlappung	68
Abb. 4.18 ICC-vs.-T-Karte des Kontrastes Angst>Neutral der SPES-Aufgabe.....	69
Abb. 4.19 ICC-Häufigkeitskarte des Kontrastes Angst>Neutral der SPES-Aufgabe.....	69
Abb. 4.20 Referenzregion des Kontrastes Angst>Neutral: rechter MTG.	71
Abb. 4.21 Referenzregion des Kontrastes Angst>Neutral: bilaterale Amygdala.	71

8 TABELLENVERZEICHNIS

Tab. 2.1 Übersicht der den jeweiligen Schritten der Wortgenerierung zugeordneten Hirnregionen nach Price (2010)	9
Tab. 2.2 Übersicht der den Basisemotionen Wut, Trauer, Angst, Freude und Ekel zugeordneten Hirnregionen nach Fusar-Poli. (2009)	14
Tab. 3.1 Kategorien der Aufgabe Wortgenerierung der ersten und zweiten Messung.....	36
Tab. 4.1 Darstellung der anatomischen Korrelate der ersten Messung der WG-Aufgabe. <i>p(fwe korrigiert)= 0,05, k> 10.</i>	51
Tab. 4.2 Darstellung der anatomischen Korrelate der zweiten Messung der WG-Aufgabe. <i>p(unkorrigiert)= 0,01, k> 10.</i>	52
Tab. 4.3 Darstellung der anatomischen Korrelate der ersten Messung der SPES-Aufgabe im Kontrast Gesichter>Kontrolle. <i>p(unkorrigiert)= 0,001, k> 10.</i>	59
Tab. 4.4 Darstellung der anatomischen Korrelate der zweiten Messung der SPES-Aufgabe im Kontrast Gesichter>Kontrolle. <i>p(unkorrigiert)= 0,001, k> 10.</i>	60
Tab. 4.5 Darstellung der anatomischen Korrelate der ersten Messung der SPES-Aufgabe im Kontrast Angst>Neutral. <i>p(unkorrigiert)= 0,01, k> 10.</i>	63
Tab. 4.6 Darstellung der anatomischen Korrelate der zweiten Messung der SPES-Aufgabe im Kontrast Angst>Neutral. <i>p(unkorrigiert)= 0,001, k> 10.</i>	64

9 ABKÜRZUNGSVERZEICHNIS

<i>AC-PC</i>	<i>Linie zwischen anteriorer und posteriorer Kommissur</i>
<i>BA</i>	<i>Brodman-Areal</i>
<i>BOLD</i>	<i>Blood Oxygen Level Dependant</i>
<i>DTI</i>	<i>Diffusion Tensor Imaging</i>
<i>EPI</i>	<i>Echo Planar Imaging</i>
<i>FD</i>	<i>Free induction decay</i>
<i>FFA</i>	<i>Fusiform Face Area</i>
<i>fMRT</i>	<i>Funktionelle Magnetresonanztomographie</i>
<i>FoV</i>	<i>Field of View</i>
<i>FWE</i>	<i>family wise error rate</i>
<i>GE</i>	<i>Gradientenecho</i>
<i>GLM</i>	<i>General Linear Model</i>
<i>GRFT</i>	<i>Feldtheorie nach Gausscher Normalverteilung</i>
<i>HbO₂</i>	<i>oxygeniertes Hämoglobin</i>
<i>Hbr</i>	<i>desoxygeniertes Hämoglobin</i>
<i>HF</i>	<i>Hochfrequenz</i>
<i>HRF</i>	<i>Hämodynamische Antwortfunktion</i>
<i>ICD</i>	<i>International Statistical Classification of Diseases and Related Health Problems</i>
<i>ICC</i>	<i>Intraklassen-Korrelationskoeffizient (Intraclass Correlation Coefficient)</i>
<i>IFG</i>	<i>inferiorer Gyrus Frontalis</i>
<i>MNI</i>	<i>Montreal Neurological Institute</i>
<i>MR</i>	<i>Magnetresonanz</i>
<i>MRT</i>	<i>Magnetresonanztomographie</i>
<i>MTG</i>	<i>medialer Gyrus Temporalis</i>
<i>ROI</i>	<i>Region of Interest</i>
<i>SMA</i>	<i>supplementary motor area</i>
<i>SNR</i>	<i>Signal to Noise ratio</i>
<i>SPES</i>	<i>subliminale Präsentation emotionaler Stimuli</i>

