



Bachelorarbeit

Die deutsche Übersetzung des Guilford-Zimmermann Temperament Survey (GZTS)

Ein Beitrag zur Validierung der übersetzten Version des mehrdimensionalen, faktorenanalytisch konzipierten GZTS-Persönlichkeitstests

Viviane Rey-Bellet

Vertiefungsrichtung Entwicklungs- und Persönlichkeitspsychologie

Referentin: Janine Schmücker, Dipl.-Psych., MSc

Zürich, Mai 2012

Diese Arbeit wurde im Rahmen des Bachelorstudienganges am Departement P der Zürcher Hochschule für Angewandte Wissenschaften ZHAW verfasst. Eine Publikation bedarf der vorgängigen schriftlichen Bewilligung durch das Departement Angewandte Psychologie.

ZHAW Zürcher Hochschule für Angewandte Wissenschaften, Departement Angewandte Psychologie, Minervastrasse 30, Postfach, 8032 Zürich.

Danksagung

Ich bedanke mich herzlich bei sämtlichen Personen, die mich in irgendeiner Form bei dieser Arbeit unterstützt haben. Ein besonderes Dankeschön geht an Frau Janine Schmücker, die sich als Referentin zur Verfügung gestellt hat. Sie nahm sich für meine Fragen stets Zeit und lieferte mir wertvolle Hinweise.

Abstract

Die vorliegende Arbeit geht der Fragestellung nach, ob zwischen der auf Deutsch übersetzten Version des Guilford-Zimmermann Temperament Survey (GZTS) und dem NEO-Fünf-Faktoren-Inventar nach Costa und McCrae (NEO-FFI) Übereinstimmungsvalidität besteht. Zur Beantwortung werden die Skalenwerte und Gesamtscores beider Persönlichkeitstests von 43 Frauen und 26 Männern, die zwischen 23 und 72 Jahren alt sind, analysiert. Die Daten werden auf entsprechende Korrelationen geprüft. Neben der Untersuchung von sieben Hypothesen werden sechs weitere Beziehungen auch hinsichtlich eines Einflusses von Drittvariablen analysiert. Weiter werden demographische Merkmale bezüglich ihrer Zusammenhänge mit dem GZTS betrachtet und mit der Originalversion verglichen.

Insgesamt können 12 statistisch bedeutsame Zusammenhänge zwischen den Skalen des GZTS und des NEO-FFI nachgewiesen werden. Ebenso korrelieren die Gesamtscores signifikant miteinander. Dabei konnte kein bedeutender Einfluss durch Drittvariablen ermittelt werden. Vor diesem Hintergrund lässt sich zwischen den untersuchten Skalen und den Gesamtscores Übereinstimmungsvalidität nachweisen. Zwei Beziehungen präsentieren sich dabei entgegengesetzt der postulierten Richtung. Auch bei den ergänzenden Analysen zeigen sich überraschende Ausrichtungen der Korrelationen. Dies wird als Anhaltspunkt gewertet, dass die deutsche Skalenbezeichnung von einer erneuten Taxierung profitieren kann.

Kongruenzen bestehen zwischen dem deutschen GZTS und der Originalversion in Bezug auf Verbindungen zwischen einzelnen Skalen und dem Alter, respektive dem Bildungsabschluss. Die für die englische Ausführung berichtete Übereinstimmung gewisser Skalen mit dem Geschlecht können hingegen bei den vorliegenden Daten nicht nachgewiesen werden. Jedoch präsentieren sich andere, für die englische Ausführung nicht beschriebene Zusammenhänge.

Inhalt

1	Einleitung	1
1.1	Ausgangslage, Zielsetzung und Fragestellung.....	1
1.2	Aufbau und Inhalt der Arbeit	1
	I THEORETISCHER TEIL	2
2	Faktorenanalytisch begründete Persönlichkeitstheorien	2
2.1	Begriffsklärungen.....	3
2.2	Theoretischer Hintergrund faktorenanalytisch begründeter Persönlichkeitstheorien ..	5
2.3	Faktorenanalyse.....	7
2.4	Hierarchisches Modell	7
2.5	Guilfords Persönlichkeitstheorie	8
2.5.1	Kritik an Guilfords Modell.....	10
2.6	Das Fünf-Faktoren-Modell (FFM).....	11
2.6.1	Kritik am Fünf-Faktoren-Modell	12
3	Persönlichkeitsdiagnostik	13
3.1	Begriffsklärungen.....	13
3.2	Verfahren der Persönlichkeitsdiagnostik	14
3.3	Persönlichkeitstest.....	15
3.4	Antwortformate und Testauswertung.....	16
3.5	Gütekriterien psychologischer Tests	17
3.5.1	Objektivität.....	17
3.5.2	Reliabilität	17
3.5.3	Validität.....	18
3.5.4	Abhängigkeit der Hauptgütekriterien.....	20
4	Die Erhebungsinstrumente und daraus resultierende Hypothesen	20
4.1	Guilford-Zimmermann Temperament Survey (GZTS).....	21
4.2	NEO-Fünf-Faktoren-Inventar nach Costa und McCrae (NEO-FFI).....	23
4.3	Fazit , Fragestellung und Hypothesengenerierung.....	25
	II EMPIRISCHER TEIL	26
5	Forschungsgegenstand und Methodik	26
5.1	Untersuchungsdesign	26
5.2	Datenerhebung	27

5.3	Datenauswertung	27
6	Darstellung der Ergebnisse	28
6.1	Identifikation von Eingabefehlern	28
6.2	Beschreibung der Stichprobe	29
6.3	Beschreibung der Skalenwerte.....	30
6.4	Ergebnisse zur Überprüfung der Hypothesen.....	32
6.5	Überprüfung weiterer Zusammenhänge	39
7	Diskussion	45
7.1	Zusammenfassung	45
7.2	Interpretation der Ergebnisse	46
7.3	Kritik und Reflexion	55
7.4	Ausblick.....	56
Literatur	58
Anhang	62

Abbildungen

Abbildung 1: Hierarchisches Modell psychischer Merkmale	8
Abbildung 2: Die sieben Persönlichkeitsbereiche nach Guilford	9
Abbildung 3: Häufigkeiten Alterskategorien	29
Abbildung 4: Häufigkeiten Bildungsabschluss	30
Abbildung 5: Streudiagramm Extraversion – Soziabilität	33
Abbildung 6: Streudiagramm Extraversion-Generelle Aktivität	34
Abbildung 7: Streudiagramm Extraversion-Nachdenklichkeit	35
Abbildung 8: Streudiagramm Neurotizismus - Emotionale Stabilität	36
Abbildung 9: Streudiagramm Neurotizismus Gelassenheit	37
Abbildung 10: Streudiagramm Verträglichkeit-Vertrauensfähigkeit	37
Abbildung 11: Streudiagramm Gewissenhaftigkeit GZTS -Gewissenhaftigkeit NEO-FFI	38

Tabellen

Tabelle 1: Beschreibung der Big Five	11
Tabelle 2: Skalen der GZTS-Originalversion	21
Tabelle 3: Erfasste Merkmalsbereiche des NEO-FFI	24
Tabelle 4: Statistische Kennwerte Skalen GZTS & NEO-FFI	30
Tabelle 5: Korrelationswerte nach Pearson der Skalenwerte GZTS & NEO-FFI	32
Tabelle 6: Korrelation nach Kendall-Tau-b Hypothese 1	33
Tabelle 7: Übersicht Ergebnisse Hypothesenprüfung	39
Tabelle 8: Korrelationskoeffizienten nach Pearson Skalen GZTS -NEO-FFI	40
Tabelle 9: Korrelation nach Pearson Summenscore GZTS-NEO-FFI	41
Tabelle 10: Korrelationen Skalen GZTS -demographische Variablen	42
Tabelle 11: Partialkorrelationen GZTS-NEO-FFI	44

Abkürzungen

GZTS	Guilford-Zimmermann Temperament Survey
NEO-FFI	NEO-Fünf-Faktoren-Inventar nach Costa und McCrae
FFM	Fünf-Faktoren-Modell
ES	Emotionale Stabilität
S	Soziabilität
V/PB	Vertrauensfähigkeit/Persönliche Beziehungen
SV	Selbstvertrauen
Gel	Gelassenheit
GA	Generelle Aktivität
D	Durchsetzungsfähigkeit
N	Nachdenklichkeit
Gew	Gewissenhaftigkeit GZTS

1 Einleitung

Fragebögen erfreuen sich in der persönlichkeitsdiagnostischen Praxis grosser Beliebtheit. Dies äussert sich nicht zuletzt in der Existenz einer Fülle entsprechender Tests, deren Bestand kontinuierlich erweitert wird. Zentrales Anliegen der Psychologischen Diagnostik ist dabei, dass diese den qualitativen Anforderungen an ein Erhebungsinstrument entsprechen. Um zuverlässige und gültige Aussagen zuzulassen, müssen dafür wissenschaftliche Gütekriterien erfüllt werden. Dabei gilt die Validität als wichtigstes und gleichzeitig empirisch am schwierigsten überprüfbares Qualitätsmerkmal.

1.1 Ausgangslage, Zielsetzung und Fragestellung

Der dieser Arbeit zugrunde liegende Guilford-Zimmermann Temperament Survey (GZTS) ist im englischsprachigen Raum ein seit Jahren bekannter Persönlichkeitstest, der bisher nicht auf Deutsch existierte. Nun wurde er durch das Institut für angewandte Psychologie der ZHAW ins Deutsche übertragen und in einem folgenden Schritt inhaltlich und testtheoretisch überarbeitet.

Im weiteren Verlauf der Testentwicklung stellt sich nun die Frage der inhaltlichen Gültigkeit dieser überarbeiteten Version. Um Aussagen hinsichtlich der kriterienbezogenen Validität generieren zu können, verfolgt die vorliegende Arbeit das Ziel, die übersetzte Fassung des GZTS in einer deutschsprachigen Stichprobe anzuwenden und mit einem etablierten Verfahren zu vergleichen.

Hierfür verfolgt diese Arbeit die Fragestellung, ob zwischen der deutschen Version des GZTS und dem NEO-Fünf-Faktoren-Inventar nach Costa und McCrae (NEO-FFI) in der deutschen Ausführung von Borkenau und Ostendorf (2008) Übereinstimmungsvalidität besteht. Für die Beantwortung werden aufgrund inhaltlicher Darlegungen Hypothesen bezüglich Korrelationen zwischen Skalen beider Verfahren statistisch überprüft und durch die Diskussion weiterer relevanter Verbindungen ergänzt. Darüber hinaus werden Parallelen zwischen der deutschen Ausführung und der Originalversion erörtert.

1.2 Aufbau und Inhalt der Arbeit

Diese Arbeit besteht aus einem theoretischen und einem empirischen Teil. Im theoretischen Abschnitt wird im zweiten Kapitel auf faktorenanalytisch begründete Persönlichkeitstheorien eingegangen. Diese werden im persönlichkeitspsychologischen Umfeld ver-

ortet, ohne auf Bereiche ausserhalb des Eigenschaftsparadigmas ausführlich einzugehen. Das Modell von Guilford sowie das Fünf-Faktoren-Modell (FFM), auf denen die eingesetzten Erhebungsinstrumente beruhen, werden ausführlicher besprochen. Leistungsmerkmale und motivationale Kennzeichen werden nicht erörtert. Ebenso werden andere faktorenanalytische Erklärungen der Persönlichkeit nicht dargelegt.

Das dritte Kapitel bespricht Merkmale und Inhalte der Persönlichkeitsdiagnostik. Dabei werden besonders psychometrische Persönlichkeitstests und deren Hauptgütekriterien beachtet. Nebengütekriterien sowie alternative persönlichkeitsdiagnostische Verfahren werden über eine Auflistung hinaus nicht weiter vertieft. Neben der groben Darlegung der Gütekriterien werden keine weiteren testtheoretischen Aspekte besprochen. Fragen der Testkonstruktion, Normierung und Eichung können im Rahmen dieser Arbeit nicht ausführlich erörtert werden.

Im vierten Kapitel werden schliesslich als Abschluss des Theorieteiles aufgrund der inhaltlichen Vorstellung der beiden eingesetzten Erhebungsinstrumente die Hypothesen generiert und vorgestellt.

Der empirische Teil veranschaulicht im fünften Kapitel den Forschungsgegenstand und das methodische Vorgehen, bevor im sechsten Kapitel die Ergebnisse präsentiert werden. In der abschliessenden Diskussion folgt im siebten Kapitel eine Zusammenfassung der theoretischen Grundlagen und der empirischen Ergebnisse. Mittels Interpretation dieser Resultate wird die Fragestellung beantwortet, bevor die methodenkritische Reflexion und Anführung weiterführender Fragen den Abschluss der Arbeit bilden.

I THEORETISCHER TEIL

Im nun folgenden Abschnitt werden faktorenanalytisch begründete Gesamtsysteme der Persönlichkeit als theoretisches Fundament der untersuchten Erhebungsinstrumente sowie zentrale Aspekte der psychometrischen Persönlichkeitsdiagnostik dargelegt. Dieser Theorieteil wird mit der Hypothesengenerierung aufgrund der inhaltlichen Vorstellung der Erhebungsinstrumente abgeschlossen.

2 Faktorenanalytisch begründete Persönlichkeitstheorien

Persönlichkeitstheorien gehen davon aus, dass individuelle Differenzen im menschlichen Verhalten und Erleben bestehen und gemessen werden können. Dafür erforschen sowohl

die Differentielle Psychologie (die sich mit der Beschreibung und Analyse von interindividuellen Unterschieden zwischen Personen und/oder Gruppen befasst) als auch die Persönlichkeitspsychologie (die verhaltensrelevante individuelle Besonderheiten von Menschen innerhalb einer bestimmten Population untersucht) Phänomene menschlicher Erlebens- und Verhaltensweisen (Asendorpf, 2011, S. 7-15).

In diesem Kapitel wird nach einleitenden Begriffsklärungen spezifisch auf faktorenanalytische Persönlichkeitstheorien eingegangen. Dabei werden insbesondere die Konzepte von Guilford und das FFM erörtert. Diese wissenschaftlich begründeten Aussagen erklären individuelle Temperamentsunterschiede anhand von Eigenschaften. Dabei werden sie zu Typen zusammengefasst, die den einzelnen Menschen für ein bestimmtes Verhalten prädisponieren.

2.1 Begriffsklärungen

Persönlichkeit

Das Wort Person kommt vom lateinischen Begriff *persona* und bedeutet auf Deutsch Theatermaske, Rolle des Schauspielers (Duden, Herkunftswörterbuch, 2007, S. 599). In der psychologischen Fachliteratur wird der Begriff unterschiedlich definiert. Gewisse Erklärungen werden sehr eng gefasst und mit Temperament gleichgesetzt. Andere erweitern den Begriff und beziehen ergänzende Faktoren wie Fähigkeiten, Interesse und Einstellungen mit ein (Schmitt & Altstötter-Gleich, 2010, S. 8). Laut dem psychologischen Wörterbuch Dorsch ist „beim derzeitigen Stand der Forschung keine Einigkeit zu erzielen“ (Häcker & Stapf, 2009, S. 741). Jedoch identifiziert Herrmann (1976) eine mehrheitliche Übereinstimmung, diese als „ein bei jedem Menschen einzigartiges, relativ stabiles und zeitlich überdauerndes Verhaltenskorrelat“ (S. 25) zu verstehen.

Persönlichkeitstheorien

Nach Pervin, Cervone und John (2005) werden Persönlichkeitstheorien entwickelt, um spezifische Charakteristika psychologischer Qualitäten und deren Beziehung untereinander zu identifizieren. Ebenso interessieren Hinweise auf Determinanten einer Persönlichkeit wie beispielsweise ein genetischer Einfluss. Schliesslich wollen Fragen nach motivationalen Aspekten beantwortet werden, um Gründe für das individuelle Verhalten und Erleben liefern zu können (S. 32-33).

Nach Dorsch lässt sich aktuell das Spektrum der Persönlichkeitstheorien zum einen anhand ihrer angewendeten Methoden (philosophische, psychoanalytische, faktorenanalytische) unterscheiden. Zum andern wird in den darin als analytische Einheit vorherrschenden Konstrukten (Eigenschaften, Faktoren, Einstellungen, Selbst) ein weiteres Trennungsmerkmal gesehen. Darüber hinaus können aufgrund der dominierenden Gesichtspunkte der Persönlichkeitsbetrachtung strukturelle, dynamische, feld- und systemtheoretische Ansätze identifiziert werden (Häcker & Stapf, 2009, S. 742-743). Konstrukte werden dabei aus einem theoretischen Zusammenhang und mit Hilfe von beobachtbaren Ereignissen erschlossen (Häcker & Stapf, 2009, S. 536).

Individuum

Das lateinische Wort *individuum* bedeutet nach Dorsch „das Unteilbare, Einmalige in raumzeitlicher und qualitativer Hinsicht“ (Häcker & Stapf, 2009, S. 464). Schmitt und Altstötter-Gleich (2010, S. 7) orten in der Auffassung des Menschen als Einzelwesen die Einheit der Differentiellen Psychologie. Diese unterscheidet sich darin von Disziplinen wie beispielsweise die Sozialpsychologie, die ganze Gruppen in den theoretischen und zu untersuchenden Fokus rücken.

Temperament

Das lateinische Wort *temperamentum* bedeutet auf Deutsch „das rechte Mass, richtiges Verhältnis gemischter Dinge“ (Duden, Herkunftswörterbuch, 2007, S. 842) und bezeichnet nach Dorsch die „vorherrschende Art und die individuelle Eigenart des Ablaufs der seelischen Vorgänge“ (Häcker & Stapf, 2009, S. 995). Dabei bezieht es sich auf Unterschiede im Bereich der „drei A der Persönlichkeit: Affekt, Aktivierung und Aufmerksamkeit“ (Asendorpf, 2011, S. 28).

Eigenschaft/Wesenszug

Die Begriffe Eigenschaft und Wesenszug (engl. *trait*) werden in der entsprechenden Fachliteratur sinnverwandt verwendet. Dabei präsentiert sich die Beständigkeit über Zeit und Situation als gemeinsamer Nenner. Nach Stemmler, Hagemann, Amelang und Bartussek (2011) wird mit der Umschreibung einer Eigenschaft als „relativ breite, zeitlich stabile Disposition zu bestimmten Verhaltensweisen, die konsistent in verschiedenen Situationen auftreten“ (S. 51) der Ansicht einer Mehrheit empirisch arbeitender Persönlichkeitsforscher entsprochen.

Typus

Nach Dorsch wird unter einem Typus im Allgemeinen entweder eine „durch einen bestimmten Merkmalskomplex charakterisierte Gruppe“ (Häcker & Stapf, 2009, S. 1032) oder „ein Mensch, der alle Merkmale seiner Gruppe in besonders ausgeprägter Weise besitzt“ (Ebda, 2009) verstanden. Sehr verbreitet präsentiert sich dabei die Klassifikation von Personen in qualitativ unterschiedliche Beurteilungskategorien, die je nach Vorliegen bestimmter Eigenschaftskombinationen definiert werden. Bekannt wurde beispielsweise die Charakterisierung nach Galens, die ausgehend von der Dominanz der Körpersäfte vom melancholischen, phlegmatischen, sanguinischen und cholерischen Temperament spricht (Stemmler, Hagemann, Amelang & Bartussek, 2011, S. 63).

Disposition

Laut Dorsch ist mit dem Begriff Disposition die Möglichkeit gemeint, bestimmte seelische und/oder körperliche Inhalte (Erlebnisweisen, Reaktionsbereitschaften) auszuleben (Häcker & Stapf, 2009, S. 221). Der Ausdruck impliziert, dass hinter einer Handlungsweise ein Persönlichkeitsmerkmal (ein Kennzeichen, das ein Individuum von andern unterscheidbar macht) steht, das eine Person für seinen Verhaltensstil empfänglich macht, dafür „disponiert“ (Schmitt & Altstötter-Gleich, 2010, S. 9).

2.2 Theoretischer Hintergrund faktorenanalytisch begründeter Persönlichkeitstheorien

Die Persönlichkeitspsychologen Raymond Bernard Cattell (1905-1998), Joy Paul Guilford (1897-1988) und Hans-Jürgen Eysenck (1916-1997) waren die ersten, deren faktorenanalytisch begründete Gesamtsysteme der Persönlichkeit bekannt wurden (Asendorpf, 2007, S. 154). Sie verfolgten dabei das hochgesteckte Ziel, die wechselseitige Abhängigkeit der verschiedenen, die Persönlichkeit konstituierenden Dimensionen aufzuzeigen. Kennzeichnend ist, dass ihre Theorien am Anfang des Entwicklungsprozesses nicht bestimmend waren, sondern vielmehr ein Ergebnis dieses Ablaufes darstellen (Amelang, Bartussek, Stemmler & Hagemann, 2006, S. 250).

Dabei wurden in den 1950-er Jahren mittels statistischer Verfahren besonders im anglo-amerikanischen Raum Typologien konzipiert. Dank nun verfügbaren Computern konnten diese Berechnungen deutlich effizienter durchgeführt werden als bis anhin. Diese Konzeptionen sind vorwiegend auf die Betrachtung einzelner Individuen ausgerichtet und

möchten der Vielfalt menschlichen Denkens, Fühlens und Verhaltens Leitbegriffe zuzuordnen, um die zugrunde liegenden „Verhaltensradikale“ zu bestimmen (Rauchfleisch, 2008, S. 17).

Das Persönlichkeitsmodell von Guilford und das FFM basieren auf dem *Eigenschaftsparadigma*, das auf der Alltagspsychologie, einem System kulturell übermittelter Vorstellungen über menschliches Verhalten und dessen Ursachen, basiert. Asendorpf (2011) definiert indes ein Paradigma als „einigermaßen zusammenhängendes, von vielen Wissenschaftlern geteiltes Bündel aus theoretischen Leitsätzen, Fragestellungen und Methoden“ (S. 13). Um Personen zu beschreiben, wird dabei auf Eigenschaftsbegriffe zurückgegriffen. Diese erweisen sich als ökonomisch, indem sie zusammengehörige Verhaltensweisen bündeln und durch einen Begriff ersetzen. Somit kann eine Person mit einem Eigenschaftsbegriff eventuell genauso treffend beschrieben werden wie mit einer langen Aufzählung verschiedener Reaktionsweisen in bestimmten Situationen.

Nach Gordon W. Allport (1897-1967), dem Pionier des eigenschaftstheoretischen Ansatzes, steuern Wesenszüge unsere Wahrnehmung der Aussenwelt. Sie tragen dazu bei, dass sich Menschen über verschiedene Situationen und die Zeit hinweg stabil verhalten (Schmitt & Altstötter-Gleich, 2010, S. 79-80). Die Annahme, dass Individuen eine allgemeine Prädisposition für eine spezifische Reaktionsweise besitzen, ist nach Pervin, Cervone und John (2005) eine von allen Eigenschaftstheorien geteilte Grundannahme. Mit dieser Vorstellung direkt verbunden ist die Meinung, dass offene Verhaltensweisen mit den dahinter liegenden Wesenszügen verknüpft sind. Es wird also nicht wie beispielsweise in der psychoanalytisch begründeten Persönlichkeitstheorie von einem Konzept des Unbewussten ausgegangen (S. 284).

Konkret beobachtbare Verhaltensweisen werden vielmehr in Kategorien zusammengefasst, welche nach Häufigkeits- und Intensitätsmerkmalen in ihrer Ausprägtheit bestimmt (also skaliert) werden können. Dabei kann jede Person eine spezifische Position auf einem gedachten, und zwischen den Polen extremer und minimaler Ausprägtheit konzipierten, Kontinuum einnehmen. Dies ermöglicht interindividuelle Vergleiche. Stemmler, Hagemann, Amelang und Bartussek (2011) sprechen dabei von Dispositionsprädikaten (Handlungsbereitschaften), die Vorhersagen über Situationen zulassen, für die bisher keine Möglichkeit der Beobachtung bestanden hatte. Neben dem ökonomischen Aspekt stellt dies einen wichtigen Punkt des Eigenschaftsparadigmas dar (S. 51-52). Die-

se Verhaltensdispositionen beziehen sich zu einem grossen Teil auf die Form, wie sich jemand verhält, also Eigenschaften des Temperaments (Asendorpf, 2007, S. 178).

2.3 Faktorenanalyse

Sowohl Guilfords Modell als auch das FFM erklären die Persönlichkeit anhand faktorenanalytisch begründeter Überlegungen. Diese Bezeichnung steht für einen Sammelbegriff verschiedener, komplexer mathematischer Verfahren, um Korrelationsmuster von Variablen analysieren zu können (Laux, 2008, S. 36). Korrelationen sind dabei Ausdruck der Enge eines Zusammenhanges. Es wird also eruiert, ob, beziehungsweise wie ein Phänomen (z.B. Zustimmung zu Aussage A) mit einem anderen (z.B. Ablehnung von Aussage B) in Verbindung steht. Damit werden Aussagen über die Art und Intensität des gemeinsamen Variierens generiert, die jedoch keine kausalen Hinweise liefern. Bei der Prüfung dieser Zusammenhangsmuster geht es darum, Gruppen identifizieren zu können, deren Merkmale enger miteinander verknüpft sind, als mit Grössen ausserhalb der jeweiligen Gruppe. Gemäss ihren korrelativen Beziehungen werden also die entsprechenden Merkmale in voneinander unabhängige Gruppen eingeteilt (Bortz, 2005, S. 512). Diese Gruppen werden *Faktor* genannt, wobei dieser Begriff für die Annahme einer gemeinsamen Ursache steht (Schmitt & Altstötter-Gleich, 2010, S. 78). Jede Variable zeichnet ihre spezifische *Faktorenladung* aus, die das Ausmass der Ähnlichkeit der Variable mit dem Faktor ausdrückt und zwischen +1 (volle Ähnlichkeit) und -1 (keine Ähnlichkeit) variieren kann (Asendorpf, 2007, S. 152).

2.4 Hierarchisches Modell

In beiden Erklärungsansätzen menschlicher Erlebens- und Verhaltensweisen ist ein hierarchisches Modell der Persönlichkeit bekannt. Diese haben sich empirisch bewährt, um die benötigte Verknüpfung zwischen Persönlichkeitseigenschaften und Handlungsweisen herzustellen (Schmitt & Altstötter-Gleich, 2010, S. 73). Die Ebenen der Hierarchie unterscheiden sich dabei anhand des Abstraktionsniveaus der darin enthaltenen Eigenschaft. Je weiter entfernt ein Wesenszug von konkreten Verhaltensweisen ist (also je abstrakter), desto höher wird sie im hierarchischen System eingeordnet. Abbildung 1 stellt den Aufbau eines hierarchischen Modells mit den zugehörigen Ebenen dar. Wenn auch die Ebenen in den verschiedenen existierenden Modellen unterschiedlich benannt werden, können darin nach Schmitt und Altstötter-Gleich (2010) allgemein akzeptierte Grundannah-

men deutlich gemacht werden. Zum einen wird das Verhalten einer Person in einer spezifischen Situation massgeblich durch ihre Persönlichkeitsmerkmale beeinflusst. Dieser Umstand lässt zum andern auch den Umkehrschluss zu, dass vom (konkreten) Verhalten auf dahinter liegende (abstraktere) Eigenschaften geschlossen werden kann. Je höher die Eigenschaft im Modell eingeordnet, desto konsistenter und stabiler wird das entsprechende Verhalten gezeigt und lässt sich somit auch zuverlässiger vorhersagen (S. 73-75).

