



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Eine theoretische und empirische Auseinandersetzung
mit den Netzwerkanalysen in der klinischen Psychologie
im Rahmen einer Re-Evaluation des Projektes „fit2work“

verfasst von / submitted by

Simon Edel BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2019

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 066 840

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium Psychologie UG2002

Betreut von / Supervisor:

Mag. Dr. Reinhold Jagsch

Inhaltsverzeichnis

Abstract	4
Einleitung	6
Theoretischer Hintergrund	7
Klassifikationssysteme als Schnittstelle zwischen Ontologie und Methodologie	8
Latente-Variablen-Modelle, ihre epistemologischen Grundlagen und ontologischen Implikationen	14
Netzwerkmodellierungen als alternative methodische und theoretisch-konzeptionelle Ansätze	18
Regularisierung	20
Analysen der Stabilität und der Akkuratheit der Schätzungen	23
Netzwerkanalysen als Kombination der Stochastik und der Graphentheorie ..	24
Die Zentralitätsmaße	25
Die Konnektivität	27
Der Impact einer Variable	29
Methodik	30
Untersuchungsgegenstand	30
Untersuchungsinstrumente	31
Fragestellungen und Hypothesen	32
Untersuchungsdesign und Stichprobenbeschreibung	33
Statistische Verfahren	36
Ergebnisdarstellung	37
Symptomnetzwerke vor und nach der Behandlung	37
Zentralitäten	40
Impact	45
Ergebnisse der Testung der ersten Hypothese	47
Symptomnetzwerke der Responder und Non-Responder	47

Ergebnisse der Testung der zweiten Hypothese	49
Diskussion	50
Korrelative Zusammenhänge – Abbildungen von Kausalstrukturen?	51
Zentralitätswerte – Indikatoren für Ansatzpunkte von Interventionen?	53
Zusammenhänge zwischen der Struktur und der Dynamik von Netzwerken	57
Das Konzept des Impact	59
Fazit	61
Literaturverzeichnis	64
Abbildungsverzeichnis	71
Tabellenverzeichnis	72
Anhang	73

Abstract

In der vorliegenden Arbeit werden die Netzwerkanalysen als neue methodische und konzeptionelle Alternative zu gängigen Praktiken der Modellierung psychologischer Konstrukte vorgestellt. Dies geschieht vor allem vor dem Hintergrund einer Diskussion der latenten-Variablen-Modelle und einer ihrer Anwendungen: der klassifikatorischen Diagnostik. Als Alternative bieten die Netzwerkanalysen Zugriff auf methodisch-analytische Zugänge, insbesondere der Graphentheorie, die mit den bisher in der Psychologie gängigen Methoden kaum nutzbar sind. Anhand einer Stichprobe von $N = 664$ TeilnehmerInnen des Projektes „fit2work“ wurden die Symptomnetzwerke der Items des BSI-53 geschätzt. Die Zuverlässigkeit der Schätzungen wurde evaluiert. Auch wurden mehrere netzwerktheoretische Inhalte teils explorativ, teils konfirmatorisch untersucht. Ein besonderer Fokus lag hierbei auf dem mathematischen Konstrukt der Konnektivität sowie dessen Verknüpfung zu inhaltlich-theoretischen Konzepten wie der Selbst-Aufrechterhaltung von psychischen Störungen oder der definitorischen Eingrenzung psychischer Krankheit. Die Schätzungen der Netzwerke ergaben komplexe Strukturen und Zusammenhänge der einzelnen Faktoren psychischer Belastung in der vorliegenden Stichprobe. Weiterhin zeigte sich auf Itemniveau ein Zusammenhang verschiedener struktureller Aspekte der geschätzten Netzwerke. Die Untersuchungen der Konnektivität ergab keine signifikanten Veränderungen dieser im Zeitraum der Teilnahme an „fit2work“ sowie keine Unterschiede diesbezüglich bei den Subgruppen der Responder und der Non-Responder. Die vorliegenden Ergebnisse werden hinsichtlich ihrer methodischen und epistemologischen Implikationen diskutiert. Hierbei wird der Fokus einerseits darauf gelegt, welche Schlussfolgerungen hinsichtlich möglicher Interventionen und Präventivmaßnahmen aus querschnittlichen Netzwerkanalysen ableitbar sind, andererseits darauf, wie das Konzept der Konnektivität erforscht wird und wie es mit theoretisch-inhaltlichen Konstrukten verknüpft werden kann. Abschließend wird resümiert, welche wissenschaftstheoretische Rolle die Netzwerkanalysen und -theorien ausfüllen können und sollen.

In this thesis, network analysis is introduced as new methodical and conceptual alternative to the common practices of modelling psychological constructs. This happens primarily against the background of a discussion of latent variable modelling and one of its applications: classificatory diagnostics. Being an alternative, network analysis offers usage of analytical methods, especially from graph theory, that are hardly accessible by statistical analysis

common in psychology. A sample of $N = 664$ participants of the project “fit2work” was used to estimate symptom networks of the BSI-53. The stability of these estimations was assessed. Also, several concepts of network theory were investigated partly explorative, partly confirmatory. Especially the mathematical concept of network-connectivity and its matching onto several theoretical constructs such as self-perpetuation of mental disorders or the definitory delimitation of mental illness is focused on. The estimation of the networks resulted in complex structures and patterns of correlation of individual factors of psychological burden in the sample. On item-level, some aggregation of structural aspects of networks presented itself. The examinations of connectivity found no significant changes in the period of participation in “fit2work” and no differences between the subsamples of the responders and non-responders. The findings are discussed regarding their methodical and epistemological implications. On the one side, this discussion focuses on the possible conclusions regarding viable interventions and preventive measures based on cross-sectional analyses. On the other side, it focuses on how connectivity is currently investigated and how it can be linked with theoretical contents. Finally, some concluding remarks concerning network theory’s role in science are presented.

Einleitung

In den vergangenen Jahren ließ sich in vielen wissenschaftlichen Disziplinen eine Entwicklung mit teils gravierenden epistemologischen Implikationen beobachten. Gemeint ist eine zunehmende Abkehr von reduktionistischen Paradigmen und eine stärker werdende Orientierung an daten-gesteuerten Modellierungen komplexer Systeme (Barabasi, 2012). Eine Folge dieser Entwicklung besteht in der zunehmenden Verbreitung der netzwerkanalytischen Zugänge in verschiedenen wissenschaftlichen Disziplinen (Newman, 2010). Diese basiert im Wesentlichen auf der Einsicht, dass die gängigen reduktionistischen Paradigmen verschiedener wissenschaftlicher Disziplinen nicht länger ausreichen, um der Komplexität und teilweise auch Ambivalenz der jeweiligen Untersuchungsgegenstände gerecht zu werden. Als beispielhaft hierfür können Entwicklungen in der Erforschung genetischer Determinanten verschiedener Krebsarten in der Medizin (Barabasi, 2012; Pomerantz & Freedman, 2011), in der der Modellierung sozialer Interaktionssysteme in der Soziologie (Newman, 2010), aber auch in der Erforschung und Modellierung verschiedener psychologischer Konzepte, insbesondere aber der Psychiatrie und klinischen Psychologie (Borsboom & Cramer, 2013) angesehen werden.

Diese Arbeit soll einen Einblick in die Anwendung der netzwerkanalytischen Zugänge in der klinischen Psychologie, ihr theoretisches Fundament sowie ihre wissenschaftstheoretischen Implikationen geben. Hierfür sollen zunächst die Grundannahmen eines aktuell vorherrschenden Paradigmas, das der Modellierung latenter Variablen, diskutiert werden. Dieses geht, vor allem im Bereich der klinischen Psychologie, mit einigen theoretisch-konzeptionellen sowie praktischen Problemen einher. Insbesondere in der stringenten Umsetzung ihrer Grundkonzeptionen in den modernen Klassifikationssystemen haben die latenten-Variablen-Modelle nicht nur die Begriffssysteme der klinischen Psychologie und Psychiatrie in eine einheitliche Nomenklatur umgewandelt. Sie prägen auch über Prozesse der Reifikation den über die historischen Ansprüche der Klassifikationssysteme weit hinausgehenden Sprachgebrauch und hierüber die Art, wie über psychische Störungen gedacht wird und wie sie erforscht werden (Hyman, 2010; Widiger & Edmundson, 2011).

Die Netzwerkmodelle stellen eine Alternative zu den latenten-Variablen-Modellen dar, die einige der wesentlichen formulierten Kritikpunkte umgeht. Als solche soll sie vorgestellt werden und ihre Anwendung am Untersuchungsgegenstand der Evaluation des Projektes „fit2work“ demonstriert werden. Weiterhin soll die Umsetzung eines der Grundkonzepte der Netzwerkanalysen, der Konnektivität, im psychologischen Bereich untersucht werden.

Hierdurch ergeben sich einige epistemologische Fragen, die theoretisch und anhand der Befunde dieser Arbeit diskutiert werden sollen.

Theoretischer Hintergrund

Die Anwendung netzwerkanalytischer Untersuchungen in der klinischen Psychologie und der Psychiatrie ist eine vergleichsweise junge Entwicklung (Borsboom, 2017; Hevey, 2018). Anhand des ersten wissenschaftlichen Artikels aus dieser Disziplin (Cramer, Waldorp, van der Maas, & Borsboom, 2010) wird ersichtlich, dass diese Entwicklung nicht zuletzt vor dem Hintergrund der Kritik an der zu diesem Zeitpunkt vorherrschenden Art der Modellierung psychopathologischer Variablen zu verstehen ist. Konkret bezog und bezieht sich diese Kritik auf die Modellierung psychopathologischer Konstrukte als latente Variablen, die auf teilweise diskussionswürdigen Grundannahmen beruht (z.B. Borsboom & Cramer, 2013). Die Kritikpunkte der Vertreter des netzwerkanalytischen Ansatzes an dieser Art der Darstellung bezogen sich zunächst in erster Linie auf psychometrische und konzeptionelle Probleme, die sich durch die Operationalisierung latenter Konstrukte anhand manifester Variablen ergeben. Allerdings liegen diesen Kritikpunkten zum Teil auch epistemologische und ontologische Fragen zugrunde, die der klinischen Psychologie und der Psychiatrie bereits seit Längerem bekannt sind (Borsboom, Mellenbergh, & van Heerden 2004; Hyman, 2010). Denn die Implikationen der netzwerkanalytischen Modellierungen beziehen sich nicht ausschließlich auf technische Aspekte der Modellierung und Operationalisierung latenter Konstrukte, sondern werfen auch davon kaum trennbare Fragen über die grundlegenden Eigenschaften psychischer Störungen auf. Die Untersuchung der Zusammenhänge zwischen diesen beiden Faktoren, also dem grundlagentheoretischen Fundament der Konzeptionen psychischer Störungen einerseits und ihrer konkreten Manifestation in Form von technisch-methodischen Ideen andererseits, kann jedoch nicht sinnvoll ohne Berücksichtigung ihres wissenschaftlichen und sozialgesellschaftlichen Kontexts stattfinden. Dieser Umstand ergibt sich daraus, dass die wissenschaftliche Psychologie eine lebendige Disziplin ist (Hayes, 1995), die in ihren Kontext eingebettet diesen widerspiegelt und ihn im Gegenzug formt. Hieraus ergeben sich zum einen Fragen nach den Faktoren, die die theoretischen Grundlagen der Modellierung psychischer Störungen maßgeblich beeinflussen, zum anderen Fragen, wie sie im Gegenzug Einfluss auf ebendiese Faktoren nehmen und vor allem nach dem Wechselspiel dieser Einflüsse und ihrer Bedeutung für den Anspruch der Modellierungen auf Legitimität. Auch diese Fragen sind der klinischen Psychologie und der Psychiatrie nicht unbekannt (Sabbarton-Leary, Bortolotti, & Broome, 2015). Die antipsychiatriische Bewegung der 1950er und folgenden Jahre kann zumindest als Beispiel dafür angesehen werden, wie

eine Kritik an methodischen Aspekten (hier in Form von mangelhaften Reliabilitäten der Diagnosen der gängigen Klassifikationssysteme) schnell zu einer generellen Infragestellung der grundsätzlichen Validität und Legitimität des Krankheitsverständnisses führen kann (Wittchen, 2011). Dem Autor dieser Arbeit erscheint es an dieser Stelle angebracht hervorzuheben, dass er sich einer Bewertung der antipsychiatrischen Bewegung entziehen möchte und hier lediglich beispielhaft aufgezeigt werden soll, dass die Verknüpfung von methodischer Kritik und Grundsatzfragen durchaus kein neues Phänomen ist.

Klassifikationssysteme als Schnittstelle zwischen Ontologie und Methodologie

Eine zentrale Rolle im Zusammenhang zwischen methodischer Operationalisierung und konzeptionellen Grundsätzen spielen die Klassifikationssysteme. Sie strukturieren die Welt der Psychopathologie (Hyman, 2010), indem sie diese mehr oder weniger wissenschaftlich fundiert kategorisieren und eine Nomenklatur definieren, auf deren Basis viele wesentliche Bestandteile der modernen klinischen Psychologie (wie zum Beispiel Forschung über verschiedene Sprach-, Kultur- und Zeiträume hinweg) erst ermöglicht werden (Widiger & Edmundson, 2011). Aus diesen beiden Aspekten, der Kategorisierung und der Definition einer Nomenklatur, ergeben sich jedoch auch Ansprüche auf Validität, Reliabilität und Legitimität. Diese sollten zwar nicht per se verworfen oder abgelehnt werden, müssen allerdings auch nicht unhinterfragt bleiben.

Eines der Kernanliegen eines der in der klinischen Psychologie gängigsten Klassifikationssysteme, dem DSM-5, besteht darin, eine gemeinsame Nomenklatur für den Umgang mit psychischen Störungen zu bieten (Widiger & Edmundson, 2011). Diese stellt eine wesentliche Voraussetzung für die systematische Erforschung psychopathologischer Erkrankungen dar. Auch ermöglicht erst ein einheitliches Begriffssystem einen sinnvollen praktischen Umgang mit den nur mittelbar zugänglichen Konstrukten der klinischen Psychologie. Dieses Kernanliegen spiegelt sich an verschiedenen Stellen in der historischen Entwicklung des DSM-Systems wider. So war das Bedürfnis nach einer standardisierten Nosologie ausschlaggebend für die Entwicklung des DSM-I (Aragona, 2015). Die zuvor vorherrschenden Systematiken psychischer Störungen basierten im Wesentlichen auf theoriegeleiteten Top-Down-Ansätzen (Sanders, 2011), weshalb es wenig verwunderlich erscheint, dass sich, entsprechend der Vielzahl an theoretischen Strömungen innerhalb der klinischen Psychologie und Psychiatrie, mehrere verschiedene Klassifikationssysteme und damit Operationalisierungen psychischer Störungen gebildet hatten (Aragona, 2015). Ein exemplarisches und breit rezipiertes Beispiel für die Vorteile einer einheitlichen Nomenklatur findet sich in den epidemiologischen Untersuchungen der Schizophrenie in den Vereinigten

Staaten von Amerika (USA) und Großbritannien in den 1970er Jahren (Hyman, 2010). So ergaben Schätzungen der Prävalenz der Schizophrenie in beiden Ländern eine fast doppelt so hohe Fallzahl in den USA, ein Umstand, der wie sich später herausstellen sollte, durch unterschiedliche Definitionsbereiche des Begriffes „Schizophrenie“ bedingt war (Hyman, 2010). Wurden in den USA auch Personen mit Krankheitsepisoden von kürzerer Dauer als sechs Monaten mit klassifiziert, wurde in Großbritannien dieser Wert als Cut-off verwendet, da hier die Schizophrenie als chronische Krankheit definiert war.

Ein maßgeblicher Einfluss auf die Abkehr von den theoriegeleiteten Ansätzen zur Klassifikation psychischer Störungen lässt sich in der Arbeit Emil Kraepelins verorten. So klassifizierte er psychische Störungen in erster Linie anhand ihres phänomenologischen Erscheinungsbildes, ein Vorgehen, das sich noch in den heutigen Klassifikationssystemen in Form eines prinzipiell deskriptiven Ansatzes widerspiegelt und das einen schulenübergreifenden Umgang mit klinisch-psychologischen Inhalten ermöglicht. Auch konzipierte Kraepelin psychische Störungen als klar voneinander abgrenzbare Entitäten, die einen sie definierenden Kern aufweisen (Maser & Cloninger, 1990, zitiert nach Widinger & Edmundson, 2011). Dieser Kern sollte gemäß Kraepelin objektiv definierbare pathologische Prozesse widerspiegeln, die sich in einer über verschiedene Personen hinweg identischen Symptomatik manifestieren (Aragona, 2015).

Dieser essentialistische Grundgedanke bestimmt bis heute die modernen Klassifikationssysteme maßgeblich. In ihnen werden psychische Störungen als latente Konstrukte definiert, die anhand ihrer wesentlichen und sie definierenden Kerne voneinander abgegrenzt und in eine Taxonomie eingeordnet werden. Diese Ordnung erfolgt nicht ausschließlich nach streng phänomenologischen oder empirischen Gesichtspunkten (so definiert sich beispielsweise die Kategorie der Persönlichkeitsstörungen nicht über die Homogenität oder zumindest Ähnlichkeit der Symptomatik, sondern über Faktoren wie die frühe Erstmanifestation oder die Persistenz der Symptome), sondern bezieht auch theoretische Konstrukte (wie etwa „Persönlichkeit“) mit ein (Widiger & Edmundson, 2011).

Darüber hinaus bringt die Modellierung psychischer Störungen als latente Konstrukte, die über manifeste Variablen zugänglich gemacht und operationalisiert werden, methodische und konzeptionelle Probleme mit sich. Auf konzeptioneller Ebene wird durch die Modellierung psychischer Störungen als latente Variablen in Kombination mit einer festgelegten diagnostischen Operationalisierung impliziert, dass a.) ein Bedeutungskern durch b.) angenommene kausale Mechanismen ein c.) klar umschriebenes Störungsbild hervorbringt.

Hier stellt sich zunächst die Frage, worin der Bedeutungskern oder die Essenz einer psychischen Störung besteht. Diese Frage scheint ontologischer Natur zu sein, allerdings wird bei genauerer Betrachtung klar, dass sie nur vor dem Hintergrund der jeweiligen Definition von Validität und damit eines erkenntnistheoretischen Gegenstandes beantwortbar ist. So ist sie unweigerlich daran gebunden, ob psychische Störungen als natürlich auftretende, objektiv nachvollziehbare „Realitäten“ verstanden werden (so wie beispielsweise physikalische Elemente als „natürliche“ Gegebenheiten verstanden werden, die sich über von menschlicher Erkenntnis unabhängigen Faktoren definieren) oder ob sie als künstliche „Hilfskonstrukte“, die nicht unabhängig vom jeweiligen Erkenntnisprozess und -interesse gedacht werden können, angesehen werden. An dieser Stelle ist anzumerken, dass mit diesen beiden Beispielen von Konzeptionen von Validität lediglich zwei Positionen einer bereits seit langem geführten und umfangreichen philosophischen Debatte dargestellt werden (Zachar, 2015). In Bezug auf die Klassifikationssysteme bleibt jedoch festzuhalten, dass spätestens seit dem DSM-III der Anspruch auf die zumindest approximative Abbildung „natürlicher“ Entitäten besteht (Zachar & Jablensky, 2015). Historisch betrachtet lässt sich dieser Anspruch auf die Arbeit von Robins und Guze (vor allem Robins & Guze, 1970) zurückführen. Der von ihnen vertretene Ansatz zur Validierung psychischer Störungen (hier am Beispiel der Schizophrenie) umschreibt klar eine schrittweise Annäherung an eine natürliche, also reale Entität, die mittels explizit formulierter Strategien wissenschaftlich zugänglich gemacht werden kann. So beschrieben Robins und Guze (1970) ein Fünf-Phasen-Modell zur Validierung, an dessen Ende, so die Hoffnung, eine klare, reliable und objektive Definition der jeweiligen psychischen Störung steht. In diesem Vorgehen lassen sich klare Bezüge zu Emil Kraepelins phänomenologisch orientierter Klassifikation erkennen, mitsamt der von Kraepelin ausgedrückten Hoffnung, eines Tages durch die Anwendung immer reliablerer und validerer wissenschaftlicher Methoden zu einem objektiven Bild der natürlichen „Essenz“ einer psychischen Störung zu gelangen (Sanders, 2011). Vor diesem historischen Hintergrund erscheint es kaum verwunderlich, dass Robins, Guze und die ihnen anhängigen Forschungsgruppen auch als *Neo-Kraepelinianer* Einzug in die Geschichtsbücher der klinischen Psychologie und Psychiatrie gefunden haben. Dies ist jedoch natürlich nicht allein der Ähnlichkeit ihrer Forderungen zu denen einer historisch bedeutsamen Persönlichkeit der Psychiatrie geschuldet, sondern in erster Linie der Tatsache, dass ihr Fünf-Phasen-Modell rasch als gängiger Standard für die Entwicklung und Validierung diagnostischer Kriterien angenommen wurde (Sanders, 2011). Die sogenannten *Feighner-Kriterien*, diagnostische Operationalisierungen für mehrere psychische Störungen (Feighner et al., 1972), resultierten

aus der direkten Anwendung des von Robins und Guze 1970 erstmals formulierten Ansatzes und prägten die weitere Entwicklung des DSM-Systems maßgeblich (Sanders, 2011). Zwar blieb der Konsens von Experten weiterhin ausschlaggebend für die Aufnahme von Störungskategorien in das DSM-III (Sanders, 2011), allerdings erhielten mit ihnen zwei wesentliche Grundkonzepte Einzug in eines der einflussreichsten Klassifikationssysteme und damit die gängige wissenschaftliche Praxis (Aragona, 2015). Diese sind zum einen die Grundkonzeption empirisch ermittelter Diagnosekriterien (und damit eine Operationalisierung latenter psychischer Störungen anhand manifester „Symptome“) und zum anderen eine Konzeption psychischer Störungen als letztendlich objektive, natürliche und „reale“ Wahrheit, also nosologischer Entitäten im engeren Sinne, deren wissenschaftliche Bestimmung in erster Linie von technischen, nicht grundsätzlichen Fragen abhängig ist. An dieser Stelle ist klar hervorzuheben, dass das DSM-System zwar nicht für sich beansprucht, gegenwärtig natürliche nosologische Entitäten abzubilden, dass allerdings bereits aus der Einleitung des DSM-5 hervorgeht, dass dies den angestrebten Idealzustand darstellt, der durch iterativ verbesserte wissenschaftliche Prozesse zugänglich gemacht werden kann (APA, 2013; Stein et al., 2010).

Der Ansatz, psychische Störungen als natürliche nosologische Entitäten zu konzipieren, manifestiert sich in verschiedenen Grundprinzipien des DSM, von denen mehrere Gegenstand anhaltender Debatten sind. Als beispielhaft hierfür kann die Kontroverse um die auch im DSM-5 weiterhin aufrecht erhaltene kategoriale Grundstruktur angesehen werden. Die Kritik an der kategorialen Struktur des Klassifikationssystems basiert im Wesentlichen auf der Einsicht, dass die meisten psychischen Störungen aus hochkomplexen und individuellen Ätiologien und Pathogenesen resultieren, deren Beschreibung durch disjunkte Kategorien als zu stark vereinfacht dargestellt wird (Kraemer, 2007; Widinger & Edmundson, 2011). Zwei sich hieraus ergebende Probleme bestehen zum einen in der Abgrenzung zwischen psychischen Störungen zu psychischer Gesundheit, zum anderen in der Abgrenzung der psychischen Störungen voneinander.

Die Abgrenzung zwischen psychischer Gesundheit und psychischer Krankheit ist eine bereits seit längerem diskutierte Grundsatzfrage der klinischen Psychologie und der Psychiatrie (z.B. Wakefield, 2007), deren Diskussion in ihrem Umfang und Detailreichtum in dieser Arbeit nur sehr verkürzt dargestellt werden kann. Hervorzuheben ist, dass eine Betrachtung von Gesundheit und Krankheit als dichotome Ausprägungen spätestens seit der *Ottawa-Charta* der Weltgesundheitsorganisation (WHO, 1986) nicht mehr als gängiger Konsens angesehen werden kann. Im Zuge dieser wurde eine Darstellung von Gesundheit und

Krankheit als mehrdimensionales Kontinuum in den offiziellen Gesundheits- und Krankheitsbegriff der WHO aufgenommen. Nicht zuletzt vor diesem Hintergrund erscheint eine Einteilung von Personen in die Kategorien „Störung liegt vor“ oder „Störung liegt nicht vor“ nur rechtfertigbar, wenn eine Loslösung von dem Anspruch, die Realität (mehr oder minder präzise) abzubilden, erfolgt. Dementsprechend erscheint der Versuch, psychische Störungen als nosologische Entitäten zu definieren, die sich klar von psychischer Gesundheit abgrenzen lassen, nur schwierig mit der gegenwärtig vorherrschenden Auffassung von Gesundheit und Krankheit in Form eines *Gesundheits-Krankheits-Kontinuums* vereinbar zu sein. Der hierbei implizierte *qualitative* Unterschied zwischen den zwei möglichen „Ausprägungen“ auf der Basis der Erreichung festgelegter Schwellenwerte (z.B. die Anzahl der Symptome, die Dauer der Symptombelastung oder die subjektiv eingeschätzte Intensität der Symptome) erscheint nicht zuletzt vor dem Hintergrund der zunehmend kritisierten Willkür bei der Definition der diagnostischen Kriterien unplausibel. Ein augenscheinliches Beispiel ergibt sich aus der Anwendung des Zeitkriteriums der DSM-5-Diagnose *Major Depressive Disorder*. Diese Diagnose kann gemäß DSM-5 nur vergeben werden, wenn eine mindestens zweiwöchige depressive Episode (mit jeweils eigenen definierenden Kriterien) vorliegt (APA, 2013). Nun erscheint es jedoch kontraintuitiv, dass es einen *qualitativen* Unterschied zwischen einer Person A mit einer depressiven Episode von zweiwöchiger Dauer und einer Person B mit einer depressiven Episode von eineinhalbwöchiger Dauer gibt (zumindest wenn man das Konstrukt „Depression“ aus dem DSM-5 nicht als Hilfskonstrukt zur besseren Kommunikation begreift, sondern als Annäherung an eine tatsächliche Realität, in der psychische Störungen natürlich vorkommende Taxa sind, also klar definierbare Grenzen haben). Ähnliche Gedankenspiele lassen sich für die klare Definition einer Mindestanzahl von Symptomen vornehmen. Weiterhin impliziert die Klassifikation, und damit die Annahme qualitativ unterschiedlicher Zustände, anhand des Kriteriums der Anzahl vorhandener Symptome in Form eines mindestens zu erreichenden Schwellenwertes, dass die eigentliche Symptomkonstellation nicht relevant ist. Hieraus lässt sich folgern, dass die Symptome a.) alle dasselbe latente Konstrukt widerspiegeln und sich b.) linear additiv verhalten (Fried, Nesse, Zivin, Guille, & Sen, 2014).

Doch nicht nur die Abgrenzung psychischer Störungen von psychischer Gesundheit, sondern auch die Abgrenzung der Störungsbilder voneinander steht im unmittelbaren Zusammenhang zur kategorialen Struktur der DSM-Klassifikation und der ihr zugrunde liegenden Annahme klar eingrenzbarer nosologischer Entitäten. Insbesondere das Phänomen der Komorbidität wird in diesem Zusammenhang immer wieder breit diskutiert (van Loo &

Romejin, 2015). Das zentrale Thema dieser Diskussion besteht in der Erklärung überzufällig häufig kovariierender Symptome verschiedener kategorialer Störungen. Auch hier bilden die rein konstruktivistische und die rein realistische Perspektive die beiden Pole eines zunehmend differenzierter werdenden Kontinuums der Erklärungsversuche für die empirisch beobachteten Kovarianzen psychopathologischer Symptome unterschiedlicher Störungen.

Aus konstruktivistischer Perspektive erscheint das Phänomen der Komorbidität vor allem auf methodischen Artefakten zu basieren, die wiederum im engen Zusammenhang mit der kategorialen Struktur des DSM-Systems stehen. Hier beschreiben van Loo und Romejin (2015) vier Faktoren, die gemäß der konstruktivistischen Kritiker die Komorbiditätsraten künstlich erhöhen. Diese bestehen in der Zunahme der Anzahl der im DSM definierten Störungsbilder, der Reduktion der Schwellenwerte bezüglich der Symptomanzahl, dem allmählichen Ausschluss der Ausschlusskriterien sowie der teilweise gravierenden Symptomüberlappung zwischen den Störungsbildern. Anzumerken ist hier, dass vor allem die ersten drei Faktoren nur einen mittelbaren Einfluss auf die postulierte künstliche Erhöhung der Komorbiditätsraten haben, indem sie die Anzahl der gestellten Diagnosen auf die eine oder andere Art erhöhen. Der vierte postulierte Faktor der Symptomüberlappung zwischen den Syndromen lässt sich zwar in den Augen des Autors der vorliegenden Arbeit einerseits nicht von der Hand weisen, allerdings stellt er nicht zwangsweise ein Argument für oder gegen die Künstlichkeit (oder artifizielle Konstruktion) von Komorbidität dar. Dies wird vor allem vor dem Hintergrund der Synthese der konstruktivistischen Position mit ihrer entgegengesetzten Position, dem Realismus, deutlich.

