
Inauguraldissertation
zur Erlangung des akademischen Doktorgrades (Dr. phil.)
im Fach Psychologie
an der Fakultät für Verhaltens- und Empirische Kulturwissenschaften
der Ruprecht-Karls-Universität Heidelberg

Neuropsychologie der Zwangsstörung
unter Berücksichtigung ihrer Symptomdimensionen:
Eine Metaanalyse

Vorgelegt von

Rebecca Leopold

Jahr der Einreichung
2014

Dekan: Prof. Dr. phil. Klaus Fiedler

Berater: Prof. Dr. phil. Matthias Backenstraß

Danksagung

Viele Menschen haben dazu beigetragen, dass diese Dissertation möglich wurde. Mein herzlicher Dank geht zu allererst an meinen Doktorvater Prof. Dr. Matthias Backenstraß für die Begeisterung, mit der er meine Idee aufgenommen hat, für die Unterstützung mit Projektmitteln, für die fruchtbare Zusammenarbeit bei der Publikation, die aus dieser Arbeit entstanden ist und für die Präzision und Sorgfalt, die er in seinen Kommentaren und Anregungen walten ließ. Herrn Prof. Dr. Willi Ecker danke ich für die wertvollen Anregungen in seinen Habilitationsschriften. Auch danke ich den internationalen Wissenschaftlern, die mit ihren Studien eine Metaanalyse erst ermöglichten und mir per E-Mail fehlende Daten zukommen ließen. Ein ganz besonderer Dank geht an Diana Byrd Burt von der University of Texas Medical School in Houston, USA, die verlorene Teile ihrer *data coding form* für mich rekonstruierte. Den Kolleginnen der Forschungsgruppe Psychotherapie und empirische Pathopsychologie, Annette und Judith, danke ich für die fröhliche Atmosphäre und für die Erfahrung, dass Wissenschaft viel Freiheit bedeutet. An meine Mitstreiterinnen Verena und Sabrina geht mein Dank für ihre tatkräftige Unterstützung beim Kodieren und für die vielen inspirierenden Diskussionen. Besonders dankbar bin ich meiner Familie und meinen Freunden, die mich in dieser Zeit begleitet haben. Mein größter Dank geht an meinen geliebten Ehemann Martin – ohne seine Ermutigung und Unterstützung wäre dieses Vorhaben nicht gelungen.

Zusammenfassung

Uneinheitliche Befunde zu neuropsychologischen Leistungen von Zwangspatienten können mit der Symptomheterogenität der Zwangsstörung zusammenhängen. Empirisch relativ gut bestätigte Symptomdimensionen sind *Waschen/Kontamination*, *Kontrollieren/Zweifel*, *verbotene Gedanken*, *Symmetrie/Ordnung* sowie *Sammelzwang*. Die Fragestellung der vorliegenden Metaanalyse lautete: Unterscheiden sich Zwangspatienten mit Schwerpunkten auf diesen Symptomdimensionen von Gesunden und voneinander hinsichtlich ihrer neuropsychologischen Leistungen? Das Ziel der Arbeit bestand darin, einen ausführlichen Überblick über die aktuelle Studienlage zur Neuropsychologie der Zwangsstörung unter Berücksichtigung ihrer Symptomdimensionen zu geben. Hierzu wurden eine systematische Recherche in Online-Datenbanken und Literaturlisten durchgeführt und führende Wissenschaftler kontaktiert. Studien wurden in die Metaanalyse aufgenommen, wenn sie neuropsychologische Testergebnisse von Zwangspatienten mit den oben genannten Symptomschwerpunkten berichteten. Von 6 069 gefundenen Publikationen wurden 38 Studien mit 1 491 Zwangspatienten und 1 091 gesunden Kontrollen eingeschlossen und sieben Metaanalysen durchgeführt. Im Ergebnis unterschieden sich die Patienten mit Waschzwang bei gleicher Intelligenz und gleichem Bildungsniveau hinsichtlich neuropsychologischer Leistungen nicht signifikant von Gesunden. Dagegen schnitten die Patienten mit Kontrollzwang in der Inhibition, im Konzeptwechsel, im verbalen und nonverbalen Gedächtnis sowie im räumlichen und verbalen Arbeitsgedächtnis signifikant schlechter ab als die gesunden Probanden. Im Vergleich beider Patientengruppen erbrachten Patienten mit Kontrollzwang größere Defizite im Planen/Problemlösen und in der Inhibition. Patienten mit Kontrollzwang zeigten gegenüber jenen mit Waschzwang zudem signifikante Defizite in anhaltender Aufmerksamkeit, im Konzeptwechsel, im Enkodieren, im verbalen und nonverbalen Gedächtnis sowie in der Verarbeitungsgeschwindigkeit. In keiner kognitiven Subdomäne oder Domäne wiesen Patienten mit Kontrollzwang bessere Leistungen auf als Patienten mit Waschzwang. Zwangspatienten mit dem Symptomschwerpunkt auf verbotenen

Gedanken erbrachten im Vergleich zu Gesunden deutlich verminderte Leistungen in Einzeltests zum Planen/Problemlösen, zum Konzeptwechsel und zum räumlichen Arbeitsgedächtnis. Allerdings ließen sich aus nur zwei eingeschlossenen Studien keine endgültigen Aussagen treffen. Hingegen zeigten Patienten mit Symmetrie-/ Ordnungszwang gegenüber Gesunden vor allem im verbalen und nonverbalen Gedächtnis signifikante Defizite. Patienten mit Sammelzwang zeigten im Vergleich zu Gesunden und zu Zwangspatienten ohne Sammelzwang ein deutlich höheres Durchschnittsalter, einen höheren Frauenanteil und eine längere Krankheitsdauer. Patienten mit Sammelzwang wiesen gegenüber Gesunden signifikante Defizite im Planen/Problemlösen, in der Inhibition, im Konzeptwechsel, im nonverbalen Gedächtnis, im räumlichen Arbeitsgedächtnis, im Entscheidungsverhalten und in der Kategorisierung auf. Der direkte Vergleich von Patienten mit versus ohne Sammelzwang ergab ein unklares Bild. Dabei blieb offen, ob, und wenn ja, in welchen kognitiven Domänen Patienten mit Sammelzwang schwerer beeinträchtigt sind als Zwangspatienten ohne Sammelzwang. Lediglich im Entscheidungsverhalten zeigten die Patienten mit Sammelzwang ein signifikantes, aber geringes Defizit, verglichen mit Zwangspatienten ohne Sammelzwang. Begrenzungen der Metaanalyse bestanden in einer relativ geringen Anzahl von Primärstudien und in kleinen Stichproben. Zwar erschienen die meisten neuropsychologischen Leistungsunterschiede von geringer praktischer Bedeutsamkeit. Jedoch liefern die Ergebnisse dieser Metaanalyse über 38 Studien hinweg empirische Evidenz dafür, dass sich die Symptomdimensionen der Zwangsstörung in ihren neuropsychologischen Profilen unterscheiden. Dies kann als Beleg für die Validität der Symptomdimensionen gelten. Methodische Forderungen, therapeutische Implikationen und Ansätze für weitere Forschungsvorhaben werden diskutiert.

Schlüsselwörter: Zwangsstörung, Symptomdimensionen, Kontrollzwang, Sammelzwang, Neuropsychologie, Metaanalyse

Inhaltsverzeichnis

Danksagung.....	1
Zusammenfassung.....	2
Inhaltsverzeichnis.....	4
Einleitung.....	6
1 Neuropsychologie der Zwangsstörung.....	8
1.1 Warum neuropsychologische Forschung?.....	9
1.2 Aufmerksamkeit.....	19
1.3 Exekutivfunktionen.....	20
1.4 Gedächtnis und Arbeitsgedächtnis.....	26
1.5 Informationsverarbeitungsgeschwindigkeit.....	31
1.6 Entscheidungsverhalten und Kategorisierung.....	32
1.7 Fazit zur Neuropsychologie.....	34
2 Symptomdimensionen der Zwangsstörung.....	35
2.1 Multivariate Verfahren.....	35
2.2 Das klinische Bild der Symptomdimensionen.....	39
2.3 Familiengenetische und Zwillingsstudien.....	45
2.4 Bildgebungsbefunde.....	46
2.5 Komorbidität.....	48
2.6 Ansprechen auf Verhaltenstherapie und Psychopharmaka.....	49
2.7 Fazit zu Symptomdimensionen.....	51
3 Forschungsstand, Ziel und Fragestellung.....	53
4 Methoden.....	59
4.1 Systematische Recherche.....	59
4.2 Datenextraktion, Studien- und Kodierungsgüte.....	61
4.3 Datenanalyse mittels Effektstärken.....	66
4.4 Integration der Effektstärken.....	72
4.5 Moderatoranalysen.....	73
4.6 Metaregression.....	74
4.7 Abschätzung der Publikationsverzerrung.....	75
5 Ergebnisse.....	78
5.1 Literaturrecherche.....	78

5.2 Studienmerkmale.....	82
5.3 Studienqualität.....	89
5.4 Teilnehmer	91
5.5 Metaanalyse 1: Patienten mit Waschzwang vs. Gesunde.....	93
5.6 Metaanalyse 2: Patienten mit Kontrollzwang vs. Gesunde.....	97
5.7 Metaanalyse 3: Patienten mit verbotenen Gedanken vs. Gesunde.....	101
5.8 Metaanalyse 4: Patienten mit Symmetrie-/Ordnungszwang vs. Gesunde.....	103
5.9 Metaanalyse 5: Patienten mit Sammelzwang vs. Gesunde	106
5.10Metaanalyse 6: Patienten mit Waschzwang vs. Kontrollzwang	111
5.11Metaanalyse 7: Patienten mit vs. ohne Sammelzwang.....	115
5.12Kombinierte Effektstärken der sieben Metaanalysen im Überblick	119
6 Diskussion	122
6.1 Methoden.....	122
6.2 Metaanalyse 1: Patienten mit Waschzwang vs. Gesunde.....	131
6.3 Metaanalyse 2: Patienten mit Kontrollzwang vs. Gesunde.....	132
6.4 Metaanalyse 3: Patienten mit verbotenen Gedanken vs. Gesunde.....	134
6.5 Metaanalyse 4: Patienten mit Symmetrie-/Ordnungszwang vs. Gesunde.....	135
6.6 Metaanalyse 5: Patienten mit Sammelzwang vs. Gesunde	136
6.7 Metaanalyse 6: Patienten mit Waschzwang vs. Kontrollzwang	138
6.8 Metaanalyse 7: Patienten mit vs. ohne Sammelzwang.....	139
6.9 Generelle Diskussion.....	141
Literaturverzeichnis.....	152
Tabellenverzeichnis.....	181
Abbildungsverzeichnis.....	182
Abkürzungsverzeichnis	184
Anhang	187
Anhang 1 Kodierung deskriptiver Daten	187
Anhang 2 Kodierung neuropsychologischer Testergebnisse	192
Anhang 3 Kodiermanual	210
Anhang 4 Skala zur Beurteilung der Studienqualität.....	211
Anhang 5 Interraterreliabilität.....	214
Erklärung gemäß § 8 Abs. 1 Buchst. b) und c) der Promotionsordnung	215

Einleitung

Gemäß der fünften Ausgabe des Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013) erfordern die vier diagnostischen Kriterien einer Zwangsstörung (*obsessive compulsive disorder*, OCD), dass Zwangsgedanken, Zwangshandlungen oder beide vorliegen (Kriterium A), dass diese zeitaufwändig sind oder klinisch bedeutsames Leiden verursachen (Kriterium B), dass sie nicht auf den Gebrauch einer Substanz oder eine medizinische Erkrankung zurückgehen (Kriterium C) und dass sie nicht durch eine andere psychische Störung besser erklärt werden können (Kriterium D). Menschen, die an einer Zwangsstörung leiden, zeigen sehr unterschiedliche Krankheitsbilder. Die Betroffenen führen Zwangshandlungen wie Waschen, Kontrollieren, Ordnen und Sammeln aus, und sie leiden an quälenden Gedanken von Ansteckung, Katastrophen, Gewalt, Sexualität und religiösen Inhalten. Man nimmt an, dass sich zwischen den Zwangssymptomen und den ihnen zugrunde liegenden neurobiologischen Mechanismen des Gehirns der Bereich der neuropsychologischen Funktionen als Bindeglied befindet (Olley, Malhi & Sachdev, 2007). So stehen z. B. sich aufdrängende Zwangsgedanken mit eingeschränkten kognitiven Inhibitionsprozessen in Verbindung, während wiederholte Zwangshandlungen möglicherweise mit der gestörten Fähigkeit zum Konzeptwechsel (*set shifting*) zusammenhängen, d. h., mit der Fähigkeit, zwischen Aufgaben, Tätigkeiten und mentalen Zuständen zu wechseln (Chamberlain, Blackwell, Fineberg, Robbins & Sahakian, 2005).

Empirische Befunde zur Neuropsychologie der Zwangsstörung insgesamt sind uneinheitlich. Einige Forscher fanden Defizite in Aufmerksamkeit, Gedächtnis und Exekutivfunktionen (Vgl. die Übersichtsarbeiten von Kuelz, Hohagen & Voderholzer, 2004; Muller & Roberts, 2005), während andere dies nicht belegen konnten (Abbruzzese, Ferri & Scarone, 1995; Moritz, Fischer et al., 2008; Moritz, Ruhe, Jelinek & Naber, 2009; Moritz, Wendt, Jelinek, Ruhe & Arzola, 2008; Simpson et al., 2006). Eine mögliche Erklärung für die unklare Befundlage besteht darin, dass die Zwangsstörung eine weniger homogene Störung ist, als es die bisherigen Definitionen der Internationalen Klassifikation psychischer Störungen

(ICD–10; Dilling, 2008) und der vierten Ausgabe des Diagnostic and Statistical Manual of Mental Disorders (DSM–IV; Saß & Houben, 1996) vorsahen.

Im DSM–5 (American Psychiatric Association, 2013) wurden Symptomdimensionen der Zwangsstörung spezifiziert: *Waschen/Kontamination*, *Kontrollieren/Zweifel*, *verbotene Gedanken*, *Ordnen/Symmetrie* und *Sammeln (hoarding)*. Diese Symptomdimensionen unterscheiden sich durch ihr klinisches Erscheinungsbild. Darüber hinaus lassen sich bei einzelnen Symptomdimensionen bestimmte familiengenetische und gehirneurophysiologische Merkmale, Häufungen komorbider Erkrankungen und unterschiedliche Behandlungserfolge beobachten (Mataix-Cols, Marks, Greist, Kobak & Baer, 2002). Da die Symptomdimensionen der Zwangsstörung in derart vielen Bereichen Differenzen aufweisen, stellt sich die Frage, ob ihnen auch unterschiedliche neuropsychologische Beeinträchtigungen entsprechen (Omori et al., 2007).

Das Ziel dieser Arbeit besteht darin, die bisherigen Forschungsbefunde über die neuropsychologischen Beeinträchtigungen von Zwangspatienten mit verschiedenen Symptom Schwerpunkten zusammenzufassen. Zur Neuropsychologie der Zwangspatienten gibt es zahlreiche Publikationen. Auch existieren Einzelstudien, die neuropsychologische Defizite von Zwangserkrankten mit bestimmten Symptom Schwerpunkten untersuchen. Jedoch wurde noch keine systematische Übersichtsarbeit zur Neuropsychologie von Zwangspatienten unter Berücksichtigung der verschiedenen Symptomdimensionen veröffentlicht. Diese Forschungslücke soll die vorliegende Arbeit schließen. Das erste Kapitel stellt theoretische Grundlagen, neuropsychologische Konstrukte und Testverfahren vor. Das zweite Kapitel beschreibt Symptomdimensionen, das dritte Kapitel den aktuellen Stand der Forschung und die sich daraus ergebenden Ziele und Fragen. Im vierten Kapitel werden die Methoden der Metaanalyse erläutert, im fünften Kapitel die Ergebnisse vorgestellt. Die Arbeit schließt im sechsten Kapitel mit einer Diskussion der Methoden und Ergebnisse sowie mit einem Ausblick auf künftige Forschung.

Um den Lesefluss zu erleichtern, wird im Text durchgehend das Maskulinum verwendet. Es sind hiermit stets weibliche und männliche Personen gemeint.

1 Neuropsychologie der Zwangsstörung

Zwangsstörungen äußern sich in Zwangsgedanken und/oder Zwangshandlungen, die der Betroffene gegen seinen Willen in stereotyper Wiederholung denken bzw. ausführen muss, die er häufig als sinnlos erlebt, die einen hohen Zeitaufwand erfordern und die Leiden oder deutliche Beeinträchtigung verursachen (Saß & Houben, 1996). Die meisten Zwangspatienten leiden sowohl unter Zwangsgedanken als auch unter Zwangshandlungen (Foa & Kozak, 1995). Die Lebenszeitprävalenz der Zwangsstörung liegt bei zwei bis drei Prozent (Ruscio, Stein, Chiu & Kessler, 2010). Männer und Frauen sind etwa gleich häufig betroffen. Allerdings scheinen Frauen häufiger Waschwänge zu entwickeln, Männer hingegen häufiger sexuelle und religiöse Zwangsgedanken (Labad et al., 2008). Für das Alter bei Ersterkrankung lässt sich eine zweigipflige Verteilung beobachten. Gehäuft finden sich erste Zwangssymptome zu Beginn der Pubertät im Alter von 12 – 14 Jahren sowie im frühen Erwachsenenalter von 20 – 22 Jahren (Shafran, 2001; Snider & Swedo, 2000).

Zwangssymptome sind stark angst- und schambesetzt. Daher begeben sich Betroffene durchschnittlich erst sechs Jahre nach Ausbrechen der Erkrankung in Behandlung (Voderholzer, Schlegl & Külz, 2011). Ohne angemessene Therapie verlaufen Zwangserkrankungen chronisch, fortschreitend oder mit schwankender Symptomatik (Voderholzer & Hohagen, 2013). Symptomfreiheit wird unbehandelt fast nie erreicht, und Spontanremissionen sind äußerst selten (Zaudig & Bossert-Zaudig, 2002). Die Behandlung gemäß der S3-Leitlinie Zwangsstörungen (Kordon, Lotz-Rambaldi, Muche-Borowski & Hohagen, 2013) erfolgt mittels kognitiver Verhaltenstherapie (Exposition mit Reaktionsverhinderung), alternativ oder zusätzlich mit Psychopharmaka (vor allem mit selektiven Serotonin-Wiederaufnahmehemmern, SSRI, und ergänzend mit Neuroleptika). In schweren Ausnahmefällen kommen als Ultima Ratio gehirnchirurgische Methoden zum Einsatz. Hierbei werden Faserverbindungen der cortico-striato-thalamischen Regelkreise irreversibel durchtrennt (Rück et al., 2008), oder es werden bei der

elektrischen Tiefenhirnstimulation Elektroden implantiert, die die Gehirnaktivität in den entsprechenden Regionen zeitweise reduzieren (Mallet et al., 2008).

In den letzten 20 Jahren hat sich die Erforschung der Zwangsstörung auf Bereiche ausgedehnt, die weit über phänomenologische und therapeutische Ansätze hinausgehen. Die neurobiologische Zwangsforschung untersucht Gehirnstrukturen und -funktionen von Patienten anhand bildgebender Verfahren, und sie steht in engem Zusammenhang mit der neuropsychologischen Zwangsforschung. Letztere untersucht kognitive Leistungen und Fehlfunktionen anhand von psychologischen Modellen und Testverfahren, und sie ist Thema der vorliegenden Arbeit. Im folgenden Absatz wird erläutert, warum neuropsychologische Merkmale bei Zwangspatienten wissenschaftlich untersucht werden. Hierzu wird, ausgehend vom Konzept der Endophänotypen und über das cortico-baso-thalamische Modell der Zwangsstörung, ein neuropsychologisches Modell der Zwangsstörung abgeleitet.

1.1 Warum neuropsychologische Forschung?

Endophänotypen als Vulnerabilitätsfaktoren. Wie die Entstehung aller psychischen Störungen lässt sich auch die Entstehung der Zwangsstörung über ein allgemeines Vulnerabilitäts-Stress-Modell erklären. Man nimmt an, dass Individuen aufgrund bestimmter Voraussetzungen (z. B. genetischer Bedingungen) eine Prädisposition für eine Erkrankung „mitbringen“. Diese Menschen weisen eine erhöhte Vulnerabilität auf. Kommen belastende Umweltbedingungen als Stressoren hinzu, kann aus der latenten Krankheitsvorstufe eine manifeste Erkrankung werden. Welche Merkmale eines Individuums Hinweise auf dessen Vulnerabilität geben, wird in der translationalen Psychiatrie erforscht. Dieser Forschungszweig verbindet Grundlagenforschung mit klinischen Behandlungskonzepten und untersucht komplexe Störungen wie Schizophrenie, Depression und die Zwangsstörung.

Ein übergeordnetes Rational für die translationale Forschung stellt das Endophänotypen-Konzept dar (Gottesman & Gould, 2003). Endophänotypen sind vermittelnde Faktoren zwischen Genen und Verhalten, „mediating factors between genes and behavior“ (Hollander, 2006, S. 737). Auf ihrer Grundlage entwickelten

Chamberlain und Menzies (2009) ein hierarchisches Modell verschiedener Ebenen, auf denen sich die Pathologie einer psychischen Störung manifestieren kann. Auf oberster Ebene befindet sich der Genotyp, d. h., die genetische Prädisposition der Erkrankung. Auf unterster Ebene steht der klinisch beobachtbare Phänotyp, d. h., das Syndrom bzw. die Diagnose. Zwischen Geno- und Phänotyp liegen verbindende Zwischenglieder, die hierarchisch organisierten Endophänotypen. Zu ihnen zählen, ausgehend vom Genotypen, die Ebenen der Proteine, der Zellen, der Organ- und Funktionssysteme und der Verhaltensweisen (vgl. Abbildung 1). Die Endophänotypen sind durch Umweltfaktoren beeinflussbar und bedingen einander wechselseitig. Sie stellen bei entsprechender Ausprägung Risikofaktoren für das Ausbrechen der klinisch beobachtbaren Symptomatik dar. Untersucht man die Zwangsstörung gezielt auf ihre Endophänotypen, dann erhält man ein umfassenderes Krankheitsverständnis. Dies ist für eine mögliche Früherkennung, den Störungsverlauf, die Prognose, die differentielle Behandlungsindikation und den Therapieerfolg von Bedeutung.

Endophänotypen der Zwangsstörung werden auf genetischer Ebene durch familiengenetische und Zwillingsstudien erforscht, auf der Ebene der Zellen und Proteine durch immunologische Studien. Auf der Ebene der Organ- und Funktionssysteme werden strukturelle und funktionelle Merkmale des Gehirns zwangserkrankter Personen untersucht, wohingegen die Ebenen der Verhaltensweisen und des klinischen Phänotypen im Rahmen klinischer Beobachtungs- und Therapiestudien erforscht werden.

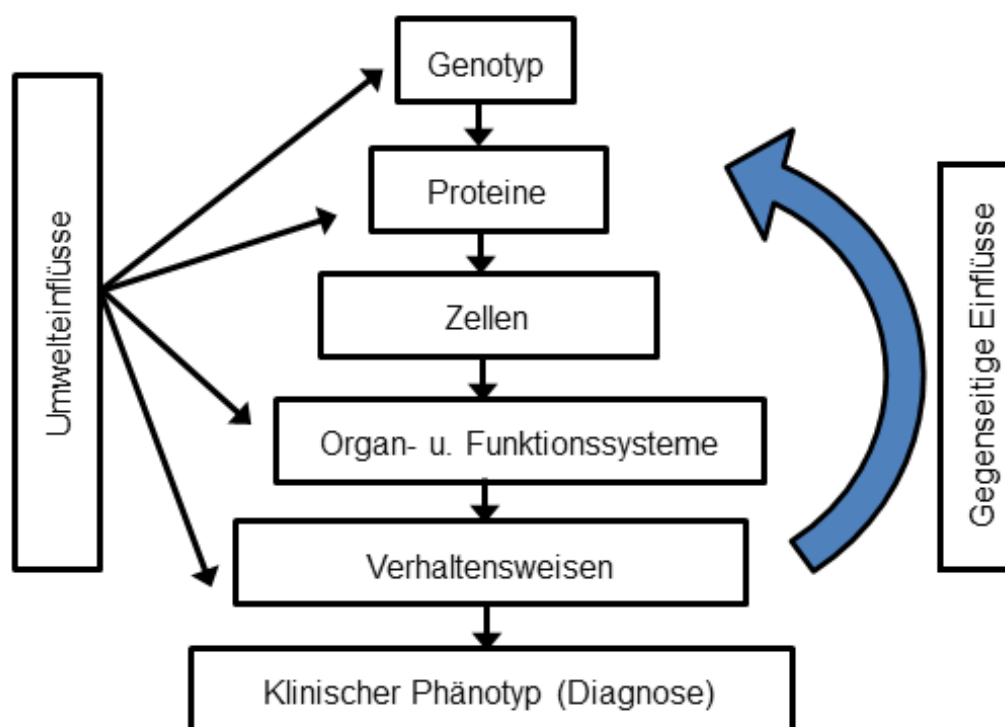


Abbildung 1. Endophänotypen. Hierarchisches Modell der Ebenen, auf denen sich pathologische Phänomene manifestieren können. Nach Chamberlain und Menzies (2009, S. 1134, übersetzt durch die Autorin).

Der neurobiologische Endophänotyp der Zwangsstörung. Überlegungen zur Neurobiologie der Zwangserkrankung entstanden aus klinischen Beobachtungen. So treten z. B. bei neurologischen Erkrankungen wie dem Gilles-de-la-Tourette-Syndrom und der Chorea Sydenham, die auf einer Schädigung der Basalganglien¹ beruhen, oft Zwangssymptome auf. Die Basalganglien sind über Regelkreise mit motorischen und präfrontalen Arealen verbunden und in zahlreiche motorische und kognitive Prozesse involviert. Zu diesen zählen z. B. zielgerichtete Bewegungen beim Greifen und das Speichern einer Information im Arbeitsgedächtnis (Girard, Tabareau, Pham, Berthoz & Slotine, 2008).

¹ Zu den Basalganglien gehören folgende subkortikale Strukturen: Der Nucleus caudatus, das Putamen und der Globus pallidus. Die Einheit von Nucleus caudatus und Putamen wird als Striatum bezeichnet (Weigel, Krauss & Jones, 2004).

Aus Beobachtungen, dass Menschen nach bestimmten Gehirnläsionen Zwangssymptome entwickelten und dass therapieresistente Zwangspatienten von neurochirurgischen Maßnahmen profitieren können, schloss man, dass neben den Basalganglien noch andere Gehirnstrukturen zur Ausbildung von Zwangssymptomen beitragen müssen. Hierzu zählt insbesondere der Frontalcortex (Mindus, Rasmussen & Lindquist, 1994). Die Bedeutung des orbitofrontalen Cortex (OFC) bei der Verhaltenssteuerung und –inhibition ist bereits seit 1848 bekannt, als der Bahnarbeiter Phineas Gage zu trauriger Berühmtheit gelangte (van Horn et al., 2012). Er erlitt einen schweren Unfall, bei dem eine Eisenstange seinen Schädel durchdrang und große Bereiche des Frontalhirns schädigte. Der zuvor gut angepasste, gewissenhafte Arbeiter überlebte den Unfall, zeigte danach aber auffällige Wesensänderungen mit eingeschränkter Planungsfähigkeit und gestörter Inhibition unangemessenen Verhaltens (Harlow, 1993). Zwar stellen die Frontalhirnverletzungen des Phineas Gage ein Extrembeispiel dar, das auf Zwangspatienten nicht direkt übertragen werden kann. Deren strukturelle und funktionelle Auffälligkeiten des Frontalhirns sind von wesentlich subtilerer Art. Gleichwohl folgen sie demselben Grundprinzip, dass Funktionen der Handlungsplanung und Verhaltensinhibition verändert sind.

Neben dem Frontalhirn und den Basalganglien spielt der Thalamus für die Zwangsstörung eine wichtige Rolle. Der Thalamus ist eine paarige, ca. drei Zentimeter lange eiförmige Struktur im Zwischenhirn, die in fast alle Teile des Cortex Projektionen aussendet. Als „Tor des Bewusstseins“ (Schandry, 2011, S. 129) sortiert der Thalamus alle ankommenden sensorischen Informationen (mit Ausnahme von Geruchseindrücken) vor und leitet sie in die verschiedenen Cortexregionen weiter. Der Thalamus ist an zahlreichen Vorgängen beteiligt, die Veränderungen in Bewusstseinszuständen, Schlaf und Wachheit, Aufmerksamkeit und Kognition zugrunde liegen (Jones, 2003).

Es lässt sich Folgendes zusammenfassen: Erstens haben Zwangspatienten Schwierigkeiten, Handlungen und Gedanken zu inhibieren. Zweitens treten bei Störungen der Basalganglien oft Zwangssymptome auf. Drittens gehen

Veränderungen des Frontalcortex mit eingeschränkter Verhaltensinhibition einher. Damit lag die Entwicklung eines neurobiologischen Modells der Zwangsstörung nahe. Dieses Modell der Zwangserkrankung (Saxena, Brody, Schwartz & Baxter, 1998) geht davon aus, dass Zwangssymptome mit gestörten neuronalen Regelkreisen des OFC zu den damit verbundener Regionen des Nucleus caudatus, des Thalamus und des anterioren Cingulums (ACC) zusammenhängen. Das Grundprinzip dieses Modells und seiner aktuellen Erweiterungen von Menzies et al. (2008) sowie von Milad und Rauch (2012) besteht darin, dass in mehreren Regelkreisen vom Frontalcortex neuronale Projektionen ausgehen in Richtung korrespondierender Zielgebiete im Striatum, von dort über direkte (exzitatorische) und indirekte (inhibitorische) Bahnen zum Thalamus, und vom Thalamus über rücklaufende Projektionen zurück in die ursprüngliche Region des Frontalhirns, wo jede Schleife begann (vgl. Abbildung 2). Bei Gesunden ist das Zusammenspiel von Aktivierung und Hemmung zwischen den frontalen Cortexgebieten, den Basalganglien und dem Thalamus weitgehend ausgeglichen, während Zwangspatienten eine Dysbalance zwischen direkten (exzitatorischen) und indirekten (inhibitorischen) Bahnen aufweisen. Bei Zwangspatienten führt ein erhöhter neuronaler Tonus im direkten Pfad zu einer dauerhaft erhöhten Aktivität in Bereichen des Frontalcortex, was die Manifestation von Zwangssymptomen begünstigen und aufrechterhalten kann (z. B. Baxter et al., 1996; Saxena & Rauch, 2000).

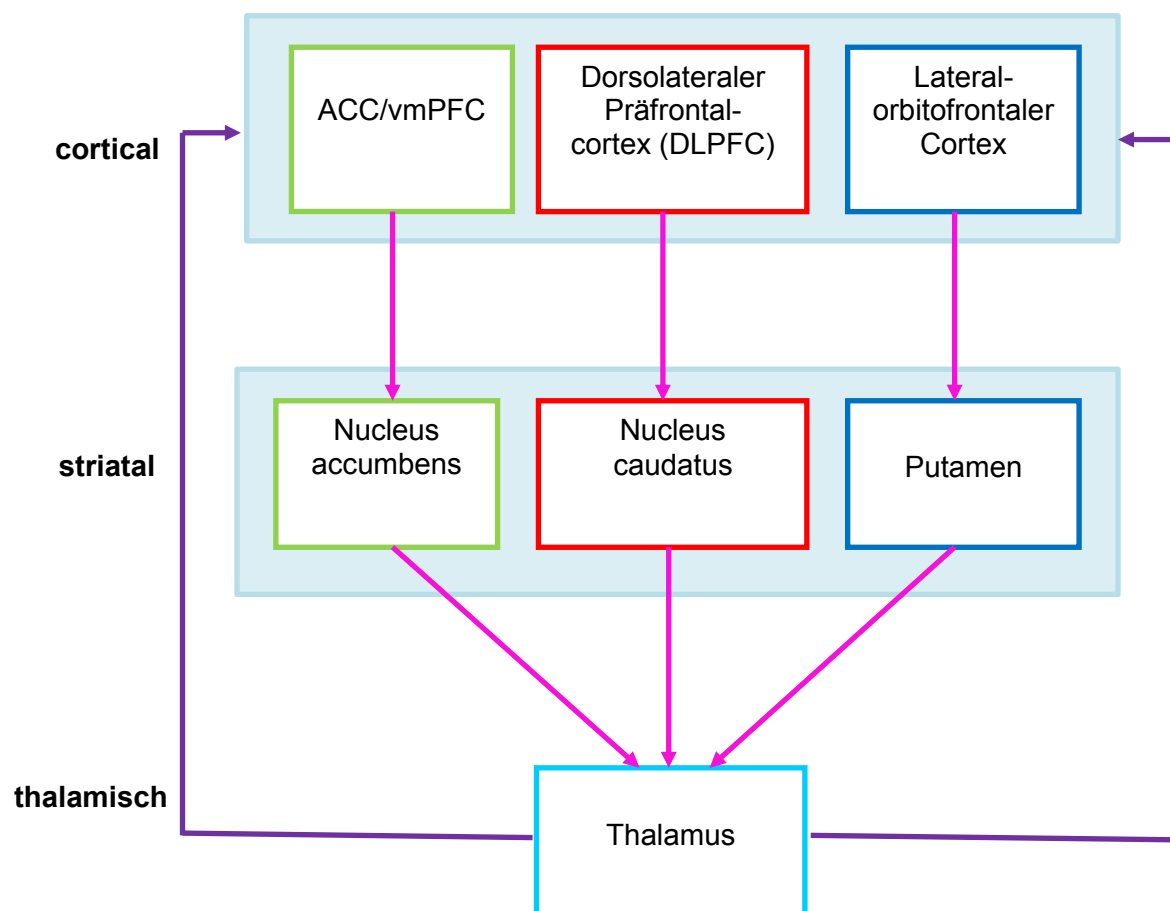


Abbildung 2. Cortico-striato-thalamisches Modell der Zwangsstörung. Nach Milad und Rauch (2012, S. 44). ACC = Anteriorer cingulärer Cortex; vmPFC = ventromedialer Präfrontalcortex.

In einer strukturellen Magnetresonanztomografie- (MRT-) Studie von Kang et al. (2004) fand sich im linken OFC von Zwangspatienten, verglichen mit Gesunden, eine signifikante Volumenminderung, die mit der Schwere der Zwangssymptomatik korrelierte: Je stärker die Zwangssymptome ausgeprägt waren, desto geringer war das Volumen in dieser Gehirnregion. Auch eine große multizentrische MRT-Untersuchung mit 412 Zwangspatienten und 368 gesunden Kontrollen zeigte, dass Zwangspatienten im Vergleich zu Gesunden signifikante Volumenminderungen im Präfrontalcortex und im ACC aufwiesen (de Wit et al., 2014). Strukturelle Befunde zum Volumen des Nucleus caudatus, der als Teil der Basalganglien an der Kontrolle von Willkürbewegungen beteiligt ist und in den

neurobiologischen Modellen der Zwangsstörung eine bedeutsame Rolle spielt, sind uneinheitlich (für eine Übersicht siehe Kathmann, 2008). Eine Metaanalyse zeigte bei Zwangspatienten ein reduziertes Gehirnvolumen im linken ACC, im linken und rechten OFC sowie erhöhte Volumina des linken und rechten Thalamus, die mit dem Schweregrad der Zwangssymptome signifikant korrelierten (Rotge et al., 2009).

Funktionelle Bildgebungsstudien bei Zwangspatienten liefern zunehmend Evidenz für veränderte Aktivierungsmuster der Bahnen zwischen dem OFC und den Basalganglien. In einer frühen Untersuchung wurde die Gehirnaktivität während der Provokation von Zwangssymptomen mittels Positronen-Emissions-Tomografie (PET) gemessen (Rauch, Jenike, Alpert & Baer, 1994): Betrachteten die Zwangspatienten Stimuli, die bei ihnen Zwangssymptome auslösten, dann war der regionale zerebrale Blutfluss der Patienten im rechten Nucleus caudatus, im linken ACC und beidseits im OFC signifikant erhöht. Hier ließ sich der im Modell angenommene Zusammenhang zwischen corticaler Aktivierung und Zwangssymptomatik direkt beobachten. Zwei Reviews erbrachten relativ konsistente Evidenz für eine erhöhte Aktivierung des Nucleus caudatus und des OFC bei Zwangspatienten (Menzies et al., 2008; Whiteside, Port & Abramowitz, 2004). Auch im Ruhezustand des Gehirns (*resting state*) wiesen Zwangspatienten im Vergleich zu gesunden Probanden eine signifikant erhöhte funktionelle Konnektivität zwischen dem Striatum und dem OFC auf, wie eine fMRI-Untersuchung von Harrison et al. (2009) ergab. Darüber hinaus ließ sich in der Studie von Harrison et al. (2009) anhand der Ausprägung der Konnektivität zwischen dem Nucleus caudatus und dem Nucleus accumbens einerseits und dem anterioren OFC andererseits der Schweregrad der Zwangssymptomatik vorhersagen. Eine weitere *resting-state* fMRI-Untersuchung (Kang et al., 2013) ergab, dass Zwangspatienten im Vergleich zu gesunden Probanden eine erhöhte funktionelle Konnektivität des Nucleus caudatus mit Teilen des cingulären Cortex und mit dem präzentralen Gyrus aufwiesen, was die Vermutung von Hyperaktivität in den striatal-corticalen Bahnen bestätigte. Untersuchungen, die Maße des lokalen Gehirnstoffwechsels von Zwangspatienten vor und nach einer Therapie erfassten, kamen ebenfalls zu interessanten Ergebnissen. So zeigte eine PET-Studie an 18 Zwangspatienten, dass sich die Raten des

Glukosestoffwechsels im Nucleus caudatus bei jenen Patienten reduzierten, die auf eine Behandlung mit SSRI oder Verhaltenstherapie gut ansprachen (Baxter et al., 1992). Eine fMRI-Studie belegte, dass die Hyperaktivität in dem Regelkreis, der mit der Zwangssymptomatik zusammenhängt und den OFC, den ACC und die Basalganglien beinhaltet, bei erfolgreicher Behandlung mit SSRI oder Verhaltenstherapie abnahm (Nakao et al., 2005). Konsistent hierzu zeigten Freyer et al. (2011) in einer weiteren fMRI-Studie anhand von zehn Zwangspatienten, dass sich deren infolge von Hyperaktivität pathologisch reduzierte Reaktionsfähigkeit des Nucleus caudatus nach erfolgreicher kognitiver Verhaltenstherapie verbesserte und dass sich die Aktivität im OFC normalisierte. Zusammenhänge zwischen veränderten Strukturen und gestörten Funktionen des Gehirns bei Zwangspatienten untersuchten Rotge et al. (2010), indem sie die Daten der oben genannten Metaanalyse zum Gehirnvolumen von Zwangspatienten (Rotge et al., 2009) mit den Daten einer Metaanalyse zu funktionellen Gehirnveränderungen bei Symptomprovokation (Rotge et al., 2008) kombinierten. Die Autoren fanden im linken lateralen OFC einen Überschneidungsbereich, der über mehrere Studien hinweg sowohl vergrößertes Volumen als auch auffällige Aktivierungsmuster aufwies. Sie schlossen daraus: “the lateral OFC could play a central role in cognitive flexibility deficits and, thus, in the genesis of pathological habits“ (Rotge et al., 2010, S. e37).

Der neuropsychologische Endophänotyp der Zwangsstörung. Die oben beschriebene Untersuchung neurobiologischer Endophänotypen der Zwangsstörung birgt Chancen, hat aber auch Nachteile. So ist die Bildgebungstechnik aufwändig und nicht überall verfügbar. Zudem lehnen viele Patienten wegen klaustrophober Ängste die Untersuchungen „in der Röhre“ ab. Eine Alternative stellen neuropsychologische Untersuchungen dar. Diese erfassen die Endophänotypen auf der Ebene beobachtbarer Verhaltensweisen. Das heißt, dass Funktionen des Gehirns anhand von psychologischen Testleistungen gemessen werden. Hiermit versucht man, „funktionelle Neurophänotypen“ (Schiepek et al., 2007, S. 386) zu bestimmen, die mit den klinischen Phänotypen korrespondieren. Vorteile neuropsychologischer Testungen liegen in ihrer unkomplizierten Anwendbarkeit und Verfügbarkeit. Es

sind lediglich ein ruhiger Raum und ein Computer bzw. Papier und Bleistift erforderlich.

Hinweise auf neuropsychologische Fehlfunktionen bei Zwangspatienten fanden sich vor allem im Zusammenhang mit veränderten Strukturen und Funktionen des Frontalhirns und der Basalganglien (Saxena et al., 1998). Man vermutet Zusammenhänge zwischen der klinischen Symptomatik (repetitive Zwangsgedanken und ritualisierte Handlungen), einer gestörten gedanklichen und verhaltensbezogenen Kontroll- und Inhibitionsfähigkeit sowie eingeschränkten Leistungen in Tests zur Inhibition, zum Konzeptwechsel und zum Einsatz von Organisationsstrategien beim Enkodieren von Informationen. Diese Zusammenhänge haben Chamberlain und Kollegen (2005) in einem Modell der Zwangsstörung (Abbildung 3) dargestellt. Das Modell lässt sich als Teil des umfassenderen Endophänotypen-Konzeptes (vgl. Abbildung 1) verstehen. Chamberlain et al. (2005) beschrieben zum einen die phänotypische Ebene des beobachtbaren Zwangsverhaltens (*syndromic behaviour*), zum anderen die endophänotypische Ebene kognitiver und verhaltensmäßiger Inhibition (*inhibitory failures*) und deren Operationalisierung anhand von neuropsychologischen Tests (*cognitive findings*):

Failures in cognitive and behavioural inhibition appear to account for many of the cognitive deficits and clinical symptoms seen in OCD. Abnormalities in neural circuitry connecting cortical to sub-cortical structures, particularly the lateral orbitofrontal loop, are likely to represent the neurobiological basis of these inhibitory failures. Such circuits are implicated in physiological inhibitory processes and are frequently found to function abnormally in OCD neuroimaging studies. (S. 412)

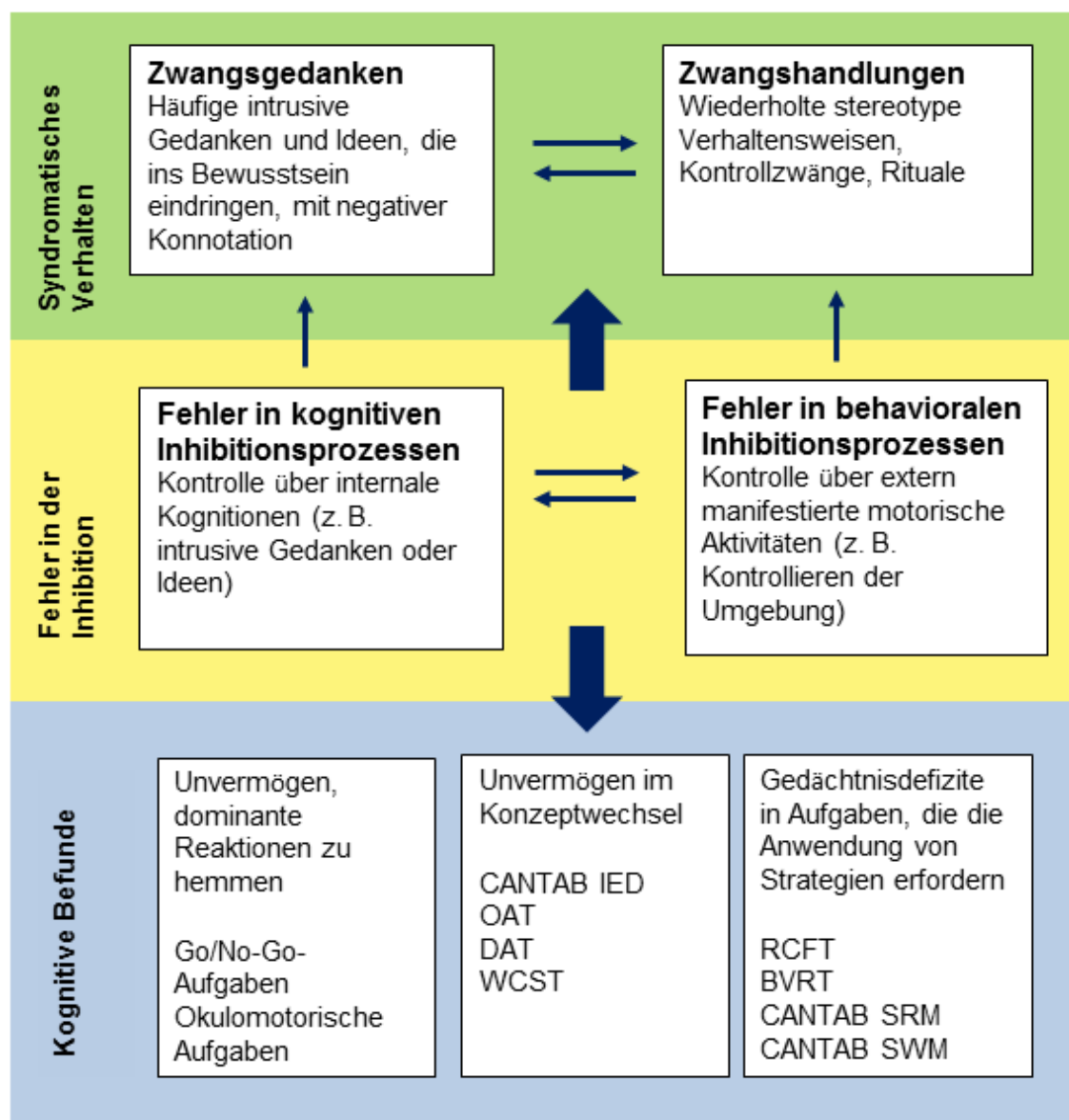


Abbildung 3. Modell zur Bedeutung gestörter Inhibitionsprozesse der Zwangsstörung. Nach Chamberlain et al. (2005, S. 412), übersetzt durch die Autorin. BVRT = Revised Visual Retention Test (Benton, 1974, zitiert nach Chamberlain et al., 2005); CANTAB = Cambridge Automated Neuropsychological Test Battery (Cambridge-Cognition, 2013; DAT = Delayed Alternation Test (Freedman & Oscar-Berman, 1986, zitiert nach Chamberlain et al., 2005); IED = Intra-Extra-Dimensional Shift; RCFT = Rey-Osterrieth Complex Figure Test (Osterrieth, 1944); WCST (Wisconsin Card Sorting Test); SRM = Spatial Recognition Memory; SWM = Spatial Working Memory.

In den Studien zu kognitiven Defiziten, die bei Zwangspatienten auftreten und mit ihren Symptomen in Verbindung gebracht werden, sind zahlreiche neuropsychologische Testverfahren eingesetzt worden. Diese Tests erfassen Leistungen in den Domänen Aufmerksamkeit, Exekutivfunktionen, Gedächtnis, Arbeitsgedächtnis, Verarbeitungsgeschwindigkeit, Entscheidungsverhalten und Kategorisierung. Die folgenden Abschnitte (1.2 bis 1.6) stellen die Auswahl jener Testverfahren vor, die bei der Sichtung der Primärstudien gefunden wurden und später im empirischen Teil der Arbeit metaanalytisch ausgewertet werden.

1.2 Aufmerksamkeit

Die menschliche Kapazität zur Informationsverarbeitung ist begrenzt, so dass nicht alle Reize aus der Umgebung ins Bewusstsein dringen können. Daher muss bei allen Prozessen der Informationsaufnahme eine Auswahl getroffen werden. Diese Auswahl geschieht vor- oder unbewusst, erfolgt in verschiedenen Sinnesmodalitäten (z. B. visuell oder auditiv) und lässt sich als Aufmerksamkeit bezeichnen. Eine aktuelle Definition beschreibt Aufmerksamkeit als Fähigkeit, sich auf eine spezifische Sache zu fokussieren, ohne sich ablenken zu lassen: “Attention refers to being able to focus on a specific thing without being distracted“ (Gillig & Sanders, 2011, S. 43). Aufmerksamkeit dient nicht nur der Auswahl von Informationen, um die Wahrnehmung zu steuern, sondern sie hat auch eine handlungsvermittelnde Funktion. Im Zusammenhang mit Zwangsstörungen sind zwei Aspekte der Aufmerksamkeit von besonderer Bedeutung: erstens die Fähigkeit, die Aufmerksamkeit auf einem Stimulus zu halten (anhaltende Aufmerksamkeit) und zweitens die Fähigkeit, die Aufmerksamkeit von einem Stimulus zu lösen (*disengagement of attention*, Eysenck & Keyne, 2005, S. 160). Während letztere inhibitorische Fähigkeiten erfordert und später im Abschnitt 1.3 bei den Exekutivfunktionen behandelt wird, kann erstere z. B. anhand folgender Testverfahren gemessen werden: Der computerbasierte Situation Awareness Test (SAT; Mueller, 2010; zitiert nach Tumkaya et al., 2013) verlangt, dass der Proband sich auf dem Bildschirm bewegende Objekte (Insekten und Eidechsen) im Blick behält und deren Identität, Position und Richtung angibt. Hingegen werden beim K–

T-Durchstreichtest (Halter, 1986; zitiert nach Bouvard, Dirson & Cottraux, 1997), einem Papier-und-Bleistift-Verfahren, auf einem zweiseitigen Blatt jeweils 20 Zeilen mit je 17 Zeichen (Buchstaben, Zahlen, Symbole) dargeboten. Der Proband soll in der linken Spalte alle Zeichen auskreuzen, die nicht mit den Zeichen der rechten Spalte übereinstimmen. Dagegen dient der Subtest Mentale Kontrolle der revidierten Wechsler Memory Scale (WMS-R; Wechsler, 1987; zitiert nach Nakao et al., 2009) zur mündlichen Erfassung anhaltender Aufmerksamkeit anhand von überlernten Zahlen- oder Buchstabenreihen, indem der Proband von 20 rückwärts zählt oder das Alphabet rückwärts spricht. Gemessen werden jeweils Reaktions- bzw. Bearbeitungszeit und Fehlerzahl.

1.3 Exekutivfunktionen

Exekutivfunktionen sind kognitive Funktionen, die eingesetzt werden, wenn Routinehandlungen nicht ausreichen und eine kontrollierte Handlungssteuerung verlangt wird (Kaiser, Mundt & Weisbrod, 2005). Dies ist nach Norman und Shallice (1986) der Fall bei Planung und Entscheidungsfindung, Fehlerkorrektur, bei neuen Handlungssequenzen, bei gefährlichen oder technisch schwierigen Abläufen sowie bei der Inhibition einer dominanten Antworttendenz. Exekutivfunktionen stehen sehr eng mit der Neurobiologie des Frontalcortex in Verbindung (Mega & Cummings, 1994). Synonym verwenden Kammer und Karnath (2006, S. 500) den Begriff „Frontalhirnfunktionen“ und bezeichnen diese als „die höchsten kognitiven Leistungen des Menschen“. Massiv gestörte Frontalhirnfunktionen infolge einer schweren Schädelverletzung wurden bereits anhand des Bahnarbeiters Phineas Gage illustriert. Bei Zwangspatienten nimmt man aufgrund der neurobiologischen Befunde im Bereich frontaler Cortexgebiete an, dass die Betroffenen Aufgaben zu Exekutivfunktionen weniger gut lösen als Gesunde (eine Übersicht geben Olley et al., 2007; Savage et al., 1999). Im Folgenden werden die für das Verständnis der Zwangsstörung wichtigsten Exekutivfunktionen beschrieben: Planen/Problemlösen, kognitive und behaviorale Inhibition, Konzeptwechsel/Regellernen und Enkodieren.

Planen/Problemlösen. Im Zusammenhang mit der Fähigkeit, das eigene Verhalten zu organisieren, wird Planen/Problemlösen mit Funktionen des

Frontalhirns, genauer des DLPFC, in Verbindung gebracht (Mega & Cummings, 1994). Zur Untersuchung wurde z. B. das Scheiben-Transfer-Paradigma eingesetzt, das je nach Autor als Tower of Hanoi (TOH; Simon, 1975; zitiert nach Kuelz et al., 2004) und Tower of London (TOL; Shallice, 1982; zitiert nach Tolin et al., 2011) bezeichnet wird. Der TOL wurde entwickelt, um Beeinträchtigungen der Planungsfähigkeit bei Patienten mit Frontalhirnläsionen zu untersuchen. Hierbei soll der Proband drei Kugeln, die auf drei Stäben verschiedener Länge sitzen, von einer definierten Anfangsposition in eine definierte Zielposition bringen. Er darf jeweils nur eine Kugel bewegen und soll mit möglichst wenigen Bewegungen die Zielposition erreichen (vgl. Abbildung 4). Die Stockings of Cambridge Task (SOC; Cambridge-Cognition, 2013) stellt eine Computerversion des TOL dar, bei der die zu versetzenden Kugeln nicht auf Stäben aufgereiht sind, sondern in stilisierten Socken an einer Wäscheleine hängen. Alle Scheiben-/Kugel-Transfer-Aufgaben erfordern die Fähigkeit, in die Zukunft voranzuplanen (Owen, 2005). Diese Fähigkeit ist bei Zwangspatienten möglicherweise beeinträchtigt (Krishna et al., 2011).

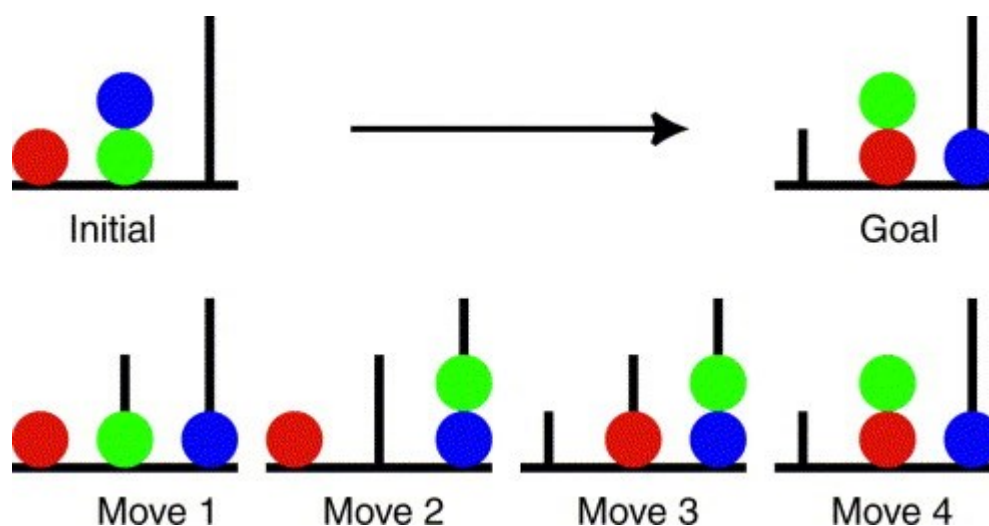


Abbildung 4. Tower of London (Polk, Simen, Lewis & Freedman, 2002, S. 73).

Kognitive und behaviorale Inhibition. Klinisch lässt sich beobachten, dass Zwangspatienten Schwierigkeiten haben, sich aufdrängende Zwangsgedanken und

die Ausführung von Zwangshandlungen zu unterdrücken. Anhand von Tiermodellen und Läsionsstudien wurde belegt, dass ein selektives Defizit in der Fähigkeit, Reaktionen auf irrelevante Stimuli zu inhibieren, mit Störungen im orbitalen Präfrontalkortex einhergeht (Rosenberg, Dick, O'Hearn & Sweeney, 1997). Die Fähigkeit zur Inhibition lässt sich nach Harnishfeger (1995, S. 184) unterscheiden in behaviorale Inhibition: „behavioral inhibition involves the (potentially intentional) control of overt behavior, such as resisting temptation, delay of gratification, motor inhibition, and impulse control” und kognitive Inhibition: “cognitive inhibition involves the control of cognitive contents of processes, and can be intentional and conscious, or unintentional and unavailable for conscious introspection” (ebenda, S. 184). Kognitive Inhibition wird mit der Color Word Stroop Task (Golden, 1978; zitiert nach Hashimoto et al., 2011) erfasst. Hierbei werden in verschiedener Schriftfarbe Wörter präsentiert, die ihrerseits andere Farben benennen. Der Proband soll die gedruckte Schriftfarbe benennen und den Impuls unterdrücken, das Wort vorzulesen (vgl. Abbildung 5). Da Lesen im Gegensatz zur Benennung von Farben hoch automatisiert ist, entsteht eine kognitive Interferenz, der Stroop-Effekt, der an einer verzögerten Reaktionszeit erkennbar ist. Die deutsche Version der Stroop-Aufgabe ist als Farbe-Wort-Interferenztest (FWIT; Bäuml, 1985; zitiert nach Exner, Martin & Rief, 2009) bekannt.

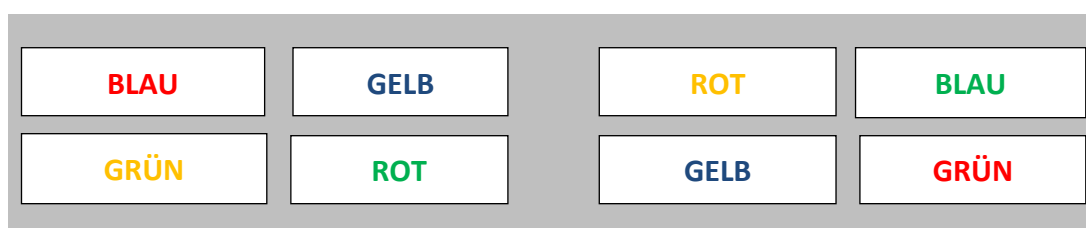


Abbildung 5. Farbe-Wort-Stroop-Aufgabe. Nach Golden (1978, zitiert nach Hashimoto et al., 2011). Der Proband soll jeweils die Textfarbe benennen und das hoch automatisierte Lesen des Farbwortes unterdrücken. Richtig wäre z. B. im ersten Feld, „rot“ zu sagen und nicht „blau“ zu lesen.

Behaviorale Inhibition bzw. *response inhibition* wird z. B. mit dem Continuous Performance Test–II (CPT–II; Conners, 2000; zitiert nach Grisham, Brown, Savage, Steketee & Barlow, 2007) erfasst, bei dem der Proband bei jedem Buchstaben, der auf dem Bildschirm erscheint und *kein X* ist, möglichst schnell eine Taste drücken soll. Ähnlich wie der CPT–II erscheint die Stop Signal Reaction Time Task (SSRTT; Logan, 1994; zitiert nach Morein-Zamir et al., 2014). Hierbei soll der Proband jedes Mal, wenn ein X oder O auf dem Bildschirm erscheint, die X- oder O-Taste auf der Tastatur drücken. Wenn jedoch auf das X oder O auf dem Bildschirm ein Dollarzeichen (Stop-Stimulus) folgt, soll der Proband den entsprechenden Tastendruck unterlassen. *Response inhibition* wird z. B. auch mit der Go/No-Go-Aufgabe (Bannon, Gonsalvez, Croft & Boyce, 2002) erfasst. Hierbei werden dem Probanden am Bildschirm mehrmals in zufälliger Abfolge zwei Stimuli dargeboten, und er soll auf einen der beiden mit Tastendruck reagieren. In späteren Durchgängen soll er die zuvor geübte Reaktion unterdrücken und auf den anderen Stimulus mit Tastendruck reagieren (Grisham, Norberg, Williams, Certoma & Kadib, 2010; Omori et al., 2007). In der Simon Spatial Incompatibility Task (Simon, 1969) und deren Computerversion (Marsh et al., 2014) wird ebenfalls die inhibitorische Kontrolle erfasst. Der Proband soll möglichst schnell mit verschiedenen Fingern eine Taste drücken, wenn ein nach rechts oder links zeigender Pfeil auftaucht. Dabei soll der Proband ignorieren, ob der Pfeil zuvor auf der rechten oder linken Bildschirmhälfte erschien. Gemessen werden jeweils Reaktionszeit, Auslassungsfehler (*omission errors*) und „zusätzliche Fehler“ (*commission errors*) wie z. B. ein zusätzlicher Tastendruck.