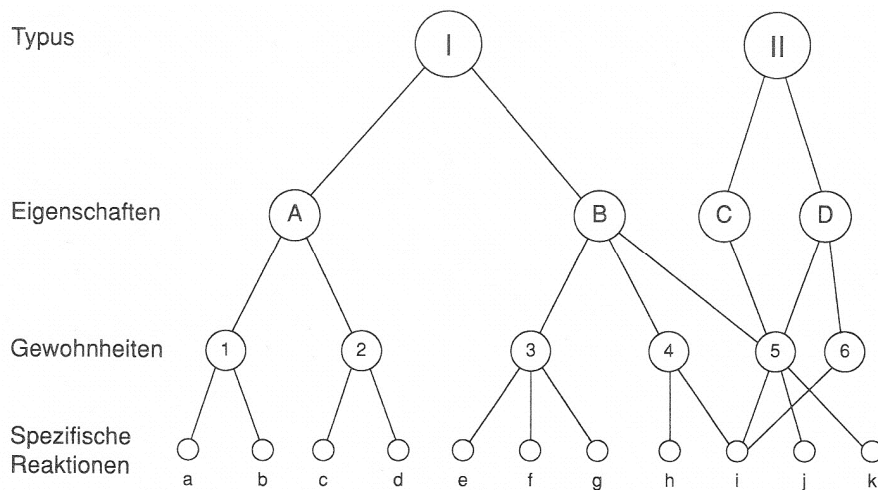


Abbildung 1: Hierarchisches Modell psychischer Merkmale (aus Netter, 2005, S. 239)

2.5 Guilfords Persönlichkeitstheorie

Guilford (1964) definiert Persönlichkeit als eine „einzigartige Struktur von Wesenszügen“ (S. 6). Diese Charakteristika sieht er in einer weiten und allgemeinen Auslegung als Ergebnis von Abstraktionen (der beobachteten Wirklichkeit). Dabei lassen unter einem hierarchischen Standpunkt erst viele Beobachtungen zusammengefasst den Schluss auf einen dahinter liegenden Wesenszug zulassen (S. 48). Diese Eigenschaftszüge teilt er in einer über den Temperamentsbereich hinaus gehenden Auffassung von Persönlichkeit in sieben Bereiche (Einstellungen, Interessen, Bedürfnisse, physiologische und morphologische Merkmale, Eignung und Temperament) ein (siehe Abbildung 2).

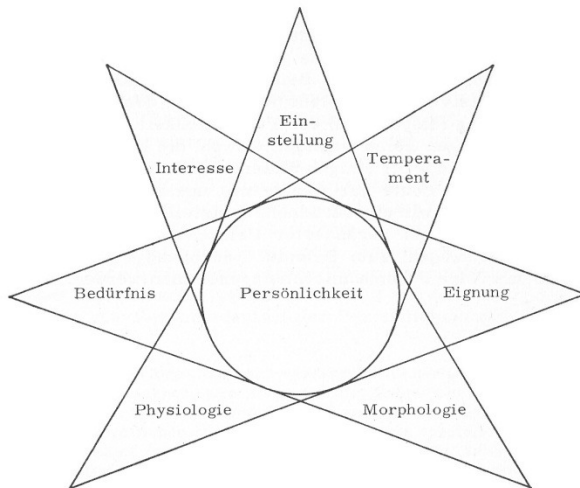


Abbildung 2: Die sieben Persönlichkeitsbereiche nach Guilford (aus: Guilford, 1964, S. 9)

Im *deskriptiven Modell* werden diese Segmente der Persönlichkeit nach Fisseni (1998, S. 324-325) in vier Klassen unterteilt:

- Somatisch: bezieht sich auf morphologische Merkmale wie Körpergröße, Haarfarbe, sowie auf physiologische Charakteristika wie beispielsweise die Körpertemperatur
- Motivational: beinhaltet Bedürfnisse, die sich auf grundlegende Wünsche beziehen wie das Bedürfnis nach Anerkennung, Interessen und Einstellungen im Sinne einer konstanten Haltung bestimmten Sachverhalten gegenüber
- Temperament: lässt Aussagen zu, wie jemand etwas tut
- Eignung: greift darauf zurück, wie gut jemand etwas kann als Voraussetzung für Tätigkeiten

Guilford kann als erster Persönlichkeitspsychologe angesehen werden, der faktorenanalytische Methoden systematisch angewendet hat, um bedeutende Schlussfolgerungen bezüglich unterschiedlicher Persönlichkeitsstrukturen zu generieren (Rushton & Irwing, 2009, S. 560). Dabei entwickelte er das *faktorenanalytische Modell* aufgrund von Antworten aus fünf kurzen Persönlichkeitsfragebögen, die Gewohnheiten und Vorlieben im interpersonalen Bereich abbilden. Diese betrachtet er als Indikatoren von Persönlichkeitsdimensionen. Von denen ausgehend identifizierte er mittels statistischer Berechnungen die psychologisch wichtige und praktisch nützliche Anzahl Dimensionen zur Persönlichkeitsbeschreibung (Guilford J. P., 1964, S. 87-88).

Er ging in diesem Zusammenhang von den folgenden 13 Faktoren des Temperamentes aus, auf denen der im Kapitel 4.1 näher vorgestellte GZTS basiert (Bartussek, 1996, S. 55-56):

- Soziale Introversion (seclusiveness)
- Gedankliche Introversion (reflectiveness)
- Depression (depression)
- Zyklische Disposition (cycloid tendency)
- Leichtlebigkeit (rhythymia, happy go lucky)
- Generelle Aktivität (general activity, general pressure for overt activity)
- Durchsetzungsfähigkeit (ascendence, social boldness)
- Männliche Gefühle und Interessen (masculinity)
- Minderwertigkeitsgefühle (inferiority feelings)
- Nervosität (nervousness)
- Sachlichkeit (objectivity)
- Freundlichkeit (agreeableness)
- Aufgeschlossenheit (cooperativeness)

Ursprünglich wollte er sein Modell nicht hierarchisch organisieren. Aufgrund zahlreicher Untersuchungen kommt er jedoch zum Schluss, dass seine 13 Faktoren nicht unkorreliert sind und schlägt ein *hierarchisches Modell* zur Strukturierung des Temperamentsbereiches vor (Guilford, Zimmermann & Guilford, 1976, S. 71). Dabei ergeben sich aus den 13 Primärfaktoren vier Dimensionen höherer Ordnung (soziale Aktivität, Intro-/ Extraversion, emotionale Stabilität und paranoide Disposition). Darüber lässt sich ein Faktor dritter Ordnung (emotionale Gesundheit) erkennen (Guilford et al., 1976, S. 30).

2.5.1 Kritik an Guilfords Modell

Amelang, Bartussek, Stemmler und Hagemann (2006) nennen eine gewisse Unbestimmtheit als Schwäche des Modells. Dabei erwähnen sie die Schwierigkeit, die Ebenen der Temperamentsfaktoren in seinem hierarchischen Modell bestimmen zu können und orten die Ursache in nicht ganz eindeutigen Definitionen (S. 252). In einer entsprechenden Untersuchung konnte Guilfords hierarchisches Modell nicht bestätigt werden (Amelang & Borkenau, 1982, S. 139).

2.6 Das Fünf-Faktoren-Modell (FFM)

Seit einigen Jahren präsentiert sich das FFM als äusserst einflussreiches Eigenschaftsmodell, das in unzähligen Studien empirisch untersucht wurde und sich auch im deutschsprachigen Raum gut bewährt hat. Es wird als Referenzmodell aufgefasst, mit dessen Hilfe Forschungsergebnisse verschiedener Traditionen vergleichbar gemacht werden sollen (Stemmler, Hagemann, Amelang & Bartussek, 2011, S. 275).

Das FFM wurde von mehreren unabhängigen Forschern im angloamerikanischen Raum entwickelt und basiert auf der lexikalischen Hypothese. Dieser liegt die Annahme zugrunde, dass sich die für die tägliche Kommunikation besonders nützlichen Merkmale oder Begriffe in der Alltagssprache niedergeschlagen haben (Laux, 2008, S. 174). Die Analyse der Sprache (lexikalischer Ansatz) soll nun die wichtigsten Persönlichkeitsmerkmale der Gesamtpersönlichkeit offenbaren (Stemmler, Hagemann, Amelang & Bartussek, 2011, S. 267). In diesem Prozess wird das ganze Lexikon einer Sprache in einem mehrstufigen Reduktionsverfahren auf einen überschaubaren Satz von Eigenschaftsbezeichnung verringert. Dafür werden bei einer grossen Anzahl Personen Fremd- und Selbstbeurteilungsbefragungen durchgeführt, die mittels Faktorenanalyse zu wenigen, möglichst unabhängigen (statistisch spricht man von orthogonalen) Faktoren zusammengefasst werden (Asendorpf, 2007, S. 154). Entsprechende Untersuchungen förderten dabei unabhängig voneinander immer wieder fünf breite und robuste Faktoren zutage. Diese Hauptfaktoren werden auch als *Big Five* bezeichnet und beschreiben weite Bereiche der Persönlichkeit auf effiziente Weise, da sie unkorreliert sind. Ihnen werden die in Tabelle 1 dargestellten Begriffe zugeordnet.

Tabelle 1: Beschreibung der Big Five (nach Asendorpf, 2011, S. 54)

Faktorenbezeichnung	Zentrale Begriffe
Neurotizismus	Nervosität, Ängstlichkeit, Gefühlsschwankungen
Extraversion	Geselligkeit, Ungehemmtheit, Aktivität
Gewissenhaftigkeit	Ordentlichkeit, Beharrlichkeit, Zuverlässigkeit
Verträglichkeit	Freundlichkeit, Hilfsbereitschaft, Wärme im Umgang mit anderen Personen
Offenheit gegenüber neuen Erfahrungen	Intellektuelle Neugier, Gefühl für Kunst und Kreativität, (korreliert positiv mit Intelligenz und vor allem Bildung)

Diese Auffassung sollte jedoch nicht insofern missverstanden werden, dass nur die fünf Faktoren bedeutsame Persönlichkeitsdimensionen sind. Kombinationen verschiedener Grössen, also eine Interpretation, dass ein Merkmal sowohl Anteile des einen als auch eines anderen Faktors integriert, sind gegebenenfalls ebenso geeignet, um Persönlichkeitsmerkmale zu beschreiben. Asendorpf (2007) nennt dabei das Beispiel der Schüchternheit, die als Verknüpfung der Dimensionen Neurotizismus und Extraversion quasi als „Resultat“ der Differenz Neurotizismus minus Extraversion aufgefasst werden kann (S. 156-157).

Bis heute gibt es keine Übereinstimmung in der Frage nach der hierarchischen Ordnung der Big Five. Costa und McCrae (1992; zit. nach Stemmler, Hagemann, Amelang & Bartussek, 2011, S. 276), zwei amerikanische Forscher, die seit Anfang der 1980-er Jahre an der Entwicklung und Erforschung des FFM beteiligt sind, legten ein hierarchisches Modell mit fünf Domänenfaktoren (Big Five) vor, die in je sechs sogenannte Facettenfaktoren unterteilt sind. Das darauf basierende NEO-Persönlichkeitsinventar, revidierte Fassung (mit der Kurzform NEO-PI-R bezeichnet) weist diese Struktur auf. Anhand der Beantwortung seiner 240 Fragen wird eine differenzierte Persönlichkeitsbeschreibung ermöglicht. Sparsamer lassen sich die Big Five mittels verschiedener existierender Kurzversionen entsprechender Persönlichkeitsfragebögen erfassen. Dabei ist jedoch zu beachten, dass diese eine geringere inhaltliche Breite erfassen (Asendorpf, 2011, S. 54-55).

2.6.1 Kritik am Fünf-Faktoren-Modell

Verschiedene forschende Personen äusserten Zweifel, ob gerade fünf Faktoren die Persönlichkeit umfassend und ökonomisch beschreiben. Bei der Konzipierung der Big Five wurden einige Bereiche wie beispielsweise Werthaltungen oder gesundheitsbezogene Eigenschaften von vornherein ausgeschlossen. Wenn diese mit berücksichtigt werden, ergeben sich mehr als die hierbei postulierten fünf Faktoren (Laux, 2008, S. 179). So wurden verschiedene Ideen publiziert, die weitere Faktoren oder die Aufspaltung eines Big Five-Faktors in zwei Domänen postulieren. Andresen (2000, S. 13) nennt einen sechsten Faktor der Risikopräferenz, Becker (1999, S. 522) schlägt einen Faktor Hedonismus/Spontanität vor. Ashton et al. (2004) führen ebenfalls sechs Faktoren an, die aufgrund lexikalischer Analysen in sieben Sprachen gefunden werden konnten (2004, S. 363).

Ebenso wird aber auch Überlegungen Aufmerksamkeit geschenkt, die sich auf weniger als fünf Faktoren berufen. Asendorpf (2007) erwähnt in diesem Zusammenhang die „Big Three“ (S. 155), welche die Big Five Faktoren Extraversion, Verträglichkeit und Gewissenhaftigkeit etwas breiter fassen und somit im Kulturvergleich besser replizierbar sind. Rushton und Irwing (2009) kommen in ihrer Metaanalyse zum Schluss, „robust evidence for the existence of a general factor of personality“ (S. 564) gefunden zu haben.

Die Frage nach der adäquaten Anzahl Faktoren lässt sich dabei nicht allein aufgrund faktorenanalytischer Berechnungen beantworten. Vielmehr verlangt sie auch eine theoretische Grundlegung, die sich beispielsweise aus der Nützlichkeitsabwägung für diagnostische Fragestellungen ergeben kann (Stemmler, Hagemann, Amelang & Bartussek, 2011, S. 275).

3 Persönlichkeitsdiagnostik

In diesem Kapitel wird nach einleitenden Erläuterungen zentraler Begriffe auf psychometrische Persönlichkeitstests im Speziellen eingegangen. Insbesondere werden dabei die Gütekriterien psychometrischer Verfahren erläutert.

3.1 Begriffsklärungen

Psychometrie

Nach Dorsch wird unter Psychometrie die Messung psychologischer Erscheinung ganz allgemein verstanden, nachdem es früher die Untersuchung der zeitlichen Verhältnisse in den seelischen Vorgängen bezeichnete (Häcker & Stapf, 2009, S. 800).

Diagnostik

Der Begriff Diagnose stammt vom griechischen Verb *dia-gi-nóskein* zu Deutsch „durch und durch erkennen“ (Duden, Herkunftswörterbuch, 2007, S. 144) ab. Zimbardo und Gericig (2008) bezeichnen Diagnostik als den „Einsatz festgelegter Testverfahren zur Bewertung der Fähigkeiten, Verhaltensweisen und Persönlichkeitseigenschaften von Personen“ (S. 326). Dabei ist die Psychologische Diagnostik eng verwandt mit der Differentiellen Psychologie, was sich in der gemeinsamen Prämisse messbarer interindividueller Unterschiede widerspiegelt (siehe Kapitel 2). Sie befasst sich mit der Theorie, der Konstruktion und der Analyse von Diagnostikverfahren, die in der Forschung genauso zur Anwen-

dung kommen wie beispielsweise beim Erstellen eines allgemeinen Persönlichkeitsbildes, bei differenzialdiagnostischen Fragen, zur Bestimmung der Leistungsfähigkeit sowie der beruflichen Eignung (Rauchfleisch, 2008).

Test

Pospeschill (2010) definiert einen psychologischen Test als „ein wissenschaftliches Routineinstrument zur Erhebung oder Messung distinkter psychologischer Merkmale auf der Basis von einzelnen Messwerten, Skalenwerten oder kategorialen Werten mit dem Ziel, den Ausprägungsgrad individueller Merkmale zu bestimmen“ (S. 16).

Persönlichkeitsfragebogen

Darunter wird ein standardisiertes Instrument verstanden, mit dem selbst- oder fremdberichtete Daten erfasst werden. Daraus werden Testwerte abgeleitet, die als Indikatoren für den individuellen Ausprägungsgrad von Persönlichkeitsmerkmalen dienen (Kubinger & Jäger, 2003, S. 332). Der im englischen Sprachgebrauch verwendete Begriff des *questionnaires* unterscheidet sich hinsichtlich seiner Qualitätsanforderungen kaum von einem psychologischen Test. Im Deutschen kann er jedoch als Sammelbegriff aufgefasst werden für vielfältige Formen schriftlicher Befragungen in verschiedenen inhaltlichen Bereichen (Jankisz & Moosbrugger, 2007, S. 2).

Verfahren

Als psychologisch-diagnostische Verfahren werden diejenigen Methoden bezeichnet, die zur Diagnose von Verhalten, Erleben oder Personenmerkmalen geeignet sind und den Gütekriterien der Psychologischen Diagnostik genügen (Kubinger & Jäger, 2003, S. 433).

3.2 Verfahren der Persönlichkeitsdiagnostik

Amelang und Schmidt-Atzert (2006) nennen folgende Gruppen von persönlichkeitsdiagnostischen Verfahren:

- Psychometrische Selbstbeschreibungsverfahren (Persönlichkeitsinventare und Fragebogen zur Erfassung von Einstellungen, Neigungen oder Interessen)
- Nichtsprachliche Persönlichkeitstests
- Objektive Persönlichkeitstests
- Projektive Verfahren

- Verhaltensbeobachtung
- Diagnostische Interviews

3.3 Persönlichkeitstest

Nach Dorsch werden *metrische* (prüfende) und *projektive* (entfaltende) Persönlichkeitstests unterschieden. Erstere dienen der „objektiven“ Erfassung der Leistungsfähigkeit anhand einzelner Funktionen oder Eigenschaften. Letztere sind auf Struktur und Charakter der Gesamtpersönlichkeit ausgerichtet und fordern die freie Entfaltung beziehungsweise Stellungnahme der Testperson (Häcker & Stapf, 2009, S. 260).

Das Brickenkamp-Handbuch unterteilt psychometrische Persönlichkeitstests in *Persönlichkeits-Struktur-Tests*, *Einstellungs-Interessentests* und *Klinische Tests*. Erstgenannte Kategorie meint eine Gruppe mehrdimensionaler Persönlichkeitstests, deren gemeinsames Kennzeichen die Messung verschiedener psychologischer Merkmale darstellt. Der zweite Begriff basiert auf sozialpsychologischen Konzepten und misst die Einstellung zu Sachverhalten oder die persönliche Meinung. Die letztgenannte Einheit wird schliesslich zu differentialdiagnostischen Zwecken im psychopathologischen Bereich eingesetzt und dient als Entscheidungshilfe bei der Auswahl von Interventionsmassnahmen und verschiedenen Therapieformen (Petermann, Holling, Leutner & Brähler, 2002, S. XIV-XV).

Nach Borkenau, Friedel und Wolfradt (2011) basieren die in der aktuellen Persönlichkeitsdiagnostik sehr verbreiteten standardisierten Fragebogenverfahren (psychometrische Persönlichkeitstests) auf der Annahme, aus der subjektiven Einschätzung liessen sich Rückschlüsse auf die Persönlichkeit der Testpersonen ziehen. In empirischen Untersuchungen liess sich auch ein Zusammenhang zwischen Selbst- und Fremdeinschätzungen der Persönlichkeitsskalen oder anderen Aussenkriterien (beispielsweise dem Berufserfolg) bestätigen. Persönlichkeitsfragebögen erfassen dabei anhand von zu beantwortenden Fragen (sogenannten Items) das typische Verhalten der untersuchten Person. Sie können universell eingesetzt und ökonomisch durchgeführt werden. Ebenso zeichnet sie, dank Standardisierung, eine sehr hohe Durchführungs- und Auswertungsobjektivität aus (S. 1-7).

Rauchfleisch (2008) warnt jedoch auch davor, psychologische Tests unkritisch einzusetzen und übersteigerte Erwartungen an die Resultate zu stellen. Er postuliert vielmehr, diese sorgfältig durchzuführen und zu interpretieren, liefern sie doch bei gezieltem Ein-

satz und auf kritischem, fundiertem Fachwissen basierend wichtige Informationen (S. 40-41). Eine zentrale Voraussetzung für aussagekräftige Resultate ortet er bei Selbstbeschreibungsv erfahren in der Introspektionsfähigkeit der getesteten Person. In diesem Zusammenhang betont er den subjektiven Charakter von Persönlichkeitstests (S. 170-171).

3.4 Antwortformate und Testauswertung

Psychometrische Persönlichkeitstests weisen unter anderem eine hohe Standardisierung auf, weil sie über gebundene (vorgegebene) Antwortformate verfügen. Testpersonen stimmen in diesem Fall einer präsentierten Aussage zu oder lehnen sie ab, Diese graduellen Beurteilungen werden schliesslich als Indikator für die Ausprägung des untersuchten Persönlichkeitsmerkmals herangezogen (Jankisz & Moosbrugger, 2007, S. 50).

Bei der Testkonstruktion können zum einen *dichotome* Antwortmöglichkeiten (mit zwei Alternativen wie z.B. trifft zu/trifft nicht zu) berücksichtigt werden. Diese generieren zwar eine weniger differenzierte Aussage, zeichnen sich dafür durch eine kurze Bearbeitungszeit und ökonomische Auswertung aus (Pospeschill, 2010, S. 50). Alternativ wird eine Testperson bei einer *Mehrfachauswahl* (Multiple-Choice) dazu aufgefordert, das Ausmass an Zustimmung oder Ablehnung graduell zu bestimmen. Solche Ratingskalen können eine unterschiedliche Anzahl Abstufungen aufweisen, wobei sich nach Pospeschill 3-7 Skaleneinheiten bewährt haben (2010, S. 51).

Bipolare Skalen weisen einen Abstufungsbereich von völliger/starker Ablehnung über einen neutralen Mittelbereich zur völligen/starken Zustimmung auf. Jankisz und Moosbrugger (2007) sehen in dieser Neutralzone auch eine Gefahr, dass sie nicht nur als mittlerer Ausprägungsbereich verstanden wird, sondern auch als Ausweichoption dient. Dies kann geschehen, wenn die Testperson den Wortlaut als unpassend empfindet, die Frage nicht versteht oder die Antwort verweigert. In diesem Fall kann das interessierende Phänomen mit einem konstrukt fremden Antwortverhalten vermischt werden, was die Interpretation verzerren kann. Dieser Umstand kann durch die Aufnahme einer Weiss-nicht-Antwortkategorie entschärft werden (S. 53-54). Verfügt die Skala über eine gerade Anzahl Stufen, entfällt die oft als normale oder typische Ausprägung interpretierte Ausweichkategorie. In diesem Fall wird von *forciertem Rating* gesprochen. Pospeschill (2010) postuliert, die adäquate Ratingform unter Einbezug des interessierenden Konstruktes und des antizipierten Antwortverhaltens sorgfältig auszuwählen, da eine unangemessene Einstufungsart Einschränkungen der Validität nach sich ziehen kann (S. 53).

Für die Auswertung werden die Antworten anhand ihrer jeweiligen Punktwerte zu einem Gesamtwert (Score) verrechnet. Bei mehrdimensionalen Persönlichkeitstests ergibt sich auf diese Weise für jede Dimension ein Skalenwert, der unter Einbezug der Normen eine Einschätzung über die quantitative Ausprägung dieses Merkmals bei der getesteten Person zulässt.

3.5 Gütekriterien psychologischer Tests

Gütekriterien stellen ein Instrument der Qualitätsbeurteilung psychologischer Tests dar, über welche im Manual (Handbuch) eines Tests Auskunft gegeben werden sollte. Die Erfüllung dieser Anforderungen wird empirisch überprüft und dargelegt. Damit unterscheidet sich ein wissenschaftliches Erhebungsinstrument beispielsweise von einer als unwissenschaftlich klassifizierten Fragensammlung, wie sie unter anderem in Zeitschriften anzutreffen sind (Pospeschill, 2010, S. 16).

Dabei hat sich zurückgehend auf Lienert und Raatz (1998, S. 7) in der wissenschaftlichen Literatur eine Unterscheidung zwischen Haupt- und Nebengütekriterien etabliert. Zu den Hauptgütekriterien werden die Objektivität, die Reliabilität und die Validität gezählt, die im Folgenden ausführlicher erörtert werden. Als Nebengütekriterium nennen Kubinger und Jäger (2003) die Skalierung, Normierung/Eichung, Ökonomie, Nützlichkeit, Zumutbarkeit, Unverfälschbarkeit und Fairness (S. 197-198).

3.5.1 Objektivität

Pospeschill (2010) sieht das Gütekriterium der Objektivität dann als gegeben, „wenn das Testergebnis unabhängig vom Untersuchungsleiter, vom Testauswerter und der Ergebnisinterpretation ist“ (S. 18). Ein psychologischer Test gilt somit als objektiv, wenn er in Bezug auf die Handhabung, Auswertung und Interpretation situationsunabhängig ist. Das bedeutet, dass unterschiedliche Testauswerter und Auswerterinnen bei den gleichen Testpersonen stets zu identischen Ergebnissen kommen (S. 20).

3.5.2 Reliabilität

Lienert und Raatz (1998) definieren das Gütekriterium der Reliabilität eines Tests als „Grad der Genauigkeit, mit dem er ein bestimmtes Persönlichkeits- oder Verhaltensmerkmal misst, gleichgültig, ob er dieses Merkmal auch zu messen beansprucht...“ (S. 9). Dieses Qualitätsmerkmal befasst sich also nicht mit inhaltlichen Aspekten. Es fragt viel-

mehr, ob, respektive wie viele, Messfehler vorhanden sind. Erfasst wird dieses Gütekriterium als Retest-Reliabilitätskoeffizient r_{tt} . Zu dessen Ermittlung wird derselbe Test nach einem bestimmten Intervall wiederholt und die Testergebnisse werden miteinander korreliert. Theoretisch kann der Reliabilitätskoeffizient zwischen 0 und 1 schwanken. Ist $r_{tt} = 0$, so ergibt sich bei einer Testwiederholung derselben Person ein komplett anderes Ergebnis, weil das Resultat nur aus Messfehlern besteht. Bei $r_{tt} = 1$ ergeben sich zwei absolut identische Bilder, da das Testergebnis vollkommen ohne Fehler entsteht. In der psychologischen Diagnostikpraxis treten jedoch beide Fälle nicht auf, sodass als grobe Faustformel Reliabilitätskoeffizienten ab $r_{tt} = 0.7$ als akzeptabel und ab 0.8 bis 0.9 als gut gelten (Pospeschill, 2010, S. 21). Niedrige Reliabilitätswerte lassen also nur unsichere diagnostische Schlüsse zu.