Der *realistischen* Perspektive zufolge reflektieren Komorbiditäten zwischen psychischen Störungsbildern gemeinsame zugrundeliegende Kausalstrukturen. Ihr zufolge signalisieren Komorbiditäten also, dass die aktuellen Erklärungen für Kovarianzen auf Symptomebene, in Form von angenommenen latenten nosologischen Störungsentitäten, die reale Kausalstruktur nur unzureichend widerspiegeln (Cramer et al., 2010; van Loo & Romejin, 2015). Als beispielhaft hierfür kann die Arbeit von Achenbach und KollegInnen angesehen werden, die mit den Konstrukten der *Internalisierung* und der *Externalisierung* zwei zugrundeliegende Dimensionen der Psychopathologie im Kindes- und Jugendalter vorschlugen, die die Kovarianz zwischen einer Vielzahl von Symptomen unterschiedlicher Diagnosegruppen erklären sollten (Widinger & Edmundson, 2011). Hinweise darauf, dass diese beiden latenten Dimensionen tatsächlich über die verschiedenen Störungsgruppen hinweg Kovarianzen in den Symptomen erklären, lieferten nicht zuletzt faktorenanalytisch gewonnene Ergebnisse (Krueger, 2002).

Allerdings lassen sich Komorbiditäten zwischen verschiedenen Störungsbildern nicht ausschließlich über kausale Zusammenhänge auf dem Abstraktionsniveau der angenommenen latent zugrundeliegenden diagnostischen Entitäten oder noch weiter abstrahierten latenten Dimensionen erklären, sondern auch auf der Ebene der Symptome. Diese Konzeption von Komorbidität, die in der Modellierung von individuellen klinischen Störungsbildern insbesondere im kognitions- und verhaltenstherapeutischen Kontext in Form von horizontalen Bedingungs-, Störungs- und Verhaltensanalysen (z.B. dem SORKC-Modell) schon länger Anwendung findet (Kanfer und Saslow stellten sie 1965 als Alternative zum traditionellen klassifikatorischen Ansatz vor), liegt auch dem ersten veröffentlichten Artikel, in dem eine netzwerkanalytische Untersuchung anhand psychopathologischer Konstrukte durchgeführt wurde, zugrunde (Cramer et al., 2010).

Latente-Variablen-Modelle, ihre epistemologischen Grundlagen und ontologischen Implikationen

Doch nicht nur die Konzeption psychischer Störungen als klar abgrenzbare nosologische und taxonomische Entitäten, sondern auch ihre Modellierung als latente Konstrukte und ihre Operationalisierung anhand manifester Variablen ist Gegenstand anhaltender Diskussionen. Diese werden zwar überwiegend auf methodischer Ebene geführt, schließen jedoch auch grundsätzliche ontologische Fragen mit ein.

Die Verwendung latenter Variablen in der Psychologie kann, analog zu Ebbinghaus' Bonmot der Psychologie als Wissenschaft mit einer kurzen Geschichte, aber einer langen Vergangenheit (Ebbinghaus, 1908), zumindest in impliziter Form auf eine lange Historie zurückblicken. So gehört die Erklärung manifest beobachtbarer Phänomene durch nicht direkt beobachtbare Gründe in Form induktiv gezogener Schlüsse zur grundlegenden kognitiven Struktur des Menschen (Sternberg & Sternberg, 2012). In Bezug auf menschliches Verhalten und Erleben, also dem Gegenstand der heutigen wissenschaftlichen Psychologie, lassen sich in der Geschichte unzählige Beispiele für induktiv gezogene Erklärungsmuster finden.

Auch wenn diese Art der Erklärung nicht formalisiert und in den modernen Begriffen der latenten und manifesten Variablen formuliert ist, so bildet sie doch das grundsätzliche Muster latenter Modellierungen ab: direkt beobachtbare Untersuchungsgegenstände werden durch zugrundeliegende theoretische Konstrukte erklärt. Konkret handelt es sich hierbei aus mathematischer Sicht um den Versuch, Kovarianzstrukturen durch kausale Einflüsse postulierter latenter Variablen zu erklären. Die mathematische Formulierung dieses Grundprinzips lässt sich auf die Arbeit Charles Spearman's zurückführen, der 1904 das Konzept der Faktorenanalyse entwickelte und damit die Grundlage für eine Vielzahl

methodischer Paradigmen schuf, die in den folgenden Jahrzehnten die Methodologie der Psychologie maßgeblich prägten (Borsboom, Mellenbergh, & van Heerden, 2003). So entwickelte sich aus der Faktorenanalyse heraus die *konfirmatorische Faktorenanalyse* und damit die Grundlage der *Strukturgleichungsmodelle*. Weiterhin fand das Grundprinzip, Zusammenhänge zwischen manifesten Variablen über angenommene latente Variablen zu erklären, Anwendung in der Entwicklung der *Item Response Theory* (zur Erklärung dichotomer manifester Variablen anhand kontinuierlicher latenter Konstrukte) und den verschiedenen Analyseverfahren latenter Strukturen wie etwa der *latent class analysis* (Borsboom et al., 2003). Gemeinsam ist diesen Techniken die Spezifikation eines Zusammenhangs zwischen zwei grundlegend unterschiedlichen Arten von Variablen, den latenten und den manifesten. Hierbei werden jedoch implizit zwei Annahmen getroffen, eine ontologischer und eine epistemologischer Natur, die sowohl für sich allein betrachtet als auch in der Anwendung in der Psychopathologie diskussionswürdig sind.

Zunächst lässt sich sagen, dass sich in dem Versuch, konkret beobachtbare Variablen durch postulierte latente Konstrukte zu erklären, eine im Kern realistische Grundhaltung (im ontologischen Sinne) gegenüber diesen latenten Konstrukten ausdrückt (Borsboom et al., 2003). Dieser Umstand ergibt sich aus der typischen Methodik, im Zuge derer die manifesten Variablen als die latente Variable reflektierend beschrieben werden. Hieraus erschließt sich auch die geläufige Bezeichnung dieser Art der Modellierung als *reflektiv*. Die Prämisse dieses Grundansatzes besteht darin, dass es eine Entität gibt, die reflektiert werden kann. Diese Entität kann nun entweder als eine reale Gegebenheit, oder aber als ein artifizielles Konstrukt, also realistisch oder konstruktivistisch, verstanden werden. Die latenten-Variablen-Modelle, insbesondere die in der Psychologie verwendeten, basieren nun im Wesentlichen darauf, dass die Ausprägung der latenten Variable aus den konkreten Werten der manifesten Variablen berechnet wird, was sich mathematisch in der Verwendung gewichteter Summenscores ausdrückt (Borsboom et al., 2003). Diese Summenscores werden im Folgenden inhaltlich interpretiert und drücken, so der Anspruch, die konkrete Ausprägung der latenten Variable im vorliegenden Fall aus. Die Logik hinter diesem Vorgehen scheint einfach zu sein. Wenn A, B und C als Funktion einer latenten Variable Z variieren, dann lässt sich mittels A, B und C, in Form eines gewichteten Summenscores, auf die Ausprägung von Z schließen. Eine ontologisch streng konstruktivistische Auffassung der latenten Variable würde diese nun zur Tautologie machen: Der gewichtete Summenscore dreier Variablen drückt den gewichteten Summenscore dieser drei Variablen, und nicht mehr, aus. Hierbei würde die „gemessene latente Variable“ (gemeint ist hier das rein mathematische Konstrukt) mit den jeweiligen

Operatoren, die für ihre Messung verwendet werden, gleichgesetzt werden, sie hätte also keinerlei Erklärungsmacht, die über die manifesten Variablen, und damit über die verwendete Operationalisierung, hinausgeht (Borsboom et al., 2003). Dies ist jedoch nicht mit dem Anspruch einer kausal wirksamen latenten Variable vereinbar. Wenn die Variable A die Variable B kausal beeinflusst, kann sie sich nicht ausschließlich aus B zusammensetzen (Borsboom et al., 2003). Da dieses Vorgehen wiederum nicht mit dem Anspruch der latenten-Variablen-Modelle und vor allem auch dem Anspruch der wissenschaftlichen Psychologie vereinbar scheint, kann geschlussfolgert werden, dass in diesen die Annahme besteht, dass die jeweils untersuchten latenten Variablen eine inhärente Bedeutung haben, die über die Erklärung der jeweils verwendeten manifesten Variablen hinausgeht. Als ein Beispiel für eine Anwendung hiervon lässt sich die adaptive Diagnostik beschreiben, die im Wesentlichen davon ausgeht, dass eine latente Variable (z.B. Intelligenz) durch eine endliche Anzahl manifester Variablen ausreichend präzise geschätzt werden kann. Die Antworten der untersuchten Personen werden hierbei durch ein latentes inhaltliches Konstrukt erklärt und nicht auf ein rein methodisches latentes Konstrukt zurückgeführt, das die Beantwortung von ausschließlich den Items, die vorgegeben wurden, erklärt.

Aus epistemologischer Perspektive erscheint eine weitere wesentliche Komponente der latenten-Variablen-Modelle interessant: die Zuschreibung einer inhaltlichen Bedeutung auf der Ebene des Individuums. So kann das konzeptionelle Vorgehen der latenten-Variablen-Modelle im Wesentlichen durch den Grundgedanken der induktiven Logik beschrieben werden: beobachtbare Kovarianzen zwischen Variablen werden durch den Einfluss postulierter zugrundeliegender Variablen erklärt. Hierbei wird die Annahme getroffen, dass die latente Variable die manifesten Variablen kausal beeinflusst, was sich auf mathematisch-formaler Ebene im Prinzip der *lokalen Unabhängigkeit* ausdrückt. Demzufolge lässt sich die Kovarianz zwischen Variablen vollständig durch eine weitere, hier latente, Variable erklären, wenn die Zusammenhänge zwischen den Variablen bei Konditionierung auf die latente Variable verschwinden (z.B. Brown, 2013). Dies kann auch als mathematische Formulierung des reflektiven Modells angesehen werden, gemäß dem die Zusammenhänge zwischen den manifesten Variablen durch eine latente Variable erklärt werden, also bei Konditionierung auf diese verschwinden.

In der praktischen Anwendung ist hierbei interessant, dass die Beschreibung von Kovarianz auf Populationsebene eigentlich nicht geeignet ist, um kausale Mechanismen auf der Individuums-Ebene zu erklären. Hierbei würde es sich um einen klassischen syllogistischen Fehlschluss handeln: auf Populationsebene erklärt das latente Konstrukt

Kovarianz zwischen zwei Items, also ist es kausal für die Antwort einer Person auf beide Items verantwortlich (für eine detaillierte und differenzierte Auseinandersetzung mit diesem Argument wird auf Borsboom et al., 2003 verwiesen). Dies wird in Bezug auf die Psychopathologie ersichtlich, wenn die konkrete Anwendung der Modellierung von Störungsbildern betrachtet wird. Diese werden mit einer gewissen Regelmäßigkeit nicht anhand von latenten oder manifesten Variablen modelliert, sondern es werden individuelle kognitive, affektive, körperliche, verhaltensbezogene, soziale, und weitere Prozesse inhaltlich abgebildet (Wittchen & Hoyer, 2011). Auch mag es sein, dass es auf Populationsebene möglich ist, gemeinsame Varianzanteile zwischen einzelnen Symptomen zu finden und diese statistisch über die Untersuchung ihrer lokalen Unabhängigkeit auf den Einfluss einer latenten Variable zurückzuführen. Allerdings scheint es in der konsequenten Anwendung auf Ebene des Individuums Beispiele zu geben, die diesem Strukturmodell nicht entsprechen. So erscheint es in hohem Maße unplausibel, dass die beiden DSM-5-Depressionssymptome *Insomnie* und *Müdigkeit* lediglich aufgrund des kausalen Einflusses eines latenten Konstruktes *Depression* kovariieren, ein kausaler Zusammenhang zwischen diesen beiden Symptomen wäre naheliegend (z.B. Fried et al., 2017). Dieses Beispiel mag plakativ sein, es spiegelt jedoch ein zugrundeliegendes generelles Problem der Verknüpfung von Strukturen auf Populationsniveau mit Strukturen auf dem Niveau des Individuums wider.

Dass dies tatsächlich ein fundamentales Problem der aktuellen klassifikatorischen Diagnostik des DSM-5 ist, wird ersichtlich, wenn man sich erneut vor Augen hält, dass in diesem der Ansatz der Symptom-Äquivalenz verwendet wird (Fried et al., 2014). Dieser Ansatz kann als eine direkte Folge der Modellierung psychischer Störungen als *latente-Variablen-Modelle* angesehen werden. So spielt es eine untergeordnete Rolle, welche Symptome manifest vorliegen, entscheidend ist in den meisten Fällen die Erreichung eines Schwellenwertes hinsichtlich der Anzahl der Symptome, ein Vorgehen, das nur sinnvoll ist, wenn diese auch tatsächlich allesamt gleichwertige Indikatoren desselben latenten Konstruktes sind. Hiermit wird jedoch ein latentes-Variablen-Modell, das streng betrachtet lediglich Kovarianzen auf Populationsebene erklärt, zur Grundlage von Entscheidungen auf individueller Ebene gemacht, ein Vorgang, der in seiner Anwendung durchaus kritisch hinterfragt werden kann. Weiterhin erscheint die Annahme einer kausal für die Ausprägung der manifesten Variablen, hier in Form der Symptome, verantwortlichen latenten Variable, der psychischen Störung, vor dem Hintergrund der zeitlichen Abfolge von Ursache und Wirkung hinterfragbar. Bei einem angenommenen reflektiven Kausalzusammenhang zwischen dem latenten Konstrukt *psychische Störung* und den manifesten Symptomen müsste

die psychische Störung den Symptomen zeitlich voran gehen. Hierfür müsste sie jedoch als eine eigenständige, von den Symptomen unabhängige Entität, existieren, was zumindest diskussionswürdig scheint.

Netzwerkmodellierungen als alternative methodische und theoretisch-konzeptionelle Ansätze

Eine mögliche Alternative zur Modellierung von Kovarianzstrukturen besteht in der Annahme von direkten Interaktionen zwischen den betrachteten Variablen in Form von Netzwerken. Dieser Modellierung liegt ein einfacher Gedanke zugrunde, der in der psychologischen Praxis auf Ebene des Individuums bereits gängig ist (z.B. Kanfer & Saslow, 1965). Wurden die Kovarianzen zwischen Variablen in der bisher üblichen Modellierung durch die putativ kausalen Einflüsse einer postulierten latenten Variable erklärt, so besteht das Grundprinzip der Netzwerkmodellierung darin, dass die betrachteten Variablen sich gegenseitig beeinflussen können. So wird beispielsweise die Kovarianz zwischen den beiden DSM-5-Depressionssymptomen *Insomnie* und *Müdigkeit* nicht darauf zurückgeführt, dass beide lediglich Indikatoren für ein und dasselbe latente Konstrukt, Depression, sind, sondern darauf, dass es einen bedeutungsvollen und inhaltlichen Zusammenhang zwischen ihnen gibt, hier in der Form, dass Insomnie zu Müdigkeit führt (z.B. Borsboom & Cramer, 2013).

Hierin spiegelt sich bereits das Grundkonzept der Modellierung und Analyse psychologischer Konstrukte als Netzwerke wider. Psychologische Konstrukte werden nicht länger als latente Variablen definiert, auf die mittels operationalisierter Entitäten zurückgeschlossen wird, sondern sie werden als Netzwerke interagierender Variablen definiert, aus deren Struktur und Dynamiken die psychologischen Konstrukte *emergent* entstehen (Schmittmann et al., 2013). Ein Konzept, das der Veranschaulichung dieses Zusammenhangs und der klaren ontologischen Abgrenzung von der Modellierung psychologischer Konstrukte als latente Variablen dienlich erscheint, ist das der Mereologie (Schmittmann et al., 2013). So lässt sich der Zusammenhang zwischen den in den Netzwerkstrukturen konkret abgebildeten Variablen, beispielsweise den „Symptomen“ einer psychischen Störung, und den psychologischen Konstrukten, beispielsweise einer psychischen Störung, nicht länger in Form eines Messmodells im Sinne einer latenten Variable, deren Ausprägung über manifeste Variablen erschlossen wird, beschreiben, sondern in Form der Mereologie: die beobachteten Variablen messen nicht ein Konstrukt, sondern sind Teil des Konstruktes und bringen es aus sich und ihren Interaktionen hervor. Als Analogie dieser Unterscheidung kann der Vergleich zwischen den Störungsmodellen der Klassifikationssysteme der klinischen Psychologie und Psychiatrie einerseits und den

Störungsmodellen der funktionalen Verhaltens- und Bedingungsanalysen aus der Praxis der klinischen Psychologie und Psychiatrie andererseits angesehen werden.

Aus dieser alternativen Art der Modellierung ergeben sich einerseits die Vermeidung einiger konzeptioneller sowie methodischer Probleme, andererseits, durch die Verwendung eines durch die Graphentheorie erweiterten Instrumentariums, Möglichkeiten zur Untersuchung von Gegenständen, die sich mit den in der Psychologie gängigen Methoden nicht darstellen lassen.

Mathematisch-konzeptionell betrachtet stellen die in dieser Arbeit untersuchten Symptomnetzwerke eine visuelle Darstellung korrelativer Zusammenhänge in Form von Graphen dar. Hierdurch ergibt sich die Möglichkeit, die eigentlich probabilistischen Inhalte in der Nomenklatur des mathematischen Gebiets der Graphentheorie darzustellen und sie mittels einschlägiger Analysemethoden zu untersuchen. Die hierfür verwendeten graphischen Modelle lassen sich zunächst in *gerichtete* und *ungerichtete* einteilen (Newman, 2010), wobei der konzeptionelle Unterschied darin besteht, dass gerichtete Netzwerke gerichtete Zusammenhänge darstellen, etwa in der Form, dass Insomnie Müdigkeit verursacht, und ungerichtete lediglich ungerichtete Zusammenhänge, beispielsweise dass es einen Zusammenhang zwischen Insomnie und Müdigkeit gibt. Mittels gerichteter Netzwerke ist es dementsprechend zumindest prinzipiell möglich, kausale Strukturen darzustellen, hierfür müssen jedoch, parallel zu John Stuart Mills Konzeption von Kausalität (Mill, 1875), einige Zusatzannahmen erfüllt sein (Pearl, 2009), die einer Anwendung in der psychologischen Forschung derzeit im Wege stehen (Epskamp, Borsboom, & Fried, 2018). So ist es für die Schätzung einer Kausalstruktur notwendig, dass eine zeitliche Struktur in den Daten vorliegt und alle postulierten Agenten des Kausalzusammenhangs in die Modellierung einbezogen werden können sowie keine zyklischen Verbindungen vorliegen (Epskamp, 2017). Da die Erfüllung dieser Annahmen im Falle psychologischer Untersuchungen bisher kaum realisierbar erscheint (Epskamp, 2017), beschränken sich diese bislang entweder auf die Modellierung gerichteter Netzwerke, die zeitliche Strukturen widerspiegeln, oder aber ungerichteter Netzwerke, auch als *Markov Random Fields* bezeichnet (Newman, 2010). Eine Form der Markov Random Fields sind die *Pairwise Markov Random Fields* (PMRF's), in denen die *paarweisen* Assoziationen zwischen zwei Variablen einen Zusammenhang abbilden, der bei Konditionierung auf alle anderen abgebildeten Variablen bestehen bleibt. Diese Verbindungen können als partielle Korrelationen verstanden werden, bei denen die Einflüsse sämtlicher anderer untersuchter Variablen herauspartialisiert werden (Epskamp, 2017). Besteht zwischen zwei Variablen in der graphischen Darstellung keine Verknüpfung,

können diese also als bedingt unabhängig (vor dem Hintergrund aller weiterer Variablen) angesehen werden (Hevey, 2018).

Die gängige Terminologie der Netzwerkanalysen stützt sich auf die der Graphentheorie, gemäß der die abgebildeten Variablen als *Nodes* und die Zusammenhänge zwischen ihnen als *Edges* bezeichnet werden (Newman, 2010). Eine Anwendungsform der PMRF's besteht nun in der Modellierung von Zusammenhängen zwischen ordinalen oder metrischen Variablen in Form von *Gauß'schen Graphischen Modellen*. Diese dienen der Darstellung von ungerichteten Zusammenhängen multivariat normalverteilter Variablen in Form von Netzwerken (Epskamp et al., 2018). Die Annahme der multivariaten Normalverteilung speist sich hier in erster Linie aus der Verwendung von Zusammenhangsmaßen, die für eine valide Schätzung der Zusammenhänge auf normalverteilte Variablen angewiesen sind, wie etwa der Pearson-Korrelation (Field, 2009). Hierbei ist hervorzuheben, dass sich diese Annahme in erster Linie auf die *Edges* beziehungsweise deren Interpretierbarkeit der Gauß'schen Graphischen Modelle bezieht. Dementsprechend kann bei Verletzungen der Annahme der multivariaten Normalverteilung bis zu einem gewissen Maße auf alternative voraussetzungsärmere Zusammenhangsmaße zurückgegriffen werden, um den verzerrenden Effekten der Voraussetzungsverletzungen entgegenzuwirken. Ein Beispiel hierfür besteht in der gängigen Verwendung von Spearman-Rangkorrelationen bei der Modellierung von Zusammenhängen zwischen ordinalskalierten Variablen oder polychorischer Korrelationen bei schiefverteilten Variablen (Epskamp, 2017).

Regularisierung.

Aus dem Grundprinzip der Netzwerkmodellierung ergibt sich jedoch ein für die psychologische Forschung nicht unproblematischer Umstand. Dieser besteht darin, dass in den PMRF's sämtliche möglichen paarweisen Interaktionen zwischen allen untersuchten Variablen analysiert und damit geschätzt werden müssen. Der Zusammenhang zwischen der Anzahl der untersuchten Variablen und der Anzahl der zu schätzenden *Edges* lässt sich kombinatorisch über den Binomialkoeffizienten $\binom{n}{k}$ darstellen. Da in den PMRF's paarweise Interaktionen modelliert werden, ergeben sich hierdurch $\binom{n}{2}$ zu schätzende Parameter, wobei n die Anzahl der Nodes des Netzwerkes wiedergibt. Hieraus ergeben sich bereits für 10 Nodes $\binom{10}{2} = 45$, für 53 Nodes $\binom{53}{2} = 1378$ zu schätzende Korrelationen, von denen aufgrund der α -Fehlerkumulierung eine Vielzahl lediglich Scheinkorrelationen darstellen (Walters, 2016). Bei dem für das klassisch-frequentistische Paradigma üblichen Vorgehen würde eine

Korrektur des α -Fehlerniveaus mittels Bonferroni-Korrektur zu einem korrigierten α -Fehlerniveau von $\alpha = .001$ für ein Netzwerk aus 10 Nodes, für ein Netzwerk aus 53 Nodes einem korrigierten α -Fehlerniveau von $\alpha = .000004$ führen. Dieser Umstand erscheint jedoch kaum mit den in der psychologischen Forschung typischerweise verwendeten Stichprobengrößen und angenommenen Effektstärken vereinbar zu sein.

Vor diesem Hintergrund wurde die Methode der *Regularisierung*, ursprünglich im Anwendungsgebiet des *Machine-Learning* entwickelt, in den Methodenkanon der Netzwerkanalysen aufgenommen (Epskamp et al., 2018). Die Regularisierung dient, ähnlich der Korrektur des Signifikanzniveaus bei multiplen Testungen im klassischen frequentistischen Paradigma, dem Ausschluss von Scheinkorrelationen (Hevey, 2018). Hierbei ist der zugrunde liegende Gedanke der gleiche: durch zufällige Variationen in der Stichprobe (Stichprobenvarianz) entstehen Kovarianzen, aus denen in weiterer Folge statistische Maße wie etwa Korrelation abgeleitet werden, die sich in weiter gezogenen Stichproben nicht finden lassen würden und damit als methodische Artefakte angesehen werden können. Im klassisch-frequentistischen Paradigma findet der p -Wert Anwendung, um wiederzugeben, wie wahrscheinlich eine empirisch vorliegende Datenstruktur ist, wenn eigentlich die Nullhypothese gilt und somit die untersuchten Variablen nur unsystematisch variieren. Hier wird durch die Spezifikation eines Schwellenwertes, oft von $p \leq .05$, schlicht festgelegt, dass statistische Aussagen mit einer gewissen Wahrscheinlichkeit nicht auf unsystematischen Variationen basieren sollen. Hierfür erfolgt eine Berechnung verschiedener Teststatistiken, deren Wahrscheinlichkeit eines Zustandekommens unter Berücksichtigung verschiedener Parameter, wie etwa der Varianzen innerhalb verglichener Gruppen, untersucht wird (Bortz & Schuster, 2010). Diese Teststatistiken spiegeln im wesentlichen *Signal-to-Noise-Ratios* wider (Carlin & Doyle, 2001). Hieraus ergibt sich im Kontext der Modellerstellung, also der Entscheidung, welche Parameter in das Modell aufgenommen werden und welche nicht, ein direkter Zusammenhang zwischen der Verwendung von p -Schwellenwerten und der Kontrolle um Scheinkorrelationen (Walters, 2016). Der gleiche Grundgedanke, wenn auch unter Verwendung anderer Methoden, liegt dem Konzept der Regularisierung zugrunde (Epskamp, 2017). Als das im psychologischen Bereich gängigste Regularisierungsverfahren kann das sogenannte *least absolute shrinkage and selection operator* (LASSO) angesehen werden (Epskamp, 2017).

Das LASSO-Verfahren stellt eine Methode zur Modellauswahl dar, die im Vergleich zu den im Bereich des Machine-Learning gängigen Verfahren mit wenig Rechenaufwand verbunden ist (Zhao & Yu, 2006). Das Grundanliegen dieser Methode besteht in der

Erklärung einer Kovarianzstruktur mittels eines möglichst sparsamen Netzwerkmodells in Form korrelativer Zusammenhangsmaße (Hevey, 2018). Hierfür werden, in Abhängigkeit von einem zu spezifizierenden Tuning-Parameter λ , die untersuchten Zusammenhänge im Betrag reduziert und, wenn sie hierdurch auf Null geschätzt werden, aus dem regularisierten Modell entfernt. Die Spezifikation des Tuning-Parameters λ sollte in der Grundkonzeption der LASSO-Regularisierung dementsprechend vor dem Hintergrund einer Abwägung zwischen der Sensitivität und der Spezifität der durchgeführten Analyse erfolgen (Hevey, 2018). Wird der Tuning-Parameter λ niedrig angesetzt, werden viele Korrelationen und damit auch viele Scheinkorrelationen ungleich Null geschätzt (Fehler erster Art), wird er hoch angesetzt, werden viele Korrelationen aus dem Modell entfernt, darunter eventuell auch tatsächliche (Fehler zweiter Art) (Epskamp, 2017). Eine gängige Praxis besteht hier in der Schätzung einer Vielzahl von Modellen unter verschiedenen Werten von λ und der Auswahl eines Modells anhand von *Fit-Indices* (Epskamp & Fried, 2018). Für die Auswahl eines der unterschiedlich „streng“ geschätzten Modelle können verschiedene Informationskriterien berechnet und herangezogen werden, üblich ist jedoch die Verwendung des *Extended Bayesian Information Criterion* (EBIC) (Epskamp, 2017). Aus Simulationsstudien geht hervor, dass die LASSO-Regularisierung unter Verwendung des EBIC generell eine hohe Spezifität aufweist, die Sensitivität jedoch mit Charakteristika des wahren Netzwerks sowie der Stichprobengröße variiert (Epskamp, 2017). Dieses Primat der Spezifität ergibt sich aus der Grundkonzeption der Regularisierung, die in der Verwendung verschiedener Techniken zur Vermeidung von Scheinkorrelationen besteht. Hierin lassen sich erneut Parallelen zur frequentistischen Nullhypothesentestung erkennen: wird die Nullhypothese beibehalten, ist dies nicht gleichbedeutend damit, dass sie als wahr angesehen wird (Wagenmakers, 2007), oder, in den Begrifflichkeiten der Netzwerkmodellierung ausgedrückt: dass eine Edge ausgeschlossen wird, kann bedeuten, dass die Nullhypothese zutrifft, oder aber, dass die Daten zu „verrauscht“ sind (Hevey, 2018).

Trotz der Parallelen zwischen den Grundideen der Konzepte der Regularisierung einerseits und der Modellschätzung anhand von p -Werten andererseits gibt es Unterschiede hinsichtlich der mathematischen Ansätze. Dies spiegelt sich auch hinsichtlich der Performanz in verschiedenen Szenarien wider. So verglichen Zhao und Yu (2006) in mehreren Simulationsstudien unter anderem die Performanz der klassischen Modellauswahl mittels Schwellenwerten (worunter auch der in der Psychologie gebräuchliche p -Wert fällt) mit der Performanz der LASSO-Regularisierung hinsichtlich der Konvergenz der Schätzungen mit einem simulierten wahren Modell anhand simulierter Stichproben. Ein wesentlicher Befund

dieser Simulationsstudien bestand in der annähernden Äquivalenz der beiden Vorgehensweisen im Bereich der sehr großen Stichproben ($n \rightarrow N$), allerdings war die LASSO-Regularisierung der Modellierung anhand Schwellenwerten im Bereich der kleinen Stichprobengrößen klar überlegen. Da in der psychologischen Forschung selten erwartet werden kann, dass die Grundpopulation asymptotisch komplett erhoben wird, kann die LASSO-Regularisierung also als nützliche Alternative zur klassischen Modellierung anhand von p -Werten angesehen werden.