<i>SPM</i>	<i>Statistical Parametric Mapping</i>
<i>STG</i>	<i>superiorer Gyrus Temporalis</i>
<i>TE</i>	<i>Time to Echo, Echo-Zeit</i>
<i>TR</i>	<i>Time of Repetition, Repetitionszeit</i>
<i>WFU</i>	<i>Wake Forest University</i>
<i>WG</i>	<i>Wortgenerierung</i>

10 ANHANG

10.1 Proband_innen- Instruktion

Allgemein

Sie erhalten eine Tastatur für die rechte Hand. Der Zeigefinger entspricht hier der linken Taste und der Mittelfinger entspricht der rechten Taste. Bitte nehmen Sie auch als Linkshänder die rechte Hand für ihre Antworten.

Versuchen Sie während der Messungen ihren Kopf still zu halten. Sie werden zwischendurch immer wieder Pausen haben und entspannen.

1) Resting State

Dieses Paradigma findet zu Beginn der Messung statt. Hier sollen Sie nachdem eine Ansagestimme über Kopfhörer die Instruktion gibt, die Augen schließen und 4 Minuten lang entspannen.

Nach 4 Minuten wird die Ansagestimme Ihnen die Instruktion geben, die Augen wieder zu öffnen und auf den Bildschirm zu blicken.

2) Verbal Fluency

Bei diesem Paradigma werden Ihnen 10 Kategorien vorgestellt, zu welchen Sie Begriffe deutlich in das Mikrofon sprechen sollen.

Sie werden jeweils die Kategorie vorgestellt bekommen und haben dann 12 Sekunden Zeit, Begriffe dazu aufzuzählen.

Kategorie: Körperteile	Hand, Fuß, Schulter, Nase, Rücken, Kopf, Stirn, Arm, Knie, Po, Ellenbogen,...
------------------------	---

Sobald das Wort 'RUHE' erscheint, ist die Sprechzeit vorbei und sie ruhen bis die nächste Kategorie vorgestellt wird.

RUHE	-----
------	-------

3a) Gesichter Encoding

Bei diesem Paradigma werden Ihnen unterschiedliche Bilder von Männern und Frauen gezeigt. Versuchen Sie, sich diese Bilder gut zu merken. Sie werden in der folgenden Aufgabe nochmals vorkommen.

Wenn Sie eine Frau auf dem Bild sehen, drücken Sie die linke Taste.

Wenn Sie einen Mann auf dem Bild sehen, drücken Sie die rechte Taste.

Wenn Sie ein gepixeltes Bild sehen, drücken Sie ebenfalls die linke Taste.

3b) Gesichter Decoding

Hier erscheinen nun zwei Gesichter gleichzeitig. Sie sollen bestimmen, welches der beiden Gesichter Sie im vorangegangenen Test schon gesehen haben.

Wenn das bekannte Gesicht auf der linken Seite ist, drücken Sie die linke Taste.