Konzeptwechsel und Regellernen. Zwangssymptome stehen auch mit der Funktion des Konzeptwechsels in Zusammenhang (Fontenelle, Marques, Engelhardt & Versiani, 2001; Harvey, 1986). Konzeptwechsel bezeichnet die Fähigkeit, den Aufmerksamkeitsfokus zwischen verschiedenen Merkmalen eines Stimulus zu wechseln, und zwar als Reaktion auf wechselnde Rückmeldungen. Der Wisconsin Card Sorting Test (WCST; Grant & Berg, 1948) stellt hier das bekannteste Verfahren dar. Die Psychologin Brenda Milner (1963) setzte den WCST gezielt bei Patienten mit Frontalhirnläsionen ein. Zur Durchführung erhält der Proband einen Stapel von

64 Karten mit Symbolen, die sich in Anzahl, Farbe und Form unterscheiden. Der Versuchsleiter deckt zu Beginn vier Karten auf. Der Proband zieht eine Karte, legt sie zu einer der vier offenen hinzu und erhält vom Versuchsleiter die Antwort „richtig“ oder „falsch“. Im folgenden Beispiel (Abbildung 6) kann die gezogene Karte (ein grüner Kreis) nach der Sortierregel „gleiche Anzahl“ dem roten Dreieck, nach der Regel „gleiche Farbe“ den zwei grünen Sternen und nach der Regel „gleiche Form“ den vier blauen Kreisen zugeordnet werden. Der Proband zieht die nächste Karte, legt sie an, erhält wieder eine Rückmeldung und ordnet so nach und nach die gezogenen Karten den liegenden Karten zu. Wie die Sortierregel lautet, soll der Proband anhand der Rückmeldungen des Versuchsleiters herausfinden. Nach einer bestimmten Zeit wechselt die Sortierregel, d. h., der Proband erhält unerwartet die Rückmeldung „falsch“ auf seine bisher richtigen Zuordnungen. Seine Aufgabe besteht nun darin, die bisherige Sortierregel aufzugeben und eine neue Regel zu lernen. Menschen mit Frontalhirndefiziten haben hierbei Schwierigkeiten, sie beharren oft auf der zuerst gelernten Regel (Robinson, Heaton, Lehman & Stilson, 1980). Der WCST existiert auch als Computerversion, Intra-Extra-Dimensional Shift (IED) der Cambridge Automated Neuropsychological Test Battery (CANTAB; Cambridge-Cognition, 2013). Gemessen werden die Gesamtzahl der Antworten, die Anzahl der gebildeten Kategorien und die Fehlerzahl.

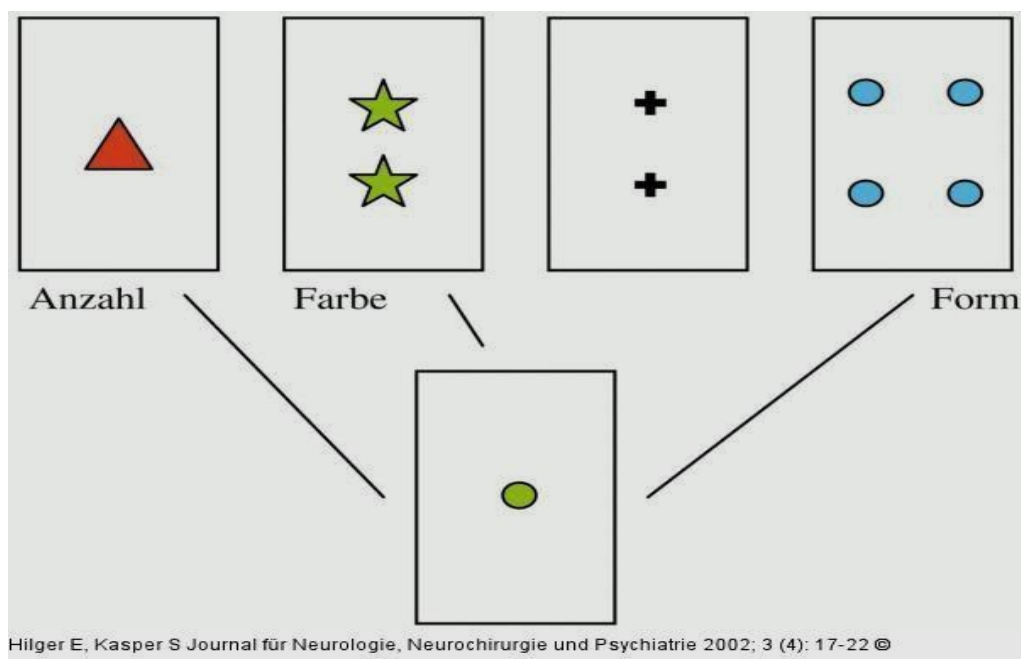


Abbildung 6. Wisconsin Card Sorting Test (WCST; Hilger & Kasper, 2002, S. 20).

Darüber hinaus wird als einfaches Papier- und Bleistift-Verfahren zum Konzeptwechsel der Trail-Making-Test, Teil B (TMT-B; Reitan, 1958) eingesetzt. Bei diesem soll der Proband alternierend Zahlen und Buchstaben in aufsteigender Reihenfolge durch Linien verbinden (1 – A – 2 – B – 3 – C etc.) und dabei geistig zwischen zwei Regeln (aufsteigende Zahlenfolge, Alphabet) hin- und herspringen. Gemessen wird die Reaktionszeit.

Enkodieren. Wenn Information gelernt, d. h. im Gedächtnis eingespeichert werden soll, dann gelingt dies umso leichter und nachhaltiger, je besser die Information organisiert ist (Bower, 1970). Eine wichtige Organisationsstrategie besteht darin, eine Struktur im zu enkodierenden Material zu finden oder selbst zu erzeugen. Beispielsweise prägt sich eine Ziffernfolge in zufälliger Form „24122014“ weniger gut ein als in gruppierter Form „24. 12. 2014“. Eine weitere Strategie besteht darin, diese Materialstruktur mit bereits vorhandener Information im Langzeitgedächtnis zusammenzuführen. Im Beispiel ergäbe dies „Heiligabend 2014“. Die Fähigkeit, Organisationsstrategien beim Enkodieren anzuwenden, wird z. B. anhand des Organisationsscores des Rey-Osterrieth Complex Figure Test

(RCFT; Osterrieth, 1944; zitiert nach Watanabe et al., 2005; vgl. Abbildung 8) geprüft. Hierbei wird erfasst, in welchem Maße der Proband beim Abzeichnen einer komplexen grafischen Figur deren einzelne Elemente gruppiert oder übergeordnete Strukturen findet. Auch bei einer Aufgabe zum visuell-räumlichen Arbeitsgedächtnis, der CANTAB SWM (*spatial working memory*; Cambridge Cognition, 1995), wird ein Strategiescore erhoben, der angibt, in wieweit der Proband eine systematische Strategie einsetzt, um über mehrere Durchgänge hinweg auf dem Bildschirm unter mehreren Quadraten dasjenige herauszufinden, hinter dem ein blaues Quadrat verborgen ist.

1.4 Gedächtnis und Arbeitsgedächtnis

Aus der klinischen Beobachtung von Patienten mit Hirnschädigungen ist bekannt, dass abgrenzbare Gedächtnisfunktionen ausfallen können, während andere Funktionen vollständig erhalten bleiben. So berichtete der Neurologe Sacks (1988) von einem Marinesoldaten, Jimmy G., der sämtliche biographischen Erinnerungen seiner letzten 30 Lebensjahre vergessen hatte. Dennoch waren Jimmy G.s kurzzeitige Behaltensleistungen unauffällig. Auch gelang es dem Patienten noch mühelos, Fertigkeiten wie Maschineschreiben und Morsen auszuführen, die er 30 Jahre zuvor erlernt hatte. Fälle, in denen Patienten mit Gehirnschädigungen deutliche Störungen im verbalen Kurzzeitgedächtnis zeigten, wohingegen ihr Langzeitgedächtnis und ihre allgemeine Intelligenz völlig intakt erschienen, wurden in der Literatur ebenfalls beschrieben (Buchner, 2006). Das Phänomen selektiver Gedächtnisausfälle führte zur Entwicklung eines „modalen Gedächtnismodells“ (Buchner & Brandt, 2008, S. 430). Demnach lässt sich das menschliche Gedächtnis grob einteilen in drei Modi: das sensorische Gedächtnis, das Kurzzeitgedächtnis und das Langzeitgedächtnis. Dieses Modell wurde inzwischen zweimal wesentlich erweitert (Baddeley, 1989; Tulving, 1993), ist aber in seinem Grundprinzip nach wie vor gültig.

Das *sensorische* Gedächtnis verarbeitet Informationen, bevor sie die bewusste Aufmerksamkeitsschwelle überschreiten. Es bildet die „Brücke zwischen der Wahrnehmung und dem, was eher einer konventionellen Vorstellung von Gedächtnis entspricht“ (Buchner & Brandt, 2008, S. 430). Nach wenigen hundert Millisekunden

„ist der größte Teil der sensorisch repräsentierten Information verloren und kann nicht mehr berichtet werden“ (Buchner, 2006, S. 447). Da das sensorische Gedächtnis aufgrund seiner extrem schnellen Zerfallszeit auch der Wahrnehmung zugeordnet werden kann, wird es in der vorliegenden Arbeit nicht berücksichtigt.

Den Begriff des *Kurzzeitgedächtnisses* erweiterte Baddeley (1989) zum Modell des Arbeitsgedächtnisses. Das Arbeitsgedächtnis enthält eine ausführende Instanz, die zentrale Exekutive. Diese steuert zwei untergeordnete Systeme, den visuell-räumlichen Skizzenblock und die phonologische Schleife.

Das *Langzeitgedächtnis* wurde von Tulving (1993) in drei Komponenten unterteilt, das episodische, semantische und prozedurale Gedächtnis. Das episodische Gedächtnis bezieht sich auf persönliches Erleben in der Vergangenheit, das in Form von erzählbaren, bildhaften oder handlungsbezogenen Episoden abgespeichert wird. Das semantische Gedächtnis hingegen beinhaltet Fakten, die von ihrem ursprünglichen Lernkontext gelöst sind. Das prozedurale Gedächtnis enthält automatisierte Handlungen, die der Introspektion kaum noch zugänglich sind. Das semantische und das prozedurale Gedächtnis seien hier der Vollständigkeit halber erwähnt, werden aber im weiteren Verlauf nicht mehr berücksichtigt, da hierzu bislang nicht genügend Studien zu Zwangspatienten mit bestimmten Symptomschwerpunkten vorliegen.

Ein Ansatz von Bredenkamp (1998) integriert das sensorische Ultrakurzzeit-, das Arbeits- und das episodische, semantische und prozedurale Langzeitgedächtnis. Bredenkamps Modell (vgl. Abbildung 7) berücksichtigt komplexere Interaktionen zwischen den einzelnen Bestandteilen. Es erlaubt auch schnelle Rückgriffe aus dem sensorischen Gedächtnis auf die Komponenten des Langzeitgedächtnisses, um Reize zu kategorisieren. Umgekehrt können auch „beiläufige“ Informationen direkt vom sensorischen Gedächtnis ins episodische Langzeitgedächtnis gelangen, ohne einen Zwischenschritt über das Arbeitsgedächtnis zu machen.

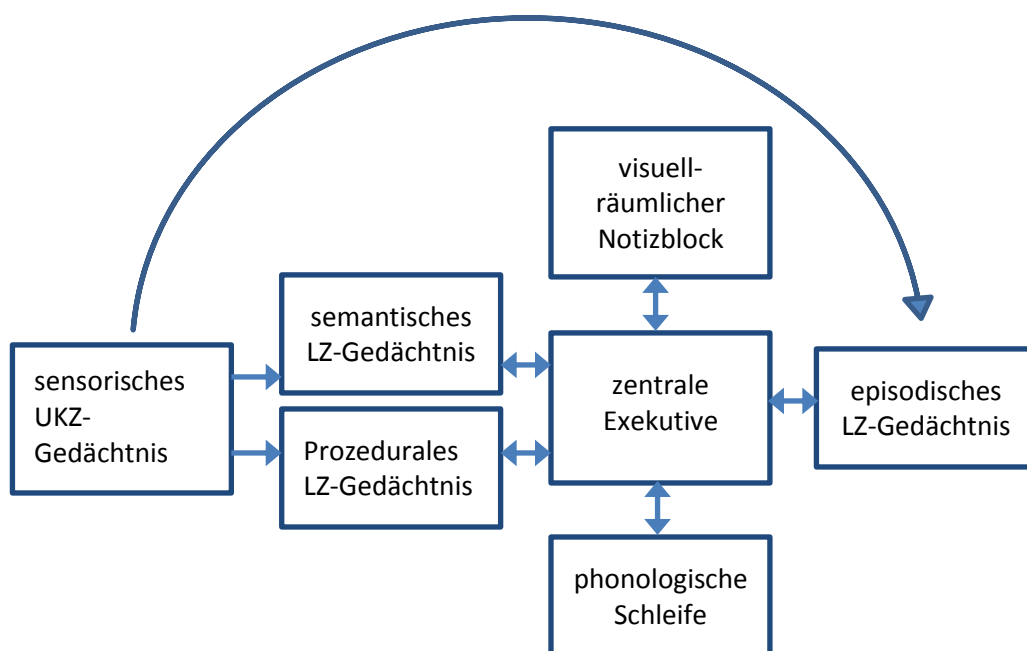


Abbildung 7. Gedächtnismodell. Nach Bredenkamp (1998, S. 63). UKZ = Ultra-kurzzeit, LZ = Langzeit.

Gedächtnisleistungen im engeren Sinne betreffen im Folgenden das episodische Langzeitgedächtnis und das Arbeitsgedächtnis. Für ein genaueres Verständnis der Zwangsstörung sind vor allem das verbale und nonverbale episodische Gedächtnis sowie das verbale und das visuell-räumliche Arbeitsgedächtnis von Bedeutung.

Das episodische Gedächtnis lässt sich in Behaltensleistungen nonverbal-bildhafter und verbaler Information unterscheiden. Episodische Gedächtnisleistungen werden mittels zweier Aufgabentypen erfasst: Während *recall*-Aufgaben ein freies Reproduzieren von Gedächtnisinhalten erfordern, überprüfen *recognition*-Aufgaben das Wiedererkennen eines zuvor präsentierten Stimulus unter anderen Stimuli (vgl. die Übersichtsarbeit von Muller & Roberts, 2005).

Episodisch-verbale Gedächtnisleistungen werden erfasst, indem man den Probanden z. B. Wortlisten vorliest. Nach einer Distraktionsaufgabe sollen sie die

Wörter aus dem Gedächtnis aufschreiben (*recall*) oder aus einer Wortliste mit Distraktoren wiedererkennen (*recognition*). Hierzu werden z. B. der verbale Subtest der Batterie d'Efficiency Mnésique (BEM 144; Signoret, 1991; zitiert nach Bouvard et al., 1997), der California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan & Ober, 1987; zitiert nach Ceschi, van der Linden, Dunker, Perroud & Brédard, 2003) und der Hopkins Verbal Learning Test (HVLT; Brandt, 1991; zitiert nach Cha et al., 2008) eingesetzt. In der revidierten Wechsler Memory Scale (WMS-R, Subtests Logical Memory LM-I und -II; Wechsler, 1987; zitiert nach Nakao et al., 2009) und im Lern- und Gedächtnistest-3 (LGT-3; Bäumlner, 1974; zitiert nach Zitterl et al., 2001) wird der Abruf episodischer Gedächtnisinhalte über das mündliche Nacherzählen vorgelesener kurzer Geschichten geprüft. Auch Paar-Assoziationslernen – das klassische Vokalbellernen – wird zur Prüfung des verbalen episodischen Gedächtnisses eingesetzt (Tuna, Tekcan & Topçuoğlu, 2005), indem man die Probanden Wortpaare lernen lässt, ihnen anschließend je ein Wort vorlegt und sie bittet, das passende Wort zu ergänzen (*recall*) oder es zwischen Distraktorwörtern wiederzuerkennen (*recognition*).

Das nonverbale episodische Gedächtnis bezieht sich auf Behaltensleistungen und auf Strategien beim Enkodieren visuell-räumlicher Stimuli, die über das Abzeichnen und freie Reproduzieren geometrischer Figuren erfasst werden. Dies erfolgt z. B. über den RCFT (Osterrieth, 1944; zitiert nach Watanabe et al., 2005; vgl. Abbildung 8), die Subtests *Visuelle Reproduktion I* und *II* der WMS-R (Wechsler, 1987; zitiert nach Nakao et al., 2009) und den Subtest *Figure Recall* der Adult Memory and Information Processing Battery (AMIPB; Coughlan & Hollows, 1985; zitiert nach Tallis, Pratt & Jamani, 1999). *Recognition* wird geprüft, indem die Probanden im Recurring Figures Test (Kimura, 1963) geometrische Figuren unter Distraktoren wiedererkennen sollen. In der Delayed-Matching-to-Sample Task (Clair et al., 2013) sollen die Probanden am Bildschirm gezeigte Stimuli wiedererkennen. Auch die nonverbalen Subtests der BEM 144 (Signoret, 1991; zitiert nach Bouvard et al., 1997) und des Lern- und Gedächtnistest-3 (LGT-3; Bäumlner, 1974; zitiert nach Zitterl et al., 2001) messen nonverbale Gedächtnisleistungen.

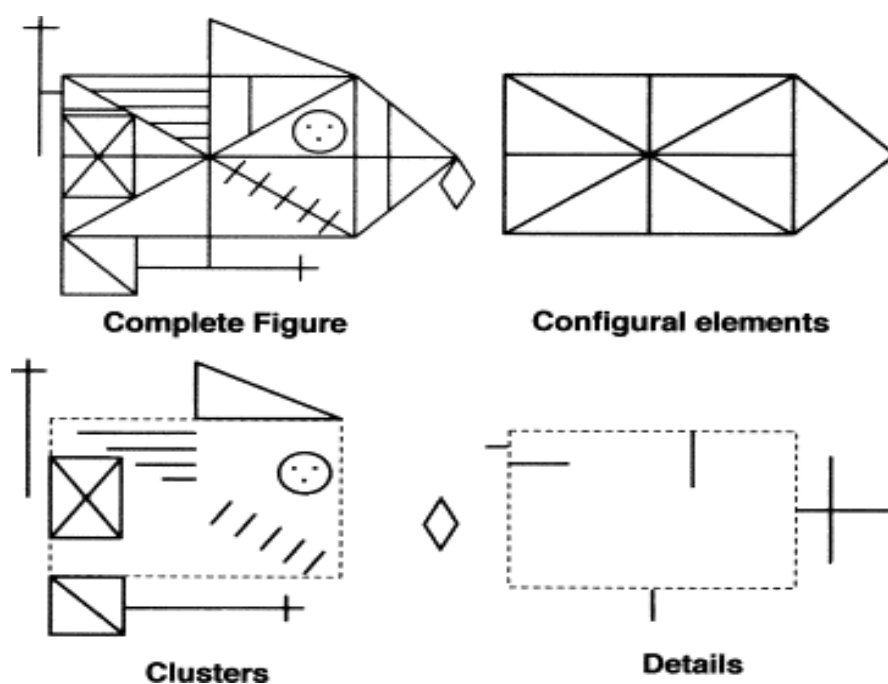


Abbildung 8. Rey-Osterrieth Complex Figure Test mit Organisationsmerkmalen der Enkodierung: Gestaltbildende Elemente (“configural elements”), Cluster und Details. (Watanabe et al., 2005, S. 565)

Leistungen des *visuell-räumlichen Arbeitsgedächtnisses* werden z. B. mit den Blockspanne-Subtests der Wechsler Adult Intelligence Scale–Revised (WAIS–R; Wechsler, 1981; zitiert nach Hashimoto et al. 2011) und der WMS–R (Wechsler, 1987; zitiert nach Exner et al., 2009) sowie der CANTAB (Cambridge-Cognition, 2013) erfasst, bei denen der Proband Würfel oder Holzblöcke in genau der Reihenfolge antippen soll, die der Versuchsleiter zuvor demonstriert hat. Auch Papier-und-Bleistift-Tests wie der Teil A des Trail Making Tests (TMT-A; Reitan, 1958), der Zahlen-Symbol-Test der WAIS–R (Wechsler, 1981; zitiert nach Hashimoto et al., 2011), der Subtest *Letter Number Sequencing* der Wechsler Adult Intelligence Scale–IV (WAIS–IV; Wechsler, 2008; zitiert nach Ayers, Saxena, Golshan & Wetherell, 2013) und die *Visual Memory Scale* (VMS) der WMS–R (Wechsler, 1987; zitiert nach Grisham et al., 2007) messen Aspekte des visuell-räumlichen Arbeitsgedächtnisses. Am Computer werden diese Leistungen erfasst, indem Stimuli (Punkte, Buchstaben etc.) in einem Raster präsentiert werden, deren Position der Proband erinnern soll (Kraft, 2006; Nakao, Nakagawa, Nakatani et al.,

2009; Rampacher, 2006). Hingegen erfordert die N-back-Aufgabe (Nakao, Nakagawa, Yoshiura et al., 2009), dass der Proband die Position von Punkten auf dem Bildschirm erinnert, die vor mehreren Durchgängen gezeigt wurden („N“ entspricht der Anzahl der Durchgänge). Auch werden Leistungen der mentalen Rotation erfasst, wenn die Probanden beim Cube Test Lücken in der räumlichen Abfolge geometrischer Figuren ergänzen sollen, während sie im Card Position Test eine Abfolge mehrerer Karten ergänzen sollen (Singh, Mukundan & Khanna, 2003).

Die *phonologische Schleife* des Arbeitsgedächtnisses speichert Informationen nur für kurze Zeitspannen (Buchner & Brandt, 2008). Leistungen des verbalen Arbeitsgedächtnisses werden über Zahlenspanne-Rückwärts-Aufgaben der WMS-R (Wechsler, 1987; zitiert nach Exner et al., 2009) und der WAIS-IV (Wechsler, 2008; zitiert nach Singh, 2003) erfasst, bei denen der Versuchsleiter Zahlenfolgen in ansteigender Länge vorspricht, die der Proband in umgekehrter Reihenfolge nachsprechen soll.

1.5 Informationsverarbeitungsgeschwindigkeit

Die Geschwindigkeit der Informationsverarbeitung (*processing speed*) umfasst sowohl perzeptuelle als auch motorische Geschwindigkeit. *Processing speed* stellt eine Funktion dar, die mit Leistungen in zahlreichen neuropsychologischen Testverfahren zusammenhängt, z. B. mit kognitiven Inhibitionsprozessen (Salthouse & Meinz, 1995). Zudem korreliert die Zeitspanne, in der Informationen verarbeitet werden, negativ mit Leistungen des Arbeitsgedächtnisses (Kunimi & Kojima, 2014). Erfasst wird die Informationsverarbeitungsgeschwindigkeit beispielsweise mit dem Zahlen-Symbol-Subtest der WAIS-R (Wechsler, 1981; zitiert nach Hashimoto et al., 2011). Hierbei wird als Vorgabe den Zahlen 1 bis 9 jeweils ein Symbol zugewiesen (vgl. Abbildung 9). Der Proband erhält eine lange Liste mit Zahlen und muss diesen möglichst schnell und genau die passenden Symbole zuordnen. Der Zahlen-Symbol-Test erfordert Geschwindigkeit, Genauigkeit, handschriftliche Fertigkeiten und Beweglichkeit und damit eine Kombination motorischer und kognitiver Funktionen. Erhoben wird zum einen die Anzahl der korrekten Symbole in einer bestimmten Zeit. Zum anderen wird die kognitive Komponente von *processing speed* gemessen. Diese

wird aus der Differenz der Ergebnisse zweier Teilaufgaben (Hauptaufgabe: Symbole den Zahlen zuordnen – Hilfsaufgabe: Symbole abzeichnen) ermittelt, um interindividuelle Unterschiede in den physischen Fertigkeiten „herauszufiltern“. Darüber hinaus wird auch mit dem TMT–A (Reitan, 1958) *processing speed* erfasst.

1	2	3	4	5	6	7	8	9
∨	⊐	÷	∧	X	⌒	⊔	⊖	⌒

2	1	3	1	4	2	1	3	5	3	2	1	4	2	1	3	1	2	4	1
⊐	∨	÷	∨	∧															

1	2	3	4	5	6	7	8	9
∨	⊐	÷	∧	X	⌒	⊔	⊖	⌒

2	1	3	1	2	1	3	1	4	2	4	2	5	1	4	3	5	2	6	2

1	6	5	2	4	7	3	5	1	7	6	3	8	5	3	6	4	2	1	8

9	2	7	6	3	5	8	3	6	5	4	9	7	1	8	5	3	6	8	2

7	1	9	3	8	2	5	7	4	1	6	7	4	5	8	2	9	6	4	3

Abbildung 9. Zahlen-Symbol-Test (Wechsler, 1981).

1.6 Entscheidungsverhalten und Kategorisierung

Entscheidungsverhalten, *decision making*, lässt sich definieren als „the ability to process environmental information in order to make advantageous decisions“ (Cavedini, Gorini & Bellodi, 2006, S. 4). Zwangspatienten tun sich oft schwer damit, Entscheidungen zu treffen. Einerseits hängt dies mit zwanghafter Langsamkeit (Hymas, Lees, Bolton, Epps & Head, 1991) und metakognitiv mit einem überhöhtem Verantwortungsgefühl (Rachman, 1993), andererseits auch mit neurobiologischen

Auffälligkeiten im Frontalcortex (Cavedini et al., 2006) zusammen. Entscheidungsverhalten wird mittels spieltheoretischer Computeraufgaben untersucht. In der Iowa Gambling Task (IGT; Bechara, Damasio, Damasio & Anderson, 1994) erhält der Proband vier Kartenstapel und ein fiktives Startkapital von 2000 Dollar. Über mehrere Durchgänge soll er je eine Karte aufdecken und erhält dafür entweder einen Gewinn oder Verlust. Ziel ist, die gewinnträchtigen Kartenstapel zu erkennen und damit einen möglichst hohen Gewinn zu erzielen. Die Cambridge Gambling Task (CGT) aus der CANTAB (Cambridge-Cognition, 2013) folgt einem ähnlichen Prinzip. Auch hier setzt der Proband zunächst Spielgeld, dann muss er raten, ob hinter dem blauen oder roten Quadrat, das auf dem Bildschirm erscheint, ein gelbes Symbol versteckt ist. Je häufiger der Proband richtig geraten hat, desto größer fällt sein Gewinn aus.

Klinische Beobachtungen von Patienten mit Sammelzwang zeigten, dass sie Schwierigkeiten haben, Entscheidungen zu treffen (Grisham, Norberg, Williams, Certoma, & Kadib, 2010; Tolin et al., 2012; Wincze, Steketee & Frost, 2007). Um speziell bei Patienten mit Sammelzwang das Entscheidungsverhalten und die Fähigkeit zur Kategorisierung zu untersuchen, wurden alltagsnahe Aufgaben entwickelt. Bei der Object Sorting Task sollten die Probanden Gegenstände wie z. B. Zeitschriften, Kleidung, leere Behältnisse, Büromaterial, Badezimmerutensilien in Kategorien ordnen, während sie bei der Personal Belonging Task eigene Gegenstände mitbringen und sortieren sollten (Wincze et al., 2007). Gemessen wurden sowohl die Bearbeitungszeit als auch die Anzahl der gebildeten Kategorien. Diese Aufgaben sind speziell auf die Problematik der Patienten mit Sammelzwang abgestimmt und damit aufgrund ihrer Alltagsnähe von hoher ökologischer Validität. Allerdings wird der Vorteil dieser idiosynkratischen Methode durch eingeschränkte Vergleichbarkeit geschmälert: Es ließen sich bislang keine Studien finden, bei denen Zwangspatienten beispielsweise mit Wasch- oder Kontrollzwang Alltagsgegenstände in Kategorien ordnen sollten.

1.7 Fazit zur Neuropsychologie

Im Zusammenhang mit Zwangsstörungen sind sieben kognitive Domänen von besonderer Bedeutung: Aufmerksamkeit, Exekutivfunktionen, Gedächtnis, Arbeitsgedächtnis, Verarbeitungsgeschwindigkeit, Entscheidungsverhalten und Kategorisierung. Aus der Fülle neuropsychologischer Testverfahren wurden hier diejenigen vorgestellt, deren an Zwangspatienten gewonnene Ergebnisse im später folgenden empirischen Teil dieser Arbeit metaanalytisch betrachtet werden sollen. Diese Testverfahren lassen sich in zwölf Subdomänen unterteilen: Aufmerksamkeitstests erfassen hier vor allem anhaltende Aufmerksamkeit. Aufgaben zu Exekutivfunktionen beinhalten Planen/Problemlösen, kognitive und behaviorale Inhibition, Konzeptwechsel/Regellernen und Organisationsstrategien bei der Enkodierung. Gedächtnistests umfassen verbale und nonverbale episodische Gedächtnisleistungen, während Aufgaben zum Arbeitsgedächtnis Leistungen des visuell-räumlichen Skizzenblocks und der phonologischen Schleife abdecken. Darüber hinaus werden Testverfahren eingesetzt, die die Informationsverarbeitungsgeschwindigkeit, das Entscheidungsverhalten und die Fähigkeit zur Kategorisierung erfassen. Die Testverfahren sind zum großen Teil standardisiert und gut bewährt, zu einem kleineren Teil speziell für Patienten mit Sammelzwang entwickelt worden.

2 Symptomdimensionen der Zwangsstörung

Zwangssymptome treten in vielfältigen Erscheinungsformen auf. Zwei Patienten mit ein und derselben Diagnose einer Zwangsstörung können sich in ihrem klinischen Erscheinungsbild vollkommen unterscheiden. Entsprechend wurden sehr früh schon Versuche unternommen, die Zwangsstörung anhand ihrer Symptomatik zu unterteilen. Eine historische Zweiteilung wurde aufgrund klinischer Beobachtungen angestellt. So unterschied bereits Falret (1869; zitiert nach Mataix-Cols et al., 2005, S. 162) nach *folie de doute*, Wahnsinn des Zweifelns, und *délire du toucher*, Wahn des Anfassens. Über ein Jahrhundert später differenzierten Steketee, Grayson und Foa (1985) Waschzwang von Kontrollzwang. In der ICD–10 (Dilling, 2008) wird die Zwangsstörung unter F42.0 (vorwiegend Zwangsgedanken oder Grübelzwang), F42.1 (vorwiegend Zwangshandlungen/Zwangsrituale) und F42.2 (Zwangsgedanken und –handlungen, gemischt) eingeordnet. Versuche, diese Einteilung mit biologischen Markern, genetischen Faktoren oder Behandlungserfolg in Verbindung zu bringen, waren wenig erfolgreich (Leckman, Mataix-Cols & Rosario-Campos, 2005a). Daher wurden Ansätze entwickelt, die Symptome der Zwangsstörung differenzierter zu erfassen.

2.1 Multivariate Verfahren

Zahlreiche Forschergruppen haben die Inhalte der Zwangssymptome mit verschiedenen Fragebögen und Experteninterviews erhoben und systematisch gruppiert (vgl. die Synopse von Ecker, 2009 sowie die Einzelarbeiten von Backenstrass, Schaller & Jäntschi, 2012; Baer, 1994; Ball, Baer & Otto, 1996; Calamari, Wiegartz & Janeck, 1999; Gönner, Schmid, Gönner, Leonhart & Ecker, 2013; Khanna, Kaliaperumal & Channabasavanna, 1990; Leckman et al., 2010; Leckman, Bloch & King., 2009; Mataix-Cols, do Rosario-Campos & Leckman, 2005; McKay et al., 2004; Sartory, Röper, Pietrowsky, Cwik & Zaudig, 2013; Storch et al., 2010; Taylor, 2005). Anhand von Symptomfragebögen und multivariaten statistischen Verfahren lassen sich Zwangssymptome in drei bis sechs Symptomdimensionen weiter ausdifferenzieren. Kathmann (2008) beschreibt in

seiner Übersichtsarbeit zum einen die Gewinnung verschiedener Faktoren durch Faktoranalysen, zum anderen die Bildung von Symptomclustern durch Clusteranalysen.

Faktoranalysen von Zwangssymptomskalen ergaben mehrere Faktoren. Je nach Autor und Stichprobe wurden drei (Baer, 1994), vier (Bloch, Landeros-Weisenberger, Rosario, Pittenger & Leckman, 2008; Meca et al., 2011), fünf (Pinto et al., 2007) oder sechs Faktoren (Foa et al., 2002; Gönner, Hahn, Leonhart, Ecker & Limbacher, 2009) berichtet. In einer Metaanalyse zur Faktorstruktur der am häufigsten eingesetzten Symptomskala, der Yale–Brown Obsessive Compulsive Scale Symptom Checklist (Y–BOCS-SC; Goodman, Price, Rasmussen & Mazure, 1989), extrahierten Bloch et al. (2008) vier Faktoren: (1) den Faktor *Symmetrie*, der Symmetriegedanken, Wiederhol-, Ordnungs- und Zählzwänge beinhaltete, (2) den Faktor *verbotene Gedanken*, der aggressive, sexuelle und religiöse Zwangsgedanken enthielt, (3) den Faktor *Kontamination/Reinigung*, zu dem Zwangsgedanken bezüglich Kontamination und Reinigungszwänge gehörten und (4) den Faktor *Sammeln*, der Zwangsgedanken und Zwangshandlungen bezüglich Sammeln beinhaltete. Kontrollzwänge hingegen ließen sich in der Metaanalyse von Bloch und Kollegen (2008) nicht eindeutig einem Faktor zuordnen. Bestanden die mit der Y–BOCS untersuchten Stichproben aus Kindern mit Zwangsstörungen, dann luden die Kontrollzwang-Items am höchsten auf dem Faktor *Symmetrie*. Hingegen luden die Kontrollzwang-Items am höchsten auf dem Faktor *verbotene Gedanken*, wenn Stichproben mit erwachsenen Zwangspatienten untersucht wurden. Über alle Studienstichproben hinweg luden die Kontrollzwang-Items am höchsten auf dem Faktor *verbotene Gedanken*. Bloch et al. (2008) schlossen 21 Studien und über 5000 Probanden ein. Acht der Studien stammten aus nicht-englischsprachigen Ländern, so dass die extrahierten Symptomdimensionen als kulturübergreifend gelten können. Die Y–BOCS (Goodman et al., 1989) wurde inzwischen zur Dimensional Y–BOCS (DY–BOCS; Rosario-Campos et al., 2006) erweitert, die anhand von 88 Items sechs Dimensionen erfasst. Dimension (1): Zwangsgedanken bezüglich Schaden aufgrund von Verletzungen, Gewalt, Aggression oder Naturkatastrophen und damit verbundene Zwangshandlungen; Dimension (2): sexuelle und religiöse

Zwangsgedanken und damit verbundene Zwangshandlungen; Dimension (3): Zwangsgedanken bezüglich Symmetrie, Ordnen, Zählen und damit verbundene Zwangshandlungen; Dimension (4): Zwangsgedanken bezüglich Kontamination und Reinigungszwänge; Dimension (5): Sammeln und Horten; Dimension (6): Verschiedene Zwangsgedanken und damit verbundene Zwangshandlungen. Kontrollzwänge werden in der DY-BOCS bei mehreren dieser sechs Dimensionen abgefragt, z. B. Kontrollieren im Zusammenhang mit sexuellen und religiösen Zwangsgedanken oder im Zusammenhang mit Kontaminationsfurcht.

In den empirischen Untersuchungen wurde nicht nur die Y-BOCS eingesetzt. Faktoranalysen anhand von Selbstrating-Fragebögen ergaben ähnliche Ergebnisse. So enthält das Maudsley Obsessive-Compulsive Inventory (MOCI; Hodgson & Rachman, 1977) ebenfalls vier Subskalen mit insgesamt 65 Items: *Kontrollieren*, *Reinigungszwang*, *Langsamkeit* und *Zweifel*. Zieht man die Anzahl der Symptomdimensionen in Betracht, dann ermöglicht der mit 18 Items recht kurze Selbstbeurteilungsfragebogen Obsessive-Compulsive Inventory-Revised (OCI-R; Foa et al., 2002) eine noch differenziertere Betrachtung. Dessen Faktorenanalysen ergaben sowohl in der englischen Originalversion (Foa et al., 2002) als auch in der deutschen Übersetzung (Gönner et al., 2009) sechs Faktoren: (1) *Waschen*, (2) *Zwangsgedanken*, (3) *Sammeln*, (4) *Ordnen*, (5) *Kontrollieren* und (6) *mentales Neutralisieren mit Zahlen*. Insgesamt lässt sich über die verschiedenen Symptomfragebögen hinweg feststellen, dass die Dimensionen *Kontamination/Reinigung*, *Symmetrie/Ordnen* und *Sammeln* in allen Studien konsistent extrahiert werden konnten, wohingegen bei den Symptombereichen aggressive, sexuelle und religiöse Gedanken sowie Kontrollzwang Überschneidungen in den Faktorenlösungen auftraten. Kompliziert wird die Zuordnung des Erlebens und Verhaltens von Zwangspatienten auch dadurch, dass die Symptomfragebögen nicht nur beobachtbares Verhalten, sondern auch Gedanken und Handlungsmotive erheben. Entsprechend unterschieden Ecker et al. (2011) zwischen Leitsymptomen und Motivdimensionen:

„Diese motivationsbezogene Binnenheterogenität symptombasierter Subtypen [17] legt es nahe, zusätzlich zu einer Typisierung über „Leitsymptome“ (z. B. Waschen, Kontrollieren, Ordnen) übergreifende motivationale „Kerndimensionen“ der Zwangsstörung anzunehmen, die unterschiedliches Erleben bei gleichem offenem Zwangsverhalten berücksichtigen, z. B. den Herd kontrollieren, bis es sich „genau richtig“ anfühlt, versus um die Gefahr eines Hausbrands abzuwenden“. (S. 62-63)

In fast allen Studien interkorrelieren die Faktoren in mittlerer Höhe. Dies spricht dafür, dass es sich im übergeordneten Sinne um eine Störung, eben um die Zwangserkrankung handelt – es sei denn, man wollte annehmen, dass sämtliche Zwangsfragebögen mangelhaft trennscharf konstruiert worden wären. Gegen letztere Annahme sprechen hierarchische Faktormodelle, die den Symptomdimensionen einen übergeordneten Generalfaktor zuschrieben (Katerberg et al., 2010; Leckman et al., 2010).

Clusteranalysen setzten Abramowitz, Franklin, Schwartz und Furr (2003) ein, um aus der Y-BOCS-SC fünf Symptomcluster zu bilden: (1) *Schaden zufügen*, (2) *Kontamination*, (3) *Sammeln*, (4) *unannehmbare Gedanken*, (5) *Symmetrie*. Da jedoch viele Zwangspatienten als „Mischtypen“ ausgeprägte Symptome mehrerer Dimensionen aufweisen, zeigen rein kategoriale, clusteranalytisch abgeleitete Typen-Modelle keine optimale Passung. Ein dimensionales Störungsmodell wird der Komplexität der Symptomatik sowie der Tatsache, dass subklinische Zwangssymptome auch in der Allgemeinbevölkerung auftreten, eher gerecht.

Insgesamt zeigen die empirischen Untersuchungen, dass bezüglich der Zwangssymptome inhaltliche Ähnlichkeiten zwischen den clusteranalytischen Kategorien und den faktoranalytischen Dimensionen bestehen. Beide Analysemethoden müssen einander nicht ausschließen. Sie lassen sich auch so betrachten, dass sie überlappende und sich ergänzende Information liefern (Leckman, Mataix-Cols & Rosario-Campos, 2005b): Während Faktoranalysen hoch

interkorrelierende Items zu Faktoren gruppieren, liefert die Clusteranalyse Itemcluster, die den Faktoren ähneln (Hasler et al., 2005).

2.2 Das klinische Bild der Symptomdimensionen

Sehr früh schon beschrieb der Arzt und Psychologe Pierre Janet (1919) Zwangsgedanken bezüglich Verbrechen und Krankheitsangst, wobei er zu letzterer auch Ansteckungs- und Kontaminationsängste zählte. Er listete Zwangshandlungen in 23 Bereichen auf, darunter: „la manie de l’ordre, la manie de la symétrie, la manie arithmétique, la manie de la répétition“ (Janet, 1919, S. 154). Reed (1985, S. 13) bezog sich auf Janet und unterschied dabei zwischen „genotypical form“ und „phenotypical content“ der Zwangsstörung. Diese Wortwahl nahm bereits einen Gedanken des Endophänotypen-Ansatzes (Chamberlain & Menzies, 2009, vgl. Abbildung 1 in Kapitel 1) vorweg. Seitdem wurde zu den Inhalten der Zwänge viel geforscht. Rasmussen und Eisen (1989) berichteten in ihrer bedeutsamen epidemiologischen Arbeit, dass 63 % der Zwangspatienten an Kontrollzwängen litten, 50 % an Waschzwängen, 36 % an Zählzwängen, 28 % an Symmetrie- und Ordnungszwängen und 18 % an Sammelzwängen. Beinahe die Hälfte der Patienten (48 %) wiesen multiple Zwangshandlungen auf.

Aktuell beschreibt das DSM–5 (American Psychiatric Association, 2013) vier Symptomdimensionen, wobei auch hier Überlappungen von Leitsymptomen (Waschzwang, Symmetriezwang, verbotene Gedanken, Kontrollieren) und einem motivationalen Kernsymptom (Schadensvermeidung) auftreten:

Cleaning (contamination obsessions and cleaning compulsions); *symmetry* (symmetry obsessions and repeating, ordering, and counting compulsions); *forbidden or taboo thoughts* (e.g., aggressive, sexual, and religious obsessions and related compulsions); and *harm* (e.g., fears of harm to oneself or others and related checking compulsions). (S. 235-236)

Sammelzwang wurde bislang im DSM–IV und in der ICD–10 als Symptom der Zwangsstörung betrachtet, wird jedoch seit 2013 im DSM–5 als eigenständige Störung, *hoarding disorder*, klassifiziert. Im Folgenden werden Waschzwang, Symmetriezwang, verbotene Gedanken, Kontrollzwang und Sammelzwang beschrieben. Je nach Klassifikationssystem lassen sich entweder fünf Symptomdimensionen der Zwangsstörung (DSM–IV und ICD–10) oder vier Symptomdimensionen der Zwangsstörung sowie eine neue Diagnosekategorie (DSM–5) betrachten.

Waschen/Kontamination. Zu dieser Dimension zählen Zwangsgedanken der Kontamination und Verschmutzung, Angst vor Ansteckung, Gefühle von Ekel sowie Zwangshandlungen von Waschen, Reinigen und Putzen, die häufig ritualisiert durchgeführt werden. Salkovskis (1985) beschrieb die Patientin J. F., eine junge Ärztin mit Kontaminationsängsten, die befürchtete, von Kosmetika Krebs zu bekommen, wenn diese dem Sonnenlicht ausgesetzt würden. Ihre Kontaminationsfurcht und die daraus folgenden Verhaltensweisen wie mehrstündiges ritualisiertes Händewaschen und das Umräumen großer Warenmengen im Geschäft, um eine „reine“, nicht krebserregende Kosmetikpackung zu erhalten, belasteten auch die Beziehung zu ihrem Verlobten.

Kontrollieren/Zweifel. Zu dieser Dimension zählen wiederholte Zwangshandlungen von Überprüfen und Kontrollieren. Häufige Auslöser sind vermeintlich nicht ausgeschaltete Herdplatten oder Bügeleisen oder abgeschlossene Türen. Die Betroffenen leiden unter pathologischem Zweifel, *la manie de doute* (Janet, 1919, S. 303), und sie weisen ein übertrieben stark ausgeprägtes Verantwortungsgefühl auf, das Rachman (1993, S. 150) als *inflated responsibility* bezeichnete. Den Patienten drängen sich Zwangsgedanken auf, dass sie Fehler gemacht, sich selbst oder andere geschädigt haben könnten bzw. dies unbedingt verhindern müssten. Ihre Kontrollhandlungen dienen dazu, Gedanken an mögliche Katastrophen und daraus resultierende Angst zu vermindern. Dieses Motiv, das mit dem übergroßen Verantwortungsgefühl zusammenhängt, wird als Schadensvermeidung, „harm avoidance“, bezeichnet (Ecker & Gönner, 2008, S. 895).

Wiederholte Kontrollhandlungen sollen zudem die unbehagliche Empfindung, eine Handlung sei nicht genau richtig durchgeführt („not-just-right-experience“; ebenda) und das Erleben von Unvollständigkeit („les sentiments d’incomplétude dans l’action“; Janet, 1919, S. 279) reduzieren. Patienten mit Kontrollzwängen weisen im Vergleich zu Patienten mit anderen Zwangssymptomen langfristig die günstigste Prognose auf (Rasmussen, Eisen & Greenberg, 2013).

Verbotene Gedanken. Diese Zwangsgedanken beziehen sich meist auf sexuelle, religiöse, aggressive oder somatische Inhalte. Sie sind quälender Natur und drängen sich den Patienten unwillentlich auf. Sie werden als eigene, aber unerwünschte Gedanken wahrgenommen und rufen Angst oder starkes diffuses Unbehagen hervor. Pinto et al. (2007, S. 55) bezeichneten sie als „taboo thoughts“. Den Betroffenen ist es kaum möglich, diese Gedanken zu ignorieren oder zu unterdrücken (Saß & Houben, 1996). So beschrieb bereits Janet (1919, S. 96) Patienten, die unter „les obsessions criminelles“ litten. Eine Frau, die von der Idee zu töten besessen war, habe wie in einer Halluzination quer vor ihren Augen ein langes Küchenmesser gesehen: „une des femmes obsédées par l’idée de tuer, voit devant elle, a gauche, une figure traversée au niveau des yeux par un long couteau de cuisine“ (Janet, 1919, S. 96). Herumliegende Messer lösen bei vielen Zwangspatienten die Angst aus, sie könnten diese Messer für eine Gewalttat benutzen. Aggressive Zwangsgedanken können z. B. auch davon handeln, das eigene Kind in der Badewanne zu ertränken oder es zu erwürgen. Die Betroffenen leiden unter großer Angst, tatsächlich einmal jemandem etwas anzutun.

Ordnen/Symmetrie. Diese Zwänge beinhalten Zwangsgedanken über Ordnung und Symmetrie sowie Verhaltensweisen wie z. B. Kosmetikartikel im Badezimmer gleichabständig anzuordnen oder die Falten in Gardinen so zu arrangieren, dass sie exakt gleich aussehen. Diese Versuche können viele Stunden in Anspruch nehmen, so dass die Betroffenen ihren normalen Alltagstätigkeiten nicht mehr nachkommen können. Auch hier finden sich häufig Nicht-genau-richtig-Empfindungen und Unvollständigkeitserleben (Ecker & Gönner, 2008). Rasmussen und Eisen (1989, S. 70) beschrieben einen Mann, der eine Bürotür exakt in der Mitte

durchschreiten und in einem exakten 90-Grad-Winkel zum Therapeuten sitzen musste: „the man who felt he had to walk through an office door exactly in the middle or sit at a 90° angle to the therapist“. Patienten mit Symmetrie- und Ordnungszwängen zeigen im Vergleich zu den anderen Symptomdimensionen den frühesten Krankheitsbeginn im Alter von ca. 13 Jahren (Kichuk et al., 2013), und sie weisen langfristig eine ungünstige Prognose auf (Rasmussen et al., 2013).

Sammeln (hoarding). Wie Frost (2013) auf der Webseite des International OCD Foundation (IOCDF)-Hoarding Center anführt, kann zwanghaftes Sammeln gelegentlich als Symptom der Zwangsstörung, *OCD-based hoarding*, betrachtet werden, während in den meisten Fällen die neue DSM–5-Diagnosekategorie *hoarding disorder* das Störungsbild der Betroffenen besser beschreibt. *OCD-based hoarding* wird nach Frost (2013) durch vier Merkmale charakterisiert: Erstens basiert der Sammelzwang auf „klassischen Zwangsgedanken“ wie Kontaminationsfurcht (beispielsweise können Gegenstände nicht entsorgt werden, weil sie kontaminiert sind und daher nicht berührt werden dürfen), abergläubischen Gedanken (wenn man etwas wegwirft, wird etwas Schreckliches passieren) oder dem Vermeiden endloser Kontrollrituale, die vor dem Wegwerfen von Dingen ausgeführt werden müssten. Zweitens empfinden die von *OCD-based hoarding* Betroffenen das Sammeln als unerwünscht und belastend. Drittens zeigen die Betroffenen keine emotionale Bindung an die gesammelten Gegenstände. Viertens findet sich kein exzessiver Erwerb von Gegenständen – es sei denn, die Betroffenen müssten Gegenstände z. B. aufgrund von Symmetrie- oder Zählzwang in bestimmter Anzahl kaufen. Sowohl von Sammelzwang als auch von Symmetrie- und Ordnungszwang Betroffene empfinden Unbehagen, wenn andere Personen ihre Besitztümer berühren oder gar bewegen (Frost & Steketee, 2010). Dass Sammelzwang gemeinsam mit anderen Zwängen auftreten kann, berichteten z. B. Pertusa, Frost und Mataix-Cols (2010): Eine 53-jährige Patientin mit Wiederhol-, Zähl- und Sammelzwängen litt unter Zwangsgedanken, dass einem geliebten Menschen etwas Schlimmes passieren könne, wenn sie Gegenstände wegwerfe. Um solchen Schaden zu vermeiden, bewahrte sie beispielsweise Essensreste und abgeschnittene Fingernägel auf. Gelegentlich gelang es ihr, an „glücksbringenden Tagen“ (dies waren für sie die

Dienstage) Gegenstände in der „glücksbringenden“ Anzahl von fünf zu entsorgen. Sie nahm nur zwei Mahlzeiten am Tag zu sich, um die Menge an zu entsorgendem Abfall und damit den möglichen Schaden, der geliebten Menschen zustoßen könnte, möglichst gering zu halten.

Die *hoarding disorder* wird gemäß DSM-5 charakterisiert durch anhaltende Schwierigkeiten, Gegenstände des eigenen Besitzes wegzuworfen, durch ein wahrgenommenes Bedürfnis, diese Gegenstände zu sammeln, durch Leid im Zusammenhang mit dem Entsorgen der Gegenstände und durch eine Anhäufung von Gegenständen, bis Wohnbereiche durch Überfüllung unbenutzbar werden. Sammeln wird dann störungswertig im Sinne der *hoarding disorder*, wenn es klinisch bedeutsames Leid oder Beeinträchtigungen in sozialen, beruflichen oder anderen Funktionsbereichen wie z. B. Sicherheit und Hygiene hervorruft. Die Prävalenz der *hoarding disorder* liegt bei 2 – 6 % (American Psychiatric Association, 2013) und damit wahrscheinlich höher als die der Zwangsstörung. Patienten mit *hoarding disorder* sprechen nicht gut auf die „klassische“ Psychotherapie der Zwangsstörung an (Abramowitz et al., 2003; Jefferys & Moore, 2008), und es wurden mittlerweile Therapieansätze speziell für Patienten mit *hoarding disorder* entwickelt (Gilliam et al., 2011; Steketee & Tolin, 2011). Während Zwangspatienten beim Unterbinden ihrer Zwänge Angst erleben, berichten Patienten mit *hoarding disorder* Trauer und Ärger, wenn sie gehortete Dinge entsorgen (Frost, 2013). Zudem leiden Zwangspatienten meist unter ihren Symptomen und empfinden sie als ichdyston, während Patienten mit *hoarding disorder* das Horten von Gegenständen als ichsynton empfinden und damit sogar positive Emotionen verbinden (ebenda). Pertusa et al. (2010) beschrieben die Patientin Frau H., eine 37-jährige Ehefrau, die alle Arten von Papieren, ungeöffnete Briefe, ungetragene Kleidungsstücke, Bücher, Zeitschriften in mehrfacher Ausfertigung und Videobänder sammelte, mit denen sie in Kontakt kam, um „die Zeit zu bewahren“, in der sie lebte. Frau H. hatte eine emotionale Bindung zu diesen Gegenständen, die sie als ihre „Kinder“ bezeichnete. Die gehorteten Objekte vermittelten ihr ein Gefühl von Identität, so dass ein Entsorgen für sie bedeutet hätte, ein Stück ihrer selbst zu verlieren.

Der Schweregrad zwanghaften Sammelns kann nicht nur im diagnostischen Gespräch exploriert, sondern auch visuell über die Clutter Image Rating Scale (Steketee & Frost, 2008) erfasst werden, bei der der Betroffene dasjenige Bild auswählt, das dem eigenen Wohnumfeld am ehesten entspricht (vgl. Abbildung 10). Als Selbstbeurteilungsfragebogen wurde das Saving Inventory-Revised (SI-R; Frost, Steketee & Grisham, 2004) entwickelt, das drei Faktoren (exzessiven Erwerb von Dingen, Schwierigkeiten beim Wegwerfen sowie Durcheinander, *clutter*) erfasst und cut-off-Werte für behandlungsbedürftigen Sammelzwang bietet.



Abbildung 10. Clutter Image Rating Scale (Steketee & Frost, 2008, S. 3). Ab einem Skalenwert von 4 empfehlen die Autoren professionelle Unterstützung.

Der Begriff der *hoarding disorder* besitzt noch keine gängige deutsche Übersetzung, und auch die Diagnosestellung nach ICD-10 erfolgt derzeit noch unter der Kategorie der Zwangsstörung. Die älteren Studien der vorliegenden Metaanalyse untersuchten (Zwangs-)Patienten mit dem Symptom Sammelzwang, während die

neueren Studien Patienten mit der Störung *hoarding disorder* einschlossen. Um diesen diagnostischen Überlappungen und der Veränderung in den Diagnosesystemen Rechnung zu tragen, werden im Folgenden Patienten mit Sammelzwang und Patienten mit *hoarding disorder* gemeinsam als Patienten mit Sammelzwang bezeichnet, da dies den breiteren Begriff darstellt.

2.3 Familiengenetische und Zwillingsstudien

Zumindest einige der Symptomdimensionen der Zwangserkrankung weisen auch eine genetische Komponente auf. Erste familiengenetische Studien deuteten darauf hin, dass Familienangehörige von Zwangspatienten mit hohen Werten auf den Dimensionen *aggressive/sexuelle Zwangsgedanken* und *Symmetrie-/Ordnungszwang* ein signifikant erhöhtes Risiko aufwiesen, selbst an einer Zwangsstörung zu erkranken (Alsobrook, Leckman, Goodman, Rasmussen & Pauls, 1999). Der Befund für die erbliche Komponente der Symmetrie-/Ordnungs-Dimension ließ sich in zwei Zwillingsstudien zu Zwangspatienten (Hasler et al., 2007) und zu Patienten mit Gilles de la Tourette-Syndrom und Zwangssymptomen (Leckman et al., 2003) replizieren. Van Grootheest, Boomsma, Hettema und Kendler (2008) untersuchten in einer weiteren, großen Zwillingsstudie 1 838 weibliche Zwillinge. Die Autoren unterzogen die Skalenwerte der Teilnehmerinnen auf dem Padua-Inventar–Revised (PI–R; van Oppen, Hoekstra & Emmelkamp, 1995) einer Faktorenanalyse und extrahierten drei Faktoren bzw. Symptomdimensionen (*Zwangsgedanken*, *Kontamination/Waschen* und *Kontrollzwang*). Anhand von Strukturgleichungsmodellen ermittelten van Grootheest et al. (2008) Zusammenhänge der Symptomdimensionen mit genetischen Faktoren sowie mit den Einflüssen geteilter Umwelt und nicht geteilter Umwelt. Nur die Symptomdimension *Kontamination/Waschen* erschien von spezifischen genetischen Faktoren beeinflusst, wohingegen die Autoren für *Zwangsgedanken* und *Kontrollzwang* keine spezifischen erblichen Einflüsse feststellen konnten. Die Forschergruppe um Katerberg et al. (2010) extrahierte aus den Y–BOCS-Angaben von 1224 Zwangspatienten fünf Faktoren (*Verbotene Gedanken*, *Kontamination/Waschen*, *Zweifel*, *Rituale/Aberglaube*, *Horten/Symmetrie*). Anschließend führten die Autoren anhand der Faktorscores von

52 Zwangspatienten und anhand der Raten ihrer ebenfalls zwangserkrankten Familienangehörigen über vier Generationen hinweg statistische Erblichkeitsanalysen durch. Hierbei zeigten sich signifikante Zusammenhänge zwischen Symmetriezwang und genetischen Faktoren. Zudem ermittelten die Autoren einerseits gemeinsame genetische Einflüsse für die Dimensionen *Kontamination* und *Symmetrie*, andererseits für *verbotene Gedanken* und *Zweifel*. Zudem fanden Katerberg et al. (2010) Belege dafür, dass alle fünf Faktoren erblich sind. Molekulargenetische Zusammenhänge der Symmetrie-/Ordnungs-Dimension wurden erstmals von Hasler, Kazuba und Murphy (2006) berichtet. Die Autoren fanden eine signifikante Korrelation des Faktors *Symmetrie/Ordnen* mit einem Mechanismus des Serotonin-Transportergens. Zudem konnte die Forschergruppe um Taj et al. (2013) nachweisen, dass ein bestimmtes Allel ausschließlich bei *Symmetrie/Ordnen*, nicht aber bei den anderen Symptomdimensionen auftrat. Taj und Kollegen (2013, S. 22) zogen daraus den Schluss: „Our finding ... contributes to the growing body of knowledge suggesting that symmetry and ordering represent a genetically unique symptom dimension of OCD“.

2.4 Bildgebungsbefunde

Bildgebungsstudien liefern ebenfalls Evidenz für die verschiedenen Symptomdimensionen der Zwangsstörung. Die Symptomdimensionen hängen einerseits mit strukturellen Auffälligkeiten, andererseits mit unterschiedlichen Aktivierungs- und Inhibitionsprozessen bestimmter Gehirnregionen zusammen, und sie ließen sich in mehreren Studien abgrenzbaren neuronalen Systemen zuordnen. Konsistent mit dem cortico-striato-thalamischen Modell der Zwangsstörung (Milad & Rauch, 2012, vgl. Abbildung 2 in Kapitel 1) ließen sich Beziehungen der einzelnen Zwangssymptome mit veränderter Aktivierung im OFC, im ACC, dem Nucleus caudatus und dem Thalamus (für eine Übersicht vgl. Leckman et al., 2009) nachweisen.

Funktionelle Unterschiede in der kortikalen Aktivierung, abhängig von verschiedenen Symptomdimensionen, untersuchten Mataix-Cols et al. (2004). Die Autoren führten mit 17 Patienten mit multiplen Zwangssymptomen und mit 17

gesunden Kontrollen eine funktionelle MRT-Studie durch und verwendeten ein Symptomprovokationsparadigma. Jeder Symptomdimension konnte ein anderes Aktivierungsmuster des Gehirns zugeordnet werden. Bei waschzwangbezogenen Stimuli zeigten die Patienten eine signifikant höhere Aktivierung in präfrontalen Regionen und im rechten Nucleus caudatus, bei kontrollzwangbezogenen Stimuli eine erhöhte Aktivierung im Putamen und im Thalamus. Hingegen wiesen die Patienten bei sammelzwangbezogenen Stimuli eine verstärkte Aktivierung im rechten OFC auf. Korrelationsanalysen innerhalb der Patientengruppe ergaben positive Zusammenhänge zwischen Scores auf den Symptomfragebogen Padua Inventory–Washington State University Revision (PI–WSUR; Burns, Keortge, Formea & Sternberger, 1996) und SI–R (Frost et al., 2004) und den neuronalen Reaktionsmustern. Die Autoren schlossen aus den Ergebnissen, dass den verschiedenen Symptomdimensionen unterscheidbare Teilaspekte cortico-basothalamischer Kreisläufe zugrunde liegen, die in kognitive und emotionale Prozesse involviert sind. Sie kamen zu folgendem Fazit: „Obsessive-compulsive disorder may be best conceptualized as a spectrum of multiple, potentially overlapping syndromes rather than a unitary nosologic entity” (Mataix-Cols et al., 2004, S. 564). Kritisch muss angemerkt werden, dass die Studie an einer kleinen Stichprobe durchgeführt wurde und dass die meisten Patienten mediziert waren, so dass sie nicht mit unbehandelten Patienten verglichen werden konnten. Der Befund dieser Studie scheint bemerkenswert, sollte aber durch Replikationen abgesichert werden.

Strukturelle Unterschiede untersuchten van den Heuvel et al. (2009) in einer MRT-Studie anhand einer größeren Stichprobe von 55 medikationsfreien Zwangspatienten und 50 gesunden Kontrollen. Die Forschergruppe führte bei den Probanden eine voxel-, d. h., bildpunkt-basierte Morphometrie des gesamten Gehirns durch und erfasste die Symptomdimensionen mit zwei Instrumenten, dem PI–R (Burns et al., 1996) und der Y–BOCS–SC (Goodman et al., 1989). Anhand multipler Regressionsanalysen untersuchten die Autoren den Zusammenhang zwischen den Symptomdimensionenwerten und den Volumina der Grauen und Weißen Substanz. Jeder Symptomdimension konnte ein unterscheidbares neurales Substrat zugeordnet werden. Werte auf der Dimension *Kontamination/Waschen* korrelierten negativ mit

dem Volumen der Grauen Substanz im bilateralen dorsalen Nucleus Caudatus. Hingegen korrelierten Scores auf der Dimension *Schadensvermeidung/Kontrollzwang* negativ mit den Volumina der Grauen und Weißen Substanz beider Temporallappen. Werte auf der Dimension *Symmetrie/Ordnung* korrelierten negativ mit dem Volumen der Grauen Substanz des beidseitigen Parietalcortex, und sie korrelierten positiv mit den Volumina der Grauen und Weißen Substanz beider medialer Temporallappen. Sämtliche Korrelationskoeffizienten lagen zwischen $r = .32$ und $r = .59$. Die Untersuchung von van den Heuvel et al. (2009) wurde methodisch sorgfältig durchgeführt. Allerdings konnten aufgrund der geringen Prävalenz keine Patienten mit Symptomschwerpunkten auf *verbotenen Gedanken* und *Sammeln* untersucht werden. Je nach Symptomdimension fanden sich Volumenvergrößerungen oder -verminderungen in bestimmten Gehirnbereichen. Dies spricht dafür, dass die Zwangsstörung keine homogene Störung darstellt. Die Autoren forderten, das derzeitige psychobiologische Modell der Zwangsstörung zu revidieren, da die Heterogenität der Störung mit berücksichtigt werden müsse. Dieser Forderung trug die Aufnahme der Symptomdimensionen ins DSM-5 (American Psychiatric Association, 2013) Rechnung.