Zeitlich nahe beieinander liegende Testwiederholungen sind jedoch bei psychologischen Tests wegen Erinnerungs- und Übungseffekten eventuell nicht aussagekräftig. Aus diesem Grund hat die von Gullikson (1950; zit. nach Moosbrugger, 2007, S. 100) begründete Klassische Testtheorie nach methodischen Auswegen gesucht, die Messgenauigkeit eines psychologischen Tests dennoch bestimmen zu können. Diese präsentiert sich im Sinne eines theoretischen Hintergrundes bis heute als eine wichtige Basis aktueller psychodiagnostischer Tests. Die *Paralleltest-Reliabilität* (bei dem gleichzeitig ein äquivalenter Test durchgeführt und mit den ersten korreliert wird), die *Splithalf-Reliabilität* (bei dem Teile eines Tests als parallele Testversion aufgefasst und miteinander korreliert werden) sowie die *innere Konsistenz* (dabei wird jedes einzelne Item als eigenständiger Testteil aufgefasst und miteinander verglichen) werden in diesem Zusammenhang als mögliche Vorgehensweisen aufgeführt (Pospeschill, 2010, S. 22-24). Letztere wird in der Regel mit Cronbachs Alpha (α) angegeben und sollte sich zwischen 0.8 und 0.9 bewegen (Borkenau, Friedel & Wolfradt, 2011, S. 3).

3.5.3 Validität

Unter Validität wird der Grad der Genauigkeit verstanden, mit der ein Test das Merkmal misst, das er vorgibt zu messen (Lienert & Raatz, 1998, S. 10). Im deutschen Sprachraum wird synonym auch der Begriff Gültigkeit verwendet. Die Validität gilt als das wichtigste, aber auch am schwierigsten zu überprüfende Gütekriterium. Dabei geht es darum zu evaluieren, ob inhaltlich das Persönlichkeitsmerkmal erfasst wird, das der Test zu messen beabsichtigt. Die Herausforderung liegt dabei darin, empirisch fassbare Kriterien zu fin-

den, mit denen das Testresultat in Verbindung gesetzt werden kann (Rauchfleisch, 2008, S. 70).

Nach Pospeschill (2010) bezieht sich die Validität also auf die Gültigkeit verschiedener Interpretationen der Testergebnisse. Dabei wird eine Inhalts-, Konstrukt-, Augenschein- und Kriteriumsvalidität unterschieden. Im Hinblick auf *inhaltliche Gültigkeit* stellt sich die Frage, ob alle Items das zu messende Konstrukt erfassen. Je klarer der Gegenstandsbereich definiert ist, umso exakter lässt sich dies angeben. Auf der Basis der *Konstruktvalidität* (die inhaltliche und kriteriumsbezogene Validität einschliesst) wird bestimmt, ob mit dem im Test gezeigten Verhalten auf zugrunde liegende Fähigkeits- oder Persönlichkeitskonstrukte geschlossen werden kann. Dies äussert sich darin, dass die gemessene Grösse mit ähnlichen Konzeptionen zusammenhängt, respektive von inhaltsfremden unabhängig ist. *Augenscheinvalidität* bezeichnet die offensichtliche Akzeptanz eines Tests durch Laien und sollte nicht mit der Inhaltsvalidität verwechselt werden. Die *Kriteriumsgültigkeit* schliesslich widmet sich der praktischen Anwendbarkeit eines Tests. Diese spezielle Form der Konstruktvalidität beschreibt, wie gut sich der Test zur Erfassung des zu messenden Kriteriums eignet. Anhand von Korrelationen der Testvariablen mit der Kriteriumsvariablen (beispielsweise ein anderer, angeblich dasselbe Konstrukt messender, valider Test) werden entsprechende Aussagen generiert (S. 24-27).

Rauchfleisch (2008) spricht in diesem Zusammenhang von *interner* (innerer) Validität. Im Gegenzug versteht er unter *externer* (äusserer) Validierung die Korrelation mit einer objektiv bewerteten Kriteriumsleistung beziehungsweise mit einem selbst- oder fremdbewerteten Schätzurteil (S. 71).

Zur Bestimmung der Konstruktvalidität können die theoretischen Konstruktionen mittels eines sogenannten nomologischen Netzwerks (das auf Korrelationen mit anderen validen Tests beruht) operationalisiert werden. Hier besteht allerdings die Gefahr eines Zirkelschlusses, wenn verschiedene Tests als valide angesehen werden, die ihre Gültigkeit aufgrund von Korrelationen untereinander bewiesen haben. Wenn dieses Netzwerk aber aus mindestens drei Messverfahren zusammengesetzt ist und unter gemeinsamer Betrachtung als Validierungsversuch angesehen wird, stellt dies eine besonders starke Methode der Validierung dar (Asendorpf, 2011, S. 51).

Indem die Validität ein sehr umfassendes Konstrukt darstellt, existiert kein einzelner Kennwert wie bei der Reliabilität. Vielmehr werden die eingangs erwähnten, unterschied-

lichen Arten von Validität zur Umschreibung der einzelnen Anwendungsbereiche herangezogen (Krohne & Hock, 2007, S. 65). Die Kriteriumsvalidität kann jedoch durch die korrelative Überprüfung des Testwertes mit einem Aussenkriterium eine statistische Kennzahl liefern. Aufgrund derer ist es möglich, das Ausmass der Gültigkeit absolut zu bestimmen. Das macht sie zu einem wichtigen Qualitätsmerkmal psychometrischer Tests und wird mit dem Kürzel r_{tc} beschrieben. Dabei steht der Buchstabe r für das Ausmass der Korrelation eines Tests (t) mit einem Kriterium (c für English criterion) (Fisseni, 2004, S. 63) Die Übereinstimmung kann anhand eines zeitlich koexistierenden Kriteriums, wie beispielsweise eines anderen Tests (sogenannte *Übereinstimmungsvalidität*) oder eines in der Zukunft liegenden Aussenkriteriums wie beispielsweise dem Arbeitserfolg (sogenannte *prognostische Validität*) bestimmt werden. Das Problem liegt jedoch darin, ein geeignetes Aussenkriterium zu finden, da dieses seinerseits nicht hundertprozentig messgenau sein wird (Kubinger, 2009, S. 64).

3.5.4 Abhängigkeit der Hauptgütekriterien

Die drei Hauptgütekriterien stehen nach Pospeschill (2010) in einem spezifischen Abhängigkeitsverhältnis. Dabei präsentiert sich die Objektivität als notwendiges, aber nicht hinreichendes Merkmal für die Reliabilität. Das heisst, dass die Resultate eines nicht objektiven Tests hauptsächlich auf Fehler bei der Durchführung, Auswertung oder Interpretation zurückzuführen sind. In diesem Verständnis wird auch die Reliabilität als notwendig, aber nicht hinreichend für die Validität betrachtet. Dies impliziert, dass Ergebnisse eines nicht reliablen Tests kein einheitliches Konstrukt messen und somit unterschiedliche Sachverhalte erfassen. Ein nicht valides Verfahren lässt also keine angemessenen Aussagen zu und es lassen sich darauf basierend keine sinnvollen Entscheide fällen (S. 27).

4 Die Erhebungsinstrumente und daraus resultierende Hypothesen

Die in der Untersuchung eingesetzten Verfahren stellen Persönlichkeits-Struktur-Tests dar. Es handelt sich um mehrdimensionale Persönlichkeitstests, die im Bereich der sogenannten normalen Persönlichkeit verschiedene, quantifizierbare Persönlichkeitsmerkmale messen (Petermann, Holling, Leutner & Brähler, 2002, S. XIV). Dabei sind sie bestrebt, die grundlegenden individuellen Unterschiede in den Eigenschaftszügen der Persönlichkeit zu ermitteln (Pervin, Cervone & John, 2005, S. 282).

Auf den inhaltlichen Darlegungen der einzelnen Skalen gründen die im letzten Abschnitt dieses Kapitels formulierten Hypothesen. Diese postulieren Zusammenhänge verschiedener Faktoren beider Verfahren, deren Überprüfungen im danach folgenden, empirischen Teil vorgestellt werden.

4.1 Guilford-Zimmermann Temperament Survey (GZTS)

Der GZTS basiert auf den im Kapitel 2.5 bereits erwähnten 13 Dimensionen des Temperamentes und wurde 1949 erstmals publiziert. Er stellt die überarbeitete Version dar von früheren, ebenfalls auf den 13 Primärfaktoren gründenden Persönlichkeitsinventaren. Diese wiesen trotz angestrebter Orthogonalität recht hohe Interkorrelationen auf, sodass die 13 Größen zu den 10 überarbeiteten und teilweise neu bezeichneten Skalen des GZTS zusammengefasst wurden (Bartussek, 1996, S. 55-57).

Diese 10 Standards werden durch jeweils 28 bis 30 Items gebildet, die Guilford, Guilford & Zimmermann (1978) vor allem aus den Vorgängerversionen generierten und denen sie die in Tabelle 2 aufgeführten Qualitäten zuordnen (S. 3-5).

Tabelle 2: Skalen der GZTS-Originalversion (eigene Darstellung nach Guilford et al., 1978)

Skalenbezeichnung	Qualitäten bei hoher Ausprägung	Qualitäten bei niedriger Ausprägung
General activity	Vital, energiegeladen, schnell handelnd, Produktion, Effizienz	Langsam/bewusst handelnd, ermüdbar, sich Zeit lassend
Restraint	Ernsthaftigkeit, bedacht, besonnen, seriös, ausdauernde Bemühungen, Selbstkontrolle	Unbekümmert, impulsiv, liebt die Aufregung
Ascendance	Durchsetzung, Führernatur, andere überzeugend	Auffälligkeiten vermeidend, ergeben, unterwürfig
Sociability	Pflegt viele Bekanntschaften, sucht den Kontakt/das Rampenlicht, greift in Unterhaltungen ein	Zurückhaltend in Gesprächen, schüchtern, pflegt wenige Bekanntschaften
Emotional stability	Optimistisch, heiter, ausgeglichene Stimmungen/Interessen/Energie	Wechselnde Stimmungen/ Interessen, pessimistisch, erregbar, tagträumend
Objectivity	Dickhäutig, sachlich	Sensibel, selbstbezogen
Friendliness	Respekt vor anderen, Erdulden feindseliger Handlungen	Streitlustig, Widerstand gegen Herrschaft
Thoughtfulness	Selbstbesinnend, reflektierend, beobachtend, mentale Ausgeglichenheit	An offensichtlichen Sachverhalten interessiert
Personal relations	Tolerant, andere respektierend, friedfertig	Überkritisch, nörgelnd, streitlustig, Selbstmitleid

Skalenbezeichnung	Qualitäten bei hoher Ausprägung	Qualitäten bei niedriger Ausprägung
Masculinity	Interesse an männlichen Aktivitäten, typisch männliche Gefühlsreaktionen	Interesse an weiblichen Aktivitäten, mitfühlend

Durchführung

Für die Beantwortung der Items wird ein dichotomes Ja-/Nein-Antwortschema vorgegeben. Ebenso wird die Möglichkeit geboten, eine Frage mit Weiss-Nicht zu beantworten, wobei die Testpersonen aufgefordert sind, diese so selten als möglich zu wählen. Für die Beantwortung wird kein Zeitlimit vorgegeben, Guilford et al. (1978) postulieren eine durchschnittliche Bearbeitungsdauer von 45 Minuten (S. 18).

Gütekriterien

Die Retest-Reliabilitätskoeffizienten der einzelnen Skalen liegen für das Ein-Jahres-Intervall zwischen $r_{tt} = 0.53$ (Skala Ascendence) und $r_{tt} = 0.80$ (Skala Masculinity). Der Koeffizient der Splithalf-Reliabilität wird mit $r_{tta} = 0.75$ (für Skalen Objectivity und Friendliness) bis $r_{tta} = 0.87$ (für Skala Sociability) angegeben (Ebda, 1978).

Für Übereinstimmungskoeffizienten von Fremd- und Selbstbeurteilungen verweisen Guilford, Zimmermann und Guilford (1976) auf eine Untersuchung mit dem „Inventory of Factors STDCR“, einer Vorgängerversion des GZTS, die sich aus fünf der 13 Primärfaktoren zusammensetzt. Die Skalen S (soziale Introversion, direkt vergleichbar mit Sociability des GZTS) weist dabei eine Übereinstimmung von $r = 0.59$ auf. Der Faktor T (gedankliche Introversion, direkt vergleichbar mit Thoughtfulness) wird mit einer Übereinstimmung von $r = 0.33$ ausgewiesen und schliesslich R (Leichtlebigkeit als gegenpoliger Vorgänger der GZTS-Skala Restraint) stimmt mit $r = 0.66$ überein in Fremd- und Selbstbeurteilungsratings (S. 55-56).

Korrelationen mit der Variable Alter werden von Guilford et al. (1976) insofern ausgewiesen, dass die Faktoren Ascendance, Sociability, General Activity und Masculinity mit dem Alter abnehmen (also signifikant negativ korrelieren), Restraint und Thoughtfulness hingegen zunehmen (signifikant positiv korrelieren) (S. 101-103).

Die Anzahl Bildungsjahre korrelieren positiv mit den Skalen Ascendance, Emotional Stability, Objectivity und Personal Relations (Guilford et al., 1976, S. 106).

Das Geschlecht korreliert mit dem GZTS dahingehend, dass Frauen zu höheren Werten bei den Skalen Restraint, Sociability, Thoughtfulness und Personal Relations tendieren (Guilford et al., 1976, S. 108).

Deutsche Skalenstruktur

Bei der überarbeiteten deutschen Version wurden die der Skala Männlichkeit zugehörigen Items vor der testtheoretischen Überprüfung komplett entfernt. Aufgrund der Reliabilitätsprüfung wurden weitere Fragen gestrichen, sodass die aktuelle, in dieser empirischen Erhebung angewendete Ausführung 232 Items beinhaltet (siehe Anhang F). Im Rahmen der faktorenanalytischen Überarbeitung präsentieren sich die folgenden neun, neu benannten, Skalen:

- Skala ES: Emotionale Stabilität (entspricht Emotional Stability): 42 Items
- Skala S: Soziabilität (entspricht Sociability): 36 Items
- Skala V/PB: Vertrauensfähigkeit/Persönliche Beziehungen (entspricht Personal Relation): 37 Items
- Skala SV: Selbstvertrauen: 22 Items
- Skala Gel: Gelassenheit (entspricht Restraint): 28 Items
- Skala GA: Generelle Aktivität (entspricht General Activity): 17 Items
- Skala D: Durchsetzungsfähigkeit: 18 Items
- Skala N: Nachdenklichkeit (entspricht Thoughtfulness): 17 Items
- Skala Gew: Gewissenhaftigkeit (entspricht Restraint): 14 Items

4.2 NEO-Fünf-Faktoren-Inventar nach Costa und McCrae (NEO-FFI)

Das NEO-FFI stellt einen Standardtest zur Messung der Big Five dar. Er wurde von Costa und McCrae als Kurzform des in Kapitel 2.6 beschriebenen NEO-PI-R konzipiert und erfasst die in Tabelle 3 dargestellten Merkmalsbereiche.

Tabelle 3: Erfasste Merkmalsbereiche des NEO-FFI (nach Borkenau & Ostendorf, 2008, S.7)

Faktor	Qualitäten bei hoher Ausprägung
Neurotizismus	Nervös, ängstlich, traurig, unsicher, verlegen, Neigung zu unrealistischen Ideen, weniger angemessener Umgang mit Stresssituationen, Neigung zu Gesundheitsorgen
Extraversion	Gesellig, aktiv, gesprächig, personenorientiert, herzlich, optimistisch, heiter, suchen Anregung und Aufregung
Offenheit für Erfahrung	Wissbegierig, kreativ, phantasievoll, unabhängig im eigenen Urteil, hohe Wertschätzung für neue Erfahrung, Abwechslung bevorzugend
Verträglichkeit	Altruistisch, mitfühlend, verständnisvoll, wohlwollend, zu zwischenmenschlichem Vertrauen, Kooperativität und Nachgiebigkeit neigend, starkes Harmoniebedürfnis
Gewissenhaftigkeit	Ordentlich, zuverlässig, hart arbeitend, diszipliniert, pünktlich, penibel, ehrgeizig, systematisch versus nachlässig und gleichgültig

Nach Borkenau und Ostendorf (2008) misst das NEO-FFI diese fünf robusten Faktoren, auf welchen wahrgenommene interindividuelle Unterschiede beruhen, in ökonomischer Weise. Es wurde als Instrument zur Erfassung der Persönlichkeitseigenschaften klinisch unauffälliger Personen entwickelt und kann sowohl zu Forschungszwecken als auch im Sinne eines Breitbandverfahrens bei Fragestellungen im pädagogischen, berufsberatenden und personalpsychologischen Kontext angewendet werden (S. 12-15).

Durchführung

Die vorgegebenen 60 Items werden auf eine fünfstufige Skala von „starke Ablehnung“ bis „starke Zustimmung“ beantwortet, wobei jeder der fünf Domänen jeweils 12 Items zugeordnet werden. Es besteht keine Zeitbegrenzung bei der Bearbeitung, der durchschnittliche Zeitaufwand wird mit 10 Minuten angegeben (Borkenau & Ostendorf, 2008, S. 36-37).

Gütekriterien

Die internen Konsistenzen (Cronbachs Alpha) der Skalen liegen durchschnittlich bei $r = 0.8$ (Borkenau & Ostendorf, 2008, S. 18) und können als gut bewertet werden. Re-Test-Reliabilitäten im Zwei-Jahres-Intervall werden mit $r_{tt} = 0.65$ (Verträglichkeit) bis $r_{tt} = 0.81$ (Gewissenhaftigkeit, Extraversion) angegeben (Ebda, 2008).

Übereinstimmungen mit Fremdbeurteilungen werden von Borkenau und Ostendorf (2008) mit $r = 0.61$ (Extraversion) bis $r = 0.49$ (Verträglichkeit) als vergleichsweise hoch ange-

sehen (S. 26-27). Als weiterer Validitätsnachweis wird die Korrelation mit Aussenkriterien angeführt und mit $r = 0.30$ zwischen Gewissenhaftigkeit und Berufserfolg angegeben (Ebda, 2008), sodass das NEO-FFI im Brickenkamp Handbuch als hinreichend valide und reliabel aufgeführt wird (Brähler, Holling, Leutner & Petermann, 2002, S. 697).

Nach Borkenau und Ostendorf (2008) korrelieren alle NEO-FFI-Skalen signifikant mit dem Alter. So tendieren jüngere Testpersonen zu niedrigeren Werten im Bereich des Neurotizismus und der Extraversion, sowie zu hohen Ausprägungen der Verträglichkeit und Gewissenhaftigkeit. Bei Metaanalysen zeigen sich jedoch auch Effekte zunehmender Ängstlichkeit und Extraversion mit steigendem Lebensalter (S. 27).

4.3 Fazit , Fragestellung und Hypothesengenerierung

Ziel dieser Arbeit ist es, einen Beitrag zur Validitätsprüfung der deutschen Version des GZTS zu leisten. Dafür geht diese Arbeit der Fragestellung nach, ob zwischen der übersetzten und testtheoretisch überarbeiteten Version des GZTS und dem NEO-FFI nach Costa und McCrae in der deutschen Fassung von Borkenau und Ostendorf (2008) Übereinstimmungsvalidität besteht.

Sowohl die Skalen des GZTS als auch diejenigen des NEO-FFI werden anhand von Eigenschaftsbegriffen erläutert. In einem alltagspsychologischen Verständnis sollten dabei ähnliche Konstrukte anhand verwandter Begriffe beschrieben werden und somit vergleichbare Dimensionen der Persönlichkeit erfassen (respektive gegensätzliche Beschreibungen sollten keine Übereinstimmungen zeigen).

Die im folgenden empirischen Teil vorgestellte Untersuchung verfolgt dabei das Ziel, gemeinsame Varianzen der Skalen der neuen Ausführung des GZTS und des NEO-FFI zu identifizieren. Die dabei zugrundeliegenden Hypothesen basieren auf den dargelegten Charakteristika der entsprechenden Faktoren und lauten folgendermassen:

- 1) Personen mit hoher Ausprägung der NEO-FFI-Skala Extraversion weisen positive Qualitäten der GZTS-Skalen Soziabilität auf
- 2) Personen mit hoher Ausprägung der NEO-FFI-Skala Extraversion weisen positive Qualitäten der GZTS-Skalen Generelle Aktivität auf
- 3) Personen mit hoher Ausprägung der NEO-FFI-Skala Extraversion weisen negative Qualitäten der GZTS-Skala Nachdenklichkeit auf

- 4) Personen mit hoher Ausprägung der NEO-FFI-Skala Neurotizismus weisen negative Qualitäten der GZTS-Skala Emotionale Stabilität auf
- 5) Personen mit hoher Ausprägung der NEO-FFI-Skala Neurotizismus weisen negative Qualitäten der GZTS-Skala Gelassenheit auf
- 6) Personen mit hoher Ausprägung der NEO-FFI-Skala Verträglichkeit weisen positive Qualitäten der GZTS-Skala Vertrauensfähigkeit/Persönliche Beziehungen auf
- 7) Personen mit hoher Ausprägung der NEO-FFI-Skala Gewissenhaftigkeit weisen positive Qualitäten der GZTS-Skala Gewissenhaftigkeit auf

II EMPIRISCHER TEIL

Nun folgt der empirische Teil dieser Arbeit. In diesem wird die Untersuchung vorgestellt, auf der diese Arbeit basiert. Nach der Darlegung des methodischen Vorgehens werden dabei die Ergebnisse präsentiert und interpretiert. Schliesslich wird dieser Abschnitt mit einer Methodenkritik und einem Ausblick auf folgende, potenzielle Forschungsfragen abgeschlossen.

5 Forschungsgegenstand und Methodik

Dieses Kapitel beinhaltet die Beschreibung des Untersuchungsdesigns, welchem die Ausführungen zur Datenerhebung folgen. Schliesslich wird anhand der Darlegung der angewendeten statistischen Verfahren auf die Datenauswertung eingegangen.

5.1 Untersuchungsdesign

Es handelt sich bei der vorliegenden Analyse um eine Querschnittsuntersuchung, bei der die Daten einer Stichprobe zu einem bestimmten Zeitpunkt erhoben und ausgewertet werden.

Um die im Kapitel 4.3 formulierten Zusammenhangshypothesen zu überprüfen, werden anhand einer Stichprobe zwei Merkmale untersucht. Dabei geht pro teilnehmende Person ein Messwertpaar in die Berechnung ein.

Die Testpersonen wurden in einem ersten Schritt mittels schriftlicher Ansprache aus dem erweiterten Bekanntenkreis der Autorin rekrutiert. Diese leiteten die Anfrage wiederum an interessierte Drittpersonen aus ihrem privaten oder beruflichen Umfeld weiter. Als

Anreiz für die Teilnahme wurde eine Rückmeldung in Form eines Persönlichkeitsprofils angeboten.

5.2 Datenerhebung

Die Daten wurden in einer Online-Umfrage in einem Zeitraum von Anfang Januar bis Mitte März 2012 erhoben. Die teilnehmenden Personen füllten dafür in anonymer Form einen Fragebogen aus, welcher die vollständigen Items der deutschen Fassung des GZTS und des NEO-FFI beinhaltete. Weiter wurden Angaben zu soziodemographischen Merkmalen (Alter, Geschlecht und Bildungsabschluss) erhoben, von denen aufgrund theoretischer Befunde ein Bezug zur Fragestellung vermutet wird.

5.3 Datenauswertung

Die ausgefüllten Fragebögen wurden von der Online-Befragungsplattform in eine Excel-datei eingelesen. Mithilfe dieser Tabellenkalkulationssoftware wurden die einzelnen Items unter Berücksichtigung ihrer spezifischen Polung den entsprechenden Skalen zugeordnet.

Für die Ermittlung der Skalenscores wurden (basierend auf dem jeweiligen Antwortformat) beim GZTS die Werte 1 (für Zustimmung) und 0 (für Ablehnung/Weiss nicht) vergeben. Die Items des NEO-FFI erhielten Werte zwischen 0 (für starke Ablehnung) und 4 (für starke Zustimmung), welche schliesslich zu den jeweiligen Skalenwerten summiert wurden. Fehlende Werte wurden dabei durch den zugehörigen Mittelwert ersetzt. Die aufbereitete Excel-datei wurde danach in die Statistiksoftware SPSS 19.0 importiert, mit welcher anschliessend die statistischen Datenanalysen durchgeführt wurden.

Zur Überprüfung der formulierten Zusammenhangshypothesen eignet sich die Ermittlung des Korrelationskoeffizienten, der die Enge des Zusammenhanges quantifiziert. Bei den vorliegenden kontinuierlichen Merkmalen wird dabei von Intervallskalenniveau ausgegangen. In diesem Fall wird die Zusammenhangsstärke mit dem Korrelationskoeffizienten (r) nach Pearson ermittelt. Die statistische Bedeutsamkeit wird bei normalverteilten bivariaten Zusammenhängen mittels eines Signifikanztestes ausgedrückt. Dieser quantifiziert die Irrtumswahrscheinlichkeit, eine nicht zutreffende Hypothese anzunehmen und wird per Konvention auf dem Signifikanzniveau α bei 5% festgelegt (Bortz & Döring,

2006, S. 494). Die ermittelten Koeffizienten werden ungerichtet auf 2-seitige Signifikanz hin überprüft.

Der Korrelationskoeffizient wird als Effektstärke interpretiert und gemäss den Konventionen nach Cohen (1988; zit. nach Sedlmeier & Renkewitz, 2008, S. 221) für die Sozialwissenschaften folgendermassen definiert:

- $r \approx 0.1$: schwacher Zusammenhang
- $r \approx 0.3$: mittlerer Zusammenhang
- $r \approx 0.5$: starker Zusammenhang

Der quadrierte Korrelationskoeffizient r^2 (Determinationskoeffizient) gibt das Kontingenz der Varianz des einen Merkmals an, das aufgrund der Daten des anderen Merkmals erklärt wird. Durch die Multiplikation des Determinationskoeffizienten mit 100 (Redundanz) wird dieser Wert als R^2 bezeichnet und anschaulich in prozentuaem Anteil dargestellt (Bortz, 2005, S. 209). Dieser Hinweis auf den Anteil erklärter Varianz (R^2) sagt jedoch nichts aus über die Richtung dieses Zusammenhanges. Diese Grösse liefert also keinen kausalen Hinweis, welches Merkmal ein anderes definiert, sondern nur, wie gross der prozentuale Anteil gemeinsamer Varianz ist.

6 Darstellung der Ergebnisse

In diesem Kapitel werden die zur Beantwortung der Fragestellung relevanten Ergebnisse aufgezeigt. Beginnend wird auf den Umgang mit Eingabefeldern und fehlenden Werte eingegangen. Danach folgt die Erörterung der univariaten deskriptiven Statistik, welchen die Ergebnisse zur Hypothesenprüfung folgen. Diese werden durch Ausführungen zu weiteren, sich für die Fragestellung als bedeutsam erweisenden, Resultaten ergänzt. Ebenso werden die vorgefundenen Zusammenhänge der deutschen GZTS-Skalen mit ausgewählten demographischen Merkmalen besprochen. Schliesslich wird auf mögliche Scheinkorrelationen eingegangen, die eine Interpretation verzerren können.