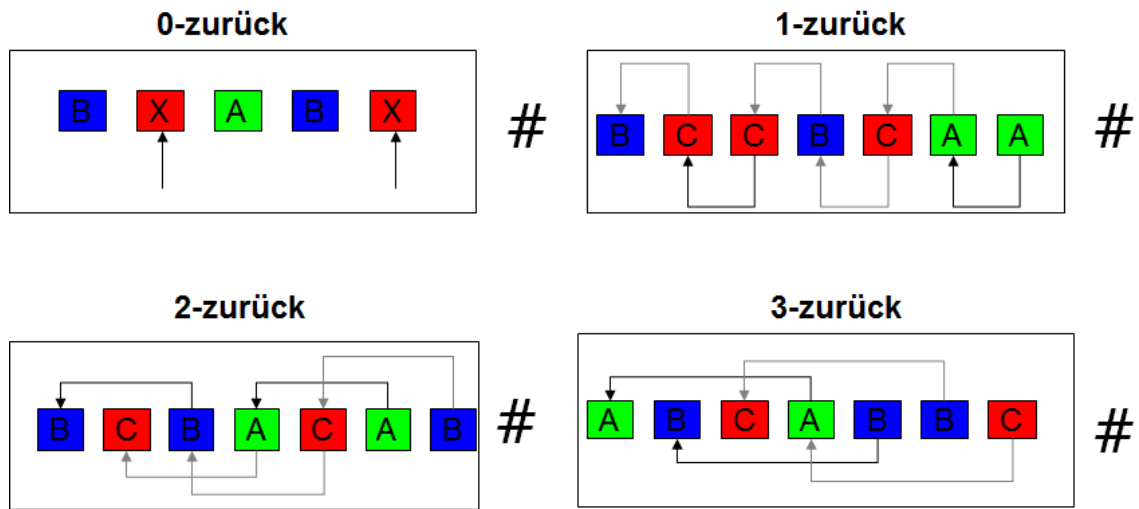
Wenn das bekannte Gesicht auf der rechten Seite ist, drücken Sie die rechte Taste.

Wenn zwei verpixelte Bilder erscheinen, drücken Sie ebenfalls die linke Taste

4) Subliminale Gesichter

Konzentrieren Sie sich auf die Gesichter.

5) N_Back



10.2 Proband_innen- Aufklärungsbogen

FACHBEREICH HUMANMEDIZIN



KLINIKUM
der Philipps-Universität
Marburg

Klinik für Psychiatrie und Psychotherapie
Philipps-Universität Marburg

Direktor: Prof. Dr. T. Kircher

Projektleiter:
Univ.-Prof. Dr. med. T. Kircher

Klinik für Psychiatrie und Psychotherapie der Philipps-Universität Marburg
Rudolf-Bultmann-Straße 8, 35039 Marburg
Tel: 06421-58-65200 (Skr.: 58-66219)
Fax:06421-58-68939

**Probanden-Information für Forschungsvorhaben mit fMRT-Untersuchungen
zur Vorbereitung der mündlichen Aufklärung durch den Arzt oder Psychologen
für die Studie: „Wirkungen von genetischen Risikovarianten für psychische
Störungen auf Gehirnstruktur und –funktion bei Gesunden und Erkrankten“**

Sehr geehrte Frau, sehr geehrter Herr,

wir möchten Sie bitten, an o.g. Studie teilzunehmen.

In diesem Informationsblatt geben wir Ihnen im Folgenden Hintergrundinformationen zur Studie, beschreiben die geplanten Untersuchungen, klären über mögliche Risiken auf, beschreiben den Umgang mit sog. „Zufallsbefunden“, konkretisieren den Umgang mit den erhobenen Daten und nennen abschließend Ausschlusskriterien für die Teilnahme an der Studie.

(1) Informationen zur geplanten Studie:

Der Hintergrund der Untersuchung sind die zunehmenden Befunde von speziellen Varianten im Erbgut (Kandidatengene), welche die Auftretenswahrscheinlichkeit von psychischen Störungen modulieren können. Schon seit längerer Zeit weiß man, dass das Auftreten von Depressionen, Bipolarer Störung oder Schizophrenie in bedeutsamer aber nicht vollständiger Weise durch das Erbgut mitbestimmt wird. Die Schwierigkeit dieser Ergebnisse liegt darin, dass einerseits nicht alle Patienten mit diesen Erkrankungen jene Genvarianten tragen und andererseits die Mehrzahl der Träger dieser Varianten psychisch gesund sind. Wir erklären dies durch ein hochkomplexes Wechselspiel von ganz verschiedenen Kandidatengenen und auch Umweltfaktoren.