2.5 Komorbidität

Wie bei allen psychischen Störungen treten auch bei der Zwangsstörung komorbide Störungen auf. Hierin unterscheiden sich die Symptomdimensionen ebenfalls. Hasler et al. (2005) untersuchten 317 Zwangspatienten und ermittelten Odds Ratios (OR)² für das Risiko komorbider Achse-I-Störungen. Die Symptomdimensionen zeigten spezifische Komorbiditäten. Patienten mit Schwerpunkt *Kontamination/Reinigungszwänge* zeigten ein erhöhtes Risiko für

² Ein Odds Ratio (OR) ist ein Quotient aus Wahrscheinlichkeiten. Das OR gibt als medizin-statistischer Wert an, wie hoch das Risiko für eine Erkrankung ist, wenn ein bestimmter Risiko- bzw. Schutzfaktor vorliegt. Bei einem OR von 1 hat der Faktor keinen Einfluss auf das Erkrankungsrisiko. Ein OR größer als 1 spricht für ein erhöhtes Krankheitsrisiko bei Vorliegen des (Risiko-)Faktors, während ein OR kleiner als 1 auf ein vermindertes Krankheitsrisiko bei Vorliegen des (Schutz-)Faktors hinweist.

Essstörungen (OR = 1.4). Im Gegensatz dazu wiesen Patienten, die auf der Dimension *Symmetrie/Ordnung* hohe Werte erzielten, erhöhte Risiken für bipolare Störung, Panikstörung und Agoraphobie auf (OR von 1.5 bis 2.2). Zwangspatienten mit vorherrschenden Symmetrie- und Ordnungszwängen waren zudem häufiger von zwanghaften Persönlichkeitsstörungen (Coles, Pinto, Mancebo, Rasmussen & Eisen, 2008) und chronischen Ticstörungen (Leckman, Grice, Boardman & Zhang, 1997) betroffen. Hingegen fand sich bei Patienten mit Sammelzwang ein hohes Risiko für alle Achse-II-Störungen, insbesondere für Cluster-C-Persönlichkeitsstörungen (Mataix-Cols, Baer, Rauch & Jenike, 2000).

2.6 Ansprechen auf Verhaltenstherapie und Psychopharmaka

Die Symptomdimensionen weisen nicht nur unterschiedlich hohe Risiken für komorbide Störungen auf. Sie erfordern auch unterschiedliche therapeutische Vorgehensweisen. In der Verhaltenstherapie der Zwangsstörung stellt die Exposition mit Reaktionsmanagement das Mittel der Wahl dar. Hierbei muss sich ein Patient mit Kontaminationsängsten und Reinigungszwängen anderen Stimuli aussetzen als ein Patient mit Kontrollzwängen (Steketee et al., 1985). Während ersterer die Angstbewältigung anhand von schmutzigen Reizen lernt, ist für letzteren die Habituation an Unsicherheit in Bezug auf potentiell angeschaltete Herde oder offengelassene Türen relevant. Entsprechend werden therapeutische Interventionen nicht an „den Zwangspatienten“, sondern an die individuelle Symptomatik des Einzelnen angepasst (Sookman, Abramowitz, Calamari, Wilhelm & McKay, 2005). Behandelt man ein Symptom erfolgreich, dann verschwindet ein zweites Zwangssymptom nicht automatisch mit. Daher sind differentielle Interventionen erforderlich, für die zuvor die Zwangssymptomatik differenziert betrachtet werden muss. Unterschiede in den Symptomschwerpunkten hängen auch damit zusammen, welche Zwangspatienten häufiger behandelt werden. So erhielten in einer Untersuchung von Ball et al. (1996) fast alle Patienten mit Waschzwang eine Psychotherapie, jedoch weniger als die Hälfte der Patienten mit Kontrollzwang. Zwangspatienten mit Zähl-, Symmetrie-, Sammel- und multiplen Zwängen waren in Therapiestudien unterrepräsentiert (ebenda). Neuere Studien bestätigten diese

Unterschiede und zeigten, dass Zwangspatienten zudem unterschiedlich gut auf eine kognitiv-verhaltenstherapeutische Behandlung ansprechen (Salomoni et al., 2009; Leckman, 2010). Während Patienten mit Kontrollzwängen gut von Verhaltenstherapie profitieren (Ball et al., 1996; Drummond, 1993), sprechen Zwangspatienten mit ausgeprägter Symptomatik auf der Dimension *verbotene Gedanken* schlechter auf Psychotherapie an (Alonso et al., 2001; Mataix-Cols, Marks et al., 2002; Starcevic & Brakoulias, 2008). Sammelzwänge wurden in den letzten Jahren verstärkt erforscht und behandelt. Trotz verbesserter Diagnostik und Therapie weisen Patienten mit Sammelzwang jedoch wenig Compliance und nur mäßige Behandlungserfolge auf und brechen häufig die Behandlung ab (vgl. die Übersicht von Leckman, Rauch und Mataix-Cols, 2007).

Die Symptomdimensionen können als Prädiktoren für das Ansprechen auf medikamentöse Behandlung der Zwangsstörung dienen. So sprechen Patienten mit Sammelzwang (Leckman et al., 2007, 2010; Mataix-Cols et al., 2005) und Patienten mit hohen Werten auf der Dimension *verbotene Gedanken* schlechter an auf SSRI als Zwangspatienten mit anderen Symptom Schwerpunkten (Alonso et al., 2001; Mataix-Cols, Rauch, Manzo, Jenike & Baer, 1999; Shavitt et al., 2006; Starcevic & Brakoulias, 2008). Wie Leckman et al. (2010) in ihrem ausführlichen Review berichteten, haben Patienten mit Symmetrie-Ordnungszwang, die oft unter komorbiden Ticstörungen leiden, häufig einen Nutzen von neuroleptischer Ergänzung der SSRI-Behandlung, während Patienten mit Kontrollzwängen generell am besten auf pharmakologische Behandlung ansprechen. Wenn weder Psychotherapie noch medikamentöse Behandlung hilft, wenden Mediziner vereinzelt die Methode der Tiefenhirnstimulation an. Zwangspatienten mit Perfektions-, Sammel- oder Symmetriezwängen profitierten bislang kaum von dieser neurochirurgischen Methode, während Patienten mit Reinigungszwängen eine durchschnittliche Besserung von 75 %, gemessen mit der Y-BOCS, erfuhren (Leckman et al., 2010).

Mataix-Cols, Rauch et al. (2002) untersuchten erwachsene Zwangspatienten im Abstand von zwei Jahren mit der Y-BOCS-SC. Die meisten Patienten behielten

über die Zeit hinweg „ihre“ Symptomdimensionen bei, d. h., dass ein Wechsel von einer Symptomdimension zu einer anderen kaum vorkam. Der stärkste Prädiktor für ein bestimmtes Symptom bestand darin, dieses Symptom bereits in der Vergangenheit gehabt zu haben. Die Partialkorrelationen über den 2-Jahres-Zeitraum waren für *Kontamination/Waschen* mit $r = .73$ am höchsten, und auch die anderen *Symptomdimensionen* (*Sammeln*, *Symmetrie*, *verbotene Gedanken*) wiesen Korrelationen von mindestens $r = .57$ auf. Es zeigten sich signifikante Unterschiede darin, welche Symptomdimensionen sich über die Jahre hinweg veränderten. Während die Dimensionen *aggressive Gedanken/Kontrollzwang*, *Symmetrie/Ordnen* und *Kontamination/Waschen* in ihrer Intensität schwankten, zeigten *Sammelzwang* und *religiöse/sexuelle Zwangsgedanken* kaum Veränderung.

Zusammenfassend kann gesagt werden, dass Patienten mit Wasch- und Kontrollzwängen sowohl auf Psychotherapie als auch auf Pharmakotherapie recht gut ansprechen, während Patienten mit verbotenen Zwangsgedanken, *Symmetrie-/Ordnungszwängen* und *Sammelzwängen* eine ungünstigere Prognose aufweisen.

2.7 Fazit zu Symptomdimensionen

Befunde aus zahlreichen Forschungsbereichen weisen darauf hin, dass die Zwangsstörung kein einheitliches Krankheitsbild darstellt. Die Symptome lassen sich über klinische Beobachtung und multivariate Verfahren auf verschiedenen Dimensionen einordnen, die individuell unterschiedlich stark ausgeprägt sein können. Empirisch gut bestätigte Symptomdimensionen sind: *Waschzwang/Kontaminationsfurcht*, *Kontrollzwang/Zweifel*, *verbotene Gedanken*, *Ordnungs-/Symmetrienzwang* und *Sammelzwang*. Letzterer nimmt *Sammelzwang* insofern eine Sonderstellung ein, als er zum einen als *OCD-based hoarding* als Symptomdimension der Zwangsstörung betrachtet wird, zum anderen neuerdings als *hoarding disorder* eine eigenständige Störungskategorie darstellt. Die Symptomdimensionen bleiben über Jahre hinweg unterschiedlich stabil, und sie weisen familiäre und genetische Besonderheiten auf. Befunde aus bildgebenden Verfahren liefern Hinweise darauf, dass den verschiedenen Symptomdimensionen verschiedene neuronale Korrelate entsprechen. Zudem unterscheiden sich die

Symptomdimensionen im Hinblick auf die Art und Häufigkeit komorbider Störungen und in Bezug darauf, wie erfolgreich sie mit Psychotherapie und Psychopharmaka zu behandeln sind.

3 Forschungsstand, Ziel und Fragestellung

Die bisherige Forschung zu den neuropsychologischen Leistungen von Zwangspatienten ergibt ein heterogenes Bild. Im Folgenden werden hierzu die wichtigsten Befunde vorgestellt, und zwar gegliedert in vier Abschnitte: Erstens Einzel- und zweitens Übersichtsarbeiten zur Neuropsychologie der Zwangspatienten, beides ohne Berücksichtigung von Symptom Schwerpunkten, drittens Einzel- und viertens Übersichtsarbeiten zur Neuropsychologie der Zwangspatienten, beides unter Berücksichtigung ihrer Symptomdimensionen. Aus den Forschungsbefunden dieser vier Bereiche werden dann das Ziel und die Fragestellung der vorliegenden Arbeit abgeleitet.

Einzelstudien von Zwangspatienten im Vergleich zu Gesunden ohne Berücksichtigung möglicher Symptom Schwerpunkte. Mehrere Forschergruppen fanden keine signifikanten Unterschiede in neuropsychologischen Leistungen von Zwangspatienten gegenüber Gesunden in anhaltender Aufmerksamkeit (Zielinski, Taylor & Juzwin, 1991) und in Exekutivfunktionen, d. h., im Planen/Problemlösen (Schmidtke, Schorb, Winkelmann & Hohagen, 1998), in kognitiver Inhibition (Moritz, Fischer et al., 2008) und im Konzeptwechsel (Abbruzzese Ferri & Scarone, 1995; Simpson et al., 2006). Auch beobachteten einige Autoren keine signifikanten Leistungsdifferenzen im verbalen Gedächtnis (z. B. Christensen, Kim, Dysken, & Hoover, 1992; Moritz, Kloss, von Eckstaedt & Jelinek, 2009) und im nonverbalen Gedächtnis (Moritz, Kloss et al., 2009; Simpson et al., 2006), während andere Studien signifikante Gedächtnisdefizite der Zwangspatienten sowohl im verbalen (Savage et al., 2000) als auch im nonverbalen Bereich (Deckersbach, Otto, Savage, Baer, & Jenike, 2000) berichteten. Dass Zwangspatienten gegenüber Gesunden in der Verarbeitungsgeschwindigkeit signifikante Minderleistungen aufwiesen, berichteten Schmidtke et al. (1998), wohingegen Basso, Bornstein, Carona und Morton (2001) Defizite von Zwangspatienten im TMT–A auf komorbide Depression zurückführten und Aronowitz, Hollander, DeCaria und Cohen (1994) im TMT–A keine signifikanten Einschränkungen fanden. Da die Vielzahl von Einzelresultaten hier aus

Platzgründen nicht vollständig beschrieben werden kann, werden im Folgenden Ergebnisse von Übersichtsarbeiten vorgestellt.

Übersichtsartikel über Studien von Zwangspatienten im Vergleich zu Gesunden ohne Berücksichtigung möglicher Symptomschwerpunkte. Zwei aktuelle Metaanalysen fassten die zahlreichen Einzelbefunde zusammen und zeigten über mehrere Studien hinweg, dass Erwachsene mit einer Zwangsstörung im Vergleich zu Gesunden signifikante, wenn auch nur mäßig ausgeprägte Defizite in den kognitiven Domänen Aufmerksamkeit (Abramovitch, Abramowitz & Mittelman, 2013), Exekutivfunktionen, verbales und nonverbales Gedächtnis, visuell-räumliche Fähigkeiten, Informationsverarbeitungsgeschwindigkeit und Arbeitsgedächtnis aufwiesen (Abramovitch et al., 2014; Shin, Lee, Kim & Kwon, 2014). In einem kritischen Übersichtsartikel berichteten Kuelz, Hohagen und Voderholzer (2004), dass Zwangspatienten gegenüber Gesunden Defizite im visuell-räumlichen Gedächtnis und, bei komplexeren Aufgaben, auch Defizite im verbalen Gedächtnis zeigten. Die Autoren vermuteten, dass diese Gedächtnisdefizite auf mangelhafte Strategien beim Enkodieren von Information zurückgingen. Allerdings steht der Befund von Kuelz et al. (2004), demzufolge die anhaltende Aufmerksamkeit bei Zwangspatienten insgesamt unbeeinträchtigt war, im Widerspruch zu den Ergebnissen der Reviews von Abramovitch et al. (2013) und Cox (1997). Letztere fand Hinweise darauf, dass Zwangspatienten Probleme mit der Aufmerksamkeitslenkung hatten und dass die Fähigkeit zur Inhibition automatischer Reaktionen mit der Schwere der Zwangssymptome korrelierte. Dagegen stellten Härting und Markowitsch (1997) in ihrem Review dar, dass Zwangspatienten überwiegend normale Leistungen in Aufmerksamkeit, Exekutivfunktionen und im verbalen Gedächtnis erbrachten und erst in komplexeren Aufgaben Defizite im visuell-perzeptiven und visuell-konstruktiven Bereich zeigten. Auch Muller und Roberts (2005) berichteten in einem narrativen Review, dass es insbesondere bei komplexeren visuellen Reizen empirische Belege für eine Beeinträchtigung der Zwangspatienten im nonverbalen Gedächtnis gebe, wohingegen die Befunde zu ihren verbalen Gedächtnisleistungen uneinheitlich ausfielen. Dagegen ergab das Review von Greisberg und McKay (2003), dass Zwangspatienten Defizite in der Anwendung

von Organisationsstrategien aufwiesen, und dass dies auch Auswirkungen auf Gedächtnisleistungen hatte. Damit übereinstimmend, führten Olley et al. (2007) als Ergebnis ihrer Übersichtsarbeit an, dass Zwangspatienten schwerpunktmäßig Defizite in Exekutivfunktionen aufwiesen und dass eine gestörte Organisation beim Enkodieren von Informationen für die Gedächtnisdefizite verantwortlich sei. Defizite in Exekutivfunktionen, vor allem im Konzeptwechsel, fand auch Henry (2006) in einer Metaanalyse über die Leistungen von Zwangspatienten im WCST.

Insgesamt erscheinen demnach die Ergebnisse der neuropsychologischen Vergleichsuntersuchungen „der Zwangspatienten“ gegenüber Gesunden inkonsistent. Dies könnte auf die Fülle von Testmethoden, auf unterschiedliche Stichproben und auf Einflüsse demografischer oder klinischer Merkmale zurückzuführen sein. Allerdings reichen diese Erklärungen nicht aus, denn Studien, die mit den gleichen neuropsychologischen Testverfahren Patienten mit Schizophrenie (Fioravanti, Carlone, Vitale, Cinti & Clare, 2005; Heinrichs & Zakzanis, 1998; Krabbendam, Arts, van Os & Aleman, 2005), unipolarer Depression (Bora, Yücel & Pantelis, 2010; McDermott & Ebmeier, 2009) und bipolarer affektiver Störung (Robinson et al., 2006) untersucht hatten, kamen zu eindeutigeren Ergebnissen. Die uneinheitlichen Resultate lassen sich möglicherweise besser damit erklären, dass die Zwangsstörung als homogene Erkrankung betrachtet wurde, deren Symptomheterogenität nicht ausreichend berücksichtigt wurde. Abramovitch et al. (2013) schlugen als Fazit ihrer Metaanalyse vor, eben diese Symptomheterogenität der Zwangsstörung zu berücksichtigen, um herauszufinden, ob Unterschiede in Symptomgruppen die kognitive Leistung beeinflussen und ob die kognitiven Defizite „der Zwangspatienten“ durch die Symptomheterogenität in den untersuchten Stichproben über- oder unterschätzt worden waren. Denn wenn, den endophänotypischen Modellvorstellungen und empirischen Befunden entsprechend, unterschiedliche neurobiologische Auffälligkeiten mit Unterschieden in der Zwangssymptomatik zusammenhängen (Chamberlain et al., 2005; McKay et al., 2004), dann sollten sich auch die neuropsychologischen Funktionen bei Zwangspatienten voneinander unterscheiden, je nachdem, welche Symptomdimensionen man betrachtet. Solchen Überlegungen folgend, führten

bislang einige Forschergruppen Studien durch. Diese Untersuchungen (Einzelstudien und Reviews) werden in den zwei folgenden Abschnitten kurz vorgestellt.

Neuropsychologische Einzelstudien von Zwangspatienten unter Berücksichtigung bestimmter Symptomschwerpunkte. Zwangsspatienten mit Waschzwang zeigten gegenüber Gesunden Defizite im nonverbalen Gedächtnis (Ceschi, van der Linden, Dunker, Perroud & Brédart, 2003). Hingegen fanden Macdonald, Antony, Macleod und Richter (1997) zwischen Zwangspatienten mit Kontrollzwang, Zwangspatienten ohne Kontrollzwang und Gesunden keine signifikanten Gedächtnisunterschiede. Konsistent zu letzterem Befund zeigten Zwangspatienten mit Kontrollzwang im Vergleich zu Zwangspatienten ohne Kontrollzwang sowie zu Gesunden gleich gute Gedächtnisleistungen in einer verzögerten Zuordnungsaufgabe, bei der sie per Tastendruck entscheiden mussten, ob zwei in zeitlichem Abstand gezeigte Stimuli identisch oder verschieden waren (Clair et al., 2013). Im Widerspruch dazu stehen die Ergebnisse von Bouvard, Dirson und Cottraux (1997), Cha et al. (2008) und Nakao et al. (2009), denen zufolge Patienten mit Waschzwang signifikant besser in nonverbalen Gedächtnisleistungen abschnitten als Patienten mit Kontrollzwang. Im Gegensatz dazu wiederum fanden Ceschi et al. (2003) zwischen Patienten mit Waschzwang und Kontrollzwang keine signifikanten Unterschiede im nonverbalen Gedächtnis. Studien, die Exekutivfunktionen bei Zwangspatienten untersuchten, ergaben, dass Patienten mit Kontrollzwang in Planungsaufgaben signifikant schlechter abschnitten als jene mit Waschzwang (Dittrich & Johansen, 2013; Nedeljkovic, et al. 2009), während Otori et al. (2007) und Penadés et al. (2007) beobachteten, dass Patienten mit Kontrollzwang Defizite in kognitiver und behavioraler Inhibition aufwiesen, verglichen mit Patienten mit Waschzwang. Personen mit Sammelzwang zeigten gegenüber Gesunden signifikante Defizite in Aufgaben zu Aufmerksamkeit (Grisham, Brown, Savage, Steketee & Harlow, 2007; Grisham, Norberg, Williams, Certoma & Kadib, 2010), in Exekutivfunktionen (Ayers et al., 2013), in nonverbalen und verbalen Gedächtnisaufgaben (Hartl et al., 2004) und im Entscheidungsverhalten (Grisham et al., 2010), nicht jedoch im verbalen Arbeitsgedächtnis (Grisham et al., 2007). Studien zu neuropsychologischen Leistungen von Zwangspatienten mit

verbotenen Gedanken oder Symmetrie- und Ordnungszwang waren vor Durchführung der Metaanalyse nicht bekannt.

Übersichtsartikel über neuropsychologische Studien von Zwangspatienten unter Berücksichtigung einzelner Symptomschwerpunkte. Über die zuvor angeführten Einzelstudien hinaus wurden bislang kaum Übersichtsartikel publiziert, die die Symptomdimensionen der Zwangsstörung berücksichtigten. Eine Ausnahme bilden zwei Reviews über Studien zur Gedächtnisleistung von Personen mit Kontrollzwang. So führten Woods, Vevea, Chambless und Bayen (2002) eine Metaanalyse zum Gedächtnis von Personen mit Kontrollzwang (*checkers*) im Vergleich zu Personen ohne Kontrollzwang (*non-checkers*) durch. Die Autoren ermittelten signifikante mittelgroße Effektstärken für freie visuelle Reproduktion, für verbale Reproduktion mit Hinweisreizen und für den Abruf von Handlungen sowie signifikante kleine Effektstärken für freie verbale Reproduktion, visuelles Wiedererkennen und Arbeitsgedächtnis. Diese Effektstärken zeigten an, dass *checkers* im Vergleich zu den *non-checkers* in den entsprechenden kognitiven Bereichen Minderleistungen erbrachten. Darüber hinaus stellten Cuttler und Graf (2009) in einem weiteren Review über 67 Studien mit 1519 Zwangspatienten und 236 Teilnehmern mit subklinischen Zwängen fest, dass Defizite im prospektiven Gedächtnis, nicht aber im retrospektiven Gedächtnis spezifisch seien für *checkers*. In beiden Überblicksartikeln wurden *checkers* mit *non-checkers* verglichen. Dabei enthielten die Gruppen der *non-checkers* sowohl Zwangspatienten verschiedener Symptomdimensionen als auch Patienten mit anderen Störungen und gesunde Probanden. Problematisch erscheint, dass diese heterogenen Vergleichsgruppen der *non-checkers* keine eindeutigen Schlussfolgerungen aus den Befunden zulassen. Weitere Übersichtsarbeiten von Studien, die neuropsychologische Testleistungen von Zwangspatienten mit anderen Symptomschwerpunkten miteinander verglichen, wurden bislang nach Kenntnis der Autorin nicht veröffentlicht.

Zusammenfassend lässt sich sagen, dass die zahlreichen Einzeluntersuchungen und Reviews, in denen Zwangspatienten als homogene Gruppe betrachtet wurden, bislang nicht zu eindeutigen Erkenntnissen über deren möglichen

neuropsychologischen Defizite führten. Stattdessen erscheinen die Resultate uneinheitlich und zum Teil sogar widersprüchlich. In den letzten Jahren wurden darüber hinaus einige Studien veröffentlicht, in denen die kognitiven Leistungen von Zwangspatienten unter Berücksichtigung ihrer Symptomdimensionen untersucht wurden. Auch hier scheint es heterogene Befunde zu geben, die noch nicht zusammenfasst wurden. Ein systematisches Review, das die aktuellen Forschungsergebnisse zur Neuropsychologie von Zwangsstörungen über mehrere Domänen hinweg und unter Berücksichtigung aller Symptomdimensionen untersucht, wurde nach Kenntnis der Autorin bislang nicht veröffentlicht. Diese Forschungslücke soll die vorliegende Arbeit schließen.

Das Ziel dieser Dissertation besteht darin, die aktuellen Forschungsergebnisse zu neuropsychologischen Leistungen von Zwangspatienten unter Berücksichtigung ihrer Symptomdimensionen in einer Metaanalyse zu quantifizieren. Aus der Überlegung heraus, dass die bisherigen heterogenen Befunde möglicherweise mit der Symptomheterogenität der Zwangsstörung zusammenhängen, und aus den genetischen, neurobiologischen, epidemiologisch-komorbiden und therapeutischen Hinweisen auf die Konstruktvalidität der Symptomdimensionen wurde folgende Frage abgeleitet: Zeigen Zwangspatienten mit Schwerpunkten auf verschiedenen Symptomdimensionen auch unterschiedliche neuropsychologische Testleistungen? Hierzu wurde ein systematisches Review mit Metaanalyse durchgeführt.

4 Methoden

Der Nomenklatur von Borenstein (2009) folgend, bezeichnet der Begriff „systematisches Review“ eine Übersichtsarbeit, in der die Literatursuche und die Auswahl der Studien nach zuvor festgelegten Regeln erfolgen. Wenn die Ergebnisse des systematischen Reviews nicht nur narrativ oder tabellarisch berichtet, sondern darüber hinaus in statistischen Berechnungen zusammengeführt werden, handelt es sich um eine Metaanalyse. Zu Beginn des Reviewprozesses liegen meist nur vage Informationen darüber vor, welche Daten gefunden und integriert werden können. Aufgrund dieses explorativen Charakters wird im Allgemeinen in Reviews auf die Formulierung von Hypothesen verzichtet (z. B. die Übersichtsarbeiten von Abramovitch, Abramowitz und Mittelman, 2013; Cuttler & Graf, 2009; Kuelz et al., 2004; Shin et al., 2014), so auch in der vorliegenden Arbeit. Vor der Durchführung wurde ein Reviewprotokoll (Egger, 2001) entwickelt, das folgende Schritte enthält: Eine systematische Recherche (Abschnitt 4.1) mit der Definition von Suchstrategie, Ein- und Ausschlusskriterien. Danach erfolgen die Datenextraktion (Abschnitt 4.2) anhand eines Kodierschemas und –manuals sowie die Prüfung der Studienqualität mittels einer hierfür entwickelten Skala. Die extrahierten Daten werden analysiert (Abschnitt 4.3) und die Effektstärken berechnet, nach ihrer Bedeutung gewichtet, auf ihre Homogenität geprüft und anhand eines statistischen Modells zusammengeführt (integriert; Abschnitt 4.4). Zuletzt werden noch Moderator- und Metaregressionsanalysen durchgeführt (Abschnitte 4.5 und 4.6) sowie ein möglicher Publikationsbias abgeschätzt (Abschnitt 4.7). Die Darstellung basiert auf der Checkliste Quality of Reporting of Meta-Analyses (QUOROM; Moher et al., 1999, S. 1897).

4.1 Systematische Recherche

Die Suchstrategie bestand aus drei Schritten: Datenbanksuche, Handsuche und E-Mail-Anfragen bei Forschern, die Studien zur Neuropsychologie von Zwangsstörungen durchführten. In den Datenbanken PsychINFO, Medline, PubMed, Google scholar, ProQuest Dissertations & Theses sowie Scirus wurde, in Anlehnung an die Metaanalysen von Bora et al. (2010) und Henry (2006), mittels folgender

Schlagworte für Titel, Abstract und Medical Subject Headings (MeSH) recherchiert: „obsessive compulsive“, „OCD“, „Zwang“, „wash*“ „check*“, „Wasch*“, „Kontroll*“, „hoarding“, „Hort*“ sowie „neuropsych*“, „neurocogn*“, „information processing“, „memory“, „attention“, „executive“, „set shifting“, „cognitive function“, „cognitive dysfunction“, „cognitive deficit“, „frontal function“, „frontal dysfunction“, „frontal deficit“. Von Hand wurden die Referenzlisten der gefundenen Übersichtsartikel (vgl. Kapitel 3, Stand der Forschung) nach relevanten Studien durchsucht. Per E-Mail wurden Experten, die Studien zur Neuropsychologie von Zwangsstörungen durchführten, kontaktiert und um unveröffentlichte Ergebnisse sowie um englischsprachige Übersetzungen ihrer Arbeiten gebeten.

Die Literatursuche wurde in zwei Wellen durchgeführt. Die erste Recherchewelle erfasste Publikationen von 1948 bis zum 31. Dezember 2011, die zweite vom 1. Januar 2012 bis zum 10. Januar 2014.

Die Primärstudien mussten folgende Einschlusskriterien erfüllen:

- Erwachsene Patienten mit der Diagnose Zwangsstörung nach einer Ausgabe der International Classification of Diseases (ICD) oder nach einer Ausgabe der Diagnostic and Statistical Manual of Mental Disorders (DSM)
- Spezifikation mindestens einer Symptomdimension der Zwangsstörung
- Gesunde Kontrollpersonen oder Zwangspatienten einer anderen Symptomdimension dienten als Vergleichsgruppe
- Mindestens eine abhängige Variable war ein Maß einer neuropsychologischen Testleistung
- Es wurden Statistiken (M , SD , F , t , r , exaktes p) angegeben, aus denen sich Effektstärken berechnen lassen
- Publikation in englischer, deutscher oder französischer Sprache.

Folgende Ausschlusskriterien wurden festgelegt:

- Studien ohne Vergleichsgruppe
- Studien mit kleiner Stichprobe ($n < 10$)

- Studien, in denen kein Symptomschwerpunkt der Zwangsstörung spezifiziert war
- Subklinische Zwangssymptome, z. B. in studentischen Stichproben
- Sekundäre Zwangssymptome infolge eines medizinischen Krankheitsfaktors, z. B. aufgrund von Gehirnläsionen, Infektionen oder dementiellen Prozessen
- Studien, die nur grafische Information, nonparametrische Tests oder ungenaue p -Werte enthielten und deren Autoren die exakten Daten nicht bis zum 30. Juni 2014 zur Verfügung stellen konnten.

4.2 Datenextraktion, Studien- und Kodierungsgüte

Datenextraktion. Die Informationen der Primärstudien wurden gemäß den Empfehlungen der Cochrane Collaboration (Higgins & Green, 2011) systematisch extrahiert. Hierzu wurde, angelehnt an die *data collection form* von Burt, Zembler und Nedrehe (1995), ein zweiteiliges Papier- und Bleistift-Kodierschema entwickelt. Im ersten Teil wurden deskriptive Daten zur Methode, zur Studienqualität, zu den Teilnehmern und zu den neuropsychologischen Domänen erhoben (Anhang 1). Im zweiten Teil wurden die neuropsychologischen Testergebnisse erfasst (Anhang 2). In Anlehnung an die Klassifikation von Abramovitch et al. (2013) wurden die neuropsychologischen Testvariablen in sieben übergeordnete kognitive Domänen und zwölf Subdomänen (Aufmerksamkeit: anhaltende Aufmerksamkeit; Exekutivfunktionen: Planen/Problemlösen, Inhibition, Konzeptwechsel, Enkodieren; Gedächtnis: verbales und nonverbales episodisches Gedächtnis; Arbeitsgedächtnis: visuell-räumliches und verbales Arbeitsgedächtnis; Verarbeitungsgeschwindigkeit; Entscheidungsverhalten; Kategorisierung) eingeteilt (vgl. Tabelle 1).

Tabelle 1. *In der Metaanalyse untersuchte neuropsychologische Testverfahren, klassifiziert nach übergeordneten Domänen und Subdomänen.*

Domäne/Subdomäne	Testverfahren der Primärstudien
Aufmerksamkeit	
Anhaltende Aufmerksamkeit	SAT, K-T-Durchstreichtest, WMS-R mentale Kontrolle
Exekutivfunktionen	
Planen/Problemlösen	TOH, TOL, CANTAB: SOC, RCFT Planungsscore
Kognitive und behaviorale Inhibition	FWIT, Color-Word Stroop Task, Simon Spatial Incompatibility Task, CPT-II, SSRTT bzw. SST, CANTAB: AGN, Go/No-Go Fehlerzahl und Trefferzahl
Konzeptwechsel/Regellernen	CANTAB IED, TMT-B, WCST: Anzahl der Kategorien, Gesamtzahl der Antworten, Gesamtzahl der Fehler
Enkodieren	RCFT Organisationsscore, SWM Strategiescore
Gedächtnis	
Verbales episodisches Gedächtnis	BEM 144 verbal, WAIS-R LM I and II, CVLT: verzögerter Abruf; HVLT: verzögerter Abruf Wortliste, Wortpaare, LGT-3 verbal
Nonverbales episodisches Gedächtnis	BEM 144: nonverbal, LGT-3 nonverbal, RCFT: verzögerter Abruf, WMS-R: visuelle Reproduktion I und II Delayed-Matching-to-Sample Task AMIPB Figure Recall, The Recurring Figures Test

Arbeitsgedächtnis

Visuell-räumliches Arbeitsgedächtnis	CANTAB SWM Fehlerzahl und Reaktionszeit, Spatial N-Back Task, WMS-R: Blockspanne, WAIS-R: Blockdesign, “spatial working memory task”, K-T Durchstreichtest, WAIS-IV: Buchstaben- Zahlen-Folge, WMS-R: VMS, Card Position Test, WAIS-R: Zahlen-Symbol-Test
Verbales Arbeitsgedächtnis	WMS-R: Zahlenspanne, “verbal working memory task”

**Verarbeitungs-
geschwindigkeit**

 WAIS-R Zahlen-Symbol-Test, TMT-A

Entscheidungsverhalten

 CANTAB CGT, IGT, Object Sorting Task,
Personal Belonging Task

Kategorisierung

 Object Sorting Task: Anzahl der Kategorien,
Bearbeitungszeit

Anmerkungen. AMIPB = Adult Memory and Information Processing Battery (Coughlan & Hollows, 1985; zitiert nach Tallis et al., 1999); BEM 144 = Batterie d'Efficiency Mnésique 144 (Signoret, 1991; zitiert nach Bouvard et al., 1997); CANTAB = Cambridge Automated Neuropsychological Test Battery mit den Subtests AGN = Affective Go/No-Go, SWM = Spatial Working Memory, CGT = Cambridge Gambling Task, IED = Intra-Extra-Dimensional Shift, SOC = Stockings of Cambridge (Cambridge-Cognition, 2013); Card Position Test (Singh et al., 2003); Color Word Stroop Test (Golden, 1978; zitiert nach Hashimoto et al., 2011); CPT-I and II = Continuous Performance Test, Teil I und II (Conners, 2000; zitiert nach Grisham et al., 2007); CVLT = California Verbal Learning Test (Delis et al., 1987; zitiert nach Ceschi et al., 2003); Delayed-Matching-to-Sample Task (Clair et al., 2013); FWIT = Farbe-Wort-Interferenztest (Bäumler, 1985; zitiert nach Exner et al., 2009); Go/No-Go Task (Tolin, Witt & Stevens, 2014); HVLT = Hopkins Verbal Learning Test (Brandt, 1991; zitiert nach Cha et al., 2008); IGT = Iowa Gambling Task (Bechara et al., 1994); K-T Durchstreichtest (Halter, 1986, zitiert nach Bouvard et al., 2012); LGT-3 = Lern- und Gedächtnistest (Bäumler, 1974; zitiert nach Zitterl et al., 2001); N-Back-Task (Nakao, Nakagawa, Yoshiura et al., 2009); Object Sorting

Task and Personal Belonging Task (Wincze et al., 2007); PBT = Personal Belonging Task (Cuttler & Graf, 2007); Recurring Figures Test (Kimura, 1963; zitiert nach Tallis et al., 1999); SAT = Situation Awareness Test (Mueller, 2010; zitiert nach Tumkaya et al., 2013); Simon Spatial Incompatibility Task (Simon, 1969); Spatial Working Memory Task and Verbal Working Memory Task (Kraft, 2006); SSRTT = Stop Signal Reaction Time Task synonym SST = Stop Signal Task (Logan, 1994; zitiert nach Morein-Zamir et al., 2014); TMT–A and B = Trail-Making Test, Teil A und B (Reitan, 1958); TOH = Tower of Hanoi (Simon, 1975; zitiert nach Kuelz et al., 2004); TOL = Tower of London (Shallice, 1982; zitiert nach Tolin et al., 2011); WAIS–R = Wechsler Adult Intelligence Scale–Revised (Wechsler, 1981; zitiert nach Hashimoto et al., 2011); WAIS–IV = Wechsler Adult Intelligence Scale–Fourth Edition (Wechsler, 2008; zitiert nach Ayers et al., 2013); WCST = Wisconsin Card Sorting Task (Milner, 1963); WMS–R LM I and II = Wechsler Memory Scale–Revised Logical Memory Version I and II (Wechsler, 1987; zitiert nach Nakao et al., 2009), VMS = Visual Memory Span (Wechsler, 1987; zitiert nach Grisham et al., 2007); Word List Task (MacDonald, Antony, MacLeod & Richter, 1997); Word Pairs Learning (Tuna et al., 2005).

Um eine möglichst hohe Objektivität und grundsätzliche Replizierbarkeit der Arbeit zu gewährleisten, wurde ein Kodiermanual entwickelt. Dieses enthält alle zu kodierenden Variablen sowie entsprechende Kodieranweisungen (Anhang 3).

Dateneingabe und -aufbereitung. Die deskriptiven Studiendaten wurden in einen SPSS-Datensatz überführt. Die Studienergebnisse der neuropsychologischen Tests wurden mittels der Software Comprehensive Metaanalysis Version 2.0 (CMA; (Borenstein, Hedges, Higgins & Rothstein, 2005) in Effektstärken transformiert.

Studiengüte. Um die Studienqualität als potentielle Moderatorvariable zu erfassen, wurde auf Grundlage der Psychotherapy Outcome Study Methodology Rating Scale von Öst (2008) und der Kriterien der Cochrane-Collaboration (Higgins & Green, 2011) eine Ratingskala der Studienqualität entwickelt (vgl. Tabelle 2). Die Skala enthielt elf Items, die mit *0 = mangelhaft*, *1 = ausreichend*, *2 = gut* bewertet wurden. Den Werten 0, 1 und 2 wurden jeweils konkrete Beispiele als Anker zugeordnet (Anhang 4). Der Summenwert der Gesamtskala wurde als Moderatorvariable Studienqualität in späteren Metaregressionsanalysen berücksichtigt.

Tabelle 2. *Skala zur Beurteilung der Studienqualität*

		0 = mangelhaft	1 = ausreichend	2 = gut
1	Einhaltung von Ethikrichtlinien	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	Klarheit der Stichprobenbeschreibung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Erfassung des Schweregrades der Störung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	Repräsentativität der Stichprobe in Bezug auf Komorbidität	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	Reliabilität der Diagnose	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	Spezifität der neuropsychologischen Variablen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7	Validität der neuropsychologischen Tests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	Anzahl der Testleiter	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	Training der Testleiter	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10	Kontrolle der Medikation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11	Statistische Methoden und Datenpräsentation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Kodierungsgüte. Die Skala zur Beurteilung der Studienqualität wurde anhand von 20 zufällig ausgewählten Studien und zwei unabhängigen Beurteilerinnen (Verfasserin dieser Arbeit und eine weitere Doktorandin der Psychologie) auf Interrater-Reliabilität geprüft. Die Zufallsauswahl der Studien erfolgte über einen Online-Zufallsgenerator (Intemodino Group, 2013). Die Angaben zu Studienautoren und –institutionen wurden auf den ausgedruckten Publikationen geschwärzt, so dass die Beurteilerinnen diesbezüglich verblindet waren. Es wurden ein zweistündiges Rater-Training durchgeführt, drei Studien probeweise kodiert und Unstimmigkeiten zwischen den beiden Raterinnen bezüglich der Beurteilung einzelner Variablen diskutiert, bis Konsens erzielt war. Diese Variablen wurden in einem zweiten Durchgang nachkodiert. Zur Bestimmung der Interrater-Reliabilität wurde für jedes einzelne Item der Intraclass-Korrelationskoeffizient $ICC(2,1)$ nach dem Modell mit

Zwei-Weg-Zufallseffekten, bei dem Rater- und Maßeffekte zufällig sind, mit absoluter Übereinstimmung (Shrout & Fleiss, 1979) berechnet.

Neun der elf Items der Skala zur Beurteilung der Studienqualität waren von ausreichender bis exzellenter Interraterreliabilität (*ICC*-Koeffizienten zwischen .75, $p = .001$ und .94, $p < .001$, vgl. Anhang 5). Für Item 3, *Erfassung des Schweregrades der Zwangsstörung*, konnte mangels Varianz kein Kennwert ermittelt werden, da alle 20 zufällig ausgewählten Studien die Y-BOCS eingesetzt hatten und damit für dieses Item mit dem Wert 2 beurteilt wurden. Für Item 6, *Spezifität der neuropsychologischen Variablen*, ergaben sich insignifikante niedrige Übereinstimmungen und Konfidenzintervalle, die den Wert Null einschlossen ($ICC = .38$, 95 %-KI - 0.38 – 0.74, $F(19,19) = 1.70$, $p = .129$). Dieses Item war nicht sinnvoll interpretierbar und wurde aus den nachfolgenden Analysen ausgeschlossen. Die endgültige Skala enthielt damit zehn Items. Ihre interne Konsistenz war zufriedenstellend (Cronbach's Alpha = .697).

4.3 Datenanalyse mittels Effektstärken

Berechnung der Effektstärken. Die neuropsychologischen Testergebnisse in den Primärstudien wurden in Effektstärken transformiert. Effektstärken sind statistische Differenz- oder Zusammenhangsmaße in der Einheit Standardabweichungen. Verwendet man Standardabweichungen als Messeinheit, dann kann man Ergebnisse über verschiedene Messungen hinweg kombinieren und vergleichen (Lipsey & Wilson, 2001), sofern es sich um dasselbe Konstrukt handelt. Effektstärken geben die Größe und die Richtung eines Zusammenhangs bzw. eines Unterschiedes an (Cooper, 1994). Für Unterschiedshypothesen sind die gängigsten Effektstärken die standardisierte Mittelwertsdifferenz Cohen's d und deren um einen Korrekturfaktor für kleine Stichproben ergänzte Ableitung Hedges' g . Für Zusammenhangshypothesen ist die häufigste Effektstärke die standardisierte Kovarianz, d. h., Pearsons Produkt-Moment-Korrelation r .

Wenn Studien ihre Ergebnisse in verschiedenen Maßen berichtet haben, dann lassen sich diese Ergebnismaße in ein gemeinsames Maß umrechnen, wenn es

inhaltlich Sinn ergibt. Als Faustregel kann gemäß Borenstein, Hedges, Higgins und Rothstein (2009, Kapitel 7) gelten: Würde der Reviewer die Studienergebnisse kombinieren, wenn die Autoren dieselben Statistiken benutzt hätten, dann sollten unterschiedliche Maßeinheiten keine Rolle spielen. So können beispielsweise mehrere Forschergruppen die gleichen Ausgangsmaße verwendet (z. B. die Symptomdimensionen des Padua-Inventars und den verbalen Gedächtnistest Logical Memory LM), dann aber die Ergebnisse entweder als Mittelwertsunterschiede zweier Gruppen (Patienten mit Kontrollzwang vs. Gesunde) oder als Zusammenhangsmaß (Korrelation zwischen der Ausprägung von Kontrollzwang und der Testleistung) berichtet haben. In diesem Falle ist eine Zusammenfassung der verschiedenen Ergebnismaße inhaltlich sinnvoll.

Da einerseits Zwangssymptome dimensional betrachtet werden können (eine Person weist auf verschiedenen Symptomdimensionen verschieden hohe Ausprägungen auf), andererseits aber die Mehrzahl der Studienautoren Gruppenvergleiche zwischen Zwangspatienten mit „reinen“ bzw. klinisch relevant vorherrschenden Symptomen (z. B. Patienten mit Waschzwang vs. Kontrollzwang) durchgeführt hatten, wird in der vorliegenden Arbeit dieser vereinfachenden Sichtweise distinkter Symptomgruppen Rechnung getragen. Die meisten eingeschlossenen Studien berichteten Gruppenmittelwerte (M , SD), so dass metaanalytisch Hedges' g berechnet wurde. Hingegen gaben einige Studien neueren Datums Korrelationskoeffizienten an. Diese wurden für die Metaanalyse in Hedges' g umgerechnet, so dass eine einheitliche Effektstärke alle Studienergebnisse vergleichbar machte. Die Effektstärke r lässt sich im Allgemeinen Linearen Modell in Cohen's d (und damit auch in das von d abgeleitete Hedges' g) überführen (vgl. Abbildung 11). Bei der Umrechnung von r nach d wird folgende Annahme getroffen: „we assume that the continuous data used to compute r has a bivariate normal distribution and that the two groups are created by dichotomizing one of the two variables“ (Borenstein et al., 2009, S. 48).

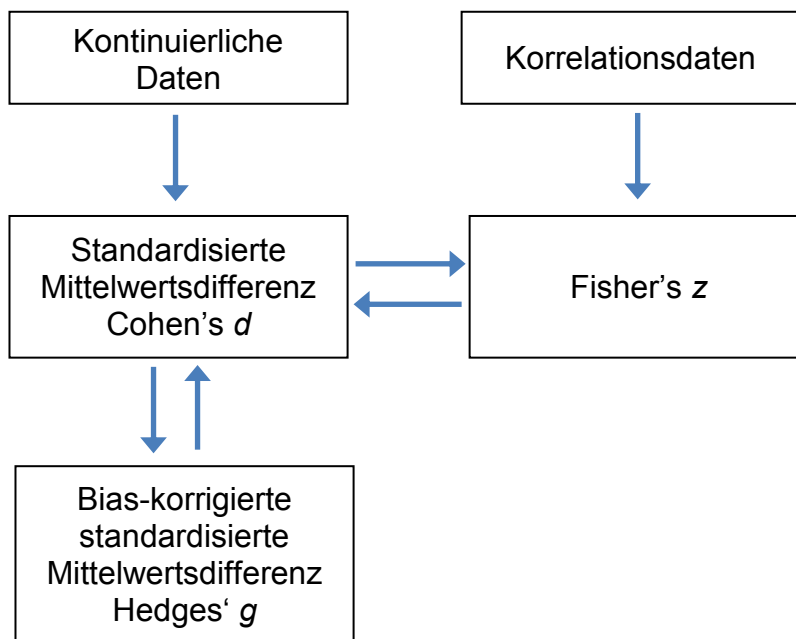


Abbildung 11. Umrechnen von Effektstärken. Nach Borenstein et al. (2009, S. 46)

Die Umrechnung von r in d soll am Beispiel der Studie von Exner, Martin und Rief (2009) illustriert werden. Exner et al. (2009) untersuchten zwei Gruppen (19 Zwangspatienten und 19 Gesunde) u. a. in Bezug auf Kontrollzwang und Gedächtnisleistung. Die Autoren erfassten die Ausprägungen beider Gruppen auf der Symptomdimension *Kontrollzwang* der deutschen Version des PI-R (Emmelkamp & van Oppen, 2000, zitiert nach Exner et al., 2009) und die verbale Gedächtnisleistung mit der WMS-R LM (Wechsler, 1987; zitiert nach Exner et al., 2009). Der mittlere Wert auf der Dimension *Kontrollzwang* lag bei den Zwangspatienten bei 13.1 ($SD = 8.3$), bei den Gesunden bei 4.9 ($SD = 3.3$). Als Ergebnis des LM-Gedächtnistests gaben die Autoren nicht die Gruppenmittelwerte an. Stattdessen berichteten sie die Korrelation ($r = -.62$, $p < .005$) zwischen den *Kontrollzwang*-Scores aller 38 Teilnehmer mit den LM-Scores aller 38 Teilnehmer. Die Autoren hatten die Variable *Kontrollzwang* a priori „dichotomisiert“, indem sie zwei Gruppen, Zwangspatienten und Gesunde, gebildet hatten. Aus den Angaben der Autoren (*Kontrollzwang*-Scores für beide Gruppen und Korrelation) kann die

Differenz der LM-Gruppenmittelwerte berechnet werden. Etwas vereinfacht³ heißt dies: Wenn in der Primärstudie aus zwei gemessenen Variablen (den PI-R-Kontrollzwang-Scores und den LM-Testscores) eine Unbekannte (die Korrelation) berechnet werden kann, dann können in der Metaanalyse durch Umformen dieser Gleichung aus den PI-R-Kontrollzwang-Mittelwerten und der Korrelation die LM-Gruppenmittelwerte berechnet werden. Die standardisierte Differenz aus den beiden LM-Gruppenmittelwerten ergibt dann Cohen's d , das in Hedges' g überführt werden kann. Auf eine kurze Formel (Borenstein, 2009, S. 48) gebracht, wird d aus r folgendermaßen ermittelt:

$$d = \frac{2r}{\sqrt{1-r^2}}$$

Einsetzen der Korrelation $r = - .62$ aus o. g. Beispiel (Exner et al., 2009) ergibt:

$$d = \frac{2 * - 0.62}{\sqrt{1 - (- 0.62)^2}} = - 1.58$$

Die Formel zeigt, dass die standardisierte Mittelwertsdifferenz einem mehr als doppelt so großen Effekt entspricht wie der Korrelationskoeffizient.

Für alle eingeschlossenen Studien wurden Effektstärken getrennt für die einzelnen neuropsychologischen Tests sowie Gesamt-Effektstärken für die übergeordneten neuropsychologischen Domänen und Subdomänen ermittelt. Gab eine Primärstudie für einen neuropsychologischen Test mehrere Ergebnisse an, dann wurde das übergeordnete Testergebnis (z. B. der Gesamtwert des Tests) verwendet. Wurden für eine neuropsychologische Subdomäne mehrere Testverfahren eingesetzt, dann wurde eines dieser Testverfahren zufällig ausgewählt (Woods et al., 2002). Die

³ Verzichtet wurde in dieser vereinfachten Darstellung auf den Zwischenschritt der Transformation von Korrelationskoeffizienten in Fisher's z , um den Wertebereich der Korrelation von -1 bis 1 in einen Wertebereich von $-\infty$ bis ∞ umzuwandeln und damit eine Differenzbildung zu ermöglichen. Auch das Problem, dass einzelne Rohwerte unstandardisiert sind, wohingegen r und d bereits standardisiert sind, wurde hier um einer besseren Verständlichkeit willen nicht berücksichtigt.

Bewertung der Größe der Effektstärken erfolgte gemäß Cohen (1992, S. 157): „the small, medium, and large ESs (or H_1 s) are $d = .20$, $.50$, and $.80$ “.

Gewichtung der Effektstärken. Studienergebnisse großer Stichproben weisen eine höhere Teststärke und geringere Varianz auf als Studienergebnisse kleinerer Stichproben. Je kleiner das N einer Einzelstudie ist, desto größer fällt der Standardfehler ihrer Effektstärke aus. Daher sind Ergebnisse großer Stichproben zuverlässiger. Entsprechend wurden die Effektstärken nach dem Stichprobenumfang gewichtet. Als Streuungsmaß wurde das inverse Varianzgewicht ermittelt, das auf dem Standardfehler der Effektstärke basiert (Lipsey & Wilson, 2001).

Effektstärken lassen sich grafisch in einem *forest plot* darstellen (Bortz & Döring, 1995; Hunter & Schmidt, 1995, vgl. Abbildung 12). Schwarze Quadrate symbolisieren Effektstärken für einzelne Studien, die Querlinie bezeichnet das jeweilige Konfidenzintervall. Die Größe der Quadrate gibt das Gewicht der Studie an. Die Gesamteffektstärke aller Studien wird durch die weiße Raute dargestellt, deren Breite dem Konfidenzintervall des Gesamteffektes entspricht.

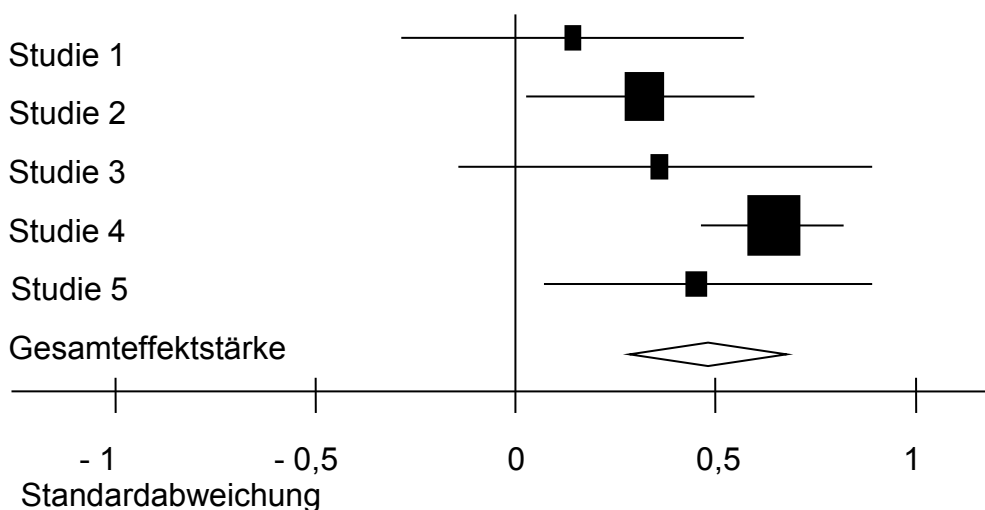


Abbildung 12. Schema eines *forest plots* nach Borenstein et al. (2009, S. 367).

Heterogenitätsprüfung der Effektstärken. Eine Zusammenfassung von Einzeleffekten zu einer Gesamteffektstärke ist nur dann sinnvoll interpretierbar, wenn die Effektstärken homogene Schätzungen eines gemeinsamen Populationsmerkmals darstellen (Lipsey & Wilson, 2001). Dies ist bei der Betrachtung neuropsychologischer Leistungen von Zwangspatienten mit verschiedenen Symptomschwerpunkten nicht ohne weiteres der Fall, da es sich um zahlreiche Konstrukte und noch zahlreichere Testverfahren handelt, die überdies bei verschiedenen Patientengruppen untersucht wurden. Die Heterogenität der einzelnen Studienergebnisse kann einerseits aufgrund inhaltlicher Überlegungen angenommen werden, andererseits anhand der statistischen Kennwerte Q , T und I^2 überprüft werden (Higgins & Thompson, 2002; Huedo-Medina, Sánchez-Meca, Marín-Martínez & Botella, 2006).

Die Statistik Q prüft als Signifikanztest die Homogenität der Verteilung der Studieneffektstärken (Cochran, 1954; Hedges & Olkin, 1985). Q bezeichnet das Verhältnis der beobachteten Varianz der Effektstärken zwischen den Studien zur Fehlervarianz innerhalb der Studien. Dem Q -Test liegt die Nullhypothese einer homogenen Verteilung der Studieneffektstärken zugrunde. Dies entspricht der Annahme, dass alle Studien eine gemeinsame Effektstärke teilen und die beobachtete Varianz ausschließlich Fehlervarianz darstellt. Wird Q signifikant, dann ist anzunehmen, dass die Effekte der Einzelstudien heterogen sind. Wird Q hingegen nicht signifikant, dann erlaubt dies nicht zwingend den Schluss, dass die Effekte homogen seien. Stattdessen kann niedrige Teststärke aufgrund geringer Anzahl der Primärstudien zur fälschlichen Annahme der Nullhypothese führen. Denn Q ist, wie alle Signifikanztests, von der Stichprobengröße bzw. von der Anzahl der Primärstudien abhängig.

Die Statistik T ist ein Schätzer für die Streuung τ (Tau) der wahren Effektstärken. T ist ein absolutes Maß in derselben Einheit wie die Effektstärken, d. h., in Standardabweichungen. Je größer T , desto größer ist das Konfidenzintervall, in das die wahren Effekte fallen würden, und desto heterogener ist die Verteilung der Effektstärken. T ist von der Anzahl der Primärstudien unabhängig.

Die Statistik I^2 ist ein deskriptives Verhältnismaß, das das Verhältnis der wahren Heterogenität zur Gesamtvarianz beschreibt. I^2 gibt an, wie viel Prozent der beobachteten Gesamtvarianz auf wahre Heterogenität zurückgehen. Je größer I^2 , desto größer der Varianzanteil in Prozent, der durch tatsächliche Unterschiede zwischen den Studien beschrieben wird. I^2 lässt sich analog zum regressionsanalytischen Kennwert R^2 (Anteil an der Gesamtvarianz, der durch die Kovariaten erklärt wird) betrachten. I^2 ist, im Gegensatz zu Q , von der Anzahl der Primärstudien unabhängig.

T und I^2 unterscheiden sich voneinander. Während T die absolute Größe der Streuung der Effektstärkenverteilung in Standardabweichungen angibt, trifft I^2 eine Aussage über den relativen, auf Heterogenität zurückgehenden Varianz-Anteil in Prozent. Wenn inhaltliche Überlegungen für Heterogenität sprechen und diese durch die Heterogenitätstests bestätigt werden (d. h., wenn Q signifikant wird und T und I^2 hohe Werte annehmen), dann sind die Effekte differenziert zu beurteilen (Higgins & Thompson, 2002; Huedo-Medina et al., 2006) und man sollte versuchen, die Heterogenität anhand von Subgruppenanalysen oder Metaregression zu erklären (vgl. Abschnitte 4.5 und 4.6).

4.4 Integration der Effektstärken

Sind die Effektstärken ausreichend homogen, so dass sie zu einem mittleren Gesamteffekt zusammengefasst werden können, dann kann dies anhand von zwei Integrationsmodellen, dem *fixed effect model* und dem *random effects model* erfolgen.

Das *fixed effect model* zeigt eine gute Passung auf homogene Primärstudien. Hierbei nimmt man an, dass die „wahre“ Effektstärke bei allen Einzelstudien identisch ist und dass die errechneten Einzeleffektstärken nur aufgrund des zufälligen Stichprobenfehlers variieren. Daher kann man bei der Gewichtung der Einzelstudien die relativ ungenaue Information der kleineren Studien weitestgehend ignorieren, da dieselbe Information, der „feste Effekt“, von den größeren Studien präziser abgebildet wird. Entsprechend unterscheiden sich im *fixed effect model* die

Einzelstudien in ihrer Gewichtung stark voneinander. D. h., Studien mit großen Stichproben werden aufgrund ihrer höheren Präzision in diesem Modell stärker gewichtet als Studien mit kleinen Stichproben.

Das *random effects model* hingegen zeigt eine bessere Passung bei heterogener Studienlage. Hierbei schätzt man nicht einen „wahren“ Effekt, sondern den Mittelwert einer Verteilung unterschiedlicher Effektstärken. Da in diesem Falle jede Studie eine andere Information über „ihre“ spezifische Effektstärke enthält, werden auch Studien mit kleinen Stichproben relativ hoch gewichtet. Dann kann die Information über diese Effektstärke zwar ungenau sein, doch ist sie in keiner anderen Studie enthalten und daher bedeutsam. Entsprechend erhält auch eine Studie mit sehr großem N kein allzu hohes Gewicht, weil der Mittelwert aller Effekte dadurch überschätzt würde. Das heißt, dass im *random effects model* die Gewichte der Studien in ihrer Bedeutung für die Metaanalyse in einem engeren Intervall liegen (Borenstein, 2009). Aufgrund der Heterogenität der neuropsychologischen Konstrukte und Tests wurde in den nachfolgenden Metaanalysen zunächst das *random effects model* gewählt. Bei ausreichender Homogenität wurden zusätzlich Effektstärken nach dem *fixed effect model* berechnet.

Gemäß den Empfehlungen Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) von Moher, Liberati, Tetzlaff und Altman (2009) wurden Effektstärken für die einzelnen neuropsychologischen Tests sowie für jede kognitive Domäne bzw. Subdomäne berechnet und in einem *forest plot* dargestellt. Für anschließende Moderator-, Metaregressions- und Publikationsbias-Analysen wurde aus jeder Primärstudie eine Effektstärke zufällig ausgewählt.

4.5 Moderatoranalysen

Je heterogener die Studienbefunde sind (Q wird signifikant und T sowie I^2 nehmen hohe Werte an, d. h., eine Gesamteffektstärke ist nicht sinnvoll interpretierbar), desto sinnvoller werden Moderatoranalysen. Diese werden durchgeführt, indem die Primärstudien in distinkte Subgruppen aufgeteilt werden. Für jede dieser Subgruppen wird eine Gesamteffektstärke berechnet. Dann werden

die Gesamteffektstärken der Subgruppen wie Effektstärken von Primärstudien behandelt und anhand des Q -Signifikanztests miteinander verglichen (Borenstein 2009). Hierbei lautet die Nullhypothese, dass die wahre mittlere Effektstärke in allen Subgruppen gleich ist. Wird die Teststatistik Q signifikant, dann wird die Nullhypothese gleicher Subgruppen verworfen.

Da möglicherweise nicht alle Testverfahren kulturfair sind, wurde in der vorliegenden Metaanalyse der Kulturkreis der Untersuchung (westlich-angelsächsisch vs. östlich-asiatisch) als Moderatorvariable angenommen. Zudem wurden mögliche Unterschiede im Datenformat der Primärstudien (Mittelwerte und Standardabweichungen vs. Korrelationskoeffizienten) als Moderator untersucht.

4.6 Metaregression

Analog zur Regressionsanalyse in Primärstudien lässt sich auch in Metaanalysen eine Metaregression durchführen. Während die Regressionsanalyse in Primärstudien die Beziehung zwischen Studienergebnis und Kovariaten untersucht, erfasst die Metaregression die Beziehung zwischen Effektstärke und Kovariaten.

In der vorliegenden Arbeit wurden demografische und klinische Variablen, die die neuropsychologischen Leistungen beeinflussen können, als Kovariaten mit aufgenommen. Für alle Metaanalysen wurden standardisierte Mittelwertsdifferenzen von Alter, Intelligenzniveau/Bildungsgrad sowie die Studiengüte (Gesamtscore der Qualitätsskala) als Kovariaten in die Metaregression aufgenommen. Dem lag die Annahme zugrunde, dass sich die Einzelstudien in diesen Variablen zufällig unterscheiden.