6.1 Identifikation von Eingabefeldern

Empirisch erfasste Datensätze können unvollständig sein. Dies geschieht, wenn Testpersonen einzelne Fragen nicht beantworten. Dies kann beispielsweise vorkommen, wenn ihnen die Frage zu persönlich scheint, oder weil sie die Antwort nicht einschätzen können. Ebenso besteht die Möglichkeit, dass sich Eingabefehler einschleichen (beispiels-

weise eine Altersangabe in Monaten anstatt der geforderten Bezeichnung in Jahren). Um Verzerrungen dadurch zu vermeiden ist es wichtig, diese Werte anhand von Häufigkeitsauszählungen zu identifizieren und gegebenenfalls zu korrigieren.

Bei der vorliegenden überschaubaren Stichprobengrösse wurde diese Kontrolle im Rahmen der Datenverarbeitung mittels Excelprogramm vor dem Import in die Statistiksoftware SPSS durchgeführt.

Die demographischen Variablen weisen bei den Häufigkeitsauszählungen keine fehlenden oder fehlerhaften Angaben auf, wie die nun folgende (und im Anhang A ausführlich beschriebene) deskriptive univariate Statistik zeigt.

6.2 Beschreibung der Stichprobe

Die Stichprobe besteht aus 43 Frauen und 26 Männern. Insgesamt haben also 69 Personen an der Befragung teilgenommen. Dabei handelt es sich um 62% weibliche und 38% männliche Testpersonen.

Die jüngste Testperson ist 23 Jahre alt, die älteste 72. Das durchschnittliche Alter beträgt 44.26 Jahre, die Standardabweichung liegt bei 12.6 Jahren. Wie in Abbildung 3 veranschaulicht, sind die meisten Teilnehmerinnen und Teilnehmer zwischen 31 und 40 Jahre alt. Am wenigsten Testpersonen beinhaltet die Alterskategorie zwischen 51 und 60 Jahren. Knapp ein Viertel der teilnehmenden Personen sind zwischen 41 und 50 Jahre alt.

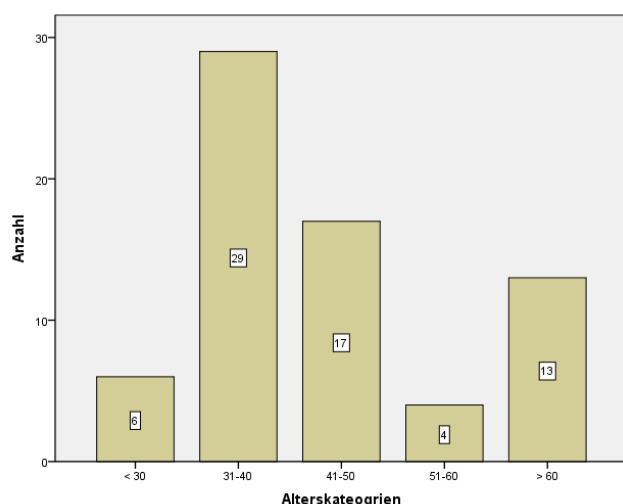


Abbildung 3: Häufigkeiten Alterskategorien

61% der teilnehmenden Personen weisen einen Bildungsabschluss auf tertiärer Stufe aus (Fachhochschule, Universität, ETH). Da die gesamte Stichprobe einen Bildungsabschluss

besitzt, verfügen wie in Abbildung 4 ersichtlich die restlichen 39% der Testpersonen über einen Abschluss auf Sekundarstufe II (Berufliche Grundausbildung, Berufsfachschulen, Mittelschulen).

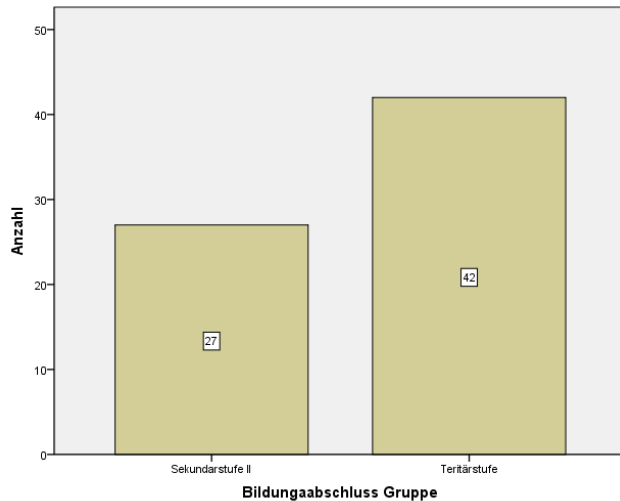


Abbildung 4: Häufigkeiten Bildungsabschluss

6.3 Beschreibung der Skalenwerte

Es handelt sich (wie aufgrund der Q-Q-Diagramme im Anhang B ersichtlich) um normalverteilte Daten. Bei diesen Summenscores wird von kontinuierlichen, intervallskalierten Werten ausgegangen.

Tabelle 4 gibt diverse statistische Kennwerte der einzelnen Skalen sowie der Gesamtscores des GZTS und NEO-FFI an.

Tabelle 4: Statistische Kennwerte Skalen GZTS & NEO-FFI

Skalenbezeichnung	N	Minimum	Maximum	Potenzielles Maximum	Mittelwert	Standardabweichung
Emotionale Stabilität	69	1	32	42	13.6	7.36
Soziabilität	69	9	32	36	24.2	4.48
Vertrauensfähigkeit	69	1	25	37	13.3	4.98
Selbstvertrauen	69	2	18	22	8.8	4.13
Gelassenheit	69	1	21	29	11.0	5.33
Aktivität	69	0	14	17	5.7	3.46
Durchsetzungsfähigkeit	69	3	17	18	11.5	2.85
Nachdenklichkeit	69	0	10	17	3.6	2.19
Gewohnhaftigkeit GZTS	69	3	14	14	8.4	2.49
Gesamtscore GZTS	69	69	131	-	100.3	13.3

Skalenbezeichnung	N	Minimum	Maximum	Potenzielles Maximum	Mittelwert	Standardabweichung
Neurotizismus	69	2	37	48	17.3	7.71
Extraversion	69	12	44	48	28.3	6.06
Offenheit	69	18	42	48	31.4	5.37
Verträglichkeit	69	21	45	48	35.1	4.59
Gewissenhaftigkeit	69	19	45	48	34.1	4.81
Gesamtscore NEO-FFI	69	110	180	-	146.2	13.2
Gültige Werte	69					

Die Skala Soziabilität weist mit 24.2 den höchsten Mittelwert bei einer vergleichsweise kleinen Standardabweichung von 4.48 auf. Der Faktor Nachdenklichkeit verfügt mit 3.6 über den niedrigsten Mittelwert und die kleinste Standardabweichung (2.19). Daneben präsentieren sich die Skalen Generelle Aktivität, Gewissenhaftigkeit und Selbstvertrauen mit Mittelwerten unter 10 und Standardabweichungen unter 5. Der Faktor Emotionale Stabilität wird mit 7.36 als Dimension mit der höchsten Standardabweichung angegeben.

Die Skalen Generelle Aktivität und Nachdenklichkeit erreichen das absolute Minimum von 0, wohingegen bei der Dimension Soziabilität keine Werte unter 9 ermittelt wurden.

Der Skalenwert Gewissenhaftigkeit schöpft den potenziellen Höchstwert von 18 aus, beim Faktor Durchsetzungsfähigkeit werden Werte bis 17 von erreichbaren 18 gemessen. Bei der Skala Nachdenklichkeit werden 10 von möglichen 17 Punkten erreicht. Die Dimension Vertrauensfähigkeit/Persönliche Beziehungen wird mit maximal 25 Punkten angegeben, bei einem potenziellen Höchststand von 37.

Die gemessenen GZTS-Gesamtscores liegen zwischen 69 und 131 und weisen einen Mittelwert von 100.3 auf. Die Standardabweichung beträgt 13.3.

Die NEO-FFI-Skala Neurotizismus weist mit 17.2 den niedrigsten Mittelwert und die grösste Standardabweichung (7.72) auf. Die Dimensionen Verträglichkeit und Gewissenhaftigkeit verfügen über die höchsten Mittelwerte (35.1 respektive 34.3) und die niedrigsten Standardabweichungen (4.61 respektive 4.94). Beide erreichen sie ein Maximum von 45 (von 48 möglichen) Punkten.

Der niedrigste evaluierte Gesamtscore des NEO-FFI beträgt 110, der höchste 180. Diese Werte weisen einen Mittelwert von 146.2 bei einer Standardabweichung von 13.2 auf.

6.4 Ergebnisse zur Überprüfung der Hypothesen

Die Validität wurde als komplexes qualitätssicherndes Konstrukt psychometrischer Persönlichkeitstest vorgestellt. Neben verschiedenen weiteren Vorgehensweisen kann dabei ein Kennwert ermittelt werden, der einen Testwert mit einem anderen auf ihren Zusammenhang hin prüft. Die Höhe dieser Korrelation wird mittels des Koeffizienten r_{tc} ausgedrückt.

Der Produkt-Moment-Koeffizient nach Pearson ist nur bei linearen Zusammenhängen aussagekräftig. Um diesen graphisch erfassen zu können, eignet sich ein Streudiagramm. Die Enge des Zusammenhanges ist durch die Positionierung der Punkte um die Gerade charakterisiert, wobei kleine Abstände einen engen Zusammenhang beschreiben. Verläuft die Gerade von links unten nach rechts oben wird ein positiver Zusammenhang dargestellt. Bei einer negativen Beziehung verläuft die Regressionsgerade entgegengesetzt von rechts unten nach links oben. Neben der ersten optischen Prüfung der Richtung und des Ausmasses der Beziehung erweist sich das Streudiagramm als ebenso hilfreich, um kurvilineare Verlaufsformen identifizieren zu können. Diese können durch Verzerrung des Koeffizienten Fehlinterpretationen nach sich ziehen. Ebenso besteht dabei die Möglichkeit, Ausreisser als weitere potenzielle Verfälscher zu erkennen (Sedlmeier & Renkewitz, 2008, S. 207-215).

In Tabelle 5 werden die Korrelationskoeffizienten zwischen den für die Hypothesen relevanten Skalen des GZTS und denjenigen des NEO-FFI dargestellt. Die aufgrund der formulierten Annahmen zu überprüfenden Zusammenhangsmasse sind dabei grau unterlegt. Ist eine solche Korrelation signifikant oder hoch signifikant, ist der Koeffizient fett markiert.

Tabelle 5 Korrelationswerte nach Pearson der hypothesenrelevanten Skalenwerte GZTS & NEO-FFI

	ES	S	V/PB	Gel	GA	N	Gew
Neurotizismus	.71**	-.20	.43**	-.58**	.255*	-.16	-.15
Extraversion	-.35**	.73**	.05	.403**	.40**	.05	-.34**
Verträglichkeit	-.12	.17	-.27*	-.01	-.22	.05	-.03
Gewissenhaftigkeit	-.02	.05	.08	-.13	.10	-.270*	.37**

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant

* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant

Überprüfung der Hypothese 1

Es wird erwartet, dass zwischen der NEO-FFI-Skala Extraversion und der GZTS-Skala Soziabilität (S) ein positiver Zusammenhang besteht.

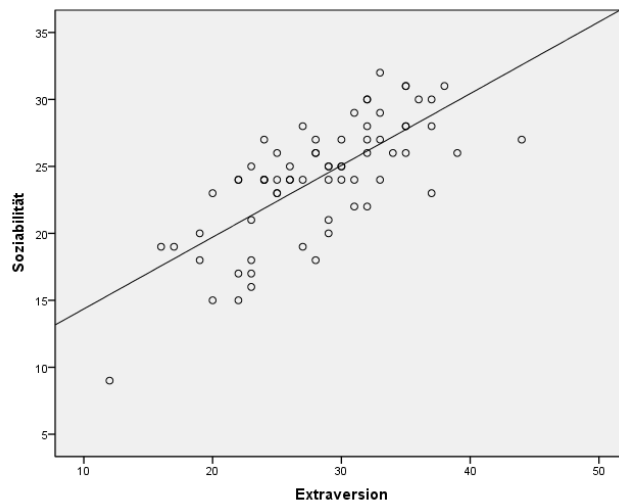


Abbildung 5: Streudiagramm Extraversion – Soziabilität

Das Streudiagramm (Abbildung 5) zeigt einen engen, positiven linearen Zusammenhang. Ebenso ist in der linken unteren Ecke ein Ausreisserwert ersichtlich. Ist dieser Wert weit weg von den übrigen Zusammenhangsgrössen, kann der Korrelationskoeffizienten nach Pearson in unzulässiger Weise verzerrt werden. Da dies zu Fehlinterpretationen führen kann, erweist sich in diesem Fall die Berechnung des Korrelationskoeffizienten nach Kendall-Tau-b als aussagekräftiger.

Dieser wird in Tabelle 6 dargestellt und liegt bei $r_{tc} = 0.56$. Er ist auf dem Niveau von α 1% sehr signifikant.

Tabelle 6: Korrelation nach Kendall-Tau-b zur Prüfung der Hypothese 1

		Soziabilität	Extraversion
Soziabilität	Korrelationskoeffizient	1	.56**
	Sig. (2-seitig)	.	0
	N	69	69
Extraversion	Korrelationskoeffizient	.56**	1
	Sig. (2-seitig)	0	.
	N	69	69

** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig)

Es besteht also gemäss Cohens Konvention (siehe Kap. 5.3) ein mittlerer bis starker, statistisch sehr bedeutsamer positiver Zusammenhang zwischen den beiden Variablen.

Der Anteil erklärte Varianz (R^2) beträgt in diesem Fall 34%. Das bedeutet, dass etwa ein Drittel der Streuung des einen Merkmals diejenige des andern erklärt.

Die Hypothese 1, dass Personen mit hoher Ausprägung der NEO-FFI-Skala Extraversion positive Qualitäten der GZTS-Skala Soziabilität aufweisen, kann bestätigt werden.

Überprüfung der Hypothese 2

Es wird postuliert, dass zwischen den Skalen Extraversion und Generelle Aktivität (GA) ein positiver Zusammenhang besteht. Das Streudiagramm (Abbildung 6) weist auf einen im Vergleich zur Hypothese 1 schwächeren, positiven, linearen Zusammenhang hin.

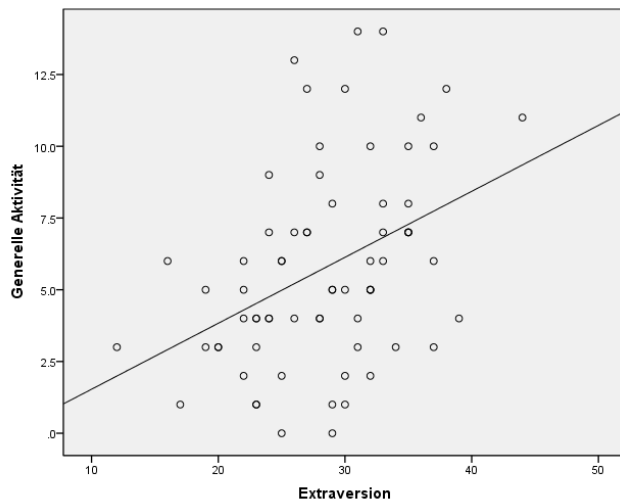


Abbildung 6. Streudiagramm Extraversion-Generelle Aktivität

Der Korrelationskoeffizient nach Pearson beträgt $r_{tc} = 0.40$ (siehe Tabelle 5) und ist ebenfalls sehr signifikant. Es besteht also ein mittlerer bis starker, positiver und statistisch bedeutsamer Zusammenhang, bei dem der Anteil erklärte Varianz 16% beträgt.

Die Hypothese 2, dass Personen mit hoher Ausprägung der NEO-FFI-Skala Extraversion hohe Werte der GZTS-Skala Generelle Aktivität zeigen, kann bestätigt werden.

Überprüfung der Hypothese 3

Es wird angenommen, dass hohe Werte der Dimensionen Extraversion mit niedrigen Werten der Dimension Nachdenklichkeit (N) einhergehen.

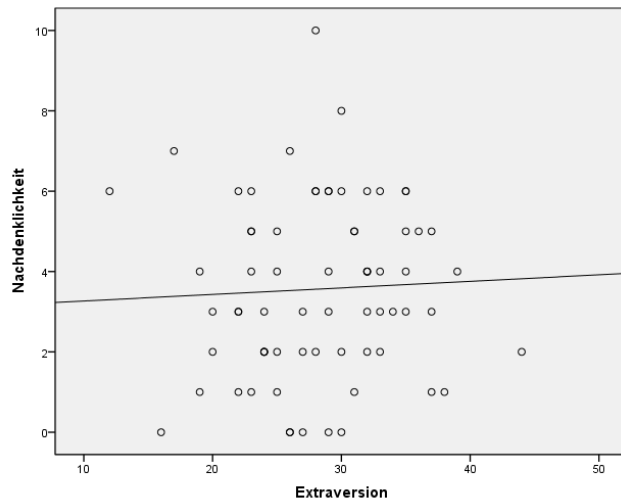


Abbildung 7: Streudiagramm Extraversion-Nachdenklichkeit

Das Streudiagramm in Abbildung 7 präsentiert sich als tendenziell kreisförmige Punktwolke mit annähernd horizontaler Gerade, was auf einen nicht nachweisbaren Zusammenhang hindeutet.

Der Koeffizient nach Pearson wird dabei mit $r_{tc} = 0.05$ angegeben und ist weder auf dem Niveau $\alpha 0.01$ noch bei $\alpha 0.05$ signifikant. Es lässt sich also kein statistisch bedeutsamer Zusammenhang feststellen.

Die Hypothese 3, dass Personen mit hoher Ausprägung der NEO-FFI-Skala Extraversion über negative Qualitäten der GZTS-Skala Nachdenklichkeit verfügen, kann nicht bestätigt werden.

Überprüfung der Hypothese 4

Es wird erwartet, dass zwischen der NEO-FFI-Skala Neurotizismus und der GZTS-Skala emotionale Stabilität (ES) ein negativer Zusammenhang besteht.

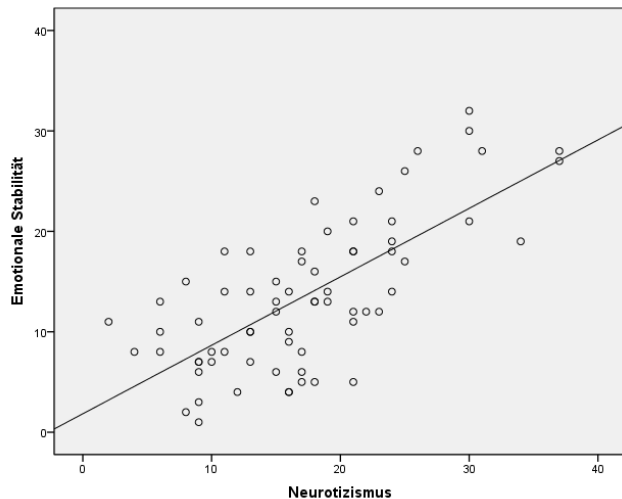


Abbildung 8: Streudiagramm Neurotizismus - Emotionale Stabilität

Das entsprechende Streudiagramm in Abbildung 8 weist auf einen engen, positiven, linearen Zusammenhang hin. Die Korrelation nach Pearson beträgt $r_{tc} = 0.71$ und wird als sehr signifikant ausgewiesen.

Es besteht also ein starker, statistisch sehr bedeutsamer positiver Zusammenhang zwischen den beiden Dimensionen. Der Anteil erklärte Varianz beträgt 50%. Somit teilen die beiden Merkmale die Hälfte ihrer Varianz, was in den Sozialwissenschaften als hoher Anteil gewertet wird.

Die Hypothese 4, dass Personen mit hoher Ausprägung der NEO-FFI-Skal Neurotizismus bei der GZTS-Skala Emotionale Stabilität eine niedrige Ausprägung aufweisen, kann nicht bestätigt werden.

Jedoch besteht ein starker, sehr signifikanter positiver Zusammenhang zwischen den beiden Variablen.

Überprüfung der Hypothese 5

Es wird damit gerechnet, dass ein hohes Ausmass an Neurotizismus mit tiefen Gelassenheitswerten (Gel) assoziiert ist.

Das Streudiagramm (Abbildung 9) weist auf einen negativen, linearen Zusammenhang hin.

Der Korrelationskoeffizient nach Pearson beträgt $r_{tc} = -0.58$ und ist auf dem Niveau 0.01 signifikant. Dies bedeutet, dass zwischen den beiden Skalen ein starker, negativer und

statistisch hoch bedeutsamer Zusammenhang besteht. Der Anteil erklärte Varianz beträgt 24%.

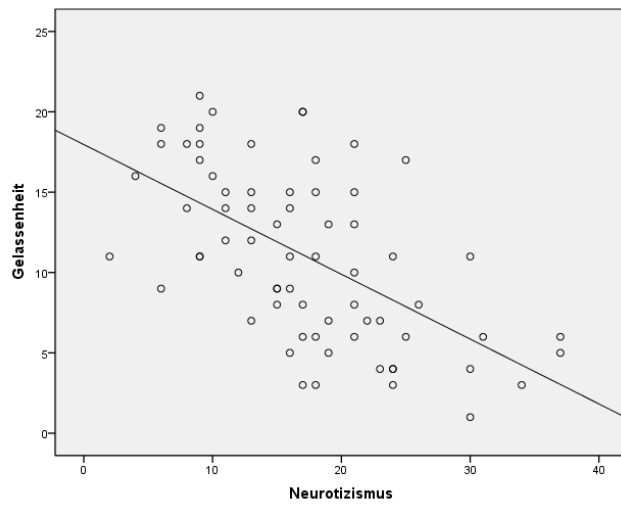


Abbildung 9: Streudiagramm Neurotizismus Gelassenheit

Die Hypothese 5, dass Personen mit hoher Ausprägung der NEO-FFI-Skala Neurotizismus negative Qualitäten der GZTS-Skala Gelassenheit aufweisen, kann bestätigt werden.

Überprüfung der Hypothese 6

Es wird erwartet, dass zwischen den Dimensionen Verträglichkeit und Vertrauensfähigkeit/Persönliche Beziehungen (V/PB) ein positiver Zusammenhang besteht.

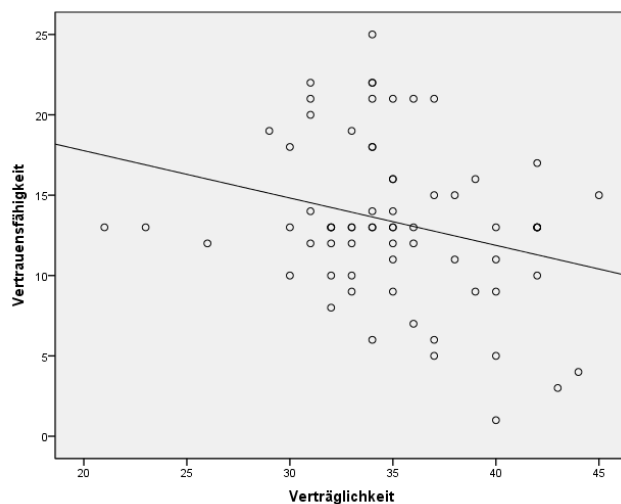


Abbildung 10: Streudiagramm Verträglichkeit-Vertrauensfähigkeit

Das Streudiagramm in Abbildung 10 weist auf einen negativen Zusammenhang hin. Der Korrelationskoeffizient nach Pearson wird mit $r_{tc} = -0.27$ angegeben und ist auf dem Ni-

veau 0.01 sehr signifikant. Es besteht also ein schwacher bis mittlerer, negativer, statistisch bedeutsamer Zusammenhang zwischen den beiden Skalen. Der Anteil erklärte Varianz beträgt 7%.

Die Hypothese 6, dass Personen mit hoher Ausprägung der NEO-FFI-Skala Verträglichkeit positive Qualitäten der GZTS-Skala Vertrauensfähigkeit/Persönliche Beziehungen aufweisen, kann nicht bestätigt werden.

Es wird jedoch ein schwacher negativer statistisch bedeutsamer Zusammenhang nachgewiesen.

Überprüfung der Hypothese 7

Es wird erwartet, dass ein positiver Zusammenhang besteht zwischen der Skala Gewissenhaftigkeit des GZTS (Gew) und derjenigen des NEO-FFI.

Das Streudiagramm (Abbildung 11) weist auf einen linearen, positiven Zusammenhang hin.

Der Korrelationskoeffizient nach Pearson beträgt $r_{tc} = 0.37$ und ist hoch signifikant. Es zeigt sich also in der Analyse ein statistisch bedeutsamer, mittlerer positiver Zusammenhang zwischen den beiden Merkmalen. Der Anteil erklärte Varianz beträgt 14%.

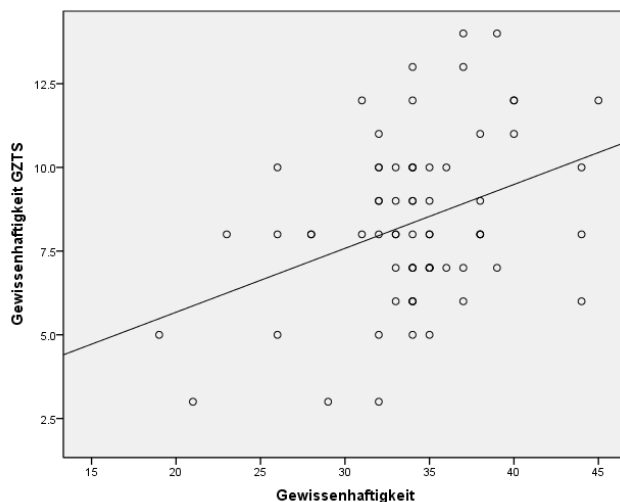


Abbildung 11: Streudiagramm Gewissenhaftigkeit GZTS -Gewissenhaftigkeit NEO-FFI

Die Hypothese 7, dass Personen mit hoher Ausprägung der NEO-FFI-Skala Gewissenhaftigkeit hohe Ausprägung der GZTS-Skala Gewissenhaftigkeit aufweisen, kann bestätigt werden.

Zusammenstellung der Ergebnisse zur Hypothesenprüfung

Zusammenfassend kann festgehalten werden, dass von den sieben formulierten Hypothesen vier bestätigt werden können und drei abgelehnt werden müssen. Von diesen drei zurückgewiesenen Annahmen präsentieren sich zwei als durchaus mindestens schwache und statistisch bedeutsame Zusammenhänge. Jedoch korrelieren sie nicht in der postulierten Richtung miteinander. Die dritte Hypothesenprüfung weist einen zu niedrigen Korrelationskoeffizienten auf, um von einem nachweisbaren Zusammenhang sprechen zu können. Ebenso kann er wegen einer zu grossen Irrtumswahrscheinlichkeit nicht als statistisch bedeutungsvoll angesehen werden. Tabelle 7 fasst diese Sachverhalte zusammen.