Diese Untersuchung findet statt im Rahmen einer Studie zur Untersuchung von Effekten genetischer Risikovarianten auf psychologische Parameter und bestimmte Gehirnparameter. Im Rahmen dieser Studie findet neben mehreren anderen Verhaltenstests auch eine Untersuchung im Kernspintomographen (fMRT) statt. Zusätzlich wird Ihnen ca. 50 ml Blut entnommen. Wir möchten damit untersuchen, welchen Einfluss Genvarianten, die als bedeutsam für psychische Erkrankungen erwiesen haben, auf Erleben und Verhalten haben und wie sie zur Entstehung von psychischen Erkrankungen beitragen können. Dazu sollen neben Erkrankten auch 500 gesunde Kontrollprobanden untersucht werden, von denen 150 Probanden auch eine Untersuchung mittels funktioneller Magnetresonanztomographie erhalten sollen.

Die Teilnahme an dieser Studie ist freiwillig, Sie können jeder Zeit und ohne Angabe von Gründen Ihre Einwilligung zurückziehen, ohne dass Ihnen daraus Nachteile entstehen. Wir würden Ihnen eine Aufwandsentschädigung von 20 Euro nach der klinischen und neuropsychologischen Untersuchung und der Blutentnahme zahlen sowie im Falle der Teilnahme an der Untersuchung mit funktioneller Magnetresonanztomographie weitere 30 Euro.

(2) Hiermit bitten wir Sie um Zustimmung zu folgenden Untersuchungen:

1. Klinische und neuropsychologische Untersuchung
2. Blutentnahme zur Serumbestimmung und Genotypisierung
3. Funktionelle Untersuchung im Kernspintomographen

1. Klinische Untersuchung:

Vor Einschluss in die kernspintomographische Untersuchung wird mit Hilfe eines strukturierten Interviews nach DSM-IV eine psychiatrische Diagnose bei Ihnen ausgeschlossen. Weiterhin findet eine neuropsychologische Testung statt, bei der Aufmerksamkeits- und Gedächtnisfunktionen untersucht werden. Diese Untersuchungen werden insgesamt ca. 2 Stunden dauern.

2. Blutentnahme zur Serumbestimmung und Genotypisierung

Eine Blutentnahme (ca. 50 ml) wird über eine kleine Kanüle aus Metall erfolgen, wie sie auch im Krankenhausalltag oder beim Hausarzt durchgeführt wird. Es handelt sich hierbei um einen in der Medizin täglich praktizierten Routineeingriff, der mit sehr geringen Risiken verbunden ist. Dieses besteht im Einzelnen aus einem Blutungsrisiko und einem Infektionsrisiko. Da die Kanüle nur wenige Sekunden in der Vene liegt, ist das Infektionsrisiko außerordentlich gering. Auch das Blutungsrisiko ist bei der Punktion einer Vene extrem gering, etwa so wie bei einer Blutentnahme. Wegen des geringen Blutdrucks einer Vene besteht hier lediglich das Risiko ungefährlicher Einblutungen.

Ihre Genvarianten werden in der Abteilung für genetische Epidemiologie am Zentralinstitut für Seelische Gesundheit in Mannheim bestimmt. Dazu müssen wir Ihre Proben in anonymisierter Form an unsere Kooperationspartner dort (Prof. Dr. Marcella Rietschel) verschicken. Erfahrungen an großen Kollektiven haben gezeigt, dass einzelne genetische Befunde, wie sie in dieser Studie erhoben werden, nur innerhalb großer Stichproben aussagekräftig sind und sich nicht zur individuellen Vorhersage eines Erkrankungsrisikos eignen. Deshalb werden wir Ihnen keine individuellen Ergebnisse Ihrer Genanalysen mitteilen.