Für die Metaanalysen, die je zwei klinische Stichproben miteinander verglichen, wurden zusätzlich folgende Variablen als Kovariaten mit aufgenommen:

- komorbide Depressivität: Standardisierte Mittelwertsdifferenzen der Scores des Beck Depressions-Inventars (BDI; Hautzinger, Bailer, Worall & Keller (1993/1996), des Beck Depressions-Inventar Revision (BDI-II; Hautzinger, Keller & Kühner, 1996/2006), der Hamilton Rating Scale for Depression

(HAM-D; Hamilton, 1960) und der Montgomery-Åsberg-Scale for Depression (MADRS; Montgomery & Åsberg, 1979)

- komorbide Angst: Standardisierte Mittelwertsdifferenzen der Scores des State-Trait-Anxiety-Inventory, Subskala State-Anxiety (STAI-S; Spielberger, Gorsuch & Lushene, 1970), der Hamilton Anxiety Rating Scale (HAM-A; Hamilton, 1959) und des Beck Anxiety Inventory (BAI; Beck, Epstein, Brown & Steer, 1988),
- der Anteil medikamentös behandelter Patienten in Prozent,
- der Schweregrad der Erkrankung (Gesamtscore der Y-BOCS),
- die Krankheitsdauer
- das Alter bei Ersterkrankung.

Eine Metaregression konnte durchgeführt werden, sobald eine Mindestzahl von $k = 4$ Studien vorhanden war (Borenstein, Hedges, Higgins & Rothstein, 2005). Es wurde angenommen, dass die untersuchten Kovariaten einen Teil der Heterogenität zwischen den Primärstudien erklärten, nicht jedoch die ganze Heterogenität. Daher wurden die Metaregressionsanalysen nach einer Variante des *random effects model* unter Einschluss von Kovariaten, dem *mixed effects model*, und der *method of moments* zur Abschätzung der Varianz zwischen den Studien durchgeführt (Kelley & Kelley, 2012). Die Metaregressionen wurden mit der Software CMA (Borenstein et al., 2005) durchgeführt, die keine simultane Aufnahme mehrerer Kovariaten in die Metaregressionsgleichung erlaubt. Daher wurde für jede Kovariate eine separate Metaregression berechnet. Da auch bei der Metaregression multiple Testungen zu einer Erhöhung falsch-positiver Ergebnisse führen können (Borenstein, 2009), wurde das Signifikanzniveau hier auf $\alpha = .01$ adjustiert.

4.7 Abschätzung der Publikationsverzerrung

Studien mit nicht-signifikanten Resultaten bleiben häufig unveröffentlicht. Hingegen werden Studien mit signifikanten Ergebnissen mit höherer Wahrscheinlichkeit veröffentlicht, mehrfach vorgestellt, in Fachzeitschriften mit höherem *impact factor* publiziert (Easterbrook, Gopalan, Berlin & Matthews, 1991),

sie gehen schneller in Druck, werden häufiger auf Englisch veröffentlicht und häufiger zitiert (Egger, 2001). Entsprechend sind sie für den Reviewer einfacher zu finden. Bezieht dieser nur jene Studien mit ein, die veröffentlicht wurden, können die Ergebnisse der Metaanalyse systematisch überschätzt werden. Es entsteht ein Publikationsbias. Diesem kann entgegengewirkt werden, indem auch „graue Literatur“ in Form von unveröffentlichten Studien, Dissertationen, Kongressbeiträgen etc. berücksichtigt wird. Zur Abschätzung des Publikationsbias können zwei *fail-safe N*-Kennwerte sowie ein grafisches Instrument, der *funnel plot*, herangezogen werden.

Das klassische *fail-safe N* (Rosenthal, 1979) schätzt, wie viele unveröffentlichte Studien mit Nulleffekten in die Metaanalyse einfließen müssten, damit die errechnete mittlere Effektstärke statistisch insignifikant wird. Je größer Rosenthals *fail-safe N*, desto geringer die Wahrscheinlichkeit, dass entsprechend viele insignifikante Studien unveröffentlicht blieben. Rosenthals *fail-safe N* trifft allerdings keine Aussage über die praktische Relevanz. D. h., Rosenthals *fail-safe N* gibt nicht an, wie viele verborgene Studien erforderlich wären, um den Effekt so zu verringern, dass er keine praktische Bedeutsamkeit mehr besitzt. Dieses Problem wird durch eine Variante von Rosenthals Formel, Orwins *fail-safe N* (Orwin, 1983) gelöst. Orwins *fail-safe N* gibt an, wie viele Studien mit einem Nulleffekt oder einem Negativeffekt erforderlich wären, um die errechnete mittlere Effektstärke so zu verringern, dass sie praktisch unbedeutsam wird. Ab wann ein Effekt als praktisch unbedeutsam gilt, wird a priori vom Reviewer definiert. Hier wurde ein Effekt von $g = 0.1$ bzw. $g = -0.1$ angenommen. Je größer Orwins *fail-safe N* wird, desto geringer ist die Wahrscheinlichkeit, dass so viele unveröffentlichte Studien mit Null- oder Negativeffekten existieren, dass der Effekt bedeutungslos wird. Ist das *fail-safe N* größer als die Anzahl der gefundenen Studien, dann ist davon auszugehen, dass kein bedeutsamer Publikationsbias vorliegt (Abramovitch et al., 2013).

Während die *fail-safe Ns* eine Aussage darüber treffen, ob der beobachtete Effekt ausschließlich auf einen Publikationsbias zurückzuführen ist, gibt der *funnel plot* an, ob ein Publikationsbias irgendeinen Einfluss auf den beobachteten Effekt

hat. Ein *funnel plot* (Schwarzer & Rücker, 2010) ist ein Streudiagramm in Form eines sich nach oben verengenden Trichters (*funnel*). Für jede publizierte Studie wird auf der x-Achse die Effektstärke und auf der y-Achse der Standardfehler der Effektstärke abgetragen. Je größer die Studienstichprobe, desto höher ist die Präzision ihrer Ergebnisse. Daher befinden sich Studien mit großen Stichproben und präziseren Ergebnissen im oberen Teil der Grafik und nahe der mittleren Effektstärke, Studien mit kleineren Stichproben und größerer Streuung im unteren Teil des Diagramms. Stellt sich der *funnel plot* symmetrisch dar, dann streuen die Resultate aller Studien gleichmäßig um den mittleren Gesamteffekt und es liegt wahrscheinlich kein Publikationsbias vor. Zeigt sich jedoch eine Asymmetrie im *funnel plot* derart, dass einseitig Studien fehlen, dann schließt man auf einen Publikationsbias und sollte den Gesamteffekt entsprechend adjustieren (Schwarzer & Rücker, 2010). Die Adjustierung im *funnel plot* kann über das Verfahren *trim and fill* von Duval und Tweedie (2000) vorgenommen werden. Hierbei wird abgeschätzt, wie viel Einfluss der Bias hatte und wie groß der beobachtete Effekt ohne den Bias wäre. *Trim and fill* bezeichnet einen iterativen Algorithmus, bei dem im *funnel plot* die extremsten Einzeleffekte entfernt (*trim*) und anschließend mit ihren „Spiegelbildern“ wieder eingefügt (*fill*) werden, so dass bei konstanter Varianz ein adjustierter Effekt entsteht. Dann enthält der *funnel plot* die beobachteten und die ergänzten Einzeleffekte sowie den beobachteten und den adjustierten Gesamteffekt. Je näher der beobachtete und der adjustierte Gesamteffekt beieinander liegen, desto geringer ist der Einfluss des Publikationsbias.

Alle Verfahren zur Abschätzung des Publikationsbias weisen eine relativ geringe Teststärke auf und sind von einer ausreichend hohen Anzahl an Primärstudien abhängig. Zudem ist es auch möglich, dass kleine Studien aus unbekanntem Gründen größere Effektstärken ergeben und dass kein Bias vorliegt. Daher empfiehlt Borenstein (2009), Asymmetrien im *funnel plot* vorsichtig zu interpretieren. Ergänzend ist anzumerken, dass der Publikationsbias zwar ein Problem in Metaanalysen darstellt, jedoch auch in narrativen Reviews und generell bei jeder Literatursuche auftritt.

5 Ergebnisse

Anhand der gefundenen Publikationen ließen sich sieben einzelne Metaanalysen durchführen. Im Folgenden werden die Ergebnisse der Literaturrecherche (Abschnitt 5.1), die Merkmale der eingeschlossenen Studien (Abschnitt 5.2), die Studienqualität (Abschnitt 5.3) und die Merkmale der Studienteilnehmer (Abschnitt 5.4) vorgestellt. Hieran schließen sich die Ergebnisse der einzelnen Metaanalysen (Abschnitte 5.5 bis 5.11) sowie eine Übersicht über die kombinierten Effekte über alle sieben Metaanalysen hinweg (Abschnitt 5.12) an. Alle Metaanalysen wurden prinzipiell nach der gleichen Methodik durchgeführt, alle Signifikanztests wurden zweiseitig mit einem Signifikanzniveau von $\alpha = 0.05$ getestet. Die einzige Ausnahme bildeten die Meta-Regressionsanalysen, die zur Korrektur der multiplen Testungen auf einem Signifikanzniveau von $\alpha = 0.01$ getestet wurden. Die Vorgehensweise ist bei der ersten Metaanalyse ausführlich beschrieben. In den Metaanalysen Nr. 2 bis Nr. 7 werden nur diejenigen Schritte beschrieben, die davon abweichen.

5.1 Literaturrecherche

Die systematische Suche wurde in zwei Wellen durchgeführt und ergab 6 069 Referenzen. Davon wurden in der ersten Recherchewelle (Publikationen von 1948 bis zum 31. Dezember 2011) 3 620 Referenzen gefunden. In der zweiten Recherchewelle (1. Januar 2012 bis zum 10. Januar 2014) wurden 2 449 Referenzen gefunden. Die Anzahl der Veröffentlichungen bis Januar 2014 stieg exponentiell an (vgl. Abbildung 13).

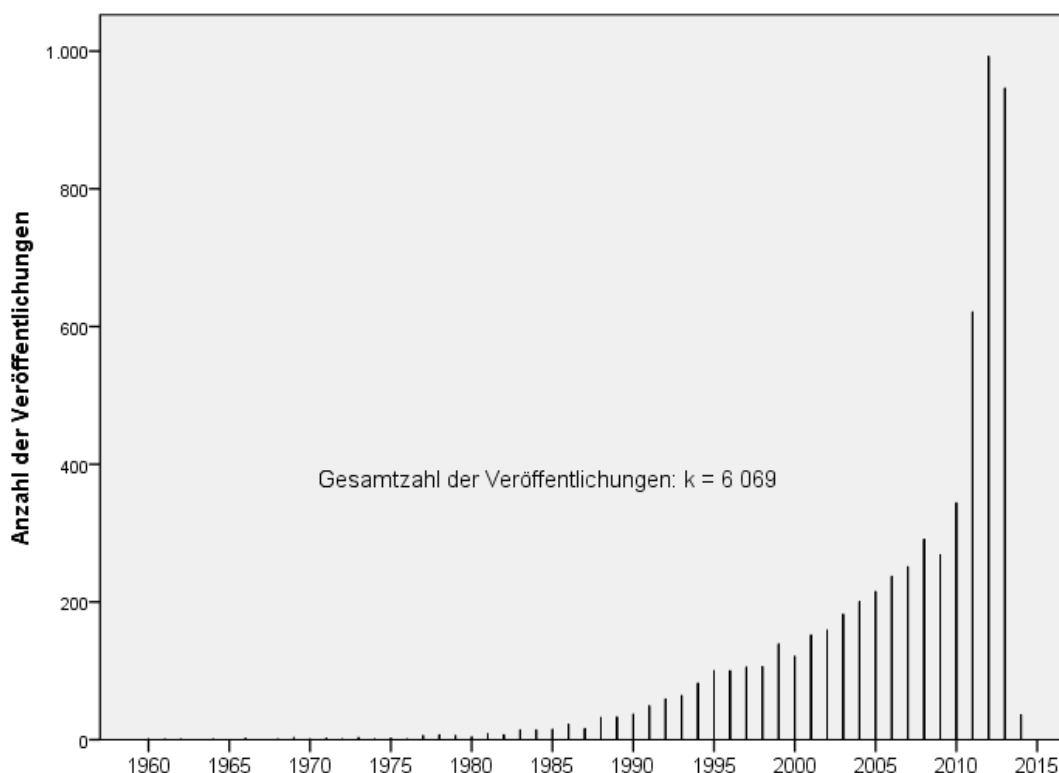


Abbildung 13. Anzahl gefundener Veröffentlichungen. Da die Recherche am 10. Januar 2014 beendet wurde, sind für das Jahr 2014 nur die Publikationen der ersten zehn Tage erfasst.

Es wurden 6 037 Titel über Datenbanken gefunden, 20 Titel über die Handsuche und 12 Titel über E-Mail-Kontakt zu Autoren. Von 17 angeschriebenen Wissenschaftlern antworteten 14, wobei keine der Antworten eine unveröffentlichte Studie enthielt, die hätte eingeschlossen werden können. Die gefundenen Titel wurden einem mehrstufigen Auswahlprozess unterzogen. Im ersten Schritt wurden die Titel geprüft und 3 838 Publikationen, die nicht den Einschlusskriterien genügten, ausgeschlossen. Im zweiten Schritt wurden die Abstracts der verbleibenden Studien geprüft und 1 455 Publikationen ausgeschlossen. Die verbleibenden 497 Publikationen wurden als Volltext besorgt, ebenfalls geprüft und 459 Publikationen ausgeschlossen. Hierzu gehörten 13 Studien, deren Autoren auch auf Nachfrage keine adäquaten Daten zur Verfügung stellen konnten. Letztlich wurden 38 Studien eingeschlossen.

Da die Schlagworte für die Recherche sehr breit gewählt waren, um möglichst keine relevante Studie zu übersehen, wurden bereits aufgrund des Titels fast zwei Drittel der gefundenen Publikationen wieder ausgeschlossen. In dieser ersten Phase der Literatursichtung war der häufigste Ausschlussgrund „keine Zwangsstörung“. Einige Schlagwortkombinationen erbrachten thematisch weit entfernte Resultate: So fand beispielsweise „hoarding AND memory“ Studien zum Sammelverhalten von Nagetieren und Singvögeln, während „wash* AND memory“ historische Untersuchungen zum ersten US-amerikanischen Präsidenten George Washington und zum Regierungssitz der USA (Washington D.C.) ergab. In der zweiten Phase der Literatursichtung wurden von den gelesenen Abstracts ca. drei Viertel ausgeschlossen, hier waren die häufigsten Ausschlussgründe „keine neuropsychologischen Tests“ und „methodische Gründe“. Zu letzteren gehörten Studiendesigns, die keine Fall-Kontroll-Studien oder keine RCTs darstellten (Studien ohne Vergleichsgruppe, Einzelfallstudien, Reviews), Publikationen z. B. auf Chinesisch oder Persisch (deren Autoren auf Nachfrage keine englische Übersetzung zur Verfügung stellten), Studien mit nicht verwertbaren Statistiken (nonparametrische Tests, Grafiken) und Studien, deren Autoren nicht auf Anfragen antworteten oder bei denen die aktuellen Kontaktdaten nicht herauszufinden waren. Von den im Volltext besorgten Studien wurden über 90 Prozent ausgeschlossen, zumeist weil keine Symptomdimensionen der Zwangsstörung spezifiziert worden waren. Die Ergebnisse der Literaturrecherche sowie der Entscheidungsprozess, welche Publikationen aufgrund welcher Kriterien ein- bzw. ausgeschlossen wurden, werden im Folgenden in einem Flussdiagramm (Abbildung 14) nach den PRISMA Standards (Moher et al., 2009) dargestellt.

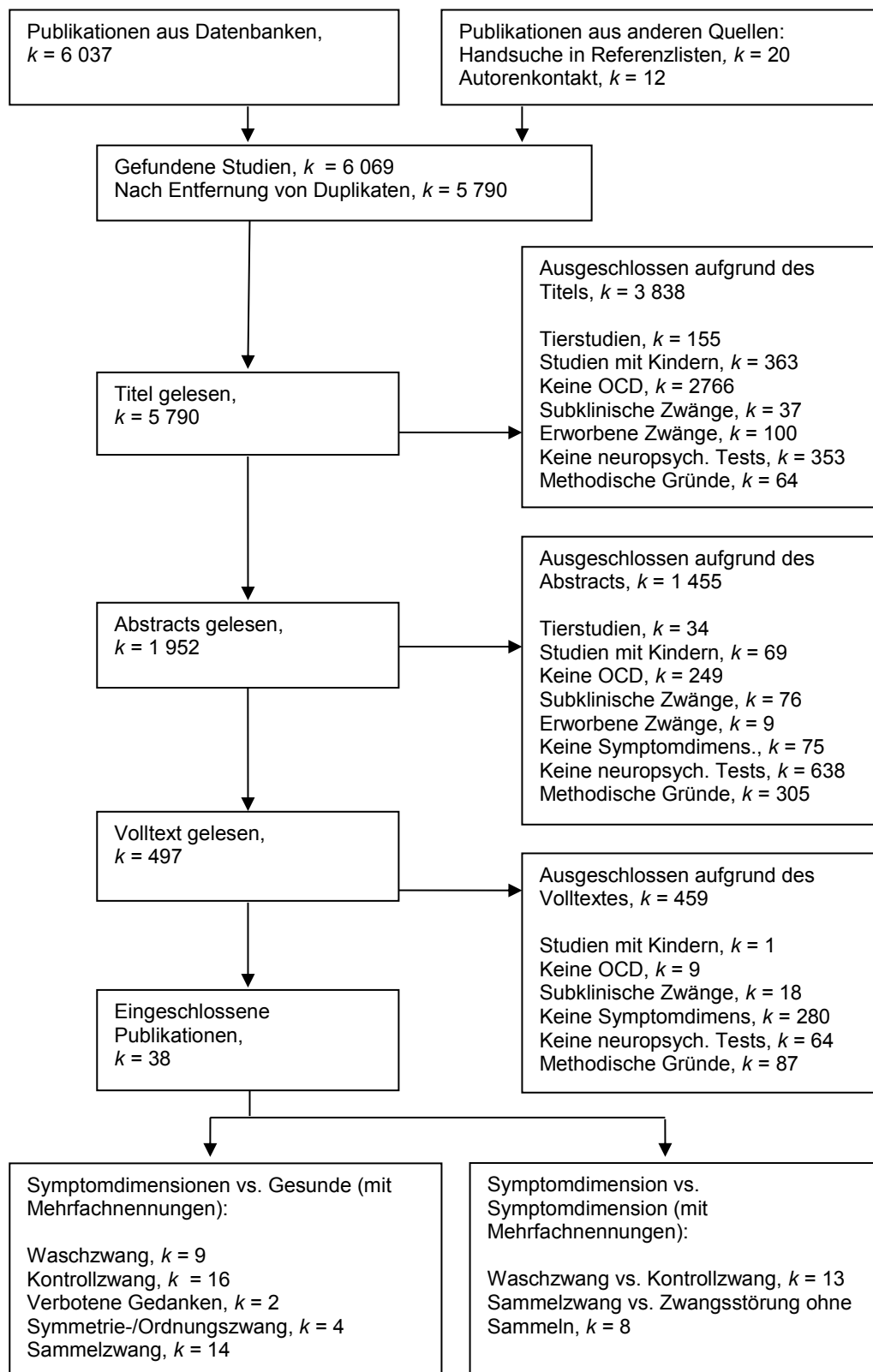


Abbildung 14. Flussdiagramm mit ein- und ausgeschlossenen Primärstudien.

5.2 Studienmerkmale

Es wurden $k = 38$ Primärstudien eingeschlossen. Die erste Welle ergab 23, die zweite Welle 15 eingeschlossene Studien. 37 Studien wurden in Datenbanken und eine Studie per Handsuche gefunden. Sechs der 38 Studien waren graue Literatur: Zwei Dissertationen (Hayward, 2012; Kraft, 2006), eine Masterarbeit (van der Meulen, 2013) sowie zusätzliche, unveröffentlichte Studienergebnisse, die die Autoren dreier Publikationen (Exner et al., 2009; Jang et al., 2010; Tumkaya et al., 2013) per E-Mail zur Verfügung stellten. Die älteste Studie stammte aus dem Jahr 1997, die jüngste aus dem Jahr 2014. Die Zahl der relevanten Veröffentlichungen nahm mit den Jahren stetig zu: Wurden in der Dekade 1991 – 2000 nur vier Studien und in der Dekade 2001 – 2010 schon 17 Studien publiziert, so waren es in den drei Jahren 2011 – 2014 ebenfalls 17 Studien (vgl. Abbildung 15).

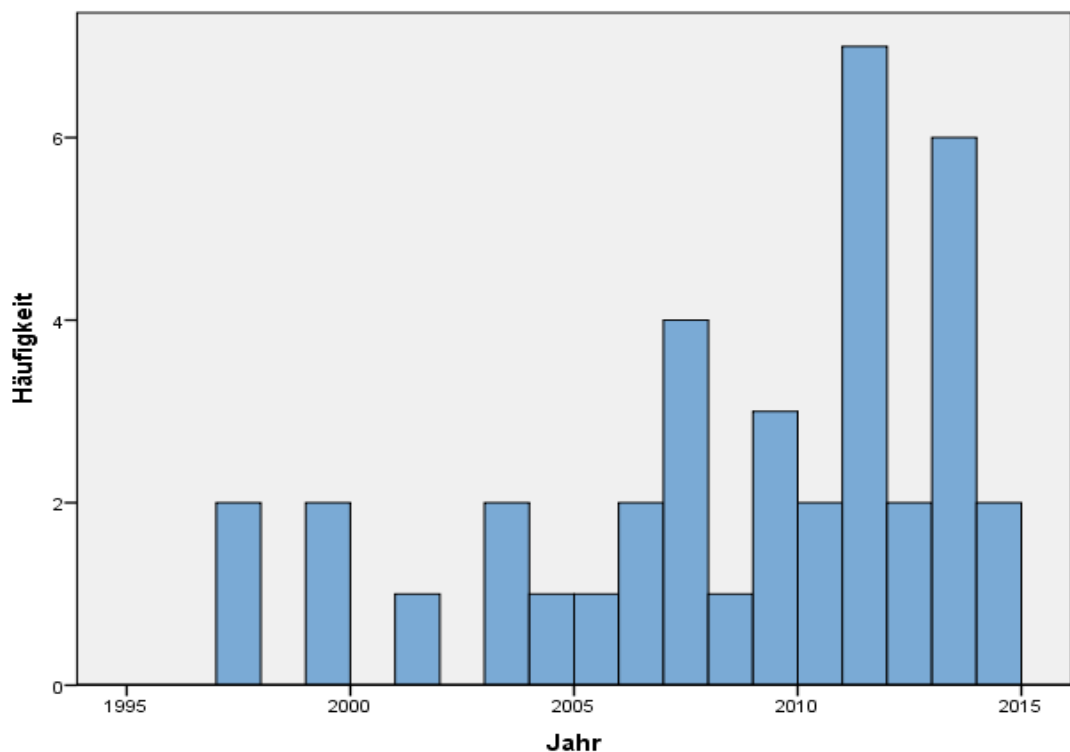


Abbildung 15. Publikationsjahre der eingeschlossenen Studien bis Januar 2014.

Aus dem westlich-angelsächsischen Kulturkreis stammten 30 Studien (78,9 %), aus dem asiatischen Kulturkreis acht Studien (21,1 %).

Die Diagnose einer Zwangsstörung wurde in allen Studien nach der dritten, vierten oder fünften Ausgabe des Diagnostic and Statistic Manual of Mental Diseases gestellt. Der Schweregrad der Zwangserkrankung wurde in 36 Studien mit validen Zwangsskalen, in einer Studie mit einem selbst entwickelten Instrument (Wincze et al., 2007) und in einer Studie mit dem SI-R (Frost et al., 2004) beurteilt. Komorbide Depression wurde in 34 der 38 Studien, komorbide Angst in 24 Studien mit mindestens einem bewährten psychometrischen Instrument erfasst. In vier Publikationen fanden sich keine Angaben über die Erfassung komorbider Störungen. Die Intelligenz der Teilnehmer wurde in 21 Studien erhoben, wobei häufig die verbale Testleistung als Schätzung für die globale Intelligenz diente.

Während 31 Studien Mittelwerte und Standardabweichungen angaben, berichteten sieben Studien Korrelationskoeffizienten (Exner et al., 2009; Grisham et al., 2007; Hashimoto et al., 2011; Jang et al., 2010; Marsh et al., 2014; Pinto et al., 2011; Tumkaya et al., 2013). Eine Übersicht über die Merkmale der 38 eingeschlossenen Studien bietet Tabelle 3.

Tabelle 3. *Eingeschlossene Studien, k = 38.*

Studie, Land	Stichprobe	Diagnose und Komorbidität	Kognitive Domäne und Intelligenztest
Ayers et al. (2013) USA	42 HOARD 25 HC	DSM-5 SI-R	Aufmerksamkeit Exekutivfunktionen
Blom et al. (2011) USA	17 OCD, 24 HOARD, 19 HC	DSM-IV OCI-R SI-R	Gedächtnis Exekutivfunktionen Entscheidungsverhalten Shipley-2
Bouvard et al. (1997) Frankreich	25 OCD (11 WASH, 14 CHECK), 20 HC	DSM-III Y-BOCS BDI	Aufmerksamkeit Gedächtnis Raven Matrizen, WAIS-R verbal
Ceschi et al. (2003) Schweiz	32 OCD (16 WASH, 16 CHECK), 16 HC	DSM-IV Y-BOCS BDI-II, STAI	Gedächtnis
Cha et al. (2008) Südkorea	47 OCD (23 WASH, 24 CHECK), 20 HC	DSM-IV Y-BOCS BDI, HAM-D, HAM-A	Gedächtnis Exekutivfunktionen WAIS-R verbal
Clair et al. (2013) Frankreich	40 OCD (23 CHECK, 17 OTHER OCD), 41 HC	DSM-IV Y-BOCS HADS, STAI	Gedächtnis
Dittrich et al. (2011) GB	68 OCD (22 WASH, 14 SYMM, 32 OTHER OCD), 66 HC	DSM-IV Y-BOCS DY-BOCS MADRS, STAI	Arbeitsgedächtnis NART
Dittrich et al. (2013) Deutschland	13 OCD 13 HC	DSM-IV Y-BOCS MADRS, STAI	Exekutivfunktionen NART

Studie, Land	Stichprobe	Diagnose und Komorbidität	Kognitive Domäne und Intelligenztest
*°Exner et al. (2009) Deutschland	19 OCD (13 CHECK, 6 OTHER OCD), 19 HC	DSM-IV Y-BOCS BDI, BSI	Exekutivfunktionen Gedächtnis Verarbeitungsgeschwind. Arbeitsgedächtnis WAIS-R
°Grisham et al. (2007) Australien	30 HOARD 30 OAD, 30 HC	DSM-IV OCI-R BDI, BAI	Exekutivfunktionen Arbeitsgedächtnis WASI
Grisham et al. (2010) Australien	19 HOARD, 17 OAD, 20 HC	DSM-IV OCI-R DASS,	Exekutivfunktionen Entscheidungsverhalten Kategorisierung WASI
Hartl et al. (2004) USA	22 HOARD, 24 HC	DSM-IV Y-BOCS BDI-II, BAI	Gedächtnis WASI
°Hashimoto et al. (2011) Japan	63 OCD (14 WASH, 10 CHECK, 1 SYMM, MULT)	DSM-IV Y-BOCS BDI-II, STAI	Exekutivfunktionen Gedächtnis Verarbeitungsgeschwi. Arbeitsgedächtnis
*Hayward et al. (2012) USA	21 HOARD 21 HC	DSM-IV OCI-R SI-R, CIR	Kategorisierung
°Jang et al. (2010) Südkorea	144 OCD, 144 HC	DSM-IV Y-BOCS BDI, BAI	Exekutivfunktionen, Gedächtnis WAIS-R
*Kraft (2006) Deutschland	36 OCD, 36 HC	DSM-IV Y-BOCS BDI	Arbeitsgedächtnis MWT-B
Lawrence et al. (2006) Großbritannien	39 OCD (20 HOARD, 19 OTHER OCD), 40 HC	DSM-IV Y-BOCS BDI, STAI	Exekutivfunktionen Entscheidungsverhalten NART

Studie, Land	Stichprobe	Diagnose und Komorbidität	Kognitive Domäne und Intelligenztest
MacDonald et al. (1997) Kanada	10 OCD CHECK, 10 OTHER OCD, 10 HC	DSM-IV Y-BOCS BDI, STAI-S	Gedächtnis SILS
°Marsh et al. (2014) USA	22 OCD (3 CHECK, 6 WASH, 5 SYMM, 7 AGGR, 1 HOARD) 22 HC	DSM-IV Y-BOCS HAM-D	Exekutivfunktionen WAIS
Morein-Zamir et al. (2014) GB	24 OCD 22 HOARD 28 HC	DSM-5 Y-BOCS OCI-R, SI-R MADRS BDI, STAI	Exekutivfunktionen Entscheidungsverhalten NART
Moritz et al. (2012) Deutschland	60 OCD (10 CHECK, 50 OTHER OCD)	DSM-IV Y-BOCS HAM-D	Aufmerksamkeit Gedächtnis Verarbeitungsgeschi. Arbeitsgedächtnis
Nakao et al. (2009) Japan	40 OCD (10 CHECK, 14 WASH, 16 MULT)	DSM-IV Y-BOCS HAM-D, STAI	Aufmerksamkeit Exekutivfunktionen Gedächtnis Arbeitsgedächtnis WAIS-R
Nedeljkovic et al. (2009) Australien	59 OCD (12 WASH, 11 CHECK, 12 AGGR, 24 MULT), 59 HC	DSM-IV Y-BOCS BDI, BAI, HAM-D, HAM-A	Exekutivfunktionen Arbeitsgedächtnis NART
Omori et al. (2007) Japan	53 OCD (27 CHECK, 26 WASH)	DSM-IV Y-BOCS BDI-II, STAI	Exekutivfunktionen Gedächtnis Verarbeitungsgeschi.

Studie, Land	Stichprobe	Diagnose und Komorbidität	Kognitive Domäne und Intelligenztest
Penadés et al. (2007) Spanien	27 OCD (10 WASH, 7 CHECK, 10 OTHER)	DSM-IV MOCI BDI	Exekutivfunktionen WAIS-III
°Pinto et al. (2011) Brasilien	32 OCD	DSM-IV Y-BOCS DY-BOCS BDI, BAI	Exekutivfunktionen WASI
Radomsky & Rachman (1999) Kanada	10 OCD WASH, 10 OAD, 20 HC	DSM-IV MOCI BDI	Gedächtnis Arbeitsgedächtnis
Singh et al. (2003) Indien	18 OCD (8 CHECK, 10 OTHER OCD), 10 HC	ICD-10 Y-BOCS MADRS, HAM-A	Arbeitsgedächtnis
Tallis et al. (1999) Großbritannien	12 OCD CHECK, 12 HC	DSM-IV PI BDI, BAI, STAI	Gedächtnis NART
Tolin et al. (2011) USA	27 HOARD, 12 OTHER OCD, 26 HC	DSM-IV Y-BOCS DASS	Exekutivfunktionen Gedächtnis WTAR
Tolin & Villavicencio (2011) USA	29 OCD 42 HOARD 36 HC	DSM-IV OCI-R HRS-I, SI-R HAM-D	Entscheidungsverhalten
Tolin, Stevens, Villavicencio et al. (2012) USA	31 OCD 43 HOARD 33 HC	DSM-5 OCI-R HRS-I, SI-R HAM-D	Entscheidungsverhalten

Studie, Land	Stichprobe	Diagnose und Komorbidität	Kognitive Domäne und Intelligenztest
Tolin et al. (2014) USA	24 OCD 24 HOARD 24 HC	DSM-5 OCI-R HRS-I, SI-R HAM-D	Exekutivfunktionen
°Tumkaya et al. (2013) Türkei	42 OCD 42 HC	DSM-IV Y-BOCS MOCI HAM-D, HAM-A	Aufmerksamkeit
Tuna et al. (2005) Türkei	17 OCD CHECK, 16 subklin. CHECK, 15 HC	DSM-IV Y-BOCS BDI, STAI	Gedächtnis
*Van der Meulen (2013) Benelux	19 OCD 22 HOARD 22 HC	DSM-IV OCI-R, SI-R BDI, BAI	Exekutivfunktionen NART
Wincze et al. (2007) USA	21 OCD, 21 HOARD, 21 HC	DSM-IV Selbst entw. Skala BDI, STAI	Kategorisierung
Zitterl et al. (2001) Österreich	27 OCD CHECK, 27 HC	DSM-IV Y-BOCS HAM-D	Gedächtnis

Anmerkungen: BAI = Beck Anxiety Inventory (Beck et al., 1988; zitiert nach Grisham et al., 2007); BDI = Beck Depression Inventory (Beck & Steer, 1993; zitiert nach Cha et al., 2008); BDI-II = Beck Depression Inventory-Revised (Beck, Steer & Brown, 1996; zitiert nach Hartl et al., 2004); CHECK = Kontrollzwang; DASS = Depression Anxiety Stress Scale (Lovibond & Lovibond, 1995, zitiert nach Tolin et al., 2011); HADS = Hospital Anxiety Depression Scale (Zigmond & Snaith, 1983; zitiert nach Clair et al., 2013); HAM-A = Hamilton Rating Scale for Anxiety (Hamilton, 1959); HAM-D = Hamilton Rating Scale for Depression (Hamilton, 1960); HC = healthy controls/gesunde Kontrollpersonen; HOARD = Sammelzwang; MADRS = Montgomery-Åsperg Depression Rating Scale (Montgomery & Åsberg, 1979); MULT = multiple Zwänge; OAD = other anxiety disorder/andere Angststörung; OTHER OCD = andere Zwangssymptome; PI = Padua Inventory (Sanavio, 1988); PI-WSUR = Padua Inventory-Washington State University Revision (Burns, Keortge, Formea & Sternberger, 1996); STAI = State-Trait Anxiety Inventory (Spielberger et al., 1970). MWT-B = Mehrfachwahl-Wortschatztest Teil B

(Lehrl, 2005; zitiert nach Kraft, 2006); NART = National Adult Reading Test (Nelson & Wilson, 1991); Shipley-2 = Shipley Institute of Living Scale-2 (Shipley et al., 2009, zitiert nach Blom et al., 2011); SILS = Shipley Institute of Living Scale (Shipley, 1940); SYMM = Symmetrie-/Ordnungszwang; WASH = Waschwang; WAIS-R = Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981; zitiert nach Hashimoto et al., 2011); WASI = Escala de Inteligencia Wechsler abreviada (Wechsler, 2005; zitiert nach Pinto et al., 2011); WTAR = Wechsler Test of Adult Reading Manual (Holdnack, 2001; zitiert nach Tolin et al., 2011). Mit * markierte Studien sind unveröffentlichte, graue Literatur. Mit ° bezeichnete Studien sind Korrelationsstudien.

5.3 Studienqualität

Für jede Studie wurde ein Gesamt-Summenwert der Skala zur Beurteilung der Studienqualität ermittelt. Da jedes der zehn Items mit null bis zwei Punkten bewertet werden konnte, waren Gesamtwerte zwischen null und 20 möglich. Die Qualität der 38 Studien lag zwischen 6 und 17 Punkten mit $M (SD) = 12.16 (2.76)$, Median = 13 und Modalwert = 13 (vgl. Abbildung 16).

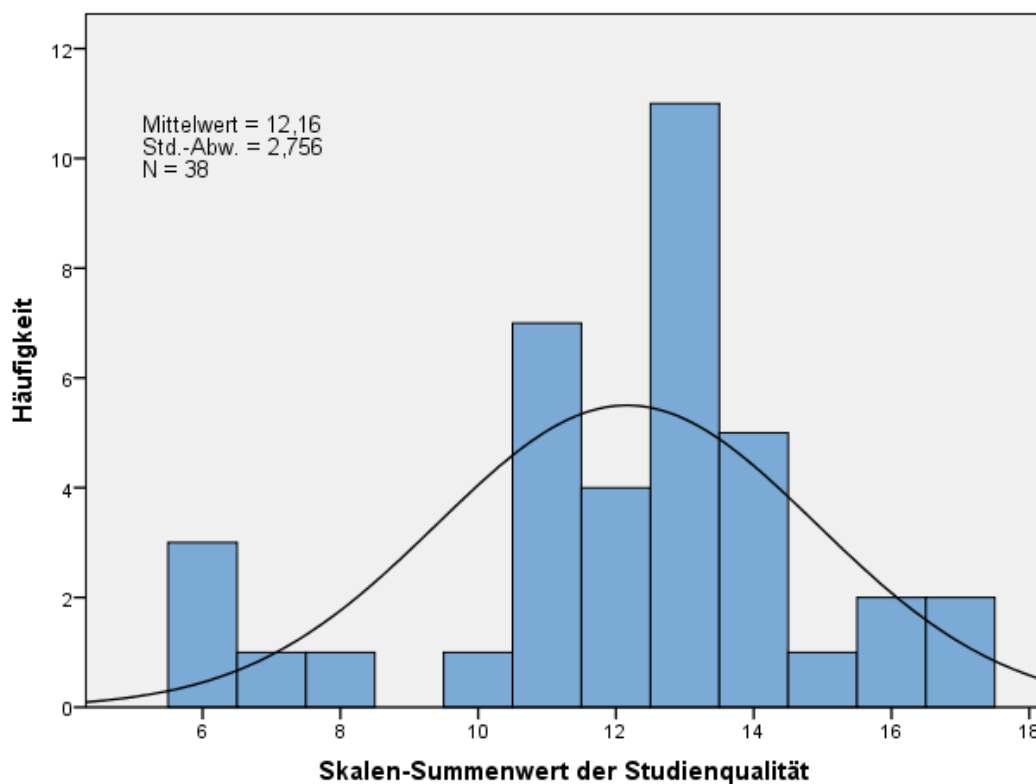


Abbildung 16. Histogramm zur Qualität der eingeschlossenen Studien.

Die einzelnen Items zur Studienqualität erbrachten folgende Ergebnisse: Die Ethikrichtlinien wurden von zehn Studien ausreichend und von weiteren 21 Studien sogar gut erfüllt, indem die Autoren das Einverständnis der Teilnehmer oder darüber hinaus die Genehmigung durch ein Ethikkomitee anführten. Die Klarheit der Stichprobenbeschreibung war bei acht Studien ausreichend, indem demografische Daten, Ein- und Ausschlusskriterien aufgeführt waren, und bei weiteren 29 Studien gut, indem zusätzlich Komorbidität und Medikationsstatus berichtet wurden. Der Schweregrad der Zwangsstörung wurde in 37 Studien gut mit einem standardisierten Instrument wie der Y-BOCS erfasst. Komorbide Störungen wurden in neun Studien ausreichend berichtet, in 19 weiteren Studien wurde Komorbidität darüber hinaus statistisch kontrolliert und damit sogar gut berücksichtigt. Die Reliabilität der Diagnose war in allen 38 Studien ausreichend, indem eines der gängigen Klassifikationssysteme (DSM oder ICD) eingesetzt wurde. Drei Studien ermittelten darüber hinaus die Interrater-Reliabilität der Diagnosen anhand von mehreren Untersuchern und gaben somit eine gute Reliabilität der Diagnose an.

Die Validität der neuropsychologischen Tests war in 13 Studien ausreichend und in 24 Studien gut, d. h., dass in 96,9 % aller Studien anerkannte und bewährte neuropsychologische Testverfahren eingesetzt wurden. Eine Untersuchung (Kraft, 2006) verwendete ein selbst entwickeltes Paradigma, über dessen Gütekriterien nichts bekannt war. Über Anzahl und Training der Testleiter berichtete die Mehrzahl der Studien nichts. Lediglich vier Studien gaben Auskunft über die Anzahl der Testleiter, sechs Studien erwähnten, dass die Testleiter zuvor ein Training erhalten hatten, und keine Studie gab an, welcher Art und Intensität dieses Training war.

Über die psychopharmakologische Behandlung der Teilnehmer wurden in 29 Studien Angaben gemacht, die Einflüsse der Medikation wurden in elf dieser Studien zudem statistisch kontrolliert. Alle eingeschlossenen Studien setzten angemessene Methoden ein, acht Studien präsentierten die Ergebnisse in ausreichender Form, 30 Studien stellten die Ergebnisse in guter Form dar.

5.4 Teilnehmer

Insgesamt nahmen $n = 1\,491$ Patienten sowie $n = 1\,091$ gesunde Kontrollen an den eingeschlossenen 38 Studien teil. Einige Probandengruppen kamen in derselben Studie in mehreren Vergleichen vor, wenn z. B. Patienten mit Waschzwang sowohl mit Gesunden als auch mit Patienten mit Kontrollzwang verglichen wurden. Aufgrund dieser Überlappungen der Teilstichproben wurden keine Signifikanztests der Gruppenunterschiede durchgeführt. Studien mit Waschzwang, Kontrollzwang und Sammelzwang waren im Vergleich zu Studien mit aggressiven Zwangsgedanken und Symmetrie-/Ordnungszwang überrepräsentiert.

Demografische Daten. Das Durchschnittsalter lag bei Patienten mit Sammelzwang um mehr als zehn Jahre höher als bei den anderen Zwangspatienten und Kontrollen. Der Frauenanteil war bei Patienten mit Sammelzwang ebenfalls höher als bei den anderen Zwangspatienten und den gesunden Kontrollen. Intelligenz und Bildung befanden sich bei allen Probanden im durchschnittlichen Bereich.

Klinische Daten. Die Krankheitsdauer war bei Patienten mit Sammelzwang fast doppelt so hoch wie bei allen anderen Zwangspatienten. Hingegen erhielten Patienten mit Sammelzwang seltener Psychopharmaka als Patienten mit anderen Zwängen. Der Schweregrad der Zwangserkrankung lag bei allen Patientengruppen im moderaten bis schweren Bereich, während gesunde Kontrollen einen subklinisch niedrigen Y-BOCS-Wert aufwiesen. Die Depressivität (BDI-Wert) war bei allen Patientengruppen relativ stark ausgeprägt, bei den gesunden Kontrollen erwartungsgemäß subklinisch niedrig. Die Ängstlichkeit (STAI-S-Wert) lag für Patienten mit Symmetrie-/Ordnungszwang am höchsten, für Patienten mit Sammelzwang niedriger und für gesunde Kontrollen erwartungsgemäß im subklinischen Bereich.

Zusammenfassend kann gesagt werden, dass sich die Patienten von den Gesunden nicht wesentlich in Bezug auf demografische Merkmale unterschieden. Allerdings wiesen alle Zwangspatienten unabhängig von den Symptomdimensionen klinisch relevante Depressions-, Angst- und Zwangswerte auf und unterschieden sich hierin systematisch von den gesunden Teilnehmern (vgl. Tabelle 4).

Tabelle 4: *Demografische und klinische Merkmale der Studienteilnehmer.*

	Wasch- zwang	Kontroll- Zwang	Verbotene Gedanken	Symm./ Ordnen	Sammel- zwang	Gesunde Kontrollen
Anzahl der Studien k	14	20	2	4	15	33
Anzahl der Probanden n	308	413	22	240	401	1 025
Alter	34.44 (4.21)	34.17 (5.65)	33.68 (3.51)	33.05 (6.07)	49.52 (7.95)	38.68 (10.57)
% Frauen	60.9 (15.0)	54.3 (23.2)	36.0 (8.5)	60.29 (17.6)	69.0 (11.9)	59.4 (16.9)
% mediziert	70.2 (29.4)	57.1 (39.5)	66.7 ($k = 1$)	65.5 (21.6)	45.6 (30.0)	0
IQ	109.62 (6.20)	108.35 (5.10)	109.21 (4.79)	112.17 (4.00)	110.09 (12.10)	114.95 (5.33)
Bildungs- jahre	13.76 (1.60)	13.62 (1.59)	15.10 ($k = 1$)	15.31 (1.69)	15.89 (1.67)	14.55 (1.73)
Krankheits- dauer	9.60 (2.77)	11.18 (3.75)	11.36 ($k = 1$)	9.75 (3.42)	20.70 ($k = 1$)	n. z.
Alter bei Erst- erkrankung	19.98 (1.50)	19.63 (2.97)	20.06 ($k = 1$)	18.92 (0.68)	16.30 ($k = 1$)	n. z.
Y-BOCS	24.44 (5.67)	24.40 (4.52)	18.28 (1.21)	22.11 (3.73)	17.36 (10.99)	1.74 (0.93)
BDI	14.19 (2.55)	15.32 (4.86)	17.37 (4.07)	15.77 (3.08)	14.71 (3.74)	4.51 (2.64)
STAI-S	51.17 (5.35)	45.07 (4.00)	k. A.	53.31 (4.54)	44.00 (6.51)	29.95 (4.00)

Anmerkung. Angegeben ist jeweils M (SD). Wurden Werte nur für eine Studie angegeben ($k = 1$), konnte keine SD berechnet werden; k. A. = keine Angabe, n. z. = nicht zutreffend; BDI = Beck Depressions Inventar; IQ = Intelligenzquotient; STAI-S = State Trait Anxiety Inventory, Subskala State; Y-BOCS = Yale-Brown Obsessive Compulsive Scale. Krankheitsdauer in Jahren.

5.5 Metaanalyse 1: Patienten mit Waschwang vs. Gesunde

Die Ergebnisse der Metaanalyse werden im folgender Reihenfolge vorgestellt: Einzeleffektstärken, kombinierte Effektstärken der Domänen bzw. Subdomänen, Heterogenitätsprüfung, Moderator- und Metaregressionsanalysen und Abschätzung des Publikationsbias. Effektstärken mit negativem Vorzeichen bedeuten, dass die Patienten mit Waschwang eine schlechtere Testleistung erbrachten als die gesunden Kontrollen. In den eingeschlossenen Studien wurden nicht alle in Kapitel 4.2 vorgestellten kognitiven Domänen und Subdomänen (vgl. Tabelle 1) untersucht; dies bildet sich entsprechend in der Metaanalyse ab.

Die Patienten mit Waschwang zeigten über neun Studien hinweg in acht von 21 Tests eine signifikant schwächere Leistung als die gesunden Kontrollen (vgl. Abbildung 17). Große signifikante Einzeleffekte ergaben sich in den verbalen Gedächtnistests HVLТ und CVLT. Darüber hinaus schnitten die Patienten mit Waschwang vor allem in Tests zum visuell-räumlichen Arbeitsgedächtnis (CANTAB SWM, Spatial Working Memory Task, Blockspanne) mit Leistungsdifferenzen von bis zu vier Standardabweichungen schlecht ab. Auch in einer Aufgabe zum verbalen Arbeitsgedächtnis wies eine große Effektstärke auf Minderleistungen der Patienten mit Waschwang hin.

Für die einzelnen Domänen und Subdomänen ergaben sich eine signifikante große Effektstärke im räumlichen Arbeitsgedächtnis, mittelgroße Effektstärken im Konzeptwechsel und im verbalen Gedächtnis und ein kleiner Effekt in anhaltender Aufmerksamkeit. Keine signifikanten Unterschiede zwischen Patienten mit Waschwang und Gesunden fanden sich hingegen im nonverbalen Gedächtnis und im verbalen Arbeitsgedächtnis.

Die Ergebnisse der Heterogenitätsprüfung zeigten, entgegen der Annahme, für anhaltende Aufmerksamkeit, Konzeptwechsel, verbales Gedächtnis und nonverbales Gedächtnis keine signifikante Heterogenität. Hingegen fanden sich signifikante Heterogenitätswerte im räumlichen Arbeitsgedächtnis und im verbalen Arbeitsgedächtnis. Daher erfolgte die Integration der Effektstärken sowohl nach dem

random effects model als auch nach dem *fixed effect model*. Wie Tabelle 5 zeigt, unterschieden sich beide Integrationsmodelle nur minimal.

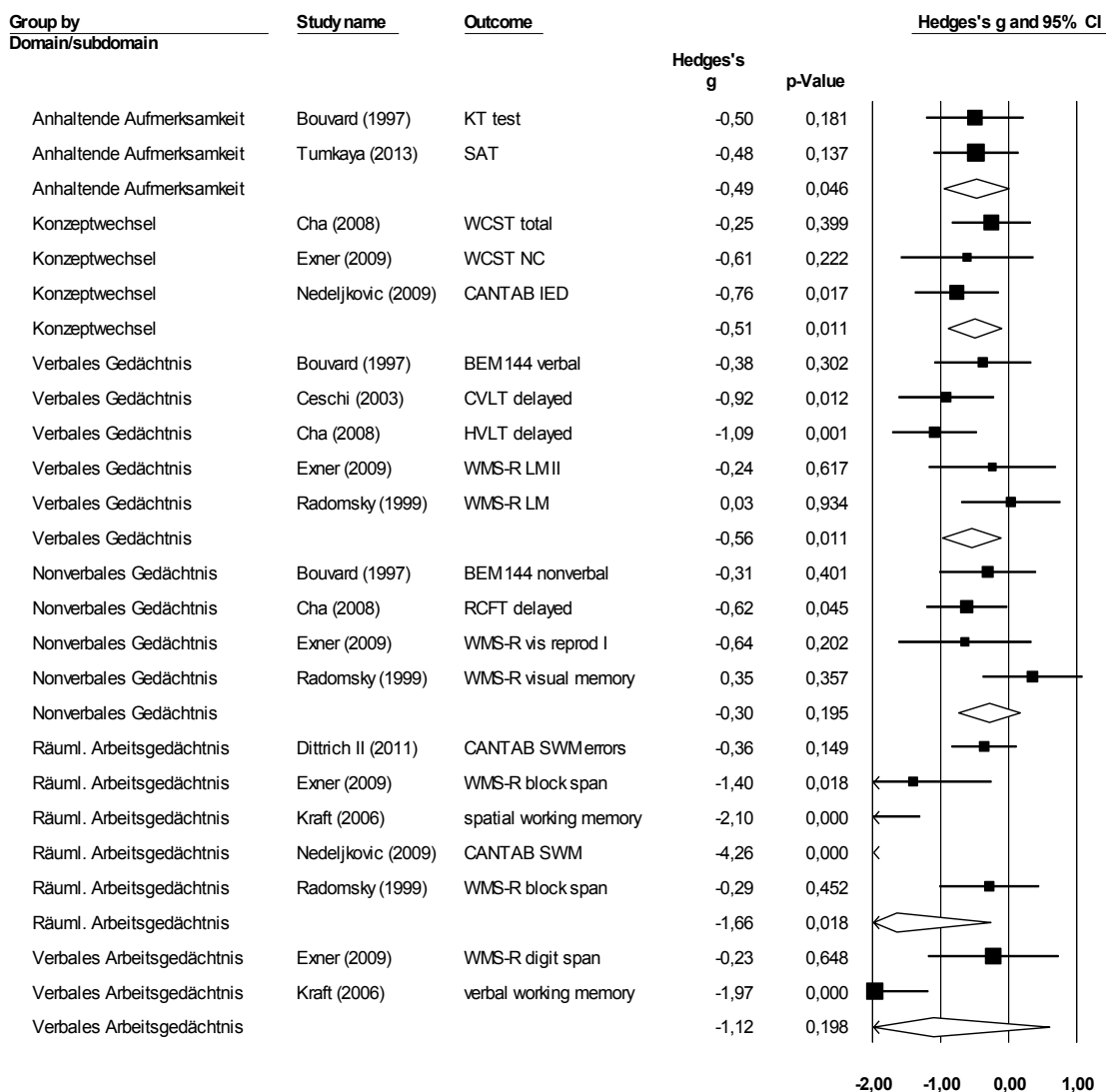


Abbildung 17. Forest plot. Patienten mit Waschzwang vs. Gesunde. Effektstärken kleiner als Null zeigen an, dass die Patienten mit Waschzwang eine schlechtere Testleistung erbrachten als die gesunden Kontrollen.

Tabelle 5. Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Patienten mit Waschwang vs. Gesunde.

Kognitive Subdomäne	<i>k</i>	Hedges' <i>g</i> (fixed) (random)	95 %- Konfidenz- intervall	<i>Q</i>	<i>df</i> (<i>Q</i>)	<i>p</i> (<i>Q</i>)	<i>I</i> ²
Anhaltende Aufmerksamkeit	2	- 0.49	- 0.97 -- - 0.01	0.00	1	.975	0.00
Konzeptwechsel	3	- 0.51	- 0.91 -- - 0.12	1.38	2	.501	0.00
Verbales Gedächtnis	5	- 0.56	- 0.99 -- - 0.13	6.78	4	.148	40.98
Nonverbales Gedächtnis	4	- 0.30	- 0.75 -- - 0.15	4.45	3	.217	32.57
Räumliches Arbeitsgedächtnis	5	- 1.66	- 3.04 -- - 0.28	63.45	4	< .001	93.70
Verbales Arbeitsgedächtnis	2	- 1.12	- 2.82 -- - 0.59	7.37	1	.007	86.44

Anmerkung. *k* = Anzahl der Studien (entspricht der Anzahl der in die Berechnung eingeflossenen Studieneffektstärken); *g* = Gesamteffektstärke für das jeweilige neuropsychologische Konstrukt; *Q* = Teststatistik für Signifikanztest der Heterogenität; *I*² = Varianzanteil, der durch Heterogenität der Studien erklärt wird.

Zwischen den Studien aus dem östlich-asiatischen und dem westlich-angelsächsischem Kulturkreis fanden sich keine signifikanten Differenzen ($Q = 0.01$, $df = 7$, $p = .937$). Auch unterschieden sich die sechs Studien, die Mittelwerte und Standardabweichungen berichteten, nicht signifikant von den zwei Korrelationsstudien ($Q = 1.87$, $df = 7$, $p = .172$).

Während die Metaregressionsanalysen keine signifikanten Einflüsse der Studienqualität ($Q = 2.26$, $df = 8$, $p = .133$), des Alters der Teilnehmer ($Q = 3.31$, $df = 8$, $p = .068$) und der Geschlechterverteilung ($Q = 0.01$, $df = 8$, $p = .904$) auf die

Effektstärken ergaben, trugen Ungleichheiten im Intelligenz- und Bildungsniveau der Probanden signifikant zu Differenzen in den gesamten neuropsychologischen Testleistungen bei ($beta = 1.01$, $intercept = -0.68$, $p < .001$). Auch beeinflussten Unterschiede im Intelligenz- und Bildungsniveau die Leistungsunterschiede im räumlichen Arbeitsgedächtnis signifikant ($beta = 1.46$, $intercept = -1.16$, $p < .001$). Mangels Daten konnten für die anderen Subdomänen (anhaltende Aufmerksamkeit, Konzeptwechsel und verbales Gedächtnis) keine separaten Metaregressionen berechnet werden.

Es ergaben sich keine Hinweise auf einen Publikationsbias (Rosenthals *fail-safe* $N = 98$, Orwins *fail-safe* $N = 58$). Der *funnel plot* (Abbildung 18) zeigte neben einem Ausreißer in die erwartete Richtung keine Asymmetrien. Entsprechend war eine Adjustierung nach *trim and fill* nicht erforderlich.

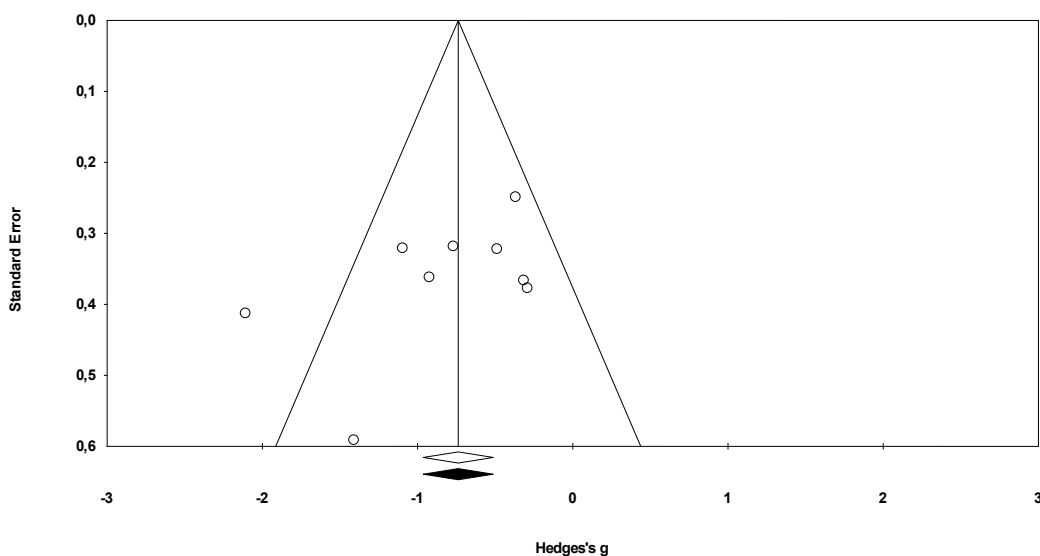


Abbildung 18. Funnel plot. Patienten mit Waschwang vs. Gesunde. Die weiße Raute stellt die beobachtete Gesamt-Effektstärke dar, die schwarze Raute die durch *trim and fill* adjustierte Effektstärke. Da keine Adjustierung erforderlich war, sind beide Gesamt-Effektstärken gleich groß.

Zusammenfassend kann gesagt werden, dass Patienten mit Waschwang gegenüber gesunden Kontrollen zwar signifikant verminderte Testleistungen in

anhaltender Aufmerksamkeit, Konzeptwechsel, verbalem Gedächtnis (jeweils mit mittelgroßen Effektstärken) und räumlichem Arbeitsgedächtnis (mit großer Effektstärke) aufwiesen. Allerdings waren diese auf signifikante Unterschiede im Intelligenz- und Bildungsniveau der Probanden zurückzuführen.

5.6 Metaanalyse 2: Patienten mit Kontrollzwang vs. Gesunde

Die Patienten mit Kontrollzwang zeigten über 16 Studien hinweg in 21 von 39 Tests eine signifikant schwächere Leistung als die gesunden Vergleichspersonen (vgl. Abbildung 19). Die größten signifikanten Einzeleffekte von über einer Standardabweichung zeigten sich in der Simon Spatial Incompatibility Task, im TMT-B, im Strategiescore und in der Fehlerzahl der SWM-Aufgabe, im HVLТ und in einer Wortlisten-Aufgabe, im RCFT, im AMIPB Figuren-Erinnern und bei der Zahlenspanne.

Für die einzelnen Domänen und Subdomänen ergaben sich signifikante große Effektstärken im Enkodieren, im verbalen und nonverbalen Gedächtnis sowie im räumlichen und verbalen Arbeitsgedächtnis. Mittelgroße Effektstärken zeigten sich in den Subdomänen Inhibition und Konzeptwechsel, während sich in der Verarbeitungsgeschwindigkeit nur ein kleiner Effekt fand. In keiner kognitiven Domäne bzw. Subdomäne schnitten die Patienten mit Kontrollzwang besser ab als die gesunden Probanden (vgl. Tabelle 6).

Die Ergebnisse der Heterogenitätsprüfung zeigten, entgegen der Annahme, für Inhibition, verbales Gedächtnis, Verarbeitungsgeschwindigkeit und verbales Arbeitsgedächtnis keine signifikante Heterogenität. Hingegen fanden sich signifikante Heterogenitätswerte im Konzeptwechsel, im Enkodieren, im nonverbalen Gedächtnis und im räumlichen Arbeitsgedächtnis. Daher erfolgte die Integration der Effektstärken sowohl nach dem *random effects model* als auch nach dem *fixed effect model*. Wie Tabelle 6 zeigt, unterschieden sich beide Integrationsmodelle insofern, als in den Domänen mit hoher Heterogenität die Effektstärken nach dem *random effects model* größer waren als nach dem *fixed effect model*, jedoch auch in breitere Konfidenzintervalle fielen.

Zwischen den fünf Studien aus dem östlich-asiatischen Kulturkreis und den elf Studien aus dem westlich-angelsächsischem Kulturkreis fanden sich keine signifikanten Differenzen ($Q = 0.60$, $df = 15$, $p = .437$). Auch unterschieden sich die zwölf Studien, die Mittelwerte berichteten, nicht signifikant von den vier Korrelationsstudien ($Q = 2.15$, $df = 15$, $p = .143$). Die Metaregressionsanalysen zeigten keine signifikanten Unterschiede im Alter ($Q = 0.90$, $df = 12$, $p = .342$), in der Geschlechterverteilung ($Q = 0.04$, $df = 10$, $p = .840$) und im Intelligenz- und Bildungsniveau ($Q = 0.72$, $df = 8$, $p = .397$). Hingegen stellte die Studienqualität einen kleinen, aber signifikanten Prädiktor für Unterschiede in den neuropsychologischen Testleistungen dar ($beta = 0.06$, $intercept = - 1.52$, $p = .012$): Je schlechter die Studienqualität war, desto geringer fielen die Effektstärken aus.

Die Kennwerte gaben keine Hinweise auf einen Publikationsbias (Rosenthals *fail-safe* $N = 294$, Orwins *fail-safe* $N = 97$). Auch der *funnel plot* (Abbildung 20) zeigte keine Asymmetrien, die auf einen Publikationsbias hindeuten würden.

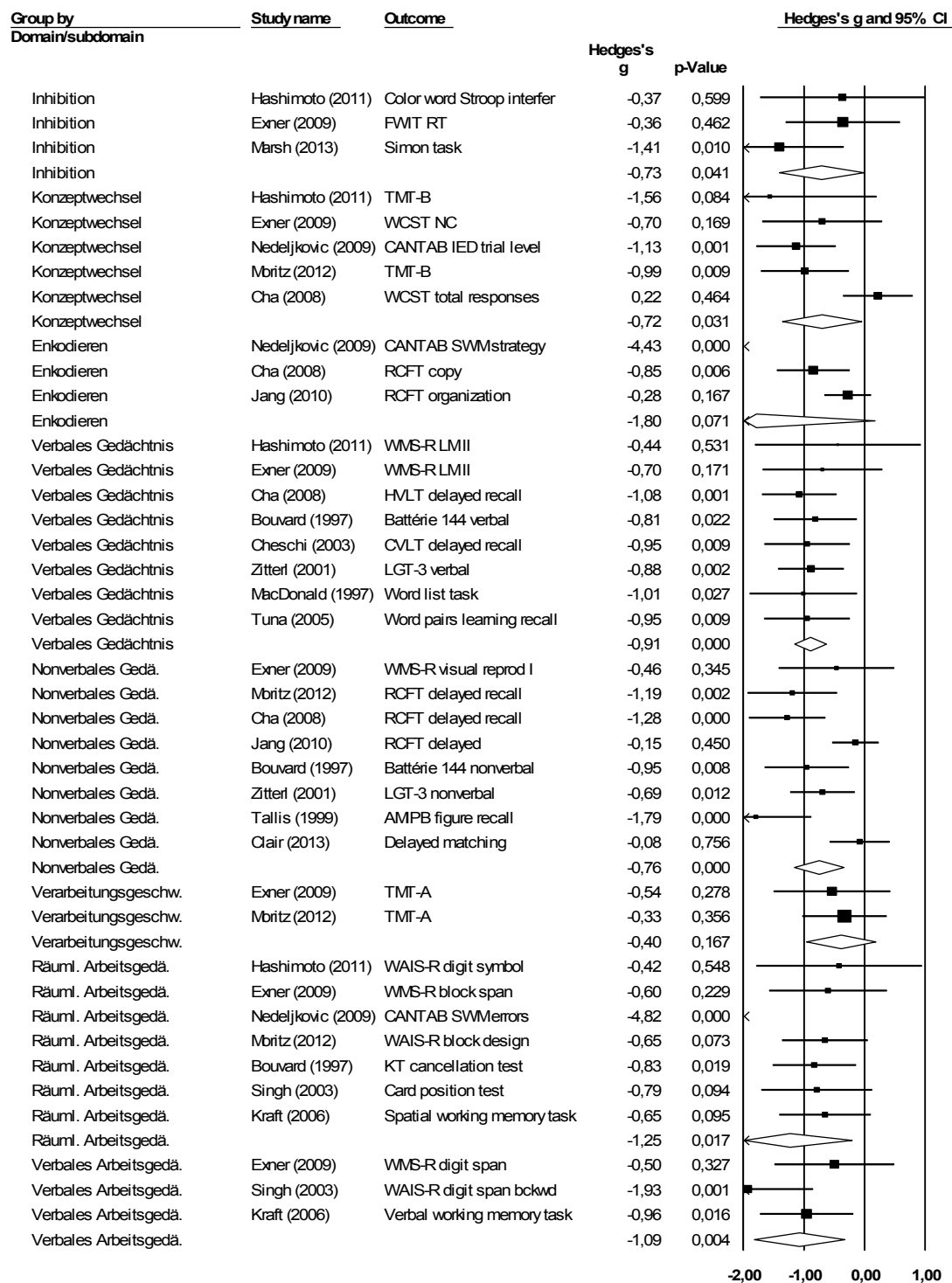


Abbildung 19. Forest plot. Patienten mit Kontrollzwang vs. Gesunde. Mit * markierte Studien sind unveröffentlichte, graue Literatur. Effektstärken kleiner als Null zeigen an, dass die Patienten mit Kontrollzwang eine schlechtere Testleistung erbrachten als die gesunden Probanden.

Tabelle 6. Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Patienten mit Kontrollzwang vs. Gesunde.

Kognitive Subdomäne	<i>k</i>	Hedges' <i>g</i> (fixed) (random)	95 %- Konfidenz- intervall	<i>Q</i>	<i>df</i> (<i>Q</i>)	<i>p</i> (<i>Q</i>)	<i>I</i> ²
Inhibition	3	-0.72 -0.73	- 1.36 -- 0.09 - 1.42 -- 0.03	2.39	2	.303	16.23
Konzeptwechsel	5	-0.59 -0.72	- 0.09 -- 0.25 - 1.37 -- 0.07	12.22	4	.016	67.25
Enkodieren	3	-0.86 -1.80	- 1.18 -- 0.65 - 3.75 -- 0.16	60.16	2	<.001	96.68
Verbales Gedächtnis	8	-0.91 -0.91	- 1.17 -- 0.65 - 1.17 -- 0.65	1.05	7	.994	0.00
Nonverbales Gedächtnis	9	-0.65 -0.84	- 0.86 -- 0.45 - 1.24 -- 0.44	27.37	8	.001	70.77
Verarbeitungs- geschwindigkeit	2	-0.40 -0.40	- 0.97 -- 0.17 - 0.97 -- 0.17	0.11	1	.737	0.00
Räumliches Arbeitsgedächtnis	7	-1.11 -1.09	- 1.44 -- 0.79 - 2.27 -- 0.22	56.74	6	<.001	89.43
Verbales Arbeitsgedächtnis	2	-1.06 -1.09	- 1.59 -- 0.53 - 1.83 -- 0.34	3.74	1	.154	46.50

Anmerkung. *k* = Anzahl der Studien (entspricht der Anzahl der in die Berechnung eingeflossenen Studieneffektstärken) *g* = Gesamteffektstärke für das jeweilige neuropsychologische Konstrukt *Q* = Teststatistik für Signifikanztest der Heterogenität; *I*² = Varianzanteil, der durch Heterogenität der Studien erklärt wird.

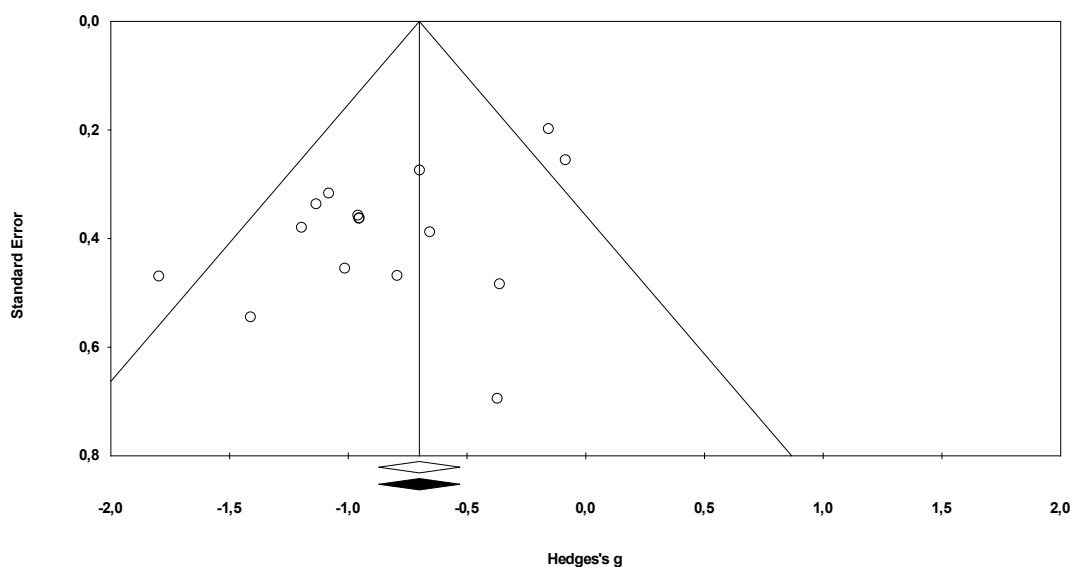


Abbildung 20. Funnel plot. Patienten mit Kontrollzwang vs. Gesunde. Der adjustierte Gesamteffekt (schwarze Raute) unterscheidet sich nicht vom beobachteten Gesamteffekt (weiße Raute).

Zusammenfassend kann gesagt werden, dass Patienten mit Kontrollzwang gegenüber gesunden Probanden verminderte neuropsychologische Testleistungen in allen untersuchten Subdomänen zeigten. Bei Aufgaben zur Inhibition und zum Konzeptwechsel fanden sich mittlere Effektstärken, wohingegen große Effektstärken im Enkodieren, im verbalen und nonverbalen Gedächtnis sowie im räumlichen und verbalen Arbeitsgedächtnis auf deutliche Defizite der Patienten mit Kontrollzwang hinwiesen. Es ergaben sich keine signifikanten Einflüsse von Kulturkreis, Alter, Geschlecht und Intelligenz/Bildungsgrad auf die Effektstärken. Hingegen führte eine geringere Studienqualität zu niedrigeren Effektstärken. Für einen Publikationsbias fanden sich keine Hinweise.

5.7 Metaanalyse 3: Patienten mit verbotenen Gedanken vs. Gesunde

Der Vergleich von Zwangspatienten mit verbotenen Gedanken mit gesunden Probanden ergab über zwei Studien hinweg fünf Effektstärken (vgl. Abbildung 21). Die Patienten mit verbotenen Gedanken zeigten in allen Tests eine geringere

Leistung als die gesunden Probanden. Vier große Effektstärken in den Tests SOC, IED, SWM und in einer Aufgabe zum verbalen Arbeitsgedächtnis wurden signifikant. Ein kombinierter Effekt konnte für das räumliche Arbeitsgedächtnis berechnet werden, hier ergab sich im *fixed effect model* ein insignifikanter Gesamteffekt ($g = -1.58$, $p = .240$), im *random effects model* hingegen ein signifikanter Gesamteffekt ($g = -1.23$, $p < .001$). Die Ergebnisse der Heterogenitätsprüfung zeigten, dass über 95 % der Gesamtvarianz auf „wahre“ Heterogenität zwischen den beiden Studien und nicht auf Zufallsfehler zurückzuführen waren (vgl. Tabelle 7).

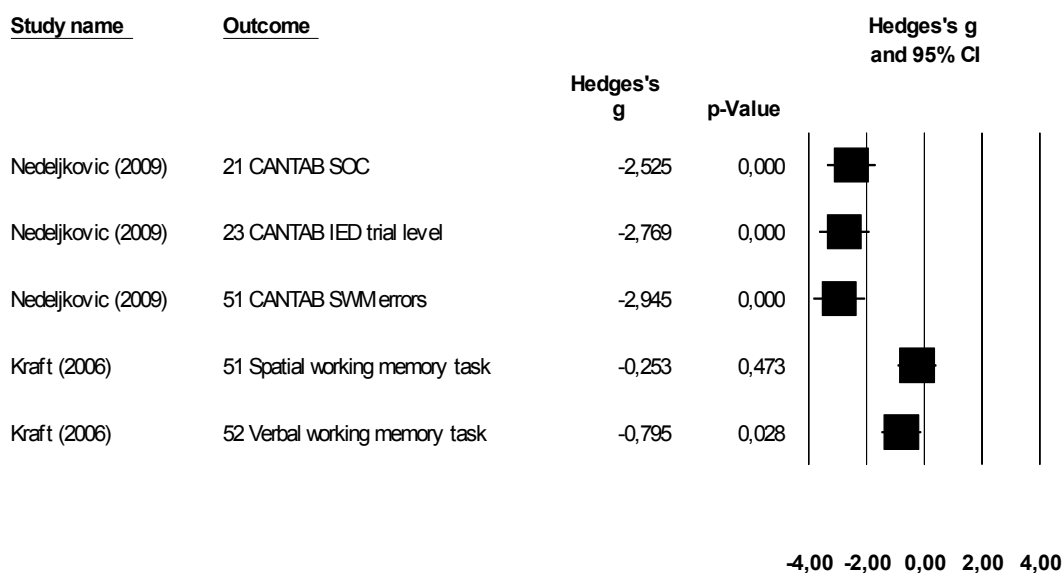


Abbildung 21. Forest plot. Zwangspatienten mit verbotenen Gedanken vs. Gesunde. Effektstärken kleiner als Null zeigen an, dass die Patienten mit verbotenen Gedanken eine schlechtere Testleistung erbrachten als die gesunden Kontrollen.

Tabelle 7. Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Zwangspatienten mit verbotenen Gedanken vs. Gesunde

Kognitive Subdomäne	<i>k</i>	Hedges' <i>g</i> (fixed) (random)	95 %- Konfidenz- intervall	<i>Q</i>	<i>df</i> (<i>Q</i>)	<i>p</i> (<i>Q</i>)	<i>I</i> ²
Räumliches	2	- 1.58	- 4.22 – 1.06	21.21	1	<.001	95.29
Arbeitsgedächtnis		- 1.23	- 1.78 – - 0.68				

Anmerkung: *k* = Anzahl der Studien (entspricht der Anzahl der in die Berechnung eingeflossenen Studieneffektstärken); *g* = Gesamteffektstärke für das jeweilige neuropsychologische Konstrukt; *Q* = Teststatistik für Signifikanztest der Heterogenität; *I*² = Varianzanteil, der durch Heterogenität der Studien erklärt wird.

Abschließend kann unter Berücksichtigung der geringen Datenmenge von nur zwei Studien festgestellt werden, dass große Effektstärken signifikant verminderte Leistungen von Zwangspatienten mit verbotenen Gedanken gegenüber gesunden Kontrollprobanden in Tests zum Planen/Problemlösen, zum Konzeptwechsel und zum räumlichen Arbeitsgedächtnis anzeigten. Anhand der kleinen Stichprobe konnten weder Moderator- und Metaregressionsanalysen noch eine Abschätzung des Publikationsbias durchgeführt werden.

5.8 Metaanalyse 4: Patienten mit Symmetrie-/Ordnungszwang vs. Gesunde

Die Patienten mit Symmetrie-/Ordnungszwang zeigten über vier Studien hinweg in drei von elf Tests eine signifikant schwächere Leistung als die gesunden Vergleichspersonen (vgl. Abbildung 22). Die größten signifikanten Einzeleffekte von über einer Standardabweichung zeigten sich im TMT-B und im WMS-R LM II.

Für die einzelnen Domänen und Subdomänen ergab sich eine große signifikante Effektstärke im verbalen Gedächtnis und eine kleine signifikante Effektstärke im nonverbalen Gedächtnis. Die Effektstärken in den Subdomänen Inhibition, Konzeptwechsel und räumliches Arbeitsgedächtnis blieben insignifikant.

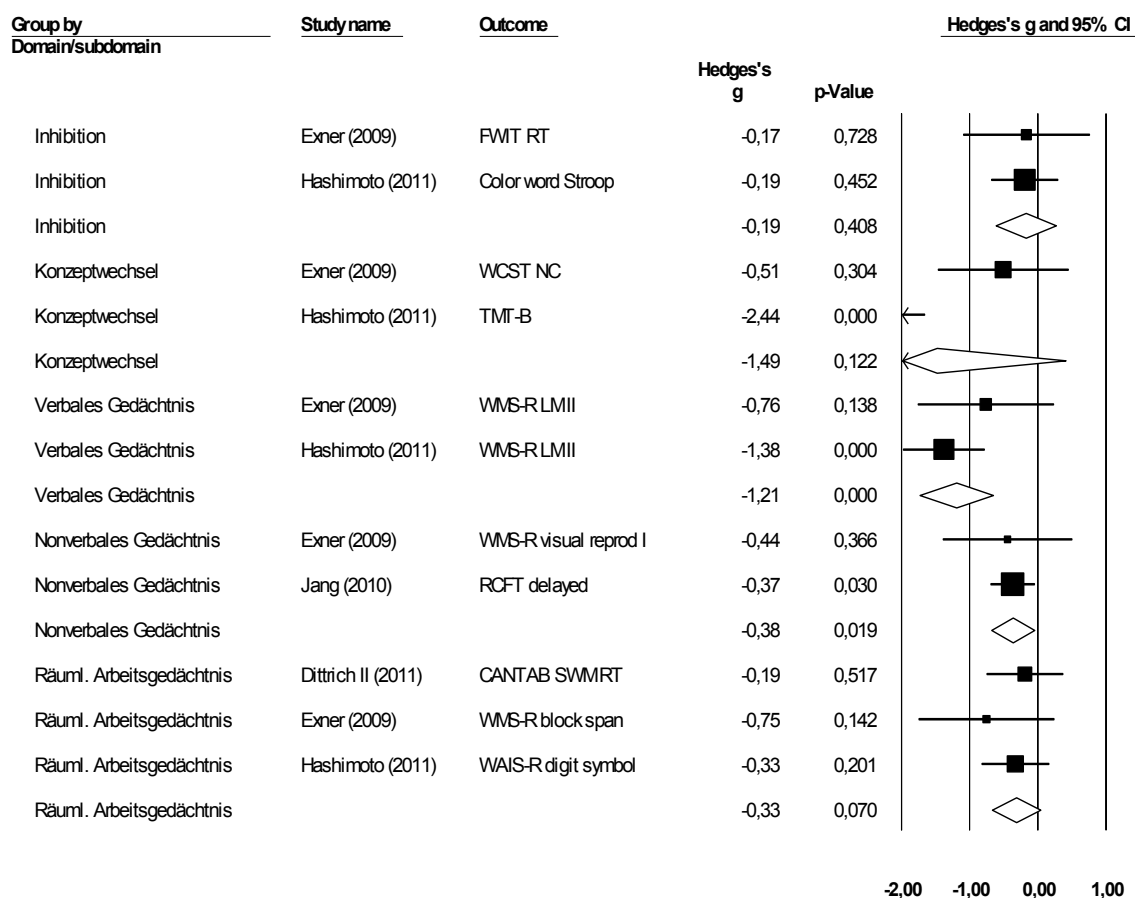


Abbildung 22. Forest plot. Patienten mit Symmetrie-/Ordnungszwang vs. Gesunde. Effektstärken kleiner als Null zeigen an, dass die Patienten mit Symmetrie-/Ordnungszwang eine schlechtere Testleistung erbrachten als die gesunden Kontrollen.

Die Ergebnisse der Heterogenitätsprüfung zeigten, entgegen der Annahme, für Inhibition, verbales Gedächtnis, nonverbales Gedächtnis und räumliches Arbeitsgedächtnis ausgeprägte Homogenität. Entsprechend lagen die kombinierten Effekte für diese Subdomänen im *random effects model* und im *fixed effect model* sehr nahe beieinander bzw. glichen sich sogar. Lediglich in den Studien zum Konzeptwechsel fand sich hohe Heterogenität, jedoch waren auch hier die Unterschiede beider Integrationsmodelle gering (vgl. Tabelle 8).

Tabelle 8. *Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Patienten mit Symmetrie-/ Ordnungszwang vs. Gesunde.*

Kognitive Subdomäne	<i>k</i>	Hedges' <i>g</i> (fixed) (random)	95 %- Konfidenz- intervall	<i>Q</i>	<i>df</i> (<i>Q</i>)	<i>p</i> (<i>Q</i>)	<i>I</i> ²
Inhibition	2	- 0.19 - 0.19	- 0.63 – 0.26 - 0.63 – 0.26	0.00	1	.962	0.00
Konzept- wechsel	2	- 1.66 - 1.49	- 2.28 – - 1.05 - 3.39 – 0.40	9.12	1	.003	89.04
Verbales Gedächtnis	2	- 1.21 - 1.21	- 1.74 – - 0.69 - 1.75 – - 0.67	1.05	1	.307	4.27
Nonverbales Gedächtnis	2	- 0.38 - 0.38	- 0.69 – - 0.06 - 0.69 – - 0.06	0.02	1	.888	0.00
Räumliches Arbeitsgedächtnis	3	- 0.33 - 1.66	- 0.68 – - 0.03 - 0.68 – - 0.03	0.91	2	.633	0.00

Anmerkung. *k* = Anzahl der Studien (entspricht der Anzahl der in die Berechnung eingeflossenen Studieneffektstärken); *g* = Gesamteffektstärke für das jeweilige neuropsychologische Konstrukt; *Q* = Teststatistik für Signifikanztest der Heterogenität; *I*² = Varianzanteil, der durch Heterogenität der Studien erklärt wird.

Zwischen den zwei Studien aus dem östlich-asiatischen Kulturkreis und den zwei Studien aus dem westlich-angelsächsischem Kulturkreis fanden sich keine signifikanten Differenzen ($Q = 0.21$, $df = 3$, $p = .644$). Auch unterschied sich die Studie, die Mittelwerte und Standardabweichungen berichtete, nicht signifikant von den drei Korrelationsstudien ($Q = 0.13$, $df = 3$, $p = .722$). Es zeigten sich keine signifikanten Unterschiede in der Studienqualität ($Q = 0.22$, $df = 3$, $p = .643$), im Alter der Teilnehmer ($Q = 0.31$, $df = 3$, $p = .577$), in der Geschlechterverteilung ($Q = 0.38$, $df = 3$, $p = .535$) und im Intelligenz- und Bildungsniveau ($Q = 0.42$, $df = 3$, $p = .517$). Aufgrund der kleinen Studienanzahl von nur vier Studien konnte ein möglicher Publikationsbias nicht abgeschätzt werden.

Abschließend kann festgestellt werden, dass Zwangspatienten mit Symmetrie-/Ordnungszwang gegenüber gesunden Probanden signifikant verminderte Testleistungen im verbalen Gedächtnis (große Effektstärke) und im nonverbalen Gedächtnis (kleine Effektstärke) zeigten. Einflüsse von Kovariaten wurden nicht signifikant.

5.9 Metaanalyse 5: Patienten mit Sammelzwang vs. Gesunde

Die Patienten mit Sammelzwang zeigten über 14 Studien hinweg in 14 von 30 Tests eine signifikant schwächere Leistung als die gesunden Vergleichspersonen (vgl. Abbildung 23). Die größten signifikanten Einzeleffekte von über einer Standardabweichung fanden sich in der Stroop-Inhibitionsaufgabe und in der Buchstaben-Zahlenfolge-Aufgabe der WAIS-IV. Ebenfalls große Einzeleffekte ergaben sich in einer Aufgabe zur Kategorisierung (Object Sorting Task) und in der VMS.

Für die einzelnen Domänen und Subdomänen ergaben sich eine große Effektstärke im räumlichen Arbeitsgedächtnis, mittelgroße Effektstärken für Inhibition und Kategorisierung sowie signifikante kleine Effektstärken im Planen/Problemlösen, im Konzeptwechsel, im nonverbalen Gedächtnis und im Entscheidungsverhalten. Keine signifikanten Unterschiede zwischen Patienten mit Sammelzwang und Gesunden fanden sich hingegen im verbalen Gedächtnis.

Die Ergebnisse der Heterogenitätsprüfung zeigten, entgegen der Annahme, für Planen/Problemlösen, Konzeptwechsel, nonverbales Gedächtnis, räumliches Arbeitsgedächtnis, Entscheidungsverhalten und Kategorisierung keine signifikante Heterogenität. Hingegen fanden sich signifikante Heterogenitätswerte in den Studien zur Inhibition und zum verbalen Gedächtnis. Daher erfolgte die Integration der Effektstärken sowohl nach dem *random effects model* als auch nach dem *fixed effect model*. Wie Tabelle 9 zeigt, unterschieden sich beide Integrationsmodelle insofern, als in den beiden Domänen mit hoher Heterogenität die Effektstärken nach dem *random effects model* größer waren als nach dem *fixed effect model*, jedoch auch in breitere Konfidenzintervalle fielen.

Da alle Studien im westlich-angelsächsischen Kulturkreis durchgeführt wurden, konnten keine kulturspezifischen Unterschiede ermittelt werden. Die 13 Studien, die Mittelwerte und Standardabweichungen berichteten, unterschieden sich nicht signifikant von der Korrelationsstudie ($Q = 0.27$, $df = 12$, $p = .601$).

Die Metaregressionsanalysen zeigten keine signifikanten Unterschiede in der Studienqualität ($Q = 0.02$, $df = 13$, $p = .883$), im Alter der Teilnehmer ($Q = 0.97$, $df = 12$, $p = .326$), in der Geschlechterverteilung ($Q = 0.09$, $df = 12$, $p = .764$) und im Intelligenz- und Bildungsniveau ($Q = 0.99$, $df = 9$, $p = .318$). Die Kennzahlen gaben keine Hinweise auf einen Publikationsbias (Rosenthals *fail-safe* $N = 110$, Orwins *fail-safe* $N = 50$). Dagegen deutete der leicht asymmetrische *funnel plot* (Abbildung 24) auf einen möglichen Publikationsbias hin. Eine Adjustierung nach *trim and fill* ergab eine ergänzte Studie, anhand derer sich der beobachtete Gesamteffekt ($g = -0.45$) sogar auf $g = -0.50$ erhöhen würde.

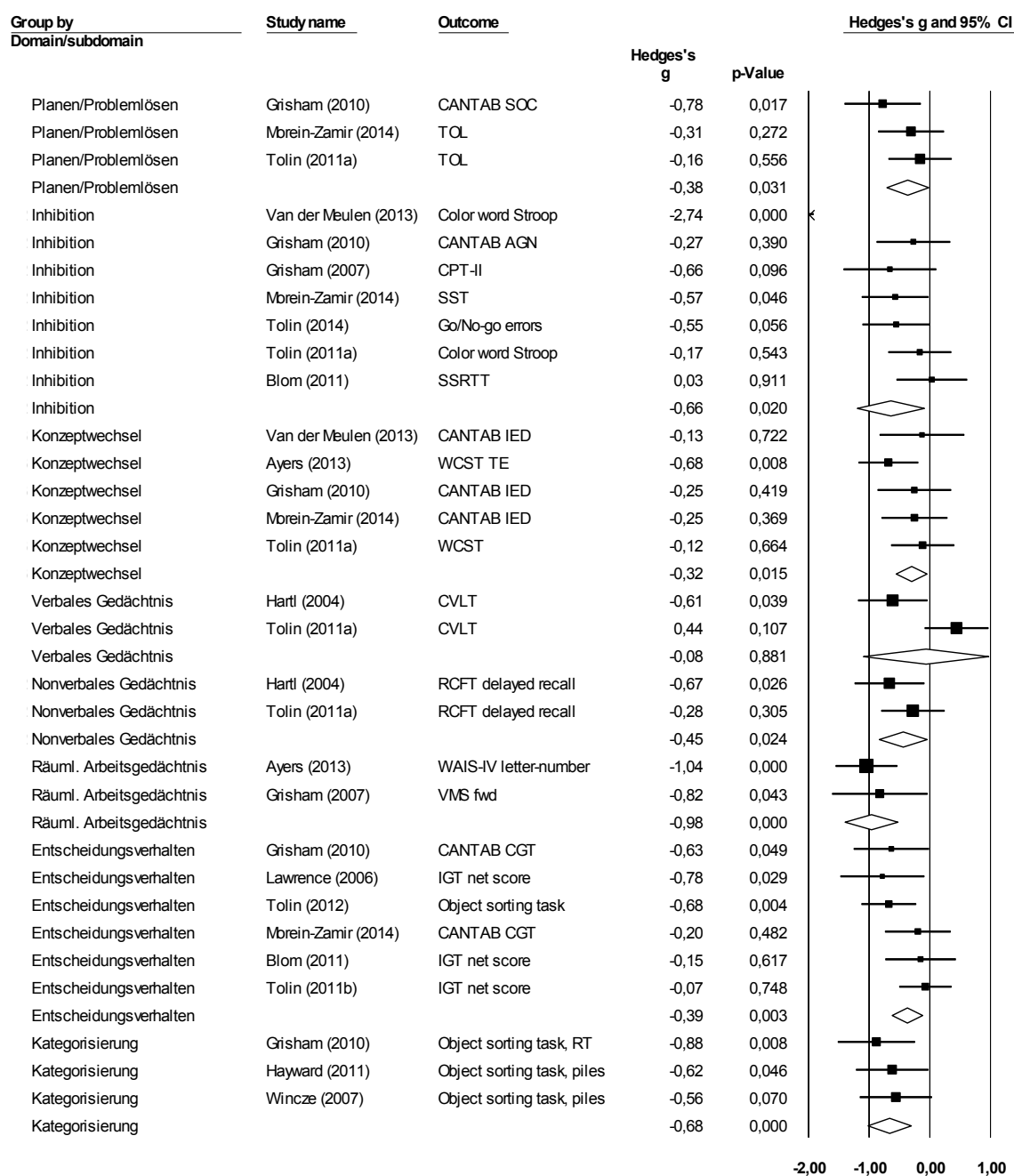


Abbildung 23. Forest plot. Patienten mit Sammelzwang vs. Gesunde. Effektstärken kleiner als Null zeigen an, dass die Patienten mit Sammelzwang eine schlechtere Testleistung erbrachten als die gesunden Kontrollen.

Tabelle 9. *Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Patienten mit Sammelzwang vs. Gesunde.*

Kognitive Subdomäne	<i>k</i>	Hedges' <i>g</i> (fixed) (random)	95 %- Konfidenz- intervall	<i>Q</i>	<i>df</i> (<i>Q</i>)	<i>p</i> (<i>Q</i>)	<i>I</i>²
Planen/ Problemlösen	3	- 0.38	- 0.71 -- 0.05	2.21	2	.332	9.30
		- 0.38	- 0.73 -- 0.03				
Inhibition	7	- 0.53	- 0.76 -- 0.30	33.02	6	< .001	81.83
		- 0.66	- 1.21 -- 0.11				
Konzept- wechsel	5	- 0.32	- 0.57 -- 0.06	2.94	4	.568	0.00
		- 0.32	- 0.57 -- 0.06				
Verbales Gedächtnis	2	- 0.04	- 0.44 -- 0.35	6.80	1	.009	85.30
		- 0.08	- 1.11 -- 0.95				
Nonverbales Gedächtnis	2	- 0.46	- 0.85 -- 0.06	0.92	1	.337	0.00
		- 0.46	- 0.85 -- 0.06				
Räumliches Arbeitsgedächtnis	2	- 0.98	- 1.41 -- 0.54	0.21	1	.650	0.00
		- 0.98	- 1.41 -- 0.54				
Entscheidungs- verhalten	6	- 0.38	- 0.60 -- 0.16	6.33	5	.276	20.95
		- 0.39	- 0.64 -- 0.14				
Kategorisierung	3	- 0.68	- 1.04 -- 0.32	0.56	2	.755	0.00
		- 0.68	- 1.04 -- 0.32				

Anmerkung. *k* = Anzahl der Studien (entspricht der Anzahl der in die Berechnung eingeflossenen Studieneffektstärken); *g* = Gesamteffektstärke für das jeweilige neuropsychologische Konstrukt; *Q* = Teststatistik für Signifikanztest der Heterogenität; *I*² = Varianzanteil, der durch Heterogenität der Studien erklärt wird.

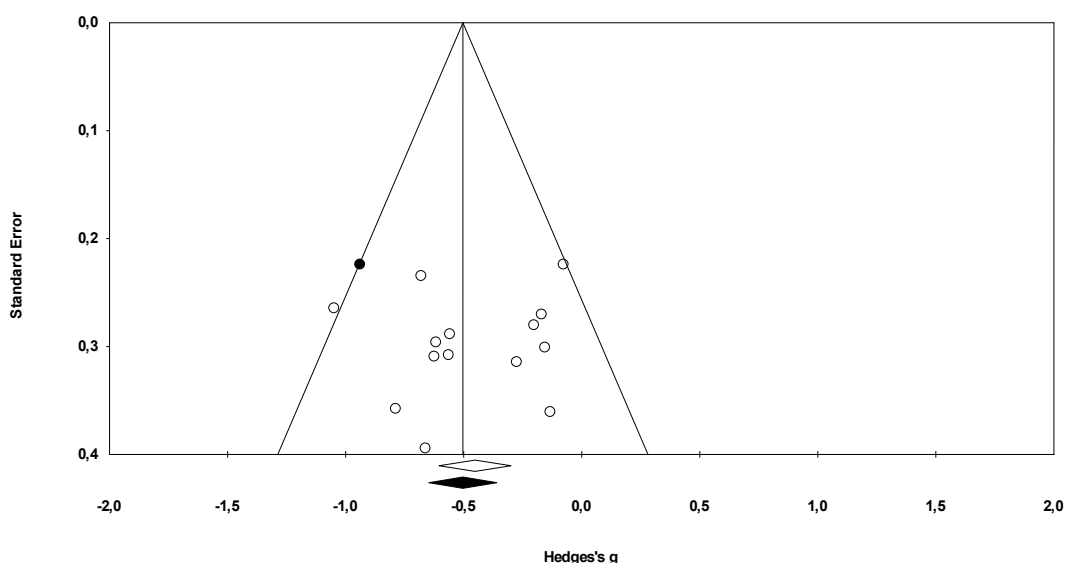


Abbildung 24. Funnel plot. Patienten mit Sammelzwang vs. Gesunde. Der adjustierte Gesamteffekt (schwarze Raute) ist größer als der beobachtete Gesamteffekt (weiße Raute).

Abschließend kann gesagt werden, dass kleine Effektstärken im Planen/Problemlösen, Konzeptwechsel, nonverbalen Gedächtnis und im Entscheidungsverhalten auftraten, während in kognitiver Inhibition und Kategorisierung mittelgroße Effektstärken und im räumlichen Arbeitsgedächtnis große Effektstärken auf entsprechende Defizite der Patienten mit Sammelzwang, verglichen mit Gesunden, hinwiesen. Die verbalen Gedächtnisleistungen von Patienten mit Sammelzwang erschienen dagegen unauffällig. Demografische Variablen hatten keinen signifikanten Einfluss auf die Effektstärken.

In den bis hier dargestellten Metaanalysen Nr. 1 bis Nr. 5 wurden Zwangspatienten mit jeweils einem Symptomschwerpunkt mit Gesunden verglichen. Bei diesen Analysen zeigten sich neuropsychologische Unterschiede zwischen den einzelner Symptom“gruppen“. Daher werden in den nun folgenden Metaanalysen jeweils Zwangspatienten untereinander verglichen. Es lagen ausreichend Daten vor, um Patienten mit Waschzwang mit Patienten mit Kontrollzwang (Metaanalyse Nr. 6) und Patienten mit Sammelzwang mit Zwangspatienten ohne Sammelzwang (Metaanalyse Nr. 7) zu vergleichen.

5.10 Metaanalyse 6: Patienten mit Waschwang vs. Kontrollzwang

Von den 13 eingeschlossenen Studien stammten jeweils drei aus Südkorea und Deutschland, zwei aus Japan, und je eine Studie aus Australien, Frankreich, Spanien, der Schweiz und der Türkei.

Die Patienten mit Waschwang zeigten über 13 Studien hinweg in 11 von 40 Tests eine signifikant bessere Leistung als die Patienten mit Kontrollzwang (vgl. Abbildung 25). Die größten signifikanten Einzeleffekte zeigten sich im TOH und der SOC-Aufgabe, gefolgt von Leistungen im SWM und der visuellen Reproduktionsaufgabe der WMS-R. In acht von zehn Domänen bzw. Subdomänen belegten positive Effektstärken, dass die Patienten mit Waschwang signifikant bessere Testleistungen erbrachten als die Patienten mit Kontrollzwang. Große Effektstärken fanden sich im Planen/Problemlösen und in der Inhibition. Eine mittlere Effektstärke ergab sich im Konzeptwechsel, wohingegen kleine Effekte bei anhaltender Aufmerksamkeit, beim Enkodieren, im verbalen Gedächtnis und im nonverbalen Gedächtnis sowie in der Verarbeitungsgeschwindigkeit auftraten. Keine signifikanten Unterschiede waren im räumlichen Arbeitsgedächtnis und im verbalen Arbeitsgedächtnis zu verzeichnen.

Die Ergebnisse der Heterogenitätsprüfung zeigten für acht der zehn kognitiven Subdomänen hohe Homogenität zwischen den Studien. Ausnahmen bildeten die Studien zum räumlichen und verbalen Arbeitsgedächtnis, die von ausgeprägter Heterogenität waren. Daher erfolgte die Integration der Effektstärken sowohl nach dem *random effects model* als auch nach dem *fixed effect model*. Wie Tabelle 10 zeigt, unterschieden sich beide Integrationsmodelle insofern, als in den Domänen mit hoher Heterogenität die Effektstärken nach dem *random effects model* größer waren als nach dem *fixed effect model*, jedoch auch in breitere Konfidenzintervalle fielen.

Zwischen den sechs Studien aus dem östlich-asiatischen Kulturkreis und den sieben Studien aus dem westlich-angelsächsischem Kulturkreis fanden sich keine

signifikanten Differenzen ($Q = 0.02$, $df = 12$, $p = .900$). Auch unterschieden sich die neun Studien, die Mittelwerte und Standardabweichungen berichteten, nicht signifikant von den vier Korrelationsstudien ($Q = 0.20$, $df = 12$, $p = .655$).

Die Metaregressionsanalysen zeigten keine signifikanten Unterschiede in der Studienqualität ($Q = 3.91$, $df = 13$, $p = .048$), im Alter der Teilnehmer ($Q = 5.12$, $df = 12$, $p = .024$), in der Geschlechterverteilung ($Q = 0.31$, $df = 12$, $p = .578$) und im Intelligenz- und Bildungsniveau ($Q = 1.67$, $df = 11$, $p = .197$), im Schweregrad der Zwangserkrankung (Y-BOCS Score; $Q = 1.26$, $df = 7$, $p = .262$), im Anteil der medizierten Patienten ($Q = 1.59$, $df = 4$, $p = .207$), in der Krankheitsdauer ($Q = 1.76$, $df = 3$, $p = .185$) und in Einflüssen komorbider Angst ($Q = 0.687$, $df = 5$, $p = .407$). Unterschiede in komorbider Depression hatten keinen signifikanten Einfluss auf die neuropsychologischen Effektstärken im Konzeptwechsel, Enkodieren, verbalen Gedächtnis und nonverbalen Gedächtnis. Allerdings zeigte sich ein signifikanter Einfluss der Gruppenunterschiede in komorbider Depression auf die Effektstärken im räumlichen Arbeitsgedächtnis ($beta = -1.42$, $intercept = 0.69$, $p < .001$, Anzahl der Studien = 3): Je stärker sich Patienten mit Waschzwang vs. Kontrollzwang voneinander in ihrer Depressivität unterschieden, desto stärker unterschieden sie sich auch in ihrer Leistung im Arbeitsgedächtnis. Mangels ausreichender Anzahl an Primärstudien konnte der Einfluss komorbider Depression auf die Testleistungen in anhaltender Aufmerksamkeit, im Planen/Problemlösen, in der Inhibition, der Verarbeitungsgeschwindigkeit und im verbalen Arbeitsgedächtnis nicht untersucht werden.

Es ergaben sich keine Hinweise auf einen Publikationsbias (Rosenthals *fail-safe* $N = 36$, Orwins *fail-safe* $N = 31$). Der *funnel plot* (Abbildung 26) zeigte eine leichte Asymmetrie, die auf einen Publikationsbias hindeuten könnte. Nach Adjustierung gemäß *trim and fill* blieb der Gesamteffekt jedoch signifikant.

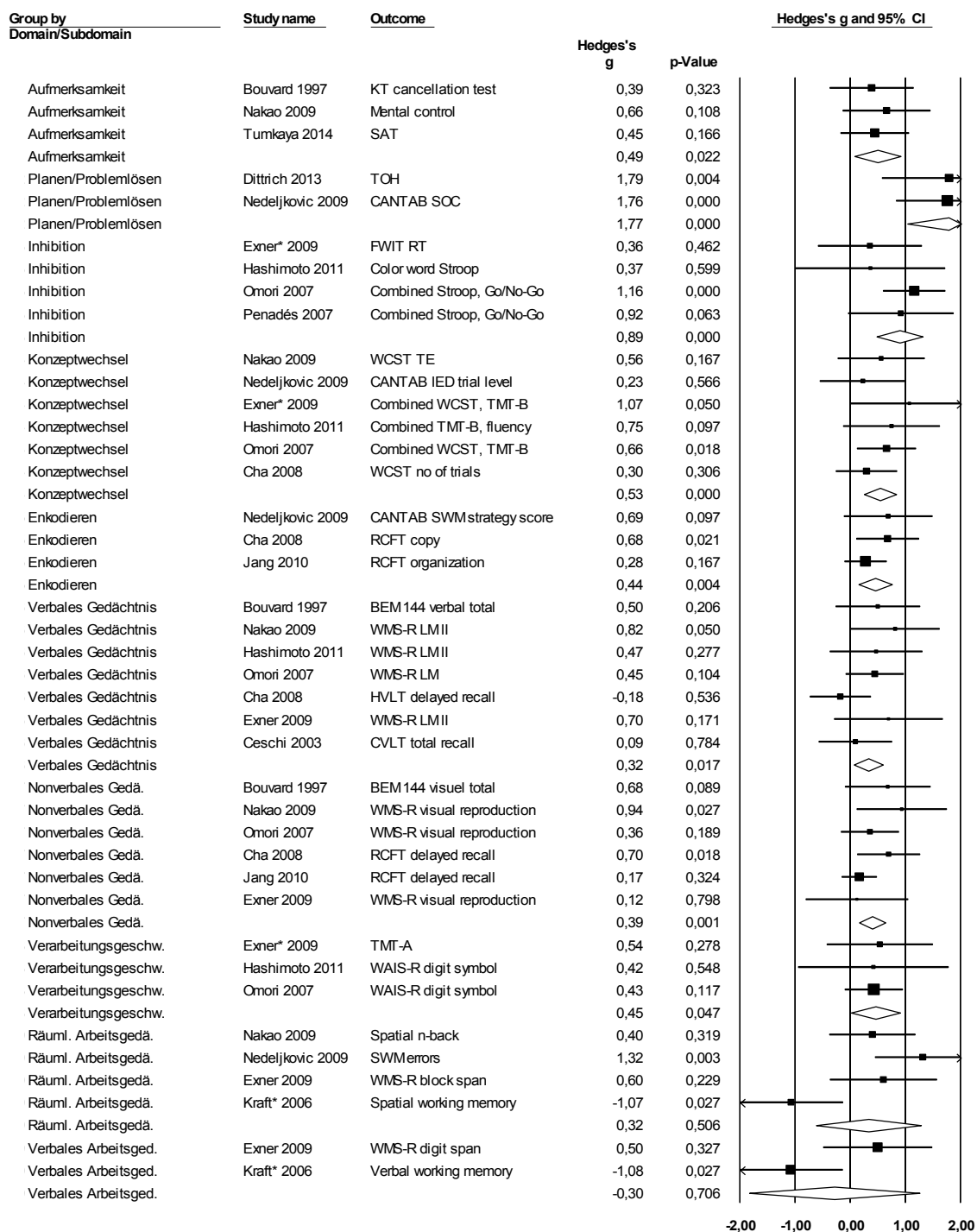


Abbildung 25. Forest plot. Patienten mit Waschzwang vs. Kontrollzwang. Effektstärken größer als Null zeigen an, dass die Patienten mit Waschzwang eine bessere Testleistung erbrachten als die Patienten mit Kontrollzwang.

Tabelle 10. *Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Patienten mit Waschwang vs. Patienten mit Kontrollzwang.*

Kognitive Subdomäne	<i>k</i>	Hedges' <i>g</i> (fixed) (random)	95 %- Konfidenz- intervall	<i>Q</i>	<i>df</i> (<i>Q</i>)	<i>p</i> (<i>Q</i>)	<i>I</i> ²
Anhaltende Aufmerksamkeit	3	0.49	0.07 – 0.91	0.26	2	.880	0.00
		0.49	0.07 – 0.91				
Planen/ Problemlösen	2	1.77	1.03 – 2.52	0.00	1	.968	0.00
		1.77	1.03 – 2.52				
Inhibition	4	0.89	0.47 – 1.31	2.63	3	.435	0.00
		0.89	0.47 – 1.31				
Konzeptwechsel	6	0.53	0.24 – 0.82	2.65	5	.754	0.00
		0.53	0.24 – 0.82				
Enkodieren	3	0.44	0.14 – 0.75	1.66	2	.437	0.00
		0.44	0.14 – 0.75				
Verbales Gedächtnis	7	0.32	0.06 – 0.58	5.93	6	.431	0.00
		0.32	0.26 – 0.58				
Nonverbales Gedächtnis	6	0.38	0.15 – 0.60	5.37	5	.372	6.97
		0.39	0.15 – 0.63				
Verarbeitungsgeschwindigkeit	3	0.45	0.01 – 0.90	0.04	2	.981	0.00
		0.45	0.01 – 0.90				
Räumliches Arbeitsgedächtnis	4	0.36	- 0.09 – - 0.80	13.51	3	.004	77.80
		0.32	- 0.63 – 1.27				
Verbales Arbeitsgedächtnis	2	- 0.32	- 1.01 – 0.36	4.44	1	.035	77.46
		- 0.30	- 1.84 – 1.25				

Anmerkung. *k* = Anzahl der Studien (entspricht der Anzahl der in die Berechnung eingeflossenen Studieneffektstärken); *g* = Gesamteffektstärke für das jeweilige neuropsychologische Konstrukt; *Q* = Teststatistik für Signifikanztest der Heterogenität; *I*² = Varianzanteil, der durch Heterogenität der Studien erklärt wird.

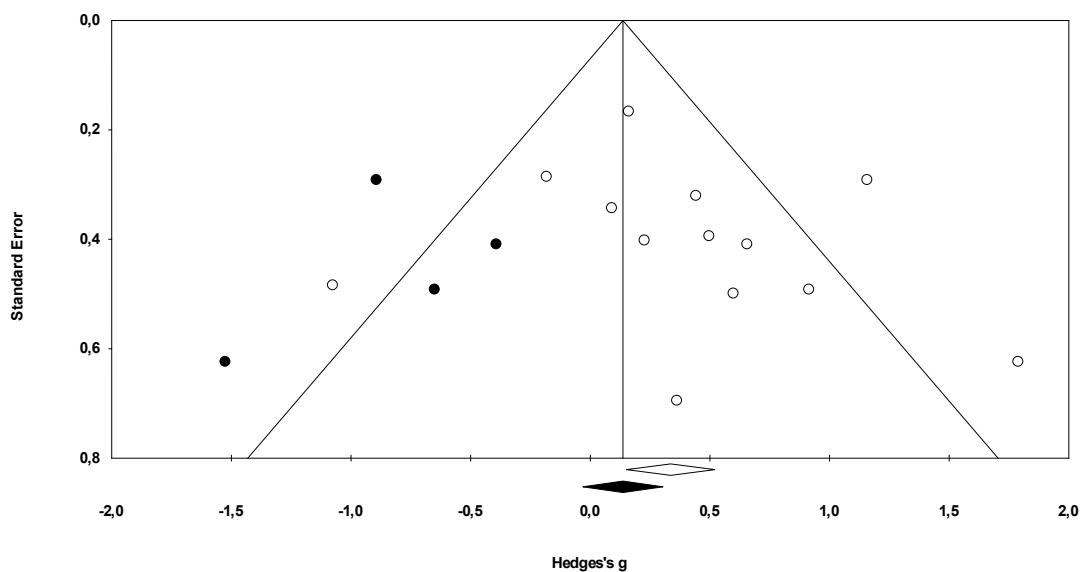


Abbildung 26. Funnel plot. Patienten mit Waschzwang vs. Kontrollzwang. Weiße Kreise stellen gefundene Studien dar, schwarze Kreise die durch *trim and fill* ergänzten Studien. Die adjustierte Gesamt-Effektstärke (schwarze Raute) fällt kleiner aus als die beobachtete Gesamt-Effektstärke (weiße Raute), bleibt aber größer als Null.

Es lässt sich zusammenfassen, dass Patienten mit Waschzwang in acht von zehn kognitiven Subdomänen signifikant besser abschnitten als Patienten mit Kontrollzwang. Die größten Defizite wiesen Patienten mit Kontrollzwang im Planen/Problemlösen, in der Inhibition und im Konzeptwechsel auf. Unterschiede in der Depressivität hatten einen signifikanten Einfluss auf die Leistungsunterschiede im räumlichen Arbeitsgedächtnis.

5.11 Metaanalyse 7: Patienten mit vs. ohne Sammelzwang

Die Patienten mit Sammelzwang zeigten über acht Studien hinweg in 10 von 13 Tests schlechtere, in drei Tests bessere Leistungen als die Zwangspatienten ohne Sammelzwang (vgl. Abbildung 27). Im Farbe-Wort-Stroop-Test, in der Iowa Gambling Task (IGT) und im IED wiesen große Effektstärken bis zu zwei Standardabweichungen auf deutlich schlechtere Leistung der Patienten mit Sammelzwang hin. Im Gegensatz dazu schnitten die Patienten mit Sammelzwang im Planungsteil des RCFT signifikant besser ab als die anderen Zwangspatienten.

In den Subdomänen Planen/Problemlösen, Inhibition und Konzeptwechsel blieben die Unterschiede zwischen Patienten mit und ohne Sammelzwang insignifikant. Lediglich im Entscheidungsverhalten wiesen die Patienten mit Sammelzwang signifikante, aber kleine Defizite auf, verglichen mit den anderen Zwangspatienten. Die Ergebnisse der Heterogenitätsprüfung zeigten, dass in fast allen Domänen (mit Ausnahme der Inhibition) die Studienheterogenität hoch war, so dass die Integration der Effekstärken nach dem *random effects model* erfolgte (vgl. Tabelle 11).

Alle acht Studien entstammten dem westlich-angelsächsischen Kulturkreis, so dass keine Aussagen über mögliche kulturspezifische Unterschiede getroffen werden konnten. Unterschiede im Datenformat (Mittelwerte und Standardabweichungen vs. Korrelationen und *F*-Statistiken; $Q = 1.74$, $df = 7$, $p = .187$) wurden nicht signifikant. Die Metaregressionsanalysen ergaben keine signifikanten Unterschiede in der Studienqualität ($Q = 2.59$, $df = 7$, $p = .107$), im Alter der Teilnehmer ($Q = 2.52$, $df = 5$, $p = .112$), in der Geschlechterverteilung ($Q = 1.26$, $df = 5$, $p = .262$), im Intelligenz- und Bildungsniveau ($Q = 0.00$, $df = 3$, $p = .975$), im Anteil medizierter Patienten ($Q = 4.61$, $df = 5$, $p = .032$) und in komorbider Depression ($Q = 0.08$, $df = 3$, $p = .783$).

Rosenthals *fail-safe* $N = 1$ wies auf einen möglichen Publikationsbias hin, was im Widerspruch zu Orwins *fail-safe* $N = 12$ und zu einem symmetrischen *funnel plot* (Abbildung 28) stand, für den eine Adjustierung nach *trim and fill* nicht erforderlich war. In der Gesamtbetrachtung erschien ein Publikationsbias damit eher unwahrscheinlich.

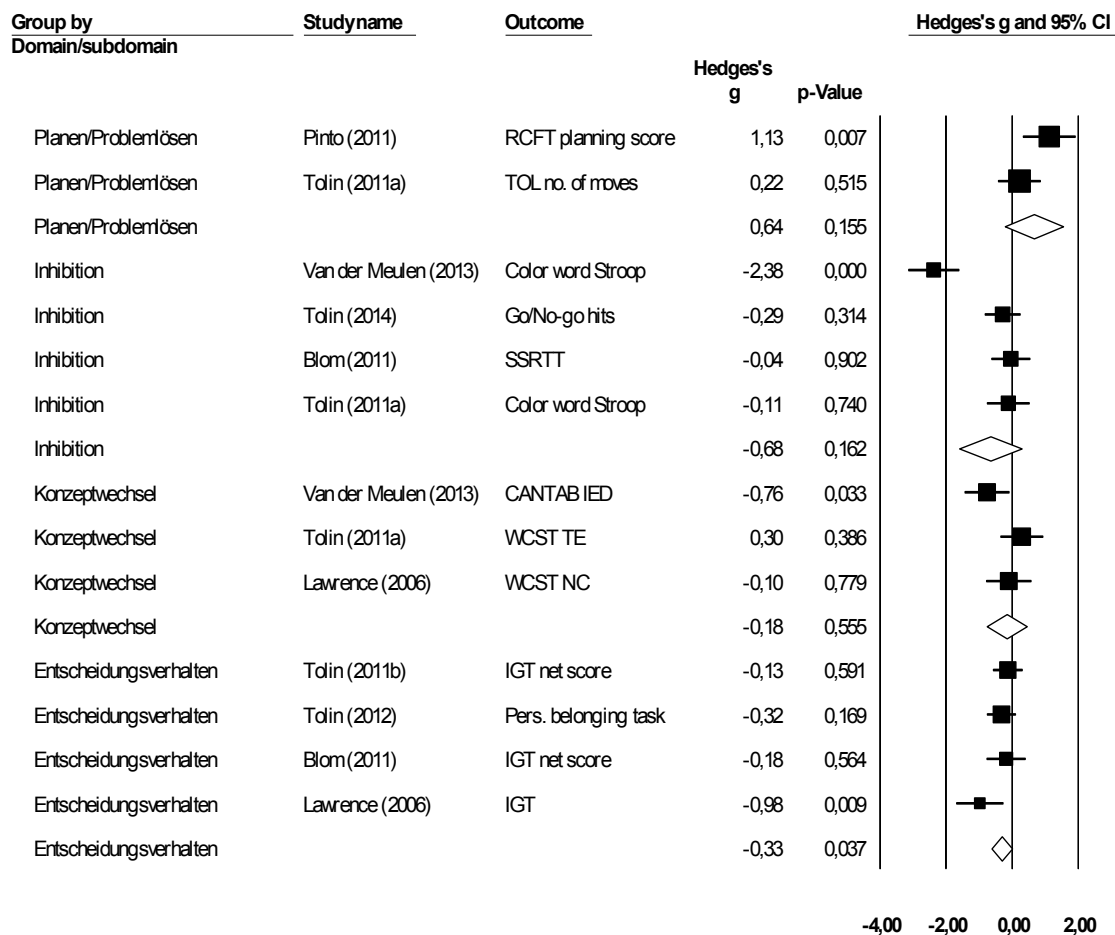


Abbildung 27. Forest plot. Patienten mit vs. ohne Sammelzwang. Effektstärken kleiner als Null zeigen an, dass die Patienten mit Sammelzwang eine schlechtere Testleistung erbrachten als die Patienten ohne Sammelzwang.

Tabelle 11. Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Patienten mit Sammelzwang vs. Zwangspatienten ohne Sammelzwang.

Kognitive Subdomäne	<i>k</i>	Hedges' <i>g</i> (fixed) (random)	95 %- Konfidenz- intervall	<i>Q</i>	<i>df</i> (<i>Q</i>)	<i>p</i> (<i>Q</i>)	<i>I</i> ²
Planen/ Problemlösen	2	0.58	0.07 – 1.10	2.81	1	.094	64.37
		0.64	- 0.24 – 1.53				
Inhibition	4	- 0.52	- 0.84 – - 0.20	25.60	3	< .001	88.28
		- 0.68	- 1.62 – - 0.27				
Konzeptwechsel	3	- 0.18	- 0.57 – 0.22	4.65	2	.098	57.01
		- 0.18	- 0.79 – 0.43				
Entscheidungs- verhalten	4	- 0.32	- 0.59 – 0.36	3.89	3	.274	22.78
		- 0.33	- 0.62 – - 0.02				

Anmerkung. *k* = Anzahl der Studien (entspricht der Anzahl der in die Berechnung eingeflossenen Studieneffektstärken); *g* = Gesamteffektstärke für das jeweilige neuropsychologische Konstrukt; *Q* = Teststatistik für Signifikanztest der Heterogenität; *I*² = Varianzanteil, der durch Heterogenität der Studien erklärt wird.

Zusammenfassend kann gesagt werden, dass sich Patienten mit Sammelzwang gegenüber Zwangspatienten ohne Sammelzwang nicht signifikant unterschieden im Planen/Problemlösen, in der Inhibition und im Konzeptwechsel, dass erstere jedoch eine signifikant verminderte Fähigkeit, Entscheidungen zu treffen, aufwiesen. Allerdings war dieser Unterschied nur gering, wie eine kleine Effektstärke belegte.

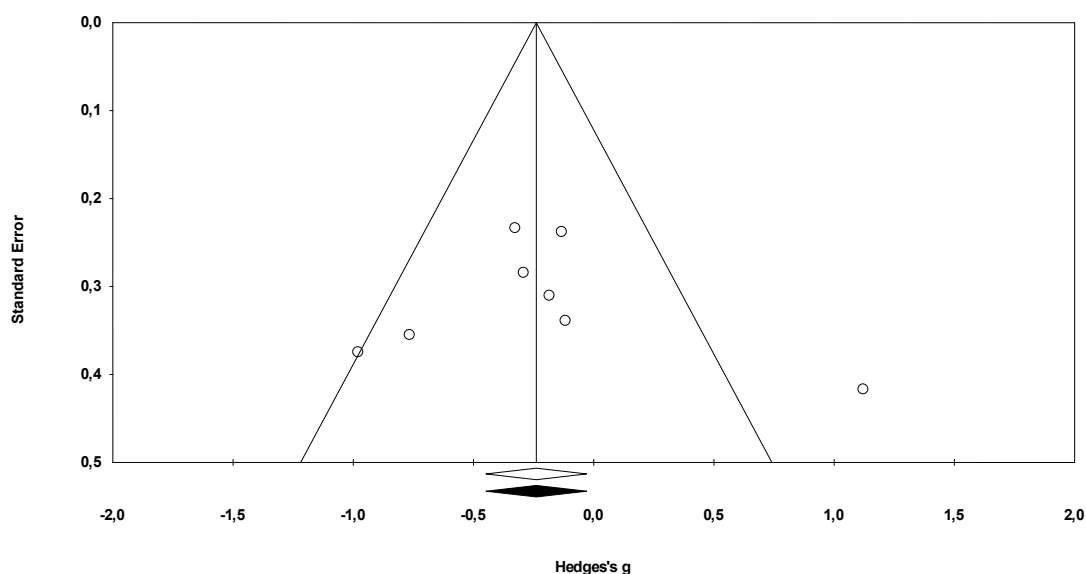


Abbildung 28. Funnel plot. Patienten mit vs. ohne Sammelzwang. Der beobachtete Gesamteffekt (weiße Raute) unterscheidet sich nicht von der durch *trim and fill* adjustierten Effektstärke (schwarze Raute).

5.12 Kombinierte Effektstärken der sieben Metaanalysen im Überblick

Eine Übersicht der Effektstärken aller sieben Metaanalysen über alle kognitiven Subdomänen hinweg (vgl. Tabelle 12) zeigt, dass je nach Symptomdimension unterschiedliche kognitive Aspekte untersucht wurden und dass entsprechend für zahlreiche Domänen und Subdomänen noch nicht genügend Daten vorlagen. So waren beispielsweise Studien zu anhaltender Aufmerksamkeit, zum Planen/Problemlösen, zum Enkodieren von Information und zur Verarbeitungsgeschwindigkeit insgesamt unterrepräsentiert. Leistungen im Entscheidungsverhalten und in der Kategorisierung wurden ausschließlich bei Patienten mit Sammelzwang untersucht, nicht jedoch bei Zwangspatienten mit anderen Symptomschwerpunkten. Für die Symptomdimension *verbotene Gedanken* konnte mangels Daten (es wurden lediglich zwei Primärstudien gefunden) nur in einer Subdomäne (räumliches Arbeitsgedächtnis) ein kombinierter Effekt ermittelt werden.

Die Übersicht zeigt weiterhin Gemeinsamkeiten und Unterschiede zwischen den verschiedenen Patientengruppen. So wiesen Zwangspatienten fast aller Symptomdimensionen (mit Ausnahme von Symmetrie-/Ordnungszwang) im Vergleich zu Gesunden große signifikante Defizite im visuell-räumlichen Arbeitsgedächtnis auf. Allerdings (und das zeigt die Tabelle nicht) gingen diese Unterschiede bei Patienten mit Waschzwang vs. Gesunde auf Unterschiede im Intelligenz- und Bildungsniveau zurück. Hingegen stellten Patienten mit vorherrschendem Kontrollzwang und mit Sammelzwang diejenigen Gruppen dar, die im Vergleich zu Gesunden in den meisten Subdomänen verminderte Testleistungen erbrachten. Innerhalb der Zwangspatienten waren es jene mit Wasch- und Kontrollzwang, die am häufigsten miteinander verglichen wurden. Hierbei zeigten Patienten mit Waschzwang in allen untersuchten Domänen und Subdomänen bessere Leistungen als Patienten mit Kontrollzwang, insbesondere im Planen/Problemlösen und in kognitiver Inhibition. Patienten mit Sammelzwang zeigten im Vergleich zu Zwangspatienten ohne Sammelzwang keine signifikanten Unterschiede in den untersuchten kognitiven Funktionsbereichen. Lediglich ein kleiner Effekt wurde signifikant: Im Entscheidungsverhalten erbrachten die Patienten mit Sammelzwang eine etwas schlechtere Leistung als Patienten ohne Sammelzwang.

Tabelle 12. Signifikante Effektstärken (*Hedges' g*, random effects, $p < .05$) der 10 kognitiven Subdomänen über alle 7 Metaanalysen.

	Waschzwang vs. Gesunde $k = 9$	Kontrollzwang vs. Gesunde $k = 16$	Verbotene Gedanken vs. Gesunde $k = 2$	Symmetrie- zwang vs. Gesunde $k = 4$	Sammelzwang vs. Gesunde $k = 14$	Waschzwang vs. Kontrollzwang $k = 13$	Sammelzwang vs. Zwang ohne Sammeln $k = 8$
	Anhaltende Aufmerksamkeit	-0.49	-	-	-	-	0.49
Planen/ Problemlösen	-	-	-	-	-0.38	1.77	n. s.
Inhibition	-	-0.73	-	n. s.	-0.66	0.89	n. s.
Konzeptwechsel	-0.51	-0.72	-	n. s.	-0.32	0.53	n. s.
Enkodieren	-	n. s.	-	-	-	0.44	-
Verbales Gedächtnis	-0.59	-0.91	-	-1.21	n. s.	0.32	-
Nonverbales Gedächtnis	n. s.	-0.84	-	-0.38	-0.46	0.39	-
Verarbeitungs- geschwindigkeit	-	n. s.	-	-	-	0.45	-
Räumliches Arbeitsgedächtni	-1.19	-1.09	-1.23	n. s.	-0.98	n. s.	-
Verbales Arbeitsgedächtni	n. s.	-1.09	-	-	-	n. s.	-
Entscheidungs- verhalten	-	-	-	-	-0.39	-	-0.33
Kategorisierung	-	-	-	-	-0.68	-	-

Anmerkung. n. s. = nicht signifikant; - = keine Effektstärke berechenbar. Die Effektstärken der Metaanalyse Nr. 1 sind auf signifikante Unterschiede im Intelligenz-/Bildungsniveau der Probanden zurückzuführen und sollten daher nicht interpretiert werden.

6 Diskussion

Die Fragestellung dieser Dissertation lautete: Unterscheiden sich Zwangspatienten mit Wasch-, Kontroll-, Symmetrie-/Ordnungszwang, verbotenen Gedanken oder Sammelzwang von Gesunden und voneinander in Bezug auf neuropsychologische Testleistungen? Auf Grundlage der gefundenen Daten wurden sieben Metaanalysen mit 38 relevanten Studien durchgeführt. Im Ergebnis zeigten die einzelnen Symptomdimensionen unterschiedliche neuropsychologische Profile. Im Folgenden werden zunächst die eingesetzten Methoden kritisch betrachtet (Abschnitt 6.1). Danach werden die sieben Metaanalysen einzeln diskutiert (Abschnitte 6.2 bis 6.8), bevor sich eine generelle Diskussion (Abschnitt 6.9) anschließt. Die Diskussion ordnet die Ergebnisse in den Kontext bisheriger Forschung ein, beschreibt Begrenzungen und Stärken der vorliegenden Arbeit und gibt einen Ausblick auf mögliche Vorhaben für die Zukunft.

6.1 Methoden

Metaanalyse. Ein systematisches Review und eine Metaanalyse sind nicht vollkommen objektiv, da der Reviewer beim Festlegen von Regeln zur Literatursuche und zur Datenaufbereitung subjektive Entscheidungen trifft. Gleichwohl handelt es sich um einen transparenten, replizierbaren Vorgang, da alle Entscheidungen klar formuliert werden. Dies stellt einen Vorteil gegenüber unsystematischen, narrativen Reviews dar, bei denen Entscheidungsprozesse intransparent bleiben oder sogar dem Reviewer nicht bewusst sind. Während narrative Reviews bei einer großen Anzahl von Primärstudien an ihre Grenzen geraten und unübersichtlich werden, entwickelt die Metaanalyse umso mehr Vorteile, je mehr Studien eingeschlossen werden. Sind die relevanten Merkmale einmal kodiert, spielt es kaum noch eine Rolle, ob man 10 oder 500 Studien analysiert (Green & Hall, 1984). Statistisch bringt die Aggregation von Daten ähnlicher Studien eine höhere Teststärke: Je umfangreicher die Stichprobe, desto höher ist die Chance, einen Effekt auch tatsächlich zu entdecken, der in der Population vorhanden ist (Borenstein, 2009). Eine Metaanalyse kann Lücken in der Forschungsliteratur

aufzeigen und damit die Forderung nach weiteren Forschungsprojekten untermauern – oder aber den Reviewer zur Erkenntnis bringen, dass auf diesem Gebiet keine weitere Forschung mehr nötig ist (Green & Hall, 1984).

Eine Kombination mehrerer Metaanalysen zu einer übergeordneten Fragestellung kann als *umbrella review* (Ioannidis, 2009, S. 489) bezeichnet werden. Welche Metaanalysen durchgeführt wurden und welche noch nicht, lässt sich in einem geometrischen Netzwerk darstellen (vgl. Abbildung 29).

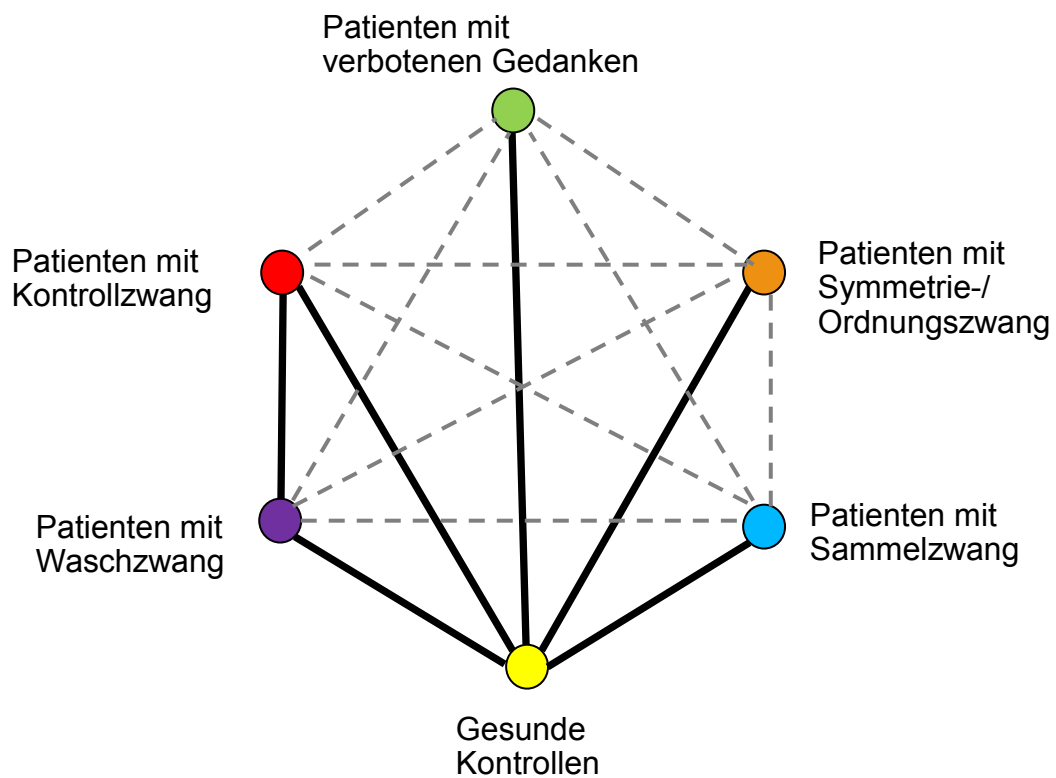


Abbildung 29. Geometrisches Netzwerk eines *umbrella reviews*. In Anlehnung an Ioannidis (2009, S. 489). Jeder andersfarbige Punkt symbolisiert eine Probandengruppe mit entsprechend vorherrschender Symptomdimension, jede Linie einen metaanalytischen Vergleich zwischen zwei Probandengruppen. Schwarze Linien bezeichnen die in der vorliegenden Arbeit durchgeführten Metaanalysen. Graue, gestrichelte Linien stellen die mangels Daten noch ausstehenden Metaanalysen dar. Der Vergleich von Zwangspatienten mit und ohne Sammelzwang ist nicht in der Grafik enthalten

Die Grafik zeigt zum einen, welche Symptomdimensionen metaanalytisch verglichen wurden. Zum anderen wird deutlich, dass viele Bereiche noch nicht untersucht worden sind. Mangels Primärstudien sind folgende Gruppenvergleiche offen geblieben: Patienten mit Waschzwang vs. Patienten mit verbotenen Gedanken; Patienten mit Waschzwang vs. Patienten mit Symmetrie-/Ordnungszwang; Patienten mit Waschzwang vs. Patienten mit Sammelzwang; Patienten mit Kontrollzwang vs. Patienten mit verbotenen Gedanken; Patienten mit Kontrollzwang vs. Patienten mit Symmetrie-/Ordnungszwang; Patienten mit Kontrollzwang vs. Patienten mit Sammelzwang; Patienten mit verbotenen Gedanken vs. Patienten mit Symmetrie-/Ordnungszwang; Patienten mit verbotenen Gedanken vs. Patienten mit Sammelzwang; Patienten mit Symmetrie-/Ordnungszwang vs. Patienten mit Sammelzwang. Zu diesen Forschungslücken könnten in Zukunft weitere Studien durchgeführt werden.

Auswahl der neuropsychologischen Konstrukte. Es wurden sieben übergeordnete neuropsychologische Domänen berücksichtigt: Aufmerksamkeit, Exekutivfunktionen, Gedächtnis, Arbeitsgedächtnis, Verarbeitungsgeschwindigkeit, Entscheidungsverhalten und Kategorisierung. Diese enthielten folgende neuropsychologischen Subdomänen: Anhaltende Aufmerksamkeit; Planen/ Problemlösen, Inhibition, Konzeptwechsel, Enkodieren, verbales und nonverbales episodisches Gedächtnis, visuell-räumliches und verbales Arbeitsgedächtnis, Verarbeitungsgeschwindigkeit, Entscheidungsverhalten und Kategorisierung. Die Entscheidung, welche neuropsychologischen Konstrukte eingeschlossen werden sollten, wurde auf Basis von Grundlagenliteratur zur Neuropsychologie (Kathmann, 2008; Tallis, 1995) sowie in Anlehnung an aktuelle Reviews zur Neuropsychologie von Zwangsstörungen (Abramovitch et al., 2013; Chamberlain et al., 2005; Härting & Markowitsch, 1997; Kuelz et al., 2004; Shin et al., 2014) getroffen. Während die meisten Reviewer jedoch Studien mit nicht-standardisierten kognitiven Testverfahren ausschlossen, wurden in der vorliegenden Arbeit auch Studien berücksichtigt, deren Autoren spezielle Aufgaben zum Kategorisieren persönlicher Objekte selbst entwickelt hatten. Hiermit sollten symptomspezifische Defizite erfasst werden.

Über die Zuordnung der neuropsychologischen Konstrukte zu den übergeordneten Domänen besteht kein ganz eindeutiger Konsens. Überschneidungen finden sich insbesondere zwischen Aufmerksamkeit, Arbeitsgedächtnis und Exekutivfunktionen. So kann das mit dem TOL erfasste visuell-räumliche Planen/Problemlösen sowohl zu den Exekutivfunktionen als auch zum räumlichen Arbeitsgedächtnis gezählt werden (Owen et al., 1990), es kann in eine separate Kategorie visuell-räumlicher Leistungen eingeordnet (Härting & Markowitsch, 1997) oder aber der Aufmerksamkeit zugerechnet werden (Abramovitch et al., 2013). Ebenso kann die Inhibitionsleistung in der Stroop-Aufgabe zur Aufmerksamkeit (Heinrichs & Zakzanis, 1998; Kuelz et al., 2004) oder zu den Exekutivfunktionen (Abramovitch et al., 2013; Bora et al., 2010) gezählt werden. Die Reihe der Überlappungen und mehrdeutigen Zuordnungen wird hier aus Platzgründen nicht fortgesetzt. Eine eindeutige neuropsychologische „Taxonomie“ ist bislang nicht bekannt und aufgrund des komplexen Zusammenspiels biologisch-struktureller und funktioneller Faktoren auch kaum realisierbar. Im Zusammenhang mit dem Versuch, die neurokognitiven Testverfahren in Domänen und Subdomänen zu organisieren, sprachen Heinrichs und Zakzanis (1998, S. 428) sogar von Willkür: „an element of arbitrariness in test organization“. Entsprechend stellt die hier vorgenommene Aufteilung keine vollständig theoriegeleitete Kategorisierung dar. Andere Forscher können zu etwas anderen Darstellungsformen kommen. Dies ist von Bedeutung für die inhaltliche Interpretation der Ergebnisse über mehrere Studien hinweg: Wenn z. B. Minderleistungen im Stroop-Test je nach Autor als Aufmerksamkeitsmängel oder aber als Defizite in den Exekutivfunktionen aufgefasst werden, dann hat dies wiederum Auswirkungen auf die theoretische Einordnung der empirischen Befunde. Hier besteht Einigungsbedarf in der Forschungsgemeinschaft.

Literaturrecherche. Die Recherche umfasste drei Suchstrategien, d. h. eine systematische Suche in den einschlägigen Datenbanken anhand einer umfangreichen Auswahl an Schlagwörtern, eine Handsuche sowie Autorenkontakte. Es wurde auch graue Literatur gefunden und eingeschlossen, so dass ein Publikationsbias so weit wie möglich minimiert wurde. Gelegentlich schließen Reviewer graue Literatur mit der Begründung aus, dass unveröffentlichtes Material, das kein *peer review*

Verfahren durchlaufen hat, von geringerer Qualität sein könnte (Weisz, Weiss, Han, Granger & Morton, 1995). Jedoch stellt ein *peer review* Verfahren weder eine Garantie für hohe Qualität dar, noch ist es das einzige Qualitätsmerkmal, denn hochwertige Forschung kann z. B. von Regierungsbehörden oder Beratungsfirmen in anderer Form veröffentlicht werden, wie Borenstein (2009) zu bedenken gibt. Der Einschluss grauer Literatur in die vorliegende Metaanalyse hat einen Zuwachs an vier verwertbaren Untersuchungen gebracht. Dies entspricht bei einer Gesamtzahl von 38 Studien immerhin mehr als zehn Prozent. Eine dieser unveröffentlichten Studien (Kraft, 2006) erbrachte Resultate, die den veröffentlichten Studienergebnissen entgegenstanden, was auf einen Publikationsbias hinweisen könnte. Ein solcher Bias wurde durch das Einschließen der unveröffentlichten Studien weniger wahrscheinlich.

Auswahl der Studien. Selbst wenn eine Metaanalyse, wie im vorliegenden Fall, nach anerkannten Standards (QUOROM, PRISMA) durchgeführt wird, bleibt noch Raum für subjektive Entscheidungen des Reviewers, die kritisiert werden können. So mussten die Primärstudien u. a. das Einschlusskriterium erfüllen, dass Zwangspatienten einer anderen Symptomdimension als Vergleichsgruppe dienten. Im Verlauf der Recherche und Datenaufbereitung über vier Jahre hinweg stellte sich jedoch heraus, dass dieses Kriterium distinkter Gruppen auf sieben Studien aus den Jahren 2007 bis 2014 nicht zutraf. Diese Studien schlossen Patienten mit multiplen Zwangssymptomen ein und quantifizierten den Zusammenhang zwischen Symptomdimensionen und neuropsychologischen Leistungen anhand von Korrelationskoeffizienten. Um die Informationen dieser Studien zu berücksichtigen, wurden deren Korrelationskoeffizienten über eine Dichotomisierung in Hedges' g umgerechnet. Dies stellt eine Vereinfachung dar, die einer dimensionalen Sichtweise der Zwangssymptomatik nicht ganz gerecht werden kann. Die Alternative hätte jedoch darin bestanden, diese sieben Studien auszuschließen, so dass die derzeitige Forschungslage unvollständig abgebildet worden wäre. In nachfolgenden Untersuchungen könnte die dimensionale Sichtweise stärker berücksichtigt werden. Dass hierzu in den nächsten Jahren genügend Daten vorliegen werden, ist aufgrund

des exponentiellen Wachstums der Publikationszahlen auf dem Gebiet der Neuropsychologie zu erwarten.

Die Publikation in englischer, deutscher oder französischer Sprache stellte ein weiteres Einschlusskriterium dar. Durch die Berücksichtigung dreier Sprachen war dieses Kriterium weiter gefasst als in anderen Metaanalysen, die nur Publikationen in englischer Sprache einschlossen (Abramovitch et al., 2013; Henry, 2006; Shin, 2014). Ein *language bias* konnte dennoch nicht ganz verhindert werden, denn zwei Studien in persischer und chinesischer Sprache (Nejati, Zabihzadeh, Maleki & Safarzadeh, 2013; Zhu, Wang & Yi, 2013) mussten ausgeschlossen werden, nachdem die per E-Mail kontaktierten Autoren keine englische Übersetzung zur Verfügung stellen konnten.

Studien mit subklinischen Stichproben wurden ausgeschlossen, da es Vergleiche zwischen Patienten und Gesunden waren, die hier metaanalytisch untersucht wurden. Zwar lassen sich Zwangssymptome auf einem Kontinuum dimensional betrachten. Allerdings erscheint eine Abgrenzung von Schweregraden subklinischer Zwangssymptome innerhalb der Gruppe der Gesunden problematisch, da hierfür noch keine klaren Kriterien existieren. Eine erste Ausnahme bilden subklinische Kontrollzwänge, für die auf der *checking*-Subskala des MOCI *cut-off*-Werte von mindestens vier (Rubenstein, Peynircioglu, Chambless & Pigott, 1993) bzw. fünf von neun möglichen Punkten (Roth & Baribeau, 1996) vorgeschlagen wurden. Cuttler und Graf (2009) sowie Woods et al. (2002) hatten an studentischen Stichproben durchgeführte Studien in ihre Reviews mit einbezogen und damit für eine erste Symptomdimension die kontinuierliche Sichtweise berücksichtigt. Hingegen untersuchte die vorliegende Metaanalyse nicht nur eine, sondern mehrere Symptomdimensionen, während die diagnostischen Abgrenzungsprobleme für die jeweiligen Symptomdimensionen innerhalb der gesunden Probandengruppe bislang unzureichend gelöst sind. Daher wurde in der vorliegenden Arbeit eine dichotome Sichtweise gewählt, die nur zwei Kategorien zuließ, entweder (klinisch relevante) Zwangssymptome oder keine (klinisch relevanten) Zwangssymptome. In Zukunft könnte z. B. die neu entwickelte DY-BOCS (Rosario-Campos et al., 2006) vermehrt

auch bei Gesunden eingesetzt werden, um subklinische Stichproben hinsichtlich aller Symptomdimensionen zu untersuchen.

Studienqualität. Zur Erfassung der Studienqualität wurde ein Instrument entwickelt, das sich an den Cochrane-Kriterien für systematische Reviews von Interventionsstudien (Higgins & Green, 2011) und an der Skala von Öst (2008) orientierte. Die hier konstruierte Skala wurde auf Item-Ebene auf ihre Interrater-Reliabilität überprüft und ergab für zehn Items gute bis exzellente Reliabilitätswerte. Ein Item, *Spezifität der neuropsychologischen Testverfahren*, wurde aufgrund sehr geringer Interrater-Reliabilität ausgeschlossen. Da die Skala für jeden Punktwert konkrete Beschreibungen und Beispiele vorsah, beide Beurteilerinnen ein Training erhielten und Unklarheiten bis zum Konsens diskutiert wurden, erscheinen methodische Mängel eher unwahrscheinlich. Für die Breite eines neuropsychologischen Konstruktes gibt es bislang keine ganz eindeutigen Kriterien. Daher kann die geringe Übereinstimmung der Beurteilerinnen bei diesem Item stattdessen an den oben bereits angeführten Überlappungen bei der Zuordnung der Testverfahren zu neuropsychologischen Konstrukten liegen.

Auch wenn eines von elf Items nicht interpretierbar war und daher ausgeschlossen wurde, so bildeten die übrigen zehn Items doch eine reliable Skala. Die Studienqualität hatte, mit einer Ausnahme, in allen hier durchgeführten Metaanalysen keinen signifikanten Einfluss auf die Effektstärken. In einer Metaanalyse wurde ihr Einfluss zwar signifikant, blieb aber so niedrig, dass er nicht als praktisch relevant gelten konnte. Allerdings ist hier anzumerken, dass das Signifikanzniveau für die multiplen Testungen der Meta-Regressionsanalysen auf $\alpha = .01$ festgelegt wurde. Zudem kann nicht ausgeschlossen werden, dass Einflüsse der Studienqualität aufgrund geringer Stichprobengröße insignifikant blieben. So zeigte eine Metaanalyse von Lipszyc und Schachar (2010) über mehrere psychische Störungen hinweg, dass die Effektstärken umso niedriger ausfielen, je schlechter die Studienqualität war. Während, dem Standard der Cochrane-Kriterien (Higgins & Green, 2011) entsprechend, in den meisten Metaanalysen zu RCTs die Studienqualität untersucht wird, blieb in Metaanalysen zu Fall-Kontroll-Studien die

Studienqualität meist unberücksichtigt. So hat keine der Metaanalysen zur Neuropsychologie von Zwangsstörungen (Abramovitch et al., 2013; Cuttler & Graf, 2009; Henry, 2006; Shin et al., 2014; Woods et al., 2002) die Qualität der Primärstudien berichtet. Dies zeigt, dass es wünschenswert wäre, wenn zukünftig in den einzelnen Fall-Kontroll-Studien die Studienqualität systematisch beschrieben und in den entsprechenden Metaanalysen als Moderatorvariable berücksichtigt würde. Eine Forschungsaufgabe könnte darin bestehen, das hier entworfene Instrument anhand einer größeren Stichprobe zu evaluieren.

Stichproben. Die in den sieben Metaanalysen verglichenen Probandengruppen stellten teilweise sich überlappende Stichproben dar. Daher wurden in der deskriptiven Statistik keine Signifikanztests durchgeführt, die zu methodisch unsauberen inferenzstatistischen Schlüssen geführt hätten. Einerseits sind die Überlappungen der Stichproben zwar aus statistischer Perspektive wenig erwünscht. Andererseits bilden sie aber die Realität ab, da viele Zwangspatienten multiple Symptome aufweisen und da viele der eingeschlossenen Studien Mehrfachvergleiche durchgeführt hatten. Die Berücksichtigung der Stichprobenüberlappungen führte damit zu einer höheren externen Validität der Metaanalyse.

Signifikanztests. In den Profilen der Einzel-Effektstärken zeigte sich, dass ein Teil der Einzeleffekte signifikant wurde, während ein anderer Teil insignifikant blieb. Es läge nahe, die Anzahl signifikanter und insignifikanter Effektstärken zu vergleichen, wie dies z. B. Heinrichs und Zakzanis (1998) in ihrer Metaanalyse durchführten. Dieser Vorgang wird als *vote counting* bezeichnet. Dem *vote counting* liegt die Annahme zugrunde, dass ein signifikanter Befund einen tatsächlichen Effekt belegt, wohingegen ein insignifikanter Befund einen Nulleffekt belegt. Diese Annahme ist, so Borenstein (2009), irreführend. Während zwar ein signifikanter Befund eindeutig einen Effekt belegt, kann ein insignifikanter Befund entweder einen Nulleffekt belegen oder aber durch zu geringe Teststärke insignifikant bleiben. Da nicht klar war, ob aus den insignifikanten Einzeleffektstärken auf Nulleffekte oder auf zu geringe Stichprobengrößen zu schließen ist, wurde in der vorliegenden Metaanalyse auf *vote counting* verzichtet. Stattdessen wurden über mehrere

neuropsychologische Domänen und Subdomänen hinweg kombinierte Effektstärken errechnet, die durch die Integration mehrerer Studienstichproben eine höhere Teststärke erreichten.

Metaregression. Mit Metaregressionsanalysen kann, analog zur Kovarianzanalyse (ANCOVA), der Einfluss von Kovariaten auf einen Effekt ermittelt werden. Eine ANCOVA ist angebracht, um zufällige Einflüsse statistisch „herauszurechnen“. Zufälliges „Rauschen“, das z. B. durch unsystematische demografische Unterschiede entsteht, lässt sich in einer Primärstudie mit einer ANCOVA bzw. in einer Metaanalyse mit einer Metaregression kontrollieren. Systematische Gruppenunterschiede hingegen lassen sich nicht „herausrechnen“. Solche systematischen Gruppenunterschiede zwischen Zwangspatienten und Gesunden finden sich bei Depressivität, Angst und Einflüssen durch medikamentöse Behandlung. Während Gesunde definitionsgemäß weder depressiv noch ängstlich noch mit Psychopharmaka behandelt sind, ist das Konstrukt der Zwangserkrankung mit Depressivität, Angst und Einflüssen von Medikamenten eng verbunden. Zwei Drittel (Tükel, Polat, Özdemir, Aksüt & Türksöy, 2002) bis 90 % der Zwangspatienten (Ruscio et al., 2010) leiden unter komorbiden Störungen, zahlreiche Zwangspatienten werden im Verlauf ihrer Erkrankung depressiv. Fast alle Zwangspatienten leiden unter Ängsten, dass etwas Verhängnisvolles geschehen könnte, entsprechend ist die Zwangsstörung im DSM-IV den Angststörungen zugeordnet. Über zwei Drittel der Zwangspatienten erhalten Psychopharmaka. Ein statistisches „Herausrechnen“ dieser systematischen Unterschiede zu Gesunden würde dazu führen, dass das Konstrukt *Zwangsstörung* nicht mehr valide und generalisierbar ist. Es handelt sich hier um ein grundsätzliches Dilemma, das häufig auftritt, wenn psychopathologische Konstrukte mit quantitativen Methoden untersucht werden: Um eine hohe interne Validität zu erreichen, wäre ein möglichst „reines“ Konstrukt, hier eine möglichst „reine“ Zwangserkrankung, wünschenswert. Dies widerspricht aber dem Wunsch nach hoher Konstruktvalidität und externer Validität, denn psychische Störungen sind komplexe Phänomene. Die Frage, wie die neuropsychologischen Leistungen von Zwangspatienten aussähen, die weder depressiv noch ängstlich noch medikamentös behandelt sind, lässt sich mittels einer

ANCOVA bzw. einer Metaregression nicht beantworten. Für eine solche Fragestellung existiert kein statistisches Verfahren (Miller & Chapman, 2001).

6.2 Metaanalyse 1: Patienten mit Waschwang vs. Gesunde

Über acht Studien hinweg zeigten Zwangspatienten mit Waschwang zunächst in 8 von 21 Einzeltests eine signifikant schwächere Leistung als die gesunden Kontrollen. Zwar erbrachten die Patienten mit Waschwang signifikant schlechtere Leistungen in anhaltender Aufmerksamkeit, im Konzeptwechsel, im verbalen Gedächtnis und im visuell-räumlichen Arbeitsgedächtnis. Allerdings ergaben Metaregressionsanalysen, dass diese Minderleistungen sowohl über die gesamten neuropsychologischen Tests als auch speziell im räumlichen Arbeitsgedächtnis auf signifikante Unterschiede im Intelligenz- und Bildungsniveau der Probanden zurückzuführen waren. Ob auch die Leistungsdefizite der Patienten mit Waschwang in anhaltender Aufmerksamkeit, im Konzeptwechsel und im verbalen Gedächtnis auf Intelligenz- und Bildungsunterschiede zurückgingen, ließ sich aufgrund der geringen Datenlage nicht ermitteln. Im nonverbalen Gedächtnis und im verbalen Arbeitsgedächtnis unterschieden sich die Patienten mit Waschwang nicht signifikant von den Gesunden. Zusammengefasst lieferte die Metaanalyse keine Evidenz dafür, dass Patienten mit Waschwang bei gleicher Intelligenz und gleichem Bildungsniveau signifikant schlechtere neuropsychologische Leistungen zeigen als Gesunde. Dieses Ergebnis zeigt, dass es sinnvoll ist, die Symptomdimensionen der Zwangsstörung differenziert zu untersuchen, anstatt alle Zwangspatienten als eine homogene Gruppe zu betrachten. Letzteres hatte, wie bereits in der Einleitung beschrieben, zu widersprüchlichen Ergebnissen geführt (z. B. Abbruzzese et al., 1995; Aronowitz et al., 1994; Basso et al., 2001; Christensen et al., 1992; Deckersbach et al., 2000; Moritz, Fischer et al., 2008; Moritz, Kloss et al., 2009; Schmidtke et al., 1998; Simpson et al., 2006; Zielinski et al., 1991), die in der Zusammenschau kaum interpretierbar erschienen.

6.3 Metaanalyse 2: Patienten mit Kontrollzwang vs. Gesunde

Kurz zusammengefasst, schnitten die Patienten mit Kontrollzwang in der Inhibition, im Konzeptwechsel, im verbalen und nonverbalen Gedächtnis sowie im räumlichen und verbalen Arbeitsgedächtnis signifikant schlechter ab als die Gesunden. Unterschiede im Intelligenz- und Bildungsniveau hatten hierauf keinen signifikanten Einfluss. Über 16 Studien hinweg erbrachten Zwangspatienten mit Kontrollzwang in 21 von 38 Einzeltests eine signifikant schwächere Leistung als die gesunden Kontrollen. Während keine Tests zum Planen/Problemlösen durchgeführt worden waren, schnitten die Patienten mit Kontrollzwang in den Subdomänen Inhibition, Konzeptwechsel, verbales und nonverbales Gedächtnis, räumliches und verbales Arbeitsgedächtnis signifikant schlechter ab als die Gesunden. Unterschiede im Intelligenz- und Bildungsniveau hatten hierauf keinen signifikanten Einfluss. In den Subdomänen Enkodierung und Verarbeitungsgeschwindigkeit fanden sich keine signifikanten Unterschiede.

Die Defizite der Patienten mit Kontrollzwang in der Inhibition und im Konzeptwechsel fügen sich widerspruchsfrei in das Modell von Chamberlain et al. (2005), das einen Zusammenhang von Kontrollritualen auf Symptomebene, mangelhaften Inhibitionsprozessen sowie Minderleistungen in den entsprechenden Testverfahren wie Go/No-Go-Aufgaben und dem WCST annahm (vgl. Abbildung 5 in Kapitel 1). Nachdem Tests wie der WCST speziell dafür entwickelt wurden, Defizite in Frontalhirnfunktionen zu erkennen (Milner, 1963), entsprechen die Leistungsdifferenzen der Patienten mit Kontrollzwang im Konzeptwechsel/Regellernen auch plausibel dem cortico-baso-thalamischen Modell (Menzies et al., 2008; Saxena et al., 1998), das Zusammenhänge zwischen perseverierendem Zwangsverhalten und Dysfunktionen u.a. in Frontalhirnregionen postuliert.

Die Patienten mit Kontrollzwang zeigten in sämtlichen erfassten Gedächtnisbereichen signifikante Minderleistungen. Damit stehen die Resultate der vorliegenden Metaanalyse im Einklang mit den Ergebnissen der Metaanalyse von Woods et al. (2002), die ebenfalls zu signifikanten Unterschieden zwischen Probanden mit und ohne Kontrollzwang im nonverbalen und verbalen Gedächtnis

kamen. Im Gegensatz zu Woods et al. (2002) fanden Cuttler und Graf (2009) für nonverbale und verbale episodische Gedächtnisleistungen keine so eindeutigen Belege. Allerdings ermittelten Cuttler und Graf (2009) keine kombinierten Effektstärken, sondern stellten ihre Ergebnisse als *vote counting* dar. Dieses Verfahren wurde bereits kritisiert, denn es kann zu der falschen Schlussfolgerung führen, insignifikante Ergebnisse als Beleg für einen Nulleffekt und nicht als Hinweis auf zu geringe statistische Teststärke anzusehen.

Als Erklärung für die verbalen Gedächtnisdefizite wird diskutiert, dass die Patienten mit Kontrollzwang beim Enkodieren weniger organisationale Strategien wie semantisches *clustering* anwenden (Deckersbach et al., 2000), die mit dem Präfrontalcortex in Zusammenhang stehen. Ob unzureichend eingesetzte Organisationsstrategien beim Enkodieren die Gedächtnisdefizite erklären können, ließ sich in dieser Metaanalyse nicht belegen, denn die Fähigkeit zum Enkodieren wurde lediglich in drei Einzelstudien untersucht. Zwei der drei Studien berichteten hierin zwar signifikante Mängel der Patienten mit Kontrollzwang, allerdings wurde in der Kombination aller drei Studienergebnisse die Gesamteffektstärke insignifikant. Hier sind weitere Untersuchungen anhand größerer Stichproben erforderlich, um eindeutigere Ergebnisse zu erhalten. Auch Bildgebungsstudien könnten weiteren Aufschluss geben. Wie Casasanto et al. (2002) in einer fMRT-Studie an gesunden Probanden herausfanden, liegt erfolgreichen vs. erfolglosen Enkodierstrategien verbaler Inhalte nicht nur eine unterschiedlich stark ausgeprägte präfrontale Aktivierung zugrunde. Darüber hinaus unterscheiden sich Probanden mit besserer vs. schlechterer verbaler Gedächtnisleistung auch in der corticalen Lateralität. So korrelierte eine bessere verbale Gedächtnisleistung mit vermehrter Aktivität in linkshemisphärischen präfrontalen Strukturen, während eine schlechtere verbale Gedächtnisleistung mit erhöhter Aktivierung in rechtshemisphärischen temporoparietalen Strukturen zusammenhing (ebenda). Aus neurobiologischer Perspektive bedeutet dies, dass den Defiziten im Enkodieren verbaler Information Unterschiede in der Aktivierung der rechten und der linken Gehirnhälfte zugrunde liegen. Ob dies speziell auch auf Patienten mit Kontrollzwang zutrifft, könnte ein weiteres Thema für zukünftige Studien sein.

6.4 Metaanalyse 3: Patienten mit verbotenen Gedanken vs. Gesunde

Zusammengefasst, zeigten Zwangspatienten mit verbotenen Gedanken deutlich verminderte Leistungen in Einzeltests zum Planen/Problemlösen, zum Konzeptwechsel und zum räumlichen Arbeitsgedächtnis. Allerdings ließen sich aus nur zwei eingeschlossenen Studien keine endgültigen Aussagen treffen.

Zwangspatienten mit verbotenen Gedanken zeigten in fünf von sechs Tests eine signifikant schwächere Leistung als die gesunden Kontrollen. Es wurden Leistungen in Aufmerksamkeit, Planen/Problemlösen und Konzeptwechsel/Regellernen untersucht, wohingegen Gedächtnisleistungen nicht erfasst worden waren. Aufgrund der geringen Studienzahl konnte nur eine kombinierte Effektstärke berechnet werden: Im visuell-räumlichen Arbeitsgedächtnis fand sich ein großer Effekt, der im *random effects model* signifikant wurde, nicht jedoch im *fixed effect model*. Da hier lediglich zwei Studieneffekte zusammengefasst wurden, von denen der eine hoch signifikant, der andere insignifikant war, kann das kombinierte Resultat aufgrund der geringen Stichprobe insignifikant geblieben sein. Um ein eindeutigeres Ergebnis zu erhalten, wären mehr Primärstudien erforderlich gewesen.

Dass bislang so wenige Untersuchungen zu Zwangspatienten mit verbotenen Gedanken publiziert wurden, ist erstaunlich. Gerade diese Symptomdimension zeichnet sich durch besonders schwierige Bedingungen aus. So weisen, wie bereits in Kapitel 2.3 angeführt, Verwandte von Zwangspatienten mit verbotenen Gedanken ein erhöhtes Risiko auf, selbst an einer Zwangsstörung zu erkranken (Alsobrook et al., 1999). Speziell für diese Symptomdimension ist eine genetische Komponente hoch wahrscheinlich. Entsprechend könnte es lohnenswert sein, bei Zwangspatienten mit verbotenen Gedanken verstärkt nach endophänotypischen Kandidatengenen zu suchen und einen möglichen Neuroendophänotypen als Vulnerabilitätsmarker für diese Symptomdimension in weiteren Studien genauer zu erforschen.

In einer aktuellen Untersuchung wiesen Brakoulias et al. (2013) darauf hin, dass bei Zwangspatienten mit verbotenen Gedanken mit einer Häufigkeit von über 20 % vermehrt Alkohol- oder Drogenabhängigkeit auftreten. Die Autoren vermuten, dass beim Ausführen von Zwangshandlungen die gleichen Pfade im Belohnungssystem des Gehirns aktiviert werden wie bei Substanzabhängigkeit. Da Patienten mit verbotenen Gedanken kaum belohnungsrelevante Zwangshandlungen ausführen, weisen sie, so die Autoren, ein erhöhtes Risiko für Abhängigkeitserkrankungen auf. Bei Substanzabhängigkeit finden sich neben Frontalhirnschädigungen vor allem Auffälligkeiten im limbischen System mit spezifischen Gedächtnisproblemen und in der rechten Hemisphäre mit umfassenden nonverbalen Defiziten (Steingass, 1999). Um solche Komorbiditäten zu erkennen, wäre bei neuropsychologischen Untersuchungen von Patienten mit verbotenen Gedanken eine besonders sorgfältige Diagnostik bezüglich aktuellen und vor allem auch bezüglich vergangenen Alkohol- und Drogenkonsums angeraten.

6.5 Metaanalyse 4: Patienten mit Symmetrie-/Ordnungszwang vs. Gesunde

Zusammengefasst, zeigten Patienten mit Symmetrie-/Sammelzwang vor allem im verbalen und nonverbalen Gedächtnis signifikante Defizite, verglichen mit Gesunden. Einschränkend ist anzumerken, dass dieses Ergebnis aufgrund von nur vier eingeschlossenen Studien zustande kam.

Über vier Studien hinweg erbrachten Zwangspatienten mit Symmetrie-/Ordnungszwang in 3 von 21 Tests eine signifikant schwächere Leistung als die gesunden Kontrollen, und zwar im TMT-B, im WMS-R LM und im RCFT. In den Subdomänen verbales Gedächtnis und nonverbales Gedächtnis wiesen die Patienten mit Symmetrie-/Ordnungszwang signifikant schlechtere Leistungen auf. Ob die reduzierten Gedächtnisleistungen mit mangelnden Organisationsstrategien im Enkodieren zusammenhängen, kann nicht beantwortet werden, da keine der eingeschlossenen Studien hierzu Testergebnisse berichtete.

Bei der Durchführung neuropsychologischer Untersuchungen von Patienten mit Symmetrie-/Ordnungszwang kann das Ordnungsniveau des Testlabors eine Rolle spielen. Probanden mit hohem Symmetrie-/Ordnungsbedürfnis empfanden im Vergleich mit Probanden mit geringem Symmetrie-/Ordnungsbedürfnis signifikant mehr Angst, wenn sie in einer unordentlichen Umgebung eine schwierige Aufgabe lösen sollten (Radomsky & Rachman, 2004). Dass Angst sich negativ auf kognitive Leistungen auswirkt, ist hinlänglich bekannt. Aktuelle neurobiologische Befunde zeigen, dass die Informationen bei Angst zum einen über eine sehr schnelle Aktivierung subcorticaler Netzwerke und der Amygdala zu Hypervigilanz führen, während etwas später der ACC und der PFC im Sinne kognitiver Vermeidung und Inhibition aktiviert werden, so dass der angstauslösende Stimulus weniger „tief“ verarbeitet wird (Hofmann et al., 2012). Zu der Tatsache, dass eine „unordentliche“ Umgebung bei Patienten mit Symmetrie-/Ordnungszwang Angst auslöst und dass dies die Testleistungen negativ beeinflussen kann, wurde in den eingeschlossenen Untersuchungen nichts erwähnt. In zukünftigen Untersuchungen sollte dies jedoch berücksichtigt und das Ordnungsniveau der Testlabors möglichst konstant gehalten werden, um eine unbeabsichtigte Symptomprovokation zu vermeiden.

6.6 Metaanalyse 5: Patienten mit Sammelzwang vs. Gesunde

Zusammengefasst, zeigten Patienten mit Sammelzwang im Vergleich zu Gesunden ein deutlich höheres Durchschnittsalter, einen höheren Frauenanteil und signifikante Defizite im Planen/Problemlösen, in der Inhibition, im Konzeptwechsel, im nonverbalen Gedächtnis, im räumlichen Arbeitsgedächtnis, im Entscheidungsverhalten und in der Kategorisierung. Die letzteren zwei kognitiven Funktionsbereiche wurden bislang, so der Stand dieser Untersuchung, nur bei Patienten mit Sammelzwang untersucht. Dies hängt mit der Psychopathologie dieses Störungsbildes zusammen. So korreliert Sammelzwang z. B. mit dem Vermeiden von Entscheidungen (Frost & Gross, 1993), und dieses Vermeidungsverhalten hängt möglicherweise mit der Angst, einen Fehler zu machen, zusammen (Frost & Hartl, 1996). Verstärkte Angst empfinden Personen mit Sammelzwang auch bei Aufgaben zum Kategorienbilden (Grisham et al., 2010).

Es fiel auf, dass sich die Patienten mit Sammelzwang in demografischen Merkmalen deutlich von den Gesunden unterschieden. So waren die Patienten mit Sammelzwang im Schnitt zehn Jahre älter. Zudem lag ihr Frauenanteil nicht wie bei den anderen Stichproben bei der Hälfte, sondern bei zwei Dritteln. Zwar zeigten sich in den Metaregressionsanalysen keine Einflüsse der demografischen Unterschiede auf die neuropsychologischen Effektstärken. Allerdings leiden Frauen mit Sammelzwang häufiger als betroffene Männer an schweren komorbiden Störungen wie bipolarer Störung und Alkoholabhängigkeit (Wheaton, Timpano, LaSalle-Ricci & Murphy, 2008), was sich wiederum auf die kognitiven Leistungen auswirken kann. Hinsichtlich dieser Unterschiede wäre zu wünschen, dass die Stichproben in zukünftigen Studien sorgfältiger parallelisiert werden.

Die Patienten mit Sammelzwang zeigten über 14 Studien hinweg in 14 von 44 Einzeltests eine signifikant schwächere Leistung als die gesunden Kontrollen. Die Patienten mit Sammelzwang wiesen signifikante Defizite im Planen/Problemlösen, in der Inhibition, im Konzeptwechsel, im nonverbalen Gedächtnis, im räumlichen Arbeitsgedächtnis, im Entscheidungsverhalten und in der Kategorisierung auf. Damit zeigten sie schwerpunktmäßig Beeinträchtigungen in den Exekutivfunktionen. Ein möglicher Zusammenhang zwischen Gedächtnisdefiziten und dem höheren Alter der Patienten mit Sammelzwang wurde in der Metaregression nicht belegt. Die Befunde dieser Metaanalyse stehen im Einklang mit einigen der Befunde, die Mataix-Cols, Pertusa und Snowden (2011) in ihrem ausführlichen Review zur Neuropsychologie und Neurobiologie von Sammelzwang berichteten. Die Autoren fanden neben Gedächtnisproblemen vor allem Defizite im Entscheidungsverhalten, in der Kategorisierung und in der Organisationsfähigkeit. Auch von anderen Forschern wurde angeführt, dass Menschen mit Sammelzwang exekutive Dysfunktionen aufweisen, die die Fähigkeit zur Informationsverarbeitung beeinträchtigen (McMillan, Rees & Pestell, 2013). Demnach scheint bei Patienten mit Sammelzwang eine deutliche Beteiligung von Frontalhirnfunktionen vorzuliegen, die zunehmend auch neurobiologisch erforscht werden. So fanden Alvarenga et al. (2012) heraus, dass hohe Werte auf der Dimension *Sammeln* mit erhöhtem Volumen der Grauen Substanz im linken superioren lateralen OFC korreliert waren, was bei den anderen

Symptomdimensionen nicht der Fall war. Dold und Aigner (2009) berichteten in ihrem Review, dass bei Symptomprovokation von Sammelzwang besonders jene Gehirnareale hoch aktiv waren, die für das Entscheidungsverhalten von Bedeutung sind (ventromediale Teile des OFC und das dorsale ACC). Auch Steketee, Frost und Kyrios (2003) berichteten, dass eingeschränktes Entscheidungsverhalten einen bedeutsamen Prädiktor für Sammelzwang darstellte, während Harrison et al. (2013) aufzeigten, dass Sammelzwang mit der Stärke der ventralen und dorsalen striatalen Verbindungen zu den frontalen Hirnregionen zusammenhing. Wie Mataix-Cols et al. (2011) in ihrem Review zusammenfassten, hängt Sammelzwang auch mit Auffälligkeiten im ventromedialen PFC, im ACC und in subcorticalen limbischen Strukturen zusammen.

6.7 Metaanalyse 6: Patienten mit Waschzwang vs. Kontrollzwang

Überraschenderweise fand sich unter den 13 Studien, die Patienten mit Waschzwang und Patienten mit Kontrollzwang hinsichtlich ihrer Neuropsychologie verglichen, keine einzige US-amerikanische Studie. Diese „nationale Forschungslücke“ ist erstaunlich, da die USA eine der führenden Forschungsnationen auf diesem Gebiet darstellen.

Zusammengefasst, zeigten Patienten mit Kontrollzwang in keiner Subdomäne oder Domäne bessere Leistungen als die Patienten mit Waschzwang. Über 13 Studien hinweg zeigten Patienten mit Kontrollzwang in 11 von 40 Tests signifikant schlechtere Testleistungen in anhaltender Aufmerksamkeit, im Planen/ Problemlösen, in der Inhibition, im Konzeptwechsel, im Enkodieren, im verbalen und nonverbalen Gedächtnis sowie in der Verarbeitungsgeschwindigkeit. Unterschiede im Intelligenz- und Bildungsniveau hatten hierauf keinen signifikanten Einfluss. Die größten Defizite zeigten Patienten mit Kontrollzwang im Planen/Problemlösen und in der Inhibition. In den Subdomänen räumliches und verbales Arbeitsgedächtnis fanden sich keine signifikanten Unterschiede.

Die Mängel „der Zwangspatienten“ im Planen/Problemlösen erklärten Chamberlain et al. (2005) damit, dass verzögerte Bearbeitungszeiten aus Strategiedefiziten beim Enkodieren, aus Aufmerksamkeitsdefiziten oder aber aus chronischem Zweifel resultieren können. Defizite in der kognitiven Inhibition stehen gemäß Chamberlain et al. (2005) in Zusammenhang mit Zwangsgedanken, während ein Versagen behavioraler Inhibition mit Zwangshandlungen zusammenhängt (vgl. Abbildung 3). Inhibitionsdefizite ließen sich hier über mehrere Studien hinweg metaanalytisch bestätigen. Bezüglich der Einschränkungen im verbalen und nonverbalen Gedächtnis wird angenommen, dass sie durch mangelhafte Organisationsstrategien zustande kommen, wie sie z. B. beim Abzeichnen der Rey-Osterrieth-Figur erforderlich sind (Kim, Park, Shin & Kwon, 2002; Kuelz et al., 2004; Savage et al., 1999).

Dass sich Patienten mit Waschzwang und Kontrollzwang in zahlreichen neuropsychologischen Funktionsbereichen signifikant unterscheiden, ist ein bedeutsames Resultat dieser Metaanalyse. Es liefert weitere empirische Evidenz für die Validität der Symptomheterogenität der Zwangsstörung. Der Befund wird dadurch untermauert, dass er im Einklang steht mit den Ergebnissen der Metaanalysen Nr. 1 und Nr. 2, in denen Patienten mit Waschzwang mit Gesunden (Kapitel 5.5 und 6.2) bzw. Patienten mit Kontrollzwang mit Gesunden (Kapitel 5.6 und 6.3) bezüglich ihrer Neuropsychologie untersucht wurden. Hierbei hatte sich gezeigt, dass Patienten mit Waschzwang bei vergleichbarer Intelligenz und gleichem Bildungsniveau nicht signifikant schlechter abschnitten als Gesunde, wohingegen Patienten mit Kontrollzwang in der Inhibition, im Konzeptwechsel, im verbalen und nonverbalen Gedächtnis sowie im räumlichen und verbalen Arbeitsgedächtnis signifikant schlechtere Testergebnisse erbrachten als Gesunde.

6.8 Metaanalyse 7: Patienten mit vs. ohne Sammelzwang

Zusammengefasst, waren die Patienten mit Sammelzwang im Schnitt mehr als zehn Jahre älter als die übrigen Zwangspatienten, und ihr Frauenanteil lag deutlich höher. Widersprüchliche und insignifikante Ergebnisse zeichneten für diese Metaanalyse insgesamt ein unklares Bild. Daher blieb offen, ob, und wenn ja, in

welchen Bereichen Patienten mit Sammelzwang kognitiv schwerer beeinträchtigt sind als Zwangspatienten ohne Sammelzwang. Lediglich im Entscheidungsverhalten zeigten die Patienten mit Sammelzwang signifikant schlechtere Leistungen als die anderen Zwangspatienten, wenn auch dieser Leistungsunterschied gering ausfiel.

Patienten mit Sammelzwang waren deutlich älter als alle anderen Patientengruppen. Dieses Ergebnis deckt sich mit der Publikationslage – es wurden zahlreiche Studien veröffentlicht, die Sammelzwang bei älteren Personen untersuchten (Ayers, Saxena, Golshan & Wetherell, 2010; Diefenbach, DiMauro, Frost, Steketee & Tolin, 2013; Kim, Steketee & Frost, 2001). Auch Steketee et al. (2003) berichteten einen Altersunterschied von über 15 Jahren zwischen Patienten mit vs. ohne Sammelzwang. Zudem lag in den eingeschlossenen Studien der Frauenanteil bei Patienten mit Sammelzwang deutlich höher als bei den anderen Patientengruppen. Da bei Patienten mit Sammelzwang Geschlechterunterschiede in der Häufigkeit komorbider Abhängigkeitserkrankungen und bipolarer Störungen und infolgedessen auch in kognitiver Leistungsfähigkeit auftreten (vgl. Abschnitt 6.6), wäre ein paarweises Parallelisieren der Stichproben auch hier wünschenswert.

Patienten mit Sammelzwang zeigten über sechs Studien hinweg in einem von 13 Tests eine signifikant bessere und in 3 von 13 Tests eine signifikant schwächere Leistung als Zwangspatienten ohne Sammelzwang. Die einzelnen Testergebnisse waren derart widersprüchlich, dass die kombinierten Effekte in drei von vier untersuchten Domänen (Planen/Problemlösen, Inhibition, Konzeptwechsel) insignifikant blieben. Lediglich im Entscheidungsverhalten zeigten die Patienten mit Sammelzwang ein signifikantes, aber kleines Defizit.

Problematisch ist zu bewerten, dass die Stichproben in den Primärstudien unterschiedlich definiert worden waren. Patienten mit Sammelzwang waren zum einen Zwangspatienten, die Sammelzwang als Symptom aufwiesen (*OCD-based hoarding*), zum anderen Patienten mit der Störung *hoarding disorder*. Darüber hinaus stellte die Gruppe der Zwangspatienten ohne Sammelzwang eine heterogene Gruppe dar, die nicht durch den Schwerpunkt auf einer Symptomdimension, sondern durch deren Abwesenheit bestimmt wurde. Daher ist es kaum sinnvoll, aus

Vergleichen mit einer derart heterogenen Gruppe eindeutige Schlussfolgerungen zu ziehen. Günstiger wäre es gewesen, Patienten mit Sammelzwang mit genau einer anderen Symptomdimension, z. B. mit Patienten mit Waschzwang zu vergleichen (vgl. Metaanalyse Nr. 6, Waschzwang vs. Kontrollzwang). Dies war aufgrund der Datenlage nicht möglich. Das heißt, dass die Ergebnisse zum Vergleich von Patienten mit und ohne Sammelzwang nur als vorläufig gelten sollten. In Zukunft könnten Studien mit klarer definierten Stichproben durchgeführt werden, in denen zum einen die *hoarding disorder* als eigenständige Diagnose, zum anderen homogenere Patientengruppen einzelner Symptomdimensionen untersucht werden.

6.9 Generelle Diskussion

In diesem letzten Kapitel werden zunächst die Ergebnisse noch einmal kurz zusammengefasst. Anschließend werden die Begrenzungen und Stärken dieser Dissertation beschrieben, bevor die Arbeit mit einem Ausblick auf die Zukunft endet.

Symptomdimensionen. Die signifikanten Ergebnisse über alle sieben Metaanalysen hinweg (vgl. Kapitel 5.12) weisen darauf hin, dass jede Symptomdimension ihr eigenes, spezifisches neuropsychologisches Profil aufweist. So fand sich keine Evidenz dafür, dass Patienten mit Waschzwang bei gleichem Intelligenz- und Bildungsniveau signifikant schlechtere neuropsychologische Leistungen zeigen als Gesunde. Im Kontrast dazu schnitten die Patienten mit Kontrollzwang in der Inhibition, im Konzeptwechsel, im verbalen und nonverbalen Gedächtnis sowie im räumlichen und verbalen Arbeitsgedächtnis signifikant schlechter ab als die gesunden Probanden. Auch Zwangspatienten mit verbotenen Gedanken zeigten deutlich verminderte Leistungen in Einzeltests zum Planen/Problemlösen, Konzeptwechsel und zum räumlichen Arbeitsgedächtnis. Allerdings ließen sich aus nur zwei eingeschlossenen Studien keine endgültigen Aussagen treffen. Hingegen wiesen Patienten mit Symmetrie-/Ordnungszwang vor allem im verbalen und nonverbalen Gedächtnis signifikante Defizite auf, verglichen mit Gesunden. Patienten mit Sammelzwang unterschieden sich von Gesunden durch ein deutlich höheres Alter, einen höheren Frauenanteil und signifikante Defizite im Planen/Problemlösen, in der Inhibition, im Konzeptwechsel, im nonverbalen

Gedächtnis, im räumlichen Arbeitsgedächtnis, im Entscheidungsverhalten und in der Kategorisierung.

Im Vergleich zu Patienten mit Waschzwang erbrachten Patienten mit Kontrollzwang in keiner Subdomäne oder Domäne die besseren Leistungen. Patienten mit Kontrollzwang zeigten Defizite in anhaltender Aufmerksamkeit, im Planen/Problemlösen, in der Inhibition, im Konzeptwechsel, im Enkodieren, im verbalen und nonverbalen Gedächtnis sowie in der Verarbeitungsgeschwindigkeit. Unterschiede in Intelligenz- und Bildungsniveau hatten hierauf keinen signifikanten Einfluss. Gegenüber Patienten mit Waschzwang wiesen Patienten mit Kontrollzwang die größten Defizite im Planen/Problemlösen und in der Inhibition auf.

Im Vergleich zu Zwangspatienten ohne Sammelzwang waren die Patienten mit Sammelzwang im Schnitt mehr als zehn Jahre älter, und ihr Frauenanteil lag deutlich höher. Eine widersprüchliche Datenlage ließ offen, ob, und wenn ja, in welchen Bereichen Patienten mit Sammelzwang kognitiv schwerer beeinträchtigt sind als Zwangspatienten ohne Sammelzwang. Lediglich im Entscheidungsverhalten schnitten die Patienten mit Sammelzwang etwas schlechter ab als die anderen Zwangspatienten.

Die Befunde der hier durchgeführten Metaanalysen können als Beleg für die Validität der verschiedenen Symptomdimensionen der Zwangsstörung gelten. Die Ergebnisse bestätigen die zu Beginn dieser Arbeit angestellten Überlegungen, warum bisherige Reviews zu so uneinheitlichen Ergebnissen kamen (Abramovitch et al., 2013; Chamberlain et al., 2005; Cox, 1997; Greisberg & McKay, 2003; Härting & Markowitsch, 1997; Henry, 2006; Kuelz et al., 2004; Muller & Roberts, 2005; Olley et al., 2007) – weil möglicherweise die Symptomheterogenität nicht beachtet wurde. In einer der aktuellsten Metaanalysen zur Neuropsychologie der Zwangsstörung (Abramovitch et al., 2013) empfahlen die Autoren, die Heterogenität der Symptomdimensionen zu berücksichtigen. Denn wenn man beispielsweise Patienten mit Waschzwang (die kaum neuropsychologisch beeinträchtigt erscheinen, wie Metaanalyse Nr. 1 zeigte) mit Patienten mit Kontrollzwang (die neuropsychologisch deutlich beeinträchtigt sind, wie die Metaanalysen Nr. 2 und Nr. 6 zeigten) als eine

einheitliche Gruppe betrachtet, dann wird dies die Leistungen der Patienten mit Waschzwang unterschätzen und die Leistungen der Patienten mit Kontrollzwang überschätzen. Die Empfehlung von Abramovitch et al. (2013) wurde hier umgesetzt; insofern reiht sich die vorliegende Arbeit in die aktuelle Forschungslage ein.

Während konsistent für alle Patientengruppen (außer für die Patienten mit Waschzwang) grundlegende Defizite im visuell-räumlichen Arbeitsgedächtnis festgestellt wurden, finden sich zahlreiche Publikationen, die Minderleistungen in den Exekutivfunktionen als übergeordnetes Defizit betrachten (Kashyap, Kumar, Kandavel & Reddy, 2013; Penadés, Catalán, Andrés, Salamero & Gastó, 2005; Savage et al., 1999; Shin et al., 2010). Von Bedeutung ist, wie Cox (1997) zu bedenken gibt, dass auch basale kognitive Fähigkeiten intakt sein müssen, um eine ungewohnte, neue Testaufgabe zu lösen. Hierzu gehört, so die Autorin, dass der Patient keine sensomotorischen Einschränkungen bei der Wahrnehmung des Materials hat, dass er die Aufgabe verstanden hat und dass er weiß, wie er reagieren soll. Entsprechend können basale Aufmerksamkeitsmängel und übergeordnete exekutive Beeinträchtigungen Folgedefizite in anderen neuropsychologischen Funktionen wie z. B. im Gedächtnis nach sich ziehen.

Nachdem die Metaanalysen zeigten, dass einerseits Patienten mit Waschzwang bei gleicher Intelligenz und gleichem Bildungsniveau genauso gute neuropsychologische Testleistungen erbrachten wie Gesunde, andererseits aber die Zwangspatienten aller anderen Symptom Schwerpunkte schlechtere Leistungen in zahlreichen kognitiven Subdomänen aufwiesen, stellt sich die Frage nach der klinisch-praktischen Relevanz dieser Unterschiede. Neuropsychologische Leistungsdifferenzen von weniger als zwei Standardabweichungen werden im Allgemeinen nicht als praktisch bedeutsam angesehen (z. B. Abramovitch et al., 2013). Legt man dieses Kriterium an, dann wird man die meisten hier gefundenen Leistungsdefizite der Zwangspatienten als praktisch irrelevant bewerten. Nichtsdestotrotz zeigen die Befunde, dass sich die Symptomdimensionen in ihrer Neuropsychologie deutlich unterscheiden. Neben genetischen, neurobiologischen,

komorbid-epidemiologischen und therapeutischen Unterschieden kann dies als ein weiterer Beleg für die Validität des Symptomdimensions-Ansatzes gelten.

Begrenzungen. Eine wesentliche Begrenzung dieser Metaanalyse bestand in der lückenhaften Datenlage. Zu zahlreichen Vergleichen zwischen Symptomdimensionen ließen sich keine Publikationen finden, wie das Schaubild zum *umbrella review* (Abbildung 29 in Abschnitt 6.1) zeigte. Zu Fragestellungen, die bislang nicht in Primärstudien untersucht wurden, ließen sich entsprechend keine Daten finden, die metaanalytisch hätten verwertet werden können. Daher blieben die Resultate insbesondere zu den Symptomdimensionen *verbotene Gedanken* und *Symmetrie-/Ordnungszwang* vorläufig. Auch konnten nicht für alle kognitiven Funktionsbereiche Effektstärken berechnet werden. Es fiel auf, dass zum Einsatz von Organisationsstrategien beim Enkodieren so wenige Studien publiziert worden waren, dass dies nur bei den Patienten mit Kontrollzwang im Vergleich zu Gesunden metaanalytisch untersucht werden konnte. Dies ist besonders erstaunlich, da in vielen theoretischen Ansätzen mögliche Defizite von Zwangspatienten im Einsatz organisationaler Strategien beim Enkodieren postuliert wurden. Hier wären weitere Studien wünschenswert.

Überwiegend kleine Probandenzahlen in den Primärstudien stellten eine Schwäche dar, denn es ist anzunehmen, dass manche Effektstärken aufgrund kleiner Stichproben insignifikant blieben. Auch die Einflüsse von medikamentöser Behandlung, Krankheitsdauer und Ersterkrankungsalter konnten mangels Daten nicht analysiert werden. Ebenso war die Studienzahl für Analysen zum Publikationsbias nicht immer ausreichend.

Eine weitere Begrenzung der Arbeit bestand darin, dass die zunehmend diskutierte dimensionale Sichtweise der Symptomdimensionen nur ansatzweise berücksichtigt wurde. In den meisten Primärstudien wurden die Zwangspatienten gemäß ihrer jeweils vorherrschenden Symptomdimension in Gruppen aufgeteilt und diese Symptom“gruppen“ miteinander verglichen. Hierbei handelt es sich um eine Vereinfachung, denn jeder Zwangspatient und auch jeder gesunde Mensch weist auf allen Symptomdimensionen unterschiedlich starke Ausprägungen auf. Die

Perspektive distinkter Symptomgruppen wurde in 31 von 38 eingeschlossenen Studien eingenommen, während sieben aktuelle Korrelationsstudien bereits die dimensionale Sichtweise einnahmen. In einer Metaanalyse werden Daten, die andere Forschergruppen zuvor generiert hatten, abgebildet. Hierbei ist es zwar möglich, unter Inkaufnahme eines Datenverlustes dimensionale Daten (Ausprägung von Symptomen) in kategoriale Daten (Symptom klinisch relevant vorhanden oder nicht) zu überführen. Hingegen ist es nicht möglich, kategoriale bzw. dichotome Daten nachträglich als Kontinuum dimensional zu betrachten. Entsprechend gilt der Kritikpunkt, dass bislang zu wenig „dimensional geforscht“ wurde, für die Primärstudien und infolgedessen auch für diese Metaanalyse. Im Zusammenhang mit der dimensionalen Betrachtung der Zwangssymptome ist noch anzumerken, dass sich über den Verlauf des Forschungsvorhabens von 2010 bis 2014 die Diagnosekriterien für Sammelzwang geändert haben. Insofern bildet diese Arbeit auch Einflüsse durch ein weiterentwickeltes Diagnosesystem ab.

Stärken. Diese Arbeit war nach Kenntnis der Autorin die erste, die alle Symptomdimensionen der Zwangsstörung und alle großen neuropsychologischen Domänen in sieben Metaanalysen integriert hat. Damit hat diese Arbeit a) einen ausführlichen Überblick über den aktuellen Forschungsstand, die bestehenden Lücken und sich daraus ergebende neue Forschungsmöglichkeiten auf diesem Gebiet gegeben, b) Evidenz dafür erbracht, dass die uneinheitlichen Befunde der bisherigen Studien auch auf die Heterogenität der Symptomdimensionen zurückzuführen sind, c) belegt, dass sich (mit Ausnahme der Patienten mit Waschzwang) Zwangspatienten aller Symptomdimensionen in ihrer Neuropsychologie signifikant von Gesunden unterscheiden und d) bestätigt, dass sich Zwangspatienten mit Waschzwang, Kontrollzwang und Sammelzwang in ihren neuropsychologischen Profilen signifikant voneinander unterscheiden.

Methodische Stärken lagen darin, dass nicht nur ein narratives oder unsystematisches Review, sondern ein systematisches Review einschließlich Metaanalyse durchgeführt wurde, das zahlenmäßige Resultate über die neuropsychologischen Defizite der Zwangspatienten erbrachte. Das methodische

Vorgehen war sorgfältig und entspricht den QUOROM- und PRISMA-Standards. Darüber hinaus wurden graue Literatur und Veröffentlichungen in mehreren Sprachen eingeschlossen, so dass ein Publikationsbias minimiert wurde. Eine weitere Stärke bestand darin, dass ein Instrument zur Beurteilung der Studienqualität entwickelt wurde. Eine solche Qualitätsbeurteilung ist in Metaanalysen zu RCTs bereits Standard (Deutsches Cochrane Zentrum, 2013), nicht jedoch in Metaanalysen zu Fall-Kontroll-Studien. Die bisher publizierten Metaanalysen zur Neuropsychologie der Zwangsstörung hatten die Studienqualität nicht systematisch beurteilt. Insofern bringt diese Arbeit ein Plus an methodischer Qualität mit.

Die Heterogenität der eingeschlossenen Studien und der neuropsychologischen Konstrukte kann ebenfalls als Stärke betrachtet werden. Zwar sind derart viele Variablen aufwändig auszuwerten. Der Aufwand ist jedoch gerechtfertigt, da hierdurch die Fragestellung über viele Studienvariationen untersucht werden konnte. Dies führte zu breiteren Ergebnissen als eine Metaanalyse, die nur Studien zu ein und demselben Testverfahren untersuchen würde. Letztendlich resultieren aus dieser Heterogenität eine zuverlässige Generalisierbarkeit der Befunde und damit eine hohe externe Validität.

Ausblick. Die Ergebnisse der vorliegenden Arbeit könnten zukünftige Vorhaben beeinflussen. Erstens lassen sich methodische Forderungen an weitere Studien ableiten, zweitens könnten sich aus den Befunden therapeutische Implikationen für die Praxis ergeben und drittens könnten die hier festgestellten Forschungslücken durch neue Studien gefüllt und theoretische Ansätze weiterentwickelt werden.

Aus methodischer Sicht wäre es wünschenswert, zukünftige Fall-Kontroll-Studien nach einheitlicheren Qualitätskriterien durchzuführen. Hierzu könnte die im Rahmen dieser Arbeit entwickelte Qualitätsskala an einer größeren Stichprobe evaluiert werden. Auf Grundlage dieser Qualitätskriterien wäre es möglich, einheitlichere Standards zur Darstellung von Ergebnissen zu entwickeln. Konkret könnte dies z. B. bedeuten, dass Autoren vollständige Resultate veröffentlichen und auch insignifikante Befunde mit korrekten p -Werten angeben. Dies würde die

Wahrscheinlichkeit erhöhen, durch metaanalytische Kombination der Ergebnisse mehrerer Einzelstudien eine größere Teststärke zu erreichen. Um darüber hinaus die Konfundierung von Zwängen mit komorbider Depressivität und Angst zu vermeiden, wären Studiendesigns denkbar, in denen Zwangspatienten bestimmter Symptomdimensionen mit Depressions- oder Angstpatienten verglichen werden. Auf diese Weise haben Ecker und Engelkamp (1995) in ihrer Untersuchung zum Handlungsgedächtnis Zwangspatienten mit anderen psychosomatischen Patienten verglichen. Zusätzlich könnte im Bereich der Neuropsychologie ein breiterer Konsens hinsichtlich der Domänen und Subdomänen sowie ihrer Zuordnung zu den einzelnen Testverfahren angestrebt werden. Eine derartige Vereinheitlichung würde in der Interpretation von Testergebnissen mehr Klarheit bringen.

Auch therapeutische Implikationen könnten aus den Befunden zur Neuropsychologie der Symptomdimensionen der Zwangsstörung abgeleitet werden. Zum einen verbessern sich bei Zwangspatienten kognitive Funktionen im Verlauf einer erfolgreichen Psychotherapie (Foa & McNally, 1986; Kuelz et al., 2006), zum anderen können aus neuropsychologischen Trainings auch bessere psychotherapeutische Verläufe resultieren (Najmi & Amir, 2010). Zudem sprechen Zwangspatienten je nach Symptomdimension besser oder schlechter auf kognitive Verhaltenstherapie an (Mataix-Cols, Marks et al., 2002). Insofern erscheint es nach den Befunden der vorliegenden Metaanalyse naheliegend, für die einzelnen Symptomdimensionen maßgeschneiderte Therapieansätze (Sookman et al., 2005) mit neuropsychologischen Elementen zu ergänzen. Insbesondere bei Patienten mit Sammelzwang können neurokognitive Defizite therapie relevant sein. Da Patienten mit Sammelzwang auf konventionelle Exposition mit Reaktionsmanagement nicht gut ansprechen, entwickelten Steketee und Frost (2006) ein Psychotherapiekonzept, das Elemente zur Verbesserung der Aufmerksamkeitsfokussierung, der Organisationsfähigkeit, des Entscheidungsverhaltens und des Planens/Problemlösens enthält. Hier sind erste Evaluationsergebnisse zu erwarten. Analog zu diesem Therapiekonzept für Personen mit Sammelzwang (Steketee & Frost, 2006) und aufbauend auf der kognitiven Therapie der Exekutivfunktionen (Müller, Hildebrandt & Münte, 2004) könnten, abgestimmt auf die jeweils vorherrschende

Symptomdimension, neuropsychologische Trainingsaufgaben in die Psychotherapie für Zwangspatienten integriert werden. Vor allem für Patienten mit Kontrollzwang wären hier Therapiemodule vorstellbar, die Fertigkeiten im Planen/Problemlösen, in kognitiver Inhibition und im Einsatz von Organisationsstrategien beim Enkodieren verbessern können.

Diese Metaanalyse hat gezeigt, bezüglich welcher Symptomdimensionen der Zwangsstörung und auch bezüglich welcher neuropsychologischer Konstrukte noch keine Studien publiziert wurden. Hier könnten zahlreiche weitere Untersuchungen durchgeführt werden. Da bestimmte neuropsychologische Defizite mögliche Endophänotypen, d. h., Vulnerabilitätsmarker mit einer erblichen Komponente, darstellen, könnte die Endophänotypenforschung zur Früherkennung des Risikos von Zwangsstörungen weiter vorangetrieben werden. Eine frühzeitige Diagnose und Therapie ist von großer Bedeutung, denn damit ließen sich eine Verschlimmerung und Chronifizierung vermeiden – insbesondere wenn man berücksichtigt, dass Zwangspatienten im Mittel erst nach über sechs Jahren eine angemessene Behandlung erhalten. Ein entsprechendes Endophänotypen-Design, nach dem Studien durchgeführt werden könnten, haben Chamberlain und Menzies (2009) bereits vorgeschlagen. Demnach könnte man Zwangspatienten und deren Angehörige ersten Grades, die genetisch bedingt ein erhöhtes Risiko aufweisen, ebenfalls an einer Zwangsstörung zu erkranken, auf Ebene der Organsysteme mittels MRT-Scans auf Besonderheiten in der grauen und weißen Substanz untersuchen (was einem neurobiologischen Endophänotypen entspräche) und ihre Leistungen in kognitiven Aufgaben überprüfen (was einem neuropsychologischen Endophänotypen entspräche). Strukturelle Auffälligkeiten und neuropsychologische Leistungsdefizite würden dann als Vulnerabilitätsmarker betrachtet, die auf ein erhöhtes genetisches Risiko hinweisen, aber nicht zwangsläufig die Manifestation einer Zwangsstörung bedeuten würden. Dies wäre insbesondere für die Dimension *Symmetrie-/Ordnungszwang* bedeutsam, da für deren genetische Komponente bereits empirische Evidenz vorliegt (vgl. Abschnitt 2.3).

In der Verschränkung von Neuropsychologie, Neurobiologie, Genforschung und Psychotherapieforschung könnten translationale Ansätze, in denen sich Grundlagenforschung und klinische Praxis miteinander verbinden, weiterentwickelt werden. Hierbei könnten mehr Kombinationsstudien durchgeführt werden, die bildgebende Verfahren, neuropsychologische Testverfahren und Symptomprovokation gemeinsam einsetzen, um in vivo ein genaueres Bild der Gehirnfunktionen und entsprechend der neuronalen Korrelate zu erhalten. Damit einhergehend, könnten die Symptomdimensionen in theoretischen Modellen der Zwangsstörung verstärkt berücksichtigt werden. Der Ansatz von Chamberlain et al. (2005) führt auf der Ebene des beobachtbaren Verhaltens Zwangsgedanken und Zwangshandlungen an. Berücksichtigt man z. B. die videogestützten Befunde von Zor, Fineberg, Eilam und Hermesh (2011), denen zufolge es deutliche qualitative Unterschiede in den Wasch- und Kontrollritualen gibt, dann ließe sich Chamberlains Endophänotypenansatz auf der Ebene der Symptome noch differenzierter betrachten. Denkbar wäre zum Beispiel, die in der vorliegenden Metaanalyse gefundenen Ergebnisse mit den zitierten genetischen und Bildgebungsbefunden (Abschnitte 2.3 und 2.4) zu kombinieren. Für Kontrollzwang wäre damit folgendes Mehr-Ebenen-Modell vorstellbar:

- Phänotyp: Kontrollzwang
- Neuropsychologischer Endophänotyp: Defizite im Planen/Problemlösen und in der Inhibition
- Biologischer Endophänotyp: erhöhte Aktivierung im Putamen und im Thalamus bei Symptomprovokation
- Genotyp: Befunde stehen noch aus.

Für Symmetrie-/Ordnungszwang hingegen könnte ein Mehr-Ebenen-Modell folgendermaßen aussehen:

- Phänotyp: Symmetrie-/Ordnungszwang
- Neuropsychologischer Endophänotyp: Defizite im verbalen und nonverbalen Gedächtnis

- Biologischer Endophänotyp: verringertes Volumen der Grauen Substanz des Parietalcortex beidseits sowie erhöhtes Volumen der Grauen und Weißen Substanz beider medialer Temporallappen
- Genotyp: Korrelation des Faktors Symmetrie/Ordnen mit einem Mechanismus des Serotonin-Transportergens; Auftreten eines bestimmten Allels ausschließlich bei Symmetrie-/Ordnungszwang; erhöhtes Risiko für Verwandte, selbst zu erkranken.

Ob es sich dabei jeweils um eine Symptomdimension oder doch um eine Kategorie handelt, ist noch nicht endgültig geklärt. Während *Kontrollzwang* auf den hier beschriebenen Ebenen im Sinne einer Dimension betrachtet werden kann, sprechen die genetischen Befunde von *Symmetrie-/Ordnungszwang* eher für eine Kategorie. Um für jede Symptomdimension ein entsprechendes Mehr-Ebenen-Modell der Endophänotypen aufstellen zu können, liegen noch nicht ausreichend Daten vor. Entsprechend bieten sich noch zahlreiche Fragestellungen an, die erforscht werden könnten.

Greift man darüber hinaus in Zukunft noch die dimensionale Sichtweise auf und nimmt verschiedene Schweregrade der einzelnen Symptomdimensionen an, dann kann man sich der Überlegung von Chamberlain et al. (2005, S. 409) anschließen, warum sich Zwangspatienten manchmal, aber nicht immer in neuropsychologischen Funktionen von anderen Probanden unterscheiden: Wenn diese Defizite mit dem Schweregrad und Fortschreiten der Störung zusammenhängen, dann könnten neuropsychologische Aufgaben auch als Marker für verschiedene Schweregrade der Zwangsstörung verwendet werden. Basierend auf der Annahme eines Kontinuums neuropsychologischer Fähigkeiten (Heinrichs & Zakzanis, 1998) könnten weitere Studien untersuchen, ob mit unterschiedlichen Schweregraden der Symptomdimensionen auch unterschiedlich große neuropsychologischer Defizite einhergehen. Diese Fragestellung wurde in den eingeschlossenen Primärstudien noch nicht untersucht, so dass sich hier ein neuer Forschungsbereich eröffnen könnte. Um herauszufinden, ob die neuropsychologischen Leistungsdefizite von Zwangspatienten verschiedener Symptomdimensionen vorübergehender oder dauerhafter Natur sind,

wären darüber hinaus Längsschnittstudien über mehrere Messzeitpunkte hinweg erforderlich.

Die Ergebnisse dieser Arbeit liefern empirische Evidenz für die Validität des Symptomdimensions-Ansatzes. Als generelles Fazit kann festgestellt werden, dass sich Zwangspatienten mit Schwerpunkten auf verschiedenen Symptomdimensionen signifikant in ihren neuropsychologischen Testprofilen unterscheiden. Besonders bedeutsam erscheint, dass über mehrere Studien hinweg Patienten mit Waschzwang kognitiv unbeeinträchtigt erschienen, während Patienten mit Kontrollzwang große Defizite im Planen/Problemlösen und in der Inhibition aufwiesen. Insgesamt legen die Resultate der hier durchgeführten Metaanalysen nahe, Patienten mit vorherrschendem Waschzwang, Kontrollzwang, verbotenen Gedanken, Symmetrie-/Ordnungszwang und Sammelzwang in ihrer symptombezogenen Verschiedenheit zu würdigen. Ein aktueller Blick auf die internationalen Forschungsaktivitäten zeigt, dass dies bereits verstärkt geschieht. Anregend wirkt hier auch die Aufnahme der Symptomdimensionen ins DSM-5. Je mehr die Zusammenhänge der Symptomdimensionen mit neuropsychologischen Leistungen sowie mit den ihnen zugrunde liegenden neuronalen Regelkreisen und den genetischen Grundlagen sichtbar werden, desto differenzierter werden sich die derzeitigen Modelle der Zwangsstörung erweitern lassen. Hieraus ergibt sich eine Fülle weiterer Möglichkeiten für die Grundlagenforschung, die Früherkennung und die klinische Therapieforschung der Zwangsstörung.

Literaturverzeichnis

- Abbruzzese, M., Ferri, S. & Scarone, S. (1995). Wisconsin Card Sorting Test performance in obsessive-compulsive disorder: No evidence for involvement of dorsolateral prefrontal cortex. *Psychiatry Research*, 58(1), 37-43.
- Abramovitch, A., Abramowitz, J. S. & Mittelman, A. (2013). The neuropsychology of adult obsessive-compulsive disorder: A meta-analysis. *Clinical Psychology Review*, 33(8), 1163-1171.
- Abramowitz, J. S., Franklin, M. E., Schwartz, S. A. & Furr, J. M. (2003). Symptom Presentation and Outcome of Cognitive-Behavioral Therapy for Obsessive-Compulsive Disorder. *Journal of Consulting and Clinical Psychology*, 71(6), 1049-1057.
- Alonso, P., Menchon, J. M., Pifarre, J., Mataix-Cols, D., Torres, L., Salgado, P. et al. (2001). Long-term follow-up and predictors of clinical outcome in obsessive-compulsive patients treated with serotonin reuptake inhibitors and behavioral therapy. *Journal of Clinical Psychiatry*, 62(7), 535-540.
- Alsobrook, I. J., Leckman, J. F., Goodman, W. K., Rasmussen, S. A. & Pauls, D. L. (1999). Segregation analysis of obsessive-compulsive disorder using symptom-based factor scores. *American Journal of Medical Genetics (Neuropsychiatric Genetics)*, 88(6), 669-675.
- Alvarenga, P. G., do Rosário, M. C., Batistuzzo, M. C., Diniz, J. B., Shavitt, R. G., Duran, F. L. S. et al. (2012). Obsessive-compulsive symptom dimensions correlate to specific gray matter volumes in treatment-naïve patients. *Journal of Psychiatric Research*, 46(12), 1635-1642.
- American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorder, third edition (DSM-III)*. Washington DC: American Psychiatric Association.
- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV)*. Washington, DC: American Psychiatric Association.

- American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5)*. Arlington, VA: American Psychiatric Association.
- Aronowitz, B. R., Hollander, E., DeCaria, C., & Cohen, L. (1994). Neuropsychology of obsessive compulsive disorder: Preliminary findings. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology*, 7(2), 81-86.
- Ayers, C. R., Saxena, S., Golshan, S. & Wetherell, J. L. (2010). Age at onset and clinical features of late life compulsive hoarding. *International Journal of Geriatric Psychiatry*, 25(2), 142-149.
- *Ayers, C. R., Saxena, S., Golshan, S. & Wetherell, J. L. (2013). Executive functioning in older adults with *hoarding disorder*. *International Journal of Geriatric Psychiatry* 28, 1175-1181.
- Backenstrass, M., Schaller, P., & Jäntsch, B. (2012). Obsessive-Compulsive Inventory-Revised (OCI-R) und Hamburger Zwangsinventar-Kurzform (HZI-K) im Vergleich: Eine Validitätsstudie. *Verhaltenstherapie*, 22(2), 106-113.
- Baddeley, A. D. (1989). *Working memory* (Nachdruck). Oxford: Clarendon Press.
- Baer, L. (1994). Factor analysis of symptom subtypes of obsessive compulsive disorder and their relation to personality and tic disorders. *Journal of Clinical Psychiatry*, 55(3, Suppl), 18-23.
- Ball, S. G., Baer, L. & Otto, M. W. (1996). Symptom subtypes of obsessive-compulsive disorder in behavioral treatment studies: A quantitative review. *Behaviour Research and Therapy*, 34(1), 47-51.
- Bannon, S., Gonsalvez, C. J., Croft, R. J. & Boyce, P. M. (2002). Response inhibition deficits in obsessive-compulsive disorder. *Psychiatry Research*, 110(2), 165-174.
- Baxter, L. R., Jr., Saxena, S., Brody, A. L., Ackermann, R. F., Colgan, M., Schwartz, J. M. et al. (1996). Brain Mediation of Obsessive-Compulsive Disorder Symptoms: Evidence From Functional Brain Imaging Studies in the Human and Nonhuman Primate. *Seminars in Clinical Neuropsychiatry*, 1(1), 32-47.

- Baxter, L. R., Schwartz, J. M., Bergman, K. S., Szuba, M. P., Guze, B. H., Mazziotta, J. C. et al. (1992). Caudate glucose metabolic rate changes with both drug and behavior therapy for obsessive-compulsive disorder. *Archives of General Psychiatry*, 49(9), 681-689.
- Bechara, A., Damasio, A. R., Damasio, H. & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50(1-3), 7-15.
- Beck, A. T., Epstein, N., Brown, G. & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, 56(6), 893-897.
- Beck, A. T., Steer, R. A. & Brown, G. K. (1996). *Beck Depressions-Inventar (BDI). Testhandbuch, zweite Auflage* (M. Hautzinger, M. Bailer, H. Worall & F. Keller, Übers.). Bern: Hans Huber (Original erschienen 1993: Manual for the Beck Depression Inventory).
- Beck, A. T., Steer, R. A. & Brown, G. K. (2006). *BDI-II Beck Depression Inventar Revision*. (M. Hautzinger, F. Keller & C. Kühner, Übers.). Frankfurt am Main: Harcourt Test Services (Original erschienen 1996: Manual for the Beck Depression Inventory-II).
- Bloch, M. H., Landeros-Weisenberger, A., Rosario, M. C., Pittenger, C. & Leckman, J. F. (2008). Meta-analysis of the symptom structure of obsessive-compulsive disorder. *The American Journal of Psychiatry*, 165(12), 1532-1542.
- *Blom, R. M., Samuels, J. F., Grados, M. A., Chen, Y., Bienvenu, O. J., Riddle, M. A. et al. (2011). Cognitive functioning in compulsive hoarding. *Journal of Anxiety Disorders*, 25(8), 1139-1144.
- Bora, E., Yücel, M. & Pantelis, C. (2010). Cognitive impairment in affective psychoses: A meta-analysis. *Schizophrenia Bulletin*, 36(1), 112-125.
- Borenstein, M., Hedges, L., Higgins, J. & Rothstein, H. (2009). *Introduction to Meta-Analysis*. Chichester: Wiley.
- Borenstein, M., Hedges, L., Higgins, J. & Rothstein, H. (2005). *Comprehensive Meta-analysis Software (Version 2)*. Englewood NJ: Biostat.

- *Bouvard, M., Dirson, S. & Cottraux, J. (1997). Etude de la mémoire de sujets obsessionnels compulsifs laveurs et vérificateurs et de sujets contrôles. *Revue Européenne de Psychologie Appliquée*, 47(3), 189-196.
- Bower, G. H. (1970). Organizational factors in memory. *Cognitive Psychology*, 1(1), 18-46.
- Brakoulias, V., Starcevic, V., Berle, D., Milicevic, D., Moses, K., Hannan, A. et al. (2013). The characteristics of unacceptable/taboo thoughts in obsessive-compulsive disorder. *Comprehensive Psychiatry*.
- Bredenkamp, J. (1998). *Lernen, erinnern, vergessen*. München: Beck.
- Buchner, A. (2006). Funktionen und Modelle des Gedächtnisses. In H.-O. Karnath & P. Thier (Hrsg.), *Neuropsychologie* (S. 437-447). Heidelberg: Springer.
- Buchner, A. & Brandt, M. (2008). Gedächtniskonzeptionen und Wissensrepräsentationen. In J. Müsseler (Hrsg.), *Allgemeine Psychologie*. Heidelberg: Spektrum.
- Burns, G. L., Keortge, S. G., Formea, G. M. & Sternberger, L. G. (1996). Revision of the Padua Inventory of obsessive compulsive disorder symptoms: Distinctions between worry, obsessions, and compulsions. *Behaviour Research and Therapy*, 34(2), 163-173.
- Burt, D. B., Zembar, M. J. & Niederehe, G. (1995). Depression and memory impairment: A meta-analysis of the association, its pattern, and specificity. *Psychological Bulletin*, 117(2), 285-305.
- Calamari, J. E., Wiegartz, P. S. & Janeck, A. S. (1999). Obsessive-compulsive disorder subgroups: A symptom-based clustering approach. *Behaviour Research and Therapy*, 37(2), 113-125.
- Cambridge-Cognition (2013). *Cambridge Neuropsychological Test Automated Battery (CANTAB)*. Zugriff am 11. 9. 2013, von <http://cantabtest.com/academic/cantabsuite/tests>
- Casasanto, D. J., Killgore, W. D. S., Maldjian, J. A., Glosser, G., Alsop, D. C., Cooke, A. M. et al. (2002). Neural Correlates of Successful and Unsuccessful Verbal Memory Encoding. *Brain and Language*, 80(3), 287-295.

- Cavedini, P., Gorini, A. & Bellodi, L. (2006). Understanding Obsessive-Compulsive Disorder: Focus on Decision Making. *Neuropsychology Review*, 16(1), 3-15.
- *Ceschi, G., Van der Linden, M., Dunker, D., Perroud, A. & Brédart, S. (2003). Further exploration memory bias in compulsive washers. *Behaviour Research and Therapy*, 41(6), 737-748.
- *Cha, K. R., Koo, M.-S., Kim, C.-H., Kim, J. W., Oh, W.-J., Suh, H. S. et al. (2008). Nonverbal memory dysfunction in obsessive-compulsive disorder patients with checking compulsions. *Depression & Anxiety*, 25(11), E115-120.
- Chamberlain, S. R., Blackwell, A. D., Fineberg, N. A., Robbins, T. W. & Sahakian, B. J. (2005). The neuropsychology of obsessive compulsive disorder: The importance of failures in cognitive and behavioural inhibition as candidate endophenotypic markers. *Neuroscience and Biobehavioral Reviews*, 29(3), 399-419.
- Chamberlain, S. R. & Menzies, L. (2009). Endophenotypes of obsessive compulsive disorder: Rationale, evidence and future potential. *Expert Review of Neurotherapeutics*, 9(8), 1133-1146.
- *Clair, A. H., N'Diaye, K., Baroukh, T., Pochon, J. B., Morgieve, M., Hantouche, E. et al. (2013). Excessive checking for non-anxiogenic stimuli in obsessive-compulsive disorder. *European Psychiatry*, 28(8), 507-513.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Coles, M. E., Pinto, A., Mancebo, M. C., Rasmussen, S. A. & Eisen, J. L. (2008). OCD with comorbid OCPD: A subtype of OCD? *Journal of Psychiatric Research*, 42(4), 289-296.
- Cooper, H. M. (Hrsg.). (1994). *The handbook of research synthesis*. New York: Russell Sage Foundation.
- Cox, C. S. (1997). Neuropsychological abnormalities in obsessive-compulsive disorder and their assessments. *International Review of Psychiatry*, 9(1), 45-60.
- Cuttler, C. & Graf, P. (2007). Sub-clinical compulsive checkers' prospective memory is impaired. *Journal of Anxiety Disorders*, 21(3), 338-352.

- Cuttler, C. & Graf, P. (2009). Checking-in on the memory deficit and meta-memory deficit theories of compulsive checking. *Clinical Psychology Review, 29*(5), 393-409.
- Deckersbach, T., Otto, M. W., Savage, C. R., Baer, L. & Jenike, M. A. (2000). The relationship between semantic organization and memory in obsessive-compulsive disorder. *Psychotherapy and Psychosomatics, 69*(2), 101-107.
- Diefenbach, G. J., DiMauro, J., Frost, R., Steketee, G. & Tolin, D. F. (2013). Characteristics of hoarding in older adults. *American Journal of Geriatric Psychiatry, 21*(10), 1043-1047.
- Dilling, H. (Hrsg.). (2008). *Internationale Klassifikation psychischer Störungen: ICD-10, Kapitel V (F); klinisch-diagnostische Leitlinien* (6., vollst. überarb. Aufl. unter Berücksichtigung der Änderungen entsprechend ICD-10-GM 2004/2008). Bern: Huber.
- *Dittrich, W. H., Johansen, T., Fineberg, N. & Landro, N. I. (2011). Cognitive performance and specific deficits in OCD symptom dimensions: II. Spatial memory and impaired recognition of visuospatial object locations. *German Journal of Psychiatry, 14*, 1-12.
- *Dittrich, W. H. & Johansen, T. (2013). Cognitive deficits of executive functions and decision-making in obsessive-compulsive disorder. *Scandinavian Journal of Psychology, 54*(5), 393-400.
- Dold, M. & Aigner, M. (2009). Neuroimaging the various symptom dimensions of obsessive-compulsive disorder. *Neuropsychiatrie, 23*(4), 193-205.
- Drummond, L. M. (1993). The treatment of severe, chronic, resistant obsessive-compulsive disorder: An evaluation of an in-patient programme using behavioural psychotherapy in combination with other treatments. *British Journal of Psychiatry, 163*, 223-229.
- Duval, S. & Tweedie, R. (2000). Trim and Fill: A Simple Funnel Plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics, 56*(2), 455-463.
- Easterbrook, P. J., Gopalan, R., Berlin, J. A. & Matthews, D. R. (1991). Publication bias in clinical research. *Lancet, 337*(8746), 867-872.

- Ecker, W. (2009). *Synopsis der Schriften für die kumulative Habilitationsleistung zum Thema Symptom- und Motivdimensionen der Zwangsstörung: Phänomenologie, diagnostische Erfassung, theoretische Konzeptualisierung und Behandlungsimplicationen*. Kumulative Habilitation, Ruprecht-Karls-Universität, Heidelberg.
- Ecker, W. & Engelkamp, J. (1995). Memory for actions in obsessive-compulsive disorder. *Behavioural and Cognitive Psychotherapy*, 23(4), 349-371.
- Ecker, W. & Gönner, S. (2008). Incompleteness and harm avoidance in OCD symptom dimensions. *Behaviour Research and Therapy*, 46(8), 895-904.
- Egger, M., Smith, G. D. & Altman, D. G. (Hrsg.). (2001). *Systematic Reviews in Health Care: Meta-Analysis in Context* (2. Aufl.). London: BMJ.
- *Exner, C., Martin, V. & Rief, W. (2009). Self-focused ruminations and memory deficits in obsessive-compulsive disorder. *Cognitive Therapy and Research*, 33(2), 163-174.
- Fioravanti, M., Carlone, O., Vitale, B., Cinti, M. E. & Clare, L. (2005). A meta-analysis of cognitive deficits in adults with a diagnosis of schizophrenia. *Neuropsychology Review*, 15(2), 73-95.
- Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G. et al. (2002). The Obsessive-Compulsive Inventory: Development and validation of a short version. *Psychological Assessment*, 14(4), 485-495.
- Foa, E. B. & Kozak, M. J. (1995). DSM-IV field trial: Obsessive-compulsive disorder. *American Journal of Psychiatry*, 152(1), 90-96.
- Foa, E. B. & McNally, R. J. (1986). Sensitivity to feared stimuli in obsessive-compulsives: A dichotic listening analysis. *Cognitive Therapy and Research*, 10(4), 477-485.
- Fontenelle, L., Marques, C., Engelhardt, E. & Versiani, M. (2001). Impaired set-shifting ability and therapeutic response in obsessive-compulsive disorder. *Journal of Neuropsychiatry and Clinical Neurosciences*, 13(4), 508-510.
- Frost, R. F. (2013). *Diagnosing Hoarding*. Zugriff am 30.10.2014, von <http://hoarding.iocdf.org/hoarding/diagnosing.aspx>

- Frost, R. O., & Gross, R. C. (1993). The hoarding of possessions. *Behaviour Research and Therapy*, 31(4), 367-381.
- Frost, R. O., & Hartl, T. L. (1996). A cognitive-behavioral model of compulsive hoarding. *Behaviour Research and Therapy*, 34(4), 341-350.
- Frost, R. O. & Steketee, G. (2010). *About Hoarding*. Zugriff am 03. 5. 2013, von <http://www.ocfoundation.org/hoarding/about.aspx>
- Frost, R. O., Steketee, G. & Grisham, J. R. (2004). Measurement of compulsive hoarding: saving inventory-revised. *Behaviour Research & Therapy*, 42, 1163-1182.
- Gilliam, C. M., Norberg, M. M., Villavicencio, A., Morrison, S., Hannan, S. E. & Tolin, D. F. (2011). Group cognitive-behavioral therapy for *hoarding disorder*: an open trial. *Behaviour Research & Therapy*, 49(11), 802-807.
- Gillig, P. M. & Sanders, R. D. (2011). Higher cortical functions: Attention and vigilance. *Innovations in Clinical Neuroscience*, 8(1), 43-46.
- Girard, B., Tabareau, N., Pham, Q. C., Berthoz, A. & Slotine, J.-J. (2008). Cortico-baso-thalamo-cortical loops. *Neural Networks* 21(4), S. 629.
- Gönner, S., Hahn, S., Leonhart, R., Ecker, W. & Limbacher, K. (2009). Identifikation der Hauptsymptome von Zwangspatienten anhand von Symptomskalen: Kriteriumsvalidität und diagnostische Genauigkeit des OCI-R. *Verhaltenstherapie*, 19(4), 251-258.
- Gönner, S., Schmid, J., Gönner, S., Leonhart, R. & Ecker, W. (2013). Das Vancouver Obsessional Compulsive Inventory—Revised (VOCI-R). Ist sein Einsatz in der Diagnostik und Therapieplanung bei Zwangspatienten sinnvoll? *Zeitschrift für Klinische Psychologie und Psychotherapie: Forschung und Praxis*, 42(1), 14-23.
- Goodman, W. K., Price, L. H., Rasmussen, S. A. & Mazure, C. (1989). The Yale-Brown Obsessive Compulsive Scale: I. Development, use, and reliability. *Archives of General Psychiatry*, 46(11), 1006-1011.
- Gottesman, I. I. & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry*, 160(4), 636-645.

- Grant, D. A. & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404-411.
- Green, B. F. & Hall, J. A. (1984). Quantitative methods for literature reviews. *Annual Review of Psychology*, 35, 37-53.
- Greisberg, S. & McKay, D. (2003). Neuropsychology of obsessive-compulsive disorder: A review and treatment implications. *Clinical Psychology Review*, 23(1), 95-117.
- *Grisham, J. R., Brown, T., Savage, C. R., Steketee, G. & Barlow, D. H. (2007). Neuropsychological impairment associated with compulsive hoarding. *Behaviour Research & Therapy*, 45(7), 1471-1483.
- *Grisham, J. R., Norberg, M. M., Williams, A. D., Certoma, S. P. & Kadib, R. (2010). Categorization and cognitive deficits in compulsive hoarding. *Behaviour Research and Therapy*, 48(9), 866-872.
- Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology*, 32, 50-55.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry*, 23, 56-61.
- Harlow, J. M. (1993). Recovery from the passage of an iron bar through the head. *History of Psychiatry*, 4(14), 271-281.
- Harrison, B. J., Pujol, J., Cardoner, N., Deus, J., Alonso, P., Lopez-Sola, M. et al. (2013). Brain corticostriatal systems and the major clinical symptom dimensions of obsessive-compulsive disorder. *Biological Psychiatry*, 73(4), 321-328.
- Harrison, B. J., Soriano-Mas, C., Pujol, J., Ortiz, H., Lopez-Sola, M., Hernandez-Ribas, R. et al. (2009). Altered corticostriatal functional connectivity in obsessive-compulsive disorder. *Archives of General Psychiatry*, 66(11), 1189-1200.
- Härting, C. & Markowitsch, H. J. (1997). Neuropsychologische Befunde der Zwangsstörung. *Fortschritte der Neurologie, Psychiatrie*, 65(11), 509-515.

- *Hartl, T. L., Frost, R. O., Allen, G. J., Deckersbach, T., Steketee, G., Duffany, S. R. et al. (2004). Actual and perceived memory deficits in individuals with compulsive hoarding. *Depression and Anxiety*, 20(2), 59-69.
- Harvey, N. S. (1986). Impaired cognitive set-shifting in obsessive-compulsive neurosis. *IRCS Medical Science: Psychology & Psychiatry*, 14(9-10), 936-937.
- *Hashimoto, N., Nakaaki, S., Omori, I. M., Fujioi, J., Noguchi, Y., Murata, Y. et al. (2011). Distinct neuropsychological profiles of three major symptom dimensions in obsessive compulsive disorder. *Psychiatry Research*, 187(1-2), 166.
- Hasler, G., Kazuba, D. & Murphy, D. L. (2006). Factor analysis of obsessive-compulsive disorder Y-BOCS-SC symptoms and association with 5-HTTLPR SERT polymorphism. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 141B(4), 403-408.
- Hasler, G., LaSalle-Ricci, V. H., Ronquillo, J. G., Crawley, S. A., Cochran, L. W., Kazuba, D. et al. (2005). Obsessive-compulsive disorder symptom dimensions show specific relationships to psychiatric comorbidity. *Psychiatry Research*, 135(2), 121-132.
- Hasler, G., Pinto, A., Greenberg, B. D., Samuels, J., Fyer, A. J., Pauls, D. et al. (2007). Familiality of factor analysis-derived Y-BOCS dimensions in OCD-affected sibling pairs from the OCD Collaborative Genetics Study. *Biological Psychiatry*, 61(5), 617-625.
- *Hayward, L. C. (2012). *Emotion and categorization: Do hoarders have difficulties with categorization, emotion, and categorizing when emotional?* Unveröffentlichte Dissertation. Bingham University, New York. ProQuest Information & Learning, US.
- Heinrichs, R. W. & Zakzanis, K. K. (1998). Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. *Neuropsychology*, 12(3), 426-445.
- Henry, J. D. (2006). A meta-analytic review of Wisconsin Card Sorting Test and Verbal Fluency Performance in obsessive-compulsive disorder. *Cognitive Neuropsychiatry*, 11(2), 156-176.

- Higgins, J. P. T. & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* [updated March 2011]. Zugriff am 18.9.2014, von <http://handbook.cochrane.org/>
- Higgins, J. P. T. & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), 1539-1558.
- Hilger, E. & Kasper, S. (2002). Kognitive Symptomatik bei schizophrener Erkrankung: Diagnostik und Pharmakotherapie. *Journal für Neurologie, Neurochirurgie und Psychiatrie*, 3(4), 17-22.
- Hodgson, R. J. & Rachman, S. (1977). Obsessional-compulsive complaints. *Behaviour Research and Therapy*, 15(5), 389-395.
- Hofmann, S. G., Ellard, K. K. & Siegle, G. J. (2012). Neurobiological correlates of cognitions in fear and anxiety: A cognitive–neurobiological information-processing model. *Cognition and Emotion*, 26(2), 282-299.
- Hollander, E. (2006). Personalized and Translational Medicine in Psychiatry and a Fond Farewell to Wayne Fenton, MD. *CNS Spectrums*, 11(10), 737-738.
- Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F. & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I² index? *Psychological Methods*, 11(2), 193-206.
- Hymas, N., Lees, A., Bolton, D., Epps, K. & Head, D. (1991). The neurology of obsessional slowness. *Brain*, 114(Pt 5), 2203-2233.
- Intemodino Group s. r. o. (2013). *Randomnumbergenerator*. Zugriff am 29. 05. 2013, von <http://randomnumbergenerator.intemodino.com/de/>
- Ioannidis, J. P. A. (2009). Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *Canadian Medical Association Journal*, 181(8), 488-493.
- Janet, P. (1919). *Les obsessions et la psychasthénie* (3. Aufl.). Paris: Félix Alcan.
- *Jang, J. H., Kim, S. H., Ha, T. H., Shin, N. Y., Do-Hyung, A., Jung-Seok, C. et al. (2010). Nonverbal memory and organizational dysfunctions are related with distinct symptom dimensions in obsessive-compulsive disorder. *Psychiatry research*, 180(2- 3), 93-98.

- Jefferys, D. & Moore, K. A. (2008). Pathological hoarding. *Australian family physician*, 37(4), 237-241.
- Jones, E. J. (2003). The Thalamus. *IBRO History of Neuroscience* (S. 1-6). Davis, California: University of California Davis.
- Kammer, T. & Karnath, H.-O. (2006). Manifestationen von Frontalhirnschädigungen. In H.-O. Karnath & P. Thier (Hrsg.), *Neuropsychologie* (S. 489-500). Heidelberg: Springer Medizin Verlag.
- Kang, D.-H., Kim, J.-J., Choi, J.-S., Kim, Y., II, Kim, C.-W., Youn, T. et al. (2004). Volumetric Investigation of the Frontal-Subcortical Circuitry in Patients With Obsessive-Compulsive Disorder. *Journal of Neuropsychiatry and Clinical Neurosciences*, 16(3), 342-349.
- Kang, D.-H., Jang, J. H., Han, J. Y., Kim, J.-H., Jung, W. H., Choi, J.-S. et al. (2013). Neural correlates of altered response inhibition and dysfunctional connectivity at rest in obsessive-compulsive disorder. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 40, 340-346.
- Kashyap, H., Kumar, J. K., Kandavel, T. & Reddy, Y. C. J. (2013). Neuropsychological functioning in obsessive-compulsive disorder: Are executive functions the key deficit? *Comprehensive Psychiatry*, 54(5), 533-540.
- Katerberg, H., Delucchi, K. L., Stewart, S. E., Lochner, C., Denys, D. A. J. P., Stack, D. E. et al. (2010). Symptom dimensions in OCD: Item-level factor analysis and heritability estimates. *Behavior Genetics*, 40(4), 505-517.
- Kathmann, N. (2008). *Neuropsychologie der Zwangsstörungen* (Band 7). Göttingen: Hogrefe.
- Kelley, G. A. & Kelley, K. S. (2012). Statistical models for meta-analysis: A brief tutorial. *World Journal of Methodology*, 2(4), 27-32.
- Khanna, S., Kaliaperumal, V. G. & Channabasavanna, S. M. (1990). Clusters of obsessive-compulsive phenomena in obsessive-compulsive disorder. *British Journal of Psychiatry*, 156, 51-54.

- Kichuk, S. A., Torres, A. R., Fontenelle, L. F., Rosário, M. C., Shavitt, R. G., Miguel, E. C. et al. (2013). Symptom dimensions are associated with age of onset and clinical course of obsessive-compulsive disorder. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 44, 233-239.
- Kim, H. J., Steketee, G. & Frost, R. O. (2001). Hoarding by elderly people. *Health & Social Work*, 26(3), 176-184.
- Kim, M.-S., Park, S.-J., Shin, M. S. & Kwon, J. S. (2002). Neuropsychological profile in patients with obsessive-compulsive disorder over a period of 4-month treatment. *Journal of Psychiatric Research*, 36(4), 257-265.
- Kimura, D. (1963). Right temporal lobe damage. *Archives of Neurology*, 8(264-271).
- Kordon, A., Lotz-Rambaldi, W., Mücke-Borowski, C., & Hohagen, F. (2013). *S3-Leitlinie Zwangsstörungen (AWMF-Registernummer 038/017) im Auftrag der Deutschen Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde (DGPPN)*. Zugriff am 21.10.2014, von http://www.dgppn.de/fileadmin/user_upload/_medien/download/pdf/kurzversion-leitlinien/S3-Leitlinie_Zwangsstörungen_lang.pdf
- Krabbendam, L., Arts, B., van Os, J. & Aleman, A. (2005). Cognitive functioning in patients with schizophrenia and bipolar disorder: A quantitative review. *Schizophrenia Research*, 80(2-3), 137-149.
- *Kraft, S. (2006). *Arbeitsgedächtnis bei Zwangsstörungen*. Unveröffentlichte Dissertation. Universität des Saarlandes, Homburg/Saar.
- *Krishna, R., Udupa, S., George, C. M., Kumar, K. J., Viswanath, B., Kandavel, T. et al. (2011). Neuropsychological performance in OCD: A study in medication-naive patients. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 35, 1969-1976.
- Kuelz, A. K., Hohagen, F. & Voderholzer, U. (2004). Neuropsychological performance in obsessive-compulsive disorder: A critical review. *Biological Psychology*, 65(3), 185-236.

- Kuelz, A. K., Riemann, D., Halsband, U., Vielhaber, K., Unterrainer, J., Kordon, A. et al. (2006). Neuropsychological impairment in obsessive-compulsive disorder--improvement over the course of cognitive behavioral treatment. *Journal of Clinical & Experimental Neuropsychology*, 28(8), 1273-1287.
- Kunimi, Mitsunobu, & Kojima, Haruyuki. (2014). The effects of processing speed and memory span on working memory. *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry*, 27(3), 109-114.
- Labad, J., Menchon, J. M., Alonso, P., Segalas, C., Jimenez, S., Jaurrieta, N. et al. (2008). Gender differences in obsessive-compulsive symptom dimensions. *Depression and Anxiety*, 25(10), 832-838.
- *Lawrence, N., Wooderson, S., Mataix-Cols, D., David, R., Speckens, A. & Phillips, M. (2006). Decision making and set shifting impairments are associated with distinct symptom dimensions in obsessive-compulsive disorder. *Neuropsychology*, 20(4), 409-419.
- Leckman, J. F., Denys, D., Simpson, H. B., Mataix-Cols, D., Hollander, E., Saxena, S. et al. (2010). Obsessive-compulsive disorder: A review of the diagnostic criteria and possible subtypes and dimensional specifiers for DSM-V. *Depression and Anxiety*, 27(6), 507-527.
- Leckman, J. F., Grice, D. E., Boardman, J. & Zhang, H. (1997). Symptoms of obsessive-compulsive disorder. *American Journal of Psychiatry*, 154(7), 911-917.
- Leckman, J. F., Bloch, M. H. & King, R. A. (2009). Symptom dimensions and subtypes of obsessive-compulsive disorder: a developmental perspective. *Dialogues in Clinical Neuroscience*, 11(1), 21-33.
- Leckman, J. F., Mataix-Cols, D. & Rosario-Campos, M. C. d. (2005a). Symptom Dimensions in OCD: Developmental and Evolutionary Perspectives. In J. S. Abramowitz & A. C. Houts (Hrsg.), *Concepts and controversies in obsessive-compulsive disorder* (S. 3-25). New York: Springer Science + Business Media.

- Leckman, J. F., Mataix-Cols, D. & Rosario-Campos, M. C. d. (2005b). Reply to Taylor: Combined Dimensional and Categorical Perspectives as an Integrative Approach to OCD. In J. S. Abramowitz & A. C. Houts (Hrsg.), *Concepts and controversies in obsessive-compulsive disorder* (S. 43-47). New York: Springer Science + Business Media.
- Leckman, J. F., Pauls, D. L., Zhang, H., Rosario-Campos, M. C., Katsovich, L., Kidd, K. K. et al. (2003). Obsessive-compulsive symptom dimensions in affected sibling pairs diagnosed with Gilles de la Tourette syndrome. *American journal of medical genetics. Part B, Neuropsychiatric genetics*, 116B(1), 60-68.
- Leckman, J. F., Rauch, S. L. & Mataix-Cols, D. (2007). Symptom dimensions in obsessive-compulsive disorder: Implications for the DSM-V. *CNS Spectrums*, 12(5), 376-387.
- Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis* (Bd. 49). Thousand Oaks, California: Sage.
- Lipszyc, J. & Schachar, R. (2010). Inhibitory control and psychopathology: A meta-analysis of studies using the stop signal task. *Journal of the International Neuropsychological Society*, 16(6), 1064-1076.
- *MacDonald, P. A., Antony, M. M., MacLeod, C. M. & Richter, M. A. (1997). Memory and confidence in memory judgments among individuals with obsessive compulsive disorder and non-clinical controls. *Behaviour Research and Therapy*, 35(6), 497-505.
- Mallet, L., Polosan, M., Jaafari, N., Baup, N., Welter, M.-L., Fontaine, D. et al. (2008). Subthalamic nucleus stimulation in severe obsessive-compulsive disorder. *The New England Journal of Medicine*, 359(20), 2121-2134.
- *Marsh, R., Horga, G., Parashar, N., Wang, Z., Peterson, B. S. & Simpson, H. B. (2014). Altered activation in fronto-striatal circuits during sequential processing of conflict in unmedicated adults with obsessive-compulsive disorder. *Biological Psychiatry* 75(8), 615-622.

- Mataix-Cols, D., Baer, L., Rauch, S. L. & Jenike, M. A. (2000). Relation of factor-analyzed symptom dimensions of obsessive-compulsive disorder to personality disorders. *Acta Psychiatrica Scandinavica*, 102(3), 199-202.
- Mataix-Cols, D., do Rosario-Campos, M. C. & Leckman, J. F. (2005). A Multidimensional Model of Obsessive-Compulsive Disorder. *American Journal of Psychiatry*, 162(2), 228-238.
- Mataix-Cols, D., Marks, I. M., Greist, J. H., Kobak, K. A. & Baer, L. (2002). Obsessive-compulsive symptom dimensions as predictors of compliance with and response to behaviour therapy: Results from a controlled trial. *Psychotherapy and Psychosomatics*, 71(5), 255-262.
- Mataix-Cols, D., Pertusa, A. & Snowden, J. (2011). Neuropsychological and neural correlates of hoarding: a practice-friendly review. *Journal of Clinical Psychology*, 67(5), 467-476.
- Mataix-Cols, D., Rauch, S. L., Baer, L., Eisen, J. L., Shera, D. M., Goodman, W. K. et al. (2002). Symptom stability in adult obsessive-compulsive disorder: Data from a naturalistic two-year follow-up study. *American Journal of Psychiatry*, 159(2), 263-268.
- Mataix-Cols, D., Rauch, S. L., Manzo, P. A., Jenike, M. A. & Baer, L. (1999). Use of factor-analyzed symptom dimensions to predict outcome with serotonin reuptake inhibitors and placebo in the treatment of obsessive-compulsive disorder. *American Journal of Psychiatry*, 156(9), 1409-1416.
- Mataix-Cols, D., Wooderson, S., Lawrence, N., Brammer, M. J., Speckens, A. & Phillips, M. L. (2004). Distinct neural correlates of washing, checking, and hoarding symptom dimensions in obsessive-compulsive disorder. *Archives of General Psychiatry*, 61(6), 564-576.
- McDermott, L. M. & Ebmeier, K. P. (2009). A meta-analysis of depression severity and cognitive function. *Journal of Affective Disorders*, 119(1-3), 1-8.
- McKay, D., Abramowitz, J. S., Calamari, J. E., Kyrios, M., Radomsky, A., Sookman, D. et al. (2004). A critical evaluation of obsessive-compulsive disorder subtypes: Symptoms versus mechanisms. *Clinical Psychology Review*, 24(3), 283-313.

- McMillan, S. G., Rees, C. S. & Pestell, C. (2013). An investigation of executive functioning, attention and working memory in compulsive hoarding. *Behavioural and Cognitive Psychotherapy*, 41(5), 610-625.
- Meca, J. S., López-Pina, J. A., López-López, J. A., Marín-Martínez, F., Rosa-Alcázar, A. I. & Gómez-Conesa, A. A. (2011). The Maudsley Obsessive-Compulsive Inventory: A reliability generalization meta-analysis. *International Journal of Clinical and Health Psychology*, 11(3), 473-493.
- Mega, M. S. & Cummings, J. L. (1994). Frontal-subcortical circuits and neuropsychiatric disorders. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 6(4), 358-370.
- Menzies, L., Chamberlain, S. R., Laird, A. R., Thelen, S. M., Sahakian, B. J. & Bullmore, E. T. (2008). Integrating evidence from neuroimaging and neuropsychological studies of obsessive-compulsive disorder: The orbitofronto-striatal model revisited. *Neuroscience and Biobehavioral Reviews*, 32(3), 525-549.
- Milad, M. R. & Rauch, S. L. (2012). Obsessive-compulsive disorder: beyond segregated cortico-striatal pathways. *Trends in Cognitive Sciences*, 16(1), 43-51.
- Miller, G. A. & Chapman, J. P. (2001). Misunderstanding Analysis of Covariance. *Journal of Abnormal Psychology*, 110(1), 40-48.
- Milner, B. (1963). Effects of different brain lesions on card sorting: the role of the frontal lobes. *Archives of Neurology*, 9, 90-100.
- Mindus, P., Rasmussen, S. A. & Lindquist, C. (1994). Neurosurgical treatment for refractory obsessive-compulsive disorder: Implications for understanding frontal lobe function. *Journal of Neuropsychiatry and Clinical Neurosciences*, 6(4), 467-477.
- Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Drummond, R. & Stroup, D. F. (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. *Lancet*, 354(9193), 1896-1900.

- Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *British Medical Journal*, 339, 332-336.
- Montgomery, S. A. & Åsberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, 134, 382-389.
- *Morein-Zamir, S., Pappmeyer, M., Pertusa, A., Chamberlain, S. R., Fineberg, N. A., Sahakian, B. J. et al. (2014). The profile of executive function in OCD hoarders and hoarding disorder. *Psychiatry Research*, 215(3), 659-667.
- Moritz, S., Fischer, B.-K., Hottenrott, B., Kellner, M., Fricke, S., Randjbar, S. et al. (2008). Words may not be enough! No increased emotional Stroop effect in obsessive-compulsive disorder. *Behaviour Research and Therapy*, 46(9), 1101-1104.
- Moritz, S., Hottenrott, B., Jelinek, L., Brooks, A. M., & Scheurich, A. (2012). Effects of obsessive-compulsive symptoms on neuropsychological test performance: Complicating an already complicated story. *Clinical Neuropsychologist*, 26(1), 31-44.
- Moritz, S., Jacobsen, D., Willenborg, B., Jelinek, L. & Fricke, S. (2006). A check on the memory deficit hypothesis of obsessive-compulsive checking. *European Archives of Psychiatry and Clinical Neuroscience*, 256(2), 82-86.
- Moritz, S., Kloss, M., von Eckstaedt, F. V. & Jelinek, L. (2009). Comparable performance of patients with obsessive-compulsive disorder (OCD) and healthy controls for verbal and nonverbal memory accuracy and confidence: time to forget the forgetfulness hypothesis of OCD? *Psychiatry Research*, 166(2-3), 247-253.
- Moritz, S., Ruhe, C., Jelinek, L. & Naber, D. (2009). No deficits in nonverbal memory, metamemory and internal as well as external source memory in obsessive-compulsive disorder (OCD). *Behaviour Research and Therapy*, 47(4), 308-315.

- Moritz, S., Wendt, M., Jelinek, L., Ruhe, C. & Arzola, G. M. (2008). No disadvantage for the processing of global visual features in obsessive-compulsive disorder. *Journal of the International Neuropsychological Society*, 14(3), 489-493.
- Muller, J. & Roberts, J. E. (2005). Memory and attention in Obsessive-Compulsive Disorder: A review. *Journal of Anxiety Disorders*, 19(1), 1-28.
- Müller, S. V., Hildebrandt, H. & Münte, T. F. (2004). *Kognitive Therapie bei Störungen der Exekutivfunktionen: Ein Therapiemanual*. Göttingen: Hogrefe.
- Najmi, S. & Amir, N. (2010). The effect of attention training on a behavioral test of contamination fears in individuals with subclinical obsessive-compulsive symptoms. *Journal of Abnormal Psychology*, 119(1), 136-142.
- *Nakao, T., Nakagawa, A., Nakatani, E., Nabeyama, M., Sanematsu, H., Yoshiura, T. et al. (2009). Working memory dysfunction in obsessive compulsive disorder: A neuropsychological and functional MRI study. *Journal of Psychiatric Research*, 43(8), 784.
- Nakao, T., Nakagawa, A., Yoshiura, T., Nakatani, E., Nabeyama, M., Sanematsu, H. et al. (2009). Duration effect of obsessive-compulsive disorder on cognitive function: A functional MRI study. *Depression and Anxiety*, 26(9), 814-823.
- Nakao, T., Nakagawa, A., Yoshiura, T., Nakatani, E., Nabeyama, M., Yoshizato, C. et al. (2005). Brain Activation of Patients with Obsessive-Compulsive Disorder During Neuropsychological and Symptom Provocation Tasks Before and After Symptom Improvement: A Functional Magnetic Resonance Imaging Study. *Biological Psychiatry*, 57(8), 901-910.
- *Nedeljkovic, M., Kyrios, M., Moulding, R., Doron, G., Wainwright, K., Pantelis, C. et al. (2009). Differences in neuropsychological performance between subtypes of obsessive-compulsive disorder. *Australian and New Zealand Journal of Psychiatry*, 43(3), 216-226.
- Nejati, V., Zabihzadeh, A., Maleki, G. & Safarzadeh, M. (2013). The Comparison of Executive Functions in Patients with Obsessive-Compulsive Disorder and Normal Individuals. *Journal of Clinical Psychology*, 5(1), 81-92.

- Nelson, H. E. & Wilson, J. (1991). *National Adult Reading Test (NART)*. Windsor, UK: NFER-Nelson.
- Olley, A., Malhi, G. & Sachdev, P. (2007). Memory and executive functioning in obsessive-compulsive disorder: A selective review. *Journal of Affective Disorders, 104*(1- 3), 15-23.
- *Omori, I. M., Murata, Y., Yamanishi, T., Nakaaki, S., Akechi, T., Mikuni, M. et al. (2007). The differential impact of executive attention dysfunction on episodic memory in obsessive-compulsive disorder patients with checking symptoms vs. those with washing symptoms. *Journal of Psychiatric Research, 41*(9), 776-784.
- Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics, 8*(2), 157-159
- Öst, L.-G. (2008). Efficacy of the third wave of behavioral therapies: A systematic review and meta-analysis. *Behaviour Research and Therapy, 46*(3), 296-321.
- Owen, A. M. (2005). Cognitive planning in humans: New insights from the Tower of London (TOL) task. In R. G. Morris & G. Ward (Hrsg.), *The cognitive psychology of planning* (S. 135–151). Hove, England: Psychology Press.
- Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E. & Robbins, T. W. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia, 28*, 1021-1034.
- Penadés, R., Catalán, R., Andrés, S., Salamero, M. & Gastó, C. (2005). Executive function and nonverbal memory in obsessive-compulsive disorder. *Psychiatry Research, 133*(1), 81-90.
- *Penadés, R., Catalán, R., Rubia, K., Andrés, S., Salamero, M. & Gastó, C. (2007). Impaired response inhibition in obsessive compulsive disorder. *European Psychiatry, 22*(6), 404-410.
- Pertusa, A., Frost, R. O. & Mataix-Cols, D. (2010). When hoarding is a symptom of OCD: A case series and implications for DSM–V. *Behaviour Research and Therapy, 48*(10), 1012-1020.

- Pinto, A., Eisen, J. L., Mancebo, M., Greenberg, B. D., Stout, R. L. & Rasmussen, S. (2007). Taboo thoughts and doubt/checking: a refinement of the factor structure for obsessive-compulsive disorder symptoms. *Psychiatry Research*, *151*(3), 255-258.
- *Pinto, P. S., Iego, S., Nunes, S., Menezes, H., Mastroiosa, R. S., Oliveira, I. R. et al. (2011). Influence of specific obsessive-compulsive symptom dimensions on strategic planning in patients with obsessive-compulsive disorder. *Revista Brasileira de Psiquiatria*, *33*(1), 40-46.
- Polk, T. A., Simen, P., Lewis, R. L. & Freedman, E. (2002). A computational approach to control in complex cognition. *Cognitive Brain Research* *15*(1), 71-83.
- Rachman, S. (1993). Obsessions, responsibility and guilt. *Behaviour Research and Therapy*, *31*(2), 149-154.
- *Radomsky, A. S. & Rachman, S. (1999). Memory bias in obsessive-compulsive disorder (OCD). *Behaviour Research & Therapy*, *37*(7), 605-618.
- Radomsky, A. S. & Rachman, S. (2004). Symmetry, ordering and arranging compulsive behaviour. *Behaviour Research & Therapy*, *42*(8), 893-913.
- *Rampacher, F. (2006). *Neuropsychologie der Zwangserkrankung: Störungsspezifische oder depressionsbedingte Defizite?* Unveröffentlichte Dissertation. Humboldt-Universität, Berlin.
- Rasmussen, S. A. & Eisen, J. L. (1989). Clinical features and phenomenology of Obsessive Compulsive Disorder. *Psychiatric Annals*, *19*(2), 67-73.
- Rasmussen, S. A., Eisen, J. L. & Greenberg, B. D. (2013). Toward a neuroanatomy of obsessive-compulsive disorder revisited. *Biological Psychiatry*, *73*(4), 298-299.
- Rauch, S. L., Jenike, M. A., Alpert, N. M. & Baer, L. (1994). Regional cerebral blood flow measured during symptom provocation in obsessive-compulsive disorder using oxygen 15-labeled carbon dioxide and positron emission tomography. *Archives of General Psychiatry*, *51*(1), 62-70.
- Raven, J. C., Kratzmeier, H. & Horn, R. (Hrsg.) (1979). *Raven-Matrizen-Test: standard progressive matrices; SPM*. Weinheim: Beltz Test.

- Reed, G. F. (1985). *Obsessional experience and compulsive behaviour: A cognitive-structural approach* (Bd. 34). Orlando: Academic Press.
- Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills*, 8, 271-276.
- Robinson, A. L., Heaton, R. K., Lehman, R. A. & Stilson, D. W. (1980). The utility of the Wisconsin Card Sorting Test in detecting and localizing frontal lobe lesions. *Journal of Consulting and Clinical Psychology*, 48(5), 605-614.
- Robinson, L. J., Thompson, J. M., Gallagher, P., Goswami, U., Young, A. H., Ferrier, I. N. et al. (2006). A meta-analysis of cognitive deficits in euthymic patients with bipolar disorder. *Journal of Affective Disorders*, 93(1-3), 105-15.
- Rosario-Campos, M. C., Miguel, E. C., Quatrano, S., Chacon, P., Ferrao, Y., Findley, D. et al. (2006). The Dimensional Yale-Brown Obsessive-Compulsive Scale (DY-BOCS): An instrument for assessing obsessive-compulsive symptom dimensions. *Molecular Psychiatry*, 11(5), 495-504.
- Rosenberg, D. R., Dick, E. L., O'Hearn, K. M. & Sweeney, J. A. (1997). Response-inhibition deficits in obsessive-compulsive disorder: An indicator of dysfunction in frontostriatal circuits. *Journal of Psychiatry & Neuroscience*, 22(1), 29-38.
- Rotge, J.-Y., Guehl, D., Dilharreguy, B., Cuny, E., Tignol, J., Bioulac, B. et al. (2008). Provocation of obsessive-compulsive symptoms: A quantitative voxel-based meta-analysis of functional neuroimaging studies. *Journal of Psychiatry & Neuroscience*, 33(5), 405-412.
- Rotge, J.-Y., Guehl, D., Dilharreguy, B., Tignol, J., Bioulac, B., Allard, M. et al. (2009). Meta-analysis of brain volume changes in obsessive-compulsive disorder. *Biological Psychiatry*, 65(1), 75-83.
- Rotge, J.-Y., Langbour, N., Jaafari, N., Guehl, D., Bioulac, B., Aouizerate, B. et al. (2010). Anatomical alterations and symptom-related functional activity in obsessive-compulsive disorder are correlated in the lateral orbitofrontal cortex. *Biological Psychiatry*, 67(7), e37-e38.

- Roth, R. M. & Baribeau, J. (1996). Performance of subclinical compulsive checkers on putative tests of frontal and temporal lobe memory functions. *Journal of Nervous & Mental Disease*, 184(7), 411-416.
- Rubenstein, C. S., Peynircioglu, Z. F., Chambless, D. L. & Pigott, T. A. (1993). Memory in sub-clinical obsessive-compulsive checkers. *Behaviour Research and Therapy*, 31(8), 759-765.
- Rück, C., Karlsson, A., Steele, J. D., Edman, G., Meyerson, B. A., Ericson, K., et al. (2008). Capsulotomy for obsessive-compulsive disorder: Long-term follow-up of 25 patients. *Archives of General Psychiatry*, 65(8), 914-921.
- Ruscio, A. M., Stein, D. J., Chiu, W. T. & Kessler, R. C. (2010). The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Molecular Psychiatry*, 15(1), 53-63.
- Sacks, O. W. (1988). *Der Mann, der seine Frau mit einem Hut verwechselte*. Reinbek bei Hamburg: Rowohlt.
- Salkovskis, P. M. (1985). Obsessional-compulsive problems: A cognitive-behavioural analysis. *Behaviour Research and Therapy*, 23(5), 571-583.
- Salomoni, G., Grassi, M., Mosini, P., Riva, P., Cavellini, P. & Bellodi, L. (2009). Artificial neural networks model for the prediction of obsessive-compulsive disorder treatment response. *Journal of Clinical Psychopharmacology*, 29(4), 343-349.
- Sanavio, E. (1988). Obsessions and compulsions: The Padua Inventory. *Behaviour Research and Therapy*, 26(2), 169-177.
- Sartory, G., Röper, G., Pietrowsky, R., Cwik, J., & Zaudig, M. (2013). Symptomkategorien der Zwangsstörung—spezifisch oder generalisiert? Ein Vergleich zwischen der Yale-Brown Obsessive-Compulsive Skala (Y-BOCS) und dem Vancouver Obsessional Compulsive Inventar (VOCI). *Zeitschrift für Klinische Psychologie und Psychotherapie: Forschung und Praxis*, 42(2), 107-117.

- Saß, H. & Houben, I. (Hrsg.) (1996). *Diagnostisches und statistisches Manual psychischer Störungen DSM-IV: Übersetzung nach der vierten Auflage des Diagnostic and Statistical Manual of Mental Disorders der American Psychiatric Association*. Göttingen: Hogrefe.
- Savage, C. R., Baer, L., Keuthen, N. J., Brown, H. D., Rauch, S. L. & Jenike, M. A. (1999). Organizational strategies mediate nonverbal memory impairment in obsessive-compulsive disorder. *Biological Psychiatry*, 45(7), 905-916.
- Savage, C. R., Deckersbach, T., Wilhelm, S., Rauch, S. L., Baer, L., Reid, T., et al. (2000). Strategic processing and episodic memory impairment in obsessive compulsive disorder. *Neuropsychology*, 14(1), 141-151.
- Saxena, S., Brody, A. L., Schwartz, J. M. & Baxter, L. R. (1998). Neuroimaging and frontal-subcortical circuitry in obsessive-compulsive disorder. *British Journal of Psychiatry*, 173(Suppl 35), 26-37.
- Saxena, S., & Rauch, S. L. (2000). Functional neuroimaging and the neuroanatomy of obsessive-compulsive disorder. *Psychiatric Clinics of North America*, 23(3), 563-586.
- Schandry, R. (2011). *Biologische Psychologie* (3. Aufl.). Weinheim: Beltz.
- Schiepek, G., Tominschek, I., Karch, S., Mulert, C., Pogarell, O. & Zaudig, M. (2007). Neurobiologische Korrelate der Zwangsstörungen – Aktuelle Befunde zur funktionellen Bildgebung. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 57(9), 379-394.
- Schmidtke, K., Schorb, A., Winkelmann, G., & Hohagen, F. (1998). Cognitive frontal lobe dysfunction in obsessive-compulsive disorder. *Biological Psychiatry*, 43(9), 666-673.
- Schwarzer, G. & Rücker, G. (2010). Statistische Methoden zur Detektion und Adjustierung von Publikationsbias. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, 104(4), 306-313.
- Shafran, R. (2001). Obsessive-compulsive disorder in children and adolescents. *Child Psychology & Psychiatry Review*, 6(2), 50-58.

- Shavitt, R. G., Belotto, C., Curi, M., Hounie, A. G., Rosário-Campos, M. C., Diniz, J. B. et al. (2006). Clinical features associated with treatment response in obsessive-compulsive disorder. *Comprehensive Psychiatry*, 47(4), 276-281.
- Shin, N. Y., Kang, D.-H., Choi, J.-S., Jung, M. H., Jang, J. H. & Kwon, J. S. (2010). Do organizational strategies mediate nonverbal memory impairment in drug-naïve patients with obsessive-compulsive disorder? *Neuropsychology*, 24(4), 527-533.
- Shin, N. Y., Lee, T. Y., Kim, E. & Kwon, J. S. (2014). Cognitive functioning in obsessive-compulsive disorder: A meta-analysis. *Psychological Medicine*, 44(6), 1121-1130.
- Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *Journal of Psychology: Interdisciplinary and Applied*, 9, 371-377.
- Shipley, W. C., Gruber, C. P., Martin, T. A. & Klein, A. M. (2009). *Shipley-2 Manual*. Los Angeles, California: Western Psychological Services.
- Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420-428.
- Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, 81, 174-176.
- Simpson, H. B., Rosen, W., Huppert, J. D., Lin, S.-H., Foa, E. B. & Liebowitz, M. R. (2006). Are there reliable neuropsychological deficits in obsessive-compulsive disorder? *Journal of Psychiatric Research*, 40(3), 247-257.
- *Singh, S., Mukundan, C. R. & Khanna, S. (2003). Working memory deficits in obsessive-compulsive disorder. *Psychological Studies*, 48(2), 69-73.
- Snider, L. A. & Swedo, S. E. (2000). Pediatric obsessive-compulsive disorder. *JAMA: Journal of the American Medical Association*, 284(24), 3104-3106.
- Sookman, D., Abramowitz, J. S., Calamari, J. E., Wilhelm, S. & McKay, D. (2005). Subtypes of Obsessive-Compulsive Disorder: Implications for Specialized Cognitive Behavior Therapy. *Behavior Therapy*, 36(4), 393-400.

- Spielberger, C. D., Gorsuch, R. L. & Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, California: Consulting Psychologists Press.
- Starcevic, V. & Brakoulias, V. (2008). Symptom subtypes of obsessive-compulsive disorder: Are they relevant for treatment? *Australian and New Zealand Journal of Psychiatry*, 42(8), 651-661.
- Steingass, H.-P. (Hrsg.). (1999). *Neuropsychologie und Sucht*. Geesthacht: Neuland.
- Steketee, G. & Frost, R. O. (2006). *Compulsive Hoarding and Acquiring: Therapist Guide (Treatments that work)*. Oxford: University Press.
- Steketee, G. & Frost, R. (2008). *Clutter Image Rating*. Zugriff am 14. 8. 2014, von <http://www.ocfoundation.org/hoarding/cir.pdf>
- Steketee, G., Frost, R. O. & Kyrios, M. (2003). Cognitive aspects of compulsive hoarding. *Cognitive Therapy and Research*, 27(4), 463-479.
- Steketee, G., Grayson, J. B. & Foa, E. B. (1985). Obsessive-compulsive disorder: Differences between washers and checkers. *Behaviour Research and Therapy*, 23(2), 197-201.
- Steketee, G. & Tolin, D. F. (2011). Cognitive-behavioral therapy for hoarding in the context of contamination fears. *Journal of Clinical Psychology, Special Issue: hoarding disorder*, 67(5), 485-496.
- Storch, E. A., Rasmussen, S. A., Price, L. H., Larson, M. J., Murphy, T. K., & Goodman, W. K. (2010). Development and psychometric evaluation of the Yale–Brown Obsessive-Compulsive Scale—Second Edition. *Psychological Assessment*, 22(2), 223-232.
- Taj, M. J, R. J., Viswanath, B., Purushottam, M., Kandavel, T., Janardhan Reddy, Y. C. & Jain, S. (2013). DRD4 gene and obsessive compulsive disorder: Do symptom dimensions have specific genetic correlates? *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 41, 18-23.
- Tallis, F. (1995). *Obsessive compulsive disorder: A cognitive and neuropsychological perspective*. Oxford, England: John Wiley & Sons.

- Tallis, F. (1997). The neuropsychology of obsessive-compulsive disorder: A review and consideration of clinical implications. *British Journal of Clinical Psychology*, 36(1), 3-20.
- *Tallis, F., Pratt, P. & Jamani, N. (1999). Obsessive compulsive disorder, checking, and non-verbal memory: A neuropsychological investigation. *Behaviour Research and Therapy*, 37(2), 161-166.
- Taylor, S. (2005). Dimensional and Subtype Models of OCD. In J. S. Abramowitz & A. C. Houts (Hrsg.), *Concepts and Controversies in Obsessive Compulsive Disorder* (S. 27-52). New York: Springer.
- Tolin, D. F., Stevens, M. C., Nave, A., Villavicencio, A. L. & Morrison, S. (2012). Neural mechanisms of cognitive behavioral therapy response in *hoarding disorder*: A pilot study. *Journal of Obsessive-Compulsive and Related Disorders*, 1(3), 180-188.
- *Tolin, D. F., Stevens, M. C., Villavicencio, A. L., Norberg, M. M., Calhoun, V. D., Frost, R. O. et al. (2012). Neural mechanisms of decision making in hoarding disorder. *Archives of General Psychiatry*, 69(8), 832-841.
- *Tolin, D. F. & Villavicencio, A. (2011). An exploration of economic reasoning in *hoarding disorder* patients. *Behaviour Research and Therapy*, 49(12), 914-919.
- *Tolin, D. F., Villavicencio, A., Umbach, A. & Kurtz, M. M. (2011). Neuropsychological functioning in hoarding disorder. *Psychiatry Research*, 189, 413-418.
- *Tolin, D. F., Witt, S. T. & Stevens, M. C. (2014). Hoarding disorder and obsessive-compulsive disorder show different patterns of neural activity during response inhibition. *Psychiatry Research: Neuroimaging*, 221(2), 142-148.
- Tükel, R., Polat, A., Özdemir, Ö., Aksüt, D. & Türksoy, N. (2002). Comorbid conditions in obsessive-compulsive disorder. *Comprehensive Psychiatry*, 43(3), 204-209.
- Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science*, 2(3), 67-70.

- *Tumkaya, S., Karadag, F., Mueller, S. T., Ugurlu, T. T., Oguzhanoglu, N. K., Ozdel, O. et al. (2013). Situation awareness in obsessive-compulsive disorder. *Psychiatry Research*, 209(3), 579-588.
- *Tuna, Ş., Tekcan, A. I. & Topçuoğlu, V. (2005). Memory and metamemory in obsessive-compulsive disorder. *Behaviour Research and Therapy*, 43(1), 15-27.
- Van den Heuvel, O. A., Remijnse, P. L., Mataix-Cols, D., Vrenken, H., Groenewegen, H. J., Uylings, H. B. M. et al. (2009). The major symptom dimensions of obsessive-compulsive disorder are mediated by partially distinct neural systems. *Brain: A Journal of Neurology*, 132(4), 853-868.
- *Van der Meulen, M. T. (2013). *Inhibition and flexibility in Hoarding and OCD: similarities in neuropsychological profile*. Unveröffentlichte Masterarbeit. Universiteit Utrecht, Niederlande.
- Van Grootheest, D. S., Boomsma, D. I., Hettema, J. M. & Kendler, K. S. (2008). Heritability of obsessive-compulsive symptom dimensions. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 147B(4), 473-478.
- Van Horn, J. D., Irimia, A., Torgerson, C. M., Chambers, M. C., Kikinis, R. & Toga, A. W. (2012). Mapping connectivity damage in the case of Phineas Gage. *PLoS ONE*, 7(5).
- Van Oppen, P., Hoekstra, R. J. & Emmelkamp, P. (1995). The structure of obsessive-compulsive symptoms. *Behaviour Research and Therapy*, 33(1), 15-23.
- Voderholzer, U. & Hohagen, F. (Hrsg.) (2013). *Therapie psychischer Erkrankungen: State of the Art 2012/2013* (8. Aufl.). München: Urban & Fischer.
- Voderholzer, U., Schlegl, S. & Külz, A. K. (2011). Epidemiologie und Versorgungssituation von Zwangsstörungen. *Der Nervenarzt*, 82(3), 273-280.
- Watanabe, K., Ogino, T., Nakano, K., Hattori, J., Kado, Y., Satoshi, S. et al. (2005). The Boston Qualitative Scoring System for the Rey–Osterrieth Complex Figure. *Brain and Development* 2, 564-569.
- Weigel, R., Krauss, J. K. & Jones, E. G. (2004). Anatomie und Nomenklatur der Basalganglien und des Thalamus. In J. K. Krauss & J. Feldmann (Hrsg.), *Tiefe Hirnstimulation* (S. 17-56). Darmstadt: Steinkopf.

- Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A. & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin*, 117(3), 450-468.
- Wheaton, M., Timpano, K. R., LaSalle-Ricci, V. H. & Murphy, D. (2008). Characterizing the hoarding phenotype in individuals with OCD: Associations with comorbidity, severity and gender. *Journal of Anxiety Disorders*, 22(2), 243-252.
- Whiteside, S. P., Port, J. D. & Abramowitz, J. S. (2004). A meta-analysis of functional neuroimaging in obsessive-compulsive disorder. *Psychiatry Research: Neuroimaging*, 132(1), 69-79.
- *Wincze, J. P., Steketee, G. & Frost, R. O. (2007). Categorization in compulsive hoarding. *Behaviour Research and Therapy*, 45(1), 63-72.
- Woods, C. M., Vevea, J. L., Chambless, D. L. & Bayen, U. J. (2002). Are compulsive checkers impaired in memory? A meta-analytic review. *Clinical Psychology: Science and Practice*, 9(4), 353-366.
- Zaudig, M. & Bossert-Zaudig, S. (Hrsg.). (2002). *Die Zwangsstörung: Diagnostik und Therapie* (2. Aufl.). Stuttgart: Schattauer.
- Zhu, X.-Z., Wang, H.-X. & Yi, J.-Y. (2013). Executive function of the patients with obsessive-compulsive disorder in conflictive task. *Chinese Journal of Clinical Psychology*, 21(2), 185-187.
- Zielinski, C. M., Taylor, M. A. & Juzwin, K. R. (1991). Neuropsychological deficits in obsessive-compulsive disorder. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology*, 4(2), 110-126.
- *Zitterl, W., Urban, C., Linzmayer, L., Aigner, M., Demal, U., Semler, B. et al. (2001). Memory deficits in patients with DSM-IV obsessive-compulsive disorder. *Psychopathology*, 34(3), 113-117.
- Zor, R., Fineberg, N., Eilam, D. & Hermesh, H. (2011). Video telemetry and behavioral analysis discriminate between compulsive cleaning and compulsive checking in obsessive-compulsive disorder. *European Neuropsychopharmacology* 21, 814-824.

Tabellenverzeichnis

Tabelle 1. <i>In der Metaanalyse untersuchte neuropsychologische Testverfahren, klassifiziert nach übergeordneten Domänen und Subdomänen.....</i>	62
Tabelle 2. <i>Skala zur Beurteilung der Studienqualität.....</i>	65
Tabelle 3. <i>Eingeschlossene Studien, k = 38.</i>	84
Tabelle 4: <i>Demografische und klinische Merkmale der Studienteilnehmer.</i>	92
Tabelle 5. <i>Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Patienten mit Waschzwang vs. Gesunde.....</i>	95
Tabelle 6. <i>Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Patienten mit Kontrollzwang vs. Gesunde.....</i>	100
Tabelle 7. <i>Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Zwangspatienten mit verbotenen Gedanken vs. Gesunde.....</i>	103
Tabelle 8. <i>Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Patienten mit Symmetrie-/ Ordnungszwang vs. Gesunde.....</i>	105
Tabelle 9. <i>Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Patienten mit Sammelzwang vs. Gesunde.....</i>	109
Tabelle 10. <i>Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Patienten mit Waschzwang vs. Patienten mit Kontrollzwang.</i>	114
Tabelle 11. <i>Kombinierte Effektstärken nach dem fixed effect model und dem random effects model mit Heterogenitätsmaßen für Patienten mit Sammelzwang vs. Zwangspatienten ohne Sammelzwang.</i>	118
Tabelle 12. <i>Signifikante Effektstärken (Hedges' g, random effects, $p < .05$) der 10 kognitiven Subdomänen über alle 7 Metaanalysen.</i>	121
Tabelle 13: <i>Interraterreliabilitäten der Items der Skala zur Beurteilung der Studienqualität.....</i>	214

Abbildungsverzeichnis

<i>Abbildung 1.</i> Endophänotypen. Hierarchisches Modell der Ebenen, auf denen sich pathologische Phänomene manifestieren können. Nach Chamberlain und Menzies (2009, S. 1134).....	11
<i>Abbildung 2.</i> Cortico-striato-thalamisches Modell der Zwangsstörung. Nach Milad und Rauch (2012, S. 44).....	14
<i>Abbildung 3.</i> Modell zur Bedeutung gestörter Inhibitionsprozesse der Zwangsstörung. Nach Chamberlain et al. (2005, S. 412)	18
<i>Abbildung 4.</i> Tower of London (Polk, Simen, Lewis & Freedman, 2002, S. 73).	21
<i>Abbildung 5.</i> Farbe-Wort-Stroop-Aufgabe. Nach Golden (1978, zitiert nach Hashimoto et al., 2011).....	22
<i>Abbildung 6.</i> Wisconsin Card Sorting Test (WCST; Hilger & Kasper, 2002, S. 20).....	25
<i>Abbildung 7.</i> Gedächtnismodell. Nach Bredenkamp (1998, S. 63).....	28
<i>Abbildung 8.</i> Rey-Osterrieth Complex Figure Test mit Organisationsmerkmalen der Enkodierung: Gestaltbildende Elemente (“configural elements”), Cluster und Details. (Watanabe et al., 2005, S. 565).....	30
<i>Abbildung 9.</i> Zahlen-Symbol-Test (Wechsler, 1981).....	32
<i>Abbildung 10.</i> Clutter Image Rating Scale (Steketee & Frost, 2008, S. 3).	44
<i>Abbildung 11.</i> Umrechnen von Effektstärken. Nach Borenstein et al. (2009, S. 46)	68
<i>Abbildung 12.</i> Schema eines <i>forest plots</i> nach Borenstein et al. (2009, S. 367).	70
<i>Abbildung 13.</i> Anzahl gefundener Veröffentlichungen.....	79
<i>Abbildung 14.</i> Flussdiagramm mit ein- und ausgeschlossenen Primärstudien.....	81
<i>Abbildung 15.</i> Publikationsjahre der eingeschlossenen Studien bis Januar 2014.	82
<i>Abbildung 16.</i> Histogramm zur Qualität der eingeschlossenen Studien.....	89
<i>Abbildung 17.</i> Forest plot. Patienten mit Waschzwang vs. Gesunde.	94
<i>Abbildung 18.</i> Funnel plot. Patienten mit Waschzwang vs. Gesunde.	96

<i>Abbildung 19.</i> Forest plot. Patienten mit Kontrollzwang vs. Gesunde.....	99
<i>Abbildung 20.</i> Funnel plot. Patienten mit Kontrollzwang vs. Gesunde.....	101
<i>Abbildung 21.</i> Forest plot. Zwangspatienten mit verbotenen Gedanken vs. Gesunde. .	102
<i>Abbildung 22.</i> Forest plot. Patienten mit Symmetrie-/Ordnungszwang vs. Gesunde. ..	104
<i>Abbildung 23.</i> Forest plot. Patienten mit Sammelzwang vs. Gesunde.	108
<i>Abbildung 24.</i> Funnel plot. Patienten mit Sammelzwang vs. Gesunde.....	110
<i>Abbildung 25.</i> Forest plot. Patienten mit Waschzwang vs. Kontrollzwang.	113
<i>Abbildung 26.</i> Funnel plot. Patienten mit Waschzwang vs. Kontrollzwang.	115
<i>Abbildung 27.</i> Forest plot. Patienten mit vs. ohne Sammelzwang.	117
<i>Abbildung 28.</i> Funnel plot. Patienten mit vs. ohne Sammelzwang.	119
<i>Abbildung 29.</i> Geometrisches Netzwerk eines <i>umbrella reviews</i> . In Anlehnung an Ioannidis (2009, S. 489).	123

Abkürzungsverzeichnis

ACC	Anteriorer cingulärer Cortex
ATT	Attention Training Technique
AGGR	Aggressive Zwangsgedanken
BAI	Beck Anxiety Inventory
BDI	Beck Depression Inventory
BDI-II	Beck Depression Inventory- 2
CANTAB	Cambridge Automated Neuropsychological Test Battery
CGT	Cambridge Gambling Task
CHECK	Kontrollzwang (checking)
COWAT-FAS	Controlled Oral Word Association Test mit den Anfangsbuchstaben F, A, S
CPT	Continuous Performance Test
CPT-II	Continuous Performance Test, Teil 2
CT	Computertomografie
DASS	Depression Anxiety Stress Scale
DLPFC	Dorsolateraler Präfrontalcortex
DSM-III	Diagnostic and Statistical Manual for Mental Diseases, third edition
DSM-IV	Diagnostisches und Statistisches Manual Psychischer Störungen, vierte Revision
DSM-5	Diagnostic and Statistical Manual for Mental Diseases, fifth edition
DY-BOCS	Dimensional Yale-Brown Obsessive Compulsive Scale
EEG	Elektroenzephalogramm
HAM-A	Hamilton Anxiety Scale
HAM-D	Hamilton Depression Scale
HC	Gesunde Kontrollgruppe (healthy controls)
HOARD	Sammelzwang (hoarding)
ICC	Intraclass Correlation Coefficient
ICD-10	Internationale Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, Revision 10
IED	Intra-Extra Dimensional Shift

IGT	Iowa Gambling Task
LZG	Langzeitgedächtnis
MADRS	Montgomery–Åsberg Depression Rating Scale
MOCI	Maudsley Obsessive-Compulsive Inventory
MRT	Magnetresonanztomografie
MULT	multiple Zwangssymptome
MWT–B	Mehrfachwahl-Wortschatztest, Teil B
NART	National Adult Reading Test
OAD	Andere Angststörung (other anxiety disorder)
OAT	Object Alternation Test
OCD	Zwangsstörung (obsessive-compulsive disorder)
OCI–R	Obsessive-Compulsive Inventory–Revised
OFC	Orbitofrontalcortex
OR	Odds Ratio
PFC	Präfrontalcortex
PI–R	Padua Inventory–Revised
PI-WUSR	Padua Inventory–Washington State University Revision
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUOROM	Quality Of Reporting Of Meta-Analyses
RCFT	Rey–Osterrieth Complex Figure Test
RCT	Randomisierte Kontrollgruppenstudie (randomized controlled trial)
RT	Reaktionszeit (reaction time)
SE	Standardfehler (standard error)
SI–R	Saving Inventory–Revised
SILS	Shipley Institute of Living Scale
SOC	Stockings Of Cambridge
SSRI	Selektive Serotoninwiederaufnahmememmer (selective serotonin reuptake inhibitor)
STAI–S	State Trait Anxiety Inventory, Subtest State-Anxiety
STAI–T	State Trait Anxiety Inventory, Subtest Trait-Anxiety
SYMM	Symmetrie-/Ordnungszwang
TABU	Verbotene Gedanken (taboo thoughts)

TMT-A	Trail Making Test, Teil A
TMT-B	Trail Making Test, Teil B
TOH	Tower of Hanoi
TOL	Tower of London
TAU	Treatment as usual
UKZ	Ultrakurzzeitgedächtnis
WAIS-R	Wechsler Adult Intelligence Scale-Revised
WASH	Waschzwang (washing)
WCST	Wisconsin Card Sorting Test
WMS-R	Wechsler Memory Scale-Revised
WTAR	Wechsler Test of Adult Reading
Y-BOCS	Yale-Brown Obsessive Compulsive Scale
Y-BOCS-SC	Dimensional Yale-Brown Obsessive Compulsive Scale

Anhang 1 Kodierung deskriptiver Daten

Kodierschema für deskriptive Daten: Metaanalyse OCD und Subgruppen

1 METHODE			
Reviewer	Name des Reviewers	1 <input type="checkbox"/> Leopold	2 <input type="checkbox"/> Oehmann
ausgefüllt_am	Datum _____		
Anm_intern	Interne Anmerkung, wird nicht publiziert		
Wave	Erhebungswelle	1 <input type="checkbox"/>	2 <input type="checkbox"/>
Source	Fundort der Referenz	1 <input type="checkbox"/> Datenbank	2 <input type="checkbox"/> Handsuche
Study_ID	Studien-Identifikationsnummer _____		
Autor	Studienautor		
Jahr	Jahr der Veröffentlichung _____		
Titel	Titel der Publikation		
Journal	Fachzeitschrift		
Land	Land, in dem die Studie durchgeführt wurde	_____	
Klassifik_system	Welches Klassifikationssystem wurde benutzt?		
	1 <input type="checkbox"/> ICD- 10	2 <input type="checkbox"/> DSM-IV	3 <input type="checkbox"/> ICD-9
	5 <input type="checkbox"/> andere	999 <input type="checkbox"/> keine Angabe	4 <input type="checkbox"/> DSM-III
Instr_schwere	Instrument, mit dem Schweregrad erfasst wurde		
	1 <input type="checkbox"/> Y-BOCS	2 <input type="checkbox"/> OCI-R	3 <input type="checkbox"/> MOCI-R
	5 <input type="checkbox"/> Padua-Inventar	6 <input type="checkbox"/> andere	4 <input type="checkbox"/> selbst entwickelte Skala
			999 <input type="checkbox"/> keine Angabe
Instr_Komorb	Instrumente, die Komorbidität erfassten	_____	
Instr_IQ	Instrumente, die Intelligenz erfassten	_____	
2 STUDIENQUALITÄT			
quali1_ethik	Einhaltung von Ethikrichtlinien	<input type="checkbox"/> 0	<input type="checkbox"/> 1 <input type="checkbox"/> 2
quali2_klare_SP	Klarheit der Stichprobenbeschreibung	<input type="checkbox"/> 0	<input type="checkbox"/> 1 <input type="checkbox"/> 2
quali3_schwere	Schweregrad der Störung	<input type="checkbox"/> 0	<input type="checkbox"/> 1 <input type="checkbox"/> 2
quali4_komorbid	Komorbidität der Stichprobe	<input type="checkbox"/> 0	<input type="checkbox"/> 1 <input type="checkbox"/> 2
quali5_diagnose	Reliabilität der Diagnose	<input type="checkbox"/> 0	<input type="checkbox"/> 1 <input type="checkbox"/> 2
quali6_spezif_neuro	Spezifität der neuropsycholog. Variablen	<input type="checkbox"/> 0	<input type="checkbox"/> 1 <input type="checkbox"/> 2
quali7_valid_test	Validität der neuropsychologischen Tests	<input type="checkbox"/> 0	<input type="checkbox"/> 1 <input type="checkbox"/> 2
quali8_n_testleiter	Anzahl der Testleiter	<input type="checkbox"/> 0	<input type="checkbox"/> 1 <input type="checkbox"/> 2
quali9_training	Training der Testleiter	<input type="checkbox"/> 0	<input type="checkbox"/> 1 <input type="checkbox"/> 2
quali10_medikontr	Kontrolle der Medikation	<input type="checkbox"/> 0	<input type="checkbox"/> 1 <input type="checkbox"/> 2
quali11_statist	Statistische Methoden und Datenpräsentation	<input type="checkbox"/> 0	<input type="checkbox"/> 1 <input type="checkbox"/> 2

3 TEILNEHMER

EG

- 1 Waschzwang
- 2 Kontrollzwang
- 3 aggressive Zwangsgedanken
- 4 Ordnungszwang
- 5 Hoarding
- 6 wenig Einsicht
- 7 early onset
- 8 tic-related

KG

- 1 gesund
- 2 Kontrollzwang
- 3 Zwang ohne Hoarding
- 4 viel (ausreichend) Einsicht
- 5 late onset
- 6 OCD ohne Tics

4 Experimentalgruppe

EG_n	Anzahl Pat. in der EG _____	MOCI-R M (SD) _____
EG_n_female	Anzahl Frauen in der EG _____	OCI-R M (SD) _____
EG_perc_female	% Frauen _____	PI M (SD) _____
EG_M_age	Alter (Mittelwert) _____ *	Alter (SD) _____ *
EG_range_age	Altersrange _____ *	
EG_Händigkeit	1 <input type="checkbox"/> alle Rechtshänder 2 <input type="checkbox"/> alle Linkshänder	3 <input type="checkbox"/> mixed 999 <input type="checkbox"/> k.A. *
EG_perc_Rechtshänder	Wieviel % der EG waren Rechtshänder?	_____ % *
EG_n_komorbid	Anzahl komorbider Patienten in der EG	_____ *
EG_perc_komorb	Wieviel % der EG waren komorbid?	_____ % *
EG_Art_komorbid	1 <input type="checkbox"/> Depression 2 <input type="checkbox"/> Angst 3 <input type="checkbox"/> Substanz 6 <input type="checkbox"/> mehrere 7 <input type="checkbox"/> Angst + Depression	4 <input type="checkbox"/> andere 5 <input type="checkbox"/> keine 999 <input type="checkbox"/> keine Angaben
EG_n_medi	Anzahl medizierter Patienten in der EG	_____ *
EG_perc_medi	Wieviel % der EG waren mediziert?	_____ % *
EG_Art_medi	Welche Medikamente erhielt die EG?	_____
EG_IQ	IQ der EG, M(SD) _____ *	
EG_edu_y	Bildungsjahre der EG, M (SD) _____ *	
EG_setting	1 <input type="checkbox"/> ambulant 2 <input type="checkbox"/> stationär	3 <input type="checkbox"/> mixed 999 <input type="checkbox"/> k.A. *
EG_Diagnose	Wie lautet die Diagnose in der EG?	_____
EG_M_Dauer_Jahre	Mittlere Krankheitsdauer	_____ *
EG_SD_Dauer_Jahre	SD der Krankheitsdauer	_____ *
EG_M_onset	Mittleres Ersterkrankungsalter	_____ *
EG_SD_onset	SD des Ersterkrankungsalters	_____ *
EG_M_Y_BOCS	Mittelwert Y BOCS, Cutoff 16	_____ *
EG_SD_Y_BOCS	SD der Y BOCS	_____ *
EG_Y_BOCS_severity	Schweregrad der Y BOCS *	
1 <input type="checkbox"/> 0- 7 leicht		
2 <input type="checkbox"/> 8- 15 mild		
3 <input type="checkbox"/> 16- 23 moderat		
4 <input type="checkbox"/> 24- 31 schwer		
5 <input type="checkbox"/> 32- 40 extrem		
6 <input type="checkbox"/> mixed		
999 <input type="checkbox"/> keine Angabe		

5 Kontrollgruppe

KG_n	Anzahl Pat. in der KG _____	MOCI-R M (SD) _____
KG_n_female	Anzahl Frauen in der KG _____	OCI-R M (SD) _____
KG_perc_female	% Frauen _____	PI M (SD) _____
KG_M_age	Alter (Mittelwert) _____ *	
KG_SD_age	Alter (SD) _____ *	
KG_range_age	Altersrange _____ *	
KG_Händigkeit	1 <input type="checkbox"/> alle Rechtshänder 2 <input type="checkbox"/> alle Linkshänder	3 <input type="checkbox"/> mixed 999 <input type="checkbox"/> k.A. *
KG_perc_Rechtshänder	Wieviel % der KG waren Rechtshänder?	_____ % *
KG_n_komorbid	Anzahl komorbider Patienten in der KG	_____ *
KG_perc_komorbid	Wieviel % der KG waren komorbid?	_____ % *
KG_Art_komorbid	1 <input type="checkbox"/> Depression 2 <input type="checkbox"/> Angst 3 <input type="checkbox"/> Substanz 6 <input type="checkbox"/> mehrere 7 <input type="checkbox"/> Angst + Depression	4 <input type="checkbox"/> andere 5 <input type="checkbox"/> keine * 999 <input type="checkbox"/> keine Angaben
KG_n_medi	Anzahl medizierter Patienten in der KG	_____ *
KG_perc_medi	Wieviel % der KG waren mediziert?	_____ % *
KG_Art_medi	Welche Medikamente erhielt die KG?	_____
KG_IQ	IQ der KG, M(SD) _____ *	
KG_edu_ys	Bildungsjahre der KG, M(SD) _____ *	
KG_setting	1 <input type="checkbox"/> ambulant 2 <input type="checkbox"/> stationär 3 <input type="checkbox"/> mixed	999 <input type="checkbox"/> k.A. *
KG_Diagnose	Wie lautet die Diagnose in der KG?	_____
KG_M_Dauer_Jahre	Mittlere Krankheitsdauer	_____ *
KG_SD_Dauer_Jahre	SD der Krankheitsdauer	_____ *
KG_M_onset	Mittleres Ersterkrankungsalter	_____ *
KG_SD_onset	SD des Ersterkrankungsalters	_____ *
KG_M_Y_BOCS	Mittelwert Y BOCS, Cutoff 16	_____ *
KG_SD_Y_BOCS	SD der Y BOCS	_____ *
KG_Y_BOCS_severity	Schweregrad der Y BOCS *	
1 <input type="checkbox"/> 0- 7 leicht		
2 <input type="checkbox"/> 8- 15 mild		
3 <input type="checkbox"/> 16- 23 moderat		
4 <input type="checkbox"/> 24- 31 schwer		
5 <input type="checkbox"/> 32- 40 extrem		
6 <input type="checkbox"/> mixed		
999 <input type="checkbox"/> keine Angabe		

6 TESTS

Domäne	Neuropsychologische Domäne	1 <input type="checkbox"/> Aufmerks. 2 <input type="checkbox"/> Gedächtnis 3 <input type="checkbox"/> Exekutivfkt 4 <input type="checkbox"/> Aufm + Ged 5 <input type="checkbox"/> Aufm + Exek 6 <input type="checkbox"/> Gedä + Exek 7 <input type="checkbox"/> alle drei Domänen		
Do_Aufmerk	Aufmerksamkeit als Domäne	0 <input type="checkbox"/> keine Aufmerksamkeit	1 <input type="checkbox"/> Aufmerksamkeit	
Do_Gedäch	Gedächtnis als Domäne	0 <input type="checkbox"/> kein Gedächtnis	1 <input type="checkbox"/> Gedächtnis	
Do_Exekut	Exekutivfunktion als Domäne	0 <input type="checkbox"/> keine Exekutivfunktion	1 <input type="checkbox"/> Exekutivfkt.	
test_sex_diff	Gab es Geschlechterunterschiede bzgl. Testleistung?	0 <input type="checkbox"/> nein	1 <input type="checkbox"/> ja	999 <input type="checkbox"/> keine Angabe
test_IQ_diff	Gab es Gruppenunterschiede in der Intelligenz?	0 <input type="checkbox"/> nein	1 <input type="checkbox"/> ja	999 <input type="checkbox"/> keine Angabe
test_edu_ys_diff	Gab es Gruppenunterschiede in der Bildung?	0 <input type="checkbox"/> nein	1 <input type="checkbox"/> ja	999 <input type="checkbox"/> keine Angabe
test_komorb_diff	Gab es Gruppenunterschiede bzgl. Komorbidität?	0 <input type="checkbox"/> nein	1 <input type="checkbox"/> ja	999 <input type="checkbox"/> keine Angabe
test_medi_doff	Gab es Gruppenunterschiede bzgl. Medikation?	0 <input type="checkbox"/> nein	1 <input type="checkbox"/> ja	999 <input type="checkbox"/> keine Angabe
andere_Match_diff	Gruppenunterschiede in anderen Matchingvariablen?			

7 ANMERKUNGEN

author_contact	Kontakt mit dem Autor gehabt?	0 <input type="checkbox"/> nein	1 <input type="checkbox"/> ja	
info_erhalten	gewünschte Information vom Autor erhalten?	0 <input type="checkbox"/> nein	1 <input type="checkbox"/> ja	2 <input type="checkbox"/> nicht zutreffend
Power	Wurde Powerberechnung durchgeführt?	0 <input type="checkbox"/> nein	1 <input type="checkbox"/> ja	999 <input type="checkbox"/> keine Angabe
Duplicate	War die Studie doppelt publiziert?	0 <input type="checkbox"/> nein	1 <input type="checkbox"/> ja	999 <input type="checkbox"/> keine Angabe

Anhang 2 Kodierung neuropsychologischer Testergebnisse

Kodierschema für neuropsychologische Variablen. Metaanalyse OCD und Subgruppen.

Aufmerksamkeit	
Orientierung, Vigilanz und geteilte Aufmerksamkeit	
TAP-Alertness	4
CPT-I	4
KT (Halter)	4
Dual task	4
Räumliche Aufmerksamkeit	
WMS-R Blockspanne	4
CANTAB spatial span	4
Corsi Block Tapping	4
n-Back	4
Verzögerte Reaktionsaufgabe	4
Testparadigma für Arbeitsgedächtnis	4
WAIS-R Zahlen-Symbol-Test	4
TMT-A	4
Verbale Aufmerksamkeit	
WAIS-R Zahlenspanne	5
WMS-R Mentale Kontrolle	5
Testparadigma für Arbeitsgedächtnis	5
Filterung, Erwartung, Bias der Aufmerksamkeit	
Dichotisches Hören	5
Spatial cueing Paradigma	5
Emotionaler Stroop-Test	5
Dot-probe-Paradigma	5

Gedächtnis	
Semantisches Gedächtnis	
WAIS-R	6
General knowledge questions	6
Episodisches allgemeines Gedächtnis	
WMS-R general memory	6
LGT- 3 general memory	6
Batterie 144 total	6
Episodisches verbales Gedächtnis	
CVLT	6
TAVEC	7
HVLT	7
DR	7
CLPT	7
Wortliste am PC	7
Dichotisches Hören	7
Paar-Assoziationslernen	7
LGT- 3	7
Batterie 144	7
WMS-R Logical Memory (LM)	7
PWMT Verbalteil	8
RAVLT	8
Episodisches nonverbales visuell-räumliches Gedächtnis	
BVMT-R	8
LGT- 3	8
Batterie 144	8
AMIPB Subtest figure recall	8
WMS-R Visuelle Reproduktion I und II	8
Cube test	8
Card position test	8
Recurring figures test	8
RCFT	9
Handlungsgedächtnis und Reality Monitoring	
Handlungssatzparadigma und Reality Monitoring-Aufgabe	10
Aufgabe zum Quellengedächtnis	10
Personal belonging task	10
Phone call reminder task	10
Gedächtnisbias	
Erinnern berührter Objekte	10
Hindsight paradigma	10
Noise rating task	10

Exekutivfunktionen	
Kognitive Inhibition	
Go/No-Go	11
CANTAB affective Go/No-Go AGN	11
CPT-II	11
Farbe-Wort-Stroop-Test	11
Neg. Priming mit Stroop	12
Neg. Priming	12
Konzeptwechsel/Regellernen	
WCST	13
CANTAB intra-extra-dimensional Set shift IED	13
CANTAB SWM	13
OAT	13
TMT-B	13
Entscheidungsverhalten	
IGT	14
CANTAB CGT	14
Visuell-räumliches Planungsvermögen	
Tower of London TOL	14
CANTAB Stockings of Cambridge SOC	15
HVOT	15
Self-ordered Pointing task	15
Kognitive Flexibilität und mentale Flüssigkeit	
Letter fluency	15
COWA	15
Regensburger Wortflüssigkeitstest	15
Category fluency	15
Animal naming test	15
Design fluency test	15
WAIS-R Zahlen-Symbol-Test	15
Kategorisierung	
Object sorting task	16
Modified object sorting task	16
Personal belonging task	16

Name des Tests	Ergebnisvariable	EG M (SD)	KG M (SD)	F	t	p	df	Andere Statistik
Verzögerte Reaktionsaufgabe								
Testparadigma für Arbeitsgedächtnis	Räumlich rehearsal Räumlich supprimiert							
WAIS-R Zahlen- Symbol-Test								
TMT-A	Response time in sec							
sonstige								
Verbale Aufmerksamkeit								
WAIS-R Zahlenspanne	Digit span fwd Digit span bckwd Digit span total							
WMS-R Mentale Kontrolle	Attention							
Testparadigma für Arbeitsgedächtnis	Verbal rehearsal Verbal supprimiert							
Filterung, Erwartung, Bias der Aufmerksamkeit								
Dichotisches Hören								
Spatial cueing Paradigma	RT							
Emotionaler Stroop-Test	Interferenz Pos val OCD interference Neg val OCD interference Pos val non-OCD interference Neg val non-OCD interference Color-word- interference							

Name des Tests	Ergebnisvariable	EG M (SD)	KG M (SD)	F	t	p	df	Andere Statistik
Dot-probe-Paradigma	Detection latencies (RT) Number of errors total Number of errors for any word type Vigilance index Orienting Disengaging Mean threat interference effect (IQ-adjusted)							

Anmerkungen: _____

Kodierschema für neuropsychologische Variablen. Metaanalyse OCD und Subgruppen.

Gedächtnis

Study ID: _____ **Rater:** Leopold Oehmann **Datum:** _____

EG N = _____	1 Wasch- zwang	2 Kontroll- zwang	3 aggress Gedanken	4 Ordnungs- zwang	5 Hoarding	6 wenig Einsicht	7 Early onset
KG N = _____	1 gesund	2 Kontroll- zwang			3 ohne Hoarding	4 viel Einsicht	5 Late onset

Name des Tests	Ergebnisvariable	EG M (SD)	KG M (SD)	F	t	p	df	Andere Statistik
Semantisches Gedächtnis								
WAIS-R	Allg. Wissen – information score Gemeinsamkeiten – similarities score							
General knowledge questions	Recall Recognition (multiple choice)							
Episodisches allgemeines Gedächtnis								
WMS-R general memory								
LGT- 3 general memory								
Batterie 144	total							

Name des Tests	Ergebnisvariable	EG M (SD)	KG M (SD)	F	t	p	df	Andere Statistik
Episodisches verbales Gedächtnis								
CVLT	Trial 1 Trial 5 Trials 1- 5 Short delay recall Short term free recall Short term cued recall Long delay recall delayed free recall delayed cued recall % recall Recognition Recognition discriminability Semantic clustering Serial clustering Retention over delay: trial 5 to SD Retention over delay: SD to LD							
TAVEC	Trial 1 recall Trial 5 recall Trial 1- 5 total correct Short delayed recall 1st list Long delayed recall 1st list Perseverations Intrusions Semantic clustering Serial clustering							
HVLT	Immediate recall Delayed recall							

Name des Tests	Ergebnisvariable	EG M (SD)	KG M (SD)	F	t	p	df	Andere Statistik
DR	Delayed recall							
CLPT	recall							
Wortliste am PC	proportions correctly recalled recognition hits recognition false alarms RT for hits RT for false alarms D'							
Dichotisches Hören								
Paar-Assoziationslernen	Recall recognition							
LGT- 3	Verbal memory							
Batterie 144	Total verbal mémorisation remémorisation oubli							
WMS-R Logical Memory (LM)	LM I total score LM I immediate recall Mean number of items recalled immediate LM II total score LM late recall Mean number of items recalled delay Verbal memory							

Name des Tests	Ergebnisvariable	EG M (SD)	KG M (SD)	F	t	p	df	Andere Statistik
PWMT Verbalteil	Item retention (recall) immediate Item retention (recall) delay							
RAVLT	Total no of words recalled after the 5 trials = Immediate recall Total no of words recalled after 30 min delay = delayed recall Long term percentage retention recognition							
Episodisches nonverbales visuell-räumliches Gedächtnis								
BVMT-R	Total recall Delayed recall Recognition hits 30-min delayed recall Total score out of 36 Copy score out of 36							
LGT- 3	Nonverbal memory							
Batterie 144	Total visuel RIE visuel RDE visuel AL visuel RL visuel AP visuel RD visuel							
AMIPB Subtest figure recall	Immediate recall Delayed recall							
WMS-R Visuelle Reproduktion I und II	Visual reproduction I (immediate) Visual reproduction II (30 min delay)							
Cube test	CT-R CT-Y							

Name des Tests	Ergebnisvariable	EG M (SD)	KG M (SD)	F	t	p	df	Andere Statistik
Card position test								
Recurring figures test	Overall score No of false positive answers							
RCFT	Organization score copy Organization score immediate recall Accuracy score copy Accuracy score immediate recall Percent recall Copy total score Delayed recall score Accuracy in delayed recall % recall – immediate to delayed Copy organization Copy Copy minus immed. Recall Copy minus delayed recall Immediate recall Delayed recall Recognition BQSS scores: Copy presence and accuracy Immediate presence and accuracy Delayed presence and accuracy Organization Organizational score CFT 3 CFT 30							

Name des Tests	Ergebnisvariable	EG M (SD)	KG M (SD)	F	t	p	df	Andere Statistik
Handlungsgedächtnis und Reality monitoring								
(Erweitertes) Handlungssatz- paradigma und Reality Monitoring Aufgabe	Free recall of motorically encoded actions Hits FA (false alarms) Pr (discrimination index) = Hits minus FA Br (response bias) = FA/1 – (hits – FA) Error type A Error type B Proportion of correct RM identifications Proportion of correct activity identifications Item recognition as old or new							
Aufgabe zum Quellengedächtnis	Correct responses							
Personal Belonging Task	Response to cue: yes Response to cue: no							
Phone call reminder task	Predicted absolute difference from 30 min Actual absolute difference from 30 min Mean checks 0- 10 min Mean checks 11- 20 min Mean checks 21- 30 min							
Gedächtnisbias								
Erinnern berührter Objekte	Free object recall dirty Free object recall clean							
Hindsight paradigma	Δz							
Noise rating task	Rating of noise volume for contamination sentences Recognition scores							

Eodierschema für neuropsychologische Variablen. Metaanalyse OCD und Subgruppen.

ExekutivfunktionenStudy ID: _____ Rater: Leopold Oehmann Datum _____

EG N = _____	1 Wasch- zwang	2 Kontroll- zwang	3 aggress Gedanken	4 Ordnungs- zwang	5 Hoarding	6 wenig Einsicht	7 Early onset
KG N = _____	1 gesund	2 Kontroll- zwang			3 ohne Hoarding	4 viel Einsicht	5 Late onset

Name des Tests	Ergebnisvariable	EG M (SD)	KG M (SD)	F	t	p	df	Andere Statistik
Kognitive Inhibition								
Go/No-Go	% of omission errors for GO stimuli % of commission errors for No-GO-stimuli							
CANTAB affective Go/No-Go AGN	Mean correct latency positive Mean correct latency negative Total omissions pos Total omissions neg							
CPT-II	HRT hit reaction time Response style D prime Perseverations % omission errors % commission errors							
Farbe-Wort-Stroop-Test	Interferenzscore Time to complete the reporting/color-word-condition Number of incorrect color names							

Name des Tests	Ergebnisvariable	EG M (SD)	KG M (SD)	F	t	p	df	Andere Statistik
Neg. Priming mit Stroop	Interferenzeffekt Errors Color-naming-latency Mean interference score NP negative priming effect RT 100 msec presentation speed RT 50 msec presentation speed RT 250 msec presentation speed							
Neg. Priming	NP location + identity NP identity NP location Average NP index Priming RT 100 msec Priming RT 250 msec Priming RT 350 msec Control RT 100 msec Control RT 250 msec Control RT 350 msec Mean errors for priming distractor 100 msec Mean errors for priming distractor 250 msec Mean errors for priming distractor 350 msec							

Name des Tests	Ergebnisvariable	EG M (SD)	KG M (SD)	F	t	p	df	Andere Statistik
Konzeptwechsel/Regellernen								
WCST	NC = SN = number of categories completed PE = perseverative errors TE = total errors LS = loss of set CL = % conceptual responses FC = NT = number of trials to complete 1 st category Number of correct responses Milner E = Milner-type errors Learning to learn NP = Non-perseverative errors Total responses							
CANTAB intra-extra-dimensional Set shift IED	IED stages completed Total errors (adjusted) Trial level							
CANTAB SWM	Between-search errors Strategy score							
OAT	Total trials Total no of errors No of perseverative errors							
TMT-B	RT Subtracted score (B minus A) in sec							

Name des Tests	Ergebnisvariable	EG M (SD)	KG M (SD)	F	t	p	df	Andere Statistik
Entscheidungsverhalten								
IGT	No of cards taken from Deck A No of cards taken from Deck B No of cards taken from Deck C No of cards taken from Deck D IGT difference $A + B - (C + D)$ Decks A + B Decks C + D Net score $(C + D) - (A + B)$							
CANTAB CGT								
Visuell-räumliches Planungsvermögen								
Tower of London TOL	No of moves to complete the task Score/12 1 st niveau Score/12 2 nd niveau RT 1 st niveau RT 2 nd niveau Total no of problems solved in minimum no of moves Mean time taken for 3 move problems Mean time taken for 4 move problems Mean time taken for 5 move problems No of 3 move problems solved with minimum no of moves No of 4 move problems solved with minimum no of moves No of 5 move problems solved with minimum no of moves							

Name des Tests	Ergebnisvariable	EG M (SD)	KG M (SD)	F	t	p	df	Andere Statistik
Kategorisierung								
Object sorting task	No of piles, personal objects No of piles, non-personal objects No of piles, personal index cards No of piles, non-personal index cards Time (sec) , personal objects Time (sec), non-personal objects Time (sec), personal index cards Time (sec), non-personal index cards							
Modified object sorting task	No of piles Time for sorting							
Personal belonging task	No of piles time							

Anmerkungen: _____

Anhang 3 Kodiermanual

Kodiermanual

1 Kodierschema für deskriptive Daten

1.1 Methode: Interne Anmerkungen, die nicht publiziert werden sollen, können hier gemacht werden.

1.2 Studienqualität: Hier ist die kommentierte Skala zur Beurteilung der Studienqualität anzuwenden.

1.3 Teilnehmer: Als EG wird diejenige Gruppe kodiert, die in der Studie als erste genannt wird. Ausnahme: Waschzwang wird immer als EG kodiert.

1.4 Experimentalgruppe und 1.5 Kontrollgruppe: Wenn in der Primärstudie nur demographische Werte für die gesamte Zwangsstichprobe, nicht aber für die einzelnen Symptomdimensionen oder Untergruppen angegeben sind, dann sind diese Werte sowohl bei EG als auch bei KG einzutragen und mit einem Sternchen zu markieren.

Wenn Durchschnittswerte zu berechnen sind (weil z. B. in einer Primärstudie das Alter für medizierte und unmedizierte Patienten getrennt angegeben wurde): Mittelwert $M = a + b / 2$ berechnen.

Die gepoolte Standardabweichung (z. B. beim Alter) lässt sich nach folgender Formel ermitteln, wenn EG und KG ein gleich großes n haben:

$$s_{pooled} = \sqrt{\frac{s_{EG}^2 + s_{KG}^2}{2}}$$

Sind EG und KG unterschiedlich groß, dann wird die gepoolte Standardabweichung nach folgender Formel berechnet (Bühner & Ziegler, 2009, Statistik für Psychologen und Sozialwissenschaftler. München, Pearson Studium. S. 254):

$$s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Dabei ist n_1 die Stichprobengröße der EG, n_2 die Stichprobengröße der KG und s_1 bzw. s_2 die Standardabweichung der EG bzw. KG.

1.6 Tests: Hier sind nur die neuropsychologischen Domänen und Moderatoreffekte einzutragen. Die Ergebnisse der neuropsychologischen Tests werden auf einem separaten Blatt kodiert.

1.7 Anmerkungen, die für die Publikation relevant sind, können hier gemacht werden.

2 Kodierschema für neuropsychologische Testergebnisse

Auf Blatt 1, dem Inhaltsverzeichnis, werden für eine erste Übersicht diejenigen Tests angekreuzt, die in der Primärstudie durchgeführt wurden.

Auf den nachfolgenden Blättern werden die einzelnen Testergebnisse mit ihren Statistiken eingetragen.

Anhang 4 Skala zur Beurteilung der Studienqualität

Die Skala wurde adaptiert nach Öst (2008) und den Kriterien der Cochrane-Collaboration (2011)

1. Einhaltung von Ethikrichtlinien (nach Cochrane)

0 = *mangelhaft*: Einhaltung von Ethikrichtlinien wurde nicht berichtet.

1 = *ausreichend*: Teilnehmer wurden informiert und erklärten schriftlich ihr Einverständnis.

2 = *gut*: Teilnehmer wurden informiert und erklärten schriftlich ihr Einverständnis, und es wurde eine Genehmigung der Ethikkommission berichtet.

2. Klarheit der Stichprobenbeschreibung

0 = *mangelhaft*: Vage Beschreibung der Stichprobe (z. B. wenn nur angegeben wurde, ob Patienten mit der Störung diagnostiziert wurden).

1 = *ausreichend*: Ausreichende Beschreibung der Stichprobe (z. B. wenn Ein-/Ausschlusskriterien, demographische Daten etc. angegeben wurden).

2 = *gut*: Gute Beschreibung der Stichprobe (z. B. Ein-/Ausschlusskriterien, demographische Daten, Bildungsniveau und Prävalenz komorbider Störungen)

3. Erfassung des Schweregrades der Störung

0 = *mangelhaft*: Schweregrad wurde nicht berichtet.

1 = *ausreichend*: Der Schweregrad der Zwangsstörung wurde eingeschätzt, ohne dass ein valides psychometrisches Verfahren eingesetzt wurde (z. B. anhand von intuitiver klinischer Expertenmeinung)

2 = *gut*: Der Schweregrad der Zwangsstörung wurde anhand eines validen psychometrischen Verfahrens (z. B. Y-BOCS) beurteilt.

4. Repräsentativität der Stichprobe in Bezug auf Komorbidität

0 = *mangelhaft*: Es wurde nicht berichtet, ob komorbide Störungen vorlagen oder komorbide Patienten wurden ausgeschlossen, so dass die Stichprobe hoch selektiv war

1 = *ausreichend*: Komorbide Patienten wurden eingeschlossen und die komorbiden Störungen wurden beschrieben

2 = *gut*: Komorbide Patienten wurden eingeschlossen und der Einfluss der Komorbidität analysiert

5. Reliabilität der Diagnose

0 = *mangelhaft*: Der diagnostische Prozess wurde nicht berichtet oder nicht mit einem strukturierten Interview erstellt.

1 = *ausreichend*: Die Diagnose wurde mit einem strukturierten Interview erstellt.

2 = *gut*: Die Diagnose wurde mit einem strukturierten Interview erstellt und es wurde eine adäquate Interrater-Reliabilität berichtet (z. B. Kappa-Koeffizient).

6. Spezifität/Breite der neuropsychologischen Variablen

0 = *mangelhaft*: Sehr breite, wenig spezifische neuropsychologische Variablen (z. B. allgemeine kognitive Leistungsfähigkeit wie Gesamt-IQ-Wert).

1 = *ausreichend*: Einigermaßen spezifische Variablen, die das neuropsychologische Konstrukt benennen. Das Konstrukt wurde von mehreren Forschergruppen mit ein und demselben Test erfasst (z. B. nonverbales Gedächtnis mit dem RCFT = Expertenkonsens). Mehrere Konstrukte können in einem Test konfundiert sein (z. B. motorische Geschwindigkeit und Aufmerksamkeit im TMT-A).

2 = *gut*: Spezifische neuropsychologische Testergebnisse wie z. B. WCST, die keine Annahmen über die (teils konfundierten) zugrunde liegenden neuropsychologischen Konstrukte treffen.

7. Reliabilität und Validität der neuropsychologischen Testverfahren

0 = *mangelhaft*: Messmethoden haben unbekannt psychometrische Eigenschaften oder Eigenschaften, die die derzeitigen Standards nicht erfüllen.

1 = *ausreichend*: Einige, aber nicht alle Messmethoden haben bekannte oder adäquate psychometrische Eigenschaften.

2 = *gut*: Alle Messmethoden haben gute psychometrische Eigenschaften. Die Testverfahren sind die besten, die für die Zwecke der Autoren verfügbar sind.

8. Anzahl der Testleiter

0 = *mangelhaft*: Keine Information über die Anzahl der Testleiter oder nur ein Testleiter, d. h. komplette Konfundierung von Testung und Testleiter.

1 = *ausreichend*: Mindestens zwei Testleiter, aber der Effekt des Testleiters auf das neuropsychologische Testergebnis ist nicht analysiert.

2 = *gut*: Drei oder mehr Testleiter und der Effekt des Testleiters auf das neuropsychologische Testergebnis wurde analysiert.

9. Training der Testleiter

0 = *mangelhaft*: Keine Informationen über Training der Testleiter.

1 = *ausreichend*: Es wird berichtet, dass die Testleiter vor der Testdurchführung trainiert wurden.

2 = *gut*: Es wird berichtet, dass die Testleiter vor der Durchführung trainiert wurden, und das Training wurde näher spezifiziert (z. B. Art und Dauer).

10. Kontrolle von Medikation

0 = *mangelhaft*: Keine Angaben darüber, ob die Probanden psychopharmakologische oder stimulierende Medikation erhielten oder keine Angaben über Art, Menge und Dauer der Medikation

1 = *ausreichend*: Art und Menge/Dauer der psychopharmakologischen oder stimulierenden Medikation wurden angegeben.

2 = *gut*: Ein möglicher Einfluss der psychopharmakologischen oder stimulierenden Medikation auf das neuropsychologische Testergebnis wurde analysiert.

11. Statistische Methoden, Datenpräsentation

0 = *mangelhaft*: Inadäquate statistische Methoden wurden eingesetzt und/oder Daten nicht vollständig dargestellt.

1 = *ausreichend*: Adäquate statistische Methoden wurden eingesetzt, aber die Daten nicht vollständig dargestellt.

2 = *gut*: Adäquate statistische Methoden wurden eingesetzt und die Daten mit *M* und *SD* dargestellt.

Andere Quellen von Datenbias (z. B. wurden Ergebnisse selektiv berichtet: Beschreibe, wie die Möglichkeit selektiver Ergebnisberichte untersucht wurde und was dabei herauskam).

Anhang 5 Interraterreliabilität

Tabelle 12: *Interraterreliabilitäten der Items der Skala zur Beurteilung der Studienqualität*

		<i>ICC</i>	95 %- Konfidenz- intervall	<i>F</i>	<i>df</i>	<i>p</i>
1	Einhaltung von Ethikrichtlinien	.939	.846 - .976	15.78	19,19	.000
2	Klarheit der Stichprobenbeschreibung	.743	.362 - .897	4.339	19,19	.001
3	Erfassung des Schweregrades der Störung	Alle Studien setzten die Y-BOCS ein, daher war aufgrund mangelnder Varianz kein ICC bestimmbar				
4	Repräsentativität der Stichprobe in Bezug auf Komorbidität	.759	.407 - .904	4.529	19,19	.001
5	Reliabilität der Diagnose	.877	.696 - .951	8.158	19,19	.000
6	Spezifität der neuropsychologischen Variablen	.376	-.382 - .738	1.696	19,19	.129
7	Validität der neuropsychologischen Tests	.762	.368 - .908	5.053	19,19	.000
8	Anzahl der Testleiter	.771	.439 - .908	4.556	19,19	.001
9	Training der Testleiter	.747	.378 - .899	4.294	19,19	.001
10	Kontrolle der Medikation	.840	.586 - .937	7.073	19,19	.000
11	Einhaltung von Ethikrichtlinien	.747	.378 - .899	4.294	19,19	.001

Erklärung gemäß § 8 Abs. 1 Buchst. b) und c) der Promotionsordnung

Erklärung gemäß § 8 Abs. 1 Buchst. b) der Promotionsordnung der Fakultät für Verhaltens- und Empirische Kulturwissenschaften der Ruprecht-Karls-Universität Heidelberg

Ich erkläre, dass ich die vorgelegte Dissertation selbstständig angefertigt, nur die angegebenen Hilfsmittel benutzt und die Zitate gekennzeichnet habe.

Erklärung gemäß § 8 Abs. 1 Buchst. c) der Promotionsordnung der Fakultät für Verhaltens- und Empirische Kulturwissenschaften der Ruprecht-Karls-Universität Heidelberg

Ich erkläre, dass ich die vorgelegte Dissertation in dieser oder einer anderen Form nicht anderweitig als Prüfungsarbeit verwendet oder einer anderen Fakultät als Dissertation vorgelegt habe.

Rebecca Leopold

Heidelberg, den