Tabelle 7: Übersicht Ergebnisse Hypothesenprüfung

Hypothese	Resultat	Entscheidung
1	$r_{tc} = 0.58^{**}/R^2 = 34\%$	bestätigt
2	$r_{tc} = 0.40^{**}/R^2 = 16\%$	bestätigt
3	$r_{tc} = 0.05/R^2 = --$	<u>nicht bestätigt</u>
4	$r_{tc} = 0.71^{**}/R^2 = 50\%$	<u>nicht bestätigt</u>
5	$r_{tc} = -0.58^{**}/R^2 = 24\%$	bestätigt
6	$r_{tc} = -0.27^{**}/R^2 = 7\%$	<u>nicht bestätigt</u>
7	$r_{tc} = 0.37^{**}/R^2 = 14\%$	bestätigt

6.5 Überprüfung weiterer Zusammenhänge

Weitere Korrelationen zwischen den GZTS- und NEO-FFI-Skalen

Zur detaillierteren Analyse der Fragestellung werden an dieser Stelle ergänzend zu den Hypothesen weitere Korrelationen zwischen den bereits untersuchten Skalenscores der deutschen GZTS-Version und dem NEO-FFI erörtert.

Tabelle 8 stellt die Korrelationskoeffizienten nach Pearson dar. Indem diese keine hypothesengeleitete Analyse darstellt, werden ausschliesslich die fett markierten, signifikanten, mittleren und starken Zusammenhänge ($r = \geq 0.3$) besprochen.

Tabelle 8: Korrelationskoeffizienten nach Pearson Skalen GZTS -NEO-FFI

	ES	S	V/PB	SV	Gel	GA	D	N	Gew
Neurotizismus	.71**	-.20	.43**	.45**	-.58**	.26*	-.24	-.16	-.15
Extraversion	-.35**	.73**	.05	-.56**	.40**	.40**	.27*	.05	-.34**
Offenheit	.04	.20	.11	-.25*	-.01	.03	.05	-.29*	.01
Verträglichkeit	-.12	.17	-.27*	0.07	-.01	-.22	-.16	.05	-.03
Gewissenhaftigkeit	-.02	.05	.08	-0.08	-.13	.10	-.04	-.27*	.37**

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant

* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant

Die entsprechenden Wertmesser des GZTS weisen dabei die folgenden, statistisch bedeutsamen und mindestens mittelstarken linearen Zusammenhänge (siehe Streudiagramme im Anhang C) mit dem NEO-FFI auf:

- Die Skala Emotionale Stabilität (ES) korreliert mit $r_{tc} = -0.35$ mit der Skala Extraversion. Dieses Ergebnis ist sehr signifikant. Es besteht also ein mittlerer, negativer und statistisch sehr bedeutsamer Zusammenhang zwischen der Skala Emotionale Stabilität und Extraversion. Der Anteil erklärte Varianz beträgt 12%.
- Die Skala Vertrauensfähigkeit/Persönliche Beziehung (V/PB) korreliert mit $r_{tc} = 0.43$ mit der Skala Neurotizismus. Dieses Ergebnis ist sehr signifikant. Es lässt sich ein mittlerer bis starker, statistisch sehr bedeutsamer Zusammenhang zwischen den Skalen Vertrauensfähigkeit und Neurotizismus nachweisen, bei dem der Anteil erklärte Varianz 18.5% beträgt.
- Die Skala Gewissenhaftigkeit des GZTS (Gew) korreliert mit $r_{tc} = -0.34$ mit der Skala Extraversion. Dieses Ergebnis ist sehr signifikant. Es besteht also ein mittlerer, negativer, statistisch sehr bedeutsamer Zusammenhang zwischen der Skala Gewissenhaftigkeit des GZTS und der Skala Extraversion. Diese beiden Skalen teilen 11.5% ihrer Varianz.
- Die Skala Gelassenheit (Gel) weist mit $r_{tc} = 0.40$ einen mittleren, statistisch sehr bedeutsamen Zusammenhang mit hohen Werten der Extraversion auf. Der Anteil erklärte Varianz beträgt 16%.
- Die Skala Selbstvertrauen (SV) weist einen mittleren bis starken, positiven Zusammenhang mit der Skala Neurotizismus auf. Der Korrelationskoeffizient beträgt $r_{tc} = 0.45$. Der Anteil erklärte Varianz beträgt 20.3%.
- Die Skala Selbstvertrauen weist einen starken, negativen Zusammenhang mit der Skala Extraversion auf. Dabei beträgt $r_{tc} = -0.56$ und der Anteil erklärte Varianz beläuft sich auf 32%.

Korrelation Gesamtscores

Um die Fragestellung nach Übereinstimmungsvalidität neben der Analyse auf Skalenebene zu erweitern, werden nun die beiden Gesamtscores auf ihren Zusammenhang hin überprüft.

Das Streudiagramm (siehe Anhang C) der normalverteilten Gesamtscores des GZTS und des NEO-FFI (siehe Q-Q-Diagramm Anhang B) weist auf einen positiven, schwach ausgeprägten linearen Zusammenhang hin.

Wird diese Analyse ungerichtet auf zweiseitige Signifikanz hin geprüft, wird der Koeffizient mit $r_{tc} = 0.22$ ausgewiesen. Die Irrtumswahrscheinlichkeit beträgt in diesem Fall 6.4% und liegt somit über dem Niveau von 0.05.

Es wird jedoch von einem positiven Zusammenhang ausgegangen, weshalb in einem weiteren Schritt ein gerichteter, einseitiger Signifikanztest durchgeführt wird. Die Irrtumswahrscheinlichkeit beträgt in diesem Fall 3.2%, womit das Ergebnis als signifikant ausgewiesen wird (siehe Tabelle 9). Es besteht also ein schwacher, positiver, signifikanter Zusammenhang zwischen den Gesamtscores beider Tests. Der Anteil erklärte Varianz beträgt 4.8%.

Tabelle 9: Korrelation nach Pearson Summenscore GZTS-NEO-FFI

		Summenscore GZTS	Summenscore NEO-FFI
Summenscore GZTS	Korrelation nach Pearson	1	.224*
	Signifikanz (1-seitig)		0.032
	N	69	69
Summenscore NEO-FFI	Korrelation nach Pearson	.224*	1
	Signifikanz (1-seitig)	0.032	
	N	69	69

* Die Korrelation ist auf dem Niveau von 0,05 (1-seitig) signifikant

Korrelationen GZTS mit demographischen Variablen

Im Kapitel 4.1 wurde über Korrelationen der Originalversion GZTS mit demographischen Variablen berichtet. Im Sinne einer weiteren Überprüfung der deutschen Version werden nun die überarbeiteten Skalen hinsichtlich eines möglichen Zusammenhangs mit den Merkmalen Alter, Bildung und Geschlecht überprüft. Dabei kann das Merkmal Bildung als dichotome, nominalskalierte Größe ebenso mit dem Korrelationskoeffizient nach

Pearson quantifiziert werden (Sedlmeier & Renkewitz, 2008, S. 235). Das nominalskalierte Merkmal Geschlecht wird mit dem Koeffizienten nach Spearman-Rho ausgewiesen.

In Tabelle 10 sind die signifikanten Korrelationen zwischen Skalen des GZTS und den soziodemographischen Merkmalen Alter, Bildung und Geschlecht aufgeführt. Signifikante Koeffizienten von im Rahmen dieser Arbeit überprüften Skalen sind fett markiert.

Tabelle 10: Korrelationen Skalen GZTS -demographische Variablen Alter, Bildung und Geschlecht

	Alter	Bildungsabschluss	Geschlecht
Emotionale Stabilität	-.14	-.11	.14
Soziabilität	-.17	.05	.11
Vertrauensfähigkeit	-.15	-.20	-.01
Selbstvertrauen	.11	-.32**	.25*
Gelassenheit	.07	.36**	-.27*
Generelle Aktivität	-.24*	.12	.15
Durchsetzungsfähigkeit	.09	.26*	-.01
Nachdenklichkeit	-.03	.10	-.14
Gewissenhaftigkeit GZTS	.33**	-.10	-.17
Neurotizismus	-.27*	-.23	.24*
Extraversion	-.23	.23	.11
Offenheit	-.23	.21	-.04
Verträglichkeit	.03	-.07	.16
Gewissenhaftigkeit	.05	.01	.09

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant

Es wird dabei ersichtlich, dass die Skala Gewissenhaftigkeit einen positiven, mittleren ($r = 0.33$) und Generelle Aktivität einen negativen, schwachen ($r = -0.24$), statistisch bedeutsamen Zusammenhang mit der Variable *Alter* aufweisen. Unter Berücksichtigung der Polung bedeutet dies, dass ältere Personen über ein geringeres Ausmass der Dimension Generelle Aktivität, sowie über ein höheres Ausmass der Dimension Gewissenhaftigkeit verfügen. Mit dem Faktor Neurotizismus zeigt sich ein schwacher ($r = -0.27$), negativer Zusammenhang, was aufgrund der Polung bedeutet, dass ältere Menschen niedrigere Neurotizismuswerte aufweisen.

Es zeigt sich mit der Variable *Bildung* ein mittlerer positiver Zusammenhang ($r = 0.36$) zur Skala Gelassenheit. Unter Berücksichtigung der Polung bedeutet dies, dass Personen mit einem tertiären Bildungsabschluss (mit 2 kodiert) über ein höheres Ausmass der Dimension Gelassenheit berichten. Weiter zeigt sich ein mittlerer negativer Zusammenhang

($r = -0.32$) zwischen der Skala Selbstvertrauen und der Bildung. Dabei verfügen Personen mit einem höheren Abschluss über stärkere Ausprägungen der Dimension Selbstvertrauen.

Zwischen dem Merkmal *Geschlecht* und der Skala Gelassenheit zeigt sich ein mittlerer, negativer Zusammenhang ($r = -0.27$). Dies bedeutet unter der Berücksichtigung der Polung, dass Männer (mit 1 codiert) höhere Werte aufweisen beim Faktor Gelassenheit. Ebenso lässt sich bei $r = 0.25$ ein schwacher Zusammenhang des Geschlechtes mit der Skala Selbstvertrauen feststellen. Somit weisen Frauen (mit 2 codiert) höhere Werte in diesem Faktor auf. Angesichts der schwachen Korrelation ($r = 0.24$) mit dem Faktor Neurotizismus kann aufgrund der Polung gesagt werden, dass bei Frauen höhere Werte in diesem Bereich nachgewiesen werden.

Überprüfung von Scheinkorrelationen

Ein aufgezeigter statistischer Zusammenhang zwischen zwei Merkmalen stellt keinen Hinweis auf eine kausale Beziehung dar. Deshalb sollten bei einem nicht experimentell erhobenen Befund Überlegungen angestellt werden, wie diese Verknüpfung zustande kommt. Dabei kann neben einer tatsächlichen (echten) Beziehung auch eine Scheinkorrelation (bei der eine Verbindung zweier Variablen durch ein Drittmerkmal erklärt wird) als Ursache in Betracht gezogen werden. Dieser Umstand lässt sich bei intervallskalierten Merkmalen mit linearem Zusammenhang statistisch mittels Berechnung von Partialkorrelationen bestimmen.

Unter dieser Vorgabe werden in Tabelle 11 die Koeffizienten der Partialkorrelationen mit den Kontrollvariablen Alter und Bildungsabschluss dargestellt. Wenn bei diesem Verfahren ein zuvor bestimmter Korrelationskoeffizient nicht sinkt, kann ausgeschlossen werden, dass die konstant gehaltene Kontrollvariable als Ursache für die Verbindung in Frage kommt (Sedlmeier & Renkewitz, 2008, S. 233).

Da jedoch Merkmale im sozialwissenschaftlichen Umfeld fast nie komplett unkorreliert sind, sollten nur diejenigen Allianzen geprüft werden, bei denen aufgrund theoretischer Überlegungen oder früherer empirischer Belege eine Interaktion vermutet wird. Aufgrund der im vorangegangenen Abschnitt angeführten Beziehungen der deutschen GZTS-Version mit den demographischen Variablen sind in der Tabelle die in Frage kommenden Skalen dargestellt. Dabei werden nur im Rahmen dieser Untersuchung bereits untersuchte

Zusammenhänge analysiert (komplette Korrelationsmatrix siehe Anhang E). Signifikante, gegenüber der Berechnung nach Pearson verminderte Koeffizienten sind dabei hervorgehoben.

Tabelle 11 Partialkorrelationen Skalen GZTS-NEO-FFI mit Kontrollvariablen Alter und Bildungsabschluss

		V/PB	SV	Gel	GA	Gew
Neurotizismus	Korrelation	.38	.46	-.55	.24	-.09
	Signifikanz (zweiseitig)	.00	.00	.00	.05	.46
Extraversion	Korrelation	.07	-.52	.38	.35	-.28
	Signifikanz (zweiseitig)	.60	.00	.00	.00	.02
Offenheit	Korrelation	.13	-.18	-.07	-.06	.11
	Signifikanz (zweiseitig)	.29	.15	.56	.65	.37
Verträglichkeit	Korrelation	-.29	.05	.02	-.22	-.05
	Signifikanz (zweiseitig)	.02	.69	.89	.08	.69
Gewissenhaftigkeit NEO-FFI	Korrelation	.08	-.08	-.15	.10	.37
	Signifikanz (zweiseitig)	.52	.53	.23	.43	.00

Bei 65 Freiheitsgraden weisen gegenüber der Kalkulation nach Pearson insgesamt *sechs* Verbindungen einen signifikanten, verminderten Koeffizienten auf. Dabei stellt sich die Verbindung Gewissenhaftigkeit GZTS – Extraversion mit $r_{tc} = -0.28$ mit einem um 0.06 kleineren Koeffizienten dar.

Die Stärke des Zusammenhanges zwischen den Skalen Vertrauensfähigkeit/Persönliche Beziehung – Neurotizismus weist mit $r_{tc} = 0.38$ einen um 0.05 verminderten Koeffizienten auf.

Die Beziehung Generelle Aktivität – Extraversion sinkt bei $r_{tc} = 0.35$ mit einer Reduktion von 0.05 im selben Umfang.

Die Zusammenhangsstärke der Beziehung Selbstvertrauen – Extraversion sinkt um 0.04 auf $r_{tc} = 0.52$.

Der Koeffizient zwischen den Skalen Gelassenheit – Neurotizismus sinkt um 0.03 auf $r_{tc} = -0.55$ und derjenige der Verbindung Gelassenheit – Extraversion nimmt um 0.02 auf $r_{tc} = 0.38$ ab.

7 Diskussion

Als Grundlage dieser Diskussion werden einleitend die vorgestellten Erklärungen des theoretischen Teils zusammengefasst. Danach werden die präsentierten Resultate interpretiert, um die Fragestellung zu beantworten. Darauf folgt eine kritische Reflexion der Methodik. Abschliessend werden im Sinne eines Ausblicks weiterführende Punkte erörtert, die gegebenenfalls ergänzende Forschungsfragen generieren können.

7.1 Zusammenfassung

Die vorliegende Arbeit befasst sich mit einem Validierungsverfahren eines neu auf Deutsch übersetzten, faktorenanalytisch konzipierten, mehrdimensionalen Persönlichkeitsfragebogens. Hierbei wird die Übereinstimmungsvalidität mit einem etablierten Verfahren überprüft. Im Theorieteil dieser Arbeit werden zum einen Aspekte faktorenanalytisch begründeter Persönlichkeitstheorien dargelegt. Zum andern werden unter einer persönlichkeitsdiagnostischen Perspektive testtheoretische Gesichtspunkte erörtert.

Dabei interessieren im zweiten Kapitel insbesondere das Persönlichkeitsmodell von Guilford und das Fünf-Faktoren-Modell, die mit den eingesetzten Erhebungsinstrumenten verbunden sind. Diese Erklärungen menschlicher Erlebens- und Verhaltensweisen basieren auf dem Eigenschaftsparadigma, welches von zeitstabilen Wesenszügen ausgeht. Für deren Identifikation wird auf eine Sprachanalyse existierender Begriffe (den sogenannten lexikalischen Ansatz) zurückgegriffen. Diese Charakteristiken sind durch die individuelle Positionierung auf einem Kontinuum zwischen extremer und minimaler Ausprägung skaliert und somit messbar, womit die Voraussetzung für eine interindividuelle Vergleichbarkeit gegeben ist. Anhand faktorenanalytischer Berechnungen werden in beiden Konzeptionen diejenigen Dimensionen der Persönlichkeit bestimmt, die für ein Verständnis menschlicher Erlebens- und Verhaltensweisen notwendig und nützlich erscheinen. Beide Modelle bedienen sich dabei einer hierarchischen Ordnung, was Rückschlüsse eines höheren Niveaus auf ein tieferliegendes zulässt. Dies bietet den Vorteil, Aussagen über noch nicht beobachtete Situationen zuzulassen.

Im dritten Kapitel wird den Gütekriterien besonderes Augenmerk gewidmet. Die Validität repräsentiert dabei die inhaltliche Gültigkeit eines Tests und erweist sich als wichtigstes, aber auch komplexestes Gütekriterium. Es steht in einem Abhängigkeitsverhältnis mit den andern Hauptgütekriterien der Reliabilität und der Objektivität, indem diese beiden

notwendige, aber nicht hinreichende Voraussetzung darstellen für die Gültigkeit eines Verfahrens.

Im vierten Kapitel werden die beiden zu vergleichenden Erhebungsverfahren vorgestellt. Aufgrund der Darlegungen der spezifischen Skaleninhalte werden Zusammenhangshypothesen zur Ermittlung fraglicher Übereinstimmungsvalidität generiert, die den Theorieteil abschliessen.

Im empirischen Teil untersucht diese Arbeit, ob zwischen den beiden dargelegten Persönlichkeitstests Parallelen bestehen. Damit stellt sich die Frage nach potenziellen Zusammenhängen zwischen simultan erfassten Merkmalen. Für die Hypothesenprüfung werden die jeweiligen Skalenwerte mittels Korrelationsberechnungen mathematisch verglichen und auf ihre statistische Bedeutsamkeit hin kontrolliert. Das Ergebnis zeigt, dass bei vier von sieben Konstellationen tatsächlich der postulierte Zusammenhang feststellbar ist. Zwei weitere erwartete Beziehungen stellen sich in der entgegengesetzten Richtung dar. Eine Hypothese muss gänzlich verworfen werden. Darüber hinaus werden die GZTS-Skalen mittels weiteren Korrelationsanalysen detaillierter betrachtet. Auf Skalenebene können sechs weitere, statistisch bedeutsame Beziehungen bestimmt werden. Ebenso korrelieren die Gesamtscores der beiden Tests signifikant miteinander.

Aufgrund theoretischer Darlegungen stehen die demographischen Variablen Alter, Geschlecht und Bildung mit den Skalenwerten in Verbindung. Entsprechende Befunde können in dieser Untersuchung teilweise bestätigt werden. Wenn nun analysierte Interaktionen durch eine Drittvariable verzerrt werden, wird von Scheinzusammenhängen gesprochen. Mittels der Berechnung von Partialkorrelationen können diese identifiziert werden. Dabei werden diejenigen Beziehungen bestimmt, deren Zusammenhangsstärke durch Konstanthaltung einer Kontrollvariable vermindert wird.

7.2 Interpretation der Ergebnisse

Die Zusammenstellung der Hypothesenüberprüfung weist auf, dass vier von sieben Hypothesen bestätigt werden. Dabei zeigen sich teilweise starke oder zumindest mittelstarke Zusammenhänge von Skalen, die aufgrund ihrer jeweiligen Beschreibung die Annahme zulassen, dass sie ähnliche Eigenschaften messen sollten. Mit Ausnahme der Verbindung Extraversion – Nachdenklichkeit (Hypothese 6) liegt die Irrtumswahrscheinlichkeit min-

destens unter 5%, sodass aufgrund der Ergebnisse (statistisch) bedeutende Aussagen generiert werden können.

Hypothese 1

Die Ergebnisprüfung der Hypothese 1 weist die erwartete hohe Übereinstimmung der Skalen Extraversion und Soziabilität auf. Beide Dimensionen werden mit Begriffen wie Geselligkeit, Personenorientiertheit und Anregungsbedürfnis in Verbindung gebracht. Aufgrund des signifikanten Ergebnisses eines mittleren bis starken Zusammenhangs lässt sich sagen, dass sich die neu konzipierte und übersetzte Version der GZTS-Skala Soziabilität (entspricht Sociability der Originalversion) Verhaltensweisen misst, die im Fünf-Faktoren-Modell als extravertiert klassifiziert sind.

Indem der Anteil erklärte Varianz bei 34% liegt lässt sich rückschließen, dass die Streuungen beider Merkmale zu 46% durch nicht gemeinsame Merkmale verursacht werden. Es werden also über die Konvergenzen hinausgehende Kapazitäten menschlicher Lebens- und Verhaltensweisen erfasst. Somit handelt es sich bei diesen beiden Skalen also nicht um identische (und somit redundante) psychologische Qualitäten.

Hypothese 2

Nach Überprüfung der Hypothese 2 lässt sich die Beziehung zwischen der NEO-FFI-Skala Extraversion und der GZTS-Skala Generelle Aktivität (entspricht in der Originalversion der Skala General Activity) als mittleren positiven Zusammenhang klassifizieren. Dies bedeutet, dass die deutsche GZTS-Skala Dimensionen der Extraversion bei der Messung miteinbezieht. So erscheint das Phänomen der Aktivität in den Skalenbeschreibungen beider Verfahren. Diejenige des GZTS postuliert, mit Merkmalen wie Energie und Vitalität Qualitäten zu erfassen, die eine aktive Lebensgestaltung implizieren. Extravertierte Menschen kennzeichnet nach den Autoren des NEO-FFI ebenso eine aktive Haltung, die sie nach Anregung und Aufregung suchen lässt.

Die 16% des Anteils erklärte Varianz erscheinen aufgrund der 84% anderweitig erklärten Streuung als nicht übermässig. Jedoch offenbart die Auflistung der zugehörigen Eigenschaften auch Differenzen. So bezieht der Faktor Generelle Aktivität auch Aspekte der Impulsivität mit ein, die beispielsweise im NEO-FFI keine explizite Erwähnung finden.

Hypothese 3

Die dritte Hypothese, dass hohe Ausprägungen der Skala Extraversion mit niedrigen Werten der GZTS-Skala Nachdenklichkeit (entspricht Thoughtfulness der Originalversion) einhergehen, konnte mittels dieser Untersuchung nicht bestätigt werden.

Der Korrelationskoeffizient fällt mit $r_{tc} = 0.05$ so tief aus, dass nicht von einem nachweisbaren Zusammenhang gesprochen werden kann.

Unter dieser Prämisse kann bei der Skala Nachdenklichkeit nicht vom postulierten Gegenpol zur Extraversion ausgegangen werden. Vielmehr scheint diese Dimension des GZTS eine von der Extraversion unabhängige Erlebens- und Verhaltensweise zu erfassen. Jedoch sei in diesem Zusammenhang darauf hingewiesen, dass die Irrtumswahrscheinlichkeit über dem 5% Niveau liegt, sodass jegliche Aussagen einer zu grossen Fehlerquelle unterliegen, als dass sie statistisch bedeutsam wären.

Hypothese 4

Eine weitere ausgeprägte und statistisch bedeutende Übereinstimmung kann aufgrund der Verifizierung der vierten Hypothese zwischen den Skalen Neurotizismus des NEO-FFI und Emotionale Stabilität des GZTS (entspricht Emotional Stability der Originalversion) nachgewiesen werden. Also kann gesagt werden, dass die übersetzte Version des GZTS damit Qualitäten registriert, die im FFM als Neurotizismus definiert werden.

Dabei bezieht sich die Beschreibung hoher Ausprägungen der GZTS-Skala auf eine Gleichmässigkeit der Stimmung, Interessen und Energie. Im NEO-FFI wird der Faktor Neurotizismus anhand stimmungslabilen Verhaltens- und Erlebensweisen wie beispielsweise Gesundheitssorgen oder mittels eines unangemessenen Umgangs mit Stresssituationen erfasst.

Insofern erstaunt die positive Richtung der Korrelation, welche entgegengesetzt postuliert wurde. Aufgrund dieses Ergebnisses lassen sich nun hohe Werte der einen Skala mit niedrigen der anderen assoziieren. Indem bei einem Koeffizienten von $r_{tc} = 0.71$ die Hälfte der Varianzen (50%) geteilt werden, weist besagter Wert auf eine im sozialwissenschaftlichen Kontext grosse Übereinstimmung des gemessenen Phänomens hin. Dies lässt die Frage aufkommen, ob die übersetzte Version der Skala Emotionale Stabilität bei hoher Ausprägung eine Neigung zu Stimmungsschwankungen misst. Eine Betrachtung der Itemebene (Anhang F, Fragen 5 und 9) erhärtet diesen Verdacht. So laden zustimmende

Antworten auf Fragen wie „Ihre Interessen ändern sich oft“ oder „manchmal sprühen Sie vor Energie und manchmal sind Sie träge“ positiv auf den Faktor. Wenn diese Überlegungen durch weitere Befunde erhärtet werden, sollte gegebenenfalls eine Überarbeitung der eventuell irreführenden Skalenbenennung in Betracht gezogen werden. Ebenso kann die Itempolung (siehe Anhang G) überprüft werden.

Hypothese 5

Bei der Formulierung der Hypothese 5, dass die Dimensionen Neurotizismus und Gelassenheit in einer Beziehung stehen, standen ähnliche Überlegungen im Vordergrund wie bei der Besprechung der vorangegangenen Annahme. Hohe Ausprägungen der GZTS-Skala Gelassenheit (entspricht Restraint der Originalversion) werden dabei aufgrund von Formulierungen wie Bedachtsamkeit und Selbstkontrolle umschrieben. Dem stehen Begriffe hoher Neurotizismuswerte gegenüber, die mit Ängstlichkeit und Unsicherheit skizziert werden. Das Ergebnis eines negativen, hoch signifikanten Zusammenhanges weist auf eine reziproke Übereinstimmung im angenommenen Bereich hin. In diesem Verständnis erfassen Ausprägungen der deutschen GZTS-Gelassenheitsskala Aspekte des Gegenpols des Neurotizismus.

Jedoch bleibt anzumerken, dass bei knapp einem Viertel (24%) geteilter Varianz nicht von identischen Dimensionen ausgegangen werden kann. Inhaltlich lässt sich das dahingehend erklären, dass beim Faktor Gelassenheit auch zurückhaltende, kontrollierte Aspekte der Erlebens- und Verhaltensweise einbezogen werden. Der anhand des Faktors Neurotizismus erfasste Bereich der Nervosität und Ängstlichkeit scheinen nur teilweise deren Gegenpol zu entsprechen.