3. Funktionelle Untersuchung im Kernspintomographen:

Die Untersuchung im Kernspintomographen gliedert sich in mehrere Teile, die einzelnen Aufgaben werden Ihnen vorher ausführlich erklärt:

1. Wir zeigen wir Ihnen nacheinander verschiedene Buchstaben und Sie sollen je nach der spezifischen Instruktion diese nur ansehen, oder einen Knopf drücken, wenn Ihnen dieser Buchstabe bereits kurz zuvor gezeigt wurde.
2. In einer anderen Aufgabe werden Ihnen Gesichter gezeigt, die Sie sich merken sollen. Dabei sollen Sie ebenfalls bestimmen, ob es sich um einen Mann oder eine Frau handelt. Direkt im Anschluss werden Ihnen jeweils zwei Gesichter gezeigt und Sie sollen entscheiden, welches dieser Gesichter Sie bereits gezeigt bekommen haben.
3. In der dritten Aufgabe werden Ihnen Gesichter, zum Teil sehr schnell hintereinander, präsentiert. Ihre Aufgabe ist es sich die Gesichter einfach aufmerksam anzuschauen und zu bestimmen, ob es sich dabei um einen Mann oder eine Frau handelt.
4. Als vierte Aufgabe sollen Sie für ca. 5 Minuten mit geöffneten Augen ruhig liegen und dabei an nichts denken.
5. In der fünften Aufgabe werden Ihnen Wörter präsentiert, die eine Kategorie bilden (z.B. „Körperteile“). Sie sollen innerhalb kurzer Zeit so viele Unterbegriffe dieser Kategorie nennen, wie Ihnen einfallen (z.B. „Arm“, „Bein“, usw.).
6. Schließlich werden strukturelle Aufnahmen Ihres Gehirns gemacht. Dabei sollen Sie für ca. 20 Minuten ruhig liegen.

Insgesamt wird diese Untersuchung im Kernspintomographen ca. 60 Minuten dauern. Sie erhalten für Ihre Teilnahme an der gesamten Studie eine Aufwandsentschädigung, 20 Euro für die klinische und neuropsychologische Untersuchung und Blutentnahme sowie 30 Euro für die Untersuchung im Kernspintomographen.

Weiterhin ist es wichtig, den Einfluss von genetischen Faktoren auch über einen längeren Zeitraum zu untersuchen. Wir würden Sie deshalb gerne nach einem Jahr erneut kontaktieren, damit diese Untersuchungen bei Ihnen wiederholt werden

können. Bei diesem späteren Untersuchungszeitpunkt erhalten Sie erneut eine Aufklärung und können die Teilnahme an der späteren Untersuchung auch ablehnen, das heißt eine Teilnahme an der jetzigen Untersuchung setzt keine erneute Teilnahme zu einem späteren Zeitpunkt voraus.

(3) Ablauf der Kernspintomographieuntersuchung und mögliche Gefahren

Das menschliche Gehirn besteht aus einer Vielzahl von Molekülen, die jeweils spezifische magnetische Eigenschaften aufweisen (sog. Kernspinresonanz). Bei Anwendung von starken Magnetfeldern können diese gemessen und im Anschluss daran die Konzentration von Molekülen im Gehirn festgestellt werden. Auf diese Weise lässt sich die Durchblutung des Gehirns ermitteln, genauer die kontinuierlichen Veränderungen des Sauerstoffgehalts im Blut innerhalb weniger Sekunden. Diese Veränderungen ergeben spezifische Muster, je nachdem, welche Sinnesreize oder Vorstellungen die Versuchsperson während der Messung innerlich verarbeitet. Auf diese Weise können unterschiedliche Funktionen des Gehirns sichtbar gemacht werden. Man spricht hier von funktioneller Magnetresonanztomographie (fMRT).

Ablauf der Untersuchung: In dieser Untersuchung werden Sie gebeten, Töne, Geräusche, Sprache oder Bilder zu verarbeiten. Um bei den Messungen im Gehirn verlässliche Ergebnisse zu erhalten, ist Ihre Kooperationsbereitschaft erforderlich.

Sie liegen während der Untersuchung auf einem Tisch, der Sie mit dem Oberkörper in die Öffnung des MR-Tomographen hineinführt. Zusätzlich wird ein Magnetspulenrahmen, ähnlich einem Helm, um Ihren Kopf gelegt.

Während der Messung entstehen unterschiedlich laute Geräusche des MR-Tomographen, die von elektrischen Umschaltvorgängen der Magnetfelder herrühren. Um deren Einfluss gering zu halten, tragen Sie einen Kopfhörer. Sie haben während der Untersuchung jederzeit die Möglichkeit, mit den Untersuchern über eine Wechselsprechanlage in Kontakt zu treten. Zusätzlich bekommen Sie einen gesonderten Alarmknopf für Notfälle, so dass sie auf Ihren Wunsch hin jederzeit aus dem MR-Tomographen herausgefahren werden können.