Analysen der Stabilität und der Akkuratheit der Schätzungen.

Zwei weitere hiermit unmittelbar im Zusammenhang stehende Konzepte sind die der Akkuratheit und der Stabilität der Schätzungen. Die Untersuchung der Stabilität und der Robustheit der Schätzungen scheint nicht nur für die Einschätzung der Interpretierbarkeit der Ergebnisse wichtig, sondern auch vor dem Hintergrund der Replikationskrise der Psychologie (Open Science Collaboration, 2015) ein wesentlicher Schritt zu sein (Epskamp et al., 2018).

Über die Akkuratheit ist darstellbar, inwiefern die ermittelten Netzwerkstrukturen, insbesondere die Edges, durch die Stichprobenvarianz beeinflusst werden (Epskamp et al., 2018). Hierfür werden anhand der Stichprobenverteilung Konfidenzintervalle der einzelnen Edges berechnet, meist auf dem 95%-Niveau. Da hierdurch auf Basis der vorliegenden Daten veranschaulicht wird, wie groß der Bereich ist, in dem der wahre Populationswert mit 95%iger Wahrscheinlichkeit liegt, wird erkennbar, wie präzise die jeweilige Edge geschätzt wird. Große Intervalle liefern hier Hinweise auf unpräzise Schätzungen, kleine Intervalle auf präzise Schätzungen. Die hierfür erforderlichen Stichprobenverteilungen der korrelativen Zusammenhänge (Edges) werden mittels Bootstrap-Verfahren erzeugt (Epskamp et al., 2018).

Die Stabilität eines Netzwerks hingegen spiegelt wider, inwiefern die Schätzungen von spezifischen Eigenschaften der Netzwerke, den Zentralitätsmaßen, durch die Stichprobenvarianz beeinflusst werden (Epskamp et al., 2018). Im Zuge dieser Untersuchung wird mittels Bootstrap-Verfahren eine Vielzahl an Netzwerkstrukturen aus Substichproben mit unterschiedlicher Größe und unterschiedlicher Zusammensetzung geschätzt. Hieraus ergeben sich für jede untersuchte Substichprobengröße, entsprechend der unterschiedlichen Stichprobenzusammensetzungen, mehrere geschätzte Netzwerke und damit auch Kennwerte. Dies ermöglicht einerseits die Berechnung korrelativer Zusammenhänge zwischen den Kennwerten der unterschiedlichen Substichproben und den eigentlich geschätzten Kennwerten (Epskamp et al., 2018). Andererseits kann dargestellt werden, inwiefern die Konfidenzintervalle der Schätzungen der Parameter auf unterschiedlichen Ausprägungen des Faktors „Substichprobengröße“ präzisere oder unpräzisere Schätzungen anzeigen. Eine

direkte Quantifizierung der Stabilität und der Akkuratheit der Schätzungen liefert der *Correlation Stability Coefficient* (CS-Koeffizient) (Epskamp et al., 2018). Der CS-Koeffizient zeigt den Anteil an Personen an, der maximal von der Untersuchung ausgeschlossen werden könnte, sodass die Korrelation der aus der Gesamtstichprobe geschätzten Werte und die der aus den Substichproben (mit jeweils unterschiedlichen Zusammensetzungen) geschätzten Werte einen gewissen Schwellenwert nicht unterschreitet. Dieser Schwellenwert ist per Konvention auf $cor = .70$ festgesetzt, kann aber prinzipiell entsprechend der angestrebten Stabilität der Untersuchungen angepasst werden. Die Konvention von $cor = .70$ ergibt sich aus der Interpretation einer solchen Korrelation als sehr starker Effekt (Cohen, 1988, zitiert nach Epskamp et al., 2018). Der CS-Koeffizient lässt sich sowohl für die einzelnen Edges als auch für die Zentralitätsmaße berechnen.

Anhand einer Simulationsstudie kamen Epskamp et al. (2017) zu dem Ergebnis, dass zur Interpretation von Netzwerkstrukturen die jeweiligen CS-Koeffizienten auf dem .70-Niveau den Wert von $CS(cor = .70) = .25$ nicht unterschreiten sollten und vorzugsweise über $CS(cor = .70) = .50$ liegen sollten.

Netzwerkanalysen als Kombination der Stochastik und der Graphentheorie.

Die Modellierung und Analyse psychologischer Konstrukte als Netzwerke kausal aufeinander einwirkender Entitäten bietet, neben der Vermeidung konzeptioneller und methodischer Probleme, die Möglichkeit der Betrachtung komplexer Phänomene auf einer basalen Ebene, beispielsweise im Bereich der Psychopathologie der Ebene einzelner Symptome. Hierdurch lassen sich Untersuchungen durchführen, die der Komplexität und Vielfältigkeit der behandelten Gegenstände gerecht werden. Als beispielhaft hierfür kann die Untersuchung von Fried et al. (2014) angesehen werden, in der Zusammenhänge zwischen Risikofaktoren der Depression mit einzelnen Depressions-Symptomen dargestellt wurden. Ein wesentlicher Befund bestand hierbei in dem klar ausdifferenzierten Zusammenhangsprofil der einzelnen Risikofaktoren, wie etwa dem Geschlecht, vorangegangenen depressiven Erkrankungen, oder aber auch dem Neurotizismus-Score der Versuchspersonen mit den einzelnen Symptomen. Diese differentiellen Effekte legen nahe, dass eine Untersuchung auf Symptomebene durchaus zu differenzierten Erkenntnissen führen kann. Auch kann die Darstellung komplexer Kovarianzstrukturen in Form von Netzwerken Aufschlüsse über Zusammenhänge wie beispielsweise Symptominteraktionen geben, die in den bisher gängigen Analysemethoden wenig Beachtung fanden.

Andererseits ermöglicht die Modellierung psychologischer Konstrukte als Netzwerke kausal aufeinander einwirkender Entitäten die Anwendung graphentheoretischer

Analysemethoden, über die sich Sachverhalte untersuchen lassen, die mit den in der Psychologie gängigen Methoden kaum oder nicht darstellbar sind. Hierbei handelt es sich in erster Linie um den Netzwerken inhärente Eigenschaften, deren inhaltliche Interpretation im theoretischen Rahmen der Psychologie möglich ist. Zwar hat die Graphentheorie ein breites Spektrum an Analysemethoden hervorgebracht, allerdings können diese aufgrund methodischer und konzeptioneller Beschränkungen nicht komplett auf „psychologische Daten“ angewandt werden. Deshalb soll in dieser Arbeit nur ein Einblick in die im psychologischen Kontext gängigerweise untersuchten Sachverhalte gegeben werden, namentlich die Zentralitätsmaße, die Konnektivität und den Impact.

Die Zentralitätsmaße.

Die Zentralitätsmaße spiegeln die Wichtigkeit einer Node in Bezug auf ihr Netzwerk wider (Newman, 2010), wodurch sie sich klar von auf den ersten Blick ähnlich erscheinenden Konzepten aus den latenten-Variablen-Modellen unterscheiden. So bieten diese in erster Linie die Möglichkeit zu untersuchen, inwiefern einzelne Variablen tatsächlich Indikatoren für eine putativ zugrundeliegende Variable darstellen und somit eine Evaluation eines Messmodells. Mit den Zentralitätsmaßen lassen sich jedoch einzelne Nodes hinsichtlich ihrer Wichtigkeit in Struktur- oder Prozessmodellen (entsprechend der Untersuchung quer- oder längsschnittlicher Netzwerkstrukturen) darstellen (Fried, Epskamp, Nesse, Tuerlinckx, & Borsboom, 2016). Eine Untersuchung der Zentralitätsmaße empfiehlt sich jedoch nicht nur aufgrund der möglichen inhaltlichen Implikationen, sondern auch aufgrund der Tatsache, dass die grafische Abbildung der Netzwerkstruktur allein kaum ausreicht, um Inferenzen über die Struktur zu ziehen (Epskamp, 2017). Dies liegt darin begründet, dass ein Netzwerk mit k Nodes eigentlich einen k -dimensionalen Raum definiert, in dem die Verknüpfungen zwischen allen Variablen grafisch abgetragen sind. Diese Abbildung von hochdimensionalen Räumen in zweidimensionale Grafiken erfolgt meist nach Algorithmen, die eine möglichst hohe Passung zwischen den metrischen Distanzen der Abbildung und den Stärken der korrelativen Zusammenhänge anstreben, beispielsweise dem *Fruchterman-Reingold-Algorithmus* (Epskamp, 2017). Im Beispiel des Fruchterman-Reingold-Algorithmus entsprechen kurze Edges stärkeren korrelativen Zusammenhängen und längere Edges schwächeren. Da hierdurch allerdings ein mehrdimensionaler Raum in einer zweidimensionalen grafischen Darstellung abgebildet wird, kommt es fast unweigerlich zu Verzerrungen, weshalb eine direkte Interpretation, beispielsweise in Form von „Node 1 befindet sich im Zentrum der Abbildung, spielt also eine wesentliche Rolle“, nicht möglich ist (Epskamp, 2017). Somit

stellt die Berechnung der Zentralitätsmaße eine wesentliche Komponente in der Analyse von Netzwerkstrukturen dar.

Typischerweise werden in psychologischen Netzwerkanalysen die drei Zentralitätsmaße *Node-Strength*, *Betweenness* sowie *Closeness* untersucht (Epskamp, 2017). Die *Node-Strength* quantifiziert die Stärke aller direkten Verknüpfungen einer Node, also aller Partialkorrelationen einer Variable. Hierbei werden sowohl positive als auch negative Edges berücksichtigt. Die *Betweenness* drückt aus, wie oft eine Node auf dem kürzesten direkten Pfad zwischen zwei anderen Knoten liegt und damit, welchen Einfluss sie bei der Propagierung von Aktivität durchschnittlich ausübt. Die Untersuchung kürzester direkter Pfade innerhalb eines Netzwerkes ist ein Bereich der angewandten Graphentheorie (Newman, 2010) und spielt beispielsweise in der Modellierung von Krankheitsprozessen (z.B. Danon et al., 2011) eine wesentliche Rolle. Die Länge der Pfade durch ein Netzwerk stellt weiterhin die Grundlage für das Zentralitätsmaß der *Closeness* dar. Diese spiegelt die durchschnittliche Nähe (als Inverse der durchschnittlichen Distanz über die jeweiligen Pfade) einer Node zu allen anderen im Netzwerk abgebildeten Nodes wider (Newman, 2010). Hierdurch ergeben sich Interpretationen des Einflusses einer Node auf das gesamte Netzwerk und anders herum, beispielsweise dass eine Node mit einer hohen *Closeness* im Schnitt am direktesten von Veränderungen im Netzwerk betroffen ist und Veränderung an dieser Node am direktesten alle anderen Knoten des Netzwerkes beeinflussen (Hevey, 2018). Es ist allerdings anzumerken, dass die *Closeness* aufgrund ihrer mathematischen Eigenschaften insbesondere im Kontext der „Wichtigkeit“ einer Node für das Netzwerk schwierig zu interpretieren ist. Dies ergibt sich aus dem sehr kleinen Wertebereich dieser Metrik für das jeweils untersuchte Netzwerk. So stellte Newman (2010) exemplarisch die *Closeness* von FilmschauspielerInnen (im Rahmen eines Netzwerkes, das die jeweiligen Zusammenarbeiten abbildete) dar, wobei er zu dem Ergebnis kam, dass der Schauspieler Christopher Lee die niedrigste *Closeness* mit einem Wert von 2.4138 und die iranische Schauspielerin Leila Zanganeh die höchste *Closeness* mit einem Wert von 8.6681 aufwies. Zwischen diesen beiden Werten lagen die Werte von circa einer halben Millionen anderen SchauspielerInnen, woraus klar hervorgeht, dass eine Differenzierung einzelner Werte oft erst im Bereich mehrerer Nachkommastellen möglich wird. Da diese oft schon durch kleine Veränderungen des Netzwerkes, wie etwa dem Hinzufügen oder Entfernen einzelner Nodes oder aber auch Edges, verändert werden, lässt sich die *Closeness* als volatil beschreiben, was ihre Interpretierbarkeit einschränkt (Newman, 2010).

Die Konnektivität.

Eine weitere inhärente Eigenschaft der Netzwerke, die sie insbesondere im Kontext der klinischen Psychologie interessant macht, besteht in ihrer Konnektivität. Mathematisch betrachtet stellt diese die Summe der Beträge aller Edges eines Netzwerkes dar. Hierdurch gibt sie auf globaler Ebene Aufschluss darüber, inwiefern Aktivität im Netzwerk propagiert wird, womit sie eine mögliche Verknüpfung zwischen der Struktur und der Dynamik eines Netzwerks bietet.

Auf konzeptioneller Ebene gibt es mehrere Ansätze zur klinisch-inhaltlichen Interpretation der Konnektivität psychopathologischer Netzwerke. So sind beispielsweise Parallelen zwischen dem Konzept der Selbst-Aufrechterhaltung eines psychopathologischen Netzwerks und der Konnektivität erkennbar. Dies ergibt sich im Wesentlichen durch die stärkere Propagierung von Aktivität in hoch konnektiven Netzwerken im Vergleich zu schwach konnektiven Netzwerken sowie durch die Bildung von Feedback-Schleifen. Ein inhaltliches Beispiel kann hier in den *Circulus-vitiosus-Modellen* gesehen werden. Solche Teufelskreismodelle finden bereits seit längerer Zeit Anwendung in der Verhaltenstherapie (Kornberger, 2017), wo sie einerseits Prozesse spezifischer psychischer Störungen beschreiben, etwa im psychophysiologischen Modell der Panikstörung (Margraf & Ehlers, 1989, zitiert nach Margraf, 2009). Andererseits können sie jedoch auch mit transdiagnostischen Inhalten der Psychopathologie, etwa im Bereich der lernpsychologischen Grundlagen mit dem Konzept der operanten Konditionierung (Rinck & Becker, 2011) in Verbindung gebracht werden.

Ein weiteres Beispiel für eine inhaltliche Interpretation ergibt sich vor dem Hintergrund der im Bereich der dynamischen Systeme untersuchten *Phasentransitionen*. Hierbei werden voneinander abgrenzbare Phasen angenommen, zwischen denen sich komplexe Systeme hin- und herbewegen (Fried et al., 2017). Ein Beispiel für eine Anwendung im psychopathologischen Kontext lieferte Borsboom (2017), indem er die Ätiologie einer hypothetischen psychischen Störung anhand von vier Phasen darstellte. In diesen Phasen kam es in seiner Darstellung, ausgehend von einem externalen Faktor, zu einer schrittweisen Aktivierung der Symptome und damit einer Ausbreitung von Aktivität. Insbesondere in der letzten der skizzierten Phasen, in der das Symptomnetzwerk aktiv ist und sich durch Selbst-Aufrechterhaltung in einem stabilen Zustand befindet, identifizierte Borsboom (2017) inhaltliche Parallelen zu den Teufelskreismodellen der klinischen Psychologie und Psychiatrie. Die Konnektivität der Netzwerke spielt hier, so Borsboom (2017), eine entscheidende Rolle einerseits für die Ausbreitung der Symptomaktivierung, andererseits für

die Annahme eines stabilen Zustandes, in dem sich die Symptome gegenseitig aufrechterhalten und verstärken. Hieraus leitete Borsboom (2017) weiterhin erste Ansätze eines Versuchs der Definition von psychischer Gesundheit und Krankheit in Begriffen der Netzwerktheorie ab. So beschrieb er psychische Gesundheit als stabilen Zustand eines niedrig konnektiven Symptomnetzwerks. Dieser Zustand ist nicht von der An- oder Abwesenheit einzelner Symptome abhängig, sondern beschreibt einen Gleichgewichtszustand, zu dem ein „gesundes System“ zurückkehrt, sollte es durch äußere Einflüsse beeinflusst werden. Parallel hierzu sieht Borsboom (2017) psychische Störungen als ebenfalls stabilen Zustand eines hoch konnektiven Netzwerkes an, in dem die einzelnen Symptome stark ausgeprägt sind. Ein Beispiel, das diese Definitionen und die darin enthaltenen Implikationen illustriert, kann in der Unterscheidung zwischen einer schweren Trauerreaktion und einer Depression gesehen werden (Borsboom, 2017). So kann eine Trauerreaktion zwar eine depressive Symptomatik auslösen, allerdings besteht der zentrale Unterschied, so Borsboom (2017), in der Rückkehr zum Ursprungszustand bei einer Reduktion des Einflusses des externen Faktors. Im Gegensatz hierzu ist die Depression durch eine Selbst-Aufrechterhaltung der Symptome auch nach Wegfallen der ursprünglich auslösenden Faktoren gekennzeichnet. Die beiden Ausprägungsformen des hier dargestellten Netzwerks lassen sich also in Begriffen der Systemtheorie als Attraktoren darstellen. Weiterhin lassen sich in dieser Konzeption von psychischer Gesundheit und Krankheit klare Bezüge zwischen der Konnektivität eines Netzwerkes zu den Konzepten der Vulnerabilität und der Resilienz erkennen.

Die empirische Untersuchung der Konnektivität von Netzwerken im klinisch-psychologischen Bereich lieferte bisher jedoch uneindeutige Ergebnisse. So fanden van Borkulo et al. (2015) zwar eine signifikant geringere Konnektivität von Personen, deren depressive Symptomatik im Folgenden remittierte, im Vergleich zu Personen, deren Symptomatik persistierte. Allerdings konnte dieser Befund in einer Untersuchung von Schweren, van Borkulo, Fried und Goodyer (2018) hinsichtlich der Eignung der Konnektivität für Prognosen hinsichtlich einer depressiven Symptomatik nicht repliziert werden. Bezüglich des Zusammenhangs zwischen der Konnektivität und dem Krankheitsstatus weisen erste Studien jedoch auf einen positiven Zusammenhang hin, der im Einklang mit Borsbooms Theorie (Borsboom, 2017) zu sein scheint. So fanden Wigman et al. (2015) Unterschiede hinsichtlich der Netzwerkstrukturen von Personen ohne klinische Diagnose, Personen mit psychotischen Störungen sowie Personen mit depressiven Störungen. Diese Unterschiede konnten jedoch leider nicht mit inferenzstatistischen Methoden abgesichert werden, weshalb sie lediglich als leichte Hinweise gesehen werden können. Pe et al. (2015) hingegen

untersuchten mittels *Experience-Sampling-Methode* (ESM) die zeitliche Dynamik von Emotionen bei Personen mit depressiver Symptomatik und einer gesunden Kontrollgruppe. Hierzu erhoben sie acht Mal täglich die aktuellen affektiven Zustände ihrer Versuchspersonen und modellierten die zeitlichen Verläufe in Form temporärer Netzwerke. Hierbei kamen sie zu dem Ergebnis, dass die Emotionsnetzwerke von Personen mit depressiver Symptomatik signifikant stärker konnektiv waren als die der Kontrollgruppe. Weiterhin konnten sie in Follow-up-Analysen zeigen, dass diese Unterschiede wohl in erster Linie in Unterschiede der Konnektivität der Netzwerke der negativen Emotionen zurückzuführen sind, hinsichtlich der Verbindungen zwischen den positiven Emotionen unterschieden sich die beiden Gruppen nicht. Diese Ergebnisse legen eine mögliche Rolle der Konnektivität bei der Verknüpfung zwischen der Struktur und der Dynamik eines Netzwerks nahe.

Der Impact einer Variable.

Ein weiteres im Rahmen der Netzwerkanalysen untersuchbares Konstrukt ist der *Impact*. Der Impact spiegelt wider, inwiefern eine einzelne Node eines Netzwerks die globale Konnektivität des Netzwerks beeinflusst (Jones, 2017). Hiermit kann er als Verknüpfung der Mikro- und der Makroebene der Netzwerkanalyse angesehen werden. Die Konzeption und die Implementierung des Impacts in den Methodenkanon der Netzwerkanalysen geht maßgeblich auf Payton Jones (2017) zurück. Die Berechnung des Impacts erfolgt nach einem eingänglichen Rational, das erstmalig im Rahmen einer Simulationsstudie auf die Analyse sozialer Netzwerke, also Netzwerke mit Personen als Nodes und Beziehungen als Edges, angewandt wurde (Jones, 2017). Um den Einfluss einer einzelnen Person auf das soziale Netz zu untersuchen, wurden zwei Netzwerke modelliert, eines mit und eines ohne die entsprechende Person, was einen Vergleich der beiden geschätzten Netzwerke hinsichtlich zentraler Eigenschaften ermöglichte. Anzumerken ist hier, dass das Entfernen einzelner Knoten teils mit gravierenden Veränderungen in den übrigen Netzwerkstrukturen einhergeht (Epskamp et al., 2018). Eine mögliche Erklärung hierfür kann insbesondere im Bereich der PMRF's in der Modellierung anhand der bedingten Abhängigkeiten und Unabhängigkeiten gesehen werden. Diese schlägt sich in der Verwendung von partiellen Korrelationen nieder, die, je nachdem welche Variablen herauspartialisiert wurden, zu stark unterschiedlichen Ergebnissen kommen können.

Auch die Untersuchung des Einflusses einer ordinalen oder metrischen Variable ist möglich, hierfür wurde ein Median-Split implementiert, anhand dem die untersuchte Stichprobe in die beiden Gruppen „höhere Ausprägung“ und „niedrigere Ausprägung“

aufgeteilt wird und die beiden hieraus geschätzten Netzwerke miteinander verglichen werden (Jones, 2017).

Methodik

Untersuchungsgegenstand

Zur Überprüfung der im Folgenden beschriebenen Hypothesen wurden bereits erhobene Daten des Forschungsprojektes „fit2work“ analysiert. Dieses Projekt besteht seit Juni 2013 und wird durch das Bundesministerium für Arbeit, Soziales, Gesundheit und Konsumentenschutz (BMASGK) gefördert (Jagsch, 2015). Ziel des Projekts ist zum einen die klinisch-psychologische und psychotherapeutische Unterstützung und Behandlung von Personen, deren Arbeitsplatz aus gesundheitlichen Gründen gefährdet ist. Zum anderen werden Personen, die ihre Erwerbstätigkeit bereits verloren haben oder die aufgrund ihrer gesundheitlichen Beeinträchtigung seit längerer Zeit nicht arbeiten konnten, bei der Reintegration in den Arbeitsmarkt unterstützt.

Zusätzlich zur unmittelbaren Unterstützung der Zielgruppe wird mit „fit2work“ weiterhin eine Struktur zur effizienten Vermittlung einer geeigneten Behandlung im Gesundheitssystem Österreichs implementiert. Hierfür wurde die Rolle der „steuernden PsychologInnen“ geschaffen (Jagsch, 2015), die die in Frage kommenden Personen im Rahmen eines Case-Managements bei der Erstellung und Umsetzung eines individuell zugeschnittenen Maßnahmenplans unterstützen.

Die im Zuge dieses Case-Managements gesetzten Maßnahmen umfassen zunächst eine Abklärung der individuellen Sachlage in Form eines Basisinformations-Gesprächs sowie einer Erstberatung durch die steuernden PsychologInnen. Bereits in der Erstberatung erfolgt eine Erhebung der jeweiligen Problemstellungen sowie der eventuell zur Verfügung stehenden Ressourcen (Weiner, 2015). Weiterhin erfolgt in diesem Schritt eine erste inhaltliche Beratung bezüglich einer möglichen weiteren Behandlung. Auch werden Hilfestellungen hinsichtlich organisatorischer Faktoren wie etwa dem Stellen von Anträgen auf Kostenübernahmen oder der Ausarbeitung eines Plans zur Wiedereingliederung mit den ArbeitnehmerInnen und ArbeitgeberInnen angeboten (Weiner, 2015).

Für den Fall, dass auf der Grundlage der in den ersten Schritten erhobenen Informationen eine weitere Beratung indiziert ist, bietet „fit2work“ im weiteren Verlauf des Beratungsangebots eine arbeitsmedizinische und/oder arbeitspsychologische Untersuchung der KlientInnen an. Im Zuge dieser werden der Gesundheitszustand der KlientInnen umfassend erfasst sowie hieraus weiterführende Maßnahmen abgeleitet. Die sich hieran

anschließende Beratung zielt einerseits auf eine weitere inhaltliche Klärung von Problemen und Möglichkeiten ab, andererseits wird auch hier eine strukturelle Unterstützung bei der Vermittlung bedarfsgerechter Maßnahmen angeboten (Weiner, 2015).

Im Zuge einer Evaluation von „fit2work“ wurde den TeilnehmerInnen zu Beginn und nach dem Abschluss ihrer Beratung eine Fragebogenbatterie vorgegeben. Diese setzte sich aus der deutschen Fassung des *Beck-Depression-Inventory II* (Hautzinger, Keller, & Kühner, 2006), dem *Brief Symptom Inventory* (Franke, 2000) sowie einem Fragebogen zum gesundheitsbezogenen Status zusammen (Jagsch, 2015). Nach Abschluss der Behandlung wurde sie durch den *Helping Alliance Questionnaire* (Bassler, Potratz, & Krauthauser, 1995) ergänzt. Weiterhin wurden den behandelnden klinischen PsychologInnen und PsychotherapeutInnen ebenfalls der *Helping Alliance Questionnaire* sowie eine Fremdbeurteilungsform des Fragebogens zum gesundheitsbezogenen Status per online-Befragung vorgegeben.

Untersuchungsinstrumente

Im Rahmen dieser Arbeit wurden die Symptomnetzwerke der TeilnehmerInnen des Projektes „fit2work“ anhand der Variablen des Brief Symptom Inventory (BSI-53) von Franke (2000) untersucht. Dieses stellt eine deutsche Kurzfassung der Symptom-Checkliste SCL-90-R dar (Jagsch, 2015), mittels der verschiedene Arten psychischer Belastungen erhoben werden können. Das BSI-53 umfasst 53 Items zur Selbstbeurteilung, die auf einer fünf-stufigen Likert-Skala (0 = „überhaupt nicht“ bis 4 = „sehr stark“) beantwortet werden. 49 dieser Items werden den neun Skalen *Somatisierung*, *Zwanghaftigkeit*, *Unsicherheit im Sozialkontakt*, *Depressivität*, *Ängstlichkeit*, *Aggressivität/Feindseligkeit*, *Phobische Angst*, *Paranoides Denken* und *Psychotizismus* zugeordnet. Die übrigen vier Items stellen sogenannte Zusatzitems dar, die Hinweise auf unspezifische psychische Belastungen geben sollen. Weiterhin lassen sich drei Kennwerte berechnen, mittels derer die psychische Belastung global dargestellt werden kann. Der *Global Severity Index* (GSI) stellt hier die globale Belastung dar, der *Positive Symptom Distress Index* (PSDI) die Intensität der Belastung in Bezug auf die tatsächlich vorhandenen Symptome und der *Positive Symptom Total* (PST) die Anzahl der vorliegenden Symptome (Franke, 2000). Hervorzuheben ist hier für diese Arbeit der Global Severity Index, der berechnet wird, indem die Summe der Summenwerte pro Skala durch die Anzahl der beantworteten Items dividiert wird.

Für den BSI liegen Normwerte für Erwachsene und Studierende vor. Die internen Konsistenzwerte der Skalen wurden von Derogatis (1993, zitiert nach Franke, 2000) in einer Stichprobe von $N = 719$ ambulanten PatientInnen ermittelt und reichen in dieser Gruppe von

Cronbachs's $\alpha = .71$ bis $\alpha = .84$ (Cronbachs- α -Werte ab $\alpha = .70$ gelten als akzeptabel, jene ab $\alpha = .80$ als gut (Field, 2009)). Für den GSI wurde in der Normstichprobe gesunder Erwachsener ($N = 600$) ein interner Konsistenzwert von $\alpha = .92$ gefunden (Franke, 2000). In einer Teilstichprobe dieser Normstichprobe ($n = 70$) wurden Test-Retest-Reliabilitätswerte der einzelnen BSI-Skalen ermittelt, die von $r = .73$ bis $r = .93$ reichen. Für den GSI lag in dieser Teilstichprobe eine Test-Retest-Reliabilität von $r = .93$ vor.

Eine inhaltliche Interpretation des BSI-53 wird laut Franke (2000) über die Berechnung von T-Werten ermöglicht. So gilt eine Person als psychisch auffällig belastet, wenn sie einen T-Wert im GSI von $T_{\text{GSI}} \geq 63$, oder aber auf zwei der Subskalen einen T-Wert von $T_{\text{Skala}} \geq 63$ aufweist.

Fragestellungen und Hypothesen

Da einerseits die Anwendung von netzwerkanalytischen Methoden im Bereich der klinischen Psychologie noch recht wenig verbreitet ist (Borsboom, 2017; Hevey, 2018) sowie andererseits komplexe Netzwerkstrukturen mit einer Vielzahl an möglichen Zusammenhängen und Interaktionen untersucht werden sollten, wurde zunächst eine explorative querschnittliche Untersuchung der Netzwerkstrukturen der psychischen Belastung einer Stichprobe von TeilnehmerInnen des Projektes „fit2work“ durchgeführt.