Hypothese 6

Die Hypothese 6, die Gemeinsamkeiten zwischen den Faktoren Verträglichkeit und Vertrauensfähigkeit/Persönliche Beziehungen postuliert, konnte nicht bestätigt werden. Es wurde ein positiver Zusammenhang angenommen, da der Faktor Verträglichkeit mit Elementen wie zwischenmenschlichem Vertrauen, sowie wohlwollender und zu Nachgiebigkeit neigender Grundhaltung assoziiert ist. Vertrauensfähigkeit/Persönliche Beziehungen (die der Skala Personal Relations der Originalversion entspricht) soll die Toleranz anderen gegenüber und eine friedfertige Grundhaltung erfassen. Unter dieser Voraussetzung

überrascht der gefundene negative Zusammenhang, der eine konträre Übereinstimmung bedeutet.

Es stellt sich analog der Ergebnisse der Hypothese 4 die Frage, ob diese Skala eher den Gegenpol der Vertrauensfähigkeit darstellt, und ob in diesem Fall die Bezeichnung allenfalls irritierend sein kann. Bei einem Anteil erklärter Varianz von 7% kann jedoch bei einer Divergenz von 91% der Streuungen nicht von einer dominanten Übereinstimmung gesprochen werden. Insofern sollte die Struktur dieser Skala anhand von korrelierenden Beziehungen zu weiteren Skalen detaillierter evaluiert werden.

Darüber hinaus lässt das Streudiagramm der Verbindung mit der NEO-FFI-Skala Verträglichkeit (siehe Anhang C) einen kurvilinearen Zusammenhang nicht komplett ausschließen. Unter der Berücksichtigung dieses Sachverhaltes, bieten sich erweiterte statistische Berechnungen an.

Hypothese 7

Die siebte Hypothese, dass die beiden Gewissenhaftigkeitsskalen übereinstimmende Bereiche messen, konnte bestätigt werden. Die entsprechenden Analysen offenbaren einen mittleren, hoch signifikanten positiven Zusammenhang. Die deutsche GZTS-Skala Gewissenhaftigkeit registriert also Aspekte der im FFM ebenfalls als gewissenhaft beschriebenen Persönlichkeitseigenschaft. Dabei werden 14% der Varianzen geteilt, was bedeutet, dass 86% der Streuung nicht identisch erklärt werden.

Die Skala Gewissenhaftigkeit des GZTS wurde bei der überarbeiteten Skalenstruktur aus dem Faktor Restraint konzipiert (analog dem Faktor Gelassenheit). Auf Itemebene werden dieser Größe Zustimmungen zu Fragen wie „Sie mögen Aufgaben, die viel Aufmerksamkeit für das Detail erfordern“ (Frage 149, Anhang E) zugeteilt. Im Sinne der Beschreibung der Originalskala Restraint wird also nach ausdauernden Bemühungen und seriöser Entscheidungsfindung gefragt. Dies verweist auf Parallelen zum NEO-FFI, welches den Faktor Gewissenhaftigkeit mit disziplinierten, ausdauernden Verhaltensweisen beschreibt. Daneben enthält der GZTS-Faktor auch Aspekte der Zurückhaltung und Besonnenheit. Diesen wird auf Itemebene anhand von Zustimmung zu Fragen wie „Sie behalten in aufregenden Situationen einen kühlen Kopf“ oder „Sie tendieren dazu Dinge zu überdenken, bevor Sie handeln“ (Fragen 125 und 20) eingeschätzt werden. Dieses kontemplative Moment wird hingegen im NEO-FFI nicht explizit erwähnt.

Überprüfung weiterer Zusammenhänge

Neben der hypothesengeleiteten Analyse konnten aufgrund der detaillierten Analyse der Fragestellung sechs ergänzende, statistisch bedeutsame und mindestens mittelstarke Zusammenhänge identifiziert werden.

So besteht zwischen dem Faktor *Emotionale Stabilität* und Extraversion ein mittelstarker, negativer Zusammenhang. Hohe Werte der Skala Emotionale Stabilität sind also mit niedrigen Extraversionswerten assoziiert. Wenn niedrig ausgebildete Extraversionseigenschaften als Abwesenheit von Optimismus und Heiterkeit verstanden werden, kann dies als bestätigender Aspekt interpretiert werden, dass hohe Ausprägungen der Emotionalen Stabilitätsskala stimmungslabile Dispositionen erfassen (wie in der Ergebnisinterpretation der Hypothese 4 dargelegt).

Die Skala *Vertrauensfähigkeit/Persönliche Beziehungen* weist einen mittleren bis starken Zusammenhang mit dem Neurotizismusfaktor auf. Die erste Dimension beschreibt dabei ein interpersonelles Verhalten, wohingegen der NEO-FFI-Faktor eher intrapsychische Vorgänge integriert. Jedoch lassen sich hypothetisch betrachtet Parallelen dahingehend aufzeigen, dass die im GZTS beschriebene, überkritische Grundhaltung mit Tendenz zu selbstmitleidigen Gedanken zu einer im NEO-FFI erwähnten Traurigkeit und unangemessenem Umgang mit Stresssituationen führen kann. Es bleibt darüber hinaus festzuhalten, dass die Tendenz zu nörgelndem Verhalten als negative Qualität der Vertrauensfähigkeitsskala beschrieben wird. Die vorliegenden Daten verweisen indes auf eine positive Beziehung zum Neurotizismusfaktor. Dies kann als ergänzender Hinweis interpretiert werden, dass die bei der Ergebnisinterpretation von Hypothese 6 angeführte, möglicherweise irreführende Skalenbezeichnung gegebenenfalls re-evaluiert werden kann.

Personen, die im GZTS über hohe Ausprägungen der *Gewissenhaftigkeit* verfügen, weisen im NEO-FFI niedrige Extraversionswerte auf, da zwischen beiden Skalen ein negativer, mittlerer Zusammenhang besteht. Somit erfasst die neustrukturierte Skala des GZTS psychische Qualitäten, die als kontrastierende Eigenschaften zu denjenigen interpretiert werden können, welche im NEO-FFI als extravertiert beschrieben sind. Inhaltlich scheint eine Vorstellung plausibel, dass dazu ein Vermeiden von Aufregung gezählt werden kann, um eine Aufgabe gewissenhaft erledigen zu können. Fraglich bleibt, ob diese Parallele beinahe einen Fünftel der Varianz erklären kann, respektive welche weiteren Eigenschaften der negative Pol der GZTS-Skala misst.

Die Skala *Gelassenheit* des GZTS weist einen mittleren Zusammenhang mit dem Faktor der Extraversion auf. Somit lässt sich sagen, dass mit der Gelassenheitsskala Eigenschaften gemessen werden, die im NEO-FFI als extravertiert definiert sind. Inhaltlich scheinen Übereinstimmungen nicht offensichtlich. So offenbaren die Erlebens- und Verhaltensqualitäten Besonnenheit und Zurückhaltung (Gelassenheit) keine direkten Parallelen zu Geselligkeit und Aktivität (Extraversion). Diesem scheinbaren Widerspruch bleibt entgegen zu halten, dass bei 16% gemeinsamer Varianz ein Spielraum von 84% anderweitig erklärtem Streuungsausmass bleibt. Gegebenenfalls verweisen hohe Ausprägungen der Extraversion auf bisher nicht beschriebene Eigenschaften der Gelassenheit. An dieser Stelle sollte auch angemerkt werden, dass das NEO-FFI ein Breitbandverfahren darstellt, welches gegebenenfalls detailliertere Facetten der GZTS-Skala nicht erfasst.

Die Skala *Selbstvertrauen* des GZTS weist mit der Neurotizismusskala des NEO-FFI einen mittleren bis starken Zusammenhang auf. Es kann also festgehalten werden, dass anhand der Dimension Selbstvertrauen Eigenschaften erfasst werden, die im NEO-FFI unter dem Faktor Neurotizismus subsummiert werden. Bei einem Anteil erklärter Varianz von 20% bleibt zwar ein beachtliches Ausmass an Streuung, die nicht durch das andere Merkmal verursacht werden. Dennoch überrascht auf den ersten Blick die positive Richtung der Beziehung. Inhaltlich lässt es sich jedoch beispielsweise aufgrund von positiv ladenden Items wie „Sie vermeiden es manchmal, mit anderen in Kontakt zu treten, aus Angst, etwas Falsches zu tun oder zu sagen“ (Frage 50, Anhang F) nachvollziehen. Unter der Berücksichtigung, dass der Faktor Neurotizismus Facetten der Unsicherheit miteinbezieht lässt sich der simultane Zusammenhang damit erklären. Zugleich führen Zustimmungen zu Aspekten der Schüchternheit wie beispielsweise bei Frage 134 („Ihre Schüchternheit hält Sie davon ab, so beliebt zu sein, wie Sie eigentlich möchten“) zu hohen Ausprägungen der Selbstvertrauensskala. Obschon sich die Richtung des Zusammenhanges inhaltlich nachvollziehen lässt, stellt sich die Frage, ob die Skalenbezeichnung in diesem Fall desorientierend ist, da Selbstvertrauen gemeinhin als Abwesenheit von Schüchternheit verstanden wird.

Weiter weist die Skala *Selbstvertrauen* einen starken negativen Zusammenhang mit dem Faktor Extraversion des NEO-FFI auf. Bei einem Anteil erklärter Varianz von fast einem Drittel muss auch von einer im sozialwissenschaftlichen Kontext starken Beziehung ausgegangen werden. Also verweist die GZTS-Skala Selbstvertrauen auf die Abwesenheit von Eigenschaften, die im NEO-FFI als extravertiert klassifiziert sind. Auf Itemebene

lässt sich dies aufgrund von Zustimmungen zu Fragen wie „Wenn Sie jemandem begegnen, warten Sie, bis er/sie sich vorstellt“ oder „Sie sind froh, wenn andere die Führung in Gruppenaktivitäten übernehmen“ (Frage 12 respektive 26, Anhang F) nachvollziehen. Diese Bevorzugung von reaktiven, zurückhaltenden Verhaltensweisen kann als Gegenpol der Aktivität interpretiert werden, die durch die Extraversionsskala erfasst wird.

Wie in Kapitel 2.6 angeführt, kann die Schüchternheit als „Schnittmenge“ der FFM-Faktoren Extraversion und Neurotizismus verstanden werden. Die Selbstvertrauens-Skala thematisiert an verschiedenen Stellen diese Erlebensqualität. So lauten weitere Items: „Ihre Schüchternheit stört Sie“ (Frage 86, Anhang F), respektive: „Es ist schon oft vorgekommen, dass Sie aufgrund Ihrer Schüchternheit gezögert haben, Verabredungen anzunehmen oder vorzuschlagen“ (Frage 114). Die vorliegenden Daten zeigen also eine Übereinstimmung mit diesem Postulat. Jedoch bestärken beide Befunde der Zusammenhangsrichtung den Sachverhalt einer gegebenenfalls irritierenden Skalenbezeichnung, die analog zu den GZTS-Skalen Emotionale Stabilität und Vertrauensfähigkeit/Persönliche Beziehungen von weiteren Analysen profitieren kann.

Die *Gesamtscores* des GZTS und des NEO-FFI korrelieren in der entsprechenden Analyse mit einem Wert von $r_{tc} = 0.22$, was als schwacher positiver, signifikanter Zusammenhang gilt. Unter dieser Prämisse kann festgehalten werden, dass die Gesamtscores beider Tests knapp 5% der Varianzen teilen und somit in sehr begrenztem Umfang eine Übereinstimmung messbar ist. Auf Skalenebene konnten fast durchwegs höhere Korrelationen festgestellt werden. Jedoch stellt sich die Frage, inwiefern ein Gesamtscore bei einem mehrdimensionalen Persönlichkeitstest diagnostisch relevant ist. So interessieren in diesem Kontext mehrheitlich die verschiedenen Ausprägungen der bestenfalls unabhängigen, einzelnen Dimensionen.

Des Weiteren muss erwähnt werden, dass eine übermäßige Übereinstimmung nicht angestrebt werden sollte, da ansonsten tautologische Schlussfolgerungen entstehen können. In diesem Fall lässt sich kein diagnostisch relevanter Mehrwert generieren, da beide Tests identische psychische Qualitäten messen. Vielmehr besteht bei einem Verfahren mit einer hohen Übereinstimmungsvalidität mit einem anderen, etablierten diagnostischen Instrument existentieller Rechtfertigungsbedarf. Schliesslich scheint die Frage legitim, weshalb ein weiterer Test existieren soll, wenn ein bereits bestehender den entsprechenden Inhalt valide und reliabel misst.

Die Skalen beider Tests korrelieren wie in Kapitel 6.5 dargelegt auch mit demographischen Variablen. Die anhand der vorliegenden Daten identifizierten Zusammenhänge zwischen dem *Alter* und den Skalen Generelle Aktivität und Gelassenheit werden auch für die Originalversion des GZTS postuliert. Unter diesem Aspekt scheint das Lebensalter auch bei der übersetzten Skala einen erwarteten Einfluss auszuüben. Die vorgestellte Übereinstimmung der Ausprägung der Skala Durchsetzungsfähigkeit und *Bildungsabschluss* lässt sich auch bei der englischen Ausgabe des GZTS finden. Analog dazu beschreiben Personen mit längerer Bildung (tertiärer Abschluss) höhere Ausprägungen der Skala Durchsetzungsfähigkeit. Die für die Originalversion des GZTS postulierten Korrelationen des *Geschlechtes* können in dieser Erhebung nicht bestätigt werden. Es kann hingegen an dieser Stelle festgehalten werden, dass das Geschlecht und der Bildungsabschluss im Zusammenhang stehen mit der Ausprägung des GZTS-Faktors Selbstvertrauen. Indem diese Dimension nicht direkt mit einer Originalskala vergleichbar ist, kann jedoch keine Parallele zur englischen Version hergestellt werden. Die anhand der vorliegenden Daten festgestellte Gegebenheit, dass Männer über ein höheres Mass an Gelassenheit berichten, wird in der Literatur zur Originalfassung nicht beschrieben.

Die Analyse der Partialkorrelationen zeigt bei sechs der in Frage kommenden Zusammenhangsmasse eine Reduktion des Korrelationskoeffizienten bei Konstanthaltung der intervallskalierten Kontrollvariablen Alter und Bildungsabschluss. Dabei sinkt dieser um den maximalen Umfang von 0.06. Die grössten Differenzen werden bei den GZTS-Skalen Gewissenhaftigkeit, Vertrauensfähigkeit/Persönliche Beziehungen und Generelle Aktivität nachgewiesen. Somit lässt sich dabei ein geringer Einfluss der Merkmale Alter und Bildung nicht ausschliessen. Die Skalen Gewissenhaftigkeit und Selbstvertrauen korrelieren darüber hinaus signifikant mit dem Geschlecht, womit auch der Einfluss dieser Grösse nicht restlos geklärt ist. Jedoch verändern sich die Werte nur minimal, sodass nicht von einem bedeutenden Einfluss dieser Drittvariablen ausgegangen werden muss.

Beantwortung der Fragestellung

Die Frage nach Übereinstimmungsvalidität eines zu überprüfenden Tests mit einem Ausserkriterium stellt eine von verschiedenen Herangehensweisen dar, die Validität eines Tests zu überprüfen. Im Bereich der einzelnen Skalen konnten diverse Zusammenhänge identifiziert werden. Auch die Gesamtscores korrelieren miteinander. Unter Berücksichtigung dieser Aspekte lässt sich der Frage auf Ebene der getesteten Skalenverbindungen

und der Gesamtwerte des GZTS und des NEO-FFI zustimmen. Für eine abschliessende Stellungnahme müssen die weiteren signifikanten Interkorrelationen auf Skalenebene analysiert werden. Ebenso würde sich eine Überprüfung anhand einer grösseren beziehungsweise homogeneren Stichprobe anbieten, um eventuell einen höheren Korrelationskoeffizienten der Gesamtscores zu erreichen.

7.3 Kritik und Reflexion

Zur Analyse der Hypothesen lag eine quantitative Methodenwahl auf der Hand. Zur Bestimmung von Übereinstimmungen zweier Merkmale haben sich die gewählten Korrelationsberechnungen bewährt. Die Prüfung von Scheinkorrelationen zeigt sich dabei als nützliches und angemessenes Instrument, um die Echtheit eines Zusammenhanges zu analysieren und die Möglichkeiten eines Fehlschlusses zu reduzieren. Indem demographische Variablen erhoben wurden, die in einer theoretischen Verbindung mit den gemessenen Konstrukten stehen, konnten diese gezielt überprüft werden. Dank eines Vergleiches dieser Beziehungen zwischen der Originalversion des GZTS und deren deutschen Ausführung konnten entsprechende Aussagen über die Beantwortung der Fragestellung hinaus herbeigeführt werden.

Im Rahmen dieser Untersuchung lieferte der Hauptteil der Analysen signifikante Ergebnisse. So konnte die übersetzte und überarbeitete deutsche Version des GZTS anhand des Referenzmodells der Big Five einer qualitätssichernden Analyse unterzogen werden. Aufgrund der Bestätigung der Mehrheit der Hypothesen lassen sich zu den überprüften Skalen fundierte Aussagen generieren, auf denen schliesslich die Beurteilung der Fragestellung basiert.

Verbesserungspotenzial lässt sich bei dieser Arbeit im Rahmen der Stichprobengrösse eruieren. In kleinen Gruppen können Ausreisser im Datensatz das Ergebnis beträchtlich verzerren. Der Umfang dieser Stichprobe lässt mit 69 teilnehmenden Personen und beinahe einer 2/3-Mehrheit weiblicher Testpersonen und tertiärer Bildungsabschlüsse nur sehr beschränkt die Bildung von Untergruppen zu, die eine angemessene Grösse ($n > 30$) für statistische Berechnungen aufweisen. Ebenso ist die Repräsentativität der anfallenden Stichprobe durch die Rekrutierung im erweiterten Bekanntenkreis der Autorin eingeschränkt. Jedoch konnte durch die persönliche Ansprache eine erfreulich hohe Rücklaufquote erreicht werden. Bei einer durchschnittlichen Bearbeitungsdauer des Fragebogens

von 45-60 Minuten schien eine grossangelegte, anonyme Rekrutierung nicht erfolgsversprechend.

Durch die Online-Befragung konnte kein motivationaler Einfluss auf die Testpersonen geltend gemacht werden. Gegebenenfalls hatte dieser Faktor einen Einfluss auf die Anzahl der Weiss-nicht-Antworten. Aus demselben Grund besteht keine Kontrolle über mögliche intervenierende Variablen oder die Kontextbedingungen. So ist beispielsweise nicht gesichert, dass die Antworten tatsächlich als Selbstbeurteilungen unter ruhigen, für die Selbstreflexion förderlichen Bedingungen zustande kamen. Dem ist entgegen zu halten, dass in diesem anonymen Setting Versuchsleiterartefakte keinen bedeutenden Einfluss ausüben.

Als Herausforderung präsentierte sich die Erhebung mit einem neu übersetzten und im Bereich der Faktorenstruktur gegenüber dem Original deutlich veränderten Test. Als Folge davon ist sicherlich die noch nicht optimale Faktorenstruktur zu nennen. Indem sich diese sehr heterogen präsentiert, sind die entsprechenden Aussagen zu den einzelnen Dimensionen beeinträchtigt. Ebenso zeigen sich widersprüchliche Befunde zwischen der inhaltlichen Beschreibung und der Skalenbenennung. In der aktuellen Form präsentiert sich die deutsche Version des GZTS als noch nicht direkt in der psychologischen Diagnostikpraxis anwendbar.

Gesamthaft lässt sich sagen, dass im Rahmen dieser Arbeit ein erster Schritt des umfangreichen und komplexen Prozesses der Validitätsprüfung erfolgen konnte.

7.4 Ausblick

Für weitere Aussagen zur Fragestellung bietet sich die Analyse ergänzender Korrelationen auf Skalenebene an. Insbesondere die GZTS-Skala Durchsetzungsfähigkeit wurde in dieser Arbeit nicht untersucht. Es wurden auch keine Angaben zu Parallelen mit der NEO-FFI-Skala Offenheit gemacht. Ferner wurden über die Hypothesenprüfung hinaus keine schwachen Zusammenhänge interpretiert. Eine Betrachtung bezüglich fehlender Übereinstimmungen der Skalen beider Tests lassen schliesslich Aussagen zur sogenannten diskriminanten Validität (also inwiefern die Skalen inhaltlich nicht denselben Inhalt messen) zu.

Um die Validität fundierter zu überprüfen, kann eine Überprüfung der Übereinstimmung von Fremd- und Selbstbeurteilungen erfolgen. Ebenso lässt sich diese mittels eines objek-

tiven Schätzurteils evaluieren, indem beispielsweise Parallelen zum Berufserfolg ermittelt werden. Ausserdem bietet sich auch die Möglichkeit, die Zufriedenheit von berufsberaterischen Resultaten von der beratenen Person einschätzen zu lassen, und diese Ergebnisse auf bedeutsame Zusammenhänge zu prüfen. Weitere potenzielle Aussenkriterien zur Validierung werden in operationalisierten, valide und reliabel messbaren psychologischen Konstrukten gesehen. Diese lassen im Sinne einer Ergänzung des FFM Aussagen zur Konstruktvalidität zu, indem sie Präzisierungen zum Inhalt der erfassten psychischen Qualitäten liefern können. In diesem Zusammenhang bietet sich beispielsweise das Phänomen der Empathiefähigkeit als Aspekt der emotionalen Intelligenz an.

Ebenso kann anknüpfend an die vorhandenen Daten anhand einer deutlich grösseren Stichprobe ein Normierungsverfahren angegangen werden. Entsprechende Profile können sich analog der GZTS-Originalversion in der psychologischen Diagnostikpraxis als nützlich erweisen, da sie eine Einschätzung des erhobenen Testwertes zulassen.

Im Sinne einer Optimierung der Skalenstruktur kann eine Homogenisierung des Faktorenumfangs angestrebt werden, um einen diagnostisch aussagekräftigeren Test generieren zu können.

Des Weiteren bietet sich die Überprüfung der Benennung und inhaltlichen Struktur derjenigen Skalen an, deren Zusammenhang mit dem NEO-FFI entgegen theoretisch hergeleiteten Hypothesen ausfiel. Im Rahmen der vorliegenden Arbeit betrifft dies namentlich die Skalen Emotionale Stabilität, Persönliche Beziehung/Vertrauensfähigkeit und Selbstvertrauen.

Die Originalversion des GZTS basiert auf einer hierarchischen Faktorenstruktur. Wenngleich deren Ordnung nicht restlos geklärt ist, bietet sich als weitere Überarbeitung der deutschen Fassung die Evaluation von Interkorrelationen der GZTS-Skalen an (siehe Anhang D), um eine gegebenenfalls korrespondierende Ordnung zu identifizieren.

Abschliessend kann festgehalten werden, dass diese Arbeit einen Beitrag zur Testevaluation leisten konnte, der Potenzial für weiterführende Untersuchungen offenbart.

Literatur

- Amelang, M. & Borkenau, P. (1982). Über die faktorielle Struktur und externe Validität einiger Fragebogen-Skalen zur Erfassung von Dimensionen der Extraversion und emotionalen Labilität. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 3 (2), 119-146.
- Amelang, M. & Schmidt-Atzert, L. (2006). *Psychologische Diagnostik und Intervention* (4., vollst. überarb. u. erw. Aufl.). Heidelberg: Springer Medizin.
- Amelang, M., Bartussek, D., Stemmler, G. & Hagemann, D. (2006). *Differentielle Psychologie und Persönlichkeitsforschung* (6. vollst. überarb. Aufl.). Stuttgart: Kohlhammer.
- Andresen, B. (2000). Six Basic Dimensions of Personality and a Seventh Factor of Genrealized Dysfunctional Personality. *Neuropsychobiology* 41, 5-23.
- Asendorpf, J. B. (2007). *Psychologie der Persönlichkeit* (4., überarb. u. aktual. Aufl.). Heidelberg: Springer Medizin.
- Asendorpf, J. B. (2011). *Persönlichkeitspsychologie* (2., überarb. u. aktual. Aufl.). Heidelberg: Springer Medizin.
- Ashton, M., Perugini, M., de Vries, R., Boies, K., Lee, K., Szarota, P. (2004). A Six-Factor Structure of Personality-Descriptive Adjectives: Solutions From Psycholexical Studies in Seven Languages. *Journal of Personality and Social Psychology*, 2, 356-366.
- Bartussek, D. (1996). Faktorenanalytische Gesamtsysteme der Persönlichkeit. In M. Amelang (Hrsg.), *Temperaments- und Persönlichkeitsunterschiede* (S. 51-105). Göttingen: Hogrefe.
- Becker, P. (1999). Beyond the Big Five. *Personality and Individual Differences*, 26, 511-530.
- Borkenau, P. & Ostendorf, F. (2008). *NEO-Fünf-Faktoren-Inventar nach Costa und McCrae* (2., neu normierte u. vollst.überarb. Aufl.). Göttingen: Hogrefe.

- Borkenau, P., Friedel, A. & Wolfradt, U. (2011). Standardisierte Persönlichkeitsfragebögen. In L. F. Hornke, M. Amelang & M. Kersting (Hrsg.), *Persönlichkeitsdiagnostik* (S. 1-72). Göttingen: Hogrefe.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler* (6., vollst. überarb. u. aktual. Aufl.). Heidelberg: Springer Medizin.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4., überarb. Aufl.). Heidelberg: Springer Medizin.
- Brähler, E., Holling, H., Leutner, D. & Petermann, F. (Hrsg.). (2002). *Brickenkamp Handbuch psychologischer und pädagogischer Tests Bd. 1* (3., vollst. überarb. u. erw. Aufl.). Göttingen: Hogrefe.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Costa, P. T. & McCrae, R. R. (1992). *Revised NEO Personality Inventory and NEO Five Factor Inventory Professional Manual*. Odessa: Psychological Assessment Resources.
- Duden. (2007). *Herkunftswörterbuch Bd. 7* (4., neu bearb. Aufl.). Mannheim: Dudenverlag.
- Fisseni, H.-J. (1998). *Persönlichkeitspsychologie. Ein Theorienüberblick* (4. überarb. u. erw. Aufl.). Göttingen: Hogrefe.
- Fisseni, H.-J. (2004). *Lerhbuch der psychologischen Diagnostik* (3., überarb.u.erw. Aufl.). Göttingen: Hogrefe.
- Guilford, J. P. (1964). *Persönlichkeit*. Weinheim: Beltz.
- Guilford, J. P., Guilford, J. S. & Zimmermann, W. S. (1978). *The Guilford-Zimmermann Temperament Survey*. Sheridan: Psychological Services.
- Guilford, J. S., Zimmermann, W. S. & Guilford, J. P. (1976). *The Guilford-Zimmermann Temperament Survey Handbook*. San Diego: Edits Publishers.
- Gulliksen, H. (1950). *Theory of Mental tests*. New York: Wiley.
- Häcker, H. O. & Stapf, K.-H. (Hrsg.). (2009). *Dorsch Psychologisches Wörterbuch* (15., überarb. u. erw. Aufl.). Bern: Huber.