Mögliche Gefahren: Die MRT-Technologie ist für den Körper nach heutigem Erkenntnisstand unschädlich. Sie basiert auf mehr als 20 Jahren Erfahrung und wird täglich in allen größeren Kliniken eingesetzt. Bekannte Risiken ergeben sich ausschließlich durch metallische Gegenständen oder Stoffe mit magnetischen Eigenschaften, die sich am oder im Körper befinden. Diese können sich erhitzen und zu Verbrennungen führen. Lose Metallteile können durch das Magnetfeld beschleunigt werden und dann zu Verletzungen führen. Daher sind Personen von der Teilnahme an der Studie ausgeschlossen, die elektrische Geräte (z.B. Herzschrittmacher, Medikamentenpumpen) oder Metallteile (z.B. Schrauben nach Knochenbruch) im oder am Körper haben.

Risiken unabhängig von den Genannten sind bislang nicht bekannt. Abgesehen von möglichen Unbequemlichkeiten, die vom stillen Liegen in der engen fMRT-Röhre resultieren, und der Lautstärke des Tomographen, sollten keine Beschwerden während der Untersuchung auftreten.

Wir möchten sie allerdings darauf hinweisen, dass über mögliche langfristige Risiken bei wiederholten MRT-Messungen bisher keine wissenschaftlich abgesicherten Ergebnisse vorliegen.

(4) Umgang mit Zufallsbefunden

Bei der Studie handelt es sich um eine Forschungsstudie. Eine neuroradiologische Befundung der MR-Bilder im Sinne einer klinisch orientierten Diagnostik findet daher nicht statt. Dennoch kann es vorkommen, dass in den MR-Bildern Signalauffälligkeiten entdeckt werden, die eine mögliche klinische Relevanz haben („Zufallsbefund“). Falls sich bei der Untersuchung Anhaltspunkte für einen Zufallsbefund ergeben, die eine fachärztliche neuro-radiologische Diagnostik empfehlenswert erscheinen lassen, würde Sie der Versuchsleiter persönlich darüber informieren und Ihnen eine fachärztlich neuro-radiologische Diagnostik empfehlen. Falls Sie über einen Zufallsbefund nicht informiert werden wollen, stellt dies ein Ausschlusskriterium für die Teilnahme an der Studie dar.

(5) Datenschutz

Ihre personenbezogenen Daten werden maschinell gespeichert und weiterverarbeitet. Die Weitergabe, Speicherung und Auswertung dieser

studienbezogenen Daten erfolgt nach gesetzlichen Bestimmungen ohne Namensnennung und setzt vor Teilnahme an der Studie Ihre freiwillige Einwilligung voraus.

(6) Ausschlusskriterien

Die Anwendung von Magnetfeldern bei der MRT-Untersuchung schließt die Teilnahme von Personen aus, die elektrische Geräte (z.B. Herzschrittmacher, Medikamentenpumpen usw.) oder Metallteile (z.B. Schrauben nach Knochenbruch, „Spirale“) im oder am Körper haben. Frauen, die schwanger sind, werden nicht als Probandinnen zugelassen. Probanden, die nicht über mögliche Zufallsbefunde informiert werden wollen, können ebenfalls nicht an der Studie teilnehmen.

Wenn weitere Fragen bestehen, werden diese gerne vom jeweiligen Untersuchungsleiter beantwortet.

**Einwilligungserklärung zur Teilnahme an dem
Forschungsvorhaben „Wirkungen von genetischen Risikovarianten
für psychische Störungen auf Gehirnstruktur und –funktion bei
Gesunden und Erkrankten“**

Bei Ihrer Bereitschaft zur Teilnahme bitten wir Sie, die Einwilligungserklärung vor der Untersuchung vollständig auszufüllen und zu unterschreiben.

Ich bestätige hiermit, dass ich durch den Untersucher, Herrn/Frau über Wesen, Bedeutung, Risiken und Tragweite der beabsichtigten Untersuchung aufgeklärt wurde und für meine Entscheidung genügend Bedenkzeit hatte. Ich habe die Probandeninformation gelesen. Ich fühle mich ausreichend informiert und habe verstanden, worum es geht. Der Untersucher hat mir ausreichend Gelegenheit gegeben, Fragen zu stellen, die alle für mich ausreichend beantwortet wurden. Ich hatte genügend Zeit, mich zu entscheiden.

Ich wurde darauf hingewiesen, dass es sich bei der Studie um eine Forschungsstudie handelt. Eine neuroradiologische Befundung der MR-Bilder im Sinne einer klinisch orientierten Diagnostik findet daher nicht statt. Dennoch kann es vorkommen, dass in den MR-Bildern Signalauffälligkeiten entdeckt werden, die eine mögliche klinische Relevanz haben („Zufallsbefund“). Mir ist bekannt, dass der Versuchsleiter mich informieren würde, falls sich bei der Untersuchung Anhaltspunkte für einen Zufallsbefund ergeben, die eine fachärztliche neuro-radiologische Diagnostik empfehlenswert erscheinen lassen.

Ich habe verstanden, dass bei wissenschaftlichen Studien persönliche Daten und medizinische Befunde erhoben werden. Die Weitergabe, Speicherung und Auswertung dieser studienbezogenen Daten erfolgt nach gesetzlichen Bestimmungen und setzt vor Teilnahme an der Studie meine freiwillige Einwilligung voraus. Ich erkläre mich damit einverstanden, dass im Rahmen dieser Studie erhobene Daten auf Fragebögen und elektronischen Datenträgern aufgezeichnet und ohne Namensnennung zum Zwecke wissenschaftlicher Auswertung analysiert werden. Ich erkläre mich damit einverstanden, dass ich zur möglichen Teilnahme an der gleichen Forschungsstudie zu einem späteren Zeitpunkt kontaktiert werde.

Ich habe eine Kopie der Probandeninformation und dieser unterschriebenen Einwilligungserklärung erhalten. Meine Einwilligung, an diesem Forschungsvorhaben als Proband teilzunehmen, erfolgt freiwillig. Ich wurde darauf hingewiesen, dass ich meine Einwilligung jederzeit ohne Angabe von Gründen und ohne Nachteile widerrufen kann.

Dürften wir Sie unverbindlich für die mögliche Teilnahme an weiteren Studien neben dieser kontaktieren?

JA NEIN

Ich willige hiermit ein, als Proband an dem Forschungsvorhaben, „Wirkungen von genetischen Risikovarianten für psychische Störungen auf Gehirnstruktur und –funktion bei Gesunden und Erkrankten“ teilzunehmen.

VOM PROBANDEN AUSZUFÜLLEN:

Name:

Geburtsdatum:

Datum: Uhrzeit:

Ort: Unterschrift:

VOM UNTERSUCHER AUSZUFÜLLEN:

Ich habe den Probanden mündlich über Wesen, Bedeutung, Reichweite und Risiken des Forschungsvorhabens aufgeklärt.

Datum: Uhrzeit:

Ort:

Untersucher:

11 AKADEMISCHE LEHRER

DAMEN/HERREN IN MARBURG

Alter	Jerrentrup	Renz
Bartsch	Kircher	Röhm
Basler	Klose	Ruchholtz
Baum	König	Schäfer
Baumann	Konrad	Schmidt
Becker	Koolman	Schulze
Bohlander	Lill	Seitz
Brehm	Löffler Lohoff	Sesterhenn
Cetin	Maier	Vogelmeier
Czubayko	Maisch	Vogt
Daut	Mandrek	Wagner
Fendrich	Michl	Waldmann
Feuser	Moll	Weihe
Gress	Mueller	Werner
Grundmann	Neubauer	Westermann
Hertl	Oertel	Wulff
Hofmannn	Pagenstecher	Zemlin
Hoyer	Preisig-Müller	