Die Zusammenhänge zwischen der Konnektivität psychologischer Netzwerke und verschiedenen theoretischen Konstrukten, wie beispielsweise der Vulnerabilität oder aber auch dem Krankheitsstatus, ist bislang noch wenig erforscht. Der theoretische Hintergrund dieser postulierten Zusammenhänge besteht in der Arbeit Borsbooms (2017). Entsprechend der hier vorgenommenen Konzeption von psychischen Krankheiten als Interaktion zwischen einerseits einer hohen Symptomausprägung und andererseits einer hohen Konnektivität des Symptomnetzwerks soll hier untersucht werden, ob die Konnektivität der BSI-53-Symptomnetzwerke signifikant abnahm. Diese Fragestellung ergibt sich auch vor dem Hintergrund, dass in bisherigen Evaluationen signifikante Verbesserungen auf allen Skalen des BSI-53 im Zeitraum der Teilnahme an „fit2work“ gefunden werden konnten (Jagsch, 2015). Hieraus ergab sich das folgende Hypothesenpaar:

Hypothese 1: Die Konnektivität nimmt von t_1 zu t_2 signifikant ab.

H_0 : Konnektivität $t_1 \leq$ Konnektivität t_2

H_1 : Konnektivität $t_1 >$ Konnektivität t_2

Weiterhin wurde untersucht, inwiefern es eine Äquivalenz zwischen der Konnektivität der BSI-53-Netzwerke und dem Ansprechen auf die Beratung und Behandlung im Rahmen

von „fit2work“ gab. Hierfür wurden die Substichproben der Responder und der Non-Responder gebildet und hinsichtlich ihrer Konnektivität zu t_1 miteinander verglichen. Da die Befundlage hierzu bisher noch uneindeutig ist (Schweren et al., 2018; van Borkulo et al., 2015) wurde das folgende ungerichtete Hypothesenpaar aufgestellt:

Hypothese 2: Die Netzwerke von Personen, deren Symptomatik sich verbessert, und solchen, deren Symptomatik sich verschlechtert oder gleich bleibt, unterscheiden sich hinsichtlich ihrer Konnektivität.

$$H_0: \text{Konnektivität}_{\text{Responder}} = \text{Konnektivität}_{\text{Non-Responder}}$$

$$H_1: \text{Konnektivität}_{\text{Responder}} \neq \text{Konnektivität}_{\text{Non-Responder}}$$

Untersuchungsdesign und Stichprobenbeschreibung

Im Zuge der explorativen Untersuchung sowie der statistischen Hypothesentestung wurden Daten der Evaluation des Projektes „fit2work“ analysiert. Hierfür wurden, entsprechend den statistischen Hypothesen, vier separate Datensätze für unterschiedlich zusammengestellte Substichproben gebildet, deren Erstellung und Deskriptivstatistiken im Folgenden beschrieben werden sollen.

Die ursprüngliche Gesamtstichprobe umfasst $N = 2143$ Personen, die ihre Daten für die Evaluation von „fit2work“ zur Verfügung gestellt haben. Diese Gesamtstichprobe setzt sich aus $n = 1290$ (60.2%) weiblichen Teilnehmerinnen und $n = 714$ (33.3%) männlichen Teilnehmern zusammen. Die Abweichung dieser Werte von der Gesamtanzahl der TeilnehmerInnen ergibt sich aus dem Umstand, dass nicht alle Personen vollständige Angaben zu ihren demografischen Daten gemacht haben. Das durchschnittliche Alter zu Beginn der Beratung beträgt $M = 44.13$ ($SD = 9.80$, Spannweite = 18-67).

Den Vorgaben des Testmanuals des Brief Symptom Inventory (Franke, 2000) folgend wurden zunächst alle Personen von der statistischen Analyse ausgeschlossen, die auf einer der verwendeten Skalen mehr als eine fehlende Antwort aufwiesen. Dieser Wert wird von Franke (2000) als Maximalgrenze für fehlende Werte angegeben, bei der eine Interpretation der Ergebnisse noch sinnvoll möglich ist. Auch wurden Personen von der Analyse ausgeschlossen, von denen Daten entweder nur zum ersten oder nur zum zweiten Testzeitpunkt vorlagen. Dies liegt darin begründet, dass die statistische Testung der Hypothesen mittels eines Permutationstests, dem *Network-Comparison-Test*, vorgenommen wurde. Mit diesem können Netzwerkstrukturen hinsichtlich mehrerer Merkmale miteinander verglichen werden, allerdings benötigt er für den Vergleich abhängiger Stichproben vollständige Datensätze, sowie möglichst gleiche Stichprobengrößen.

Weiterhin wurden einzelne fehlende Werte, sofern sie nicht bereits zu einem Ausschluss von der Analyse geführt hatten, mittels dem auf der *Maximum-Likelihood-Methode* basierenden *Expectation-Maximization-Algorithmus* (EM-Algorithmus) imputiert. Die Notwendigkeit für eine Imputation ergibt sich aus der Art und Weise der Schätzung der untersuchten Netzwerke. Diese können im Wesentlichen als Darstellung partieller Korrelationen angesehen werden, bei denen alle anderen interessierenden Variablen zugleich kontrolliert werden. Dementsprechend würde ein paarweiser Ausschluss fehlender Werte bei der Berechnung der partiellen Korrelationen dazu führen, dass die Daten von Personen mit einzelnen fehlenden Werten nicht in die Berechnung einfließen können. Für die Imputation wurde der EM-Algorithmus verwendet, da er einerseits an weniger strenge Annahmen gebunden ist als andere Schätzmethoden und andererseits valide Ergebnisse hinsichtlich der Lagemaße, Varianzen sowie Kovarianzen einer Stichprobe ermöglicht (Allison, 2002, zitiert nach Baltes-Götz, 2013). Nachteilig wirkt sich die Verwendung des EM-Algorithmus auf die Schätzung der Standardfehler aus, die tendenziell unterschätzt werden. Da jedoch die Schätzung der Netzwerke nicht auf „klassischen“ inferenzstatistischen Verfahren beruht, sondern Algorithmen aus dem Bereich des Machine-Learning (in Form der Regularisierung mittels LASSO) und den Bayes'schen Verfahren (in Form der Verwendung des EBIC's) heranzieht und somit ein Einfluss der Verzerrung der Standardfehler als gering angesehen werden kann, wurde diese Methode gewählt.

Die sich aus diesen Schritten ergebende Stichprobe umfasst insgesamt $N = 664$ Personen, deren Daten für die Berechnungen der Symptomnetzwerke vor und nach der jeweiligen Behandlung verwendet wurden. Von diesen 664 Personen sind 424 weiblich (63.9%), das durchschnittliche Alter dieser Gruppe beträgt $M = 44.82$ Jahre ($SD = 9.51$, Spannweite = 19-59 Jahre). 240 (36.1%) der Personen sind männlich, hier beträgt das durchschnittliche Alter $M = 44.85$ Jahre ($SD = 9.61$, Spannweite = 19-61). Die Verteilungen der Familienstände sowie der höchsten abgeschlossenen Ausbildung finden sich in Tabelle 1.

Für die Testung der zweiten Hypothese wurden weiterhin zwei Substichproben gebildet, anhand derer untersucht werden sollte, ob, und wenn ja inwiefern, sich Personen mit einer Verbesserung der Symptomatik von Personen mit einer Verschlechterung der Symptomatik hinsichtlich der Eigenschaften der jeweiligen Symptomnetzwerke unterscheiden.

Tabelle 1

Soziodemografische Merkmale der Analytestichprobe

		Häufigkeit	Prozent
Familienstand	ledig	222	33.4
	verheiratet	290	43.6
	geschieden	140	21.1
	verwitwet	7	1.1
	keine Angabe	5	0.8
Höchste abgeschlossene Ausbildung	keine Ausbildung abgeschlossen	6	0.9
	Pflichtschule	76	11.4
	Berufsschule/Lehre	275	41.5
	Berufsbildende mittlere Schule	109	16.4
	AHS/BHS	79	11.9
	Kurzstudium, Spezialausbildung	65	9.8
	Universität/FH	52	7.8
	Keine Angabe	2	0.3

Anmerkung. $N = 664$.

Die hierfür notwendige Unterteilung wurde über das Globalmaß GSI operationalisiert, indem dieses zunächst für beide Messzeitpunkte berechnet wurde. Im Folgenden wurde eine Differenzvariable berechnet, indem die GSI-Werte zum ersten Erhebungszeitpunkt von denen zum zweiten Erhebungszeitpunkt subtrahiert wurden. Hierdurch erfolgt eine einfache Darstellung der Veränderung im Globalscore, wobei negative Werte einer Verbesserung und positive Werte einer Verschlechterung entsprechen. Die Substichprobe der Non-Responder wurde aus allen Personen gebildet, die sich hinsichtlich ihres GSI-Wertes nicht veränderten oder verschlechterten. Die Substichprobe der Responder wurde gebildet, indem die 20 Prozent mit der stärksten Verbesserung hinsichtlich des GSI-Wertes ausgewählt wurden. Dieses Vorgehen liegt darin begründet, dass versucht wurde, einen Kompromiss zwischen einer inhaltlichen Bedeutsamkeit der Gruppenbildung (in Form einer Unterscheidung zwischen Verbesserung und Nicht-Verbesserung) einerseits und den Anforderungen der Analysemethoden (in Form gleicher Gruppengrößen) andererseits zu finden.

Die sich hieraus ergebende Stichprobe der Responder umfasst insgesamt $n = 133$ Personen. Hiervon sind 92 Personen weiblich (69.2%) und 41 männlich (30.8%). Das durchschnittliche Alter dieser Substichprobe beträgt $M = 42.53$ Jahre ($SD = 9.24$, Spannweite

= 20-59). Die Substichprobe der Non-Responder umfasst $n = 129$ Personen. Von diesen sind 79 Personen weiblich (61.2%) und 50 männlich (38.8%). Das Durchschnittsalter beträgt $M = 46.45$ Jahre ($SD = 8.27$, Spannweite = 20-58).

Statistische Verfahren

Sowohl für die explorative Untersuchung als auch für die Testung der aufgestellten Hypothesen wurden Verfahren aus dem Methodenkanon der Netzwerkanalysen verwendet. Die Daten für die netzwerkanalytischen Untersuchungen wurden mit Hilfe des R-Packages `foreign` (R Core Team, 2018) in R importiert. Aus den Rohdaten wurden zunächst polychorische Korrelationsmatrizen gebildet, aus denen im Anschluss mit dem R-Package `qgraph` (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012) die Netzwerkstrukturen geschätzt wurden. Hierfür wurde einheitlich die gLASSO-Regularisierung verwendet, der Parameter γ des hierbei verwendeten EBIC wurde durchgängig auf $\gamma = 0.50$ gesetzt. Weiterhin wurde die Stabilität der sich hieraus ergebenden Netzwerke mit Hilfe des R-Packages `bootnet` (Epskamp et al., 2018) untersucht. Hierbei wurden einerseits die Stabilität der Edges über die Berechnung von 95%-Konfidenzintervallen, andererseits die Correlation-Stability-Coefficients berechnet. In beiden Fällen wurden die Berechnungen anhand $m = 1000$ Bootstrap-Stichproben durchgeführt. Die Zentralitätsanalysen wurden ebenfalls mit Funktionen aus dem Package `qgraph` (Epskamp et al., 2012) durchgeführt. Weiterhin wurde untersucht, ob die Berechnungen der Zentralitäten sowie der Impact-Werte in einem Zusammenhang mit den Varianzen der jeweiligen Items stehen. Hierfür wurden über das R-Basispaket die Varianzen der Items berechnet. Deren Verteilungen sowie die der netzwerkanalytischen Kennwerte wurden mittels Shapiro-Wilk-Tests auf Abweichungen von der Normalverteilung überprüft. Die Korrelationen der Kennwerte der einzelnen Items mit ihren jeweiligen Varianzen wurden zu t_1 mittels Spearman-Rangkorrelationen, zu t_2 mittels Pearson-Korrelationen durchgeführt. Die Testung der aufgestellten Hypothesen erfolgte mit Hilfe des R-Packages `Network Comparison Test` (van Borkulo et al., 2017). Anhand des hierin implementierten Permutationstests wurden 100 Iterationen durchgeführt. Die Veränderungen der Globalscores GSI zwischen t_1 und t_2 wurden mittels eines Wilcoxon-Vorzeichen-Rangtests durchgeführt. Dies liegt darin begründet, dass im Vorfeld durchgeführte Shapiro-Wilk-Tests und eine visuelle Untersuchung der Verteilungen der GSI-Werte ergaben, dass diese signifikant von Normalverteilungen abwichen. Die Berechnungen der Impact-Werte wurde mit Hilfe des R-Packages `NetworkTools` (Jones, 2019) durchgeführt. Hierfür wurden zunächst einfache Impact-Berechnungen durchgeführt. Aus im Folgenden eingehend beschriebenen Gründen wurde

eine alternative Berechnungsmethode der Impact-Werte konzipiert, die auf händischen Berechnungen basierte. Weiterhin wurden Impact-Berechnungen, die durch Bootstrap-Verfahren stabilisiert wurden, durchgeführt. Hierbei wurde, aufgrund der hohen Rechenlast, auf die Informationen von $m = 250$ Bootstrap-Ziehungen zurückgegriffen. Die Übersichtsgrafiken dieser Arbeit wurden mit ggplot2 (Wickham, 2016) erstellt, um die Grafiken in höherer Qualität als PDF-Dokumente exportieren zu können, wurde das R-Package Cairo (Urbanek & Horner, 2019) verwendet.

Ergebnisdarstellung

Im Folgenden werden die Ergebnisse der netzwerkanalytischen Untersuchung der BSI-53-Symptomnetzwerke (im Folgenden: Symptomnetzwerke) einer Stichprobe der KlientInnen von „fit2work“ dargestellt. Diese Darstellung gliedert sich, entsprechend den Hypothesen, in zwei Teile. Zunächst sollen die Symptomnetzwerke der Gesamtstichprobe vor und nach der Behandlung vorgestellt werden. Im Anschluss erfolgt die Darstellung der Symptomnetzwerke der „guten Responder“ und der „schlechten Responder“ vor der Behandlung.

Symptomnetzwerke vor und nach der Behandlung

Vor der Darstellung der Symptomnetzwerke selbst sollen die Ergebnisse der Untersuchungen hinsichtlich der Akkuratheit und Stabilität ihrer Schätzungen vorgestellt werden. Dieser Schritt ermöglicht nicht zuletzt eine Einschätzung, inwiefern die Ergebnisse der Netzwerkanalysen reliabel und robust sind, ein Vorgehen, das insbesondere vor dem Hintergrund der Replikationskrise in der Psychologie (Open Science Collaboration, 2015) angemessen erscheint.

Die Berechnungen der CS-Koeffizienten erfolgten jeweils für ein Symptomnetzwerk der BSI-53-Items zu t_1 und zu t_2 auf Basis von Bootstrap-Berechnungen mit je $m = 1000$ Ziehungen. Untersucht wurde hiermit jeweils die Stabilität der Edges sowie der *Node-Strengths*. Die Ergebnisse der Koeffizienten legen nahe, dass zu t_1 die Schätzungen der Edges mit $CS(cor = .70) = .75$ im sehr guten, die Schätzungen der Node-Strengths mit $CS(cor = .70) = .67$ im guten Bereich der Stabilität liegen. Zu t_2 wurden die Edges ebenfalls mit $CS(cor = .70) = .75$ sehr stabil, die Node-Strengths mit $CS(cor = .70) = .52$ stabil geschätzt. Insgesamt lassen sich die Schätzungen als stabil und damit als gut interpretierbar (Epskamp et al., 2018) beschreiben.

Weiterhin wurden für die Schätzungen der Edges mittels Bootstrap-Verfahren (mit je $m = 1000$ Ziehungen) Konfidenzintervalle berechnet. Diese finden sich anbei der jeweiligen

Darstellung der Edges. Eine grafische Darstellung derjenigen gebootstrapteten Konfidenzintervalle, die nicht den Wert Null enthalten, finden sich in den Abbildungen 1 und 2.

Die geschätzten Netzwerke zu t_1 und t_2 sind in den Abbildungen 3 und 4 dargestellt. Anzumerken ist, dass hier lediglich alle Edges, die $r \geq .10$ sind, abgebildet wurden. Die Darstellungen sämtlicher Zusammenhänge finden sich im Anhang (Abb. 5 und Abb. 6). Weiterhin findet sich in Tabelle 2 und Tabelle 3 eine Aufstellung der einzelnen Edges, die $r \geq .30$ sind, also gemäß Cohen zumindest mittlere Effektstärken repräsentieren (Cohen, 1988). Tabellarische Darstellungen aller Edges, die $r \geq .10$, sind finden sich im Anhang (Tabelle 4 und Tabelle 5). Auf Darstellungen aller Edges wurde aufgrund des Umfangs (1431 korrelative Zusammenhänge) sowie des als eher gering einzuschätzenden Informationsgehalts der Verknüpfungen mit $r < .10$ verzichtet.

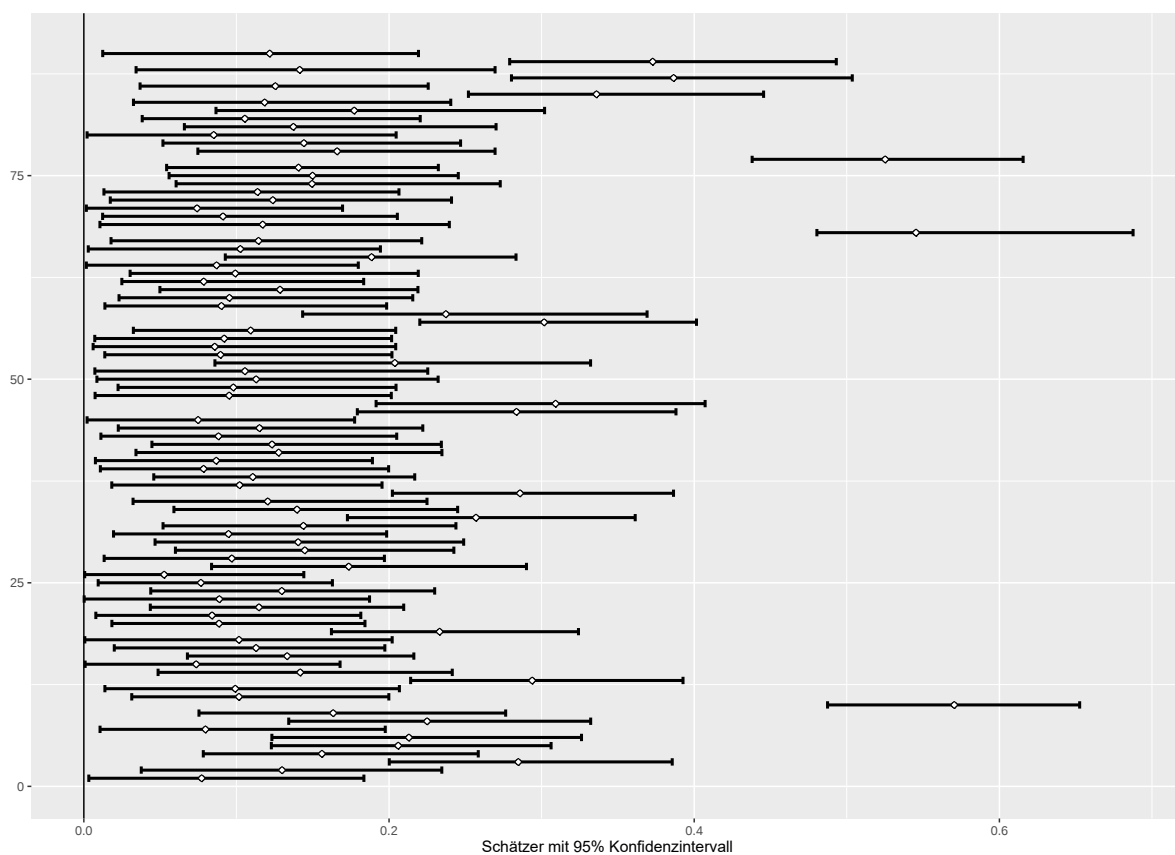


Abbildung 1. Bootstrap-Schätzungen zu t_1 , deren Konfidenzintervalle nicht Null enthalten.

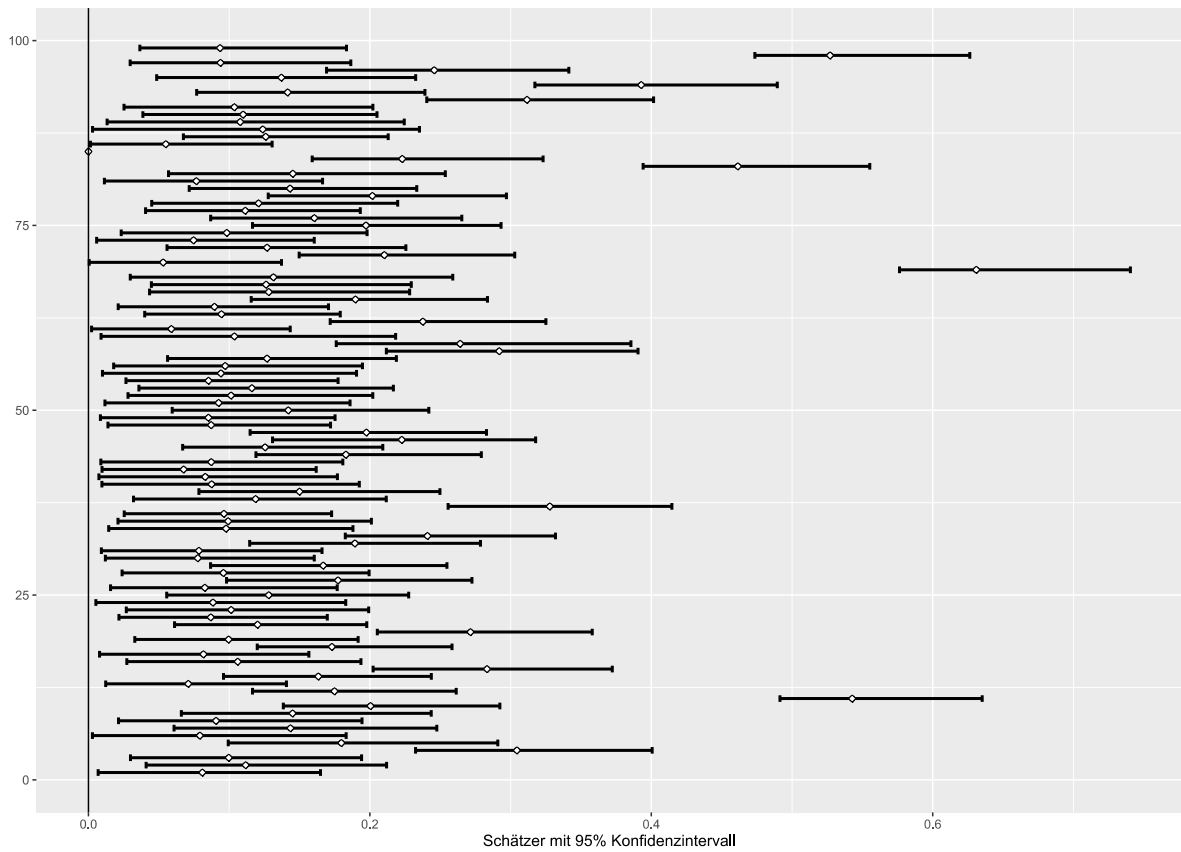


Abbildung 2. Bootstrap-Schätzungen zu t_2 , deren Konfidenzintervalle nicht Null enthalten.

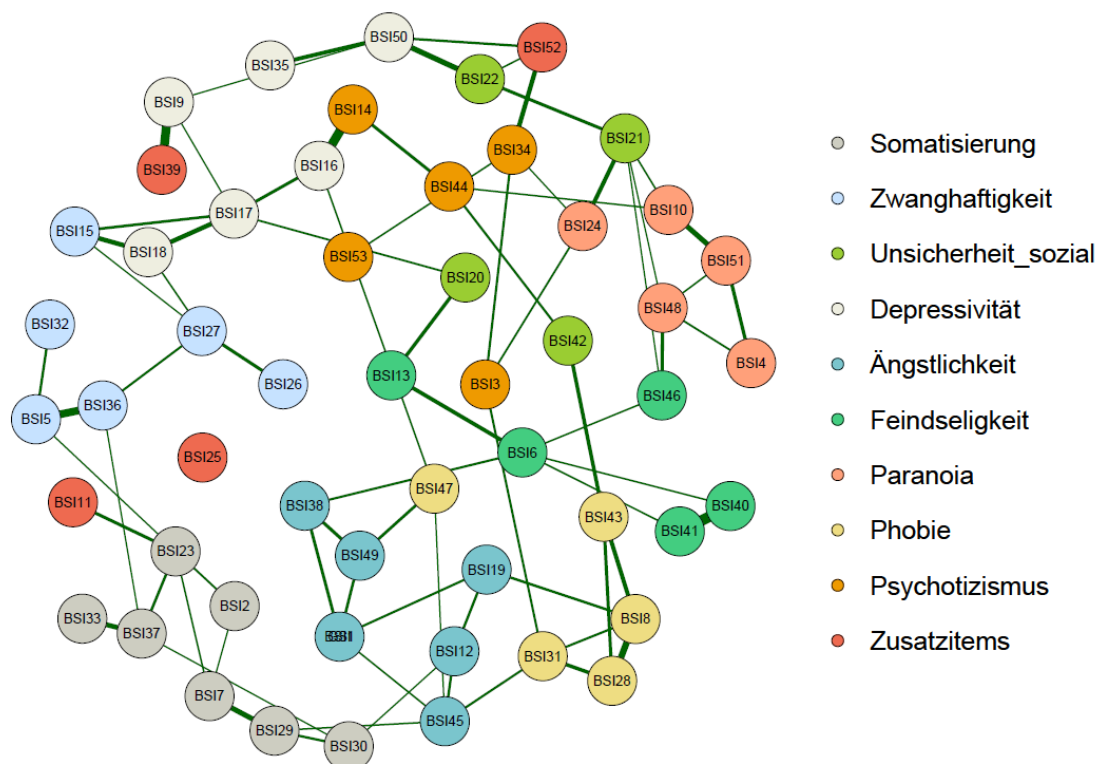


Abbildung 3. Symptomnetzwerk des BSI-53 zu t_1 .

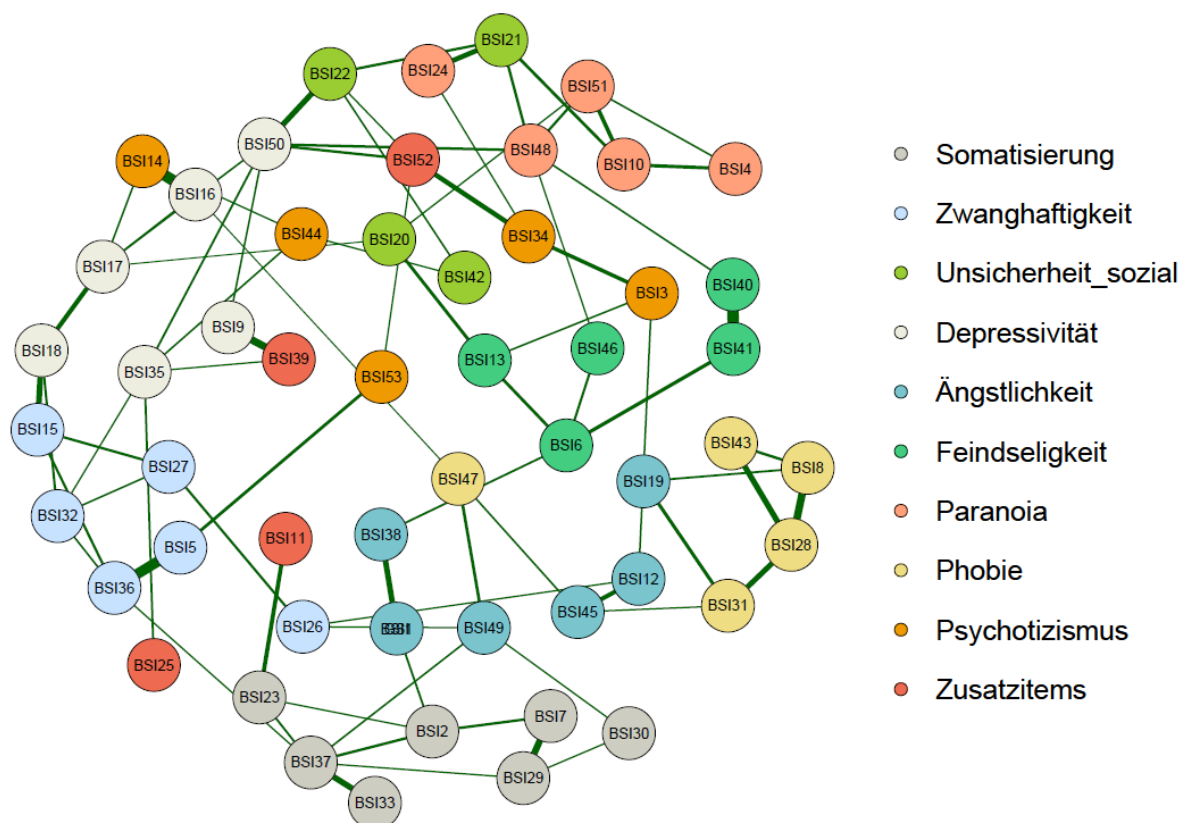


Abbildung 4. Symptomnetzwerk des BSI-53 zu t_2 .

Zentralitäten

Auf Itemniveau wurden Analysen hinsichtlich der Zentralitätskennwerte der Node-Strength, der Betweenness sowie der Closeness durchgeführt. Die Ergebnisse dieser Analyse für beide Messzeitpunkte finden sich in Tabelle 6. Weiterhin wurde untersucht, inwiefern die Zentralitätsmaße der Items mit den jeweiligen Varianzen linear zusammenhängen. Hierfür erfolgte zunächst eine Überprüfung der Annahme der Normalverteilung für die drei Zentralitätsmaße jeweils für t_1 und t_2 mittels Shapiro-Wilk-Test. Diese ergab, dass zu t_1 die Verteilung der Node-Strengths nicht signifikant von einer Normalverteilung abwich ($W = .98$, $p = .342$), die der Betweenness ($W = .94$, $p = .014$) sowie der Closeness ($W = .94$, $p = .011$) jedoch schon. Weiterhin wick die Verteilung der Itemvarianzen zu t_1 signifikant von einer Normalverteilung ab ($W = .94$, $p = .020$). Zu t_2 wick keine der Verteilungen der Zentralitätsmaße statistisch signifikant von einer Normalverteilung ab (Node-Strength: $W = .98$, $p = .649$; Betweenness: $W = .97$, $p = .123$; Closeness: $W = .98$, $p = .498$). Auch die Verteilung der Itemvarianzen zu t_2 wick nicht statistisch signifikant von der Normalverteilung ab ($W = .96$, $p = .054$).

Tabelle 2

Darstellung der Edges zu t_1 mit $r \geq .30$

Itemnr.	Itemformulierung	r	SD	95% KI
40	Der Drang, jemanden zu schlagen, zu verletzen oder ihm Schmerz zuzufügen	.66	.04	[.58, .74]
41	Der Drang, Dinge zu zerbrechen oder zu zerschmettern			
14	Einsamkeitsgefühle, selbst wenn Sie in Gesellschaft sind	.56	.04	[.50, .64]
16	Einsamkeitsgefühle			
9	Gedanken, sich das Leben zu nehmen	.55	.04	[.47, .63]
39	Gedanken an den Tod und ans Sterben			
5	Gedächtnisschwierigkeiten	.48	.04	[.40, .56]
36	Konzentrationsschwierigkeiten			
8	Furcht auf offenen Plätzen oder auf der Straße	.40	.04	[.32, .49]
28	Furcht vor Fahrten in Bus, Straßenbahn, U-Bahn oder Zug			
22	Minderwertigkeitsgefühle gegenüber anderen	.34	.04	[.26, .43]
50	Das Gefühl, wertlos zu sein			
7	Herz- oder Brustschmerzen	.32	.04	[.24, .40]
29	Schwierigkeiten beim Atmen			
10	Das Gefühl, dass man den meisten Leuten nicht trauen kann	.32	.04	[.23, .40]
51	Das Gefühl, dass die Leute Sie ausnutzen, wenn Sie es zulassen würden			
33	Taubheit oder Kribbeln in einzelnen Körperteilen	.30	.05	[.21, .39]
37	Schwächegefühl in einzelnen Körperteilen			

Anmerkung. Die Items wurden dem BSI-53 entnommen; r = partielle polychorische Korrelation; SD = Standardabweichung; KI = Konfidenzintervall.

Tabelle 3

Darstellung der Edges zu t_2 mit $r \geq .30$

Itemnr.	Itemformulierung	r	SD	95% KI
40	Der Drang, jemanden zu schlagen, zu verletzen oder ihm Schmerz zuzufügen	.58	.05	[.48, .69]
41	Der Drang, Dinge zu zerbrechen oder zu zerschmettern			
14	Einsamkeitsgefühle, selbst wenn Sie in Gesellschaft sind	.57	.04	[.49, .65]
16	Einsamkeitsgefühle			
5	Gedächtnisschwierigkeiten	.53	.04	[.44, .62]
36	Konzentrationsschwierigkeiten			
8	Furcht auf offenen Plätzen oder auf der Straße	.39	.06	[.28, .50]
28	Furcht vor Fahrten in Bus, Straßenbahn, U-Bahn oder Zug			
9	Gedanken, sich das Leben zu nehmen	.39	.05	[.28, .49]
39	Gedanken an den Tod und ans Sterben			
7	Herz- oder Brustschmerzen	.35	.05	[.25, .45]
29	Schwierigkeiten beim Atmen			
33	Taubheit oder Kribbeln in einzelnen Körperteilen	.31	.05	[.22, .40]
37	Schwächegefühl in einzelnen Körperteilen			
15	Das Gefühl, dass es Ihnen schwerfällt, etwas anzufangen	.30	.04	[.21, .39]
18	Das Gefühl, sich für nichts zu interessieren			
28	Furcht vor Fahrten in Bus, Straßenbahn, U-Bahn oder Zug	.30	.05	[.19, .41]
43	Abneigung gegen Menschenmengen, z.B. beim Einkaufen oder im Kino			

Anmerkung. Die Items wurden dem BSI-53 entnommen; r = partielle polychorische Korrelation; SD = Standardabweichung; KI = Konfidenzintervall.

Aus diesem Grund wurden die im Folgenden berichteten Untersuchungen auf lineare Zusammenhänge für alle Zentralitätsmaße zu t_1 mittels Spearman-Rangkorrelationen durchgeführt, die Berechnungen zu t_2 mittels Pearson-Korrelationen.

Zu t_1 korrelierten weder die Node-Strength ($r(51) = .05$, $p = .717$) noch die Betweenness ($r_s(51) = -.02$, $p = .872$) oder die Closeness ($r_s(51) = .08$, $p = .553$) signifikant mit der jeweiligen Varianz der einzelnen Items. Auch zu t_2 war keine der Korrelationen der Zentralitätsmaße mit der Varianz auf dem korrigierten α -Fehlerniveau von $p < .008$ statistisch signifikant (Node-Strength: $r(51) = -.07$, $p = .644$; Betweenness: $r(51) = .01$, $p = .950$; Closeness: $r(51) = -.04$, $p = .800$), auch bei einem liberaleren α -Fehlerniveau von $p < .05$ erreichte keiner der Zusammenhänge statistische Signifikanz.

Auch wurde überprüft, inwiefern es Zusammenhänge zwischen den jeweiligen Zentralitätsmaßen zu t_1 und t_2 gibt. Hierfür wurden für die Node-Strength eine Pearson-Korrelation, für die Betweenness und die Closeness Spearman-Rangkorrelationen berechnet. Diese ergaben für den Zusammenhang zwischen den Node-Strengths zu t_1 und t_2 eine Korrelation von $r(51) = .85$, $p < .001$, für den Zusammenhang zwischen den Betweenness-Werten eine Korrelation von $r_s(51) = .35$, $p = .005$ sowie für den Zusammenhang zwischen den Closeness-Werten eine Korrelation von $r_s(51) = .39$, $p < .002$. Eine grafische Darstellung der Zentralitätsmaße zu t_1 und t_2 findet sich in Abbildung 7.

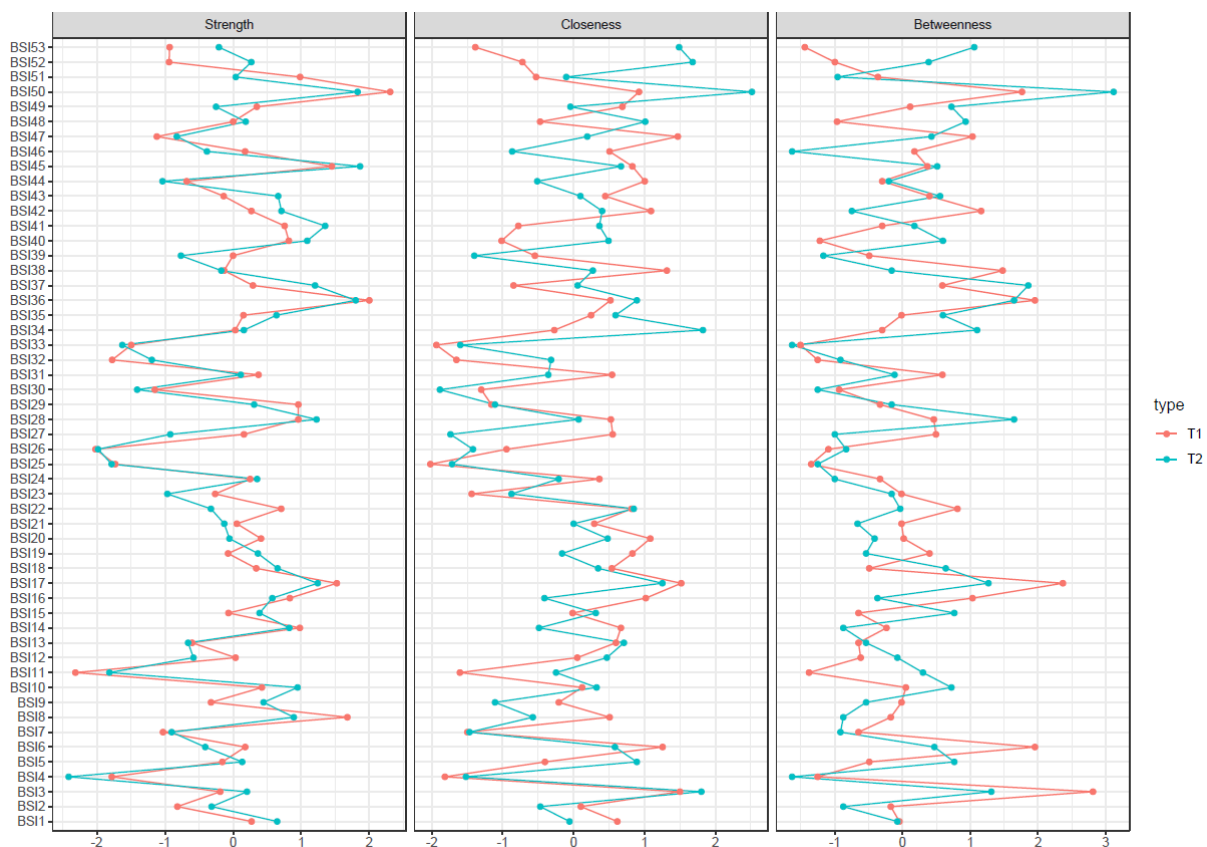


Abbildung 7. Grafische Darstellung der Zentralitätsmaße zu t_1 und t_2 .

Tabelle 6

Zentralitätsmaße des BSI-53-Netzwerks zu t_1 und t_2

Itemnr.	t_1			t_2		
	Strength	Betweenness	Closeness	Strength	Betweenness	Closeness
BSI1	1.03	92	0.000874099	1.08	82	0.000883944
BSI2	0.87	84	0.000843488	0.95	44	0.000861898
BSI3	0.96	272	0.000926669	1.02	148	0.000983387
BSI4	0.73	16	0.000729312	0.66	8	0.000805758
BSI5	0.97	64	0.000813346	1.01	122	0.000934686
BSI6	1.02	218	0.00091211	0.94	108	0.000918196
BSI7	0.84	54	0.000748204	0.87	42	0.000808532
BSI8	1.23	84	0.00086765	1.12	44	0.00085627
BSI9	0.94	94	0.000824994	1.05	60	0.000827738
BSI10	1.05	98	0.000844573	1.12	120	0.000904396
BSI11	0.65	8	0.000741945	0.74	100	0.000873787
BSI12	1.00	56	0.000840515	0.91	82	0.000912108
BSI13	0.90	54	0.000872955	0.90	60	0.000924963
BSI14	1.13	80	0.000877118	1.11	44	0.000860952
BSI15	0.98	54	0.000836584	1.05	122	0.000903764
BSI16	1.11	160	0.00089801	1.07	68	0.000865038
BSI17	1.21	244	0.00092775	1.16	146	0.000953998
BSI18	1.04	64	0.000869518	1.08	116	0.000905552
BSI19	0.98	120	0.000886789	1.04	60	0.00087837
BSI20	1.05	96	0.000901793	0.98	66	0.000912688
BSI21	1.00	94	0.000854754	0.97	54	0.000887061
BSI22	1.09	146	0.00088608	0.95	84	0.000932404
BSI23	0.95	94	0.000751817	0.86	78	0.00084022
BSI24	1.03	74	0.000859062	1.04	38	0.000875736
BSI25	0.74	10	0.000717143	0.74	26	0.000795316
BSI26	0.70	26	0.000781091	0.72	46	0.000811109
BSI27	1.01	126	0.000870248	0.86	38	0.000794136
BSI28	1.13	124	0.000868669	1.16	164	0.000890837
BSI29	1.13	74	0.000768403	1.03	78	0.00082768
BSI30	0.82	36	0.000759891	0.80	26	0.000786165
BSI31	1.04	132	0.000869701	1.01	80	0.00086793
BSI32	0.73	16	0.000739046	0.83	42	0.000870049
BSI33	0.77	0	0.000722175	0.77	8	0.000801584
BSI34	1.00	76	0.000821225	1.01	138	0.000984599
BSI35	1.01	94	0.00085206	1.08	114	0.000918678
BSI36	1.28	218	0.00086834	1.24	164	0.000934739
BSI37	1.03	132	0.000786993	1.16	174	0.000890033
BSI38	0.97	188	0.000915708	0.97	78	0.00090155
BSI39	0.99	64	0.000804906	0.89	30	0.000812122
BSI40	1.11	18	0.000776966	1.14	114	0.000913345
BSI41	1.10	76	0.000791038	1.18	94	0.000906503
BSI42	1.03	168	0.000902369	1.09	50	0.000908545
BSI43	0.97	120	0.000864009	1.08	112	0.000892166
BSI44	0.89	76	0.000897148	0.85	76	0.000859619
BSI45	1.20	118	0.000886731	1.25	110	0.000922779
BSI46	1.02	106	0.00086761	0.94	8	0.000840667
BSI47	0.83	160	0.000924894	0.88	106	0.000897319
BSI48	0.99	34	0.000809306	1.02	130	0.000941071
BSI49	1.04	102	0.000878344	0.96	120	0.000884639
BSI50	1.33	206	0.000892316	1.25	234	0.001021555
BSI51	1.13	72	0.000805934	1.00	40	0.000881475
BSI52	0.85	32	0.000794445	1.03	104	0.000976822
BSI53	0.85	4	0.000754921	0.96	136	0.000966611

Impact

Um den Einfluss einzelner Nodes auf die globale Konnektivität der Netzwerke zu untersuchen, wurden die Impact-Werte mittels Bootstrap-Verfahren berechnet. Aufgrund des hohen Rechenaufwandes für diesen Schritt wurden $m = 250$ Ziehungen per Bootstrap-Verfahren durchgeführt. Die Impact-Werte zu t_1 sowie die jeweils zugehörigen Konfidenzintervalle finden sich in der Tabelle 7, eine grafische Aufbereitung in Abbildung 8. Weiterhin wurde überprüft, inwiefern die errechneten Impact-Werte in einem Zusammenhang mit den Varianzen der jeweiligen Items stehen. Auch hierfür wurde zunächst die Annahme der Normalverteilung der Impact-Werte mittels Shapiro-Wilk-Test überprüft. Diese ergab, dass die Verteilung der Impact-Werte nicht statistisch signifikant von einer Normalverteilung abwich ($W = .98, p = .636$). Die Spearman-Rangkorrelation der gebootstrapteten Impact-Werte mit der Varianz der jeweiligen Items war zu t_1 statistisch nicht signifikant ($r_s(51) = -.09, p = .510$).

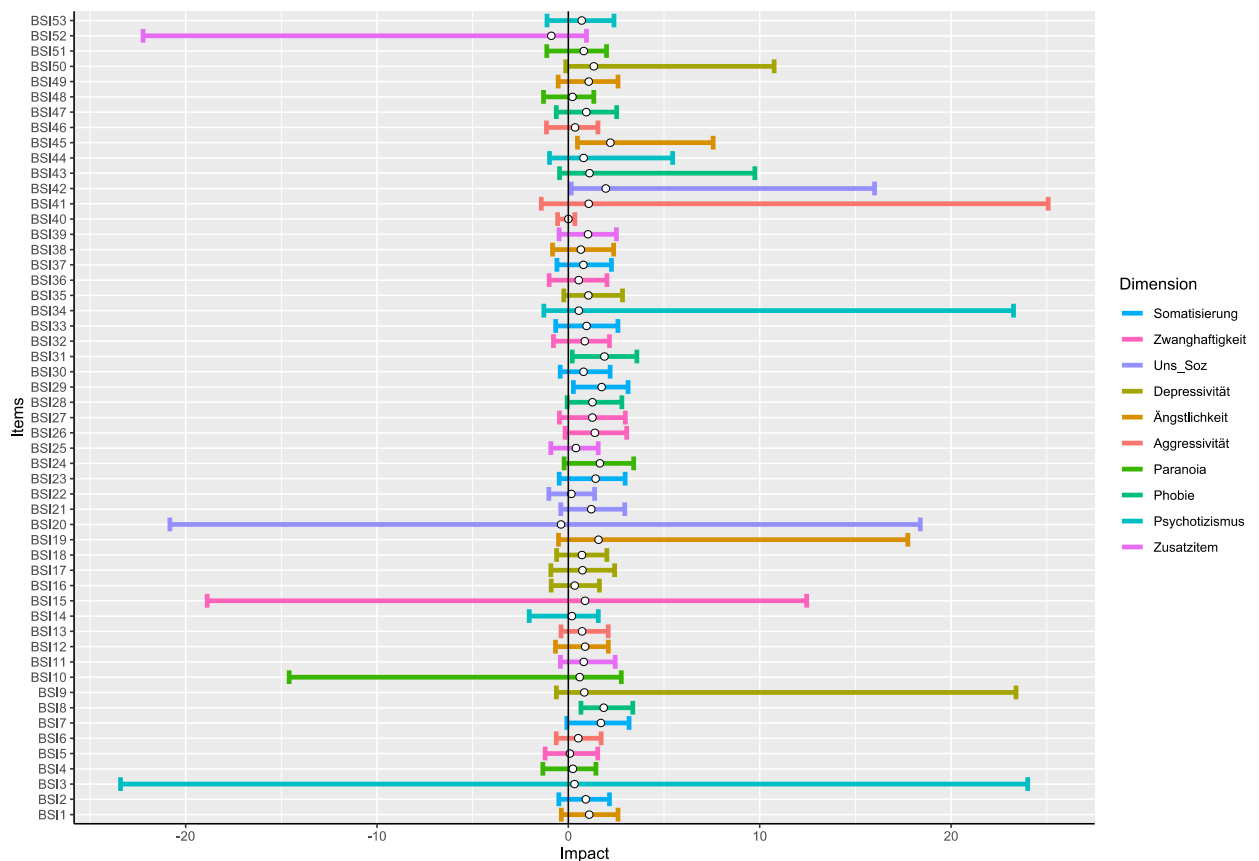


Abbildung 8. Grafische Darstellung der Impact-Werte zu t_1 mit den dazugehörigen Konfidenzintervallen der Items des BSI-53.

Tabelle 7

Impact-Werte des BSI-53 zu t₁

Itemnummer	Impact	95% KI
BSI1	1.09	[-0.37, 2.60]
BSI2	0.91	[-0.50, 2.15]
BSI3	0.32	[-23.40, 24.00]
BSI4	0.24	[-1.33, 1.44]
BSI5	0.08	[-1.22, 1.53]
BSI6	0.52	[-0.63, 1.71]
BSI7	1.70	[-0.08, 3.17]
BSI8	1.85	[0.66, 3.37]
BSI9	0.84	[-0.63, 23.40]
BSI10	0.59	[-14.59, 2.78]
BSI11	0.80	[-0.42, 2.45]
BSI12	0.87	[-0.67, 2.09]
BSI13	0.72	[-0.39, 2.09]
BSI14	0.19	[-2.05, 1.57]
BSI15	0.87	[-18.87, 12.46]
BSI16	0.34	[-0.90, 1.62]
BSI17	0.74	[-0.91, 2.42]
BSI18	0.71	[-0.61, 2.01]
BSI19	1.58	[-0.52, 17.74]
BSI20	-0.39	[-20.83, 18.39]
BSI21	1.20	[-0.40, 2.95]
BSI22	0.16	[-1.02, 1.37]
BSI23	1.43	[-0.48, 1.97]
BSI24	1.65	[-0.22, 3.41]
BSI25	0.40	[-0.92, 1.56]
BSI26	1.38	[-0.17, 3.05]
BSI27	1.26	[-0.47, 2.80]
BSI28	1.26	[-0.06, 2.80]
BSI29	1.74	[0.36, 3.12]
BSI30	0.79	[-0.43, 2.18]
BSI31	1.89	[0.21, 3.58]
BSI32	0.86	[-0.78, 2.15]
BSI33	0.95	[-0.66, 2.59]
BSI34	0.54	[-1.28, 23.27]
BSI35	1.05	[-0.24, 2.83]
BSI36	0.54	[-1.01, 2.02]
BSI37	0.79	[-0.60, 2.25]
BSI38	0.65	[-0.83, 2.37]
BSI39	1.03	[-0.49, 2.52]
BSI40	0.00	[-0.56, 0.34]
BSI41	1.06	[-1.42, 25.08]
BSI42	1.96	[0.15, 16.00]
BSI43	1.11	[-0.46, 9.75]
BSI44	0.80	[-0.99, 5.45]
BSI45	2.20	[0.48, 7.57]
BSI46	0.35	[-1.15, 1.55]
BSI47	0.94	[-0.63, 2.52]
BSI48	0.21	[-1.30, 1.33]
BSI49	1.07	[-0.53, 2.60]
BSI50	1.33	[-0.14, 10.75]
BSI51	0.80	[-1.13, 1.99]
BSI52	-0.89	[-22.22, 0.95]
BSI53	0.70	[-1.12, 2.38]

Ergebnisse der Testung der ersten Hypothese

Der Vergleich der Konnektivitäten der BSI-53-Netzwerke zu t_1 und t_2 erfolgte über den Network-Comparison-Test. Dieser ergab für das Netzwerk zu t_1 eine globale Konnektivität von 25.13 und von 25.47 für das Netzwerk zu t_2 . Der Unterschied zwischen den beiden Netzwerken hinsichtlich ihrer Konnektivität war mit $p = .290$ statistisch nicht signifikant. Der Vergleich der GSI-Scores zwischen t_1 und t_2 wurde mittels eines Wilcoxon-Vorzeichen-Rangtests durchgeführt und fiel signifikant aus ($z = -16.74, p < .001$).

Symptomnetzwerke der Responder und Non-Responder

Weiterhin wurde im Rahmen dieser Arbeit eine Subgruppenanalyse durchgeführt. Hierfür wurden zwei Gruppen gebildet, die Responder und die Non-Responder, und jeweils ein Symptomnetzwerk geschätzt. Durch den Vergleich dieser beiden Netzwerke wurden sich Aufschlüsse über den Zusammenhang zwischen strukturellen Aspekten und der Netzwerkdynamik erhofft.

Auch im Zuge dieser Untersuchungen wurden Stabilitätsanalysen durchgeführt. Die Berechnungen der CS-Koeffizienten erfolgten auch hier auf der Basis von Bootstrap-Berechnungen mit je $m = 1000$ Ziehungen. Die Werte dieser Koeffizienten legen nahe, dass bei der Berechnung des Netzwerks der Responder sowohl die Edges mit $CS(cor = .70) = .05$ als auch die Zentralitätsmaße mit $CS(cor = .70) = .05$ hinsichtlich ihrer Stabilität gemäß gängiger Richtlinien in einem nicht mehr interpretierbaren Bereich liegen. Auch bei der Berechnung des Netzwerks der Non-Responder liegen die Stabilitäten der Schätzungen der Edges mit $CS(cor = .70) = .05$ und der Zentralitätsmaße mit $CS(cor = .70) = .05$ in einem Bereich, in dem eine Interpretation nicht substantiiert wäre.

Weiterhin wurden auch für diese beiden Netzwerke mittels Bootstrap-Verfahren mit je $m = 1000$ Ziehungen Konfidenzintervalle für die Edges berechnet. Diese finden sich wieder anbei der jeweiligen Darstellung der Edges.

Die geschätzten Netzwerke der Responder und der Non-Responder finden sich in den Abbildungen 9 und 10.

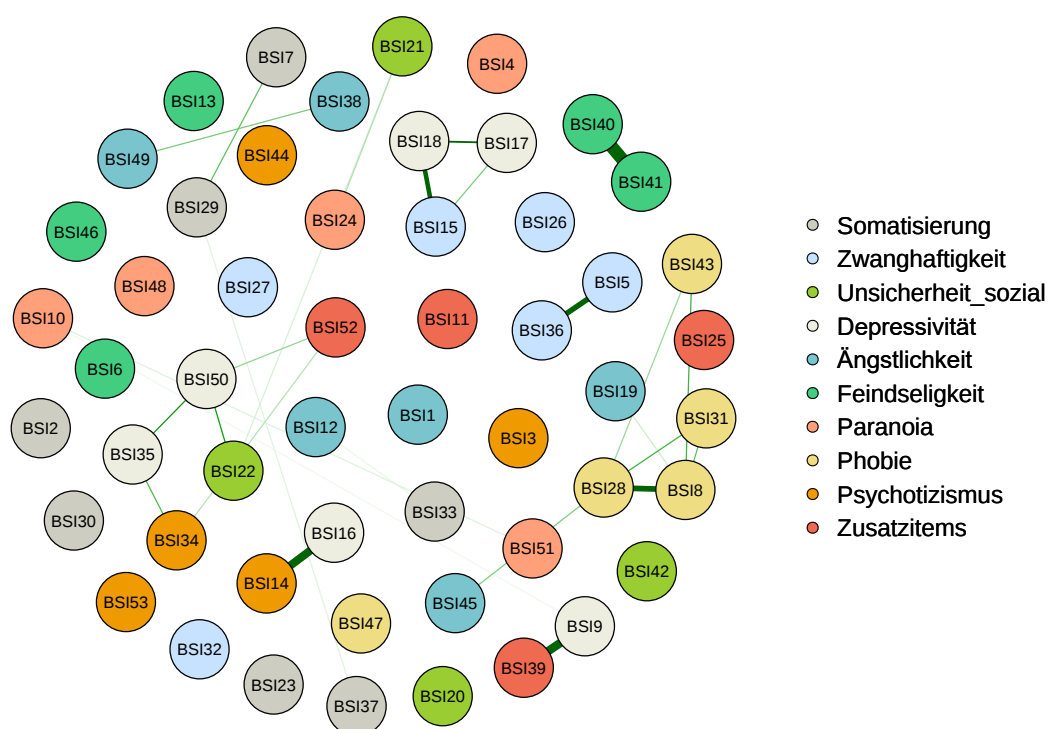


Abbildung 9. Symptomnetzwerk des BSI-53 der Responder zu t_1 .

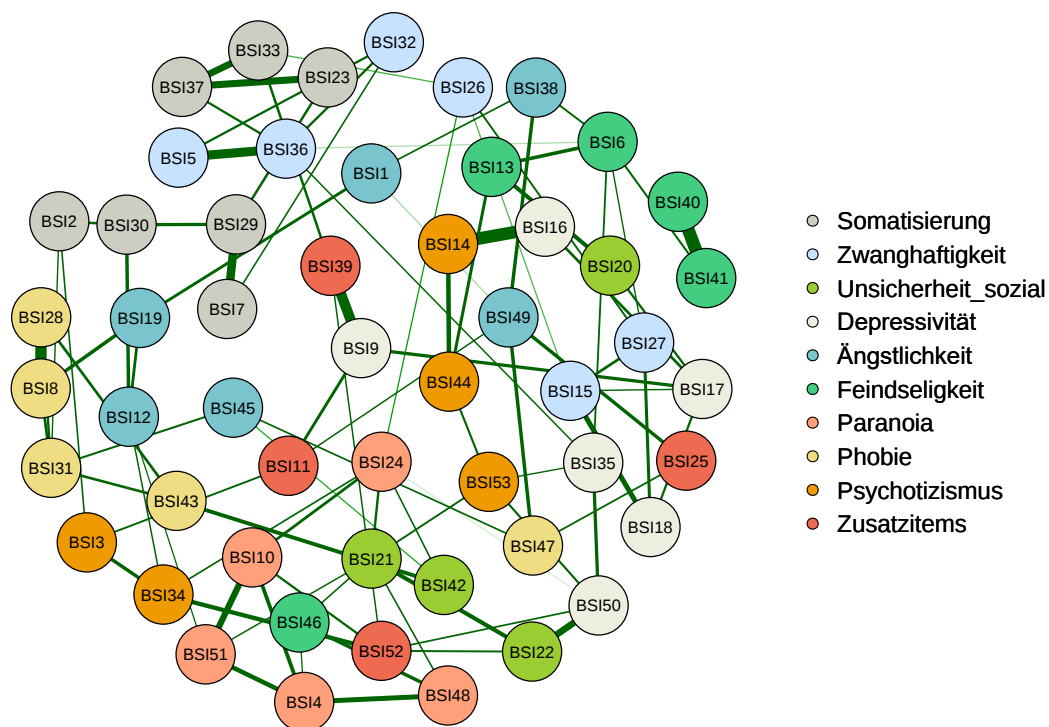


Abbildung 10. Symptomnetzwerk des BSI-53 der Non-Responder zu t_1 .

Auch hier sind, aus ähnlichen Überlegungen wie bei den Abbildungen der Netzwerke zu t_1 und t_2 , lediglich die Edges dargestellt, die stärker als $r = .10$ sind. Abbildungen der Netzwerke mit allen Edges finden sich im Anhang (Abb. 11 und Abb. 12). Weiterhin finden sich in Tabelle 8 und Tabelle 9 Aufstellungen der einzelnen Edges. Da im geschätzten Netzwerk der Responder lediglich neun Edges größer oder gleich $r \geq .10$ waren und keine Edge größer gleich $r \geq .30$ war, werden hier alle Edges mit $r \geq .10$ tabellarisch dargestellt. Aus dem Netzwerk der Non-Responder werden hier alle Edges mit $r \geq .30$ dargestellt, eine Tabelle mit allen Edges mit $r \geq .10$ findet sich unter Tabelle 10 im Anhang.

Ergebnisse der Testung der zweiten Hypothese

Anhand der Subgruppen der Responder und der Non-Responder wurde untersucht, inwiefern die Netzwerkkonnektivität einen Einfluss auf den weiteren Verlauf der Behandlung hat. Hierzu wurden die beiden geschätzten Netzwerke zu t_1 mittels eines Network-Comparison-Tests hinsichtlich ihrer Konnektivität miteinander verglichen. Dieser ergab für das Netzwerk der Responder eine globale Konnektivität von .77, für das Netzwerk der Non-Responder von .62. Die beiden Netzwerke unterschieden sich hinsichtlich ihrer Konnektivität mit $p = .360$ nicht signifikant voneinander.

Da diese Ergebnisse vor dem Hintergrund der gravierenden visuellen Unterschiede zwischen den beiden Netzwerkstrukturen unplausibel erschienen, wurden die globalen Konnektivitäten noch einmal händisch berechnet. Hierbei ergaben sich für das Netzwerk der Responder eine globale Konnektivität von 2.63, für das Netzwerk der Non-Responder von 25.24.

Tabelle 8

Darstellung der Edges der Responder zu t_1 mit $r \geq .10$

Node 1	Node 2	r	SD	95% KI
BSI40	BSI41	.30	.10	[.09, .50]
BSI9	BSI39	.24	.09	[.05, .43]
BSI14	BSI16	.23	.11	[.02, .44]
BSI8	BSI28	.20	.09	[.03, .37]
BSI5	BSI36	.19	.09	[.01, .38]
BSI15	BSI18	.18	.08	[.00, .35]
BSI17	BSI18	.14	.07	[.00, .27]
BSI35	BSI50	.13	.07	[-.01, .27]
BSI22	BSI50	.12	.07	[-.01, .26]

Anmerkung. Die Items wurden dem BSI-53 entnommen; r = partielle polychorische Korrelation; SD = Standardabweichung; KI = Konfidenzintervall.

Tabelle 9

Darstellung der Edges der Non-Responder zu t_1 mit $r \geq .30$

Node 1	Node 2	r	SD	95% KI
BSI40	BSI41	.52	.10	[.31, .72]
BSI14	BSI16	.51	.06	[.29, .64]
BSI8	BSI28	.47	.06	[.33, .60]
BSI9	BSI39	.43	.07	[.28, .59]
BSI5	BSI36	.39	.09	[.22, .56]
BSI7	BSI29	.37	.07	[.22, .51]
BSI23	BSI37	.31	.08	[.15, .48]
BSI33	BSI37	.30	.08	[.13, .47]
BSI22	BSI50	.30	.07	[.17, .43]

Anmerkung. Die Items wurden dem BSI-53 entnommen; r = partielle polychorische Korrelation; SD = Standardabweichung; KI = Konfidenzintervall.

Diskussion

Im Rahmen dieser Arbeit wurde eine netzwerkanalytische Untersuchung der Symptomnetzwerke einer Stichprobe der TeilnehmerInnen des Projektes „fit2work“ durchgeführt. Diese hatte in erster Linie explorativen Charakter, allerdings wurden auch einige spezifische Phänomene konfirmatorisch geprüft. Hierbei handelt es sich um eine Untersuchung der Veränderungen einer globalen Eigenschaft der Netzwerkmodelle, der Konnektivität, im Zeitraum einer Interventionsmaßnahme. Auch wurden Zusammenhänge zwischen der Konnektivität und dem weiteren Symptomverlauf untersucht, um Einblicke in den Zusammenhang zwischen dem mathematischen Konzept der Konnektivität und dem inhaltlichen Konzept der Resilienz der Symptomnetzwerke zu erhalten.

Im Zuge der Analysen wurden auch einige Untersuchungen hinsichtlich der verwendeten Methodik durchgeführt. Hierbei handelt es sich zunächst um die Stabilitätsanalysen, die mittels Bootstrap-Verfahren eine Einschätzung der Zuverlässigkeit und der Präzision der Ergebnisse ermöglichen. Hier ist festzuhalten, dass die Schätzungen der Netzwerke zu t_1 und t_2 gemäß der gängigen Konventionen (Epskamp et al., 2018) insgesamt als stabil angesehen werden können. Insbesondere die Schätzungen der Edges zeigten sich sowohl zu t_1 als auch zu t_2 mit Correlation-Stability-Koeffizienten von $CS(cor = .70) = .75$ im sehr guten Bereich, was für eine hohe Stabilität der Schätzung der Netzwerkstruktur spricht. Auch die Schätzungen der Node-Strength liegen gemäß gängiger Konventionen im guten Bereich. Dieser Befund spiegelt sich auch in der Abwesenheit einer signifikanten Korrelation

zwischen den verschiedenen Zentralitätsmaßen und den Varianzen der jeweiligen Items wider. Dieser Untersuchungsschritt wurde vorgenommen, um zu überprüfen, ob es Hinweise auf einen Einfluss der differentiellen Variabilität auf die Schätzungen der Netzwerkstruktur gibt. Dem lag der Gedanke zugrunde, dass die untersuchten Items teilweise unterschiedliche Varianzen aufwiesen, was sich wiederum in ihren Kovariationen und folglich Korrelationen niederschlagen könnte. Da diese im Wesentlichen die Grundlage für die Schätzung der Netzwerkstruktur bilden, könnte ein solches methodisches Artefakt zu einer starken Verzerrung sämtlicher weiterhin durchgeführter Analyseschritte führen. So ist es beispielsweise vorstellbar, dass ein Item, das mit einem ausgeprägten Bodeneffekt gemessen wird, wenig Varianz aufweist und folglich weniger mit anderen Items kovariieren kann. Hieraus würden sich im Rahmen der Netzwerkmodellierung nur wenige Edges mit diesem Item ergeben, folglich hätte es eine geringe Node-Strength. Ein Einfluss dieses Phänomens würde die Interpretierbarkeit der Zentralitätsmaße und darüber der Netzwerkstruktur erheblich negativ beeinträchtigen (Terluin, de Boer, & de Vet, 2016). Die Tatsache, dass die Korrelationen zwischen sämtlichen Zentralitätsmaßen, sowohl zu t_1 als auch zu t_2 , mit den jeweiligen Itemvarianzen gemäß Cohen (1988) unterhalb der Grenze eines kleinen Effekts lagen (alle $r < .10$) und vom gängigen Signifikanzniveau von $p \leq .05$ weit entfernt waren (alle $p \geq .55$), deutet darauf hin, dass keine linearen Zusammenhänge zwischen den Itemvarianzen und den Zentralitätsmaßen vorlagen. Dies lässt für sich allein genommen zwar nicht automatisch den Schluss zu, dass die abgebildeten Korrelationen inhaltlich substantiell sind und die Netzwerkstruktur ausschließlich tatsächliche Wirkmechanismen querschnittlich widerspiegelt. Allerdings kann anhand der nicht signifikanten Korrelationen zumindest der Einfluss eines wesentlichen methodischen Artefakts ausgeschlossen werden.

Korrelative Zusammenhänge – Abbildungen von Kausalstrukturen?

Aus der Abbildung 3 gehen die korrelativen Zusammenhänge zwischen den Items des BSI-53 hervor. Hieraus werden mehrere Punkte inhaltlicher und epistemologischer Natur ersichtlich. Zunächst lässt sich beschreiben, dass die einzelnen Items auch ohne eine Modellierung ihrer zugrundeliegenden Dimensionen grob entsprechend der postulierten latenten Variablen, sprich Skalen des BSI-53, angeordnet sind. Dies kann vor dem Hintergrund der Tatsache, dass für die graphische Darstellung des Netzwerks der Fruchterman-Reingold-Algorithmus verwendet wurde, als interessanter Befund angesehen werden. Der Fruchterman-Reingold-Algorithmus platziert die einzelnen Nodes entsprechend der Stärke ihrer Zusammenhänge näher oder weiter voneinander entfernt. Damit entsteht zumindest theoretisch eine direkte Äquivalenz zwischen der mathematischen Distanz zu der

metrischen Distanz zweier Nodes. Vor diesem Hintergrund scheint die Tatsache, dass die Items mehrerer Skalen nahe beieinander liegen, als Hinweis auf ein gewisses Clustering dieser Items interpretierbar zu sein. Diese Deutung der metrischen Distanzen ist jedoch problematisch, da es sich bei der grafischen Abbildung des Netzwerks um eine auf zwei Dimensionen reduzierte Fassung eines hochdimensionalen Raumes handelt (Epskamp, 2017). Hierbei kommt es bereits bei Netzwerken mit wenigen Nodes fast zwangsläufig zu Verzerrungen, im hier behandelten Netzwerk mit 53 Nodes sollten also Interpretationen, die sich auf rein visuelle Faktoren stützen, nur mit höchster Vorsicht vorgenommen werden. Weiterhin wäre lediglich zu mutmaßen, wodurch diese Struktur bedingt wird. So ist es zwar vorstellbar, dass die Items einzelner Skalen gemeinsame Bezugspunkte haben, durch die sich Muster in den Kovarianzstrukturen ergeben. Es bleibt jedoch fraglich, ob diese tatsächlich latente Variablen im Sinne eines übergeordneten Konstruktes, das die gemessenen Variablen kausal bedingt, darstellen. So wäre es möglich, dass die Items in einigen Fällen höhere Kovarianzen aufweisen, weil sie sehr ähnliche Inhalte erfassen. Als beispielhaft hierfür kann die Korrelation zwischen den Items „Einsamkeitsgefühle, selbst wenn Sie in Gesellschaft sind“ und „Einsamkeitsgefühle“ gesehen werden. Ein Versuch, diese Verknüpfung auf einen Zusammenhang zwischen den beiden postulierten latenten Konstrukten „Depressivität“ und „Psychotizismus“ zurückzuführen, würde die fast gleiche wörtliche Formulierung ignorieren und damit implizieren, dass es sich hierbei um klar voneinander abgrenzbare Konstrukte handelt und das eine Item nicht in einem von diesem Modell unabhängigen Zusammenhang mit dem latenten Konstrukt des jeweils anderen Modells steht. Die Annahme, dass „Einsamkeitsgefühle“ Depressivität reflektieren, „Einsamkeitsgefühle, selbst wenn Sie in Gesellschaft sind“ jedoch nicht, erscheint wenig intuitiv.

Dieser Grundgedanke ergibt sich auch bei Betrachtung der Tabelle 2, in der die Zusammenhänge des Symptomnetzwerks zu t_1 , die größer oder gleich $r \geq .30$ sind, abgebildet sind. Hier wird aus den Itemformulierungen ersichtlich, dass es unterschiedliche Erklärungsansätze für das Auffinden korrelativer Zusammenhänge gibt. So sind beispielsweise im Fall des Zusammenhangs zwischen „Gedächtnisschwierigkeiten“ und „Konzentrationsproblemen“ mehrere Erklärungsansätze vorstellbar, von einem kausalen Wirkmechanismus zwischen diesen beiden Variablen bis hin zur Annahme einer latenten Variable, die sich in diesen beiden Variablen manifestiert. Auch ist es vorstellbar, dass die Items kovariieren, weil sie mit ähnlichen Formulierungen die gleichen Variablen abfragen (wie beispielsweise die bereits angeführten „Einsamkeitsgefühle“ und „Einsamkeitsgefühle, selbst wenn Sie in Gesellschaft sind“). Hieran offenbart sich ein epistemologisches Problem

nicht nur der Netzwerkanalysen, sondern der mathematischen Modellierung und Inferenzstatistik im Allgemeinen, das im Folgenden noch besprochen werden soll.

Zentralitätswerte – Indikatoren für Ansatzpunkte von Interventionen?

Im Zuge dieser Untersuchung wurde weiterhin analysiert, welche Nodes, und damit welche Aspekte der psychischen Belastung, einen hohen Stellenwert in dem Belastungsnetzwerk einnehmen. Hierfür wurden die Zentralitätsmaße der Node-Strength, der Betweenness sowie der Closeness berichtet, die in Tabelle 6 abgebildet sind. Aus den Ergebnissen dieser Analysen gehen mehrere Dinge hervor. Zunächst ist auffallend, dass einige Items sowohl hinsichtlich ihrer Node-Strength als auch hinsichtlich ihrer Betweenness die zentralsten Items des Netzwerks zu t_1 bilden. So reichen insgesamt sieben Items aus, um auf beiden Metriken die jeweils stärksten fünf abzubilden. Zusammenhänge zwischen den Zentralitätsmaßen sind nicht unüblich in der netzwerkanalytischen Literatur (z.B. Fried et al., 2016) und können für sich genommen bereits Implikationen mit sich bringen. Diese ergeben sich vor allem vor dem Hintergrund der unterschiedlichen mathematischen Konzeptionen der Node-Strength und der Betweenness. So spiegelt die Node-Strength die Stärke der Korrelationen einer Node wider, die Betweenness quantifiziert hingegen, wie oft die Node auf dem kürzesten direkten Pfad zwischen zwei anderen Nodes liegt. Hiermit scheint der Zusammenhang zwischen diesen beiden Zentralitätsmaßen einen Einblick in den Zusammenhang verschiedener struktureller Aspekte des Netzwerks zu geben, was bei der Generierung von Hypothesen sinnvollerweise genutzt werden könnte. So wäre es beispielsweise vorstellbar, dass gezielte Interventionen an Nodes mit einer hohen Node-Strength und einer hohen Betweenness eine größere Wirkung auf das gesamte Krankheitssystem haben als Interventionen an Nodes mit nur einer hohen Node-Strength. Es wäre zu erwarten, dass deren Effekt eher lokal wirksam wäre, während eine Intervention auf Basis beider Maße auch einen hemmenden, also hier positiven, Effekt auf die Vermittlung von Aktivität im Netzwerk hätte.

Dieser Grundgedanke basiert jedoch auf der wesentlichen Annahme, dass sich aus querschnittlichen Untersuchungsdesigns längsschnittliche Implikationen ableiten lassen können. Diese Annahme lässt sich, wenn unhinterfragt, sowohl im generellen als auch in Bezug auf die Netzwerkanalyse als problematisch beschreiben. So ist es beispielsweise schwierig zu versuchen, einzig aus der Node-Strength Ziele für Interventionen abzuleiten. Dies liegt darin begründet, dass in querschnittlichen Schätzungen von Netzwerken lediglich ungerichtete Zusammenhänge dargestellt werden können. Der Schluss, dass eine Intervention auf einer Node mit hoher Node-Strength viele andere Nodes beeinflussen würde, muss also

nicht zwingend korrekt sein, alternativ ist es auch vorstellbar, dass die Kausalrichtung andersherum ist, die betroffene Node also von den umliegenden aufrechterhalten wird und dementsprechend besonders resistent ist, oder aber Wechselwirkungen zwischen den einzelnen Nodes vorliegen (Rodebaugh et al., 2018). Dementsprechend ist eine längsschnittliche Untersuchung der Zusammenhänge zwischen der Node-Strength und verschiedenen Interventionszielen (wie beispielsweise dem Rückgang einzelner Symptome) unabdingbar, um zu substantiierten und präzisen Hypothesen und entsprechenden Befunden bezüglich der Zusammenhänge zwischen Netzwerkstrukturen und Netzwerkdynamiken kommen zu können.

Eine empirische Untersuchung der Frage, ob die querschnittlich geschätzte Zentralität einer Variable den längsschnittlichen Zusammenhang zwischen einer Veränderung in dieser Variable und Veränderungen im ganzen Netzwerk vorhersagt, führten Rodebaugh et al. (2018) durch. Hierfür schätzten sie zunächst ein querschnittliches Netzwerkmodell der Items der *Liebowitz Social Anxiety Scale* (Liebowitz, 1987) und untersuchten die Items hinsichtlich ihrer Zentralität. Im Folgenden untersuchten sie anhand einer zweiten Stichprobe, inwiefern Veränderungen einzelner Items mit Veränderungen aller anderen Items korrelierten. Bei der anschließende Untersuchung des Zusammenhangs zwischen der querschnittlichen Zentralität eines Items und der längsschnittlichen „Wichtigkeit“ eines Items für die globale Symptomverbesserung konnten sie erste Hinweise darauf finden, dass ein gewisser Zusammenhang zwischen der längsschnittlichen Dynamik des Netzwerks einerseits und der querschnittlichen Struktur des Netzwerks bestehen könnte (Rodebaugh et al., 2018). Allerdings gingen mit diesem Zusammenhang mehrere starke methodische Limitationen einher, sodass die Autoren zu der abschließenden Bewertung kamen, dass die Verwendung von querschnittlich geschätzten Maßen für längsschnittliche Designs, wie etwa eine Intervention, auch unter sehr guten Bedingungen (z.B. stabile Schätzungen, große Stichproben) nicht immer zielführend sei und eingehender untersucht werden sollte.

Hieran wird ersichtlich, dass es anscheinend lediglich vereinzelt und mit stark eingeschränkten Geltungsbereichen möglich ist, aus querschnittlichen Befunden Implikationen für Maßnahmen abzuleiten. Allerdings wurden diesbezüglich in der bisherigen Netzwerkliteratur lediglich klinisch-psychologische Interventionen in Betracht gezogen, deren Ableitung jedoch aus den oben beschriebenen Gründen problematisch war. Insbesondere die Tatsache, dass anhand querschnittlicher Untersuchungen keinerlei Aussagen über Wirkungsrichtungen getroffen werden können, stand hier einer Nutzung dieser für die Anwendung im Wege, da unklar ist, welche inhaltlichen Gründe für das Zustandekommen der

hohen Zentralität verantwortlich sind. Zwei Möglichkeiten zum Umgang mit diesem Problem sollen an dieser Stelle kurz perspektivisch skizziert werden.

Zunächst wäre es naheliegend, die Ergebnisse von querschnittlichen Untersuchungen zu verwenden, um eine Vorauswahl an Konstrukten zu treffen, die aufgrund verschiedener Charakteristika eine wesentliche Rolle im Netzwerk spielen. Im Folgenden könnten kurze längsschnittliche Erhebungen durchgeführt werden, um die Interaktionsmechanismen der betreffenden Nodes sowie der jeweils umliegenden Nodes näher zu beleuchten. Hierfür würde sich die Verwendung von Experience-Sampling-Methoden anbieten, da sich hierdurch ein differenziertes Bild der Wirkmechanismen ergeben könnte. Dies liegt darin begründet, dass es mittels ESM möglich ist, Fragebögen in sehr kurzen Zeitabständen vorzugeben, wodurch die Erhebung (und damit auch die statistische Modellierung) näher an die tatsächlich ablaufenden Prozesse kommen würde. Eine Analogie könnte hier in dem Versuch gesehen werden, spezifische therapeutische Prozesse eines klinisch-psychologischen Maßnahmenpakets zu untersuchen. Durch die Erhebung und folglich Modellierung zu lediglich zwei Zeitpunkten, vor und nach den Maßnahmen, ließen sich weit weniger differenzierte Aussagen treffen, als bei mehreren Erhebungen (beispielsweise nach jeder spezifischen Interventionsmaßnahme). Ähnlich hierzu könnte die Verwendung von Experience-Sampling-Methoden basierten Daten zu einem differenzierteren Bild der Wirkmechanismen der einzelnen Nodes führen. Die Ergebnisse einer vorherigen querschnittlichen Untersuchung könnten hier nutzbringend sein, da durch die Vorauswahl einzelner Nodes einerseits die Belastung für die ProbandInnen durch die Befragung geringer gehalten würde sowie andererseits die Modellierung der zeitlichen Abläufe spezifischer und damit mit einer höheren Testmacht durchgeführt werden könnte. Dies liegt darin begründet, dass es im Bereich der Netzwerkschätzungen einen positiven Zusammenhang zwischen der Anzahl der modellierten Nodes und der mindestens notwendigen Stichprobengröße gibt (Epskamp, 2017). Somit wäre es beispielsweise erheblich ökonomischer, ein längsschnittliches Netzwerkmodell aus sechs Variablen zu schätzen als aus 20 Variablen. Dementsprechend könnten hierdurch die Stärken der quer- und längsschnittlichen Netzwerkanalysen zur Geltung gebracht werden und eine wesentliche Schwäche, der Bedarf nach einer großen Stichprobe, teilweise ausgeglichen werden.

Eine weitere Möglichkeit zur Verwendung querschnittlicher Befunde bezüglich der Zentralitäten einzelner Variablen könnte in der Gestaltung präventiver Maßnahmen liegen. Dies wäre jedoch in den Augen des Autors dieser Arbeit nur sinnvoll möglich, wenn hierbei sowohl die Node-Strength als auch die Betweenness berücksichtigt würden. So ist es vorstellbar, dass das zugrundeliegende Problem der Unbekanntheit der Wirkmechanismen, die

zu einer hohen Node-Strength führen, nur dann wirklich problematisch ist, wenn man von einem aktiven Netzwerk ausgeht, was konkret bedeuten würde, dass die abgebildeten Symptome ausgeprägt vorliegen. Dies würde ein Ansetzen an einer einzelnen zentralen Node ohne Wissen um die Wirkmechanismen hinter dieser Zentralität zu einem wenig ethischen Unterfangen machen, da die hierfür verwendeten Ressourcen vielleicht besser in einer Intervention mit gesicherten Wirkmechanismen und entsprechend höherer Wahrscheinlichkeit eines positiven Outcomes investiert wären. Im Gegensatz hierzu könnte ein schwach ausgeprägtes Symptom mit einer hohen Node-Strength und hohen Betweenness ein optimaler Kandidat für präventive Maßnahmen sein. Dies wird ersichtlich, wenn man erneut die drei möglichen Szenarien bezüglich der Wirkmechanismen hinter der Node-Strength betrachtet. Szenario a.) Die hohe Node-Strength des Symptoms kommt dadurch zustande, dass es die umliegenden Symptome beeinflusst. Hier ergibt es sich fast von selbst, dass diese Node ein lohnendes Ziel für präventive Maßnahmen wäre. Es könnten einerseits, aufgrund der hohen Node-Strength, lokale Effekte im Symptomnetzwerk erzielt werden, andererseits, aufgrund der hohen Betweenness, auch die Ausbreitung von Aktivität vermindert werden. Szenario b.) Die hohe Node-Strength des Symptoms kommt dadurch zustande, dass es durch die umliegenden Symptome beeinflusst wird. Auch hier wäre eine Intervention vor der Manifestation der Symptome, also eine primäre Prävention, sinnvoll. Dies ist dadurch bedingt, dass das betrachtete Symptom durch andere Symptome vergleichsweise einfach aktiviert werden könnte (hohe Node-Strength) und weiterhin die Weiterleitung von Aktivität ermöglichen würde (hohe Betweenness). So wäre es beispielsweise vorstellbar, dass präventive Maßnahmen, die an Brückensymptomen, also Symptomen, die zwischen zwei Symptomclustern stehen und diese verbinden, der Manifestation einer komorbiden Störung entgegenwirken könnten. Szenario c.) Es gibt wechselseitige Beeinflussungen zwischen dem fokalen Symptom und den umliegenden Symptomen. In diesem Fall würden beide Gründe für präventive Maßnahmen, die an diesem Symptom ansetzen, sprechen: sowohl die Tatsache, dass hierdurch ein lokaler Effekt erzielt werden könnte, als auch die Tatsache, dass dieses Symptom ein hohes Risiko der Aktivierung aufweisen könnte (und sich die Aktivität folglich „ungestörter“ im Symptomnetzwerk ausbreiten könnte), sprechen hierfür.

Es ist jedoch klar hervorzuheben, dass es sich bei diesen entworfenen Szenarien um Idealtypen handelt, und es nicht unwahrscheinlich scheint, dass in der Realität alle drei möglichen Wirkmechanismen zugleich eine Rolle spielen. Dementsprechend wäre es für die Verwendung der Zentralitätsmaße in Form präventiver Maßnahmen essenziell, dass diese auf Nodes abzielen, die zum Zeitpunkt der Maßnahme nicht aktiv sind. Der Grundgedanke

hierbei bestünde kurz gesagt in der Identifizierung von Nodes, die nicht aktiv werden sollten, wobei sowohl die Information der Node-Strength als auch die der Betweenness eine eigenständige und wesentliche Rolle spielen.

Zusammenhänge zwischen der Struktur und der Dynamik von Netzwerken

Weiterhin wurden im Rahmen der vorliegenden Arbeit zwei Hypothesen untersucht, die sich beide auf den Zusammenhang zwischen der globalen Struktur und der Dynamik der Netzwerke bezogen. Die erste Hypothese diente hierbei der Untersuchung, ob sich die Netzwerkstrukturen der ProbandInnen der Stichprobe während der Beratungen im Rahmen von „fit2work“ signifikant veränderten. Hierfür wurden das Symptomenetzwerk zu t_1 und das Symptomenetzwerk zu t_2 mittels Network-Comparison-Test bezüglich ihrer Konnektivität verglichen. Dieser ergab, dass sich die beiden Netzwerke nicht signifikant hinsichtlich ihrer Konnektivität voneinander unterschieden. Hierbei ist anzumerken, dass bereits in bisherigen Evaluationen von „fit2work“ signifikante Unterschiede zwischen t_1 und t_2 auf allen Subskalen des BSI-53 sowie des GSI-Scores gefunden wurden (Jagsch, 2015). Auch in dieser Stichprobe wurde, hier mittels eines Wilcoxon-Vorzeichen-Rangtests, eine signifikante Verbesserung der GSI-Werte von t_1 zu t_2 gefunden. Hieran wird ersichtlich, dass anhand der Konnektivität keine Aussagen über die Manifestation von Symptomen möglich sind, sondern dass sie lediglich Aufschlüsse über die Zusammenhänge zwischen ihnen zulässt. Dies entspricht so jedoch auch klar dem Anspruch und der Konzeption der netzwerkanalytischen Methoden und sollte dementsprechend nicht als Schwäche gewertet werden. Eine Deutung dieser Befunde kann jedoch nicht zuletzt vor dem theoretischen Hintergrund der Netzwerktheorie im Bereich der Psychopathologie erfolgen. Gemäß dieser besteht die Verknüpfung der mathematisch-statistischen Kenngröße der Konnektivität mit einer inhaltlichen Theorie der Psychopathologie in den Konzepten der Vulnerabilität und der Resilienz. So verwendet Borsboom (2017) nicht nur die Begriffe Konnektivität und Vulnerabilität/Resilienz beinahe äquivalent, auch spielen sie, neben der Stärke der Symptomausprägung, eine zentrale Rolle in seiner Definition von psychischer Gesundheit und Krankheit. Im Rahmenmodell der Netzwerktheorie ließen sich die Ergebnisse also dahingehend interpretieren, dass zwischen den beiden Messzeitpunkten der Evaluation von „fit2work“ eine signifikante Verbesserung des Globalscores stattgefunden hat (und in detaillierteren Evaluationen des Projektes auch eine Verbesserung auf den einzelnen Skalen des BSI-53 aufgezeigt werden konnte), es aber keine Auswirkungen auf die generelle Vulnerabilität gegeben zu haben scheint.

Hieraus ergeben sich Implikationen für die Rückfallprophylaxe, die auf dem Grundgedanken basieren, dass sich die Aktivität in einem hoch konnektiven

Symptomnetzwerk schneller und einfacher ausbreitet als in einem niedrig konnektiven Netzwerk. Bei der Konzeption präventiver Ansätze und Strategien könnten die Befunde der in Tabelle 6 beschriebenen Zentralitätsanalyse und die Befunde der in Tabelle 7 beschriebenen Impactanalyse wertvolle Beiträge liefern, um die Konnektivität des Netzwerkes zu verringern.

Die zweite Hypothese diente der Untersuchung, ob es einen Zusammenhang zwischen der Konnektivität der Symptomnetzwerke und dem weiteren Verlauf der psychischen Belastung gab. Hierfür wurden anhand der Veränderungen im GSI-Wert zwei Substichproben gebildet. Die Gruppe der Non-Responder wurde aus allen Personen gebildet, die sich hinsichtlich ihres GSI-Wertes verschlechterten oder nicht veränderten. Die Gruppe der Responder wurde aus den 20 Prozent der Personen gebildet, die sich am stärksten verbesserten. Ein Network-Comparison-Test, der die beiden Symptomnetzwerke dieser beiden Gruppen miteinander verglich, kam zu dem Ergebnis, dass sie sich nicht hinsichtlich ihrer globalen Konnektivität unterschieden. Dieser Befund spricht an sich gegen eine Verknüpfung zwischen der Konnektivität und Konzepten wie der Selbst-Aufrechterhaltung von psychischen Krankheiten. Allerdings sollte auch dieser Befund nur vor dem Hintergrund der statistischen Methoden, durch die er zustande gekommen ist, interpretiert werden. So wird beispielsweise aus den Stabilitätsanalysen der Netzwerkschätzungen der Responder und der Non-Responder ersichtlich, dass die beiden Netzwerke und damit auch die aus ihnen gezogenen Inferenzen nicht wirklich sinnvoll interpretiert werden können. Weiterhin ist auffällig, dass es rein visuell gravierende Unterschiede zwischen den beiden Netzwerkstrukturen gibt. Auch lagen die vom Network-Comparison-Test mitberechneten Konnektivitäten der beiden Netzwerke sehr nahe beieinander, was mit den grafischen Abbildungen sowie den berechneten Edges nur schwierig zur vereinbaren scheint. Aus diesem Grund entschloss sich der Autor dieser Arbeit, die Konnektivitäten selbstständig zu berechnen, was zu stark abweichenden Werten führte. Da diese Werte eher mit den grafischen Abbildungen und den Werten der einzelnen Edges vereinbar scheinen und zudem entsprechend der Definition der Konnektivität gebildet wurden, wäre die Vermutung naheliegend, dass es sich hierbei um ein technisches Artefakt handelt, das auf der Instabilität der Schätzungen beruht. Dies macht eine Interpretation der Ergebnisse des Network-Comparison-Tests sowie der hier geschätzten Netzwerkstrukturen im Wesentlichen gegenstandslos.

Das Konzept des Impact

Weiterhin wurde das Konzept des Impacts an dem BSI-53-Symptomnetzwerk zu t_1 untersucht. Hierfür wurden die Impact-Werte zunächst auf reguläre Art geschätzt. Allerdings kam es bei der mehrmaligen Wiederholung dieser Berechnung zu auffälligen Unregelmäßigkeiten. Diese konnten darauf zurückgeführt werden, dass die Anwendung des Impacts auf ordinale oder metrische Daten auf einem Median-Split der Stichprobe auf der jeweils untersuchten Variable beruht. Hierdurch sollen die Schätzung zweier Netzwerke ermöglicht werden, eines mit hoher Ausprägung auf der betrachteten Node, eines mit niedriger Ausprägung. Hierbei werden jedoch die Daten der Personen, die genau auf dem Median liegen, per Zufallsprinzip in eine der beiden Gruppen eingeteilt, was, je nach Einteilung, zu stark voneinander abweichenden Ergebnissen führen kann. Aus diesem Grund wurde weiterhin versucht, die Berechnung ohne Median-Split durchzuführen. Hierfür wurden pro Variable fünf Subgruppen gebildet (entsprechend der jeweiligen Ausprägung) und für jede einzelne Subgruppe ein Netzwerk geschätzt sowie die Konnektivität berechnet. Hiermit wurde versucht, die nachteiligen Effekte der künstlichen Dichotomisierung anhand des Medians sowie der beschriebenen zufälligen Aufteilung entgegenzuwirken. Dabei sollten die Konnektivitäten nicht inferenzstatistisch miteinander verglichen werden (der Vergleich mehrerer Gruppen ist derzeit leider noch nicht in den R-Code des Network-Comparison-Tests implementiert), sondern lediglich anhand der grafischen Darstellung auf einen Trend untersucht werden, wodurch zumindest eine grobe Einschätzung des Zusammenhangs zwischen einzelnen Variablen und der Konnektivität des Netzwerks ermöglicht werden sollte. Hierbei ergaben sich jedoch erneut methodische Probleme. So waren die einzelnen Faktorstufen stark schiefverteilt und pro Faktorstufe nicht ansatzweise ausreichend Personen vorhanden, um die jeweiligen Netzwerke stabil genug schätzen zu können. Aus diesem Grund wurde dieses Vorgehen zwar für die im Rahmen dieser Arbeit durchgeführten Analysen ausgeschlossen, allerdings scheint es zunächst eine sinnvolle Ergänzung zum herkömmlichen Vorgehen der Impact-Berechnung zu sein. Dies sollte natürlich noch umfassend geprüft werden, was jedoch nicht Gegenstand dieser Arbeit ist.

Aufgrund der Tatsache, dass das einfache Vorgehen der Impact-Berechnung zu teils stark unterschiedlichen Ergebnissen führt, wurde beschlossen, die Berechnungen mit der Bootstrap-Methode zu kombinieren, sodass der Effekt der zufälligen Aufteilung ausgeglichen werden konnte. Hierdurch konnten einerseits die jeweiligen Berechnungen mehrere hundert Mal wiederholt werden und die jeweils entstehenden Kennwerte gemittelt werden, um stabilere Schätzer zu erhalten, andererseits konnten Konfidenzintervalle für die Schätzungen

berechnet werden, wodurch die Verteilungsbreiten der Schätzungen direkt ersichtlich wurden. Hierbei ergaben sich fünf Nodes, deren Konfidenzintervall nicht den Wert Null enthielten. Bei Betrachtung der zugehörigen Items („Furcht auf offenen Plätzen oder der Straße“, „Schwierigkeiten beim Atmen“, „Notwendigkeit, bestimmte Dinge, Orte, oder Tätigkeiten zu vermeiden, weil Sie hierdurch erschreckt werden“, „Starke Befangenheit im Umgang mit anderen“ sowie „Schreck- oder Panikanfälle“) fällt auf, dass diese überwiegend dem Störungskomplex der Angststörungen zugeordnet werden könnten. Tatsächlich entsprechen vier der Items Diagnosekriterien verschiedener Angststörungen (größtenteils im Bereich der Agoraphobie und/oder der Panikstörung), und das fünfte Item, „Schwierigkeiten beim Atmen“, stellt zwar kein Diagnosekriterium dar, ist jedoch so häufig eine Komponente bei verschiedenen Angststörungen, dass es als Symptom Einzug in gängige Modelle und Behandlungsansätze dieser gefunden hat (z.B. Heinrichs, Alpers, & Gerlach, 2009; Margraf & Schneider, 1990). Dieser Befund legt eine tragende Rolle dieser Items für die Konnektivität nahe, was sie als Ziele für Interventionen sowie präventive Maßnahmen prädestinieren würde. Allerdings sollten auch diese Befunde durch genauere Untersuchungen der Wirkmechanismen, die ihnen zugrunde liegen, abgesichert werden. Aus methodischer Sicht ist an diesem Ergebnis sowie am Konzept des Impacts im Allgemeinen interessant, dass hierdurch dargestellt werden kann, wie einzelne Nodes die globale Konnektivität ihres Netzwerks beeinflussen. Dies könnte Implikationen für die Methodik der Netzwerkanalysen im psychopathologischen Bereich mit sich bringen. Hier wurden Zusammenhänge zwischen verschiedenen Konstrukten und der Konnektivität bisher größtenteils auf syndromaler Ebene untersucht, beispielsweise in der Form der Frage, ob Personen mit einer diagnostizierten Depression eine höhere Konnektivität als Personen ohne diagnostizierte Depression aufweisen. Hierdurch wurde die Konnektivität als Funktion eines ganzen Syndroms modelliert. Allerdings legt die Berechnung von Impact-Werten, die positiv, negativ oder null sein können, nahe, dass es durchaus auf Symptomebene differentielle Effekte geben kann. Hier wäre es naheliegend, die Grundkonzeption der Netzwerkanalysen, also die Betrachtung komplexer Wirkzusammenhänge auf Symptomebene, auch auf die Untersuchung der Konnektivität anzuwenden. Dies hätte gegenüber dem aktuell gängigen Vorgehen den wesentlichen Vorteil, dass hierdurch nicht, wie derzeit üblich, die differentiellen Effekte der einzelnen Symptome auf die Konnektivität unter dem Konstrukt einer einzelnen Störung subsumiert werden oder bei der Analyse unberücksichtigt bleiben, weil sie einer Störung zugeordnet werden, die nicht untersucht wird. Zur Illustration dieses Grundgedankens soll kurz ein beispielhaftes Szenario entworfen werden.

Gegeben seien zwei psychische Störungen mit jeweils zwei Symptomen, A, B, C, und D. Werden nun die Effekte der einen psychischen Störung auf die Konnektivität auf herkömmliche Art und Weise modelliert, werden die einzelnen Effekte der zugehörigen Symptome A und B aggregiert dargestellt werden. Wenn diese nun Impact-Werte mit unterschiedlichem Vorzeichen haben, wird eine Untersuchung anhand des gemittelten Wertes nur verzerrte Ergebnisse liefern können. Weiterhin würde hierbei Symptom C, das eventuell auch einen hohen Impact-Wert aufweist, nicht berücksichtigt werden, da es nicht dem untersuchten Störungsbild zugeordnet wird.

Hierin lassen sich Ähnlichkeiten zu genau den Aspekten der statistischen Modellierung erkennen, die von den Proponenten der Netzwerkanalysen im psychologischen Bereich kritisiert werden (z.B. Fried & Nesse, 2015). Hier ist es vorstellbar, dass die konsequente Umsetzung der netzwerkanalytischen Grundgedanken eine Bereicherung des zugehörigen Methodenkanons darstellen könnte und dass sie zu differenzierteren und damit substantiierteren Ergebnissen hinsichtlich der Zusammenhänge zwischen den Netzwerkstrukturen und der Netzwerkdynamik führen könnte.

Fazit

Ziel dieser Arbeit war es, einen Einblick in die theoretische Konzeption der netzwerkanalytischen Ansätze in der klinischen Psychologie zu geben sowie ihre Anwendung an einem Beispiel zu demonstrieren. Hierbei sollten zunächst die praktischen Aspekte der Methodik vorgestellt und erläutert werden. Da sich diese und weiter die Vorteile der Netzwerkanalysen nicht zuletzt vor der Kritik an der in der klinischen Psychologie gängigen Art der Modellierung, der formativen Modellierung latenter Variablen, erschließt, sollte dieses methodische Paradigma in seinem theoretischen Fundament und seinen Anwendungen, hier in erster Linie in Form des Klassifikationssystems DSM-5, dargestellt und diskutiert werden.

Die wesentlichen Endpunkte der theoretischen und empirischen Untersuchung dieser Arbeit lassen sich einerseits im technisch-methodischen Teil verorten, andererseits im epistemologischen. Zunächst macht es den Eindruck, als stellen die Netzwerkanalysen eine wertvolle alternative Methodik zur Modellierung psychologischer Prozesse und Variablen dar. Die alternativen Techniken ihres Methodenkanons beruhen teilweise auf ähnlichen Grundgedanken wie die in der Psychologie weithin bekannten. Ein Beispiel hierfür ist in der Verwendung einer Regularisierung, um die Befunde um Scheinkorrelationen zu kontrollieren, zu sehen. Teilweise beruhen sie jedoch auch auf Grundgedanken, die sich aus dem

alternativen Grundverständnis ergeben. Beispielhaft hierfür könnte die Verwendung graphenanalytischer Methoden gesehen werden, die sich aus der Grundkonzeption psychologischer Phänomene als Netzwerke interagierender Variablen – Nodes – ergibt. Insgesamt ergeben sich hierdurch methodische Erweiterungen, die interessante Werkzeuge zur differenzierten Erforschung komplexer Wirkzusammenhänge darstellen können. Als solche müssen sie jedoch als das erkannt werden, was sie sind: einerseits statistische Verfahren, andererseits Grundlage für eine zwar noch in den Kinderschuhen steckende (Fried & Cramer, 2017), aber zunehmend umfangreicher und komplexer werdende Theorie verschiedener psychologischer Konstrukte. Als solche sind die Implikationen, die man unter der Verwendung ihrer Techniken ableiten kann, immer aus zumindest zwei Richtungen beschränkt. In ihrer Funktion als statistisches Werkzeug können durch sie Inferenzen immer nur so weit detailliert und damit wahrheitsgetreu ermittelt werden, wie es die Faktenlage zulässt. So wurden die Prozesse, die in dieser Arbeit erforscht werden sollten, beispielsweise über größtenteils querschnittliche Analysen untersucht. Die längsschnittlichen Untersuchungen, die durchgeführt wurden, die Hypothesentestungen, wurden anhand Erhebungen über große Zeiträume durchgeführt. Mit diesem Hintergrund haben sie einen anderen Geltungsbereich und erfassen längerfristige Effekte, als es beispielsweise mit täglichen Messungen oder sogar ESM-Erfassung möglich wäre. Dies muss jedoch nicht nachteilig gesehen werden, wenn man sich die Grundansprüche, die aus den verschiedenen Arten der Untersuchungsdesigns hervorgehen, vor Augen hält. Auf einer rein technischen Ebene implizieren die Ergebnisse dieser Arbeit, dass ein einzelnes im Rahmen der Netzwerkanalysen untersuchtes Konstrukt, das der Konnektivität, derzeit nicht so beforscht wird, wie man es bei konsequenter Umsetzung der Grundgedanken der Netzwerktheorie vermuten könnte. So scheint es, als würde sie derzeit rein auf syndromaler und damit einer recht globalen Ebene variieren. Die Ergebnisse dieser Arbeit legen jedoch nahe, dass sie als Funktion einzelner Symptome dargestellt werden können. Dieser Befund leitet sich jedoch nicht nur aus empirischen Ergebnissen ab, sondern auch aus der theoretischen Auseinandersetzung mit der Netzwerktheorie. In ihrer Funktion als Theorie, in ihrer Ähnlichkeit zur Systemtheorie vielleicht sogar Metatheorie, dient sie aus epistemologischer Sicht als Muster oder Schablone zur Deutung von letztendlich mathematischen Kennwerten. Als solches und vor dem Hintergrund der alternativen Methoden, Forschungsgegenstände und Grundkonzeptionen könnte sie als wissenschaftliches Paradigma verstanden werden, das im steten Austausch mit alternativen Paradigmen, wie dem latenten-Variablen-Modell, steht. Dies wird beispielsweise ersichtlich, wenn man sich vor Augen führt, welche Konstrukte

gängigerweise in den Netzwerkanalysen untersucht werden: die Konstrukte der latenten-Variablen-Modelle, etwa die DSM-5-Kriterien der Major Depressive Disorder. Dies muss jedoch nicht als Schwäche verstanden werden, da sich im Endeffekt nicht sagen lässt, ob das eine Paradigma oder das andere per se besser oder schlechter ist und beide einen berechtigten Anspruch auf Praktikabilität und Legitimität haben. Dieser Anspruch ergibt sich jedoch nur, wenn sie als das verstanden werden, was sie sind: als schlussendlich kommunikative Strukturen zur Erschließung selbst festgesetzter und damit umgrenzter Inhalte.

Literaturverzeichnis

- Allison, P.D. (2002). *Missing Data*. Thousand Oaks: Sage Publications.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5. Auflage). Washington, DC: Author.
- Aragona, M. (2015). Rethinking received views on the history of psychiatric nosology: Minor shifts, major continuities. In P. Zachar, D. St. Stoyanov, M. Aragona & A. Jabelensky (Hrsg.), *Alternative perspectives on psychiatric validation: DSM, ICD, RDoC, and beyond* (S. 27-46). Oxford: Oxford University Press.
- Baltes-Götz, B. (2013). *Behandlung fehlender Werte in SPSS und Amos*. Zugriff am 02.07.2019. Verfügbar unter <https://www.uni-trier.de/fileadmin/urt/doku/bfw/bfw.pdf>
- Barabási, A. L. (2012). The network takeover. *Nature Physics*, 8, 14-16. doi:10.1038/nphys2188
- Bassler, M., Potratz, B., & Krauthauser, H. (1995). Der "Helping Alliance Questionnaire" (HAQ) von Luborsky. Möglichkeiten zur Evaluation des therapeutischen Prozesses von stationärer Psychotherapie. *Psychotherapeut*, 40(1), 23-32.
- Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, 16(1), 5-13. doi:10.1002/wps.20375
- Borsboom, D., & Cramer, A. O. J. (2013). Network Analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 9(1), 91-121. doi:10.1146/annurev-clinpsy-050212-185608
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, 110(2), 203-219. doi:10.1037/0033-295X.110.2.203
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, 111(4), 1061-1071. doi:10.1037/0033-295X.111.4.1061
- Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Auflage). Berlin: Springer.
- Brown, T. A. (2013). Latent variable measurement models. In T. D. Little (Hrsg.), *The Oxford handbook of quantitative methods: Volume 2 statistical analysis* (S. 257-280). New York: Oxford University Press.

- Carlin, J. B., & Doyle, L. B. (2001). Statistics for clinicians 4: Basic concepts of statistical reasoning: Hypothesis tests and the *t*-test. *Journal of Paediatrics and Child Health*, 37(1), 72-77. doi:10.1046/j.1440-1754.2001.00634.x
- Cohen, S. (1988). *Statistical power analysis for the behavioral sciences* (2. Auflage). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, 33, 137-150. doi:10.1017/S0140525X09991567
- Danon, L., Ford, A. P., House, T., Jewell, C. P., Keeling, M. J., Roberts, G. O., . . . Vernon, M. C. (2011). Networks and the epidemiology of infectious disease. *Interdisciplinary Perspectives on Infectious Diseases*. 1-28. doi:10.1155/2011/284909
- Derogatis, L. R. (1993). *BSI Brief Symptom Inventory: Administration, scoring, and procedures manual* (4. Auflage). Minneapolis: National Computer Systems.
- Ebbinghaus, H. (1908). *Abriss der Psychologie*. Leipzig: Veit & Comp.
- Epskamp, S. (2017). *Network Psychometrics*. Dissertation, Universität Amsterdam.
- Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, 50(1), 195-212. doi:10.3758/s13428-017-0862-1
- Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, 48(4), 1-18.
- Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, 23(4), 617-634. doi:10.1037/met0000167
- Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A. Jr., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry*, 26(1), 57-63. doi:10.1001/archpsyc.1972.01750190059011
- Field., A. (2009). *Discovering statistics using SPSS* (3. Auflage). London: Sage Publications Ltd.
- Franke, G. H. (2000). *BSI: Brief Symptom Inventory von L.R. Derogatis (Kurzform der SCL-90-R) – deutsche Version: Manual*. Göttingen: Beltz Test.

- Fried, E. I., & Cramer, A. O. J. (2017). Moving forward: Challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science*, 12(6), 999-1020. doi: 10.1177/1745691617705892
- Fried, E. I., Epskamp, S., Nesse, R. M., Tuerlinckx, F., & Borsboom, D. (2016). What are 'good' depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *Journal of Affective Disorders*, 189, 314–320. doi:10.1016/j.jad.2015.09.005
- Fried, E. I., & Nesse, R. M. (2015). Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. *BMC Medicine*, 13:72, 1-11. doi:10.1186/s12916-015-0325-4
- Fried, E. I., Nesse, R. M., Zivin, K., Guille, C., & Sen, S. (2014). Depression is more than the sum score of its parts: individual DSM symptoms have different risk factors. *Psychological Medicine*, 44(10), 2067-2076. doi:10.1017/S0033291713002900
- Fried, E. I., van Borkulo, C. D., Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: a review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, 52(1), 1-10. doi:10.1007/s00127-016-1319-z
- Hautzinger, M., Keller, F., & Kühner, C. (2006). *BDI-II. Beck Depressions Inventar Revision Manual*. Frankfurt: Harcourt Test Services.
- Hayes, N. (1995). *Psychology in Perspective*. London: Palgrave.
- Heinrichs, N., Alpers, G. W., & Gerlach, A. L. (2009). *Evidenzbasierte Leitlinie zur Psychotherapie der Panikstörung und Agoraphobie*. Göttingen: Hogrefe.
- Hevey, D. (2018). Network Analysis: A brief overview and tutorial. *Health Psychology and Behavioral Medicine*, 6, 301-328. doi:10.1080/21642850.2018.1521283
- Hyman, S. E. (2010). The diagnosis of mental disorders: The problem of reification. *Annual Review of Clinical Psychology*, 6, 155-179. doi:10.1146/annurev.clinpsy.3.022806.091532
- Jagsch, R. (2015). *Endbericht Evaluierung des arbeitsbezogenen Therapieerfolgs der klinisch-psychologischen und psychotherapeutischen Behandlung im Rahmen des Projekts „fit2work“*. Wien: Universität, Institut für Angewandte Psychologie: Gesundheit, Entwicklung und Förderung.

- Jones, P. (2017). *Impact: A new statistic for network analysis (networktools)*. Zugriff am 23.07.2019. Verfügbar unter <https://cran.r-project.org/web/packages/networktools/vignettes/Impact.pdf>
- Jones, P. (2019). *networktools: Tools for Identifying Important Nodes in Networks. R package version 1.2.1*. Zugriff am 28.07.2019. Verfügbar unter <https://CRAN.R-project.org/package=networktools>
- Kanfer, F. H., & Saslow, G. (1965). Behavioral analysis: An alternative to diagnostic classification. *Archives of General Psychiatry*, 12(6), 529-538. doi:10.1001/archpsyc.1965.01720360001001
- Kornberger, M. W. (2017). Circulus-vitiosus-Modelle in der Verhaltenstherapie: Aktueller Stand, kritische Betrachtung und mögliche Alternativen. *Verhaltenstherapie & psychosoziale Praxis*, 49(1), 59-67.
- Kraemer, H. C. (2007). DSM categories and dimensions in clinical and research contexts. *International Journal of Methods in Psychiatric Research*, 16, 8-15. doi:10.1002/mpr.211
- Krueger, R. F. (2002). Psychometric perspectives on comorbidity. In J. E. Helzer & J. J. Hudziak (Hrsg.), *American Psychopathological Association series. Defining psychopathology in the 21st century: DSM-V and beyond* (S. 41-54). Arlington, VA: American Psychiatric Publishing, Inc.
- Liebowitz, M. R. (1987). Social phobia. *Modern Problems of Pharmacopsychiatry*, 22, 141–173. doi:10.1159/000414022
- Maser, J. D., & Cloninger, C. R. (1990). Comorbidity of anxiety and mood disorders: Introduction and overview. In J. D. Maser & C. R. Cloninger (Hrsg.), *Comorbidity of mood and anxiety disorders* (S. 3–12). Washington, DC: American Psychiatric Press.
- Margraf, J. (2009). Hintergründe und Entwicklung. In J. Margraf & S. Schneider (Hrsg.), *Lehrbuch der Verhaltenstherapie: Band 1: Grundlagen, Diagnostik, Verfahren, Rahmenbedingungen* (3. Auflage) (S. 3-46). Heidelberg: Springer.
- Margraf, J., & Ehlers, A. (1989). Etiological models of panic – Psychophysiological and cognitive aspects. In R. Baker (Hrsg.), *Panic disorder: Theory, research and therapy* (S. 205-231). Oxford: Wiley & Sons.
- Margraf, J., & Schneider, S. (1990). *Panik: Angstanfälle und ihre Behandlung* (2. Auflage). Berlin: Springer.

- Mill, J. S. (1875). *A system of logic, ratiocinative and inductive: Being a connected view of the principles of evidence, and the methods of scientific investigation* (9. Auflage). London: Longmans, Green, Reader and Dyer.
- Newman, M. E. J. (2010). *Networks: An introduction*. New York: Oxford University Press.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), 943-951. doi:10.1126/science.aac4716
- Pe, M. L., Kircanski, K., Thompson, R. J., Bringmann, L. F., Tuerlinckx, F., Mestdagh, M., ..., Gotlib, I. H. (2015). Emotion-network density in major depressive disorder. *Clinical Psychological Science*, 3(2), 292-300. doi:10.1177/2167702614540645
- Pearl, J. (2009). *Causality* (2. Auflage). New York: Cambridge University Press.
- Pomerantz, M. M., & Freedman, M. L. (2011). The genetics of cancer risk. *Cancer Journal*, 17(6), 416-422. doi:10.1097/PPO.0b013e31823e5387
- R Core Team. (2018). *foreign: Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', R package version 0.8-71*. Zugriff am 15.05.2019. Verfügbar unter <https://CRAN.R-project.org/package=foreign>.
- Rinck, M., & Becker, E. S. (2011). Lernpsychologische Grundlagen. In H.-U. Wittchen & J. Hoyer (Hrsg.), *Klinische Psychologie & Psychotherapie* (S. 107-128). Berlin: Springer.
- Robins, E., & Guze, S. B. (1970). Establishment of diagnostic validity in psychiatric illness: Its application to schizophrenia. *American Journal of Psychiatry*, 126(7), 983-987. doi:10.1176/ajp.126.7.983
- Rodebaugh, T. L., Tonge, N. A., Piccirillo, M. L., Fried, E., Horenstein, A., Morrison, A. S., ..., Heimberg, A. G. (2018). Does centrality in a cross-sectional network suggest intervention targets for social anxiety disorder? *Journal of Consulting and Clinical Psychology*, 86(10), 831-844. doi:10.1037/ccp0000336.
- Sabbarton-Leary, N., Bortolotti, L., & Broome, M. R. (2015). Natural and para-natural kinds in psychiatry. In P. Zachar, D. St. Stoyanov, M. Aragona & A. Jabelensky (Hrsg.), *Alternative perspectives on psychiatric validation: DSM, ICD, RDoC, and beyond* (S. 76-93). Oxford: Oxford University Press.
- Sanders, J. L. (2011). A distinct language and a historic pendulum: the evolution of the Diagnostic and Statistical Manual of Mental Disorders. *Archives of Psychiatric Nursing*, 25(6), 394-403. doi:10.1016/j.apnu.2010.10.002

- Schmittmann, V. D., Cramer, A. O. J., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2011). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, 31(1), 43-53. doi:10.1016/j.newideapsych.2011.02.007
- Schweren, L., van Borkulo, C. D., Fried, E., & Goodyer, I. M. (2018). Assessment of symptom network density as a prognostic marker of treatment response in adolescent depression. *JAMA Psychiatry*, 75(1), 98-100. doi:10.1001/jamapsychiatry.2017.3561
- Stein, D. J., Phillips, K. A., Bolton, D., Fulford, K. W. M., Sadler, J. Z., & Kendler, K. S. (2010). What is a mental/psychiatric disorder? From DSM-IV to DSM-V. *Psychological Medicine*, 40(11), 1759-1765. doi:10.1017/S0033291709992261
- Sternberg, R. J., & Sternberg, K. (2012). *Cognitive Psychology* (6. Auflage). Wadsworth, CA: Cengage Learning, Inc.
- Terluin, B., de Boer, M. R., & de Vet, H. C. W. (2016). Differences in connection strength between mental symptoms might be explained by differences in variance: Reanalysis of network data did not confirm staging. *PLoS ONE*, 11(11). doi:10.1371/journal.pone.0155205
- Urbanek, S., & Horner, J. (2019). Cairo: R graphics device using cairo graphics library for creating high-quality bitmap (PNG, JPEG, TIFF), vector (PDS, SVG, PostScript) and display (x11 and Win32) output. R package, version 1.5-10. Zugriff am 27.07.2019. Verfügbar unter <https://CRAN.R-project.org/package=Cairo>
- van Borkulo, C., Boschloo, L., Borsboom, D., Penninx, B. W. J. H., Waldorp, L. J., & Schoevers, R. A. (2015). Association of symptom network structure with the course of longitudinal depression. *JAMA Psychiatry*, 72(12), 1219-1226. doi:10.1001/jamapsychiatry.2015.2079
- van Borkulo, C., Boschloo, L., Kossakowski, J. J., Tio, P., Schoevers, R. A., Borsboom, D., & Waldorp, L. J. (2017). Comparing network structure aspects: A permutation test. *Working Paper*. doi:10.13140/RG.2.2.29455.38569
- van Loo, H. M., & Romejin, J. W. (2015). Psychiatric comorbidity: fact or artifact? *Theoretical Medicine and Bioethics*, 36(1), 41-60. doi:10.1007/s11017-015-9321-0
- Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of *p* value. *Psychonomic Bulletin & Review*, 14(5), 779-804. doi:10.3758/BF03194105

- Wakefield, J. C. (2007). The concept of mental disorder: Diagnostic implications of the harmful dysfunction analysis. *World Psychiatry*, 6(3), 149-156.
- Walters, E. (2016). The p -value and the problem of multiple testing. *Reproductive BioMedicine Online*, 32(4), 348-349. doi:10.1016/j.rbmo.2016.02.008
- Weiner, P. (2015). *Informationsblatt „Ablauf fit2work Personenberatung“*. Zugriff am 02.07.2019. Verfügbar unter <http://fit2work.at/artikel/downloads-materialien-3>
- Weltgesundheitsorganisation. (1986). *Ottawa-Charta zur Gesundheitsförderung*. Zugriff am 20.07.2019. Verfügbar unter http://www.euro.who.int/__data/assets/pdf_file/0006/129534/Ottawa_Charter_G.pdf
- Widiger, T. A., & Edmundson, M. (2011). Diagnoses, dimensions, and DSM-5. In P. E. Nathan & D. H. Barlow (Hrsg.), *The Oxford handbook of clinical psychology* (S. 256-280). New York: Oxford University Press.
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York: Springer.
- Wigman, J. T., van Os, J., Borsboom, D., Wardenaar, K. J., Epskamp, S., Klippel, A., ..., Wichers, M. (2015). Exploring the underlying structure of mental disorders: cross-diagnostic differences and similarities from a network perspective using both a top-down and a bottom-up approach. *Psychological Medicine*, 45(11), 2375-2387. doi:10.1017/S0033291715000331
- Wittchen, H.-U., & Hoyer, J. (2011). Diagnostische Klassifikation psychischer Störungen. In H.-U. Wittchen & J. Hoyer (Hrsg.), *Klinische Psychologie & Psychotherapie* (S. 27-55). Berlin: Springer.
- Zachar, P. (2015). Psychiatric disorders: natural kinds made by the world or practical kinds made by us? *World Psychiatry*, 14(3), 288-290. doi:10.1002/wps.20240
- Zachar, P., & Jablensky, A. (2015). Introduction: The concept of validation in psychiatry and psychology. In P. Zachar, D. St. Stoyanov, M. Aragona & A. Jabelensky (Hrsg.), *Alternative perspectives on psychiatric validation: DSM, ICD, RDoC, and beyond* (S. 3-26). Oxford: Oxford University Press.
- Zhao, P., & Yu, B. (2006). On model selection consistency of LASSO. *Journal of Machine Learning Research*, 7(12), 2541-2563.

Abbildungsverzeichnis

Abbildungsnr.	Abbildungstitel	Seite
Abbildung 1	Bootstrap-Schätzungen zu t_1 , deren Konfidenzintervalle nicht Null enthalten.	38
Abbildung 2	Bootstrap-Schätzungen zu t_2 , deren Konfidenzintervalle nicht Null enthalten.	39
Abbildung 3	Symptomnetzwerk des BSI-53 zu t_1 .	39
Abbildung 4	Symptomnetzwerk des BSI-53 zu t_2 .	40
Abbildung 5	Symptomnetzwerk des BSI-53 zu t_1 mit sämtlichen Korrelationen.	73
Abbildung 6	Symptomnetzwerk des BSI-53 zu t_2 mit sämtlichen Korrelationen.	73
Abbildung 7	Grafische Darstellung der Zentralitätsmaße zu t_1 und t_2 .	43
Abbildung 8	Grafische Darstellung der Impact-Werte zu t_1 mit den dazugehörigen Konfidenzintervallen der Items des BSI-53.	45
Abbildung 9	Symptomnetzwerk des BSI-53 der Responder zu t_1 .	48
Abbildung 10	Symptomnetzwerk des BSI-53 der Non-Responder zu t_1 .	48
Abbildung 11	Symptomnetzwerk des BSI-53 der Responder zu t_1 mit sämtlichen Korrelationen.	74
Abbildung 12	Symptomnetzwerk des BSI-53 der Non-Responder zu t_1 mit sämtlichen Korrelationen.	74

Tabellenverzeichnis

Tabellennr.	Tabellentitel	Seite
Tabelle 1	Soziodemografische Merkmale der Analysestichprobe	35
Tabelle 2	Darstellung der Edges zu t_1 mit $r \geq .30$	41
Tabelle 3	Darstellung der Edges zu t_2 mit $r \geq .30$	42
Tabelle 4	Darstellung der Edges zu t_1 mit $r \geq .10$	75/76
Tabelle 5	Darstellung der Edges zu t_2 mit $r \geq .10$	77/78
Tabelle 6	Zentralitätsmaße des BSI-53-Netzwerks zu t_1 und t_2	44
Tabelle 7	Impact-Werte des BSI-53 zu t_1	46
Tabelle 8	Darstellung der Edges der Responder zu t_1 mit $r \geq .10$	49
Tabelle 9	Darstellung der Edges der Non-Responder zu t_1 mit $r \geq .30$	50
Tabelle 10	Darstellung der Edges der Non-Responder zu t_1 mit $r \geq .10$	79/80

Anhang

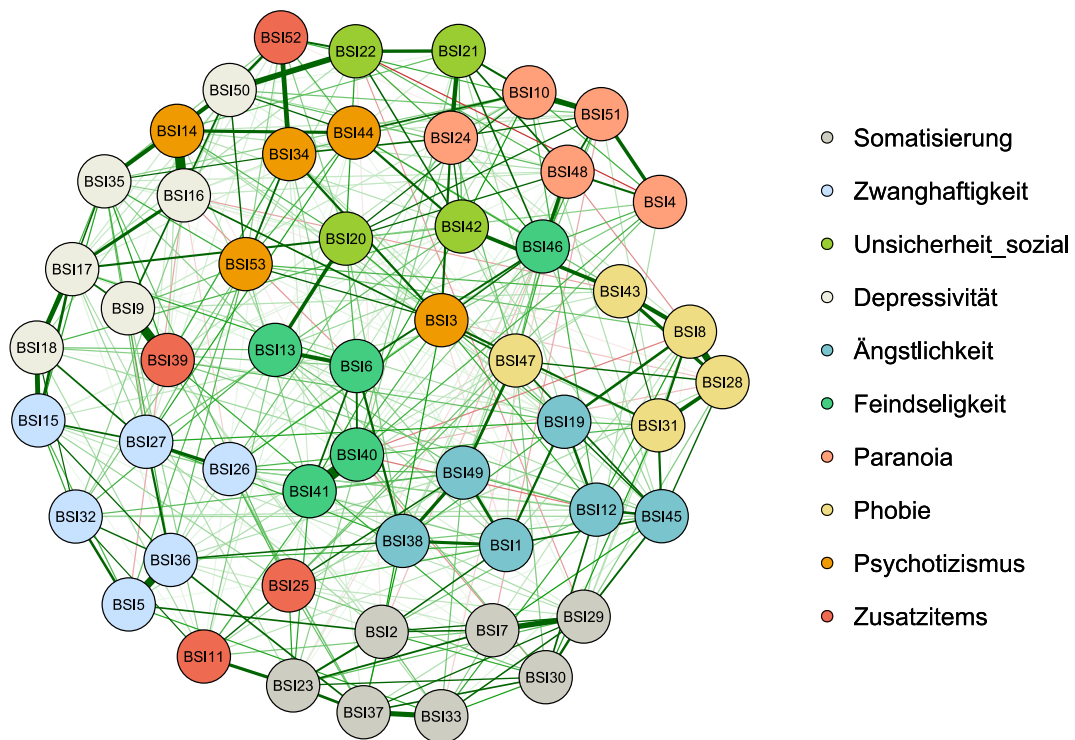


Abbildung 5. Symptomnetzwerk des BSI-53 zu t_1 mit sämtlichen Korrelationen.

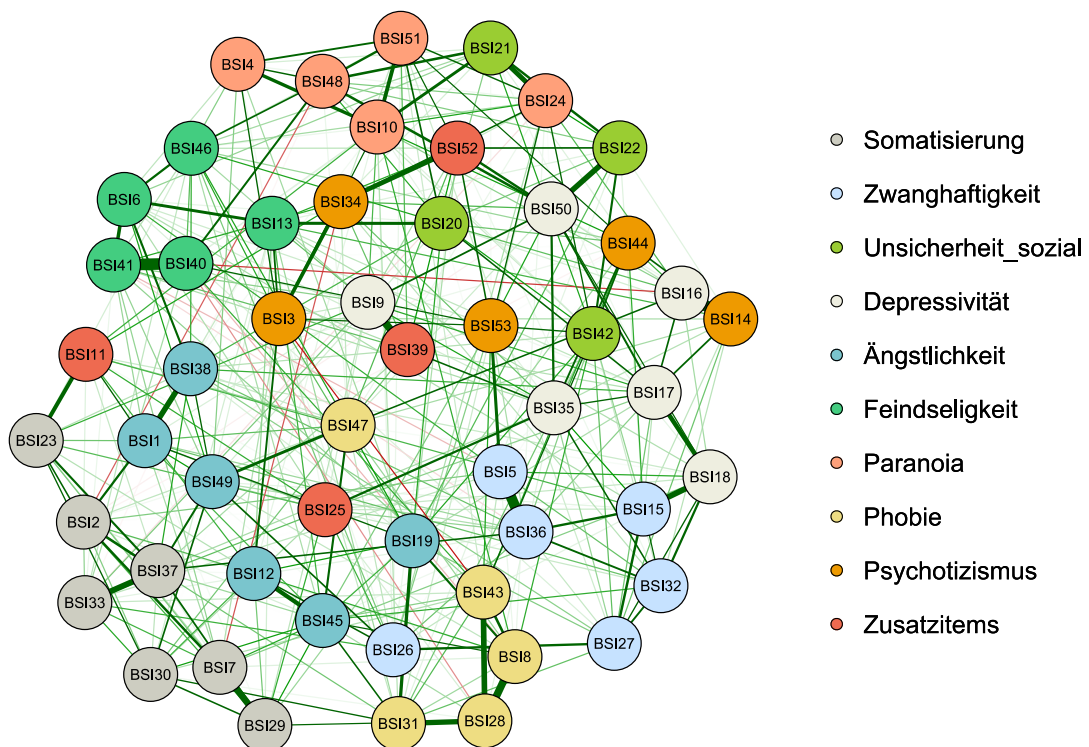


Abbildung 6. Symptomnetzwerk des BSI-53 zu t_2 mit sämtlichen Korrelationen.

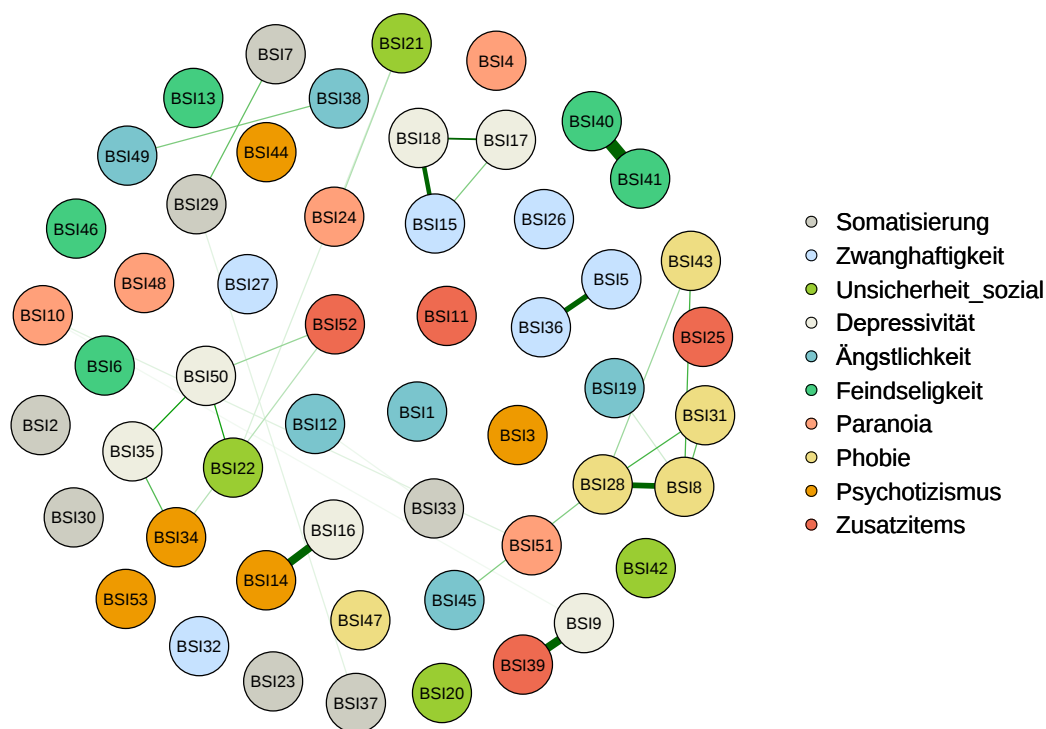


Abbildung 11. Symptomnetzwerk des BSI-53 der Responder zu t_1 mit sämtlichen Korrelationen.

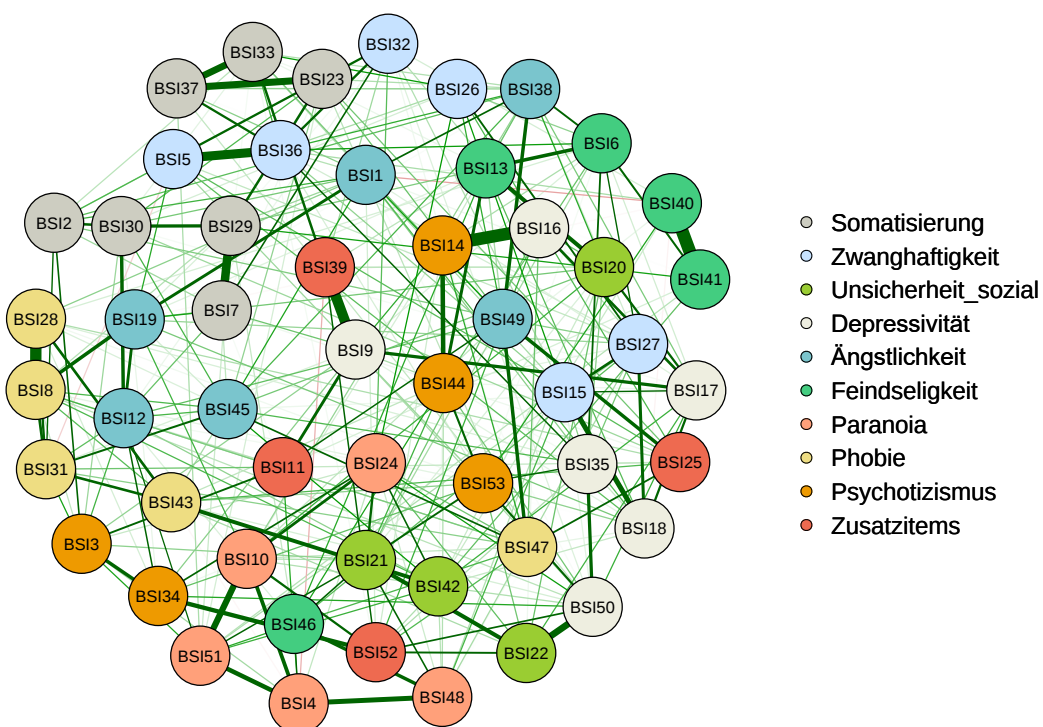


Abbildung 12. Symptomnetzwerk des BSI-53 der Non-Responder zu t_1 mit sämtlichen Korrelationen.

Tabelle 4 (Teil 1)

Darstellung der Edges zu t_1 mit $r \geq .10$

Node 1	Node 2	r	SD	95% K
BSI40	BSI41	.66	.04	[.58, .74]
BSI14	BSI16	.56	.04	[.50, .64]
BSI9	BSI39	.55	.04	[.47, .63]
BSI5	BSI36	.48	.04	[.40, .56]
BSI8	BSI28	.40	.04	[.32, .49]
BSI22	BSI50	.34	.04	[.26, .42]
BSI7	BSI29	.32	.04	[.24, .40]
BSI10	BSI51	.32	.04	[.23, .40]
BSI33	BSI37	.30	.05	[.21, .39]
BSI15	BSI18	.29	.04	[.20, .37]
BSI17	BSI18	.28	.04	[.21, .36]
BSI34	BSI52	.28	.05	[.18, .39]
BSI21	BSI24	.26	.04	[.18, .33]
BSI8	BSI43	.26	.04	[.17, .34]
BSI35	BSI50	.25	.04	[.17, .33]
BSI6	BSI13	.24	.04	[.16, .32]
BSI42	BSI43	.23	.04	[.15, .30]
BSI28	BSI31	.22	.05	[.13, .32]
BSI13	BSI20	.22	.04	[.14, .29]
BSI4	BSI51	.21	.04	[.13, .30]
BSI46	BSI48	.21	.04	[.12, .29]
BSI38	BSI49	.20	.04	[.12, .28]
BSI26	BSI27	.20	.04	[.12, .28]
BSI28	BSI43	.20	.04	[.12, .28]
BSI21	BSI22	.20	.04	[.11, .28]
BSI11	BSI23	.20	.05	[.10, .29]
BSI16	BSI17	.19	.03	[.12, .26]
BSI14	BSI44	.19	.04	[.12, .26]
BSI1	BSI38	.19	.04	[.10, .27]
BSI47	BSI49	.18	.05	[.09, .27]
BSI1	BSI49	.17	.04	[.09, .26]
BSI15	BSI17	.17	.04	[.10, .24]
BSI23	BSI37	.16	.04	[.08, .25]
BSI8	BSI19	.16	.04	[.08, .24]
BSI5	BSI32	.16	.05	[.06, .25]
BSI12	BSI45	.16	.04	[.07, .24]
BSI12	BSI19	.15	.05	[.06, .25]
BSI50	BSI52	.15	.04	[.07, .23]
BSI29	BSI30	.15	.05	[.06, .24]
BSI3	BSI34	.14	.06	[.03, .26]
BSI1	BSI19	.14	.04	[.06, .23]
BSI42	BSI44	.14	.04	[.06, .23]
BSI8	BSI31	.14	.05	[.05, .23]
BSI6	BSI38	.14	.04	[.07, .21]
BSI27	BSI36	.14	.04	[.07, .21]

Tabelle 4 (Teil 2)

Darstellung der Edges zu t_1 mit $r \geq .10$

Node 1	Node 2	r	SD	95% KI
BSI31	BSI45	.14	.04	[.06, .21]
BSI3	BSI31	.14	.05	[.05, .23]
BSI3	BSI24	.14	.05	[.04, .23]
BSI4	BSI48	.13	.04	[.05, .22]
BSI17	BSI20	.13	.03	[.06, .20]
BSI10	BSI21	.13	.04	[.04, .21]
BSI2	BSI23	.13	.05	[.04, .22]
BSI6	BSI46	.12	.04	[.04, .21]
BSI22	BSI52	.12	.05	[.03, .21]
BSI6	BSI40	.12	.06	[.00, .24]
BSI6	BSI41	.12	.05	[.01, .22]
BSI48	BSI51	.12	.04	[.04, .19]
BSI29	BSI45	.12	.04	[.03, .20]
BSI7	BSI23	.12	.04	[.03, .20]
BSI34	BSI53	.11	.05	[.01, .22]
BSI18	BSI27	.11	.04	[.03, .20]
BSI15	BSI27	.11	.04	[.03, .19]

Anmerkung. r = partielle polychorische Korrelation; SD = Standardabweichung, KI = Konfidenzintervall

Tabelle 5 (Teil 1)

Darstellung der Edges zu t_2 mit $r \geq .10$

Node 1	Node 2	r	SD	95% KI
BSI40	BSI41	.58	.05	[.48, .69]
BSI14	BSI16	.57	.04	[.49, .65]
BSI5	BSI36	.53	.04	[.44, .62]
BSI8	BSI28	.39	.06	[.28, .50]
BSI9	BSI39	.39	.05	[.28, .49]
BSI7	BSI29	.35	.05	[.25, .45]
BSI33	BSI37	.31	.05	[.22, .40]
BSI15	BSI18	.30	.04	[.21, .39]
BSI28	BSI43	.30	.05	[.19, .41]
BSI22	BSI50	.29	.05	[.20, .39]
BSI1	BSI38	.29	.05	[.20, .39]
BSI28	BSI31	.28	.05	[.18, .39]
BSI21	BSI24	.27	.05	[.17, .36]
BSI34	BSI52	.26	.06	[.14, .37]
BSI17	BSI18	.24	.04	[.16, .32]
BSI12	BSI45	.23	.05	[.13, .33]
BSI11	BSI23	.22	.05	[.12, .33]
BSI10	BSI51	.21	.05	[.12, .31]
BSI3	BSI34	.21	.06	[.09, .33]
BSI6	BSI41	.19	.05	[.09, .30]
BSI4	BSI10	.19	.05	[.09, .28]
BSI19	BSI31	.19	.05	[.08, .29]
BSI13	BSI20	.18	.05	[.08, .28]
BSI5	BSI53	.17	.05	[.07, .27]
BSI10	BSI21	.17	.05	[.08, .26]
BSI6	BSI13	.17	.05	[.07, .27]
BSI47	BSI49	.17	.05	[.06, .27]
BSI21	BSI48	.15	.05	[.06, .24]
BSI8	BSI43	.15	.06	[.03, .27]
BSI2	BSI37	.15	.05	[.06, .24]
BSI48	BSI50	.15	.05	[.06, .25]
BSI50	BSI52	.15	.05	[.05, .25]
BSI21	BSI22	.15	.05	[.05, .24]
BSI2	BSI7	.15	.05	[.05, .25]
BSI15	BSI27	.15	.05	[.05, .24]
BSI48	BSI51	.14	.04	[.05, .23]
BSI15	BSI36	.14	.04	[.07, .22]
BSI26	BSI27	.14	.05	[.04, .23]
BSI18	BSI32	.14	.05	[.04, .23]
BSI6	BSI46	.14	.05	[.03, .24]
BSI1	BSI2	.14	.05	[.04, .23]
BSI25	BSI35	.13	.05	[.03, .23]
BSI35	BSI50	.13	.04	[.05, .22]
BSI23	BSI37	.13	.04	[.05, .22]
BSI8	BSI19	.13	.05	[.04, .23]

Tabelle 5 (Teil 2)

Darstellung der Edges zu t_2 mit $r \geq .10$

Node 1	Node 2	r	SD	95% KI
BSI6	BSI38	.13	.05	[.04, .22]
BSI45	BSI47	.13	.06	[.02, .24]
BSI22	BSI42	.13	.05	[.03, .22]
BSI17	BSI50	.13	.04	[.04, .21]
BSI40	BSI48	.13	.06	[.01, .24]
BSI37	BSI49	.12	.05	[.03, .22]
BSI27	BSI32	.12	.05	[.02, .22]
BSI3	BSI12	.12	.06	[.01, .23]
BSI4	BSI51	.12	.05	[.02, .22]
BSI35	BSI44	.12	.05	[.02, .22]
BSI32	BSI36	.12	.04	[.03, .20]
BSI3	BSI13	.12	.05	[.01, .23]
BSI9	BSI50	.12	.05	[.01, .22]
BSI14	BSI17	.12	.04	[.03, .20]
BSI29	BSI37	.11	.05	[.02, .20]
BSI14	BSI44	.11	.05	[.01, .21]
BSI46	BSI48	.11	.05	[.01, .21]
BSI20	BSI51	.11	.04	[.02, .20]
BSI42	BSI44	.11	.05	[.01, .21]
BSI16	BSI17	.11	.04	[.02, .20]
BSI26	BSI49	.11	.05	[.01, .21]
BSI30	BSI49	.11	.05	[.01, .20]
BSI22	BSI52	.11	.04	[.02, .20]
BSI35	BSI39	.11	.05	[.01, .20]
BSI24	BSI34	.11	.05	[.01, .20]
BSI31	BSI45	.11	.05	[.01, .20]
BSI2	BSI23	.11	.05	[.01, .20]
BSI32	BSI35	.10	.05	[.01, .20]
BSI29	BSI30	.10	.05	[.01, .20]
BSI36	BSI37	.10	.04	[.02, .18]
BSI12	BSI26	.10	.05	[.01, .20]
BSI52	BSI53	.10	.05	[.00, .20]
BSI16	BSI47	.10	.05	[.00, .20]
BSI17	BSI20	.10	.04	[.02, .18]

Anmerkung. r = partielle polychorische Korrelation; SD = Standardabweichung, KI = Konfidenzintervall

Tabelle 10 (Teil 1)

Darstellung der Edges der Non-Responder mit $r \geq .10$

Node 1	Node 2	<i>r</i>	<i>SD</i>	95% KI
BSI14	BSI16	.51	.05	[.40 , .62]
BSI40	BSI41	.47	.10	[.28 , .66]
BSI8	BSI28	.43	.06	[.31 , .55]
BSI9	BSI39	.42	.07	[.29 , .56]
BSI5	BSI36	.39	.08	[.24 , .55]
BSI7	BSI29	.36	.07	[.22 , .50]
BSI33	BSI37	.35	.07	[.21 , .49]
BSI34	BSI52	.31	.08	[.16 , .47]
BSI14	BSI44	.30	.06	[.18 , .42]
BSI22	BSI50	.30	.07	[.16 , .43]
BSI4	BSI51	.28	.07	[.13 , .43]
BSI23	BSI37	.27	.07	[.14 , .41]
BSI10	BSI51	.26	.06	[.13 , .38]
BSI16	BSI17	.25	.06	[.12 , .38]
BSI47	BSI49	.25	.07	[.11 , .39]
BSI42	BSI43	.24	.07	[.10 , .38]
BSI31	BSI45	.20	.06	[.08 , .33]
BSI12	BSI30	.20	.07	[.06 , .35]
BSI10	BSI24	.20	.06	[.08 , .32]
BSI2	BSI30	.20	.07	[.06 , .34]
BSI35	BSI50	.20	.06	[.08 , .32]
BSI15	BSI18	.20	.07	[.05 , .34]
BSI6	BSI40	.20	.07	[.05 , .34]
BSI6	BSI13	.19	.07	[.05 , .33]
BSI4	BSI48	.19	.07	[.04 , .33]
BSI21	BSI22	.18	.06	[.05 , .32]
BSI29	BSI30	.18	.07	[.03 , .33]
BSI46	BSI48	.18	.07	[.04 , .31]
BSI18	BSI27	.18	.08	[.02 , .33]
BSI21	BSI24	.17	.07	[.04 , .31]
BSI36	BSI38	.17	.07	[.04 , .31]
BSI26	BSI49	.17	.06	[.05 , .30]
BSI38	BSI49	.17	.06	[.04 , .30]
BSI8	BSI19	.17	.06	[.06 , .28]
BSI17	BSI18	.17	.06	[.05 , .29]
BSI1	BSI19	.16	.07	[.02 , .30]
BSI31	BSI43	.16	.07	[.03 , .30]
BSI36	BSI37	.16	.06	[.04 , .28]
BSI28	BSI31	.16	.07	[.01 , .31]
BSI34	BSI53	.16	.07	[.02 , .30]
BSI13	BSI20	.15	.07	[.01 , .30]
BSI33	BSI39	.15	.06	[.03 , .28]
BSI5	BSI32	.15	.07	[.01 , .29]
BSI4	BSI10	.15	.07	[.01 , .29]
BSI25	BSI49	.15	.07	[.01 , .29]

Darstellung der Edges der Non-Responder mit $r \geq .10$

Node 1	Node 2	r	SD	95% KI
BSI2	BSI3	.15	0.08	[.01, .30]
BSI8	BSI31	.15	0.07	[.01, .28]
BSI1	BSI38	.14	0.07	[.01, .28]
BSI26	BSI27	.14	0.07	[.01, .28]
BSI44	BSI46	.14	0.07	[.01, .28]
BSI12	BSI19	.14	0.07	[.01, .27]
BSI43	BSI53	.14	0.07	[.00, .27]
BSI15	BSI27	.14	0.07	[.00, .27]
BSI25	BSI37	.13	0.07	[.00, .27]
BSI28	BSI43	.13	0.07	[.00, .27]
BSI22	BSI52	.13	0.07	[.01, .26]
BSI8	BSI43	.13	0.07	[.01, .26]
BSI12	BSI34	.13	0.07	[.02, .27]
BSI9	BSI14	.12	0.07	[.01, .26]
BSI3	BSI34	.12	0.08	[.04, .28]
BSI17	BSI20	.12	0.06	[.01, .24]
BSI9	BSI11	.12	0.08	[.03, .27]
BSI10	BSI19	.12	0.06	[.01, .23]
BSI15	BSI17	.12	0.06	[.00, .24]
BSI32	BSI36	.12	0.06	[.01, .25]
BSI12	BSI45	.12	0.07	[.01, .25]
BSI45	BSI47	.12	0.07	[.02, .25]
BSI15	BSI35	.11	0.06	[.00, .23]
BSI42	BSI51	.11	0.05	[.01, .22]
BSI4	BSI46	.11	0.07	[.03, .24]
BSI2	BSI7	.11	0.06	[.02, .24]
BSI50	BSI53	.11	0.06	[.01, .23]
BSI25	BSI30	.11	0.06	[.01, .23]
BSI24	BSI50	.11	0.06	[.00, .22]
BSI6	BSI27	.11	0.05	[.01, .21]
BSI29	BSI41	.10	0.06	[.02, .23]
BSI13	BSI38	.10	0.06	[.01, .22]
BSI13	BSI44	.10	0.06	[.02, .22]
BSI11	BSI49	.10	0.06	[.02, .22]
BSI20	BSI48	.10	0.06	[.02, .22]
BSI6	BSI35	.10	0.05	[.01, .21]

Anmerkung. r = partielle polychorische Korrelation; SD = Standardabweichung, KI = Konfidenzintervall