- Herrmann, T. (1976). *Lehrbuch der empirischen Persönlichkeitsforschung* (3., neubearb. Aufl.). Göttingen: Hogrefe.
- Jankisz, E. & Moosbrugger, H. (2007). Planung und Entwicklung von psychologischen Tests und Fragebogen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 27-72). Heidelberg: Springer.
- Krohne, H. W. & Hock, M. (2007). *Psychologische Diagnostik Grundlagen und Anwendungsfelder*. Stuttgart: Kohlhammer.
- Kubinger, K. (2009). *Psychologische Diagnostik* (2., überarb. u. erw. Aufl.). Göttingen: Hogrefe.
- Kubinger, K. & Jäger, R. (Hrsg.). (2003). *Schlüsselbegriffe der Psychologischen Diagnostik*. Weinheim: Beltz.
- Laux, L. (2008). *Persönlichkeitspsychologie* (2. überarb. u. erw. Aufl.). Stuttgart: Kohlhammer.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Aufl.). Weinheim: Psychologie Verlags Union.
- Moosbrugger, H. (2007). Klassische Testtheorie. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 99-112). Heidelberg: Springer Medizin.
- Netter, P. (2005). Eigenschaften. In H. Weber, T. Rammsayer & (Hrsg.), *Handbuch der Persönlichkeitspsychologie und Differentiellen Psychologie* (S. 231-243). Göttingen: Hogrefe.
- Pervin, L. A., Cervone, D. & John, O. P. (2005). *Persönlichkeitstheorien* (5., vollst. überarb. u. erw. Aufl.). München: Ernst Reinhardt.
- Petermann, F., Holling, H., Leutner, D. & Brähler, E. (2002). Zur Konzeption des Handbuchs. In E. Brähler, H. Holling, D. Leutner & F. Petermann (Hrsg.), *Brickenkamp Handbuch psychologischer und pädagogischer Tests Bd. 1* (3., vollst. überarb. u. erw. Aufl., S. XI-XXV). Göttingen: Hogrefe.
- Pospeschill, M. (2010). *Testtheorie, Testkonstruktion, Testevaluation*. München: Ernst Reinhardt.

- Rauchfleisch, U. (2008). *Testpsychologie. Eine Einführung in die Psychodiagnostik* (5. Aufl.). Göttingen: Vandenhoeck & Ruprecht.
- Rushton, P. & Irwing, P. (2009). A General Factor of Personality in 16 sets of the Big Five, the Guilford-Zimmermann Temperament Survey, the California Psychological Inventory and the Temperament and Character Inventory. *Personality and Individual Differences*, 47, 558-564.
- Schmitt, M. & Altstötter-Gleich, C. (2010). *Differentielle Psychologie und Persönlichkeitspsychologie*. Weinheim: Beltz.
- Sedlmeier, P. & Renkewitz, F. (2008). *Forschungsmethoden und Statistik in der Psychologie*. München: Pearson Studium.
- Stemmler, G., Hagemann, D., Amelang, M. & Bartussek, D. (2011). *Differentielle Psychologie und Persönlichkeitsforschung* (7., vollst. überarb. Aufl.). Stuttgart: Kohlhammer.
- Zimbardo, P. G. & Gerrig, R. J. (2008). *Psychologie* (18., aktualis. Aufl.). München: Pearson Studium.

Anhang

Anhang A	Vollständige deskriptive Statistik demographische Variablen
Anhang B	Q-Q-Diagramme Skalenscores GZTS und NEO-FFI
Anhang C	Streudiagramme weiterer Korrelationen Skalen GZTS und NEO-FFI
Anhang D	Vollständige Korrelationsmatrix nach Pearson Skalen GZTS und NEO-FFI
Anhang E	Vollständige Korrelationsmatrix Partialkorrelationen Skalen GZTS und NEO-FFI
Anhang F	Vollständige Auflistung der deutschen GZTS-Items
Anhang G	Skalenstruktur und Itempolung der deutschen GZTS-Fassung

Anhang A Vollständige deskriptive Statistik demographische Variablen

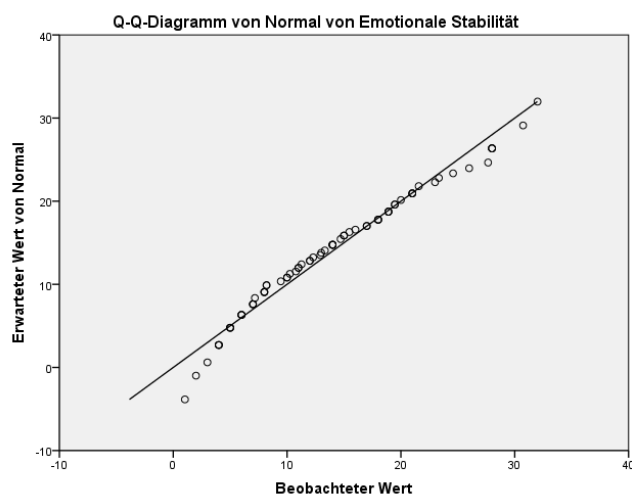
Alter		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	männlich	26	37.7	37.7	37.7
	weiblich	43	62.3	62.3	100
	Gesamt	69	100	100	

Geschlecht		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	männlich	26	37.7	37.7	37.7
	weiblich	43	62.3	62.3	100
	Gesamt	69	100	100	

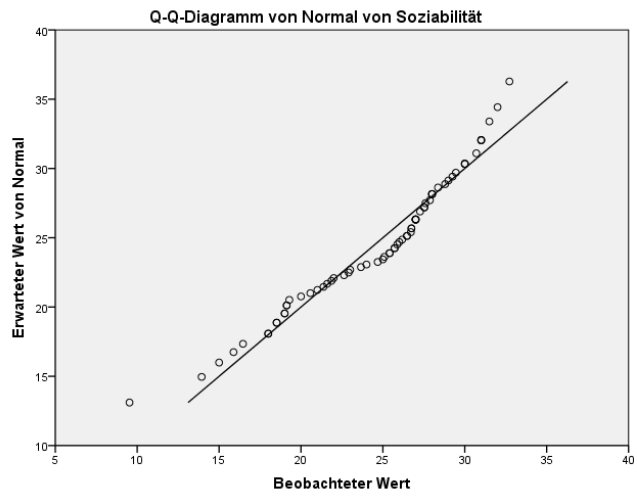
Bildungsabschluss		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Sekundarstufe II	27	39.1	39.1	39.1
	Tertiärstufe	42	60.9	60.9	100
	Gesamt	69	100	100	

Anhang B Q-Q-Diagramme Skalenscores GZTS und NEO-FFI

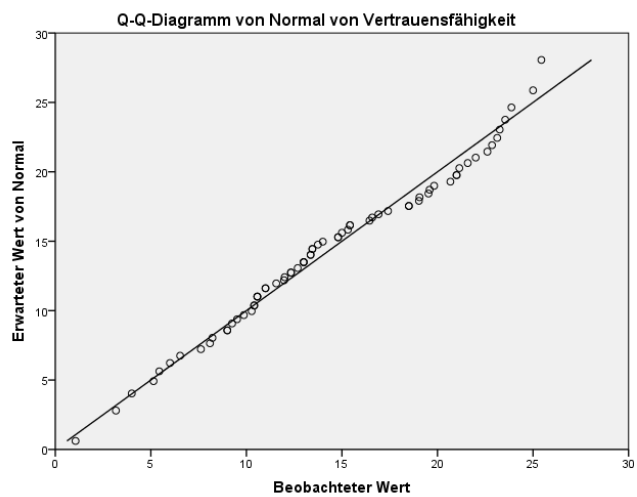
Emotionale Stabilität



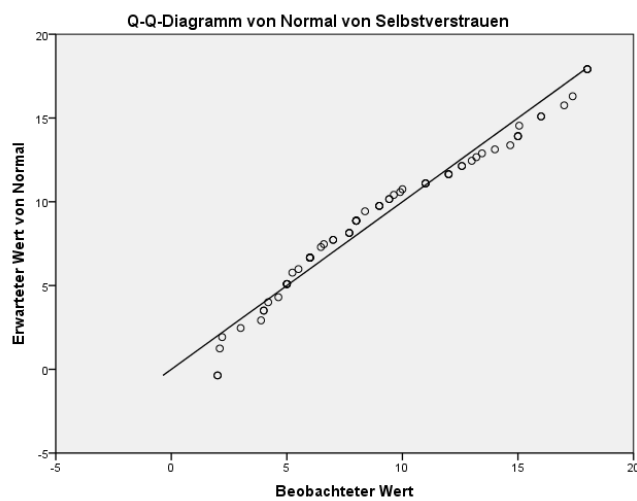
Soziabilität



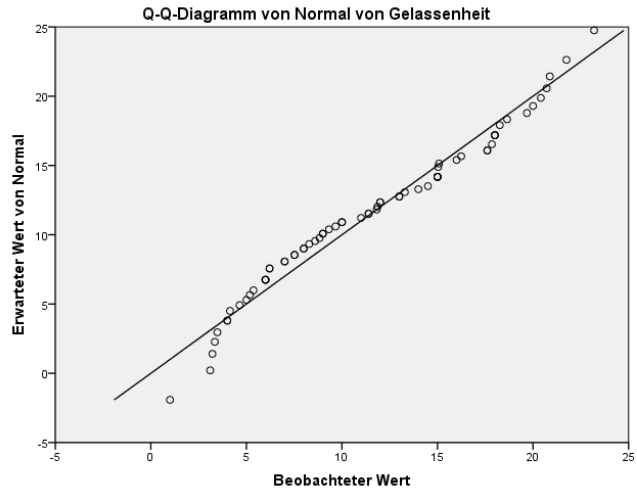
Vertrauensfähigkeit/Persönliche Beziehungen



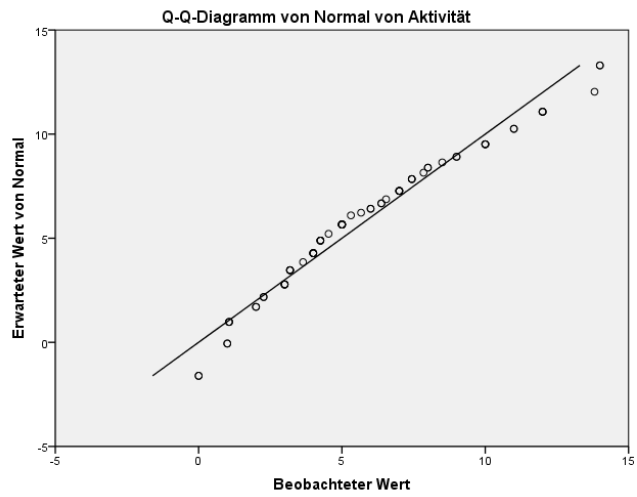
Selbstvertrauen



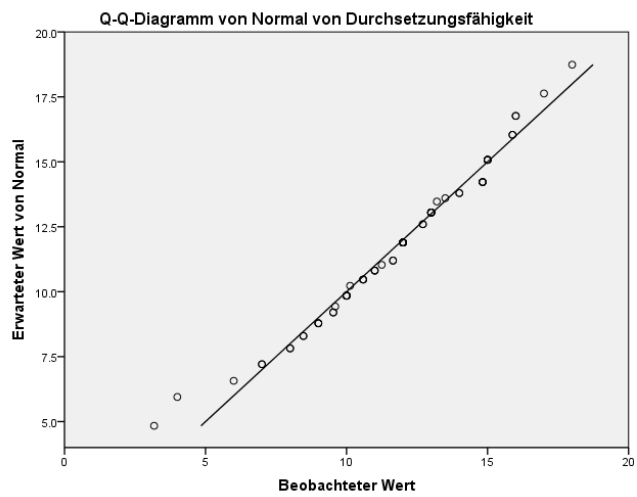
Gelassenheit



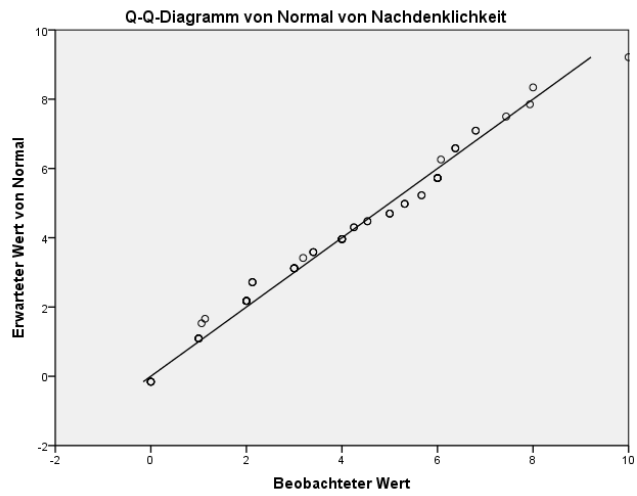
Generelle Aktivität



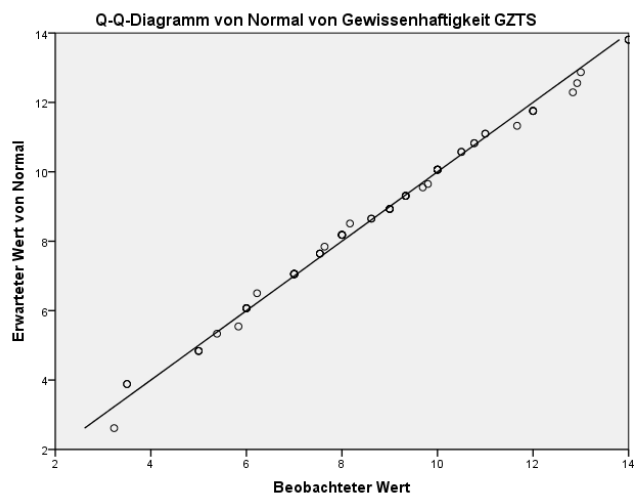
Durchsetzungsfähigkeit



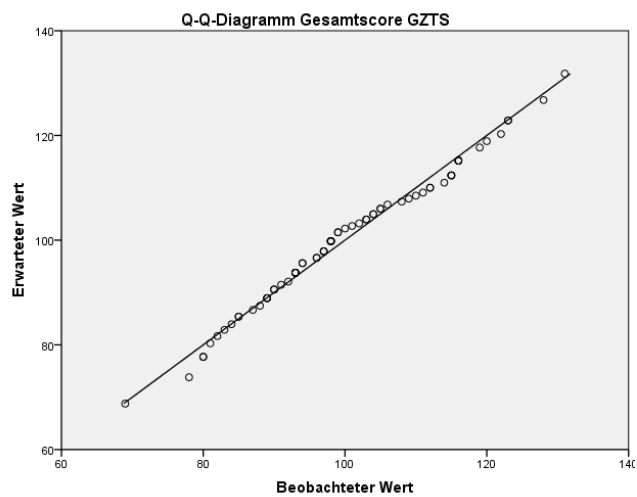
Nachdenklichkeit



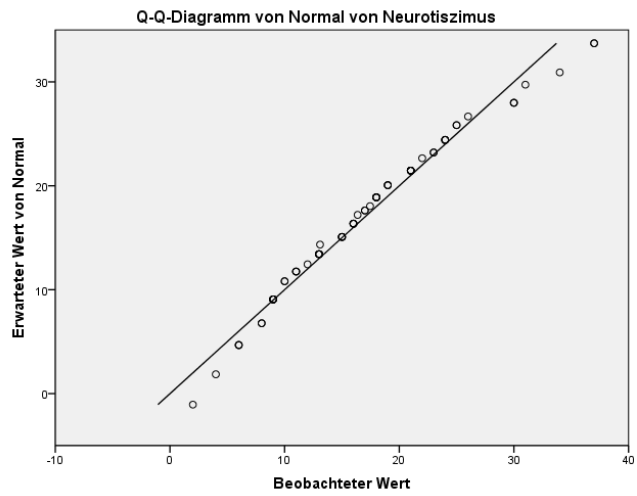
Gewissenhaftigkeit GZTS



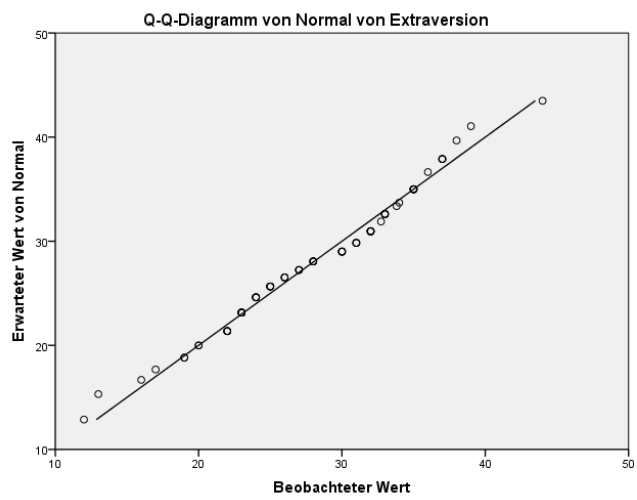
Gesamtscore GZTS



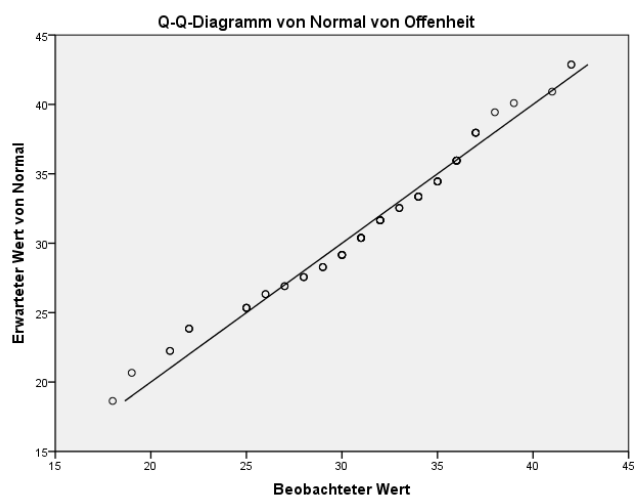
Neurotizismus



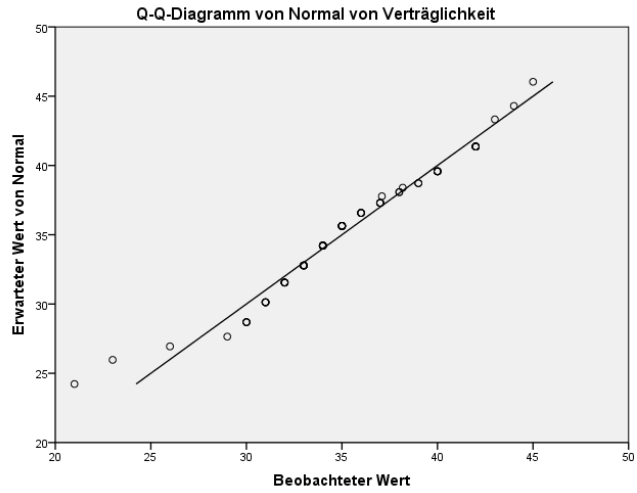
Extraversion



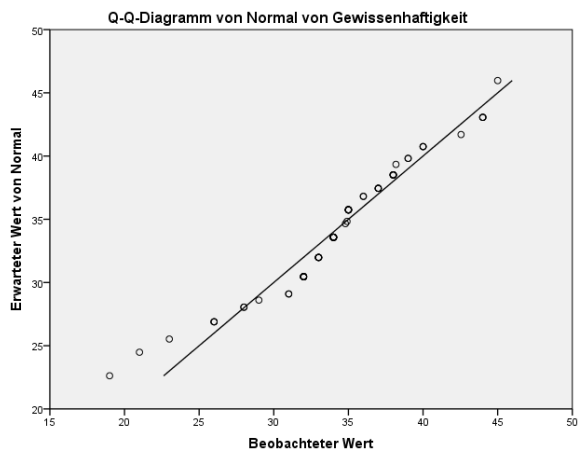
Offenheit



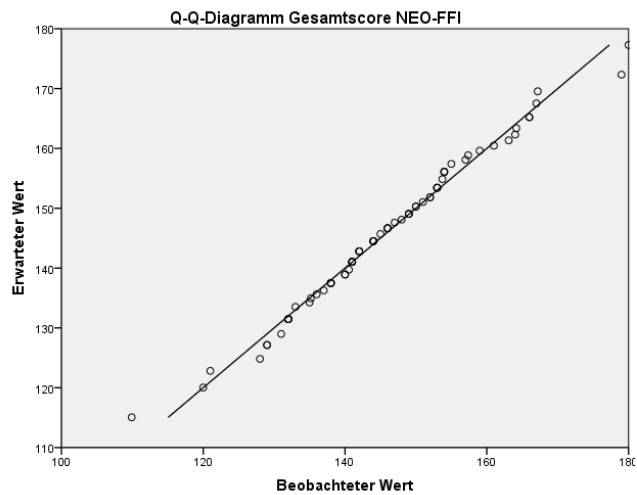
Verträglichkeit



Gewissenhaftigkeit NEO-FFI

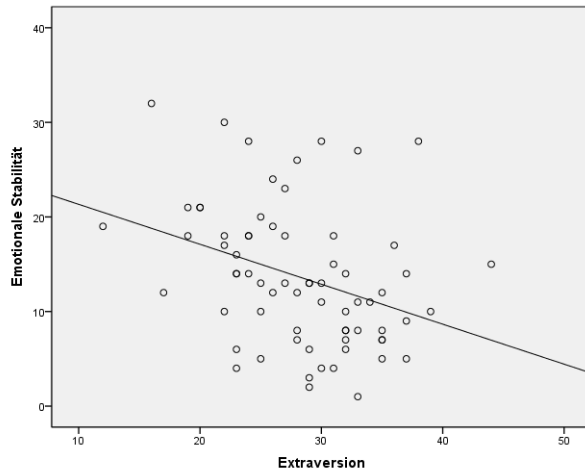


Gesamtscore NEO-FFI

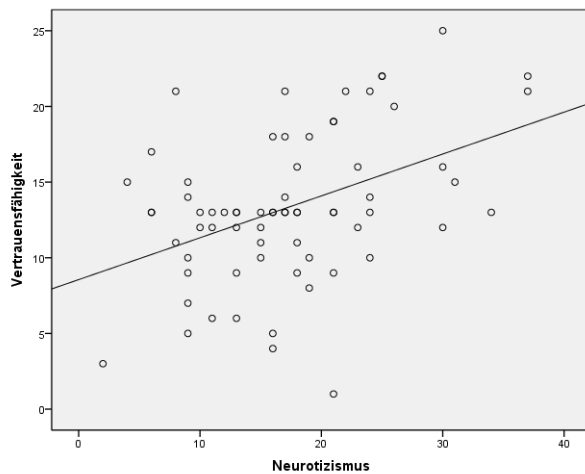


Anhang C Streudiagramme weiterer Korrelationen Skalen GZTS und NEO-FFI

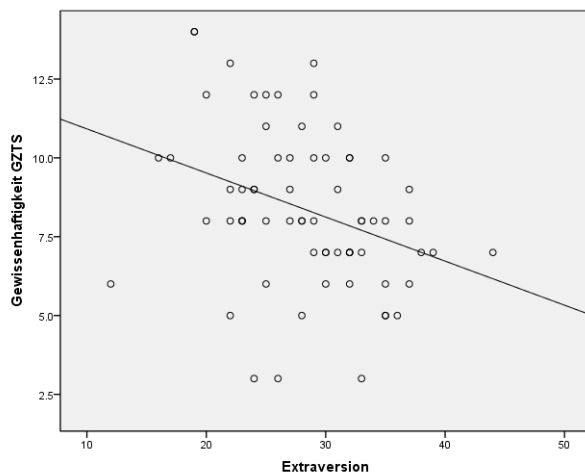
Emotionale Stabilität – Extraversion



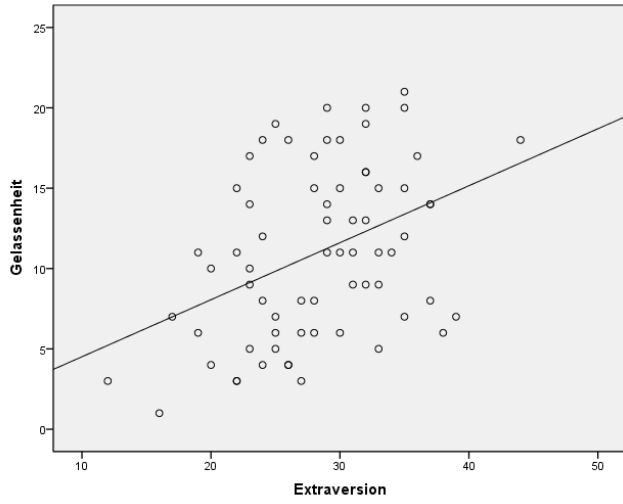
Vertrauensfähigkeit/Persönliche Beziehungen – Neurotizismus



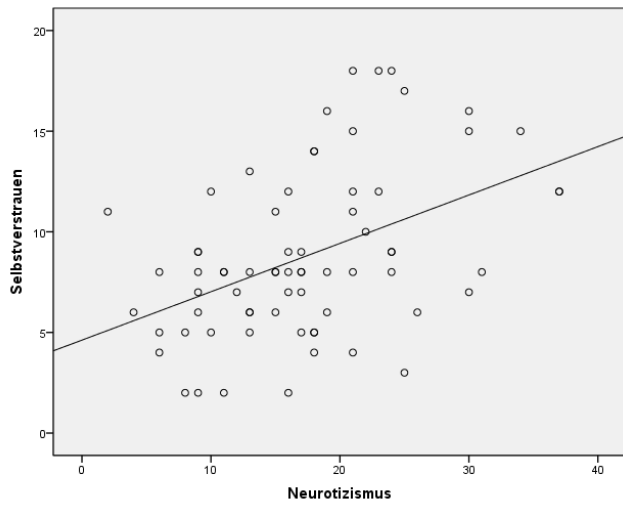
Gewissenhaftigkeit GZTS – Extraversion



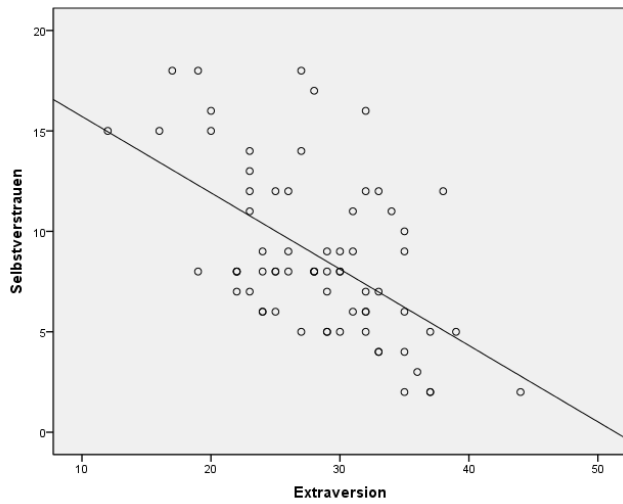
Gelassenheit – Extraversion



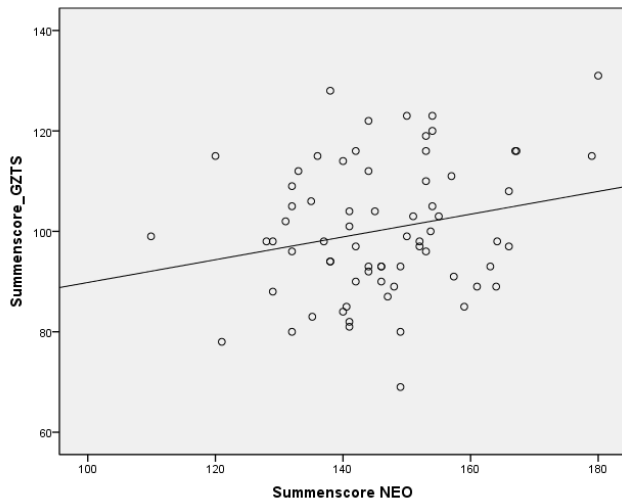
Selbstvertrauen – Neurotizismus



Selbstvertrauen – Extraversion



Gesamtscores GZTS – NEO-FFI



Anhang D Vollständige Korrelationsmatrix nach Pearson Skalen GZTS – NEO-FFI

	ES	S	V/PB	SV	Gel	GA	D	N	Gew	1	2	3	4
S	-.13												
V/PB	.51**	.14											
SV	.40**	-.42**	.02										
Gel	-.49**	.30*	-.14	-.40**									
GA	.25*	.39**	.39**	-.19	.07								
D	-.24*	.20	-.12	-.38**	.25*	.19							
N	-.41**	-.11	-.32**	.11	.34**	-.16	-.04						
Gew	.10	-.28*	.04	.15	-.07	-.31**	-.16	-.31**					
1	.71**	-.20	.43**	.45**	-.58**	.26*	-.24	-.16	-.15				
2	-.35**	.73**	.05	-.56**	.40**	.40**	.27*	.05	-.34**	-.33**			
3	.04	.20	.11	-.25*	-.01	.03	.05	-.29*	.01	-.12	.26*		
4	-.12	.17	-.27*	.07	-.01	-.22	-.16	.05	-.03	-.14	.19	.20	
5	-.02	.05	.08	-.08	-.13	.10	-.04	-.27*	.37**	.00	.10	-.04	-.10

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant

* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant

Legende NEO-FFI:

- 1 Neurotizismus
- 2 Extraversion
- 3 Offenheit
- 4 Verträglichkeit
- 5 Gewissenhaftigkeit NEO-FFI

**Anhang E Vollständige Korrelationsmatrix Partialkorrelationen Skalen GZTS
– NEO-FFI mit Kontrollvariablen Alter und Bildungsabschluss**

		ES	S	V	SV	Gel	GA	D	N	Gew
Neurotizismus	Korrelation	.71	-.25	.38	.46	-.55	.24	-.17	-.15	-.09
	Sign. (2-seitig)	.00	.04	.00	.00	.00	.05	.17	.23	.46
	Freiheitsgrade	65	65	65	65	65	65	65	65	65
Extraversion	Korrelation	-.38	.73	.07	-.52	.38	.35	.25	.02	-.28
	Sign. (2-seitig)	.00	.00	.60	.00	.00	.00	.04	.89	.02
	Freiheitsgrade	65	65	65	65	65	65	65	65	65
Offenheit	Korrelation	.04	.16	.13	-.18	-.07	-.06	.02	-.33	.11
	Sign. (2-seitig)	.77	.20	.29	.15	.56	.65	.87	.01	.37
	Freiheitsgrade	65	65	65	65	65	65	65	65	65
Verträglichkeit	Korrelation	-.12	.18	-.29	.05	.02	-.22	-.15	.06	-.05
	Sign. (2-seitig)	.32	.14	.02	.69	.89	.08	.24	.63	.69
	Freiheitsgrade	65	65	65	65	65	65	65	65	65
Gewissenhaftigkeit	Korrelation	-.02	.05	.08	-.08	-.15	.10	-.05	-.27	.37
	Sign. (2-seitig)	.90	.66	.52	.53	.23	.43	.71	.03	.00
	Freiheitsgrade	65	65	65	65	65	65	65	65	65

Anhang F Vollständige Auflistung der deutschen GZTS-Items

1. Sie beginnen neue Projekte mit grosser Begeisterung.
2. Sie würden lieber eine Aktivität planen als daran teilzunehmen
3. Sie haben schon öfters die Verantwortung für die Organisation eines Projektes oder einer Gruppe übernommen.
4. Sie unterhalten gerne Gäste.
5. Ihre Interessen ändern sich oft.
6. Sie leben nach dem Motto: „Essen, Trinken und fröhlich sein, denn morgen könnte schon alles vorbei sein.“
7. Wenn Sie merken, dass ein neu gekauftes Produkt defekt ist, zögern Sie normalerweise nach einer Rückerstattung oder einem Ersatz zu fragen.
8. Es fällt Ihnen leicht, neue Bekanntschaften zu schliessen.
9. Manchmal sprühen Sie vor Energie und manchmal sind Sie sehr träge.
10. Sie handeln gerne schnell.
11. Andere nehmen Sie als seriöse, ersthafte Person wahr.
12. Wenn Sie jemand neues begegnen, warten Sie bis sich er/sie sich vorstellt.
13. Sie nehmen gerne an sozialen Aktivitäten teil.
14. Manchmal fühlen Sie sich grundlos unglücklich oder elend.
15. Sie sind oft so gehetzt, dass Sie früher oder später ausgelaugt sein könnten.
16. Sie gehen gerne auf Parties.
17. Wenn sich Ihre Meinung grundlegend von der eines Vortragenden unterscheidet, tendieren Sie dazu, es der Person während oder nach dem Vortrag mitzuteilen.
18. Sie denken über vergangene Misserfolge nicht lange nach.
19. Sie wundern sich oft, woher andere ihre überschüssige Energie haben.

20. Sie tendieren dazu, Dinge zu überdenken bevor Sie handeln.
21. Sie vermeiden es, mit Verkaufsstellen über Preise zu verhandeln.
22. Sie arbeiten ungern alleine an einem isolierten Arbeitsplatz.
23. Es fällt Ihnen oft schwer, nachts einzuschlafen, weil Sie ständig über die Ereignisse des Tages grübeln.
24. Sie sind oft in Eile, selbst wenn genügend Zeit vorhanden ist.
25. Sie bevorzugen Arbeiten, die Aufmerksamkeit für das Detail erfordern.
26. Sie sind froh, wenn andere die Führung in Gruppenaktivitäten übernehmen.
27. Sie genießen es, neue Leute kennen zu lernen.
28. Es muss viel passieren, um Sie aus der Ruhe zu bringen.
29. Sie arbeiten langsamer und sorgfältiger als die meisten Ihres Alters.
30. Sie sind eine unbekümmerte Person.
31. Wenn sich jemand unfair verhält, zögern Sie, ihn/sie darauf hinzuweisen.
32. Es stört Sie, wenn andere Sie bei Ihrer Arbeit beobachten.
33. Sie blicken optimistisch in die Zukunft.
34. Es ist Ihnen wichtig, genug Zeit zu haben, um zu entspannen.
35. Sie nehmen das Leben sehr ernst.
36. Sie gehen gerne zu Vorstellungsgesprächen.
37. Sie würden gerne Gastgeber oder Gastgeberin für Parties in einem Club sein.
38. Sie sind eine griesgrämige, mürrische Person.
39. Sie sind die Art von Person, die immer „auf dem Sprung“ ist.
40. Sie sehnen sich oft nach Aufregung.
41. Der Gedanke, vor anderen eine Rede zu halten, erzeugt bei Ihnen Unbehagen.
42. Es fällt Ihnen leicht, ein Gespräch mit Fremden zu beginnen.
43. Sie haben oft grundlos ein schlechtes Gewissen.
44. Sie werden voller Energie wahrgenommen.
45. Manchmal fällen Sie spontane Entscheidungen, die Sie später bereuen.
46. Sie sind von Natur aus freundlich, sodass sich andere schnell in Ihrer Gegenwart wohl fühlen.
47. Sie sind oft verträumt.
48. Sie handeln jeweils schnell.
49. Sie haben die Angewohnheit, etwas zu beginnen und dann schnell das Interesse an der Sache wieder zu verlieren.
50. Sie vermeiden es manchmal, mit anderen in Kontakt zu treten, aus Angst, etwas Falsches zu tun oder zu sagen.
51. Sie erleben oft grundlose Stimmungsschwankungen mit Höhen und Tiefen.
52. Sie scheinen immer viel Elan und Vitalität zu besitzen.
53. Leute, die sich über Dinge Sorgen machen, sind für Sie schwierig zu verstehen.
54. Wenn ein Angestellter in einem Geschäft andere zuerst bedient, die nach Ihnen an der Reihe wären, machen Sie ihn darauf aufmerksam.
55. Ohne zahlreiche soziale Kontakte wären Sie sehr unglücklich.
56. Manchmal sieht Ihre Zukunft sehr düster aus.
57. Manchmal wünschen Sie sich, dass andere ein wenig langsamer wären, um Ihnen die Möglichkeit zu geben, aufzuholen.

58. Viele Ihrer Freunde denken, dass Sie Ihre Arbeit viel zu ernst nehmen.
59. Sie zögern, in ein Meeting zu gehen, wenn Sie wissen, dass alle Augen auf Sie gerichtet sein werden.
60. Sie fühlen sich fast immer gesund und stark.
61. Es scheint so, als fehle Ihnen manchmal der nötige Antrieb, um gleich viel wie andere zu erledigen.
62. Sie treffen spontane Entscheidungen.
63. Sie können anderen gut etwas vormachen, wenn Sie sich in Schwierigkeiten befinden.
64. Wenn Sie jemandem vorgestellt werden, fällt es Ihnen schwer, Gesprächsstoff für eine gute Konversation zu finden.
65. Sie fühlen sich einsam, auch wenn Sie unter Leuten sind.
66. Sie sind in der Lage, aussergewöhnlich viele Stunden zu arbeiten, ohne sich müde zu fühlen.
67. Bei einem Unfall leisten Sie aktiv Hilfe.
68. Es fällt Ihnen schwer, neue Freunde zu finden.
69. Ihre Stimmung wechselt manchmal von fröhlich zu traurig oder umgekehrt, ohne zu wissen warum.
70. Sie ergreifen jeweils die Initiative, um eine langweilige Party lebendiger zu gestalten.
71. Sie nehmen gerne am Vereinsleben teil.
72. Manchmal denken Sie sehr langsam und manchmal sehr schnell.
73. Sie erledigen Ihre Aufgaben gerne langsam und bedacht.
74. Sie sind eine sorglose, unbekümmerte Person.
75. Sie beschwerten sich nicht, wenn Sie in einem Restaurant minderwertiges Essen serviert bekommen.
76. Um sich für eine Stelle zu bewerben, würden Sie lieber einen Brief schreiben als sich persönlich vorzustellen.
77. Sie sind oft niedergeschlagen.
78. Sie neigen dazu, von einer Arbeit zur nächsten zu eilen, ohne sich richtig auszu-ruhen.
79. Sie sind so besorgt um die Zukunft, dass Sie die Gegenwart nicht richtig genießen können.
80. Wenn Sie sich zu jemandem hingezogen fühlen, den Sie nicht kennen, versuchen Sie trotzdem, ihn/sie kennen zu lernen, auch wenn es schwierig sein könnte.
81. Sie neigen dazu, Ihre Bekanntschaften auf ein paar wenige zu beschränken.
82. Sie denken selten näher über Ihre vergangenen Fehler nach.
83. Sie sind weniger energiegeladener als viele andere, die Sie kennen.
84. Sie analysieren Ihre Gedanken und Gefühle nicht im Detail.
85. Wenn Sie in Sitzungen anderer Meinung als jemand anders sind, dann teilen Sie dies mit.
86. Ihre Schüchternheit stört Sie.
87. Es stört Sie manchmal, dass ein sinnloser Gedanke Ihnen immer wieder durch den Kopf geht.

88. Sie erledigen ihre Aufgaben jeweils in Eile.
89. Sie sind ein Verkaufstalent.
90. Sie sind oft der „Stimmungsmacher“ an einer Party.
91. Bei der Arbeit oder beim Spielen ist es für andere schwierig, mit Ihnen mitzuhalten.
92. Sie können einem Vortrag lange zuhören, ohne müde zu werden.
93. Sie können sich besser mündlich als schriftlich ausdrücken.
94. Ihre Stimmung ist in der Regel einigermaßen gleichbleibend.
95. Sie mögen es nicht, bei der Arbeit gehetzt zu werden.
96. Es fällt Ihnen schwer „Nein“ zu sagen, wenn Ihnen jemand unbedingt etwas verkaufen möchte, das Sie nicht wirklich wollen.
97. Sie haben wenige Freunde, mit denen Sie sich wirklich entspannen und es sich gut gehen lassen können.
98. Sie bleiben auch bei Problemen stets fröhlich.
99. Andere sagen Ihnen Sie sollen gelassener werden.
100. Sie gehören zu den Menschen, die mehr rauchen oder trinken als Sie sollten.
101. Wenn sie in der Öffentlichkeit Leuten begegnen, bei denen Sie das Gefühl haben, Sie schon einmal getroffen zu haben, sprechen Sie sie darauf an.
102. Sie bevorzugen es, alleine zu arbeiten.
103. Enttäuschungen kümmern Sie so wenig, dass Sie nicht zweimal darüber nachdenken.
104. Sie bewegen sich langsam und bedacht.
105. Sie sind in Gegenwart wichtiger Leute gehemmt.
106. Andere nehmen Sie als sehr soziale Person wahr.
107. Weil Sie sich viele Sorgen machen, haben Sie schon öfters schlaflose Nächte verbracht.
108. Sie arbeiten so lange an einer Aufgabe, bis sie abgeschlossen ist.
109. Ihnen fallen gute Ausreden ein, wenn Sie welche brauchen.
110. Ihre Tagträume sind oft über Dinge, die niemals wahr werden können.
111. Sie rennen oft die Treppe hinauf und nehmen zwei Stufen auf einmal.
112. *Ihre Pflichten halten Sie selten davon ab, es sich gut gehen zu lassen.*
113. Sie mögen es, eine grosse Verantwortung auf sich zu nehmen, wie zum Beispiel ein neues Projekt zu organisieren.
114. Es ist schon oft vorgekommen, dass Sie aufgrund Ihrer Schüchternheit gezögert haben, Verabredungen anzunehmen oder vorzuschlagen.
115. Sie lassen sich in ihrer Stimmung leicht von Ihrem Umfeld beeinflussen.
116. Andere sind oft von der grossen Menge an Arbeit, die Sie leisten können, beeindruckt.
117. Sie fühlen sich im Allgemeinen sorgenfrei.
118. Es fällt Ihnen schwer, einen Verkäufer loszuwerden, dessen Anliegen Sie nicht interessiert.
119. In Gesprächen sind Sie eher derjenige der zuhört, als derjenige der spricht.
120. Sie denken oft, dass das Leben lebenswert ist.

121. Es ärgert Sie, wenn Sie an einer Kreuzung warten müssen, weil ein langer Güterzug vorbei fahren muss.
122. Sie sagen oft, was Sie gerade denken.
123. Sie halten gerne Reden.
124. Sie sind gerne mit anderen Leuten zusammen.
125. Sie behalten in aufregenden Situationen einen kühlen Kopf.
126. Andere nehmen Sie als lebhaft Person wahr.
127. Sie fühlen sich besser, wenn Sie Ihre Wut herauslassen können.
128. Sie versuchen möglichst jeden Ärger mit anderen zu vermeiden.
129. Andere scheinen Ihre Gesellschaft zu genießen.
130. Sie fühlen sich manchmal grundlos müde und lustlos.
131. Es ist schwierig zu verstehen, wieso manche Leute so langsam sind und so wenig erledigt bekommen.
132. Sie wetten gerne, auch wenn Sie es sich nicht leisten können.
133. Wenn jemand Unwahrheiten über Sie erzählt, sprechen Sie die Person so bald wie möglich darauf an.
134. Ihre Schüchternheit hält Sie davon ab, so beliebt zu sein, wie Sie eigentlich möchten.
135. Sie sorgen sich in der Regel nicht um mögliche Missgeschicke.
136. Sie werden in der Regel für Ihre Taten vollumfänglich belohnt.
137. Sie würden gerne bestimmten Leuten die Meinung sagen.
138. Sie würden eher einen Abend zu Hause mit einem Buch verbringen als auf eine Party zu gehen.
139. Sie würden vieles an der Natur des Menschen ändern, wenn Sie könnten.
140. In Gruppenaktivitäten kommen Sie nicht zu kurz.
141. In den meisten Fällen ist es wichtig, dass Sie bekommen was Sie wollen, auch wenn Sie hart dafür kämpfen müssen.
142. Die meisten Angestellten eines öffentlichen Amtes stellen das öffentliche Interesse über ihr eigenes Interesse.
143. Andere sprechen hinter Ihrem Rücken über Sie.
144. Geld ist wichtig, weil es Macht verleiht.
145. Es fällt Ihnen leicht, sich natürlich zu verhalten, wo auch immer Sie sind.
146. Die meisten Leute sind dumm.
147. Andere Leute schreiben Ihnen oft die Schuld für Dinge zu, für die Sie nicht verantwortlich sind.
148. Sie mögen es nicht, in einem Wettbewerb zu verlieren.
149. Sie mögen Aufgaben, die viel Aufmerksamkeit für das Detail erfordern.
150. Sie langweilen sich oft, wenn das Thema in Gesprächen nichts mit ihren Hobbies, Erfahrungen oder Interessen zu tun hat.
151. Die meisten Leute wissen was zu tun ist, ohne dass es ihnen gesagt wird.
152. Sie sind in bestimmten Dingen heikel.
153. Es gibt jemanden, den Sie gerne in die Schranken weisen würden.
154. Sie sind mehr an Sport interessiert als an intellektuelle Themen.

155. Sie identifizieren sich oft mit einer der Hauptcharaktere in Filmen oder Theaterstücken.
156. Sie zögern, anderen zu sagen, sie sollen sich um ihre eigenen Angelegenheiten kümmern.
157. Sie haben es sich zur Regel gemacht, Ihre vergangenen Taten kritisch zu analysieren.
158. Die meisten Orte müssten ihre Verkehrsregeln verbessern.
159. Sie geraten in Streitereien oder Diskussionen, die Sie eigentlich nicht verursachen wollten.
160. Sie nehmen es Freunden oder Familienmitgliedern übel, wenn sie Ihnen Befehle erteilen.
161. Sie tendieren dazu, über Dinge zu philosophieren.
162. Die meisten Leute sind nur ehrlich, aus Angst davor erwischt zu werden.
163. Sie neigen dazu, viel über sich nachzudenken.
164. Sie hätten gerne oft neugierigen Leuten gesagt, sie sollen sich um ihre eigenen Angelegenheiten kümmern.
165. Sie sind oft in Gedanken versunken.
166. Zu viele Leute versuchen, so viel wie möglich zu nehmen und so wenig wie möglich zurückzugeben.
167. Sie überwinden beschämende und erniedrigende Erfahrungen sehr schnell.
168. Die meisten Leute versuchen bei der Arbeit ihr Geld wert zu sein.
169. Fast alles was passiert, scheint irgendwie mit Ihnen zusammenzuhängen.
170. Sie möchten oft die tieferliegenden Gründe für die Handlungen anderer erfahren.
171. Manche Leute tun oder sagen Dinge, um Sie absichtlich zu verärgern.
172. Einige Leute können so unhöflich werden, dass Sie sie am liebsten in die Schranken weisen möchten.
173. Manchmal haben Sie das eigenartige Gefühl, nicht mehr sie selbst zu sein.
174. Mit anspruchlosen Aufgaben setzen Sie sich nicht auseinander.
175. Sie mögen es, über komplexe Problemstellungen nachzudenken.
176. Die meisten Leute lernen schnell, damit sie nicht denselben Fehler zweimal machen.
177. Andere verletzen Sie ohne es zu merken, weil Sie jeweils Ihre Gefühle verbergen.
178. Es ist für Sie befriedigend, wenn Sie andere dazu bringen können, das zu tun, was Sie möchten.
179. Sie nehmen sich oft die Zeit, um lange über Dinge nachzudenken.
180. Sie haben alle Belohnungen im Leben bekommen, die Sie verdienen.
181. Sie haben das Gefühl, dass einige Leute insgeheim versuchen, Sie auszustechen.
182. Sie finden es spannend, andere zu beobachten.
183. Sie haben bis jetzt wenige falsche Leute getroffen.
184. Sie fragen andere sehr oft um Rat.
185. Sie versuchen ständig, sich zu verbessern.
186. Die meisten Leute folgen blind einem Anführer.
187. Es ist allgemein schwierig, Sie zu kränken.

188. Wenn jemand versuchen würde, Ihnen ein gesetzmässiges Recht abzuerkennen, wird er/sie es mit Ihnen zu tun bekommen.
189. Sie meiden eher komplizierte Probleme, die eine Denkleistung erfordern.
190. Es gab Zeiten, da hatten Sie das unangenehme Gefühl, andere könnten Ihre Gedanken lesen.
191. Anstatt einen Streit zu beginnen, ist es besser, „die andere Wange auch hinzuhalten“.
192. Sie versuchen zu errahnen, was andere denken, wenn sie mit Ihnen sprechen.
193. Sie haben Ihre Glückssträhne schon hinter sich.
194. Sie wurden schon mehr als einmal schwerstens beleidigt.
195. Wenn Sie die Handlungen anderer missbilligen, sagen Sie es ihnen gleich.
196. Nach einer kritischen Situation, denken Sie in der Regel über etwas nach, was Sie hätten tun sollen, aber nicht getan haben.
197. Wenn eine Sache richtig gemacht werden soll, müssen Sie sie selbst in die Hand nehmen.
198. Sie sind über die Moral Ihrer Generation besorgt.
199. Grosse Unternehmen sind eine gute Sache.
200. Wenn etwas schief läuft, beunruhigt es Sie nur leicht.
201. Sie sorgen dafür, dass andere Sie nicht ausnützen.
202. Sie neigen dazu, über Ihre Vergangenheit zu grübeln.
203. Einige Leute stecken Ihre Nase mehr in Angelegenheiten, als Ihnen recht ist.
204. Sie haben oft das Gefühl, dass jemand über Sie persönlich spricht.
205. Es stört Sie wenn in einer Organisation die Dinge anders laufen, als Sie es sich vorstellen.
206. Sie analysieren gerne Ihre Gedanken und Gefühle.
207. Die meisten Leute verschleiern Ihren Wettbewerbsgeist mit Höflichkeit.
208. Es gibt Tage, da läuft bei Ihnen alles schief.
209. Sie verspüren den Drang, etwas Begeisterung in die Runde zu bringen, wenn es langweilig wird.
210. Sie setzen Pläne lieber in die Tat um, als die Zeit mit der Erarbeitung der Pläne zu verwenden.
211. Das Bildungssystem in diesem Land ist im Allgemeinen gut.
212. Sie belasten oft andere mit Ihren Sorgen.
213. Sie hätten gerne genug Geld und Macht, um diejenigen zu beeindrucken, die denken, sie seien besser als Sie.
214. Sie befinden sich häufig in einer Vermittlerrolle.
215. Sie sind empfindsam.
216. Sie mussten schon oft für das kämpfen, was Ihrer Meinung nach das Richtige ist.
217. Die meisten Leute möchten mehr bekommen, als sie geben.
218. Einige Leute haben schon zu Unrecht kritisch über Sie gesprochen.
219. Die Meinungen der meisten Leute sind wertlos.
220. Sie neigen dazu, introspektiv zu sein, das heisst, sich selbst zu analysieren.
221. Sie machen sich oft Gedanken über den Sinn des Lebens.

222. Manche Leute machen Ihnen absichtlich das Leben schwer.
223. Kritik stört Sie nur wenig.
224. Es ärgert Sie, wenn Sie sehen, wie jemand in einem Job pfuscht, von dem Sie wissen, wie er perfekt erledigt werden könnte.
225. Sie neigen dazu, in der Gegenwart zu leben und sich keine Gedanken über die Vergangenheit oder die Zukunft zu machen.
226. Um voranzukommen, werden die meisten Leute ab und zu einer Notlüge greifen.
227. Wenn Sie etwas verlieren, haben Sie oft den Verdacht, dass es jemand anders entweder gestohlen oder verlegt haben könnte.
228. Manche Handlungen anderer Leute scheinen Sie immer wieder zu verärgern.
229. Es gibt im Ganzen zu viele Angestellte, die einen höheren Lohn verdient hätten als Ihre Vorgesetzten.
230. Es gibt Zeiten, da scheint jeder gegen Sie zu sein.
231. Wenn jemand beim Anstehen vordrängelt, sprechen Sie ihn/sie darauf an.
232. Sie wundern sich oft, warum sich manche Leute auf eine gewisse Weise verhalten.
233. Fast alle Leute versuchen das Richtige zu tun, wenn sie eine Chance dazu bekommen.

Item 112 unter Einhaltung der ursprünglichen Nummerierung bei Skalenstrukturierung gelöscht, sodass gesamthaft 232 Items resultieren

Anhang G Skalenstruktur und Itempolung des deutschen GZTS

ES	S	V/PB	SV	Gel	GA	D	N	Gew
5	2	137	1	6	10	7	22	11
9	4	139	3	16	15	21	38	20
14	8	141	12	18	24	31	84	25
19	13	142	17	28	34	54	129	29
23	27	143	26	30	39	75	154	66
40	33	144	32	35	44	85	157	73
43	42	147	36	37	48	96	170	104
45	46	148	41	53	78	118	174	108
47	52	153	50	58	88	122	182	119
49	55	155	59	63	91	127	189	125
51	62	158	64	74	95	128	190	138
56	67	159	76	82	116	133	191	149
57	68	162	86	90	122	156	192	175
60	70	164	89	98	131	188	193	185
61	71	166	92	99	150	195	205	
65	80	171	105	103	169	201	206	
69	81	172	113	109	197	210	220	
72	93	178	114	117		231		
77	97	180	123	135				
79	101	181	134	136				
83	102	183	140	152				
87	106	186	145	160				
94	111	194		167				
100	120	198		187				
107	124	203		199				
110	126	204		200				
115	146	207		215				
130	151	213		223				
132	168	216		225				
161	176	217						
163	184	218						
165	209	222						
173	211	226						
177	214	228						
179	219	229						
196	233	230						
202		232						
208								
212								
221								
224								
227								

Negativ gepolte Items

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig und ohne Benützung anderer als der angegebenen Hilfsmittel verfasst habe.

Unterschrift: