

Universität Koblenz-Landau

Fachbereich 8: Psychologie

Diplomarbeit

**Konstrukt- und Kriteriumsvalidität
der schriftlichen Weisheitstestverfahren
3d-WS, SAWS, PWS, TWR und AWR**

Im Juli 2008

Marina Hoffmann

E-Mail: Hoffmann.Marina@web.de

1. Gutachter: Prof. Dr. Manfred Schmitt
2. Gutachter: Dr. Christine Altstötter-Gleich

Inhaltsverzeichnis:

Erklärung	5
Abkürzungsverzeichnis	5
1 Einleitung	8
2 Was ist Weisheit aus psychologischer Sicht?	11
2.1 Einführung in die Entstehungsgeschichte der psychologischen Weisheitsforschung.....	12
2.1.1 Entwicklungsmodelle der Persönlichkeit der Persönlichkeitspsychologie.....	12
2.1.2 Entwicklungsmodelle der Persönlichkeit der Entwicklungspsychologie	13
2.1.3 Entwicklungsmodelle der Persönlichkeit der Klinischen Psychologie.....	14
2.2 Laientheorien zu Weisheit	14
2.2.1 Weisheitsverständnis.....	14
2.2.2 Profil einer weisen Personen.....	15
2.2.3 Gelebte Weisheit.....	16
2.2.4 Ausblick auf die explizite Weisheitsforschung.....	17
2.3 Weisheitsdefinitionen	18
3 Weisheitstestverfahren	20
3.1 Leistungsmessung.....	21
3.1.1 Allgemeine Weisheit.....	22
3.1.2 Selbstbezogene Weisheit	24
3.2 Selbsteinschätzungsfragebögen	25
3.2.1 3dimensionale-Weisheitsskala (3d-WS).....	26
3.2.2 Selbstadministrative Weisheitsskala (SAWS).....	27
3.2.3 Weisheitsentwicklungsskala	29
3.2.4 Analytische und Synthetische Weisheit.....	29
3.3 Multimodale Erhebungsverfahren.....	30
3.3.1 Das Weisheitskonzept von Wink & Helson sowie Helson & Srivastava	30
3.3.1.1 Praktische Weisheitsskala (PWS).....	31
3.3.1.2 Praktisches Weisheitsrating (PWR).....	32
3.3.1.3 Transzendentes Weisheitsrating (TWR).....	32
3.3.1.4 Allgemeines Weisheitsrating (AWR).....	32
3.3.2 Das Weisheitskonzept nach Hartmann	33

3.4	Kategoriensysteme von Weisheitstestverfahren	34
3.4.1	Gegenstand.....	34
3.4.2	Orientierung.....	34
3.4.3	Modalität.....	35
3.5	Ausblick zur Validierung der schriftlichen Weisheitstests.....	36
4	Studien zur Validität der Weisheitstestverfahren.....	37
4.1	Kriteriumsvalidierung	38
4.2	Konvergente und Diskriminante Validierung	41
4.2.1	Soziodemographische Korrelate	42
4.2.1.1	Geschlecht.....	42
4.2.1.2	Alter und Lebenserfahrung	42
4.2.1.3	Scheidung	44
4.2.1.4	Berufliche Orientierung	44
4.2.2	Kognitive Korrelate	45
4.2.2.1	Intelligenz	45
4.2.2.2	Ego-Entwicklung	47
4.2.2.3	Kognitiver Denkstil	48
4.2.2.4	Weitere kognitive Konstrukte.....	48
4.2.3	Persönlichkeitskorrelate.....	49
4.2.3.1	Fünf-Faktoren-Modell der Persönlichkeit	49
4.2.3.2	Psychologisches Feingefühl.....	50
4.2.3.3	Persönliches Wachstum	51
4.2.3.4	Adaptivität	53
4.2.3.5	Soziale Erwünschtheit	53
5	Fragestellungen und Hypothesen.....	55
5.1	Voraussetzungen der Validierung.....	55
5.1.1	Formulierung der Hypothesen zur Validierung der einzelnen Weisheitstestverfahren.....	55
5.1.2	Voranalysen zu den Testverfahren.....	56
5.1.3	Voranalysen zu den Teilnehmergruppen	57
5.2	Kriteriumsvalidierung über die Bestimmung der Inneren Validität	58
5.3	Kriteriumsvalidierung über die Bestimmung der Äußeren Validität	59
5.4	Konstruktvalidierung anhand der Konvergenten Validität.....	60
5.5	Konstruktvalidierung anhand der Diskriminanten Validität	60

5.6	Weitere Analysen.....	62
6	Versuchsplanerisches und Methodisches Vorgehen	64
6.1	Stichprobenumfang	64
6.2	Stichprobenziehung.....	66
6.2.1	Rekrutierung der Nominatoren	66
6.2.2	Rekrutierung der Gruppe der Weisheitsnominierten	67
6.2.3	Rekrutierung der Kontrollgruppe.....	67
6.2.4	Rekrutierung der Gruppe der Psychotherapeuten	68
6.3	Verwendete Instrumente.....	68
6.3.1	Dreidimensionale-Weisheitsskala (3d-WS).....	69
6.3.2	Selbstadministrative Weisheitsskala (SAWS)	69
6.3.3	Praktische Weisheitsskala (PWS).....	70
6.3.4	Transzendentes Weisheitsrating (TWR)	70
6.3.5	Allgemeines Weisheitsrating (AWR)	71
6.3.6	Washington University Sentence Completion Test (WUSCT)	71
6.3.7	Psychologisches Feingefühl (PF).....	72
6.3.8	Persönliches Wachstum (PW) & Umweltkontrolle (UK).....	72
6.3.9	Allgemeine Intelligenz (HAWIE-R).....	72
6.3.10	Demographischer Fragebogen	73
6.4	Vorgehen in der Datenerhebung.....	73
6.4.1	Telefonische Kontaktaufnahme	74
6.4.2	Schriftliche Erhebung	74
6.4.3	Aufbau des Fragebogenheftes.....	75
6.4.4	Mündliche Erhebung.....	76
6.5	Methodische Umsetzung der statistischen Auswertung.....	77
6.5.1	Interne Konsistenz und Interraterreliabilität	77
6.5.2	Korrelationskoeffizienten	77
6.5.3	Mittelwertsvergleiche	79
6.6	Statistische Hypothesen.....	80
6.6.1	Statistische Hypothesen der Inneren Validität.....	80
6.6.2	Statistische Hypothesen der Äußeren Validität	81
6.6.3	Statistische Hypothesen der Konvergenten Validität	82
6.6.4	Statistische Hypothesen der Diskriminanten Validität	82
6.7	Voranalysen zu den Testverfahren	82

6.7.1	Mittelwerte der Testverfahren.....	82
6.7.2	Reliabilitätsanalysen der Testverfahren.....	84
6.7.3	Anpassung der Auswertungs- und Bewertungskategorien von TWR & AWR	84
6.7.4	Faktorbildung.....	86
6.7.5	Einflüsse der Antwortlänge auf WUSCT	87
6.7.6	Einflüsse der Verbalen Intelligenz auf TWR, AWR und WUSCT	87
6.8	Voranalyse der Teilnehmergruppen.....	88
6.8.1	Alter	88
6.8.2	Beruf	89
6.8.3	Bildungsniveau und Berufstätigkeit.....	89
6.8.4	Verbale Intelligenz.....	90
6.8.5	Psychotherapeutische Behandlung.....	90
6.8.6	Familienstand.....	90
7	Ergebnisse der Validierungsstudie	92
7.1	Kriteriumsvalidierung anhand der Inneren Validität	92
7.2	Kriteriumsvalidierung anhand der Äußerer Validität.....	95
7.3	Konstruktvalidierung anhand der Konvergenten Validität.....	98
7.4	Konstruktvalidierung anhand der Diskriminanten Validierung.....	101
7.5	Weitere Analysen.....	104
8	Diskussion	112
8.1	Kriteriumsvalidierung anhand der Inneren Validität	112
8.1.1	Multiple Validitätskoeffizienten	112
8.1.2	Bivariate Korrelationen.....	114
8.1.3	Zusammenfassung der Befunde zur Inneren Validität.....	114
8.2	Kriteriumsvalidierung anhand der Äußerer Validität.....	115
8.3	Konstruktvalidierung anhand der Konvergenten Validität.....	117
8.3.1	Psychologisches Feingefühl und Persönliches Wachstum als Indikatoren der Persönlichkeitsreife.....	118
8.3.2	Ego-Entwicklung als Indikator der Persönlichkeitsreife	119
8.3.3	Zusammenfassung der Befunde der Konvergenten Konstruktvalidierung.....	120
8.4	Konstruktvalidierung anhand der Diskriminanten Validität	120
8.4.1	Umweltkontrolle als Indikator der Adaptivität.....	120
8.4.2	Zusammenfassung der Befunde zur Typologie der Suchenden und der Erfolgstypen.....	121

8.4.3	Der Wortschatztest als Indikator der Intelligenz	123
8.4.4	Alter	123
8.4.5	Geschlecht.....	124
8.4.6	Zusammenfassung der Befunde der Diskriminanten Konstruktvalidierung...	124
8.5	Schlüsse auf die Inhaltsvalidität der Weisheitstestverfahren.....	125
8.6	Begrenzungen der Studie.....	125
8.6.1	Begrenzungen des Studiendesigns.....	126
8.6.2	Diskussion der Teilnehmergruppen bezüglich der Rekrutierung, Zusammensetzung und Ergebnisgeneralisierung	128
8.6.3	Diskussion der Weisheitstestverfahren.....	131
8.6.4	Begrenzungen des Methodischen Vorgehens	136
8.7	Ausblick.....	139
9	Zusammenfassung.....	142
10	Literaturverzeichnis.....	145
11	Tabellenverzeichnis.....	159
12	Abbildungsverzeichnis.....	161
13	Anhang	162
A	Faltblatt Weisheitsnominierungsverfahren	162
B	Anschreiben.....	163
B.1	Computervermitteltes Anschreiben für die Nominiertengruppe.....	163
B.2	Anschreiben – spezifisches für Kontrollgruppe.....	165
B.3	Anschreiben - spezifisches für Psychotherapeuten	165
C	Das Fragebogenheft.....	166
D	Material für das Telefoninterview	178
D.1	Gesprächsleitfaden für das Telefoninterview	178
D.2	Dokumentation des Telefoninterviews	180
E	Bewertungskategorien für TWR und AWR	181
E.1	TWR 181	
E.2	AWR 181	
F	Tabellen zur weiteren Übersicht der statistischen Auswertung.....	183
F.1	Matchingprofiltabellen.....	183
F.2	Mittelwertstabellen der Weisheitstests	186
F.3	Korrelationstabellen der Weisheitstests	188
G	Rückmeldung für die Teilnehmer	191

**Die Weisheit eines Menschen misst
man nicht nach seiner Erfahrung,
sondern nach seiner Fähigkeit,
Erfahrungen zu machen.**

George Bernhard Shaw

Erklärung

Hiermit versichere ich gemäß § 18 Abs. 8 der Diplomprüfungsordnung Psychologie der Universität Koblenz-Landau, Campus Landau, in der Fassung vom 18.02.1993, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe. Die Arbeit hat in gleicher oder ähnlicher Form noch keinem anderen Prüfungsausschuss vorgelegen.

Landau, den 1. Juli

Marina Hoffmann

Abkürzungsverzeichnis

Dieses Verzeichnis enthält die verwendeten Abkürzungen für Testverfahren dieser Arbeit in alphabetischer Reihenfolge.

3d-WS	- 3dimensionale-Weisheitsskala (Ardelt, 2003, 2006)
ACL	- Adjective Check-List (Gough & Heilbrun, 1983)
AWR	- Allgemeines Weisheitsrating (Helson & Srivastava, 2001, 2002)
HAWIE-R	- Hamburg Wechsler Intelligenz Test für Erwachsene – Revision (Tewes, 1995)
HS	- latenter Faktor der Testverfahren TWR, AWR und PWS (Helson & Srivastava, 2001, 2002)
IPIP	- International Personality Item Pool (Goldberg, et al., 2006)
PF	- Psychologisches Feingefühl (Goldberg, 2006)
PW	- Persönliches Wachstum (Ryff, 1989; Ryff & Keyes, 1995)
PWB	- Psychologisches Wohlbefinden (Ryff, 1989; Ryff & Keyes, 1995)
PWR	- Praktisches Weisheitsrating (Helson & Srivastava, 2001, 2002)
PWS	- Praktische Weisheitsskala (Wink & Helson, 1997)
SAWS	- Selbstadministrative Weisheitsskala (Webster, 2003, 2007)
TWR	- Transzendentes Weisheitsrating (Wink & Helson, 1997)
UK	- Umweltkontrolle (Ryff, 1989; Ryff & Keyes, 1995)
WUSCT	- Washington University Sentence Completion Test (Loevinger, 1976)

1 Einleitung

Die vorliegende Diplomarbeit ist eine Konstrukt- und Kriteriumsvalidierung von schriftlichen Weisheitstestverfahren. In der expliziten Weisheitsforschung wurde bisher sehr wenig vergleichende Forschung betrieben. Vielmehr fand eine zunehmende Differenzierung der Weisheitstheorien, Operationalisierungen und Erhebungsmethoden statt. Ein empirischer Vergleich von Testverfahren stellt einen Schritt zur Konvergenz der einzelnen Forschungsrichtungen dieses kleinen Forschungsbereiches dar, auf den weitere folgen müssen.

Es fällt schwer das komplexe Konstrukt Weisheit in eine allgemein akzeptierte Theorie oder Definition zu überführen. Weisheit wird übereinstimmend als zeitstabiler Persönlichkeitstrait verstanden (Ardelt, 2003; Brown & Greene, 2006; Hartmann, 2005; Helson & Srivastava, 2001, 2002; Staudinger, Lopez & Baltes, 1997; Sternberg, 1998; Takahashi & Overton, 2002; Webster, 2003, 2006, Wink & Helson, 1997). Einige der wichtigsten Definitionen werden im Abschnitt 2.3 vorgestellt, die sich vor allem aufgrund des theoretischen Hintergrunds der einzelnen Weisheitsforscher unterscheiden. Nach Takahashi und Overton (2002) sind die inhaltlich unterschiedlichen Aspekte dieser Theorien gleichwertig, und deren Eigenwert für die Weisheitsforschung sei anzuerkennen.

Um den jeweiligen Wert von Testverfahren als Operationalisierungen von Weisheitsdefinitionen zu prüfen, bedarf es der vergleichenden Forschung. Beim Vergleich der Tests an einer Stichprobe, erlaubt das Gütekriterium der Validität diese Beurteilung. Der Begriff Weisheit wird im theoretischen Teil dieser Arbeit oftmals mit unterschiedlichen Definitionen verschiedener Forscher erläutert. Ihnen ist jedoch gemeinsam, dass Weisheit im Unterschied zur Klugheit eine menschliche Grundhaltung beschreibt, die eine durch allgemeine Lebenserfahrung und Einsicht basierende mentale Reife darstellt (vgl. Brockhaus, 2004). Diese Begriffsbestimmung wird jeweils dann unter Weisheit in dieser Arbeit verstanden, wenn eine andere Definition im jeweiligen Kontext fehlt.

In Kapitel 3 werden alle bekannten Weisheitstestverfahren anhand der Kategorien Leistungsmessung, Selbsteinschätzungsfragebogen und kombinatorische Vorgehensweisen bei multimodalen Erhebungsverfahren beschrieben. Die in Interviews erhobenen Leistungstests Allgemeine und Selbstbezogene Weisheit sind im Vergleich zu schriftlichen Erhebungsverfahren in unökonomischer Durchführung und Auswertung. Zu prüfen ist daher, ob die schriftlichen Weisheitsskalen der 3dimensionalen-Weisheit (3d-WS), der Selbstadministrativen Weisheit (SAWS), und der Praktischen Weisheit (PWS) sowie die beiden schriftlichen Ratingver-

fahren der Transzendenten (TWR) und Allgemeinen Weisheit (AWR) Konstrukt- und Kriteriumsvalidität beweisen, und somit echte Alternativen zu aufwendigen mündlichen Erhebungsverfahren darstellen. Zur Beantwortung dieser Fragestellung werden im empirischen Teil dieser Arbeit die schriftlichen Weisheitstests erstmals verglichen.

Das zentrale Konstrukt einer Vergleichsstudie ist die Validität. Die Validität ist das Gütekriterium der Gültigkeit eines Tests, mit der das Persönlichkeitsmerkmal erfasst wird, das zu messen beabsichtigt war. Nach Lienert und Raatz (1998) erlaubt der Testwert eines validen Tests den unmittelbaren und fehlerfreien Rückschluss auf den Ausprägungsgrad des Merkmals. Die Aufgabe der vorliegenden Arbeit ist daher, die Übereinstimmung der einzelnen Weisheitstests hinsichtlich des theoretischen Konstruktes Weisheit als auch anhand von Kriterien zu beurteilen. Die Aspekte der Inhalts-, Konstrukt-, und Kriteriumsvalidität werden nun nacheinander vorgestellt.

Inhaltsvalidität bedeutet, dass alle Aspekte des zu erfassenden Konstruktes in dem zu beurteilenden Test operationalisiert sind. Inhaltsvalidität kann nicht anhand von objektiven Kriterien bestimmt werden. Vielmehr ist sie bei der Testkonstruktion beispielsweise durch Expertenratings abzusichern. In den Kapiteln 2 und 3 werden Ausführungen zur inhaltlichen Validität der Weisheitstestverfahren vorgestellt. Auf die schriftlichen Verfahren bezogen, ist nicht von einer perfekten Inhaltsvalidität zu sprechen. Da die Inhaltsvalidität die Größenordnung der Konstrukt- und Kriteriumsvalidierung wesentlich mitbestimmt, kann sie in einer Validierungsstudie der Konstrukt- und Kriteriumsvalidität nicht vollkommen ausgeblendet werden.

Die Existenzberechtigung eines psychologischen Konstruktes gründet sich auf dessen Eigenständigkeit gegenüber anderen wissenschaftlich etablierten Konstrukten. Eine Verortung eines Konstruktes in das Beziehungsgefüge wissenschaftlich belegter Konstrukte, auch nomological network genannt (Cronbach, 1971), wird über eine Konstruktvalidierung durchgeführt. Die Konstruktvalidität wird über die Prüfung konvergenter und diskriminanter Annahmen bestimmt. Konvergente Validität ist gegeben, wenn theoretische Beziehungen zu verwandten Konstrukten empirisch nachgewiesen sind. Diskriminante Validität ist gegeben, wenn theoretische Unabhängigkeit zu theoretisch nicht verwandten Konstrukten empirisch nachgewiesen ist (Campbell & Fiske, 1959). Zur Konstruktvalidierung der Weisheitstests in der vorliegenden Studie wurden daher folgende etablierte Konstrukte ausgewählt: Persönlichkeitsreife, Adaptivität, Intelligenz, Alter und Geschlecht. Kapitel 4 führt diese Konstrukte ein und präsentiert ihren wissenschaftlichen Diskurs im Zuge der Weisheitsforschung.

Kriteriumsvalidität beschreibt die Übereinstimmung eines theoretischen Konstruktes

und dessen empirischer Testung mit objektiven, externen Außenkriterien, sprich beobachtbaren Merkmalen, oder mit subjektiven Kriterien, wie Einschätzungen von Experten oder Laien (Michel & Conrad, 1982). Ist die Ausprägung des Kriteriums gleichzeitig gegeben, spricht man von Konkurrenter Validierung im Gegensatz zur Prädiktiven Validierung, bei der das Kriterium erst zu einem späteren Zeitpunkt erkannt und gemessen werden kann. Die Konkurrente Validität ist über die Methode der bekannten Gruppen bestimmbar. Existieren bekannte Gruppen, die sich im Hinblick auf das Zielkriterium unterscheiden, so sollte ein Test ebenfalls zwischen diesen Gruppen unterscheiden können, um Konkurrente Validität aufzuweisen (Schnell, Hill & Esser, 1999). In der vorliegenden Studie wird dieses Herangehen wie folgt realisiert. Durch Laiennominierungen werden Personen mit hoher Weisheit direkt als Nominerte für die Teilnahme an der Studie vorgeschlagen, und daraufhin ein gematchte Kontrollpersonen rekurtiert. Als indirektes Kriterium wird der Beruf Psychotherapeut¹ verwendet, da Personen dieser Berufsgruppe berufliche Expertise in der Bewältigung menschlicher Probleme besitzen. Ausführungen zum theoretischen Hintergrund der Wahl dieser äußeren Kriterien finden sich in den Abschnitten 4.1, und 6.

Die theoretischen Hypothesen der Konstrukt- und Kriteriumsvalidierung dieser Studie werden in Kapitel 5 ausführlich dargestellt. Sollten nach Lienert und Raatz (1998) validitätsähnliche Testverfahren vorliegen, so kann über intraindividuelle Vergleiche zwischen den Tests eine Validitätskontrolle erfolgen. „Validitätskoeffizienten gelten zunächst nur für ein bestimmtes Validitätskriterium und für die Population, aus der die Validierungsstichprobe stammt (Lienert & Raatz, 1998, S. 224).“ Die Ergebnispräsentation in Kapitel 7, und Diskussion in Kapitel 8, werden den Erkenntnisgewinn und die Tragweite der Ergebnisse der empirischen Arbeit erleuchten und Raum für weitere Forschungsmöglichkeiten aufzeigen, siehe dazu Kapitel 9.

Weisheitsforschung stellt ein interessantes aber aufwendiges Forschungsgebiet dar, da Weisheitsforschung kostenintensiv im Bezug auf Zeit, menschliche Ressourcen, und Geld ist (Jordan, 2005). Die zentrale Problematik liegt darin, dass Weisheit ein äußerst rares und nicht extern erkennbares Persönlichkeitsmerkmal ist. Die besondere Herausforderung dieser Diplomarbeit ist darüber hinaus, gleich fünf Testverfahren zu validieren. In Anbetracht dieser Aufgabenstellung, ist eine ausführliche theoretische und methodische Darstellung der Testverfahren notwendig. In Konsequenz erreicht diese Prüfungsarbeit einen hohen Seitenumfang.

¹ Zur Erleichterung des Leseflusses wird nachfolgend auf die gleichzeitige Verwendung von männlichen und weiblichen Bezeichnungen verzichtet. Die männlichen Bezeichnungen umfassen somit jeweils Frauen und Männer gleichermaßen.

2 Was ist Weisheit aus psychologischer Sicht?

Erste, empirisch geprüfte, Antworten auf diese Frage entstammen der Entwicklungspsychologie in ihrem Interesse an positiven Seiten des Alterns, und der Strömung der Positiven Psychologie in ihrem Interesse an den guten Aspekten des menschlichen Miteinanders (Seligman & Csikszentmihalyi, 2000). Zu einer Übereinkunft, wie Weisheit zu Definieren, Operationalisieren und zu Erfassen ist, sind die führenden Weisheitstheoretiker der Psychologie bisher nicht gekommen (Birren & Svenson, 2005). Beispielsweise wird im Buch *“Classification of Character Strengths and Virtues“*, von Peterson und Seligman (2004) die Wertekategorie Weisheit und Wissen wie folgt verstanden:

“For example, the virtue of wisdom can be achieved through such strengths as creativity, curiosity, love of learning, open-mindedness, and what we call perspective—having a “big picture” on life. These strengths are similar in that they all involve the acquisition and use of knowledge, but they are also distinct. Again, we regard these strengths as ubiquitously recognized and valued, although a given individual will rarely, if ever, display all of them (Peterson & Seligman, 2004, p. 13)”.

In der selben Publikation wird die in der Definition benannte, hierarchisch untergeordnete Charakterstärke perspective, ebenfalls Weisheit genannt und so definiert: „Perspective [wisdom]: Being able to provide wise counsel to others; having ways of looking at the world that make sense to oneself and to other people (Peterson & Seligman, 2004, p. 29)“. Anhand dieser Definitionen wird deutlich, dass Weisheit als höherer Wert oder als Charakterstärke eines Menschen verstanden, bedeutende Unterschiede hervorbringt. Zum besseren Verständnis solcher Differenzen, folgt eine Einführung in die Entwicklung der psychologischen Weisheitsforschung in Unterkapitel 2.1. Nur durch eine gute theoretische Fundierung und ausführliche Validierung kann geklärt werden, was Weisheit aus psychologischer Sicht ist.

Die ersten Weisheitsstudien erforschten Lagentheorien zu Weisheit zum Zwecke der externen Validierung. Dieser Forschungsabschnitt wird Unterkapitel 2.2. gewidmet, während 2.3 Schwerpunkte von Weisheitsdefinitionen präsentiert. Diese Forschungsfelder werden als implizite Weisheitsforschung bezeichnet. Aufbauend auf die implizite Weisheitsforschung entstanden unterschiedliche Ansätze zur Messung von Weisheit, die in Kapitel 3 beschrieben werden. Eine schematische Zusammenfassung dieser vielfältigen Definitionen und Klassifikationsschemata von Weisheitstestverfahren wird in Kapitel 3.4 eingeführt. Diese Darstellung schließt die Fragestellung der Inhaltsvalidität von Weisheit. Kapitel 4 präsentiert empirische Befunde der Konstrukt- und Kriteriumsvalidierung, die zu einer umfassenden Beantwortung der Frage, was Weisheit aus psychologischer Sicht ist, notwendig sind.

2.1 Einführung in die Entstehungsgeschichte der psychologischen Weisheitsforschung

Weisheit im Alter ist das erste Weisheitsmodell und wurde von Granville Stanley Hall, dem ersten Präsidenten der APA, im Jahre 1922 erstellt (Seligman & Csikszentmihalyi, 2000; Staudinger, 1996). Aufgrund fehlender Weiterführung finden sich die Wurzeln der Weisheitsforschung jedoch größtenteils in späteren Modellen der Persönlichkeitsentwicklung. Micklers Dissertation (2004) über die Nähe von Persönlichkeitsreife und Weisheit folgend, werden die drei Traditionen der Entwicklungstheorien reifer Persönlichkeit, der (1) Persönlichkeitspsychologie, der (2) Entwicklungspsychologie und der (3) Klinischen Psychologie, nacheinander vorgestellt. Aus allen drei Traditionen werden Konstrukte aufgeführt, die für Weisheit grundlegend sind, wie beispielsweise Kenntnis und Integration der eigenen Persönlichkeit, Kognitiven Komplexität, Selbsttranszendenz und Ausrichtung auf das Wohlergehen anderer.

Die Messbarkeit von *Weisheit* wurde in der Psychologie beispielhaft von Gordon W. Allport angezweifelt, der empfiehlt, Konzepte, die Werte und Charakter betreffen, der Philosophie zu überlassen, während sich die Persönlichkeitspsychologie auf messbare, adaptive Traits beschränken sollte, wozu ein Konstrukt wie Weisheit nicht zu zählen sei (Peterson & Seligman, 2004). Die heutige explizite Weisheitsforschung ist jedoch Teil der Traitforschung der Persönlichkeitspsychologie (Hartmann, 2005; Staudinger, Maciel, Smith & Baltes, 1998; Sternberg, 1998). Weisheit beweist in empirischen Studien Stabilität sowie Eigenvarianz und entwickelt sich auf der Grundlage mehrerer Persönlichkeitseigenschaften (Ardelt 2003; Brown & Greene, 2006; Hartmann, 2005; Helson & Srivastava, 2001; Staudinger et al. 1997; Takahashi & Overton 2002; Wink & Helson, 1997; Webster 2003, 2006).

2.1.1 Entwicklungsmodelle der Persönlichkeit der Persönlichkeitspsychologie

Es gibt viele Modelle, Definitionen und Operationalisierungen von Persönlichkeitstheoretikern die den hoch entwickelten, reifen Menschen thematisieren. An dieser Stelle wird die Selbstverwirklichte Persönlichkeit von Abraham Maslow herausgegriffen. Nach Maslow ist die Selbstverwirklichte Persönlichkeit anhand von 15 Eigenschaften dem Ideal der Persönlichkeitsentwicklung nahe. Eine genauere Ausführung findet sich in Fisseni (2003). Was für Durchschnittsmenschen konfliktvolle Dichotomien, Polaritäten oder philosophische Dilemmas sind, werden von selbstverwirklichten Menschen auf höherer Ebene integriert und aufgelöst (Maslow, 1999). Maslows Ansatz ist einer der Vorläufer der Positiven Psychologie (Seligman & Csikszentmihalyi, 2000) und beeinflusste Loevingers Modell der Ego-Entwicklung

sowie Ryffs Modell des Psychologischen Wohlbefindens, die häufig in Weisheitsstudien verwendet werden, siehe Kapitel 4.2.2.2 und 4.2.3.3.

Aus der Persönlichkeitsforschung entstammt die Zweiteilung erwachsener Reifekonzepte in Persönlichkeitsreife und Adaptivität. Persönlichkeitsreife erfordert eine Ausrichtung auf Werte und das Gemeinwohl, sowie dem Streben nach Selbstverwirklichung. Dagegen erfordert Adaptivität Extraversion, positives Selbstkonzept, interne Kontrollüberzeugung und positive soziale Beziehungen (Argyle, 1987; Lindenberger & Baltes, 1997; Ryan & Deci, 2001; Ryff, 1989; Ryff & Keyes, 1995; Wink & Helson, 1987). Beide Konstrukte sind zwei lediglich leicht positiv korrelierte Dimensionen.

Persönlichkeitswachstum und Persönlichkeitsreife sind zentrale Bedingungen zur Entwicklung von Weisheit (Ardelt, 2003; Clayton & Birren, 1980; Helson & Srivastava, 2001; Holliday & Chandler, 1986; Orwoll & Perlmutter, 1990; Webster, 2003). Weisheit integriert diese diversen Eigenschaften auf hohem Niveau. Die Trennung zwischen Adaptivität und Persönlichkeitsreife findet sich in Kapitel 4.2.3.3 und 4.2.3.4 bei der Beschreibung konvergenter und diskriminanter Zusammenhänge von Weisheit und wird empirisch verwendet.

2.1.2 Entwicklungsmodelle der Persönlichkeit der Entwicklungspsychologie

Einige Modelle der Entwicklungspsychologie verstehen Weisheit als höchste Stufe der Persönlichkeitsentwicklung. Die Arbeiten der folgenden Theoretiker sind besonders bedeutsam.

Nach Erik Erikson (Fisseni, 1998) steht Weisheit in Form einer integrierten Persönlichkeit am Ende der acht psychosozialen Krisen, in denen bestimmte Egokompetenzen zu entwickeln sind. Die letzte Stufe erfordert die Integration und Akzeptanz des bisherigen Lebens und Todes. Daraus entstehe eine Loslösung, die es erlaubt, Probleme von einer holistischen und abstrakten Perspektive zu bearbeiten. Eriksons Theorie diente Webster (2003) zur Konstruktvalidierung seines Weisheitsfragebogens, siehe 3.2.2.

Loevingers Konzept der Ego-Entwicklung ist ein 9-stufiges Modell. Die neunte, integrierte Stufe steht für Weisheit durch die höchste Charakterentwicklung, den kognitiven und interpersonalen Stil sowie der Ebene der bewussten Auseinandersetzung (Loevinger, 1976). Auf die besondere Stellung von Ego-Entwicklung wird in 3.4 und 4.2.2.2 eingegangen.

Labouvie-Vief (Labouvie-Vief & Medler, 2002) unterscheidet zwischen Affektkomplexität, der Reife zuzuordnen, und Affektoptimierung, der Adaption zuzuordnen. In ihrem 4-stufigen Modell der kognitiven und emotionalen Entwicklung ist die Autonome Stufe die höchste. Forschungsergebnisse der expliziten Weisheitsforschung stützen diese Befunde (Helson & Srivastava, 2001; Kunzmann & Baltes, 2003).

2.1.3 Entwicklungsmodelle der Persönlichkeit der Klinischen Psychologie

Aufgrund der Defizitorientierung der klinischen Psychologie zählen Modelle psychischer Gesundheit nicht zu den zentralen Forschungsthemen, weshalb theoretische Abhandlungen über Persönlichkeitsreife und Weisheit nicht selbstverständlich sind. Jedoch finden sich Theoretiker der psychoanalytischen und humanistischen Tradition, wie Jung (1930), Fromm (1954), Rogers (1961) und Ellis (1977), die sich auf Reife und Persönlichkeitsentwicklung als Antezedent von Weisheit beziehen und somit Grundlagen für die Weisheitsforschung liefern.

Die deutlichsten Hinweise auf Weisheit beinhaltet die Arbeit von Carl Rogers (1961). Die bezeichnenden Eigenschaften seiner fully functioning person sind Offenheit für Erfahrungen, Akzeptanz anderer, Selbstvertrauen, Selbstständigkeit, Komplexität, Prozess-sein und die Bereitschaft für Veränderung. Diese Beschreibung ähnelt nicht nur Maslows Selbstverwirklichter Persönlichkeit, sondern viele dieser Aspekte finden sich in Weisheitstheorien wieder, was beispielsweise in Abschnitt 4.2.3.1 gezeigt wird.

2.2 Laientheorien zu Weisheit

Nach Baltes und Smith (1990) sollte eine theoretische Weisheitsdefinition einerseits große Übereinstimmung mit philosophischen Theorien aufweisen, die in Unterkapitel 2.3.1 behandelt werden, und andererseits Aspekte von Laientheorien integrieren können. Die Erforschung des Laienverständnisses von Weisheit stellt den Beginn der empirischen Weisheitsforschung dar und wird bis heute weiterbetrieben (Clayton & Birren, 1980; Glück, Bluck, Baron & McAdams, 2005; Sternberg, 1985; Takahashi & Bordia, 2000). Die Forschungsarbeiten zum Laienverständnis von Weisheit sind in drei Bereiche unterteilbar, die jeweils einen anderen Zugang zu den Laientheorien wählen: Erstens, die Erforschung von Definitionen oder Determinanten des Weisheitsverständnisses der Allgemeinbevölkerung, siehe Abschnitt 2.2.1. Zweitens, Nominierungsstudien zu Weisheit, Abschnitt 2.2.2, um aus den Nominierungen Rückschlüsse über die zu Grunde liegenden Weisheitstheorien der Laien zu ziehen. Drittens, die Forschungsrichtung der Gelebten Weisheit in Abschnitt 2.2.3, die in Laienberichten nach Indikatoren für persönliche Weisheit sucht.

2.2.1 Weisheitsverständnis

Dieser Ansatz erforscht das allgemeine Verständnis von Weisheit und deren Facetten in der Allgemeinbevölkerung direkt durch Befragung. Die große Bandbreite an positiven Eigenschaften, die von Laien mit Weisheit in Verbindung gebracht wird, erfordert es, Dimensionen

der Weisheit aus diesen Aussagen zu extrahieren.

Bluck und Glück (2003) fassen die Weisheitskriterien impliziter Studien von Clayton und Birren (1980), Holliday und Chandler (1986), Jason et al. (2001) sowie Sternberg (1985) mittels Literaturreview zu fünf Dimensionen zusammen: (1) Kognitive Fähigkeit, (2) Einsicht, (3) Reflektive Einstellung, (4) Sorge um Andere und (5) Reale-Welt-Problemlösefähigkeiten. Die Bedeutung dieser Dimensionen soll hier knapp dargestellt werden. Bluck und Glück vereinen unter kognitive Fähigkeit, kristalline und fluide Intelligenz sowie Problemlösefähigkeit. Die zweite Komponente, Einsicht, beschreibt die Motivation und Fähigkeit, durch Intuition und Perspektivenübernahme sich Selbst und den Kern von Problemen zu erkennen. Reflektive Einstellung umfasst Besonnenheit sowie Perspektivenübernahmefähigkeit. Die vierte Dimension, Sorge um Andere ist gekennzeichnet durch Interesse und Verständnis für Andere, dass auf Fairness, Freundlichkeit und Respekt aufbaut. Die letzte Komponente, Reale-Welt-Problemlösefähigkeit, steht für die Anwendung der beschriebenen Fähigkeiten, was Urteilsfähigkeit, Sensibilität und soziale Kompetenz in Beratungssituationen voraussetzt und Wissen um die Grenzen der eigenen Möglichkeiten erfordert.

Implizite Theorien erlauben lediglich Vermutungen über die interne und externe Validität eines empirischen Konzeptes. Eine Besonderheit impliziter Laintheorien ließ sich dementsprechend nicht durch die explizite Weisheitsforschung replizieren. Viele Menschen assoziieren Weisheit mit hohem Alter (Clayton & Birren, 1980; Baltes, Staudinger, Maercker & Smith, 1995; Jason et al., 2001). In einer Studie von Clayton und Birren (1980) zeigte sich, dass ein positiver Zusammenhang zwischen Alter und Weisheit lediglich vom jüngeren und Erwachsenen mittleren Alters angenommen wird. Clayton und Birren erklären die differenten Erwartungen an die Variable Alter durch die veränderte Urteilsbasis der Alten, die den vermuteten linearen Zuwachs von Weisheit am deutlichsten erkennen sollten. Empirisch wird die Scheinkorrelation von *Alter* und *Weisheit* durch die vermittelnden Variablen Lebenserfahrung und Verarbeitung von Lebenserfahrungen gebildet. In Abschnitt 4.2.1.1 werden die Ergebnisse der expliziten Weisheitsforschung im Bezug auf diese Variablen dargestellt.

2.2.2 Profil einer weisen Personen

Die zweite Herangehensweise zur Erforschung des Weisheitsverständnisses von Laien ist, Laien in Nominierungsstudien zu befragen, wer nach ihrem Verständnis als weise zu bezeichnen sei. Aus diesen Studien lassen sich folgende Profileigenschaften von Personen zusammenfassen, die nominiert wurden. Der typische Nominierte ist demnach männlich, über 50 Jahre alt, und bekleidet eine Mentorenrolle (Orwoll & Perlmutter, 1990). Dieses Profil lässt

sich in seiner Zusammensetzung wie folgt erklären. Dass Menschen generell dazu neigen, jeweils ältere Personen als Weise zu nominieren (Denney, Dew & Kroupa, 1995), folgt der bereits beschriebenen Annahme, *Weisheit* sei positiv mit dem *Alter* korreliert. In der Studie von Jason et al. (2001) beschreiben die Nominatoren ihre Nominierten als einsichtige und spirituelle Mentoren oder Führungspersonen, die den Nominatoren Unterstützung in der persönlichen Lebenszielfindung oder Werteorientierung boten. Obwohl nur wenige Menschen denken, Weisheit habe etwas mit dem Geschlecht einer Person zu tun, werden jeweils konsistent etwa doppelt so viele Männer wie Frauen nominiert (Jason et al. 2001, Denney et al. 1995; Sowarka, 1989). Eine Erklärung dieser Ungleichverteilung kann durch die geringere Präsenz älterer Frauen in gesellschaftlich beachteten Mentorenrollen gegeben werden.

Das genannte Profil eines Weisen lässt sich somit durch die Konfundierung der Variablen Alter und Geschlecht mit Personen in einer einflussreichen Mentorenrolle erklären, die einen bedeutsamen Eindruck bei den Nominatoren hinterlassen. Obwohl der Stellenwert der beiden demographischen Faktoren Geschlecht und Alter anzuzweifeln ist, verbietet sich jedoch der Rückschluss, Nominierungen als nicht valide zu bezeichnen. Es erscheint jedoch wahrscheinlich, dass einige weise Personen auf Grund dieser Faktoren bekannter sind und somit eher nominiert werden als andere.

2.2.3 Gelebte Weisheit

Die zentrale Annahme der Forschungsrichtung zur Gelebten Weisheit ist, dass jeder Mensch in seinem Leben Ereignisse durchlebt, durch die an Weisheit hinzugewonnen wird. Studien untersuchen die Art solcher Erfahrungen in autobiographischem Selbstberichten (Bluck & Glück, 2002). Das Urteil darüber, was als weise angesehen wird, wird einzig von den Teilnehmern selbst bestimmt. Diese autobiographischen Erzählungen entsprechen nach Glück, Bluck, Baron und McAdams (2004) theoretischen und empirischen Arbeiten, die nahe legen, dass Weisheit (1) eine Kombination aus experimentellem oder praktischem Wissen, Kognition, Affekt und Aktion ist (Ardelt, 1997; Clayton & Birren, 1980; Labouvie-Vief, 1990; Sternberg, 1998); (2) kontextabhängig ist (Kekes, 1983; Staudinger & Baltes, 1996a; Sternberg, 1998) und durch (3) Lebenserfahrung erlangt wird (Baltes & Smith, 1990; McKee & Barber, 1999; Sternberg, 1998).

In einer deutschsprachigen Studie von Bluck und Glück (2004), benennen die Teilnehmer durchschnittlich vier Ereignisse in denen sie weise gehandelt haben. Die bedeutendste dieser Situationen und die daraus resultierende Handlung zeigen thematische Unterschiede in den befragten Altersgruppen, die sich nach Geschlecht und Bildungsniveau gleichen. Im All-

gemeinen wurden negative Situationen erinnert, die in ein positiv beurteiltes Ergebnis durch die eigene Handlung überführt wurden. Für die Gruppe der Jugendlichen (15-20 Jahre) war demnach Empathie und Unterstützung die häufigste gelebte Weisheitsart, bei den jungen Erwachsenen (30-40 Jahre) rückt Selbstvertrauen und Selbstbestimmtheit in den Mittelpunkt, die Älteren (60-70 Jahre) benennen Wissen und Flexibilität am häufigsten. In einer zweiten, amerikanischen Studie (Glück et al., 2003) zeigten sich andere Ergebnisse. Die Altersspanne der US-Amerikaner lag zwischen 30 und 72 Jahren. Im Gegensatz zur deutschsprachigen Teilstichprobe gleichen Alters, wurde hier vor allem der Bereich der Empathie und Unterstützung genannt, was möglicherweise durch den das Fehlen der Instruktion zur Auswahl der bedeutendsten Situation aus den salienten Erinnerungen erklärbar ist.

Insgesamt können nahezu alle Jugendlichen und Erwachsenen Ereignisse berichten, in denen sie weise gehandelt haben. Die Situation und die erforderlichen Handlungsweisen variieren zwischen den Altersgruppen, was durch die verändernden Lebensaufgaben im Laufe des Erwachsenenlebens begründet wird (Bluck & Glück, 2002; Glück et al. 2004).

2.2.4 Ausblick auf die explizite Weisheitsforschung

Zusammenfassend lässt sich sagen, dass eine große Übereinstimmung über das Verständnis von Weisheit in der erwachsenen Bevölkerung herrscht, und nahezu jeder Proband mindestens eine Erinnerung an Situationen eigener Weisheit benennen kann. Zusätzlich zeigen Laientheorien von Weisheit große Übereinstimmung zu den empirischen Arbeiten der expliziten Weisheitsforschung, die im nächsten Kapitel vorgestellt werden.

Anwendungsbeispiele von Laientheorien in der expliziten Forschung gibt es in unterschiedlicher Weise. Eine Beurteilung der Gelebten Weisheitserkenntnisse mittels theoriegeleiteter Kriterien systematisierten Wink und Helson (1997) in Form des Tests der Transzendenten Weisheit, siehe Abschnitt 3.3.1.3. Mit dem Weisheitsverständnis von Laien arbeiteten Webster und Brown. Webster (2003) belegte die Kriteriumsvalidität seines SAWS durch signifikante Unterschiede in Gesamtpunktwerten je nachdem, ob die Teilnehmer möglichst weise oder möglichst töricht antworten sollten, siehe Abschnitt 3.2.2. Brown (2004) bildete mittels Grounded Theory aus Laientheorien seine Weisheitsdefinition, die daraufhin zur Entwicklung der Weisheitsentwicklungsskala diente, siehe 3.2.3. Nominierungsverfahren, in Kapitel 4.1 beschrieben, wurden zur externen Validierung in der expliziten Forschung von Ardel (2003) und der Berliner Forschungsgruppe angewendet (Baltes et al. 1995). Diese Studien beweisen, dass die empirische Weisheitsforschung Übereinstimmungen mit den Laientheorien von Weisheit haben.

2.3 Weisheitsdefinitionen

“Wisdom is a value-laden subject. Trying to comprehend its dimensions and meanings, which are difficult to define, explain, and operationalize, has stirred the imaginations of scientists and artists, among others, for millennia (Achenbaum, 2004, p. 300)”. Weisheit zu ergründen beschäftigte Philosophen diverser Epochen. In historischer Reihenfolge sind exemplarisch Ptahhotep, Laotse, Konfuzius, Buddha, Sokrates, Plato, Aristoteles, Salomon, Augustinus, Thomas von Aquin und Kant zu nennen, deren Theorien in mehreren Überblicksarbeiten zusammengetragen wurden (Assmann, 1991; Baltes, 2007; Birren & Svensson, 2005; Tügel 2006). Baltes extrahiert mittels einer solchen kulturhistorischen Analyse mehrere Determinanten von Weisheit (Baltes, 2007; Kunzmann & Baltes, 2005). Weisheit zeige sich demnach in schwierigen Problemen der Lebensführung und Sinnfindung, sie repräsentiere außergewöhnliches Wissen, Urteilsfähigkeit und die Fähigkeit, Ratschläge erteilen zu können. Weisheit stellt für Baltes die perfekte Integration von Wissen und Charakter, Geist und Werten dar. Sie koordiniert und fördert individuelles sowie gesellschaftliches Wachstum und erfordert daher Balance und Mäßigung. Zu Weisheit gehört die Erkenntnis über die Grenzen des Wissens und die Ungewissheiten der Welt und des Seins. In der Übersichtsarbeit zu Laientheorien der Weisheit von Bluck & Glück (2005), zeigen sich fünf interkorrelierte Dimensionen. Diese einzelnen Dimensionen wurden in obigem Kapitel 2.2.1 näher beschrieben.

Aus der Literaturübersicht wird erkennbar, dass das psychologische Weisheitskonzept ein dimensionales Konstrukt ist, das sich aus miteinander korrelierenden Facetten zusammensetzt. Das multidimensionale Verständnis von Weisheit ist mittlerweile anerkannte Forschungsmeinung, jedoch fehlt der Konsens über deren einzelne Dimensionen.

Die Definitionen der expliziten Weisheitsforschung spezialisieren sich meist auf einzelne Aspekte. Um die mannigfaltigen psychologischen Weisheitsdefinitionen zu verdeutlichen, werden deren einzelne Ansatzpunkte im Überblick aufgezählt: (1) Vereinigung von Kognition und Affekt (Birren & Fischer, 1990, Blanchard-Fields, Brannan & Camp, 1987; Clayton, 1976 nach Birren und Svensson, 2005; Csikszentmihalyi & Rathunde, 1990, Labouvie-Vief, 1990), (2) Unterscheidung zu Psychometrischer Intelligenz (Clayton, 1982), (3) Problemlösen (Arlin, 1990; Dittmann-Kohli, 1984; Kitchener & Brenner, 1990), (4) Integration von multiplen Dimensionen des Selbst (Clayton & Birren 1980; Holliday & Chandler, 1996; Kramer, 1990, Pascual-Leone, 1990; Orwoll & Perlmutter, 1990), (5) Wissen um, und die Reaktion auf die Grenzen des Menschlichen und des Wissens (Kitchener, 1983; Meacham, 1983; Taranto, 1989), (6) Balance zwischen Inter-, Intra-, und Extrapersonalen Interessen über die Zeit und Umwelt zum Wohle der Allgemeinheit (Sternberg, 1998, 2004a), (7)

Unterschiede zwischen westlicher und östlicher Kultur (Takahashi, 2000).

Exemplarisch wird die Forschungsdiskussion zum letzten Punkt kurz dargestellt. Nach Baltes und Kunzmann (2004), betont die westliche säkularisierte Philosophie die Analytische Weisheit, expertiseorientiert, urteilsfähig, kompetent im Umgang mit Schwierigkeiten oder unklaren Lebenssituationen. Die östliche Philosophie betont die Synthetische Weisheit, in Vereinigung von Geist und Werten in einer weisen Person, die Balance von Kognitiven, Affektiven und Reflektiven Komponenten der Weisheit erreicht (Ardelt, 2003, 2004a). Aus diesen Wurzeln entsprang die Diskussion zwischen den Forschern, ob und wie Weisheit in Menschen zu erfassen sei. Baltes und Kunzmann (2004) beschreiben ihre Position wie folgt:

“If one only follows studying wise persons, one cannot benefit from the rich historical work on wisdom conducted in other disciplines, especially Western secularized philosophy. Moreover, an immediate devotion to wisdom as wise persons would lay the foundation for a theory of wisdom in everyday or common-sense psychology rather than in scientific theory in the narrow sense (Baltes, & Kunzmann, 2004, p. 291)”.

Ardelt (2004a) fordert hingegen, dass Weisheit als Begriff zu weisen Menschen referieren soll, anstatt zu Expertise in den fundamentalen Pragmatismen des Lebens, wie er von der Berliner Gruppe verstanden wird. Weisheit sei mehr als ein Aspekt des Wissens, aber Wissen sei nur ein Aspekt von Weisheit, fassen Blanchard-Fields & Norris (1995, nach Ardelt, 2004b) diese Argumentation griffig zusammen.

Ausführliche Beschreibungen über historische und psychologische Weisheitstheorien geben Birren und Svensson (2005). Da Weisheit als mindestens so komplex wie das Konstrukt Intelligenz angesehen wird, ist nach Birren und Svensson (2005) anzunehmen, dass es noch lange Zeit benötigen wird, um heute schon bestehende Forschungsfragen beantworten zu können und sich entweder auf eine allgemein akzeptierte Definition von Weisheit zu verständigen, oder analog zur Intelligenz verschiedene getrennte Unterbereiche zu definieren.

“Psychology’s goal cannot be to claim a complete understanding of wisdom. Rather, its goal must be to discover more about what comprises the wise individual, to confirm through empirical investigation that explicit theories represent valid conceptions of the construct, and to identify other facets of social and cognitive functioning are related to wisdom (Jordan, 2005, p. 165).”

3 Weisheitstestverfahren

“If wisdom is indeed one of the highest forms of knowledge (Clayton & Birren, 1980) and an ideal of human development (Staudinger & Baltes, 1996) then measuring wisdom must involve sophisticated methods (Jordan, 2005, p. 18).” Einzig der Vergleich der Verfahren ermöglicht es, den Stellenwert theoretischer Ansätze und Methoden der Weisheitsforschung zu klären und somit einen wissenschaftlichen Beweis zur Inneren Validität von Weisheitsmessungen zu erbringen.

Erste Messungen von Allgemeiner Weisheit erfolgen von der Berliner Forschergruppe des Max-Planck-Institutes für Bildungsforschung über Interviews. Ihr Ansatz gilt der Beurteilung von mündlichen Reaktionen auf Szenarien an einem analytischen Ideal der Weisheit. Aus dieser Forschungsrichtung stammt das Interviewtestverfahren zur Selbstbezogenen Weisheit (Mickler, 2004; Staudinger et al. 2005).

Einen anderen Ansatz verfolgen Selbstauskunftsfragebögen, die Weisheit indirekt über Persönlichkeitseigenschaften erfassen. Die schriftlichen Verfahren zeichnen sich im Vergleich zu den Interviewverfahren durch hohe Ökonomie in Durchführung und Auswertung aus.

Eine weitere Möglichkeit das heterogene Konstrukt Weisheit zu erfassen ist, über multimethodale und multimodale Instrumente verschiedene Aspekte von Weisheit abzudecken. Die gemeinsame Varianz dieser Messungen lässt sich somit auf ein übergeordnetes Konstrukt zurückführen, das ein breiteres Verständnis von Weisheit abbildet, als es den Untertests alleine möglich wäre. Diese Herangehensweise wählen Helson und Srivastava (2001, 2002), die somit Praktische Weisheit, Transzendente Weisheit und Allgemeine Weisheit zu einem übergeordneten Konstrukt vereinen. Hartmann (2000, nach 2004) wählt ebenfalls diese Vorgehensweise. Über drei verschiedene Instrumente werden weise Produkte, Prozesse und Persönlichkeitseigenschaften erfasst.

In den folgenden Abschnitten 3.1 bis 3.3 werden die objektiven Interviewverfahren, die Fragebogenverfahren und die Kombinationen mehrerer Erhebungsmethoden und Weisheitskategorien vorgestellt. Es wird jeweils die Definition von Weisheit, der theoretische Hintergrund der Verfahren, sowie deren Administration beschrieben, bevor auf Kritik zu dem jeweiligen Verfahren eingegangen wird. Die im empirischen Teil dieser Arbeit verwendeten Instrumente werden im Methodenteil näher beschrieben. In Abschnitt 3.4 wird zum Zwecke der Systematisierung der Testverfahren mittels wichtiger Kategorien der Weisheitstheorien das Klassifikationsmodell in Tabelle 3.1 eingeführt.

3.1 Leistungsmessung

Zur Beurteilung der Weisheit von Produkten eignet sich die Leistungsmessung. Für eine derartige Messung ist es notwendig, das zu beurteilende Konstrukt reliabel und störungsfrei in einer spezifischen Situation zu erfassen, während Konkurrenzkonstrukte möglichst auszuschließen sind. Vorteile solch verzerrungsfreier Leistungsmessungen in Objektivität, Reliabilität und Validität gegenüber Persönlichkeitsinventaren sind jedoch mit dem höheren zeitlichen und ressourcenbezogenem Aufwand in Durchführung und Auswertung konfundiert (Kunzmann & Baltes, 2005).

Instrumente, die empirisch-experimentelle Leistungsmessungen von weisen Produkten unternehmen, sind in der Allgemeinen Weisheit nach dem Berliner Weisheitsmodell und in der Messung der Selbstbezogenen Weisheit zu finden (Staudinger et al. 2005). Allgemeine Weisheit beinhaltet Erkenntnisse über das Leben im Allgemeinen, ohne Betroffenheit der eigenen Person, während sich Selbstbezogene Weisheit mit den Einsichten einer Person über das eigene Leben beschäftigt (Staudinger, 2005).

Beide Instrumente stammen von den Hauptvertretern der expertisenzentrierten, analytischen Ausrichtung der Weisheitsforschung und entsprechen sich in Durchführung und Bewertung. Die Probanden werden aufgefordert, mittels der Technik des lauten Denkens auf die Situation des jeweiligen Szenarios zu antworten, während sie alleine sind. Die Antworten werden per Tonband aufgezeichnet und nach wissensbezogenen Basiskriterien und kontextbezogenen Metakriterien gescort. Szenarien die schwierige Lebenssituationen fiktiver Personen beschreiben, dienen der Messung der Allgemeinen Weisheit. Fragen der Introspektion in die eigenen Gewohnheiten als Freund oder im Kontakt mit Fremden, beziehen sich auf die Selbstbezogene Weisheit. Beide Verfahren sind konstrukt- und kriteriumsvalidiert. Für Allgemeine Weisheit sind Validitätsbelege in einer großen Anzahl von Studien zu finden, während Befunde zur Selbstbezogenen Weisheit bisher lediglich in einer Studie vorliegen (Mickler, 2004).

Zu diskutieren bleibt, ob diese gemessene Leistung wirklich charakteristisch für die Weisheitsleistung der Person ist. Leistungsmessungen sind klassischer Weise darauf ausgerichtet, die Höchstleistungen, also die bestmögliche Performanz der Probanden zu erfassen. Die Weisheitsforschung macht hier jedoch eine Ausnahme. Baltes und Smith (1990) streben nach der Messung der typischen Performanz der Teilnehmer. Die gemessene Leistung entspricht ihrer Meinung nach der im Alltag anzutreffenden Leistung der Probanden. Die externe Validität dieser Annahme ist jedoch nicht empirisch geprüft. Es wäre ebenso möglich, dass Probanden bei der Bewältigung von Problemen in ihrem persönlichen Umfeld besondere Be-

mühungen unternehmen, um eine bestmögliche, weise, Lösung zu finden oder zumindest zu erwägen. Um diesen Bedenken Rechnung zu tragen, wären Kontrollgruppen einzuführen, in der die persönliche Relevanz erhöht wird, und in der eine möglichst weise Antwort gefordert wird. Bisher gibt es lediglich Hinweise darüber, wie weit die vermeintlich typische Leistung von der bestmöglichen Leistung der Person abweicht. In den Studien zur Allgemeinen Weisheit zeigten erleichternde Untersuchungsbedingungen sowie Interaktion mit anderen, dass in einem gewissen Rahmen Potential zur Verbesserung der Leistung vorhanden ist (Smith, Staudinger & Baltes, 1994; Staudinger, 1996; Staudinger & Baltes, 1996).

3.1.1 Allgemeine Weisheit

Die Forscher des Max-Planck Institutes für Bildungsforschung in Berlin entwickelten und erforschten das Berliner Weisheitsmodell zur Messung von Allgemeiner Weisheit in Studien von 1990 bis 2004. Die Forschungsmotivation der Berliner Gruppe entstammt dem Interesse, positive Aspekte des Alterns zu identifizieren. Weisheit wurde daher als mögliche Stärke älterer Personen ausgewählt und erforscht. Ihre analytische Weisheitstheorie, die Weisheit als Ideal von Geist und Werten ansieht, erklärt, dass Menschen lediglich schwache Träger von weisheitsbezogenem Wissen sind (Baltes & Kunzmann, 2004).

In ihrer Forschung streben Baltes und Smith (1990) nach hoher Kohärenz mit dem kulturellen Phänomen der Weisheit. Weitere Wurzeln liegen in der Entwicklungspsychologie der Lebensspanne und vor allem im Expertisekonzept der Kognitionspsychologie. Bezüge bestehen zu der neopiagetischen Tradition der postformalen Operationen und dem dialektischen Denken. Postformales bzw. dialektisches Denken ist charakterisiert durch die Fähigkeit, Wahrheitskriterien vor dem Hintergrund pluralistischer Perspektiven zu entwickeln.

Allgemeine Weisheit wird definiert als Expertise im der fundamentalen Pragmatik des Lebens, die Denken, Urteilen und Handeln vereint, sowie Wissen und Werte, Geist und Charakter. Die fünf Weisheitskriterien des Allgemeinen Weisheitsparadigmas sind (1) Faktisches Wissen und (2) Prozedurales Lebenswissen, (3) Lebensspannen-Kontextualismus, (4) Wert-Relativismus, sowie (5) Erkennung und Umgang mit Ungewissheit (Staudinger & Baltes, 1996). Die beiden Basiskriterien: Faktuelles und Prozedurales Wissen über die Pragmatismen des Lebens entsprechen Einteilungen von Expertenwissen. Das Basiskriterium Faktuelles Wissen berücksichtigt generelles und spezifisches Wissen über menschliche Entwicklung, soziale Beziehungen und Normen, Lebensprobleme, die Koordination von Bedürfnissen und Wohlbefinden sowie Breite und Tiefe der Problembearbeitung im jeweiligen Produkt. Das Basiskriterium Prozedurales Wissen beinhaltet die Verwendung von Strategien und Heuristi-

ken bezüglich Beratung, Entscheidungsfindung, Lebensplanung und Schadensbegrenzung. Die drei weisheitsspezifischen Meta-Kriterien ermöglichen eine elaborierte Analyse der jeweiligen Situation: Wissen über die Kontextbezogenheit des Lebens, Werterelativismus und Bewusstsein und Management von Ungewissheiten. Die Metakriterien Kontextbezogenheit und Werterelativismus erlauben es, durch Relativierung der Ist-Situation und Erweiterung des herangezogenen Kontextes neue Blickwinkel in die Problembetrachtung mit aufzunehmen. Dazu zählen beispielsweise kulturelle Unterschiede von Lebenskontexten und Werten, die in eine Evaluation eines Problems zu integrieren sind. Ungewissheit, das dritte Metakriterium, diskutiert die Begrenztheit des verfügbaren Wissens und zeichnet sich durch elaborierte Problemerkennung und Problemlösung aus.

Zur Messung werden Vignetten verwendet, die Dilemmata der Bereiche existentielle Lebensbewältigung, Lebensplanung und Lebensrückblick grob skizzieren (Baltes et al., 1995). Die beiden Vignetten „Jemand erhält einen Telefonanruf von einem guten Freund. Dieser sagt, er könne nicht mehr weiter, er werde sich das Leben nehmen. Was könnte man/die Person in einer derartigen Situation bedenken und tun? (Staudinger & Baltes, 1996b, S.63-64).“ oder „A 14-year-old girl wants to move away from home right away. What could one / she consider and do? (Kunzmann & Baltes, 2003, p. 1108)“ zeigen dies. Wegen des existentiellen Charakters der ersten Vignette, wird diese Lebensmanagementaufgabe als bestes Exemplar zur Weisheitsmessung angesehen, die zusätzlich die größere Eigenvarianz aufweist (Maercker, Böhmig-Krumhaar & Staudinger, 1998; Staudinger et al., 1994).

Das Berliner Weisheitsmodell zur Messung Allgemeiner Weisheit wurde mehrfach in seiner Fünf-Faktorenstruktur bestätigt, inhaltlich validiert und beweist gute Interraterreliabilitäten der Beurteilung, sowie Interne Konsistenz der einzelnen Szenarien (Baltes & Kunzmann, 2004; Baltes et al. 1995; Kunzmann & Baltes, 2003; Smith & Baltes, 1990; Smith et al., 1994; Staudinger & Baltes, 1996b; Staudinger et al., 1997; Staudinger et al., 1998; Staudinger, Smith & Baltes, 1992). Diese Studienergebnisse werden in Kapitel 4 geschildert.

Obwohl das Allgemeine Weisheitsinterview die beste empirische Absicherung genießt, sind einzelne Kritikpunkte zu nennen. Im Folgenden wird zunächst die Kritik von Ardelts beschrieben, die sich als die Hauptkritikerin des Berliner Weisheitsmodells etablierte, bevor auf Kritik anderer Weisheitsforscher eingegangen wird.

Ardelts (2004a) Kritikpunkte an den Vignetten des Modells beziehen sich auf deren externe Validität. Die sehr kurzen Vignetten führten zu einem Mangel an persönlicher Relevanz und Beurteilungsgrundlage. Nach Ardelts entsteht eine weise Antwort umso wahrscheinlicher, je mehr inhaltliche Informationen bezüglich der Person und der jeweiligen Situation

den Probanden vorliegen. Reale Situationen, die weises Handeln von den Probanden fordern, zeigen diese Eigenschaften, viele Hintergrundinformationen und persönliche Relevanz.

Bezüglich der Kriteriumsvalidität des Weisheitsmodells schließt Ardelt (2004b), dass das Berliner Weisheitsmodell nicht Weisheit, sondern Expertise in Fundamentaler Lebenspragmatik erfasse. Die Ergebnisse fehlender Überlegenheit von Personen die für ihre Weisheit nominiert wurden gegenüber klinischen Psychotherapeuten zeigen dies genauso wie das generell schlechte Abschneiden dieser Gruppe auf der verwendeten Skala. Nur drei der 22 Nominierten erreichten einen Testwert oberhalb des Mittelpunktes der siebenstufigen Weisheitsskala, wobei eine Antwort erst ab einem Punktwert von 5 als weise klassifiziert wird (Baltes et al. 1995; Baltes & Staudinger, 2000). Die niedrigen Werte dieser Teilnehmergruppe überraschen vor allem vor dem Hintergrund des aufwendigen Delphi-Nominierungsprozederes. Eine ausführliche Beschreibung der jeweiligen Studienergebnisse findet sich in den Abschnitten zur Kriteriumsvalidität und zum Einfluss beruflicher Spezialisierung in Kapitel 4.

Websters (2003) Kritik am diesem Verfahren beruht im Mangel der empirischen Absicherung der Ergebnisse gegen Einflüsse *sozial* erwünschter Antworttendenzen. Eine Erklärung für das gute Abschneiden der klinischen Psychotherapeuten sieht er in darin begründet, dass diese die Beurteilungskriterien errahnen und daher bessere Ergebnisse liefern. Brown und Greene (2006) fordern empirische Belege dafür, dass Personen mit hohem weisheitsbezogenem Wissen, dieses auch in Handlungen übertragen.

3.1.2 Selbstbezogene Weisheit

Selbstbezogene Weisheit ist eine Adaption des Modells zur Allgemeinen Weisheit auf selbstbezogene Themen und zeigt daher große inhaltliche und konzeptionelle Ähnlichkeit zum Berliner Paradigma (Staudinger et al. 2005; Mickler, 2004). Selbstbezogene Weisheit, als hohe Stufe der Persönlichkeitsreife verstanden, stellt nach Staudinger et al. (2005) die Vereinigung von Reifetheorien der Persönlichkeitsperspektive und der Expertiseorientierung in der Theorie zur Allgemeinen Weisheit dar, zu denen das Berliner Paradigma und Sternbergs Balance Theorie (1998) gezählt werden können.

Selbstbezogene Weisheit versteht sich daher als Einsicht in das eigene Leben und die eigene Person. Eine Messung erfolgt über mündliche Produkte, die nicht mit konkretem Verhalten gleichzusetzen sind. Selbstbezogene Weisheit ist definiert durch fünf Kriterien: Zwei Basiskriterien (1) Selbstwissen und (2) Wachstums- und Bewältigungswissen, sowie drei Metakriterien (3) Zusammenhangswissen, (4) Selbstrelativierung und (5) Ambiguitätstoleranz. Die Kriterien der Selbstbezogenen Weisheit sind an denen der Allgemeinen Weisheit orien-

tiert, durch die Einbeziehung von Theorien reifer Persönlichkeit ist der Schwerpunkt eines jeden Kriteriums auf die jeweilige Person selbst ausgerichtet. Nähere Ausführungen gibt Mickler (2004).

Viele Kritikpunkte des Berliner Paradigmas lassen sich auf Selbstbezogene Weisheit übertragen. Die Beurteilungskriterien, die Durchführung und Auswertung sind weitestgehend identisch. Der zentrale Unterschied beider Instrumente liegt in der Fragestellung der Selbstbezogenen Weisheit: „Wie bin ich als Freund/ Freundin? (Mickler, 2004, S.158).“ Hier ist im Gegensatz zur Allgemeinen Weisheit selbstbezug gegeben, jedoch mangelt es der Frage an Präzisierung und Brisanz, um die Selbstoffenbarung der Probanden nicht zu gefährden. Ob damit wirklich eine weisheitsrelevante Thematik ausgewählt wurde und gleichzeitig genügend Offenheit von den Probanden in der Bearbeitung erreicht wurde, bleibt anzuzweifeln.

3.2 Selbsteinschätzungsfragebögen

Interviewverfahren stellen hohe Anforderungen an zeitliche Ressourcen aller Beteiligten. Eine ökonomischere Methode ist dagegen die schriftliche Erhebung von Daten. In wie fern die Methode der Fragebogenmessung für das Konstrukt Weisheit angemessen ist, hängt vor allem von der konzeptionellen Fassung des Konstruktes ab. „Although it might be difficult or even impossible to measure wisdom per se through a standardized self-administered questionnaire, it is hypothesized that wisdom can be assessed indirectly through indicators that are essential elements of the latent variable wisdom (Ardelt, 2003, p. 276).“

Variablen, die von den Theoretikern der bisher vorgestellten Interviewverfahren, als Kovariaten oder Vorläufer benannt wurden, bilden in den indirekten Fragebogenmethoden die indikativen Variablen für Weisheit der Person (Staudinger & Baltes, 1996b). In Micklers (2004) Interviewverfahren zur Produktmessung von Selbstbezogener Weisheit, werden beispielsweise Lebenserfahrung, Reminiszenz und Reflektion als Vorläufer von Weisheit bezeichnet, während Webster (2003) diese als Dimensionen des multidimensionalen Modells zur Beschreibung von selbstbezogener Weisheit einer weisen Person nutzt. Wie nicht anders zu vermuten, sind die den einzelnen Fragebogen zu Grunde liegenden Definitionen von Weisheit nicht deckungsgleich. Die Zuordnung zu den in Abschnitt 2.3.2 vorgestellten Kategoriensystemen erleichtert es, konzeptionelle Ähnlichkeiten und Unterschiede der einzelnen Fragebögen zu erkennen und diese beurteilen zu können.

Fragebogeninstrumente zur Erfassung von Weisheit sind Ardelts (2003) 3d-WS, Websters (2003, 2006) SAWS, Takahashi und Overtons (2002) Messung von Synthetischer

Weisheit, sowie Brown und Greenes (2006) Weisheitsentwicklungsskala. Bisher ist die Validierung der Fragebogeninstrumente noch nicht so weit fortgeschritten, wie die Validierung der Interviewverfahren.

Die beiden in dieser Studie verwendeten selbstadministrativen Fragebögen zur Erfassung selbstbezogener Weisheit, der 3d-WS und der SAWS werden zu Beginn beschrieben. Die Instrumente von Brown und Greene sowie von Takahashi und Overton werden im empirischen Teil nicht verwendet und daher nur knapp in den Abschnitten 3.2.3 und 3.2.4 eingeführt. Diese Auslassung ist in ihrer Länge, der fraglichen Übertragbarkeit auf erwachsene Stichproben, oder der Instrumentauswahl begründet. Für detaillierte Informationen zu beiden Tests wird auf die Publikationen von Brown und Greene (2006) und Takahashi und Bordia (2000), sowie Takahashi und Overton (2002) verwiesen.

3.2.1 3dimensionale-Weisheitsskala (3d-WS)

Ardelts Forschungsarbeiten behandeln Weisheit als Lebensziel des erfolgreichen Alterns (1995), und in Verbindung zu Lebenszufriedenheit (1997, 1998, 2000a, 2000b). Menschen erreichen nach Ardel (2004a) weder das Weisheitsideal, noch zeigen sie in verschiedenen Bereichen und Zeitpunkten die gleiche Menge Weisheit. Mit der 3dimensionalen Weisheitsskala (3d-WS) erforscht Ardel daher, wie Nahe Menschen diesem Ideal kommen (2003).

Der 3d-WS als Fragebogen Selbstbezogener Weisheit erfasst Weisheit indirekt über Persönlichkeitsvariablen. Im Gegensatz zu den analytisch ausgerichteten Interviewverfahren, folgt Ardelts 3d-WS der östlichen, integrierten Weisheitstradition. Die drei Dimensionen (1) Kognitiv, (2) Reflektiv und (3) Affektiv des 3d-WS wählt Ardel entsprechend zu Clayton und Birrens (1980) Weisheitsaspekten. Für Ardel sind diese Dimensionen notwendige und hinreichende Bestimmungsstücke der Weisheit, die sich über den gemeinsamen latenten Faktor der Dimensionen bestimmt. Um einer Person Weisheit zuzusprechen, muss daher jede der drei Komponenten in hohem Maße ausgeprägt, auftreten (Ardelt 2003; Clayton & Birren, 1980). Der 3d-WS bildet mit 39 Items somit ein sparsames Instrument zur Weisheitsmessung.

Die (1) Kognitive Dimension des 3d-WS erfasst die Fähigkeit, die komplexen, tieferen Bedeutungen und Gegensätze des inter- und intrapersonellen Lebens zu erkennen (Blanchard-Fields & Norris, 1995, nach Ardel, 2004b; Chandler & Holliday, 1990; Kekes, 1983; Sternberg, 1990). Dazu gehören die positiven sowie die negativen Eigenschaften der menschlichen Natur, die Begrenztheit von Wissen und die Ungewissheit zukünftiger Entwicklungen (Ardelt, 2004b). Ardel grenzt diese kognitive Fähigkeit der Elaboration von der Fähigkeit zur Perspektivenübernahme ab, die durch die (2) Reflektive Dimension beschrieben wird (Chandler

& Holliday, 1990; Clayton 1982; Csikszentmihalyi & Rathunde 1990; Kramer 1990; Orwoll & Achenbaum 1993; Rathunde 1995; Taranto 1989). Die (3) Affektive Dimension der Sympathie und des Mitgefühls entwickelt sich, wenn aus den Fähigkeiten der beiden anderen Dimensionen Einsicht in das Selbst und in Andere entsteht (Csikszentmihalyi & Rathunde 1990; Levitt 1999; Pascual-Leone 1990). Dazu ist die Überwindung von Subjektivität, Ichbezogenheit und Voreingenommenheit wichtig (Ardelt, 2004b).

Der 3d-WS (Ardelt, 2003) hat gute Retest-Reliabilität bei akzeptabler bis guter interner Konsistenz. Die Inhaltsvalidität wurde durch Übereinstimmung der Weisheitsdefinitionen von Laien mit den Dimensionen des 3d-WS bestimmt. 63,5% nennen die Affektive Dimension, 53,9% die Kognitive Dimension sowie 25,5% die Reflektive Dimension, was Ardelt als zufriedenstellend beurteilt. Im Zuge der Kriteriumsvalidierung erreichte die Nominierten-gruppe signifikant bessere Werte als deren Nominatoren. Ebenfalls gute Werte erzielt die Konvergente und Diskriminante Validierung. Mittlere Korrelationen finden sich zu Sinn im Leben und Umweltkontrolle, entsprechende negativ korrelieren Depressivität und Angst vor dem Tod. Unkorreliert ist der 3-WS mit Alter, Geschlecht, ethnologischer Abstammung und Einkommen. Der 3d-WS zeigt unter Verwendung des Inventory of Desirable Responding (Paulhus, 1991; Paulhus & John, 1998) keine Einflüsse von Sozialer Erwünschtheit.

Staudinger et al. (2005) kritisieren am 3d-WS die Korrelation zur Umweltkontrollskala von Pearlin und Schooler (1978), dass dieser eher Adaptivität als Weisheit messe. In diesen Diskurs wurde mittels Abschnitt 2.1.1 über die Trennung des Reifekonzeptes Adaptivität von Weisheit eingeführt. Eine empirische Klärung steht bisher noch aus.

In den Ausführungen zur Fragebogenentwicklung bleibt nach Brown und Greene (2006) unklar, ob Ardelt unterschiedliche Stichproben für die Auswahl und Überprüfung der Items gewählt habe. Wäre nur eine Stichprobe verwendet worden, so sei ein guter Fit der Faktorenstruktur nichts Überraschendes.

3.2.2 Selbstadministrative Weisheitsskala (SAWS)

Webster entwickelt (2003) und verbessert (2007) die Selbstadministrative Weisheitsskala (SAWS). Die SAWS ist den synthetischen, selbstbezogenen Testverfahren zuzuordnen. Weisheit zeige sich in der jeweiligen Intention, Kompetenz und Anwendung von Erkenntnissen aus kritischen Lebensereignissen (Webster 2003, 2007). Weisheit steht daher für Selbstaktualisierung, optimales Wachstum und die Erschließung eigener und fremder Potentiale.

Webster konzipiert Weisheit entsprechend einer breiten Forscheröffentlichkeit als multidimensionales Konstrukt (Ardelt, 1997; Baltes & Staudinger, 2000; Birren & Fischer,

1990; Clayton & Birren 1980; Chandler & Holliday 1990; Labouvie-Vief, 1990; Orwoll & Perlmutter, 1990; Sternberg, 1998; Taranto, 1989; Wink & Helson, 1997). Weisheit wird im SAWS durch eine konjunktive Verknüpfung determinierender Dimensionen erfasst, wenn auf allen Dimensionen sehr hohe Wertungen erzielt wurden. Mittels Literaturstudium extrahiert Webster fünf Dimensionen der SAWS: (1) kritische Lebensereignisse, (2) Offenheit, (3) Emotionsregulationsfähigkeit, (4) Humor und (5) Reminiszenz & Reflektivität. Nach Webster (2003) bilden die fünf Dimensionen einen unvollständigen Cluster interkorrelierter Weisheitskomponenten, die nun einzeln vorgestellt werden.

Die erste Dimension, (1) kritische Lebensereignisse, erfasst die Fähigkeit diese in den eigenen Erfahrungsschatz zu integrieren und aus ihnen zu lernen. (2) Offenheit steht für die Fähigkeit alternative Sichtweisen, Informationen und Lösungen zu suchen und zu erkennen, die auch die eigene Psyche betreffen. Die Dimension (3) Emotionsregulationskompetenz, ist nach Rogers (1961) eine Fähigkeit der mentalen Gesundheit und Persönlichkeitsreife, aber auch der Weisheit (Ardelt, 1997, 2003; Birren & Fischer, 1990; Clayton & Birren 1980; Chandler & Holliday, 1990; Kramer 1990; Orwoll & Perlmutter 1990; Sternberg, 1998). Emotionsregulationskompetenz bezeichnet hier sowohl das Management als auch die Sensitivität für Klassen, Nuancen und Verbindungen des Spektrums menschlicher Emotionen. (4) Humor ist in Weisheitstheorien selten, obwohl dessen Stellenwert für Persönlichkeitsreife und Coping bekannt ist (Taranto, 1989). Humor, als reifer Abwehrmechanismus zur Stressreduktion operationalisiert und für die Förderung von prosozialem Bindungsverhalten angewendet, wird als Zeichen für Weisheit verstanden. (5) Reminiszenz und Reflektivität von schwierigen Lebensereignissen zur Entwicklung von Weisheit wird allgemein anerkannt (Birren & Svensson, 2005). Die Auseinandersetzung mit Erinnerungen, Introspektionen und erlangten Einsichten ermöglicht deren Integration in einen größeren Rahmen, was grundlegend für die weitere Reifung und die Bewältigung zukünftiger Herausforderungen ist.

Sehr hohe Diskriminante Validität zeigt der SAWS zu dem Laienverständnis von Weisheit und Torheit. Belege für Konvergente Validität finden sich moderaten Korrelationen zu Eriksons Wachstumszielen der Generativität und Ego-Integrität (Webster, 2003).

Brown und Greene (2006) kritisieren, dass Webster die Skala offenbar ohne empirische Analyse oder Expertenbefragung zusammengestellt habe. Ebenso seien die Korrelationen zu Generativität und Ego-Integrität des Modells von Erikson zu gering. In der Kritik von Brown und Greene steht auch die Facette Humor als Dimension von Weisheit. Sie sei theoretisch nur wenig begründet und würde auch in Laientheorien nicht genannt. Zusätzlich wies die niedrige Faktorenladung von Humor, auf deren Randposition im SAWS hin.

3.2.3 Weisheitsentwicklungsskala

Brown (2004; Brown, Stevens, Trojano & Schneider, 2002) stellt mittels der explorativen Methode Grounded Theory ein Modell zur Weisheitsentwicklung bei Collegestudenten auf und überprüft dieses an einer großen Collegestichprobe (Brown & Green, 2006). Die Faktorenstruktur der empirischen Daten zeigt die große Übereinstimmung. 71 Items wurden den folgenden sieben Faktoren zugeordnet (1) Selbstkenntnis, (2) Emotionsregulationsfähigkeit, (3) Altruismus, (4) Inspirierende Beschäftigung, (5) Urteilsfähigkeit, (6) Lebenswissen und (7) Lebensfertigkeiten. Inhaltliche Parallelen ihrer Konzeption sehen Brown und Greene zu diversen Weisheitstheoretikern, zu denen unter anderem Ardelt, Baltes und Smith, Holliday und Chandler, Sternberg und Webster zu nennen sind.

Der Online-Test des Modells bewies große Inhaltsvalidität und moderate bis hohe Faktorenladungen. Die Indikatoren der Modellgüte erreichen aufgrund der sehr großen Stichprobe von 1188 Studenten einen mangelnden Fit. Einflüsse von sozialer Erwünschtheit auf die einzelnen Items wurden mittels der Kurzform des Marlowe-Crowne Social Desirability Scale von Reynolds (1982) ausgeschlossen.

Insgesamt konnte über die Collegestudenten mit einem Altersdurchschnitt von 21 Jahren ein breites Modell der Weisheit erschlossen werden. Aufgrund der Länge und des Fehlens einer Studie mit älteren Probanden, wurde diese Skala nicht in die vorliegende empirische Studie aufgenommen. Die Weisheitsentwicklungsskala bietet sich aber als Partner für weitere Validierungen an, da die Bandbreite des Instrumentes vielversprechend erscheint.

3.2.4 Analytische und Synthetische Weisheit

Takahashi und Overton (2002) messen Analytische und Synthetische Weisheit im kulturellen Vergleich. Die synthetische Sichtweise betont dialektisches Denken, Selbstaktualisierungstendenz, Transformation des Bewusstseins und Selbsttranszendenz, während die analytische Sichtweise, Wissen und Kognitive Komplexität betont. Ihr Ansatz besagt, dass beide Aspekte gleichberechtigt das Verständnis von Weisheit im Erwachsenenalter bilden. Nach Takahashi und Overton vereinen die Arbeiten von Achenbaum und Orwoll (1991), Ardelt (1997), Blanchard-Fields und Norris, 1995, nach Ardelt, (2004b), Csikszentmihalyi und Rathunde (1990), Kramer (1990), sowie Wink und Helson (1997) beide Arten von Weisheit.

Ihre Instrumentauswahl zur Messung von synthetischer und analytischer Weisheit ist jedoch unzureichend. Die synthetische Komponente enthielt Daten zur Selbstverwirklichten Persönlichkeit, Empathie, sozialen Beziehungen und negativen Emotionen, wobei Takahashi und Overton mit diesen Instrumenten keine umfassende Abbildung synthetischer Weisheit

einfordern. Zur Bestimmung der analytischen Komponente von Weisheit wurde lediglich das Vocabulary subset und Similarities Subset des Wechsler Adult Intelligence Scal-Revisited erhoben. Eine Gleichsetzung von Intelligenz mit analytischer Weisheit wird analytischer Weisheit nicht gerecht. Die Befunde (Takahashi, 2000; Takahashi & Bordia, 2000; Takahashi & Overton, 2002, 2005) sind daher nicht valide den einzelnen Konzeptionen von Weisheit zuzuordnen und werden deshalb nicht näher diskutiert.

3.3 Multimodale Erhebungsverfahren

Um eine valide Messung von Weisheit zu erhalten, bietet es sich an, verschiedene Erhebungsmethoden und Weisheitsbereiche abzudecken, was dem Multitrait-Multimethod-Ansatz von Campbell und Fiske (1959) entspricht. Diesen Weg wählen Helson und Srivastava, 2001, aufbauend auf Wink und Helson (1997), als auch Hartmann (2005). Beide Konzepte werden in den nachfolgenden Abschnitten beschrieben.

3.3.1 Das Weisheitskonzept von Wink & Helson sowie Helson & Srivastava

Wink und Helson (1997) trennen zwischen Praktischer Weisheit und Transzendenter Weisheit. Praktische Weisheit vereint pragmatisches und soziales Wissen, während sich Transzendente Weisheit durch metaphysisches, spirituelles, durch Reflexion erworbenes Wissen auszeichnet. Die Trennung geht auf philosophische Unterscheidungen zurück (Taranto, 1989) und separiert soziales und praktisches Wissen von transzendtem Wissen (Orwoll & Perlmutter, 1990). Beide Konzepte entstammen unterschiedlichen psychologischen Traditionen. Praktische Weisheit ist ein Feld der kognitiven Lebensspannenforschung, die Weisheit als gute Urteilsfähigkeit und Expertise in den fundamentalen Lebenspragmatiken bezeichnet (Staudinger & Baltes, 1996b). Transzendente Weisheit ist dagegen mit der psychoanalytischen Tradition von Persönlichkeitsentwicklung (Erikson, 1968; Labouvie-Vief, 1990) assoziiert, siehe Abschnitt 2.1.3.

Wink und Helson (1997) nehmen darüber hinaus an, dass sich beide Weisheitsaspekte unabhängig voneinander entwickeln können. Sie gründen ihr Genesemodell auf Achenbaum und Orwolls (1991) Trennung zwischen intrapersonalen, interpersonalen und transpersonalen Bereichen von Weisheit, die jeweils affektive, kognitive und konative Entwicklung erfordern. Praktische Weisheit entwickelt sich durch Lebensereignisse und interpersonale Entwicklung, was Empathie, Verständnis und reife Beziehungsgestaltung beinhaltet. Transzendente Weisheit erfordert Entwicklung im transpersonalen Bereich, was Selbsttranszendenz, Erken-

nen der Grenzen des Wissens und philosophisches oder spirituelles Commitment beinhaltet.

Wink und Helson (1997) verwendeten zu deren Messung zwei Methoden des Selbstberichtes. Ein Persönlichkeitsfragebogen erfasste Praktische Weisheit, während über einen narrativen Erfahrungsbericht Transzendente Weisheit erfasst wurde.

Helson und Srivastava (2001) ergänzten das Erhebungsdesign der Weisheitstests um eine Aufgabe der Allgemeinen Weisheit, die anstatt via Interview über schriftliche Befragung erhoben wird. Damit veränderte sich das Forschungsinteresse weg von der Prüfung der Eigenständigkeit der einzelnen Weisheitsbereiche (Wink & Helson, 1997), hin zur Prüfung der Existenz eines den Konstrukten übergeordneten Verständnisses von Weisheit. Die latente Modellierung des neuen übergeordneten Konstruktes als geteilte Varianz aller drei Weisheitskonzepte folgt somit einem Multitrait-Multimethod-Denken. Der gemeinsame latente Faktor Weisheit vereint die gegensätzlichen Kategorien Allgemeine und Selbstbezogene Weisheit, Analytische und Synthetische Weisheit, Persönlichkeitsmessung und leistungsbezogene Fremdbeurteilung.

Empirisch zeigte sich der latente Faktor Weisheit hoch durch das Allgemeine Weisheitsrating (AWR) definiert. Der Erklärungsbeitrag der beiden anderen Weisheitstests, Praktische Weisheitsskala (PWS) und Transzendentes Weisheitsrating (TWR), erreichte nur den moderaten Bereich (Helson & Srivastava, 2001).

Eine weitere Änderung von Helson und Srivastava (2002) betrifft das Erhebungsdesign der Praktischen Weisheit. Die dichotome Praktische Weisheitsskala (PWS) wird durch Fremdratings nach einem qualitativen Interview mit dem Proband geratet, und als Praktisches Weisheitsrating bezeichnet (PWR). Durch diese Veränderung erwiesen sich im Gegensatz zu den früheren Studien Praktische Weisheit und Transzendente Weisheit positiv korreliert. Nun folgen die Beschreibungen dieser Erhebungsinstrumente, deren Darstellung in Methodenteil 6.3.4 und 6.3.5 ergänzt wird.

3.3.1.1 Praktische Weisheitsskala (PWS)

Ein Teil des Fragebogens Adjektive Checkliste (ACL, Gough & Heilbrun, 1983), der aus 17 dichotomen Items besteht, fassen Wink und Helson (1997) als Instrument für Praktische Weisheit in der Praktischen Weisheitsskala (PWS) zusammen. Im iterativen Reviewverfahren von Persönlichkeitsforschern des Institute of Personality and Social Research wurden 13 indikative und vier kontraindikative Items ausgewählt: „clear thinking, fair minded, insightful, intelligent, interest wide, mature, realistic, reasonable, reflective, thoughtful, tolerant, understanding, wise, und immature, intolerant, reckless, und shallow (Wink & Helson, 1997, p. 3)“. Die ausgewählten Begriffe erfüllen Achenbaums und Orwolls (1991) Forderung nach

intrapersonellen und besonders interpersonalen Charakteristiken und beschreiben eine kognitive Informationsverarbeitung die klar, offen, komplex, einsichtig, realitätsorientiert, benevolent ist, und nicht durch Vorurteile oder Impulsivität gestört wird. Zusammen mit dem Selbstbewusstsein sich diese Fähigkeiten zuzuschreiben, zeugen diese Eigenschaften von Praktischer Weisheit.

Die Daten der PWS wurden retrospektiv aus bereits gesammelten Daten des ACL extrahiert und zu einem Kennwert aufsummiert. Die PWS hat akzeptable interne Konsistenzen für unterschiedliche Teilnehmergruppen und die Werte der PWS zeigen bei wiederholter Messung nach mehreren Jahrzehnten positive Korrelationen (1997, 2001).

3.3.1.2 Praktisches Weisheitsrating (PWR)

Das Praktische Weisheitsrating erfolgt über eine Fremdbeurteilung (Helson & Srivastava, 2002). Nach einem mehrstündigen qualitativen Interview mit den Probanden wird deren Praktische Weisheit durch den Interviewer geratet. Detailliertere Informationen fehlen nicht nur zur Thematik des geführten Interviews, sondern auch bezüglich der Urteilsbildung. Eine Einschätzung der Reliabilität und Validität der Fremdbeurteilungen oder der Expertise der Interviewer fehlen ebenso wie Daten zur Vergleichbarkeit der Teilnehmerwerte zwischen beiden Erhebungsmethoden.

3.3.1.3 Transzendentes Weisheitsrating (TWR)

Wink und Helson (1997) messen mittels der Frage: „Many people hope to become wiser as they grow older. Would you give an example of wisdom you have acquired and how you came by it? (1997, p. 6)“ Transzendente Weisheit. Die schriftlichen Ausführungen des ausgewählten Lebensbeispiels werden in einer einzelnen Bewertung zwischen 1 und 5 eingestuft. Positiv bewertet werden Freiheit von Selbstbezug, philosophische oder spirituelle Einsicht, sowie das Bewusstsein über die Grenzen und den kontextuellen Charakter von Wissen und Lebenssituationen.

In den empirischen Studien zeigten sich gute Interraterreliabilitäten und eine nahezu normalverteilte Streuung der Messwerte in diesem Range in der untersuchten Stichprobe. Diese Herangehensweise des Selbstberichtes zur Messung von Weisheit wird ebenfalls von anderen Weisheitsforschern verwendet (Hartmann, 2004; Bluck & Glück 2004).

3.3.1.4 Allgemeines Weisheitsrating (AWR)

Um ein übergeordnetes Konstrukt von Weisheit zu erfassen, verwenden Helson und Srivastava (2001, 2002) eine Aufgabe des Berliner Weisheitsmodells. Das Szenario berichtet über den

Telefonanruf eines guten Freundes der sich das Leben nehmen möchte. Diese Vignette wurde ausgewählt, da sie den Kern der Weisheit am besten erfasse und Weisheitsnominierte in dieser Aufgabe anderen Gruppen überlegen waren (Baltes et al., 1995). Als Erhebungsmethode wurde nicht das Interview, sondern die schriftliche Bearbeitung des Szenarios gewählt.

Die Kriterien zur Bestimmung des Weisheitsgrades der Antworten sind deren (1) Kognitive Differenziertheit, die Anwendung von (2) Prozeduralem Wissen, (3) Moralischer Komplexität und (4) Emotionalem Verständnis. (1) Kognitive Differenziertheit beleuchtet verschiedene Gründe und Kontexte einer Sachlage, die zur elaborierten Beurteilung von Lösungswegen notwendig sind. (2) Prozedurales Wissen beinhaltet Handlungswissen in verschiedenen Situationen. (3) Moralische Komplexität wird offenbart, wenn der moralische Gehalt einer Situation differenziert diskutiert wird. (5) Emotionales Verständnis zeigt sich durch emotionale Offenheit, Verständnis, Mitgefühl und Respekt gegenüber anderen (Helson, 2006, persönliche Mitteilung, 03.10. 2006). Während die ersten drei Kriterien die Basiskriterien des Berliner Weisheitsmodells und Werterelativismus relativ gut widerspiegeln, ist das Kriterium des (3) Emotionalen Verständnisses neu. Dieses Kriterium verkörpert die emotionale Komponente der Weisheit, wie sie in Orwolls Weisheitsdefinition zu finden ist, an der sich Helson und Srivastava (2002) orientieren. Durch dieses Kriterium erhält die analytisch orientierte Aufgabe einen synthetischen Aspekt.

3.3.2 Das Weisheitskonzept nach Hartmann

Hartmann (2005) kombinierte direkte und indirekte Verfahren aus einer längsschnittlichen Studie um Weisheit in Prozessen, Produkten und in Persönlichkeitseigenschaften zu messen. Sie prüft, ob Weisheit über diese drei Modalitäten konvergent erfassbar ist. Es handelt sich bei diesem Vorgehen um ein Multitrait-Multimethod-Design.

Zur Messung weiser Prozesse wurden aus den Beobachterdaten des California Q-Sort (Block, 1978; nach Hartmann, 2005) zwölf charakteristische Persönlichkeitseigenschaften nach Expertenurteil ausgewählt; ein Beispiel ist: „Is able to see to the heart of important problems (Hartmann, 2005, p. 187)“. Diese Daten haben moderate Reliabilität bei akzeptabler bis sehr guter Interner Konsistenz. Ähnlich zu TWR, diente eine offene Frage zu erlangter Weisheit der Produktmessung. Die Antworten wurden nach intrapersonalen, interpersonalen und transpersonalen Dimensionen (Achenbaum & Orwoll, 1991) von zwei Ratern geratet. Zur Beschreibung einer weisen Persönlichkeit fanden 70 Items des CPI (Gough, 1964) mit akzeptabler Interner Konsistenz und moderater Retest-Reliabilität Verwendung.

Genauere inhaltliche Angaben zu jedem der einzelnen Tests fehlen. Hartmann (2005)

berichtet lediglich, dass die einzelnen Messinstrumente konvergieren. „Wise products are generated by wise persons using wise processes (Hartmann, 2005, p. 187).“ Die Herangehensweise von Hartmann ist viel versprechend. Wegen der fehlenden Angaben und der Unzugänglichkeit der Originalstudie aus dem Jahre 2000 ist der Stellenwert der verwendeten Instrumente nicht weiter bestimmbar.

3.4 Kategoriensysteme von Weisheitstestverfahren

Um eine Zusammenschau der theoretischen Schwerpunkte der Testverfahren zu erleichtern, bedarf es theoretisch fundierter Vergleichskategorien, die eine differenzierte aber sparsame Repräsentation der Verbindung von Weisheitsdefinitionen und Weisheitsmessverfahren erlauben. Zur Systematisierung von Weisheitstestverfahren bieten sich folgende inhaltliche oder modale Kriterien an.

3.4.1 Gegenstand

Eine erste inhaltliche Trennung zwischen den Verfahren separiert je nach dem Gegenstand, der im Test beschrieben wird. Handelt es sich um den Probanden selbst, so liegt eine Messung selbstbezogener Weisheit vor. Handelt es sich um andere Personen, so liegt eine Messung Allgemeiner Weisheit vor.

Staudinger et al. (2005) schlagen vor, zwischen Selbstbezogener und Allgemeiner Weisheit zu differenzieren. Selbstbezogene Weisheit sei die höchste Form der Selbstkenntnis und Selbsteinsicht, während der Fokus der Allgemeinen Weisheit auf andere Personen gerichtet sei. Zu Verfahren der Allgemeinen Weisheit zählt das gleichnamige Interview sowie Helsons und Srivastavas Adaption der Allgemeinen Weisheit im AWR und Wink und Helsons PWS, während 3d-WS, SAWS, TWR Selbstbezogene Weisheit erfassen.

3.4.2 Orientierung

Eine weitere inhaltliche Trennung betrifft die philosophische Orientierung der verwendeten Weisheitsdefinition. Je nach Orientierung, ob kognitiv betont oder ganzheitlich, ließe sich so zwischen Analytischer Weisheit und Synthetischer Weisheit trennen.

Die Unterscheidung zwischen Analytischer und Synthetischer Weisheit ordnet der Analytischen Weisheit die Interviews zur Allgemeinen Weisheit und Selbstbezogenen Weisheit und den PWS zu. 3d-WS, SAWS, TWR und AWR repräsentieren eher synthetische, also ganzheitlich orientierte Weisheit (Takahashi, 2000).

3.4.3 Modalität

Neben diesen inhaltlichen Kategorien gibt es die Möglichkeit, die verschiedenen Weisheitstests nach ihrer Modalität zu systematisieren. Modalitäten betreffen die verschiedenen Detektionsebenen auf denen Weisheit messbar ist und in unterschiedlicher Form besteht. Eine derartige Kategorisierung gibt Hartmann (2004), indem sie (1) weise Prozesse, (2) weise Produkte und (3) weise Personen trennt.

(1) Weise Prozesse stellen das kognitive Substrat von Weisheit dar, postformales, komplexes, dialektisches Denken. Zu dessen Erfassung wird von Hartmann (2004) Loevingers projektives Verfahren zur Bestimmung der Ego-Entwicklung (WUSCT) benannt. Staudinger et al. (2005) verwenden dieses Instrument sogar als Maß für Selbstbezogene Weisheit, während andere Studien es als einen Prädiktor von Weisheit verwenden (Helson & Srivastava, 2001). Ausführungen zu diesem Konzept und Test finden sich in Abschnitt 4.2.2.2 und 6.3.6.

Von der Prozesskomponente der Weisheit ist die Produktkomponente zu unterscheiden. Als (2) weises Produkt werden Aussagen zu Fragestellungen erachtet, die eine weisheitsrelevante Thematik beinhalten. Es handelt sich hierbei nicht mehr um die Ebene der Denkprozesse, sondern um deren Resultate, Einschätzung und Handlungsanweisungen für die sich der Proband entschieden hat. Hierzu sind die Leistungsmessungen Interviewverfahren der Allgemeinen Weisheit, der Selbstbezogenen Weisheit sowie TWR und AWR zu zählen.

Klassischer Weise spiegelt die Trennung zwischen (2) weisen Produkten und (3) weisen Personen die Trennung zwischen Leistungsmessung und persönlichkeitsorientierter Bestimmung der Weisheit wieder. Mit weisen Personen beschäftigen sich Nominierungsstudien sowie Verfahren die über ausgewählte Persönlichkeitscharakteristika eine indirekte Identifikation von weisen Personen vornehmen. Solche Tests sind 3d-WS, SAWS und PWS.

Ähnlich zur Einteilung nach modalen Kriterien, bieten Staudinger und Baltes (1996b) eine inhaltlich determinierte Einteilung der Weisheitskomponenten an. Sie unterteilen die drei Bereiche (1) Postformales Denken, (2) Intelligenz sowie Expertise und 3) weise Persönlichkeit. Die Kategorisierung von Hartmann erlaubt es, Kriterien wie (2) Intelligenz und Expertise zur Beurteilung von (2) weisen Produkten heranzuziehen, ist aber nicht auf diese beschränkt. Aus theoretischer Sicht ist demnach die Systematisierung von Hartmann der von Staudinger und Baltes überlegen.

Tabelle 3.1 visualisiert die Klassifizierung nach Gegenstand, Orientierung und Modalität, ausgewählter Weisheitstestverfahren. Sie enthält die Interviewverfahren Allgemeine Weisheit und Selbstbezogene Weisheit, die Fragebogen 3d-WS und SAWS und das dreigeteilte Weisheitskonzept von Helson und Srivastava.

Tabelle 3.1. Kategorisierung ausgewählter Weisheitstestverfahren.

Messinstrumente		Kategorien		
		Modalität	Orientierung	Gegenstand
Allgemeine Weisheit	(Interview)	1	3	5
Selbstbezogene Weisheit	(Interview)	1	3	6
3-dimensionale Weisheitsskala	(3d-WS)	2	4	6
Selbstadministrative Weisheitsskala	(SAWS)	2	4	6
Praktische Weisheitsskala	(PWS)	2	3	5
Transzendentes Weisheitsrating	(TWR)	1	4	6
Allgemeines Weisheitsrating	(AWR)	1	4	5

Anmerkungen. Die Nummerierung kodiert die Kategorienbezeichnung. Modalität: 1 = Produktmessung, 2 = indirekte Erfassung über weise Persönlichkeitseigenschaften; Orientierung: 3 = analytisch, 4 = synthetisch; Gegenstand: 5 = Allgemeinheit, 6 = Selbst.

3.5 Ausblick zur Validierung der schriftlichen Weisheitstests

Die im empirischen Teil der Arbeit zu validierenden schriftlichen Fragebogen 3d-WS, SAWS und PWS sowie die narrativen Selbstauskünfte TWR und AWR zeichnen sich durch höchste Sparsamkeit im Bezug auf den Itemumfang aus. Wie beschrieben, weichen die Definition der Weisheit der zu vergleichenden Tests von einander ab. Ihr Schwerpunkt liegt jedoch auf der Synthetischen und Personenbezogenen Weisheit, siehe obige Tabelle 3.1. Zu diesen Tests gibt es erste, in einzelnen Punkten zufriedenstellende bis gute, Validierungsergebnisse, obwohl die genannten Kritikpunkte Diskussions- und Verbesserungsbedarf offenbaren. Über die Konstrukt- und Kriteriumsvalidierung sollen einige Befunde der mündlichen Erhebungsverfahren der Weisheit auf die schriftlichen übertragen, repliziert werden, um eine Einordnung dieser Tests und dieser Erhebungsmethode zu erleichtern. Im nächsten Kapitel werden die Validierungsstudien der diversen Weisheitstests inhaltlich systematisiert vorgestellt.

4 Studien zur Validität der Weisheitstestverfahren

Die Zielsetzung empirischer Forschung ist es, hoch konsistente Konstrukte zu identifizieren, die theoretisch und empirisch eindeutige Beziehungen zu anderen spezifischen Konstrukten aufweisen, um Erklärungen oder Prognosen ursächlich auf einzelne voneinander getrennte Bestimmungsstücke aufzubauen (Michel & Conrad, 1982). Die partielle Inkompatibilität von Reliabilität und Validität beschreibt den Konflikt, dass ein valides Messinstrument ein eher breites Konstruktverständnis widerspiegelt, weshalb dessen Reliabilität, vor allem die Interne Konsistenz, eingeschränkt ist (Lienert & Raatz, 1998). Zum Beleg der Validität des breiten Konstruktes Weisheit und seiner Messinstrumente, werden in diesem Kapitel die bisherigen Studien der Kriteriums- und Konstruktvalidierung vorgestellt.

Die Bedeutung und Relevanz des Konstruktes Weisheit wurden in Kapitel 2 und 3 anhand von Definitionen, des historischen Stellenwertes, des zeit- und epochenunabhängigen Charakters von Weisheit und deren hohen gesellschaftlichen Nutzen, belegt (Csikszentmihalyi & Rathunde, 1990). Nachdem die theoretische Fundierung der einzelnen Verfahren zur Messung von Weisheit beschrieben wurde, die deren inhaltliche Validität stützen, werden nun die Bereiche der Kriteriums- und Konstruktvalidität thematisiert, zur Begriffsdefinition siehe Kapitel 1. Das Konstrukt Weisheit bedarf der Einordnung und Abgrenzung gegenüber einer Reihe anderer theoretisch und empirisch gesicherter Konstrukte, wie Intelligenz und Persönlichkeitsreife, um seine Eigenständigkeit zu belegen. In empirischen Studien konnten nur wenige der zur konvergenten Validierung ausgewählten Konstrukte einen größeren Anteil von Varianz in den Weisheitstests erklären, was für die Eigenständigkeit dieses psychologischen Konstruktes spricht. In Unterkapitel 4.2 werden die herangezogenen Konstrukte zur Konstruktvalidierung inhaltlich gruppiert beschrieben und jeweils konvergenter oder diskriminanter Validität zugeordnet.

Die Kriteriumsvalidität von Weisheitsmessungen, also die Übereinstimmung mit externen Merkmalen, wurde meist durch Nominierungsverfahren zur Stichprobenziehung der Validierungsstudie abgesichert, dank dessen Hilfe sich viele der raren weisen Personen in einer solchen Stichprobe wieder finden sollten, siehe Abschnitt 4.1.1. Dieses Vorgehen entspricht einer Äußeren Validierung an einem subjektiven Kriterium. Unter der gegebenen Voraussetzung, dass Weisheit valide über einen der existierenden Tests gemessen wird, ist es ebenso möglich, die Konstruktvalidität anderer Weisheitstestverfahren an diesem Test zu prüfen, was einer Inneren Validierung entspricht.

4.1 Kriteriumsvalidierung

Als Maßnahme zur Kriteriumsvalidierung in der Weisheitsforschung wurden bisher meist Nominierungsverfahren eingesetzt. In einem Nominierungsverfahren ernennen Nominatoren in Übereinstimmung mit ihren impliziten Theorien Träger des jeweiligen Merkmals. Solange die Nominatoren keinen speziellen Expertisestatus im Hinblick auf Weisheitsdiagnostik haben, sind ihre Einschätzungen somit als Laienurteile zu bezeichnen. Nach Michel und Conrad (1982), ist eine Kriteriumsvalidierung auch unter Zuhilfenahme von Laieneinschätzungen möglich, obwohl die optimaleren Kriterien in entweder eindeutig beobachtbaren externen Kriterien zu finden wären, oder durch Expertenbeurteilung gewonnen würden, genannt Äußere Validierung. Eine weitere Möglichkeit ist die Methode der Bestimmung der Inneren Validität zur Kriteriumsvalidierung, was den Vergleich verschiedener Testverfahren desselben Konstruktes miteinander bedeutet, aber bisher nicht in der Weisheitsforschung realisiert wurde. Im Folgenden werden daher die theoretischen Grundlagen einer Nominierung und die empirischen Studien zur Kriteriumsvalidierung bei Weisheitstestverfahren vorgestellt.

Lucinda Orwoll und Marion Perlmutter (1990) postulieren drei Kriterien, welche die Zulässigkeit von Nominierungsverfahren im Bereich der Weisheitsforschung überprüfen. Das erste Kriterium prüft, ob die Probanden in der Lage sind, weise Personen zu identifizieren. Dies ist gegeben, da als abgesichert gilt, dass die Weisheit einer Person von anderen erkannt werden kann (Ardelt, 2003; Denney et al. 1995; Saum-Aldehoff, 2001). Dank der Forschung zu Laientheorien von Weisheit, die in Abschnitt 2.2 beschrieben wurde, ist gesichert, dass sich Definitionen von Weisheit in der Allgemeinbevölkerung gleichen. Zur Verbesserung der Ergebnisse schlagen Orwoll und Perlmutter (1990) vor, entweder mehrere oder geschulte Nominatoren zu befragen. Der zweite Punkt hinterfragt, ob die Nominierungen akkurat sind. Nach ihrem Vorschlag sollen Peer-Nominierungen getroffen werden, somit wäre persönliche Bekanntheit ein begünstigender Faktor für eine angemessene Beurteilung. Das dritte Kriterium stellt die Frage der Findung einer angemessenen Kontrollgruppe zu den Nominierungen. Die Nominatoren wenden für die Nominierung ihr persönliches Modell von mit Weisheit zusammenhängenden Konstrukten und mit Weisheit unvereinbaren Konstrukten an. Weisheitsnominierungen werden somit auch aufgrund von Variablen impliziter Theorien erstellt, die in empirischen Studien irrelevant für Weisheit waren, dazu gehören höheres Alter, höhere Bildung und männliches Geschlecht. Diese Informationen sollten laut Orwoll und Perlmutter (1990) zur Bestimmung einer Kontrollgruppe für Nominierte Beachtung finden.

Aus Studien zu den impliziten Weisheitstheorien von Laien können verschiedene Charakteristika von Nominierten extrahiert werden, siehe Abschnitt 2.2.2. Perlmutter, Adams,

Nyquist und Kaplan (1998, nach Orwoll und Perlmutter, 1990) berichten, dass 78% der Studienteilnehmer Weisheit mit Alter in Verbindung bringen, und die meisten Nominierten 50 Jahre und älter waren. Interessanterweise stieg das Alter der Nominierten mit dem Alter des Nominatoren, was ebenfalls in der Studie von Denney, Dew und Kroupa (1995) belegt wurde. Das Durchschnittsalter von Nominierten liegt in den Studien bei etwa 60 Jahren, berichten Jason et al. (2001) und Baltes et al. (1995). Obwohl nach Perlmutter nur wenige Menschen denken, Weisheit habe etwas mit Geschlecht zu tun, so werden doch jeweils etwa 1,5 mal mehr Männer als Frauen nominiert (Jason et al. 2001, Denney et al. 1995, Sowarka, 1989). Die Nominatoren in der Studie von Jason et al. (2001), benennen ihre nominierten Weisen als Führungspersonen, einsichtig und spirituell. Sie unterstützten die Nominatoren bei deren Lebenszielfindung oder in deren Werteorientierung.

Aus Studien zur expliziten Weisheitsforschung von Ardel (2003) und Baltes et al. (1995) ist belegt, dass die Nominierten im Vergleich zu Kontrollgruppen im Durchschnitt höhere Werte erreichen. Einschränkend muss jedoch gesagt werden, dass nicht alle Nominierten als weise diagnostiziert wurden. Die Rekrutierung der Nominatoren und das Verfahren der Nominierung unterscheiden sich zwischen den Forschern. Zunächst wird das aufwendigere, mehrstufige Nominierungsverfahren von Baltes et al. (1995) vorgestellt.

Baltes et al. (1995) stellten ein Delphipanel aus Journalisten zusammen, die für ihre Teilnahme als Nominatoren am Nominierungsprozess bezahlt wurden. Die vielfältigen ökonomischen und verfahrenstechnischen Vorteile von Delphiverfahren werden in Rowe und Wright (1999), sowie Häder und Häder (1998) beschrieben. Zur Absicherung einer akkuraten Beschreibung wählen Baltes et al. journalistische Vertreter unterschiedlicher Medien und politischer Richtungen aus Berlin aus, da diese mit Persönlichkeiten des öffentlichen Lebens im Berliner Raum genügend vertraut sind, um deren Weisheit zu beurteilen, auf die aus ethischen Gründen die Zielgruppe beschränkt wurde. Aus den Antworten der Frage, wen sie als Weisen bezeichnen würden, wurde nach der ersten Befragungsrunde eine Gesamtliste mit den Nominierten der einzelnen Journalisten erstellt. In der zweiten Runde des Delphiverfahrens beurteilten die Journalisten alle Nominierten in deren persönlichen Weisheitsgrad und ihrer persönlichen Bekanntheit zu diesen Personen. Nur jene Personen wurden abschließend in die Gruppe der Nominierten aufgenommen, die von allen ihnen persönlich bekannten Journalisten eine positive Einschätzung der persönlichen Weisheit erhielten. Aus ursprünglich gut 150 Nominierungen der ersten Liste wurden somit 14 unter 80jährige Persönlichkeiten ausgewählt, die alle an der anschließenden Studie teilgenommen haben.

Diese Studie (Baltes et al. 1995) beinhaltete vier Teilnehmergruppen, eine Nominier-

tengruppe, zwei Gruppen von Psychotherapeuten und eine Kontrollgruppe von Akademikern aus nicht sozialen Bereichen. Der Altersrange der vierzehn Nominierten reicht von 41 bis 79 Jahren bei einem Mittel von 64 Jahren. Die älteren klinischen Psychologen (Altersrange 60-84 Jahre, $M = 66$ Jahre) und älteren Kontrollakademikern (60-80 Jahre, $M = 68$ Jahre) zeigten ähnliche Werte, während die Gruppe der jungen klinischen Psychologen deutlich jünger war (25-35 Jahren, $M = 29$ Jahre).

Die Personen aus allen vier Gruppen beantworteten jeweils zwei Weisheitsaufgaben, ein Lebensplanungsproblem und die existentiellen Lebensmanagementaufgabe, in der ein guter Freund im Telefongespräch seinen Selbstmord ankündigt. Die Weisheitsleistungen der Gruppe der Nominierten waren im Durchschnitt mit den Leistungen der klinischen Psychologen vergleichbar. Die Nominierten konnten jedoch in der für Weisheit prototypischen Aufgabe des existentiellen Lebensmanagements (Maercker et al., 1998) die besten Werte erreichen. Vorteile der Nominierten gegenüber der Kontrollgruppe lagen im Mittel bei weniger als einem Skalenpunkt der siebenstufigen Skala. Es befanden sich überdurchschnittlich viele der Weisheitsnominierten in den oberen 20% der Gesamtwerte Verteilung im Vergleich zur Kontrollgruppe. Unter den Nominierten befanden sich sowohl Personen, die relativ hohe Wertungen mit fünf auf der 7-stufigen Skala erreichten, als auch solche, die unterdurchschnittliche Werte erzielten. Somit war nicht jeder Nominierte weiser als einzelne Personen der Kontrollgruppen, aber im Mittel erreichten die Nominierten höhere Werte.

Ardelt (2004b) kritisiert, dass aufgrund des aufwendigen Nominierungsverfahrens durchaus höhere Werte in der Gruppe der Nominierten zu erwarten gewesen wären. Die geringen positiven Ergebnisse des Nominierungsverfahrens verdeutlichten, dass das Berliner Modell mit den verwendeten Fragestellungen und Beurteilungskategorien nicht das Weisheitsbild der Nominatoren widerspiegeln. Eine weitere, aber konservativere Interpretationsmöglichkeit ist jene, dass das Berliner Modell sehr strenge Kriterien an eine weise Antwort anlegt und daher selbst nominierte Personen selten hohe Werte erreichen.

Das von Ardelt (2003) verwendete Verfahren zur Validierung des 3d-WS, löst die Problematik einer akkuraten Nominierung durch die von Orwoll und Perlmutter (1990) vorgeschlagene Peernominierung. Die Mitglieder von sich über längere Zeit regelmäßig treffenden Hobbygruppen für Rentner schätzten ihre Weisheit gegenseitig ein. In ihrer Studie erzielten die Nominierten signifikant bessere Werte als die nicht nominierten Probanden, jedoch ist die Differenz mit lediglich 0,14 Skalenpunkten auf der fünfstufigen Skala nur gering aber signifikant. Die Mittelwerte beider Gruppen sind als überdurchschnittlich zu bezeichnen. Die Nominierungen korrelierten hoch mit der Unterskala der Reflektiven Dimension der Weisheit, je-

doch nur moderat mit den anderen Unterskalen. Als etwas bessere Beurteilungsmethode der Weisheit der Teilnehmer stellte sich ein qualitatives, mehrstündiges Interview heraus, das lediglich an einer Teilgruppe durchgeführt wurde. Hier korrelierten alle Unterskalen hoch mit der Einschätzung des Interviewers. Ein an der gleichen Teilgruppe durchgeführtes Laienrating der Transkripte lieferte keine signifikanten Ergebnisse (Ardelt, 2003).

Als weitere Methode zur Kriteriumsvalidierung bietet es sich an, andere Weisheitstestverfahren zu verwenden. Bisher wurde dies nicht getan. In der empirischen Studie dieser Arbeit, werden demnach erstmals verschiedene Weisheitinstrumente miteinander verglichen. Da es sich in den theoretischen Konzeptionen jeweils um eigenständige Fassungen von Weisheit geht, kann nicht erwartet werden, dass die Instrumente wirklich das Selbe messen, was empirisch Korrelationen in der Höhe der Reliabilität der Messung gleichkommen würde. Da die zugrunde liegenden Weisheitstheorien nicht identisch sind, müssen deren Operationalisierungen Unterschiede aufweisen, die wiederum die mögliche Korrelation zwischen den einzelnen Testverfahren herabsetzen.

4.2 Konvergente und Diskriminante Validierung

Dank der Breite des Konstruktes Weisheit gibt es eine Vielzahl von verwandten Konstrukten zu denen empirische Beziehungen nachzuweisen sein sollten. Folgende vermutete Beziehungen von Allgemeiner Weisheit zu anderen Konstrukten verdeutlichen dies:

„For instance, exceptional understanding of self and others, together with having learned from experience, good communication skills, a commitment to personal growth, open mindedness, emotional maturity, and empathy, but also good reasoning ability, sagacity, and perspicacity have been mentioned as prerequisites and correlates of wisdom (Staudinger et al. 1998, p. 3)“.

Da es sich bei den mit Weisheit in Bezug stehenden Konstrukten um theoretisch voneinander getrennte Konzepte handelt, erklären diese nur wenig Varianz der Weisheit. Um Studienergebnisse der Weisheitstests unmittelbar zu vergleichen, müssten diese mit den gleichen konvergenten und diskriminanten Korrelaten gearbeitet haben. Da dies leider nicht der Fall ist, wird im Folgenden eine Auswahl von Konstrukten vorgestellt, die in einzelnen Studien Verwendung zur Konstruktvalidierung der Instrumente gefunden haben. Es wird dabei auf eine Gruppierung nach konvergenten oder diskriminanten Beziehungen zu Weisheit verzichtet, dafür werden die soziodemographischen Variablen, in Abschnitt 4.2.1, und kognitiven Variablen, in Abschnitt 4.2.2, sowie persönlichkeits-theoretischen Konstrukte, in Abschnitt 4.2.3, inhaltlich geschlossen präsentiert.

Einschränkend sei auf den Befund eingegangen, welcher möglicherweise die Ergeb-

nisse der Beziehung zwischen Weisheit und weiteren Konstrukten verzerrt. In den Studien konvergenter und diskriminanter Validierung von Weisheitstests erzielen nur wenige Probanden hohe Werte. So erreichen in der Studie von Smith und Baltes (1990) von 240 Teilnehmern nur elf das prototypische Niveau einer weisen Antwort. Angaben zu Beziehungen von Weisheit werden daher nicht auf den weisen Personenkreis beschränkt, sondern beziehen sich auf das gesamte Spektrum von Teilnehmern. Ob diese berichteten Zusammenhänge identisch mit den Zusammenhängen des oberen Wertebereiches gemessener Weisheit sind, wird nicht berichtet und bleibt daher offen.

4.2.1 Soziodemographische Korrelate

Alter und Lebenserfahrung sind zentrale Begriffe in Laintheorien zu Weisheit. Ob sie jedoch empirisch signifikante Korrelate sind, wird in Abschnitt 4.2.1.1 geklärt. Die berufliche Spezialisierung ist ein soziodemographisches Korrelat der persönlichen Weisheit, siehe Abschnitt 4.2.1.2. Soziale, beratende oder therapeutische Berufe erfordern es, sich mit Problemen anderer Menschen auseinanderzusetzen, was ein Anwendungsbereich der Weisheit ist.

Zu beachten ist, dass Kohorteneffekte die Zusammenhänge soziodemographischer Variablen beeinflussen. So ist beispielsweise im höheren Alter, Geschlecht mit formaler Bildung korreliert. Formale Bildung ist wiederum ein Indikator für kognitive Entwicklung und daher mit Weisheit korreliert (Sternberg, 2004b). Aus diesem Grund sollten soziodemographische Korrelate in ihrer Gesamtheit und mit Vorsicht interpretiert werden.

4.2.1.1 Geschlecht

Eine der wichtigsten soziodemographischen Variablen in der empirischen Psychosozialforschung ist das Geschlecht. Die empirischen Studien zu Allgemeiner Weisheit zeigen keine Geschlechterunterschiede (Staudinger & Baltes, 1996). In den Fragebogenstudien von Ardelt (2003) und Webster (2006) schnitten Frauen hingegen leicht, aber signifikant, besser ab als Männer. Dieser Befund wird mit Vorteilen der Frauen in den jeweils vorhandenen affektiven Facetten begründet, aber nicht näher belegt oder reproduziert (Webster, 2006). Die Millsstudie von Wink und Helson (1997) und Helson und Srivastava (2001, 2002) beinhaltete ausschließlich weibliche Teilnehmer. Zusammenfassend sind keine bedeutsamen Geschlechtsunterschiede in Weisheitsstudien zu finden.

4.2.1.2 Alter und Lebenserfahrung

Es gibt drei konkurrierende theoretische Modelle, die den Zusammenhang von Weisheit und Lebensalter thematisieren (Jordan, 2005). Das Positive Model postuliert, dass mit Alter auch

Weisheit einhergeht. Dieser erwartete positive Zusammenhang, wie er von Theoretikern und Laien angenommen wird (Birren & Fischer, 1990; Heckhausen, Dickson & Baltes, 1989; Kitchner & Brenner, 1990; Labouvie-Vief, DeVoe & Bulka, 1989; Labouvie-Vief, 1990; Orwoll & Perlmutter, 1990; Sternberg, 1986; Taranto, 1989), zeigt sich in empirischen Studien nicht. Negativ sieht das Verfallsmodell nach Meacham (1990) den Zusammenhang von Alter und Weisheit. Meacham postuliert, dass nur Kinder wirklich weise seien. Empirische Daten sprechen gegen dieses Modell, und favorisieren das dritte, das kristalline Modell. Ähnlich zur kristallinen Intelligenz, soll das individuelle Niveau von Weisheit vom frühen Erwachsenenalter bis ins hohe Alter stabil bleiben. Dies erscheint plausibel, da es über alle 533 Teilnehmer des Berliner Paradigmas keine Korrelation zwischen Alter und Weisheit für Erwachsene über 25 Jahren gibt (Baltes et al., 1995; Baltes & Staudinger, 2000; Baltes, et al., 1995; Smith et al., 1994; Smith & Baltes, 1990; Staudinger et al., 1992; Staudinger, 1999). Lediglich jenseits der 80 lässt sich ein negativer Alterstrend beobachten, der über hirnganischen Abbau begründet wird (Lindenberger & Baltes, 1997). Ardelt (2003) belegt hingegen eine negative Beziehung zwischen Alter und ihrem 3d-WS für über 51jährige.

Forschung mit Jugendlichen zeigte, dass die Hauptzeit des Zugewinns an Weisheit zwischen 15 und 25 Jahren liegt (Pasupathi, Staudinger, & Baltes, 2001). Die deutlichsten Zugewinne im Berliner Modell sind in den Metakriterien des relativierenden Denkens und dem *Denken über Ungewissheiten* nachweisbar. Nach Richardson und Pasupathi (2005) sind die dafür wichtigen Entwicklungsbereiche die Intellektuelle Kapazität, Selbstreflektiertes und deduktives Denken, die Persönlichkeits- und Moralische Entwicklung, sowie Unvoreingenommenheit und Toleranz. Ein Konstrukt, das diese Bereiche umfasst ist die Ego-Entwicklung nach Loewinger, siehe Abschnitt 4.2.2.3. In der Jugend sind intelligenzbezogene und persönlichkeitsbezogene Faktoren wichtiger für die Entwicklung von Weisheit als Variablen des Intelligenz-Persönlichkeits-Interface. Da Intelligenz und Persönlichkeitsvariablen im Erwachsenenalter große Stabilität aufweisen, sind Interfacevariablen, wie Kognitiver Denkstil, in Paragraph 4.2.2.2, entscheidender für die Weisheit einer erwachsenen Person (Pasupathi et al., 2001).

Lebenserfahrungen erscheinen für den Gewinn von Weisheit grundlegend (Clayton & Birren, 1990), aber wie dieser Zusammenhang aussieht, ist bisher nicht abschließend geklärt. Theoretisch eröffnen mehr Lebenserfahrungen die Möglichkeit zu größerer Weisheit, allerdings ist ein linearer Zusammenhang nicht zwingend. In der Studie von Staudinger et al. (1992) ist die Anzahl von Lebenserfahrungen kein Prädiktor für Weisheit. Nach Kramer (2000) regen harte Zeiten zur Reflektion über Werte und das Leben an sich an, was jedoch

nicht bedeutet, dass jeder Mensch an Weisheit gewinnt, der schwierige Zeiten erlebt hat (Ardelt, 1998). Weisheit wird durch die erfolgreiche Lösung von Krisen erlangt (Ardelt, 1997; 2000; 2004b), wobei Offenheit für neue Erfahrungen eine wichtige Rolle spielt (Ardelt, 1997; Helson & Srivastava, 2001, 2002; Wink & Helson, 1997), siehe Abschnitt 4.2.3.1. Der Umgang mit Erfahrungen erscheint somit wichtiger als die reine Erfahrung an sich und moderiert somit den Zusammenhang zwischen den Variablen Lebenserfahrung oder Alter und Weisheit.

4.2.1.3 Scheidung

Die Millsstudie von Helson und Srivastava (2002), erforscht die längsschnittliche Entwicklung von Frauen in ihrer Praktischen und Transzendentalen Weisheit. Es zeigte sich, dass geschiedene Frauen zwischen den Erhebungszeitpunkten in ihrem Alter von 27 bis 54 Jahren signifikant mehr Zuwachs an Praktischer Weisheit erlangten als Frauen in einer Ehe. Eine Erklärung dieses Befundes ist, dass negative Lebenserfahrungen, wie Beziehungsprobleme oder Scheidungen, für Weisheit förderlich sind (Birren & Fischer, 1990; Bursik, 1991; Kramer 1990; Taranto, 1989).

4.2.1.4 Berufliche Orientierung

Berufliche Praxis in der Bearbeitung menschlicher Probleme ist aus mehreren Gründen als Prädiktor für Weisheit geeignet. Erstens, durch beruflichen Umgang mit sozialen Problemen, ist es leichter möglich, Allgemeine Weisheit zu erlangen. Zweitens, könnten diese Berufe das Bedürfnis steigern, mit sich selbst im Reinen zu sein, und somit Selbstbezogener Weisheit dienlich sein. Drittens, besteht die Möglichkeit, dass solche Berufe von Menschen erwählt werden, die dank ihrer Persönlichkeit, Motive oder Einstellungen generell höhere Werte von Allgemeiner oder Selbstbezogener Weisheit aufweisen.

Empirische Studien zur Erforschung des Zusammenhangs von beruflicher Orientierung und Weisheit wurden vor allem für Allgemeine Weisheit durchgeführt. Der Einfluss eines Berufes mit Ratgeberfunktion in schwierigen Lebenslagen, wie Familienrecht, Seelsorge oder klinische Psychologie wurde Smith et al. (1994), und Staudinger et al. (1992) untersucht. Für klinische Psychologen zeigte sich, dass sie signifikant häufiger in den Top 20% der Gesamtstichprobe waren. Staudinger und Baltes (1995) stellten fest, dass berufliche Orientierung über Persönlichkeitsvariablen hinaus mit 14% an eingeständiger Varianz den größten einzelnen Erklärungsbeitrag an weisheitsbezogenem Wissen erklärte. In einer späteren Studie von Staudinger et al. (1998) erklärte die berufliche Spezialisierung in klinischer Psychologie mit 25% den größten Varianzanteil an weisheitsbezogenem Wissen auf. Altersunterschiede in weisheitsbezogenem Wissen ließen sich bei den Psychotherapeuten nicht finden, d. h. dass

mehr berufliche Erfahrung nicht mit mehr Weisheit einhergeht.

Die Befunde zur Weisheit von Psychotherapeuten (Baltes et al., 1995; Smith et al. 1994; Staudinger et al., 1992) wurden von Helson und Srivastava (2002) repliziert. Sowohl in praktischer als auch in transzendentaler Weisheit waren sie Frauen ähnlicher akademischer Bildung überlegen. Das latente Konstrukt Weisheit lies sich durch Offenheit für Erfahrungen und eine therapeutische Ausbildung vorhersagen. Die Befunde belegen, dass Psychotherapeuten eine besonders interessante Untersuchungsgruppe für Weisheitstests sind.

4.2.2 Kognitive Korrelate

Dass Kognitive Variablen in Weisheitstheorien eine bedeutende Rolle spielen, wird besonders in der theoretischen Tradition der Analytischen Weisheit offensichtlich. Kognitive Facetten der Weisheit werden von allen empirischen Weisheitsforschern angenommen (Ardelt, 2003; Baltes, & Smith, 1990; Brown & Greene 2006; Hartmann, 2000 nach Hartmann 2004; Staudinger et al., 2005; Takahashi, 2000; Webster, 2003; Wink & Helson, 1997) und sind Bestandteil von Lagentheorien zu Weisheit (Ardelt, 2003; Bluck & Glück, 2003). Gerade auf Grund der Betonung kognitiver Aspekte sollte eine Validierung von Weisheitstestverfahren die Beziehung zwischen Weisheit und Intelligenz, dem bedeutendsten Konstrukt kognitiver Kompetenz, erhellen. Clayton (1982) und Ardelt (2004a) erklären die Trennung von akademischer Intelligenz und Weisheit auf theoretischer Ebene. Auch Li und Kunzmann (2004) argumentieren, dass die Beziehungen zwischen Faktenwissen und Prozeduralen Wissen, den Basiskriterien des Berliner Weisheitsmodells, nicht alleine über klassische Intelligenzbereiche wie Abstraktionsfähigkeit oder Logisches Denken zu erklären sei. Empirisch ist die Beziehung zu Weisheit diskriminanter Art, wie in Abschnitt 4.2.2.1 näher beschrieben wird.

Auch in anderen Studien zu Kompetenzbereichen und Leistung ist Intelligenz nicht die beste Erklärungsvariable zwischen kognitiver Leistungsfähigkeit und erzielten Ergebnissen (Mackintosh, 1998). Alternative Erklärungen des Zusammenhang von kognitiver Kompetenz und Weisheit bieten Loevingers Konstrukt der Ego-Entwicklung in Abschnitt 4.2.2.2 und das Konzept der Denkstile nach Sternberg, siehe Abschnitt 4.2.2.3. Weitere kognitive Variablen, die für Weisheit indikativ seien könnten, werden in Abschnitt 4.2.2.4 erörtert.

4.2.2.1 Intelligenz

Nach Stern ist Intelligenz: „die allgemeine Fähigkeit eines Individuums, sein Denken bewusst auf neue Forderungen einzustellen; sie ist allgemeine geistige Anpassungsfähigkeit an neue Aufgaben und Bedingungen des Lebens (Stern, 1912, S. 3)“. Welcher Zusammenhang zwischen Intelligenz und Weisheit besteht, wird zuerst theoretisch dann empirisch diskutiert.

Nach Vivian Clayton (1982) und Monika Ardelt (2000b, 2003, 2004b) ist Weisheit nicht mit intellektuellem Wissen gleichzusetzen. Beide Wissensarten zeigen Unterschiede die den Inhalt und die Persönlichkeit des Trägers betreffen (Chandler & Holliday, 1990; Clayton & Birren, 1980; Csikszentmihalyi & Rathunde 1990; Sternberg, 1990). Intellektuelles Wissen ist auf die Beherrschung der Realität und deren Anforderungen ausgerichtet, während weisheitsbezogenes Wissen durch ein tieferes Verständnis zur Akzeptanz der Realität und der Begrenztheit von absolutem Wissen führt. Der Ansatz von Intellektuellem Wissen ist entdeckend, theoretisch, abstrakt und linear, während weisheitsbezogenes Wissen wiederentdeckend, angewandt, konkret und dialektisch ist. Weisheitsbezogenes Wissen ist nicht zeitgebunden, sondern universell. Es entwickelt sich auf der Basis persönlicher Erfahrungen, was somit in gewisser Weise vom Alter, Offenheit, Transzendenz und Erfahrungen der Unvorhersagbarkeit und Ambiguität determiniert wird, anstatt durch faktenorientiertes Studium und strikter wissenschaftlicher intelligenzgeleiteter Denkweise (Ardelt, 2004b). Die vorherrschende Wissensart beeinflusst auch die Persönlichkeit des Trägers (Ardelt, 2004b), der durch sein Weisheitsbezogenes Wissen seine Selbstbezogenheit überwinde und Sympathie für andere zeige, während der intellektuell Intelligente sich anderen überlegen fühle. Bei Kontrollverlusten reagiere der Weise mit Gelassenheit anstatt mit negativen Gefühlen.

Diesen Argumenten folgend, sollte Weisheit nur gering mit psychometrischer Messung von Intelligenz korrelieren. In den meisten Validierungsstudien zu Weisheit fehlen Psychometrische Intelligenztests. Im Berliner Modell wurden Klassische Intelligenztests zur Bestimmung von fluider und kristalliner Intelligenz erhoben. Der Einfluss der Intelligenzmaße ist sehr gering und circa halb so groß wie der Einfluss von Persönlichkeitsvariablen (Staudinger et al., 1997). Fluide und kristalline Intelligenz konnten nur 2% an eigenständiger Varianz von Allgemeiner Weisheit erklären. Auf bivariater Ebene wurden 16% vor allem durch die Intelligenzmaße Advanced Progressive Matrices und HAWIR-R erklärt. Die Ergebnisse des HAWIE-R als Maß für semantisches Wissen zeigen, dass keine Verzerrung der Weisheitsmessung durch Unterschiede im semantischen Wissen erfolgte. In einer weiteren Studie von Staudinger et al. (1998) war einzig der Advanced Progressive Matrices und Thrustones Letter Series Form B-L für fluide Intelligenz signifikant, jedoch in geringer Höhe mit weisheitsbezogenem Wissen korreliert. Das Instrument zur Messung von selbstbezogener Weisheit zeigte sich mit formaler Bildung und Verbaler Intelligenz leicht korreliert (Mickler, 2004). Der geringe Zusammenhang beider Größen ist möglicherweise wegen des eingeschränkten Ranges von Intelligenzwerten erklärbar, der durch die Auswahl von Teilnehmern mit höherem Bildungsniveau in den Studien des Berliner Modells verursacht wurde. Es ist daher möglich, dass

der Einfluss von Intelligenz auf Weisheit in der Population größer ist, als er in diesen Studien erscheint.

4.2.2.2 Ego-Entwicklung

Für Loevinger (1998) stellt Ego-Entwicklung das Hauptmerkmal zur Beschreibung von Personen gleich nach Intelligenz dar. Ego-Entwicklung sei der Mastertrait, der für die Organisation der Persönlichkeit verantwortlich ist, und vier Bereiche umfasst: (1) Charakterentwicklung, vor allem Impulskontrolle und moralische Orientierung, (2) Kognitiver Stil (3) Interpersonaler Stil (4) Themen bewusster Auseinandersetzung. Durch stetige Reifungsprozesse in diesen Kategorien entwickelt sich die Persönlichkeit. Loevinger (1976) entwickelt ein quasi-strukturelles Stufenmodell. Die neun Stufen repräsentieren qualitative Unterschiede und stehen Piagets kognitivem Entwicklungsmodell nahe. Der Washington University Sentence Completion Test (WUSCT) zur Messung von Ego-Entwicklung ist ein in vielen Studien validiertes und normiertes projektives Instrument, in dem über spontane Satzergänzungen auf die Persönlichkeitsreife geschlossen wird.

Mit Kramer (1990) argumentieren Wink und Helson (1997), dass Ego-Entwicklung eine Voraussetzung für Weisheit sei, da Weisheit die Urteilsfähigkeit darstelle, die mit hohen Ego-Entwicklungsstufen assoziiert wird. Persönliche Reife, gemessen mittels Ego-Entwicklungstest, zeigt einen hohen Grad intrapsychischer Differenzierung und Autonomie (Manners & Durkin, 2001) und ist unabhängig von sozialer Anpassung, Subjektivem Wohlbefinden, Lebenszufriedenheit und positivem Affekt (Helson & Wink 1987, McCrae & Costa, 1983). Ego-Entwicklung ist diskriminant von sozioökonomischen Status, kristalliner Intelligenz und der Wortflüssigkeit als Teil der Verbalen Intelligenz verschieden (Cohn & Westenberg, 2004; Labouvie-Vief & Diehl, 1998; McCrae & Costa, 1980; Newman, Tellegen, & Bouchard, 1998). 94% aller theoretischen Beziehungen des Konzepts Ego-Entwicklung bleiben nach Konstanthalten für Einflüsse von Intelligenz erhalten (Cohn & Westenberg, 2004).

In der Studie von Helson und Srivastava (2001) zeigte sich Ego-Entwicklung moderat mit Weisheit korreliert. Von Staudinger et al. (2005) wurde der Satzergänzungstest von Loevinger sogar als Maß für Selbstbezogene Weisheit bezeichnet. Wenn Ego-Entwicklung als Maß für Weisheitsmessung aufgefasst werden soll, so ist es der Typologie von Hartmann zufolge (2000, nach Hartmann, 2005) ein Instrument zur Erfassung von weisen Prozessen. Jedoch wird der WUSCT zur Messung von Ego-Entwicklung eher zur Konvergenten Validierung von Weisheitstests herangezogen, anstatt zur Kriteriumsvalidierung benutzt zu werden. Der WUSCT zeigt moderat positive Beziehungen zu Selbstbezogener Weisheit (Mickler, 2004) und Allgemeiner Weisheit (Pasupathi et al. 2001). Weitere Studien mit Loevingers

Ego-Entwicklungstest fanden in der Weisheitsforschung bisher nicht statt.

Die empirisch gefundene Höhe des Zusammenhangs legt Ego-Entwicklung eine konvergente Beziehung zu Weisheit nahe, anstatt Äquivalenz. Da bisher Werte über die Beziehungen der Weisheitstestverfahren untereinander fehlen, ist nicht auszuschließen, dass dieser Schluss falsch sein könnte.

4.2.2.3 Kognitiver Denkstil

Ein neueres Konstrukt, das als Interfacevariable den Raum zwischen Persönlichkeitsvariablen und Intelligenz füllt, ist der kognitive Denkstil. Mittels der breiten Palette an 13 Denkstilen, die Sternberg (1988, 1997) entwickelte, lassen sich bisher nicht erklärbare Zusammenhänge der Bereiche Intelligenz und Persönlichkeit quantifizieren. Nach Grigorenko und Sternberg (1995) vereinen Denkstile Komponenten aus verschiedensten Bereichen wie persönlichkeitszentrierte Ansätze, Theorien zur Mentalen Selbstführung, Kognitionszentrierte Ansätze und Aktivitätszentrierte Ansätze. Die Denkstile sind in die Bereiche Funktion, Form, Niveau, Reichweite und Neigung einteilbar. 35% Varianz des Weisheitsmaßes Allgemeine Weisheit konnte durch Kreativität und folgende fünf Denkstile erklärt werden: External, Konservativ, Oligarisch, Monarchisch, und Judikativ (Staudinger et al., 1997). Weise Denkstile berücksichtigen Prioritäten und beurteilen Handlungen nach der Nützlichkeit für das Allgemeinwohl, in dem sie sensitiv für die Bedürfnisse der Beteiligten sind.

4.2.2.4 Weitere kognitive Konstrukte

In weisheitsrelevanten Aufgaben sind schlecht definierte Probleme zu lösen, und nicht klar strukturierte Aufgaben wie in Psychometrischen Intelligenztests, weshalb eher Praktische statt Psychometrische Intelligenz mit Weisheit zusammenhängt (Dittmann-Kohli, 1984; Kitchener, 1983). Gute komplexe Problemlöser zeichnen sich durch die Überlegenheit in der Identifikation von Einflussfaktoren und deren komplexen Beziehungen aus, die es ihnen ermöglicht, die Realität in ein akkurateres Modell zu übertragen und bessere Prognosen für Zustände von dynamischen Systemen zu treffen. Eine effiziente Aktion des Problemlösers erfordert kognitive, emotionale, personale und soziale Fähigkeiten und Wissen. Weise Personen können nach Csikszentmihalyi und Nakamura (2005) komplexe Probleme besser repräsentieren und lösen.

“To the contrast the cut and dired puzzles that characterize items in an IQ test, work problems often are (a) unformulated or in need of reformulation, (b) of personal interest, (c) lacking in information necessary for solution, (d) related to everyday expertise, (e) poorly defined, (f) characterized by multiple ‘correct’ solutions, each with liabilities as well as assets, and (g) characterized by multiple methods for picking a problem solution (Sternberg, et al. 1995, p. 913)”.

Forschungsergebnisse zu erfolgreichen Managern können an dieser Stelle Hinweise auf kognitive Konstrukte liefern, die mit der komplexen Problemlösefähigkeit verwandt sind. Die

Vorteile erfolgreicher Manager sind auf Tacit Knowledge, Implizites Lernen und Praktische Intelligenz zurückzuführen (Klemp & McClelland, 1986, nach Mackintosh, 1998; Reber, Walkenfeld & Hernstadt, 1991; Sternberg, Wagner, Williams & Horvath, 1995), die unabhängig von Psychometrischer Intelligenz sind. Die Fähigkeiten erscheinen für Weisheit relevant, jedoch fehlen bisher Forschungsergebnisse in diese Richtung.

4.2.3 Persönlichkeitskorrelate

Das klassische psychometrische Instrument zur Beschreibung der Persönlichkeit ist das Fünf-Faktoren-Modell der Persönlichkeit. Es ist für die Eigenständigkeit des Konstruktes Weisheit wichtig, sich von diesem Persönlichkeitskonstrukt abzugrenzen, siehe dazu Abschnitt 4.2.3.1. Zur Auswahl von häufig verwendeten persönlichkeits-theoretischen Konstrukten in der Konstruktvalidierung von Weisheitsmaßen gehören Psychologisches Feingefühl, Abschnitt 4.2.3.2, und Persönliches Wachstum, Abschnitt 4.2.3.3. Diese Konstrukte stehen in konvergenter Beziehung zu Weisheit, während Adaptivität, auch Soziale Reife genannt, Abschnitt 4.2.3.4, auf theoretischer und empirischer Ebene diskriminante Beziehung zu Weisheit einnimmt. Da die wissenschaftliche Diskussion um Soziale Erwünschtheit die Interpretation dieser als Persönlichkeitsvariable erlaubt, die das Verhalten einzelner Teilnehmer beeinflussen kann, stellt sich die Frage wie wichtig Soziale Erwünschtheit für die Weisheitsforschung ist, siehe Abschnitt 4.2.3.5.

4.2.3.1 Fünf-Faktoren-Modell der Persönlichkeit

Das Fünf-Faktoren-Modell der Persönlichkeit ist eines der am häufigsten verwendeten Psychometrischen Maße zur Bestimmung von Charaktereigenschaften einer Person (McCrea & John, 1992; Borkenau & Ostendorf, 1989). Persönlichkeit im Erwachsenenalter in entlang von fünf sehr konstanten Persönlichkeitsdimensionen organisiert, die entweder angeboren sind, oder sich bereits früh im Leben entwickeln (McCrae & Costa, 1984). Der Ansatz umfasst Extraversion, Neurotizismus, Gewissenhaftigkeit, Verträglichkeit und Offenheit für Erfahrungen. In einer neuen Studie von Allemand, Zimprich und Hertzog (2007) bestätigte sich abermals die Stabilität der fünf Faktoren, wobei mit steigendem Alter Neurotizismus und Offenheit abnehmen.

Vor allem Offenheit wird mit Weisheit in Verbindung gebracht, da Offenheit mit Ego-Entwicklung (Costa & McCrae, 1978 aus 1980) und mit Rogers fully functioning person in Beziehung steht (1961; sieh Abschnitt 2.1.1). Offenheit wird von Kramer (2000), als wichtigste Persönlichkeitseigenschaft von Weisheit identifiziert, die Komponenten von psychologischem Feingefühl und der Neugier auf neue Erfahrungen enthält (Costa, Zonderman, &

McCrae, 1991; McCrae & Costa; 1997). Offenheit hängt mit Allgemeiner und Selbstbezogener Weisheit in gleichem Maße zusammen (Mickler, 2004). Das Berliner Weisheitsmodell belegt die positive Beziehung zu Offenheit in vielen Studien (Maercker et al. 1998; Pasupathi & Staudinger, 2001; Staudinger et al., 1997; Staudinger et al., 1998; Staudinger & Pasupathi, 2003). In einer dieser Studien konnten die Persönlichkeitsvariablen Offenheit und Psychologisches Feingefühl einen bedeutsamen Teil der Varianz von Weisheit erklären, wobei der bivariate Zusammenhang von Offenheit mit Allgemeiner Weisheit bei $r = .42$ lag (Staudinger et al., 1997). In der Studie von Helson und Srivastava (2002) zeigte sich ein positiver Zusammenhang von Offenheit für Erfahrungen, beruflicher Beschäftigung im psychotherapeutischen Bereich und Weisheit für Frauen. Websters SAWS (2003) zeichnet sich durch eine eigene Unterskala Offenheit aus.

Der Faktor Verträglichkeit hat eine leicht positive Korrelation zu Allgemeiner Weisheit, da Weisheit einen empathischen und rücksichtsvollen Umgang mit Anderen erwarten lässt. Die leicht negative Korrelation zu Neurotizismus deutet darauf hin, dass Weisheit als Vorhandensein von Seelischer Gesundheit zu verstehen ist (Kunzmann & Baltes, 2003). Zu den Faktoren Extraversion und Gewissenhaftigkeit zeigen sich keine Zusammenhänge. Selbstbezogene Weisheit zeigt ähnliche Korrelationsmuster zum Fünf-Faktoren-Modell, wie die hier dargestellten Ergebnisse für Allgemeine Weisheit (Mickler, 2004).

4.2.3.2 Psychologisches Feingefühl

Eine Skala für Psychologisches Feingefühl misst Interesse an sowie Reaktionsfähigkeit auf innere Bedürfnisse, Motive und Erfahrungen bei sich und anderen (Gough, 1964). Attribute des Konstruktes Psychologisches Feingefühl werden als Aspekte reifer Persönlichkeit bezeichnet, wozu beispielsweise sensibel und empathisch, außergewöhnliches Verständnis und Intuition zählen (z.B. Clayton & Birren, 1980; Holliday & Chandler, 1986).

Empirisch zeigen sich Zusammenhänge zwischen Psychologischem Feingefühl, Ego-Entwicklung (Helson & Roberts, 1994; Helson & Wink, 1987; Westenberg & Block, 1993) und Wink und Helsons (1997) Praktischer und Transzendentaler Weisheit. Die bivariate Korrelationen zum Allgemeiner Weisheit liegt bei $r = .28$ (Pasupathi & Staudinger, 2001; Staudinger et al. 1997). In der Validierungsstudie zu Selbstbezogener Weisheit (Mickler, 2004) war Psychologisches Feingefühl als Reifeindikator jedoch nicht signifikant mit dieser korreliert. Der fehlende Zusammenhang wird dadurch erklärt, dass es schwieriger sei, Einsichten in die eigene Person zu erlangen als in allgemeine psychologische Phänomene.

4.2.3.3 Persönliches Wachstum

Positive lebenslange Persönlichkeitsentwicklung hin zu Persönlichkeitsreife ist typisch und normativ (Maslow, 1999; Vaillant, 1977). Persönlichkeitsreife wird als ein Vorläufer von Weisheit betrachtet, siehe Kapitel 2.1.1. Ein Konstrukt zur Messung von Persönlichkeitsreife, das in der Weisheitsforschung Verwendung gefunden hat, ist Ryffs Modell des Psychologischen Wohlbefindens (PWB) (Ryff 1989; Ryff & Keyes, 1995).

Ihre sechs Unterskalen stellen Adaptive und Wachstumsorientierte Reifeaspekte dar. Die drei Facetten Autonomie, Gute Beziehungen, und Lebensziel vereinen Elemente aus beiden Bereichen. Selbstakzeptanz und Umweltkontrolle sind adaptiver Reife zuzuordnen und zeigen moderate Beziehungen zu Lebenszufriedenheit. Persönliches Wachstum indiziert Persönlichkeitsreife und ist nur gering mit Lebenszufriedenheit korreliert (Ryff 1989, Ryff & Keyes 1995). Die beiden Unterskalen Persönliches Wachstum (PW) und Umweltkontrolle (UK) des PWB sind nur moderat miteinander korreliert (Ryff, 1989; Ryff & Keyes, 1995; Schmutte & Ryff, 1997, Bauer, McAdams & Sakaeda, 2005). Reifes Erwachsenen Verhalten spiegelt sich in individuellen Kombinationen dieser Reifeaspekte, und wurde wie folgt in Validierungsstudien der Weisheit verwendet.

PW indiziert Wachstumsorientierte Persönlichkeitsreife, während UK als Indikator für Adaptivität steht. Persönlichkeitsreife wird durch ein hohes Niveau intrapsychischer Analyse und Differenziertheit erreicht. Adaptivität, siehe nächster Abschnitt, zeichnet sich durch die Fähigkeit aus, Anforderungen der Gesellschaft in Form von Normen und sozialen Beziehungen erfolgreich zu meistern (Staudiner et al., 2005). Diese unterschiedlichen Reifekonzepte offenbarten in der Mills Studie von Helson und Srivastava (2001) differentielle Beziehungen zu verschiedenen Reifekonstrukten wie Kompetenz, Generativität, Ego-Entwicklung und ihrem Weisheitskonzept.

Helson und Srivastava (2001) beziehen sich auf die Arbeit von Helson und Wink (1987) und entwickeln vier Typen unterschiedlicher erwachsener Reife aus der Kombination von UK und PW, siehe Abbildung 1. Es zeigte sich, dass der Typus der Suchenden, mit hohen Werten in PW und in niedrigen UK-Werten, die signifikant höchsten Werte im latenten Weisheitskonstrukt, einem faktorenanalytischen Verbund der Testverfahren TWR, AWR und PWR, erreichten. Die Beziehungen von PW zu Weisheit und Ego-Entwicklung waren signifikant positiv (Helson & Srivastava, 2001). Der Typus der Erfolgreichen, welche sich durch hohe UK-Werte und PW-Werte auszeichnen, gilt als die günstigste Form des erwachsenen Reifeverhaltens. Trotzdem erreichen diese nicht die höchsten Weisheitswerte. Die Erfolgstypen werden zwar als generativ, lebenszufrieden und positiv gestimmt beschrieben, doch fehlt

es ihnen an Zeit zur Integration von Einsichten zu tieferen Erkenntnissen. Die Suchenden zeichnen sich dadurch aus, besondere Ambitionen hinsichtlich weiterer Persönlichkeitsentwicklung zu haben, somit auch mehr Zeit für derartige Erkenntnisvertiefungen zu investieren. Im Vergleich dazu investieren die Erfolgreichen mehr Anstrengungen und Zeit in ihre gesellschaftlichen Ambitionen. Die in der folgenden Abbildung ebenfalls enthaltenen Reifekonzepte der Bewahrer und Verarmten zeigen keine hohen Weisheitswerte oder positiven Zusammenhänge zu Persönlichkeitsreifeindikatoren.

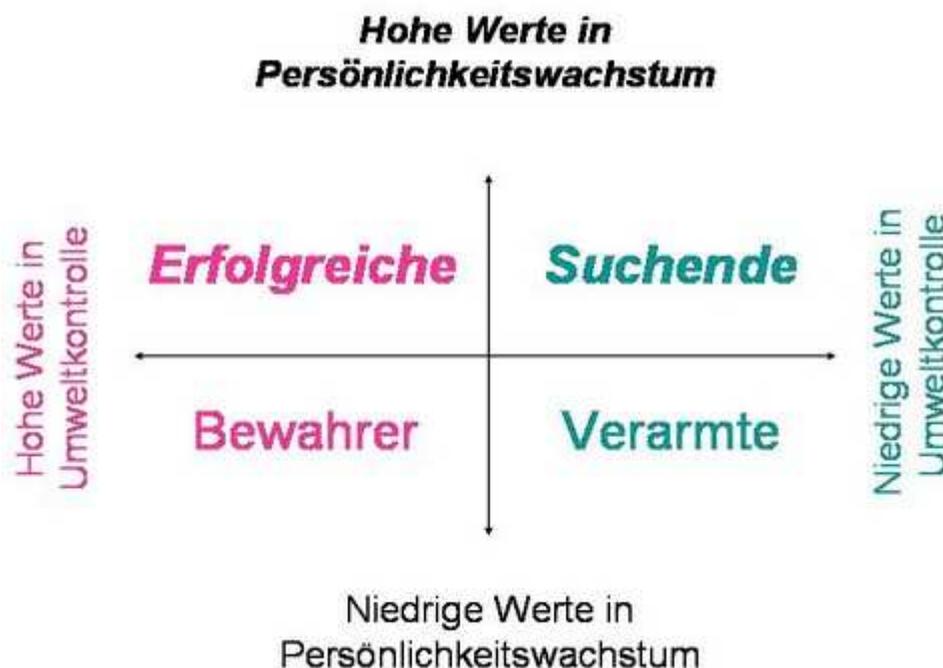


Abbildung 4.1. Typologie von Persönlichkeitsreife nach Helson und Srivastava (2001).

Einen Interaktionseffekt zwischen Alter und der Reifeorientierung stellen Kunzmann, Stange und Jordan (2005) in einer Studie zur Allgemeinen Weisheit fest. Während Adaptivität einen linear positiven Alterstrend aufweist, sind die Werte der Wachstumsorientierten Persönlichkeitsreife in der mittleren Altersgruppe stabil, können jedoch im Alter abfallen. Dieser nicht-lineare Alterstrend ist typisch für Indikatoren von Selbstbezogener und Allgemeiner Weisheit (Staudinger et al. 2005). Desweiteren bezeichnen Staudinger et al. (2005) PW und Lebenssinn als Unterskalen Selbstbezogener Weisheit und stützen sich auf die Studien von Baltes et al. (1995), Kunzmann und Baltes (2003), Kunzmann et al. (2005) und Mickler (2004).

4.2.3.4 Adaptivität

Nach Helson und Srivastava (2001) sollte Weisheit und Adaptivität nicht linear positiv zusammenhängen, siehe vorherigen Abschnitt. Weisheit geht weniger mit Sozialer Anerkennung und Erfolg einher als mit einer relativen Distanz zu dieser (Staudinger et al., 2005). Die im vorherigen Abschnitt vorgestellte Studie von Helson und Srivastava (2001) belegt eine umgekehrt u-förmige Beziehung von Weisheit und Adaptivität. Die Unterskala Umweltkontrolle (UK) des PWB fungierte dabei als Indikator für Adaptivität. In der Studie zeigte UK keine signifikante Korrelation zu Weisheit oder Ego-Entwicklung.

In Micklers Validierung der Selbstbezogenen Weisheit (2004) wurden eine diskriminante Beziehungen zur Unterskala UK nachgewiesen. Diesen diskriminanten Zusammenhang bezweifeln Staudinger et al. (2005) in der Beziehung zu Ardelts 3d-WS. Sie bezeichnen den 3d-WS als Maß für Adaptivität aufgrund des Korrelationsmusters zu Beherrschung. Um der Fragestellungen auf den Grund zu gehen, ob der 3d-WS eher Weisheit und Persönlichkeitsreife oder Adaptivität erfasst, könnte somit einerseits die Unterskala UK des PWB herangezogen werden (Helson & Wink, 1987; Ryff & Keyes, 1995; Schmutte & Ryff, 1997), anderserseits auch mit der Typologie von Helson und Srivastava (2001) gearbeitet werden.

4.2.3.5 Soziale Erwünschtheit

Soziale Erwünschtheit ist ein Konzept, das Verzerrungen im Antwortverhalten von Teilnehmern begünstigt (Lienert & Raatz, 1998). Soziale Erwünschtheit wird in zwei Kategorien eingeteilt, die Selbsttäuschung, eine mangelnden Fähigkeit sich akkurat selbst zu beschreiben, und Impression Management, das das Ziel verfolgt, sich anderen gegenüber möglichst positiv, bzw. sozial Erwünscht, darzustellen (Paulhus, 1984, nach Egloff & Schmukle, 2002). Die Sichtweise, das Sozial Erwünschtes Antwortverhalten eine Messung verunreinige, wurde in den letzten Jahren kontrovers diskutiert. Sozial Erwünschte Antworttendenzen sind nach Dier Sandvik, Pavot und Gallagher (1991) ein substanzieller Teil der Persönlichkeit, anstatt eine Quelle von Fehlervarianz. Neuere Forschungsarbeiten stützen diese Sichtweise (Borkenau & Ostendorf, 1992; McCrae & Costa, 1983; Lönnqvist et al., 2007; Piedmont, McCrea, Riemann & Angleitner, 2000; Smith, 1997).

Es stellt sich daher die Frage, ob Soziale Erwünschtheit und Weisheit gemeinsame Konstruktvarianz teilen. Eine gemeinsame Varianzquelle lässt sich aus Items wie „In Diskussionen sachlich bleiben ... Eigene Fehler eingestehen (Stöber, 2002, S. 232)“ erkennen. Diese Items beschreiben sozial Erwünschte, aber seltene Verhaltensweisen, die weises Verhalten widerspiegeln. So sind über die Hälfte der 16 Items der Sozialen Erwünschtheitsskala SDS-17

von Stöber (2002) Indikatoren weisen Verhaltens, was vor allem aufgrund des dichotomen Antwortformates Missverständnisse in der Interpretation des Zusammenhangs von Weisheit und Sozialer Erwünschtheit fördern kann.

Innerhalb der Weisheitstests gibt es objektive, nichtdurchschaubare Interviewverfahren und subjektive Fragebögen, bei denen das Ziel einzelner Items oder des gesamten Fragebogens erkannt und beeinflusst werden kann. Der PWS erfragt mit dem Item weise die direkteste subjektive Beurteilung Persönlicher Weisheit. Zu den indirekten subjektiven Skalen gehören der 3d-WS und der SAWS. Eine Sonderstellung hat das Transzendente Weisheitsrating von Wink & Helson (1997), das direkt nach erlangter Weisheit fragt, jedoch durch das offene Antwortformat die Möglichkeit der Manipulation verringert.

Um den Einfluss Sozial Erwünschter Antworttendenzen zu verstehen, ist es wichtig, zwischen direkter und indirekter Selbsteinschätzung der eigenen Weisheit zu trennen. Eine direkte Selbstzuschreibung von Weisheit ist nicht gesellschaftlich erwünscht. Dem Weisheitsverständnis der Gesellschaft entspricht es, das sich der Weise selbst nicht als weise bezeichnet. Dies kann mit der paradoxen Erkenntnis erklärt werden, dass die eigene Unwissenheit immer bewusster wird, je mehr Wissen eine Person gesammelt hat. Die soziale Norm verpflichtet daher zu einer bescheidenen Einschätzung oder Verneinung der eigenen Weisheit, genauso wie die Erkenntnistheorie. Im zweiten Fall, einer indirekten Weisheitsmessung über assoziierte Persönlichkeitseigenschaften ist die Soziale Norm eine andere. Schon Edwards (1953) belegt, je positiver eine Eigenschaft beurteilt wird, desto mehr Personen schreiben sich diese zu. Eine Sozial Erwünschte Verzerrung würde daher zur indirekten Zuschreibung von Weisheit führen.

Einzig für den 3d-WS wurde die Unabhängigkeit von Sozialer Erwünschtheit empirisch belegt (Ardelt, 2003). Eine objektive Beurteilung der Weisheit einer Person sollte daher entweder über diesen Test, Leistungsmessung, Fremdbeurteilung oder mittels multimethodaler Erhebung getroffen werden. Die Qualität von Fremdbeurteilungen von Weisheit lassen sich beispielsweise über Nominierungsstudien oder Interraterreliabilitäten abschätzen, was im jeweiligen Falle zu diskutieren ist.

5 Fragestellungen und Hypothesen

Dieses Kapitel präsentiert die Übertragung der Validierungsziele dieser Arbeit in Hypothesen. Das Ziel ist eine Kriteriums- und Konstruktvalidierung schriftlicher Weisheitstestverfahren, des SAWS (Webster, 2003, 2006), 3d-WS (Ardelt, 2003, 2006) sowie PWS, TWR, und AWR (Wink & Helson, 1997; Helson & Srivastava, 2001, 2002). Die einzelnen Teilziele der Kriteriums- und Konstruktvalidierung sind wie folgt eingeteilt. Zur Kriteriumsvalidierung zählt die Innere Validität der Weisheitstestverfahren und die Äußere Validität zu den externen Kriterien bei gleichzeitiger Erhebung, sprich, unter Konkurrenter Validität (5.2 und 5.3). Die Konstruktvalidierung umfasst die Konvergente Validität zu theoretisch verwandten Konstrukten sowie die Diskriminante Validität zu theoretisch getrennten Konstrukten (5.4 und 5.5).

Vorab müssen grundlegende Voraussetzung für die verwendeten Testverfahren und Teilnehmerstichproben geprüft werden, siehe Abschnitt 5.1. In den anschließenden Kapiteln folgen die Beschreibungen der versuchsplanerischen Umsetzung der Studie, die Ergebnispräsentation und die Darstellung der zusätzlichen Analysen.

5.1 Voraussetzungen der Validierung

Auf den Voranalysen der einzelnen Testverfahren bauen die eigentlichen Fragestellungen dieser Studie auf. Hinsichtlich der einzelnen Testverfahren sind Voraussetzungen zu prüfen, die deren Reliabilität und die Beeinflussung der Testergebnisse betreffen. Voranalysen der Teilnehmergruppen befassen sich mit deren Vergleichbarkeit, sie sollten sich beispielsweise nicht in Verbaler Intelligenz unterscheiden. Zunächst folgt jedoch eine Erläuterung zur Formulierung der zu testenden Hypothesen.

5.1.1 Formulierung der Hypothesen zur Validierung der einzelnen Weisheitstestverfahren

Da es sich in dieser Arbeit um eine Validierung schriftlicher Verfahren zu einem psychologischen Konstrukt, der Weisheit, handelt, wird angenommen, dass sich die im Folgenden postulierten konvergenten oder divergenten Beziehungen jeweils bei allen Weisheitstests zeigen. Eine solche integrative Hypothese ist daher streng genommen nur bestätigt, wenn sich die postulierten Beziehungen bei allen Weisheitstests zeigen. Es wäre ebenso möglich, keine integrativen Hypothesen zu formulieren, sondern jeden zu testenden Zusammenhang für alle

Weisheitstestverfahren einzeln zu berücksichtigen, und in einem zweiten Schritt, bei Bestätigung dieser basalen Hypothesen, auf die Gültigkeit der Aussagen für alle Weisheitstests zu verweisen. Dieses Vorgehen ist dem integrierten überlegen, da bei der Falsifikation der integrierten Hypothese nicht geklärt ist, welche und wie viele Tests die Falsifikation verursachten. Aus diesem Grund wird das zweite, detaillierte Vorgehen bei der Ergebnisprüfung gewählt.

Dies würde jedoch für die folgenden Hypothesenabschnitte bedeuten, dass ein zu testender Zusammenhang in theoretischer, empirischer und statistischer Hypothese jeweils für SAWS, 3d-WS, TWR, AWR und PWS einzeln formuliert werden müsste. Die Formulierung: die einzelnen Weisheitstestverfahren ersetzt daher die Ausformulierung der Hypothese für alle einzelnen Testverfahren. Dieses Vorgehen wird ausdrücklich deshalb gewählt, um Redundanz für den Leser zu reduzieren. Bezieht sich eine Fragestellung nicht auf alle Weisheitstestverfahren, so wird dies eindeutig formuliert.

5.1.2 Voralysen zu den Testverfahren

Als Grundlage für die Hypothesen über die Testverfahren wird deren Reliabilität im Sinne der Internen Konsistenz zur Prüfung der Testgenauigkeit bestimmt. Die empirische Prüfgröße der Internen Konsistenz ist Cronbachs Alpha. Akzeptable Werte liegen über .70, gute Werte über .90 (Bortz & Döring, 2005). Bei den Ratingverfahren TWR und AWR ist die Auswertungsobjektivität und Reliabilität der Beurteilung der Antworten zu prüfen. Dies kann über deren Interraterreliabilität, die empirische Prüfgröße ist hier Cohens Kappa, bestimmt werden. Cohens Kappa hat zwischen 0,40-0,75 zufrieden stellende Werte (Diehl & Arbinger, 2003). Weichen die Urteile der Rater von einander ab, wird eine Konsensfindung angestrebt.

Bei TWR, AWR und WUSCT werden Effekte der Antwortlänge auf deren Beurteilung kontrolliert. Eine längere Antwort impliziert mehr Inhalt; ob dieser jedoch einer höheren Merkmalsausprägung entspricht, ist zu prüfen. Es ist zu Erörtern, ob Antwortlänge eine systematische Störgröße ist, somit auszupartialisieren wäre, oder ob diese wahre Konstruktvarianz erklärt und daher nicht auszupartialisieren ist. Vermutlich ist Antwortlänge bei vielen sehr kurzen Antworten oder hoher Streuung der Antwortlänge keine systematische Störvariable. Dagegen ist sie eine derartige Störvariable bei langen Antworten oder bei sehr geringer Streuung. Für AWR und TWR legen Helson und Srivastava (2001, 2002) in der amerikanischen Studie eine Mindestantwortlänge von 70 Worten fest.

Eine weitere Variable, die als systematische Störgröße des TWR, AWR und des WUSCT fungieren könnte, ist die Verbale Intelligenz der Teilnehmer. Es erscheint plausibel, dass es eine hohe Verbale Intelligenz erleichtert, über komplexe Fragestellungen schriftlich zu

philosophieren. In bisherigen Studien sind Allgemeine oder Selbstbezogene Weisheit (Mickler, 2004), TWR, AWR und PWS (Helson & Srivastava, 2001, 2002), und Ego-Entwicklung (Manners & Durkin, 2001) nicht bedeutsam durch Verbale Intelligenz beeinflusst, siehe Kapitel 3 und 4.

Zur Hypothesenprüfung für 3d-WS (Ardelt, 2003) und die drei Tests TWR, AWR und PWS von Helson und Srivastava (2001) sind Faktorwerte zu bilden, die die einzelnen Dimensionen und Untertests zusammenfassen. Diese Faktorwerte werden über die Extraktion des ersten Faktors, durch eine Hauptachsenanalyse der Korrelationsmatrix gebildet.

5.1.3 Voranalysen zu den Teilnehmergruppen

Für die Fragestellung einer Validierung von Weisheitstests sollten Teilnehmer untersucht werden, die dieses Merkmal möglichst in hohem Maße aufweisen. Aufgrund eines fehlenden objektiven externen Kriteriums der Weisheit, wird über die Methode der Laiennominierung, siehe Kapitel 2, eine Gruppe von Personen identifiziert, die im Fremdurteil eines Bekannten als weise eingeschätzt werden. Die Interne Validität einer quasiexperimentellen Untersuchung wird erhöht, wenn in gepaarten Gruppen die relevanten Störvariablen parallelisiert sind (Bortz, 1999). Zu diesen Nominierten werden daher Kontrollpersonen nach Alter, Geschlecht, Beruf und Bildungsniveau gepaart, um den Einfluss dieser möglichen systematischen Störvarianzquellen konstant zu halten. Eine dritte für Weisheitsmessverfahren interessante Gruppe sind Psychotherapeuten, da deren berufliche Spezialisierung ihnen sowohl theoretisches Wissen als auch praktische Erfahrung in der Unterstützung der Problembewältigung anderer Menschen eröffnet. Bei der Wahl der Psychotherapeuten wird eine Übereinstimmung der Alters- und Geschlechtsverteilung zu den gepaarten Gruppen angestrebt. Die zu prüfende Vergleichbarkeit der Gruppen hinsichtlich der genannten Variablen soll sicherstellen, dass Gruppenunterschiede nicht auf den Einfluss dieser Variablen zurückgeführt werden können. Kapitel 6 bietet genauere Informationen zum Stichprobendesign.

Da die Weisheitsforschung zur Entwicklung eines effektiven verhaltenstherapeutischen Therapieverfahrens der Depression führte (Schippan, Baumann & Linden, 2004), ist als Umkehrschluss möglich, dass eine absolvierte Psychotherapeutische Behandlung die Weisheit des Klienten fördert. Es ist daher interessant, die Häufigkeiten von Personen mit Psychotherapeutischen Behandlungen in den gepaarten Gruppen zu vergleichen. Sollten signifikant unterschiedliche Verteilungen vorliegen, sind weitere Analysen über die Rolle dieser Variable im Gefüge zwischen Persönlichkeitsreife und Weisheit notwendig.

5.2 Kriteriumsvalidierung über die Bestimmung der Inneren Validität

Zum Beleg der Inneren Validität der Weisheitstests wird ein multipler Validitätskoeffizient (Lienert & Raatz, 1998) bestimmt, der auf der multiplen Korrelation der einzelnen Tests basiert. Die Größenordnung eines solchen Validitätskoeffizienten sollte über $R = .70$ liegen, während in der Praxis Werte von etwa $.60$ akzeptabel sind und in einer Testbatterie auch ein erheblich niedrigerer Koeffizient eines Tests sinnvoll erscheinen kann. Da Weisheit ein sehr breites Konstrukt unterschiedlicher Definitionen ist, daher viel Varianz in diversen Facetten vereinigt, wird für diese Studie eine Höhe von $R = .60$ erwartet. Werden die Weisheitstests dieser Studie als Testbatterie verstanden, erfasst die multiple Korrelation der einzelnen Tests, deren Stellenwert in der Gruppe.

Hypothese 1: Der multiple Korrelationskoeffizient der einzelnen Weisheitstestverfahren erreicht einen Wert von $.60$.

„Ein solcher multipler Validitätskoeffizient RTC lässt hoffen, dass sich die auf irrelevante Gemeinsamkeiten zurückführbaren Kovarianzanteile wechselseitig aufheben, so dass der multiple Validitätskoeffizient ebenso zu bewerten wäre, wie der äußere Validitätskoeffizient (Lienert & Raatz, 1998, S. 223).“

Neben dieser zentralen Fragestellung der Inneren Validität sind aufgrund theoretischer Annahmen verschiedene Hypothesen über Zusammenhänge zwischen einzelnen Tests zu testen. Aufgrund der Ergebnisse von Wink und Helson (1997) ist zu prüfen, ob die beiden Tests Praktische Weisheitsskala und Transzendentes Weisheitsrating nicht korrelieren. Im Falle fehlgeschlagener Replikation, muss auf spätere Studien von Helson und Srivastava (2001, 2002) verwiesen werden, in denen Praktische Weisheit in ein Fremdeinschätzungsverfahren eine signifikant positive Korrelation zu TWR aufwies, siehe Abschnitte 3.3.1.1 und 3.3.1.2.

Hypothese 2: Die Testverfahren PWS und TWR sind nicht im mittleren oder hohen Maße miteinander positiv korreliert.

Des Weiteren lässt sich anhand der Kategorisierung von Weisheitstestverfahren aus Kapitel 2.3.3 ableiten, dass der SAWS und 3d-WS höher miteinander korrelieren als mit den anderen Weisheitstests dieser Studie. Da SAWS und 3d-WS identisch verortet sind, sich daher in gleicher Weise auf mindestens einer Kategorie von TWR, AWR und PWS unterscheiden, sollte eine Korrelation zwischen beiden identisch verorteten Tests höher ausfallen als zu diesen auf theoretischer Ebene getrennten Testverfahren.

Hypothese 3: Die Testverfahren SAWS und 3d-WS korrelieren signifikant höher miteinander als mit den drei Weisheitstests PWS, TWR und AWR.

5.3 Kriteriumsvalidierung über die Bestimmung der Äußeren Validität

Die Äußere Validität der Testverfahren ist gegeben, wenn diese mit einem Kriteriumsmerkmal der Weisheit in bedeutendem Maße korrelieren. Da es kein äußeres objektives Kriteriumsmerkmal gibt, das die Weisheit einer Person anzeigt, ist die psychologische Forschung auf subjektive Fremdeinschätzungen der Weisheit, also Laiennominierungen, als Kriterium angewiesen, siehe Abschnitte 2.2 und 4.1. Ein Laiennominierungsverfahren dient der Bestimmung des externen Kriteriums der Fremdeingeschätzten Weisheit, Abschnitt 6.3. Im Sinne der Konkurrenten Validität sollte die Gruppe dieser Nominierten mindestens genauso weise sein wie Kontrollpersonen, die mit ihnen nach Alter, Geschlecht, Beruf und Bildungsniveau gepaart wurden, für die jedoch keine Einschätzung ihrer Weisheit vorliegt. Als Replikation von Forschungsergebnissen des Berliner Modells sollte die Gruppe der Nominierten signifikant bessere Ergebnisse in den Weisheitstests erzielen als die Kontrollgruppe. Sollte dies zutreffen, so spricht dies für die Validität des Äußeren Kriteriums und der Testverfahren.

Hypothese 4: Die Nominierten haben ein signifikant höheres Mittel auf den jeweiligen Weisheitstestverfahren als die Kontrollgruppe.

In Studien von Helson und Srivastava (2001) und Baltes et al. (1995), wurde deutlich, dass Psychotherapeuten für die externe Kriteriumsvalidierung von Weisheitstests interessant sind. Die Mills-Studie zeigte einerseits, dass sich Weisheit durch den Beruf Psychotherapeut vorhersagen lässt. In der Studie von Baltes et al. andererseits, waren die Psychotherapeuten signifikant weiser als Kontrollpersonen. Demzufolge ist die nachstehende Hypothese ableitbar.

Hypothese 5: Die Psychotherapeuten haben ein signifikant höheres Mittel auf den einzelnen Weisheitstestverfahren als die Kontrollgruppe.

Im Vergleich beider äußerer Kriterien, der Nominierung und des Berufes Psychotherapeut, wird erwartet, dass die Nominierten höhere Weisheitswerte erreichen. Aufgrund der geringen Vorteile der Nominierten gegenüber den Psychotherapeuten in Teilen der Studie von Baltes et al. (1995), ist der erwartete Effekt jedoch klein. Bestätigt sich diese Hypothese, so ist sie als Beleg für die Stärke des externen Kriteriums der Fremdeingeschätzten Weisheit zu verstehen. Findet sich diese Überlegenheit nicht, so stützt dies die Studie von Helson und Srivastava zur Bedeutung der beruflichen Spezialisierung in Psychotherapie (1995).

Hypothese 6: Die Nominierten haben ein signifikant höheres Mittel auf den jeweiligen Weisheitstestverfahren als die Psychotherapeuten.

5.4 Konstruktvalidierung anhand der Konvergenten Validität

Als Beleg der Konvergenten Validität gilt, wenn eine positive Beziehung zwischen zwei theoretisch verwandten Konstrukten empirisch bewiesen wird und der gemeinsame Varianzanteil zwischen 10% und 50% erreicht. Konvergente Korrelationen liegen daher im Bereich von .30 und .70. Nach Unten grenzt dieser Bereich diskriminante Beziehungen ab, nach Oben dagegen validitätsähnliche Bereiche, die gleiche anstatt getrennte Konstrukte erfassen.

Persönlichkeitsreife und Weisheit sind aus theoretischer Sicht konvergent miteinander verbunden. Den Abschnitten 4.2.2.3 und 4.2.3 folgend, ist Ego-Entwicklung als kognitives Konstrukt anzusehen und Psychologisches Feingefühl und Persönliches Wachstum als eher persönlichkeitsbezogene Indikatoren der Persönlichkeitsreife. Es ist somit die Existenz von moderaten Korrelationen zwischen den Konstrukten der Persönlichkeitsreife und den einzelnen Testverfahren der Weisheit zu prüfen, die die Verwandtschaft und Eigenständigkeit beider Konstrukte stützt.

Hypothese 7: Ego-Entwicklung korreliert moderat positiv mit den einzelnen Weisheitstestverfahren.

Hypothese 8: Psychologisches Feingefühl korreliert moderat positiv mit den einzelnen Weisheitstestverfahren.

Hypothese 9: Persönliches Wachstum korreliert moderat positiv mit den einzelnen Weisheitstestverfahren.

5.5 Konstruktvalidierung anhand der Diskriminanten Validität

Die Eigenständigkeit eines Konstruktes wird durch diskriminante Beziehungen zu nicht theoretisch verwandten Konstrukten belegt, sollte die entsprechende Korrelation nicht über .30 hinausgehen. Zu diesen theoretisch nicht verwandten Konstrukten der Weisheit zählen die Persönlichkeitsvariable Adaptivität, die Kognitive Fähigkeit Verbale Intelligenz, sowie die demographischen Variablen Alter und Geschlecht. Diese verkörpern die zentralen theoretischen Annahmen über diskriminante Konstrukte der Weisheit, siehe Kapitel 4.

Dem Abschnitt 4.2.3.4 entsprechend, wird die Hypothese geprüft, dass die einzelnen Weisheitstestverfahren unabhängig vom Reifekonzept der Adaptivität seien, welches sich von wachstumsorientierten Persönlichkeitsreifekonzepten unterscheidet. Staudinger et al. (2005) vertreten jedoch die Gegenthese für den Weisheitstest 3d-WS. In deren Validierungsstudie

(2003) korrelierte sie moderat mit einem Maß der Umweltkontrolle. Bisher liegen keine weiteren empirischen Belege für oder gegen diese Behauptung vor.

Hypothese 10: Die Korrelation von Adaptivität und den einzelnen Weisheitstestverfahren liegt im diskriminanten Bereich.

Die Abschnitte des Kapitels 4 zu Intelligenz belegen, dass Psychometrische Intelligenz keinen bedeutsamen Einfluss auf Weisheit zu haben scheint. Demnach sollten die Weisheitstests keine Korrelation zu kristalliner, Verbaler Intelligenz nach HAWIE-R aufweisen. Der dazu verwendete Wortschatztest wird dank der hohen Interkorrelation mit der Gesamtskala des HAWIE-R auch als Maß Allgemeiner Intelligenz angesehen.

Hypothese 11: Die Korrelation von Verbale Intelligenz und den einzelnen Weisheitstestverfahren liegt im diskriminanten Bereich.

In einer Erwachsenenstichprobe sollten die Weisheitstestverfahren keine moderate oder große Beziehung zur demographischen Variablen Alter aufweisen, wie dies Kapitel 4 verdeutlicht. Beispielhaft finden sich keine Altersunterschiede ab 25 Jahren aufwärts bis ins hohe Alter über die Studien des Berliner Paradigmas. Die Hypothese 12 prüft daher diese diskriminante Korrelationsbeziehung für die vorliegenden einzelnen Weisheitstests.

Hypothese 12: Die Korrelation zwischen dem Alter der Teilnehmer und den einzelnen Weisheitstestverfahren liegt im diskriminanten Bereich.

Die Beziehung zwischen der Personenvariablen Geschlecht und Weisheit wird aus theoretischer Sicht eindeutig als diskriminant angesehen. Würden sich jedoch signifikante Vorteile der Frauen in den zu validierenden Weisheitstests zeigen, so wären diese bei 3d-WS und SAWS Replikationen der Befunde von Webster (2003) und Ardelt (2003), siehe Kapitel 4. Ob dies jedoch in der vorliegenden Stichprobe und bei anderen Weisheitstestverfahren ebenso der Fall ist, bleibt in Hypothese 13 zu testen. Wird diese Hypothese bestätigt, so erfassen die Testverfahren ein Weisheitskonzept ohne geschlechtsspezifischen Bias, was angesichts der Annahme der diskriminanten Beziehung durchaus wünschenswert ist.

Hypothese 13: Die Korrelation zwischen dem Geschlecht der Teilnehmer und den einzelnen Weisheitstests liegt im diskriminanten Bereich.

5.6 Weitere Analysen

Im Folgenden werden Hypothesen hergeleitet, die über die Kernzielsetzung der Validierung der schriftlichen Weisheitsmessverfahren hinausgehen. Ihre Bedeutung für die Interpretation der Kernhypothesen liegt in der Vertiefung des Verständnisses für die Ergebnisse der einzelnen Teilnehmergruppen. Es wird daher eine deskriptive Beschreibung anstatt einer strikt hypothesengeleiteten Analyse für diesen Zweck erfolgen. Um den nachgeordneten Stellenwert dieser Weiteren Analysen zu betonen, werden weder als Hypothesen gekennzeichnet noch im Abschnitt der statistischen Hypothesen beschrieben.

Es wird analysiert, ob sich die Teilnehmergruppen bezüglich der Inneren Validität, einem Teilbereich der Kriteriumsvalidität, unterscheiden, wozu signifikante Korrelationsunterschiede zwischen den Gruppen in einzelnen Weisheitstestverfahren existieren müssten. Es würde für das Äußere Kriterium der Nominierung sprechen, wenn die Nominierten hohe Werte der Interkorrelation der Weisheitstests aufweisen. Wären hingegen wesentlich geringere Korrelationen als in der Gesamtgruppe für die Nominierten zu berichten, so müsste die fehlende Innere Validität in Anbetracht der Äußeren Validität des Nominierungskriteriums diskutiert werden. Bestünden unterschiedlich hohe Validitätskoeffizienten zwischen den Gruppen, so wäre von Differentieller Validität der Testverfahren zu sprechen.

WA 1: Für jede Teilnehmergruppe werden die Interkorrelationen der Weisheitstests auf bivariater Ebene miteinander verglichen.

Die theoretische Trennung der Konstrukte der Konvergenten und Diskriminanten Validierung müssen auf empirischer Ebene geprüft werden. Die einzelnen bivariaten Korrelationen werden für die Gesamtgruppe als auch für die einzelnen Teilnehmergruppen getrennt präsentiert.

WA 2: Für Gesamtgruppe und die drei Teilnehmergruppen werden die Interkorrelationen der Testverfahren zur konvergenten und diskriminanten Validierung auf bivariater Ebene berichtet.

Über die bisher etablierte Gruppeneinteilung der Teilnehmer hinaus geht die Analyse der Typologie zu Erwachsenen Reifeverhalten und dessen besonderer Beziehung zu Weisheit von Helson und Srivastava (2001). Diese Analyse stellt eine Ergänzung zur Konvergenten und Diskriminanten Validierung in den Kernhypothesen dar.

Das zur konvergenten Validierung verwendete Skala Persönliches Wachstum und das zur diskriminanten Validierung verwendete Skala Umweltkontrolle bilden über Trennung am jeweiligen Median die vier Typen nach Vorbild von Helson und Srivastava (2001). Interessant

für die Validierung von Weisheitstests sind insbesondere die Vergleiche zwischen dem Typus der Suchenden, welche einzig im PW hohe Werte erzielen und den Erfolgstypen, welche in beiden Tests hohe Werte aufweisen. In der betreffenden Studie zeigten die Suchenden höhere Werte auf der faktorenanalytischen Kombination der Tests TWR, AWR und PWR als die Erfolgstypen. Darüber hinaus zeigten die Suchenden die konvergenten Beziehungen, die aus theoretischer Ableitung für Weisheit zutreffen sollten, siehe Kapitel 4. Übereinstimmend mit dem oben beschriebenen Herangehen bei den drei etablierten Teilgruppen, wird ebenfalls für diese Stichprobenunterteilung die Mittelwerte und Interkorrelationen der Weisheitstestverfahren präsentiert.

WA 3: Für die Suchenden und die Erfolgstypen werden die Mittelwerte der Weisheitstestverfahren verglichen.

WA 4: Für die Suchenden und die Erfolgstypen werden die Interkorrelationen der Weisheitstests auf bivariater Ebene miteinander verglichen.

6 **Versuchsplanerisches und Methodisches Vorgehen**

Dieses Kapitel beschreibt das versuchsplanerische und methodische Vorgehen dieser Validierungsstudie. Es handelt sich hierbei um eine querschnittliche Studie. Die Stichprobe hat eine besondere Bedeutung für die Kriteriumsvalidierung. Ihre Daten präsentiert Abschnitt 6.6. Zunächst werden Überlegungen zur Bestimmung eines geeigneten Stichprobenumfangs in Kapitel 6.1 angestellt. Die drei ad hoc gebildeten Teilnehmergruppen dieser Studie sind an die Validierungsstudie von Balthes et al. (1995) angelehnt, die Nominierte, Kontrollpersonen und Psychotherapeuten umfasste, siehe Kapitel 4.1. Die Rekrutierung der entsprechenden Personen wird im zweiten Abschnitt beschrieben. Den methodischen Gesichtspunkten der verwendeten Instrumente widmet sich Abschnitt 6.3. Eine Schilderung der Kontaktaufnahme sowie der schriftlichen und mündlichen Datenerhebung erfolgt in Abschnitt 6.4. Die Abschnitte 6.5 und 6.6 widmen sich der Auswertungsstrategie und den statistische Hypothesen der Forschungshypothesen.

6.1 **Stichprobenumfang**

Um Testverfahren für das rare Persönlichkeitsmerkmal Weisheit zu validieren, sollten sich möglichst viele Probanden in der Stichprobe befinden, die dieses Merkmal auch aufweisen. Eine Validierung über eine Zufallsstichprobe liefere daher Gefahr, keine weisen Personen zu erfassen, und somit nichts über Weisheit aussagen zu können. Eine Zufallsstichprobe ist einem Nominierungsverfahren zur Identifizierung von weisen Personen in der Eignung für eine Validierungsstudie daher unterlegen. Es ist ebenso nicht realistisch, eine sehr große Gruppengröße über ein Nominierungsverfahren eines sehr raren Merkmals wie Weisheit zu erwarten. Deshalb werden die Überlegungen über eine realistisch anzustrebende Stichprobengröße auf bisherige Studien und die folgende ethische Erwägung aufgebaut.

Lienert und Raatz (1998) empfehlen, bei fehlenden repräsentativen Stichproben einen Validitätskennwert aus dem Gruppenvergleich zwischen extrem hohen und niedrigen Merkmalsausprägungen – in diesem Falle für das Merkmal Weisheit – zu ermitteln. Für die Rekrutierung von Versuchspersonen ist es jedoch ethisch nicht vertretbar, eine Nominierung in die Extremgruppe fehlender Weisheit zu erfragen. Eine solche negative Bewertung kann für den derart Nominierten eine unberechenbare Bedrohung des Selbstwertes darstellen, weshalb auf einen Extremgruppenvergleich verzichtet wird.

Nach Lienert und Raatz (1998) steigt die Bedeutsamkeit eines Validitätskennwertes mit der Anzahl der Teilnehmer und der Qualität der Stichprobe. Bei einer repräsentativen Stichprobe seien 30 Personen die Untergrenze für eine Validierungsstudie. Die Validierungsstudie zu Allgemeiner Weisheit von Baltes et al. (1995) brachte über ein Nominierungsverfahren mit lediglich 14 Nominierten signifikante, wenn auch leichte Gruppenunterschiede zur Kontrollgruppe hervor, siehe Kapitel 4.1. Selbst eine kleine Stichprobe kann daher Ergebnisse zur Analyse der Validität von Weisheitstestverfahren liefern.

Da die Anzahl der eingehenden Nominierungen nicht willkürlich beeinflussbar ist, wurde auf eine vorab festgelegte Größenordnung der Teilnehmerstichproben verzichtet. Als Mindestwert für die Stichprobengröße wurde jedoch die Menge von 30 Teilnehmern je Gruppe angestrebt, somit die Teilnehmeranzahl der oben genannten Studie verdoppelt.

Die Stichprobenumfangsbestimmung erfolgt daher nicht über die statistisch elegante Methode der Poweranalyse. Über eine Poweranalyse wird der optimale Stichprobenumfang zu einem praktischen Maß der Signifikanz bestimmt, die sich in einer kleinen, mittleren oder großen Effektstärke ausdrücken kann. Hierbei gilt, je größer der Effekt, desto weniger Teilnehmer sind notwendig, um diesen zu belegen. Genauso kann die Effektstärke bestimmt werden, die mittels des erreichten Stichprobenumfangs nachweisbar war. Auf die vorliegende Studie angewendet, ergeben sich folgende zu erwartende Effektstärken, wenn ein erwarteter Stichprobenumfang von mindestens 90 Personen, ein Alphafehlerniveau von $\alpha = .05$ und einer Teststärke von $1 - \beta = .80$ zu Grunde gelegt wird.

Für Mittelwertsunterschiede abhängiger Stichproben spielt die Korrelation zwischen den Messwertreihen eine entscheidende Rolle. Zur Entdeckung eines mittleren Effektes $d' = 0,50$ bei geringer Korrelation von $r = .20$ sind 43 Personen erforderlich. Kleine Effekte benötigen hier 276 Personen. Die Prüfung von Korrelationen eines mittleren Effektes ($r = .30$) erfordern 68 Personen. Ein kleiner Effekt ($r = .10$) wird erst ab 618 Personen offenbart. Beim Vergleich von Korrelationsverhältnissen erfordert der Beleg eines mittleren Effektes von $q = .30$ jeweils 140 Personen. Ein großer Effekt ($q = .50$) benötigt jeweils 54 Personen. Bei der Testung von multiplen Korrelationen sind realistischer Weise nur mittlere Effekte entdeckbar, für die bei 5 Prädiktoren 91 Personen erfasst werden müssen. Zusammenfassend können daher nur mittlere oder große Effekte der praktischen Signifikanz in der vorliegenden Studie erwartet werden (Bortz & Döring, 2006).

6.2 Stichprobenziehung

Nachfolgend wird das Rekrutierungsverfahren der einzelnen Teilnehmergruppen vorgestellt. Die Zielsetzung der Stichprobenziehung ist, ad hoc eine Gruppe mit möglichst hoher Merkmalsausprägung in Weisheit über Nominierung zu rekrutieren. Dazu werden in der zweiten Gruppe nach Alter, Geschlecht und beruflichem Werdegang gematchte Kontrollpersonen rekrutiert. Für die dritte Gruppe der Psychotherapeuten gelten die Geschlechts und Altersverteilungen der Nominierten, an ihre fremdeingeschätzte Weisheit wird keine Bedingung geknüpft. Psychotherapeuten lieferten bisher aufgrund ihrer Ausbildung und beruflichen Praxis in Studien zu Weisheitstestverfahren interessante Ergebnisse, die im Hinblick der Validität der Verfahren diskutiert werden (Baltes et al., 1995; Helson & Srivastava, 2002; Staudinger et al., 1998), siehe Abschnitt 4.2.1.4. Beim Vergleich der drei Teilnehmergruppen sind Aussagen über die Validität des Außenkriteriums der Nominierung und der Testverfahren möglich.

6.2.1 Rekrutierung der Nominatoren

Zur Kriteriumsvalidierung der Testverfahren diene als Außenkriterium eine hoch positive Fremdeinschätzung der persönlichen Weisheit von Personen, die die Nominierengruppe bilden. Um eine solche Gruppe zu erhalten, sind Nominatoren notwendig. Aufgrund der in Kapitel 2 geschilderten Ergebnisse der Laienforschung zu Weisheit, ist zu erwarten, dass keinerlei Kriterien die Eignung der Nominatoren beeinflussen, weshalb diese nach Zugänglichkeit, also ad hoc, rekrutiert werden können.

Über ein Schneeballverfahren kann ein möglichst großer Kreis potentieller Nominatoren angesprochen werden. Ein Schneeballverfahren sieht vor, dass die als Nominatoren angesprochenen Personen den Nominierungsauftrag an ihre Bekannten weitergeben. Der Nominierungsauftrag lautete: „Kennen Sie eine Person in Ihrem Bekanntenkreis, die Sie als weise bezeichnen würden?“. Dieser Nominierungs- und Rekrutierungsauftrag wurde entweder im persönlichen oder computervermittelten Kontakt an die potentiellen Nominatoren erteilt, und durch ein Informationsfaltblatt, siehe Anhang A, unterstützt. Das Faltblatt informierte über den Inhalt der Studie, die Modalitäten einer Nominierung sowie über die Möglichkeiten zur Kontaktaufnahme zur durchführenden Diplomandin und der leitenden Institution.

An der Universität Koblenz-Landau sowie im Bekanntenkreis der Diplomandin wurden über 200 dieser Faltblätter ausgegeben und etwa 50 Emails versendet, die das Email-Schneeballverfahren initiierten. Bei einem völligen Versagen des Schneeballverfahrens käme eine konservative Schätzung auf circa 10.000 potentielle Nominierte, wenn ein durchschnittlicher Bekanntenkreis von 40 Personen je angesprochenen Nominator vorhanden wäre. Das

Schneeballverfahren lässt jedoch keine Rückschlüsse über die Gesamtmenge an potentiellen Nominierten zu, die den Nominierungen zu Grunde lag.

Zur Verringerung der Hemmschwelle einer Nominierung konnte je nach persönlicher Vorliebe die Nominierung im persönlichen Kontakt mit der Diplomandin, sowie per Telefon oder Emailkontakt auf anonyme Weise erfolgen. Um anonyme Nominierungen zu gewährleisten, wurden keine soziodemographischen Daten der Nominatoren gesammelt.

Es wurde darauf verzichtet, nur mehrfach nominierte Personen in die Nominierten-gruppe aufzunehmen, was in der Studie von Baltes et al. (1995) Bedingung war. Ein solches Vorgehen verlangt eine Vorabselektion von potentiellen Teilnehmern, für die es jedoch keine wissenschaftlich begründeten Kriterien gibt. Eine eingegangene Nominierung ist per definitionem ein positives Fremdurteil über das subjektive Kriterium der wahrgenommenen Weisheit eines Nominierten. In sofern wird Orwoll und Perlmutter (1990) Kriterium der Peernominierung erfüllt, während Mehrfachnominierungen keine Bedingung sind, siehe 2.2.2.

6.2.2 Rekrutierung der Gruppe der Weisheitsnominierten

Die Aufklärung über die Nominierung und die Studie erfolgte im telefonischen Kontakt, siehe Kapitel 6.3. Auf diesem Wege wurde das Einverständnis der Nominierten zur Teilnahme eingeholt. Insgesamt gingen 73 Nominierungen ein, wobei die Teilnahmequote der Nominierten bei 52 % lag. Die Nominierten wurden jeweils von einem Nominator vorgeschlagen.

6.2.3 Rekrutierung der Kontrollgruppe

Für die Zusammenstellung der Kontrollgruppe ist ausschlaggebend, dass diese sich im Mittel sowie in der Varianz verschiedener potentieller Moderatorvariablen von der Gruppe der Nominierten nicht unterscheidet (Orwoll & Perlmutter, 1990). Eine versuchsplanerische Methode, um dieses Ziel zu erreichen, ist das Matching von Teilnehmern dieser Variablen.

Zur Erstellung des Rekrutierungsprofils für das Matching wurden die im Erstkontakt die soziodemographischen Variablen Alter, Geschlecht, Beruf und Bildungsniveau ausgewählt. Die Rekrutierung der Kontrollpersonen wurde ebenfalls ad hoc durchgeführt, d.h. die der Diplomandin bekannten Personen wurden auf Passung zu einem Rekrutierungsprofil geprüft und angesprochen, ob sie wiederum eine Person kennen, die einem der gesuchten Matchingprofile entspreche. Die Güte des Matchings der Kontrollgruppe zur Nominierengruppe zeigt Anhang F.1, der die soziodemographischen Merkmale jeder Gruppe umfasst. Die Kontaktaufnahme erfolgte in gleicher Weise wie zu den Teilnehmern der Nominierungsstichprobe, siehe dazu Abschnitt 6.3.

6.2.4 Rekrutierung der Gruppe der Psychotherapeuten

Die Psychotherapeuten wurden über eine Krankenkassenliste niedergelassener psychologischer sowie ärztlicher Psychotherapeuten im Rhein-Neckarraum rekrutiert. Die Matchingkriterien für diese Gruppe waren Geschlecht und Alter der Teilnehmer der Nominierten-Gruppe. Die notwendige Menge älterer Psychotherapeuten konnte nicht über die verwendete Liste praktizierender Ärzte und Psychotherapeuten gedeckt werden, weshalb eine altersbezogene Gleichverteilung der weiblichen und männlichen Psychotherapeuten angestrebt wurde. Dieses Vorgehen ist aufgrund der Befunde der Studie von Baltes et al. (1995) gerechtfertigt, da weder Alter noch berufliche Erfahrung einen Einfluss auf die Allgemeine Weisheit der Psychotherapeuten haben.

6.3 Verwendete Instrumente

Zur Konstruktvalidierung der fünf Weisheitstestverfahren 3d-WS, SAWS, PWS, TWR und AWR, werden die konvergenten Konstrukte Ego-Entwicklung, Persönliches Wachstum und Psychologisches Feingefühl sowie die diskriminanten Konstrukte Adaptivität und Intelligenz verwendet. Die Reihenfolge dieser Instrumente in der Tabelle 6.1 entspricht deren methodischer Darstellung in den nächsten Abschnitten. Einzig das Intelligenztestverfahren HAWIE-R wird in mündlicher Erhebung erfasst. Die Beschreibung des demographischen Teils des Fragebogenhefts ergänzt die Darstellung der verwendeten Instrumente. Die schriftliche und mündliche Erhebung wird in Abschnitt 6.4 vorgestellt. Auf eine Ausführung der Definitionen der Zielkonstrukte der einzelnen Instrumente wird in sofern verzichtet als diese bereits in vorherigen Kapiteln 3 und 4 erfolgte.

Die drei im amerikanischen oder kanadischen Englisch entwickelten Instrumente PWS, SAWS, und die Unterskala Psychologisches Feingefühl des IPIP wurden von zwei unabhängigen muttersprachlichen Übersetzern mittels der Backwardmethode ins Deutsche übersetzt. In Fällen der Diskrepanz der übersetzten Version mit der Ursprungsversion wurde eine Übereinkunft der Übersetzer über eine angemessene Übersetzung erlangt. Die Adjektivliste des PWS zeigte dabei die größten Probleme, da im Sprachgebrauch typischere als die ursprünglichen Adjektive in der Rückübersetzung gewählt wurden. Dies ist einerseits durch die Kontextfreiheit des ACL (Gough & Heilbrun, 1983), und andererseits durch das Fehlen eines exklusiven deutschen Äquivalentes erklärbar.

Tabelle 6.1. Verwendete Testverfahren.

Konstrukt	Instrument/Subskala	Abkürzung
Schriftliche Erhebung		
Weisheit	3dimensionale-Weisheitsskala	3d-WS
	Selbstadministrative Weisheitsskala	SAWS
	Praktisches Weisheitsskala	PWS
	Transzendentes Weisheitsrating	TWR
	Allgemeines Weisheitsrating	AWR
Persönlichkeitsreife	Ego-Entwicklung	WUSCT
	Psychologisches Feingefühl	PF
	Persönliches Wachstum	PW
Adaptivität	Umweltkontrolle	UK
Mündliche Erhebung		
Intelligenz	Wortschatztest	HAWIE-R

6.3.1 Dreidimensionale-Weisheitsskala (3d-WS)

Die 3-dimensionale Weisheitsskala von Ardelt (2003) eröffnet das Fragebogenheft der Studie. Die 3d-WS ist ein 39 Items umfassender Fragebogen Selbstbezogener Weisheit mit Kognitiven, Affektiven und Reflektiven Dimensionen. Weisheit wird als latenter Faktor höherer Ordnung aus den drei Dimensionen gebildet. Das fünfstufige Antwortformat der 3d-WS umspannt für die Items 1-24 die Pole von Trifft ganz bestimmt für mich zu bis Trifft nicht für mich zu, während die Items 25-36 die Antwortkategorien von Stimme sehr zu bis Stimme überhaupt nicht zu aufweisen. Bei der Reliabilitätsprüfung der 3d-WS zeigte sich in einem zehnmonatigen Intervall eine Retest-Reliabilität ($r_{tt} = .85$) und eine Interne Konsistenzen ($\alpha = .71$ bis $.85$). Die 3d-WS, in amerikanischem Englisch entwickelt, wurde von Ardelt (2006) für eine Online-Erhebung einer deutschen Zeitschrift ins Deutsche übersetzt und in dieser Version verwendet.

6.3.2 Selbstadministrative Weisheitsskala (SAWS)

Die Selbstadministrative Weisheitsskala von Webster (2003, 2006) wird über fünf Dimensionen gemessen: Kritische Lebensereignisse, Offenheit für Erfahrungen, Emotionsregulationskompetenz, Humor und Reminiszenz und Reflektiertheit. Das Antwortformat ist sechsstufig

mit den Polen Trifft sehr auf mich zu und Trifft überhaupt nicht auf mich zu. In der 2003er Version mit 6 Items je Unterskala, gibt der Cut-Off-Wert von 152 von 180 möglichen Punkten die Grenze zu Weisheit an. In der erweiterten 2006er Version bilden 8 Items je Unterskala insgesamt 40 Items. Der Cut-Off-Wert gleichen Verhältnisses liegt bei 203 von 240 Punkten. Die neuere Version zeigte eine sehr gute Interne Konsistenz und exzellente Rest-Reliabilität ($r_{tt} = .838$; $\alpha = .904$) in einem Zeitraum von 2 Wochen.

6.3.3 Praktische Weisheitsskala (PWS)

Die dichotomen Items der Praktischen Weisheitsskala nach Wink und Helson (1997) bestehen aus 13 indikativen und vier kontraindikativen Items: Nüchtern denkend, unparteiisch, einsichtig, intelligent, vielseitig interessiert, reif, realistisch, vernünftig, reflektiert, gedankenvoll, tolerant, verständnisvoll, weise, unreif, intolerant, leichtsinnig, oberflächlich. Die Reihenfolge der Items im Fragebogenheft folgt einer zufälligen Verteilung.

Diese Items entstammen ursprünglich der Adjektiv Check List (Gough & Heilbrun, 1983), die aufgrund ihrer Länge mit 300 Items nicht vollständig eingesetzt werden konnte. Das Item ehrlich wurde als Einstiegsitem eingefügt, welches nicht in den Gesamtscore der PWS Skala eingeht. Diese Eigenschaft wurde ausgewählt, da sie die Absicht der Teilnehmer, die anderen Items wirklich ehrlich zu beantworten, positiv beeinflussen kann.

Die Werte der Internen Konsistenz liegen zwischen $\alpha = .74$ und bei $\alpha = .79$. Über einen Zeitraum von mehreren Jahrzehnten ergeben sich Korrelation von $r = .40$, was auf die relative Stabilität dieses Traits schließen lässt (Wink & Helson, 1997).

6.3.4 Transzendentes Weisheitsrating (TWR)

Die Frage des Transzendentalen Weisheitsratings nach Wink und Helson (1997) lautet: „Viele Leute hoffen, mit steigendem Alter weiser zu werden. Würden Sie ein Beispiel von Weisheit geben, die Sie erworben haben und beschreiben, wie es dazu gekommen ist?“. Um den Aufforderungscharakter des offenen Antwortformates zu unterstreichen, wird den Teilnehmern eine ganze Seite zur Bearbeitung dieser Aufgabe zugestanden. Die Ergebnisverteilung der 138 Teilnehmer zeigte Annäherung an die Normalverteilung. Das fünfstufige Bewertungsschema wurde in der Studie von Wink und Helson (1997) von Persönlichkeitspsychologen nach zwei Stunden Training mit guter Interraterreliabilität von .81 bis .86 gescort. In dieser Studie wurde das Scoring aller TWR-Antworten von der Diplomandin und einer Diplompsychologin durchgeführt. Die Ausführungen des Bewertungsschemas zeigt E.1 im Anhang.

6.3.5 Allgemeines Weisheitsrating (AWR)

Helson und Srivastava (2001, 2002) adaptieren eine Allgemeine Weisheitsaufgabe: „Stellen Sie sich vor, Sie erhalten einen Anruf eines guten Freundes. Dieser erzählt Ihnen, er könne nicht so weiter machen und er wolle sich das Leben nehmen. Was würden sie denken und tun?“ in ihr offenes Antwortformatmodell des TWR und erstellen somit das Allgemeine Weisheitsrating. Es werden gute bis sehr gute Bewertungsübereinstimmungen von .80 bis .96 erreicht. Das Scoring dieser Aufgabe übernimmt die Diplomandin, wobei ein Drittel der Antworten von einer Diplompsychologin gegengescort wurde. Das Bewertungsschema umfasst vier Kriterien, für die jeweils fünf Punkte vergeben werden können: Kognitive Differenziertheit, Prozedurales Wissen, Moralische Komplexität und Emotionales Verständnis. Das Bewertungsschema E.2 ist für jedes Kriterium intervallskaliert. Die Analyse des AWR stützt sich auf den Summenscore über diese Kriterien.

6.3.6 Washington University Sentence Completion Test (WUSCT)

Ego-Entwicklung von Loevinger (1976) ist ein Indikator der Persönlichkeitsreife, siehe Kapitel 4.2.2.2. Loevingers projektiver Satzergänzungstest WUSCT enthält dazu Satzanfänge der folgenden Art: Eine Ehefrau solle..., oder Was mich in Schwierigkeiten bringt ist.... Eine ausführliche Beschreibung der Kodierung von Ego-Entwicklung findet sich im Manual von Hy und Loevinger (1996).

Den Satzergänzungstest gibt es als Vollversion mit 36 Items und in Kurzversionen. Besonders die 18 Item Version erzielt eine gute bis zufrieden stellende Interne Konsistenz und wird daher hier verwendet. Die deutschsprachigen Items ($\alpha = .86$) entstammen der Validierungsstudie von Mayer (2003), welche kohärente Ergebnisse mit bisherigen Studien lieferte (Lorr & Manning, 1978; Manners & Durkin, 2001). Aufgrund der Diskussion von Instruktionseffekten von Drewes und Westenberg (2001), wurde die Standardinstruktion: Ergänzen Sie die folgenden Satzanfänge möglichst spontan gewählt.

Das Scoring des Gesamtwertes folgt der von Loevinger entwickelten Ogive-Regel (Hy & Loevinger, 1996), um jeden Probanden auf seiner optimalen Leistungsebene zu bewerten. Die Ogive-Regel verringert Einflüsse von Antworten die unterhalb des jeweiligen Leistungsniveaus liegen, aber erwartungsgemäß auftreten. Zum Scoring bemerkt Holt (1980), dass Psychologiestudenten anhand des Manuals von Hy und Loevinger (1996) gute Interraterreliabilitäten (.76 bis .85) mit Ratern erreichen, die in diesem Bereich trainiert wurden. Das Scoring der Diplomandin erfolgt daher entsprechend des Manuals, in Grenzfällen wird durch Übereinkunft mit zwei Psychologiestudenten ein adäquates Scoring getroffen.

6.3.7 Psychologisches Feingefühl (PF)

Psychologisches Feingefühl wird mittels der Competence Subskala des International Personality Item Pool (Goldberg, et al., 2006) erhoben. Dieses zehn Items umfassende Instrument wurde konstruiert, um Goghs (1964) Unterskala Psychological Mindedness des California Psychological Interview (CPI) zu ähneln. Die IPIP Unterskala hat eine bessere Interne Konsistenz mit $.74$ als die Unterskala des CPI mit $a = .55$. Eines der Items lautet: Ich kann mit guten Lösungen aufwarten. Das Antwortformat ist eine sechsstufige Skala mit den Polen Trifft sehr auf mich zu bis Trifft überhaupt nicht auf mich zu.

6.3.8 Persönliches Wachstum (PW) & Umweltkontrolle (UK)

Zur Erhebung der Konstrukte Persönlichkeitsreife und Adaptivität auf möglichst ökonomische Art wurden die Unterskalen Persönliches Wachstum (PW) und Umweltkontrolle (UK) des Fragebogens zum Psychologischen Wohlbefinden von Ryff (1989) ausgewählt, die in mehreren Studien indikativ für diese Reifekonzepte waren, siehe dazu Kapitel 4.2. Die Skala Umweltkontrolle ist indikativ für Adaptivität, was sich im Item: Die Regelung meiner Finanzen und meiner persönlichen Angelegenheiten gelingt mit im Großen und Ganzen gut zeigt. Die Skala Persönliches Wachstum ist indikativ für Persönlichkeitsreife, was das Item: Ich habe das Gefühl, dass ich mich im Laufe der Zeit persönlich sehr viel weiter entwickelt habe belegen.

Die beiden Unterskalen PW und UK korrelieren nur gering miteinander ($r = .35$) (Schmutte & Ryff, 1997). Jede Skala hat 14 Items und wird sechs Abstufungen von Trifft sehr auf mich zu bis Trifft überhaupt nicht auf mich zu beantwortet. Die Interne Konsistenz der Unterskala UK liegt bei $a = .80$, während die Unterskala PW den wenig akzeptablen Wert von $a = .64$ in der Weisheitsstudie von Staudinger et al. (1997) erreicht.

6.3.9 Allgemeine Intelligenz (HAWIE-R)

Der Wortschatztest des HAWIE-R (Tewes, 1991) misst die Kenntnis von Wörtern und ihrer Bedeutung. Er wird häufig anstatt mehrerer Tests der HAWIE-R Testbatterie zur Messung Allgemeiner Intelligenz genutzt. Der Wortschatztest zeigt in Weisheitsstudien eine Interne Konsistenz von $a = .90$ (Mickler, 2004; Staudinger et al., 1997; Staudinger & Pasupathi, 2003). Die Erhebungsinstruktionen und die Bewertung der Antworten folgen dem standardisierten Testmanual von Tewes (1991).

Zur Verkürzung der Durchführung wurde nicht die vollständige Reihenfolge aller 32 möglichen Items angewendet. Es wurde solange auf einfache und durchschnittlich schwere

Items verzichtet, wie die Testung annehmen lässt, dass die Probanden über diesem Niveau liegen. Daher folgte auf das Einstiegsitem Apfel – Was bedeutet der Begriff Apfel? die ausgewählten Items 14. Oase, 18. Mandant, 21. Prestige und 25. Adäquat. Wurde eines dieser Items nicht gelöst, wurde das Leistungsniveau entsprechend der sonst ausgelassenen Items näher bestimmt. Bei richtiger Beantwortung dieser Items wurde die Manualreihenfolge fortgesetzt, mit: Nihilismus, Sequenz, Pragmatisch, Determinante, Thorium, Konkordant, Geoid. Der Wortschatztest wird im Telefoninterview durchgeführt, was der üblichen Interviewsituation sehr nahe kommt. Die Antworten der Teilnehmer wurden anhand des Manuals von Tewes (1991) bewertet und simultan dokumentiert, siehe Anhang D.1.

6.3.10 Demographischer Fragebogen

Die demographischen Daten der Erhebung lassen sich in mehrere Kategorien einteilen. Bei den Variablen Alter, Anzahl der Schul-, Studiums- und Ausbildungsjahre, Beruf und Anzahl der Kinder sind keine geschlossenen Antwortalternativen vorgegeben, sondern die Teilnehmer geben ihre Antwort über zur Verfügung stehende Lücken an. Dagegen werden die Variablen Geschlecht der Teilnehmer, Bildungsniveau, Berufstätigkeit, Familienstand und Psychotherapeutische Behandlung über Antwortalternativen erfasst. Das Bildungsniveau wurde über folgende vorgegebene Kategorien erfragt: kein Schulabschluss, Hauptschule/Volksschule, Mittlere Reife, Abitur, Hochschulabschluss, Promotion, Habilitation. Die Angabe der aktuellen Berufstätigkeit konnte entlang der Alternativen berufstätig/beurlaubt, in Ausbildung, in Rente, nicht berufstätig und sonstige erfolgen, zu welcher genauere Angaben gemacht werden konnten. Ausprägungen des Familienstands sind ledig, verheiratet, geschieden, verwitwet, und das letzte, dichotome, Item erfragt, ob sich die Teilnehmer bereits in einer psychiatrischen oder psychotherapeutischen Behandlung befunden haben oder nicht.

6.4 Vorgehen in der Datenerhebung

Nachfolgend wird die telefonische Erst-Kontaktaufnahme mit den Teilnehmern, der schriftliche Teil und der mündliche Teil der Datenerhebung beschrieben. In der schriftlichen Erhebung über das Fragebogenheft werden Weisheit, Persönlichkeitsreife, Adaptivität und demographische Variablen erfasst, während sich das spätere Telefoninterview der mündlichen Erhebung der Intelligenz widmet, vergleiche Tabelle 2. Auf die konzeptionelle Durchführung der einzelnen Erhebungen gehen die Abschnitte 6.4.2 und 6.4.3 ein. Abschnitt 6.4.1 dient der Beschreibung der telefonischen Erstkontaktaufnahme.

6.4.1 Telefonische Kontaktaufnahme

Entsprechend der oben beschriebenen Rekrutierung wurde mit den Teilnehmern aller drei Gruppen telefonisch Kontakt aufgenommen. Im Erstgespräch wurden Informationen über das Ziel der Studie, die Teilnahme und den Rahmen der Untersuchung gegeben, die nun vorgestellt werden.

Das Ziel der Validierung von Fragebogenverfahren zur Weisheitsmessung wurde je nach Wissens- und Interessesstand des Teilnehmers erläutert. Mögliche Aspekte dieser Erklärung waren, eine Definition von Weisheit, die Sinnhaftigkeit der Messung von Weisheit und Vorteile von Fragebogeninstrumenten. Als Definition von Weisheit wurde die Begriffsbestimmung des Brockhaus für Psychologie den Teilnehmern verwendet, die lautet: „Weisheit: ist eine hohe Kompetenz in fundamentalen Lebensfragen, die sich in höchstem Wissen und höchster Urteilsfähigkeit im Umgang mit schwierigen Problemen der Lebensplanung, Lebensgestaltung und Lebensbedeutung ausdrückt (Starke-Perschke et al., 2001, S. 682)“.

Die Teilnehmer wurden entsprechend ihrer Zugehörigkeit zu einer Teilnehmergruppe darüber informiert, welche Bedeutung ihre Teilnahme für die gesamte Untersuchung habe. Des Weiteren wurden Informationen über den formalen Rahmen der Studie, die zweigeteilte Datenerhebung mittels Fragebogenheft und anschließendem Telefonat, die Dauer der Erhebungen, den Schutz der Persönlichen Kontakt- und Testdaten der Teilnehmer und den Verwendungszweck der Daten gegeben. Die angekündigte nicht personalisierte Ergebnisrückmeldung über die Validität der Testverfahren, siehe Anhang G, steigerte einerseits die Teilnahmebereitschaft und versicherte andererseits die Anonymität der Datenauswertung. Bei Zustimmung zur Teilnahme wurden je nach Vorliebe die postalischen oder elektronischen Kontaktdaten der Teilnehmer zur Übersendung des Anschreibens, siehe Anhang A und B, sowie des eigentlichen Fragebogens, siehe Anhang C, verwendet.

6.4.2 Schriftliche Erhebung

Zum Bereich der schriftlichen Korrespondenz gehören das Anschreiben und das Fragebogenheft, dessen Instrumente in Abschnitt 6.3 methodisch vorgestellt wurden. An dieser Stelle erfolgt die Schilderung der formalen Informationen zu beiden Dokumenten.

Bei der postalischen Versendung wurde der Fragebogen und das jeweilige Anschreiben mit einem frankierten Rücksendeumschlag, adressiert an Prof. Dr. Manfred Schmitt, Universität Koblenz-Landau, versendet. Die Portokosten wurden von dem Fachbereich übernommen, so dass den Teilnehmern keinerlei Gebühren entstanden.

Das Anschreiben informierte über Ziel, Umfang und Dauer der Studie mit Fragebo-

generhebung und Telefoninterview. Für das Fragebogenheft benötigten drei Personen unterschiedlicher Expertise jeweils circa eine Stunde. Für das Telefonat wurde eine Dauer von 15 Minuten erwartet. Die Anleitung zur Bearbeitung des Fragebogenheftes ist im jeweiligen Anschreiben integriert und versichert die Vertraulichkeit der angegebenen Kontaktinformationen und Testergebnisse. Die Instruktion forderte die Teilnehmer dazu auf, nicht die Antwortalternative zu markieren, die sie als indikativ für Weisheit einschätzten, sondern davon ungeachtet jene Antwortalternative auszuwählen, die die Teilnehmer persönlich am besten beschreibe.

Die elektronisch versendeten Seiten des Anschreibens und des Fragebogens gleichen den postalischen. Um zu gewährleisten, dass die Handhabung des elektronischen Fragebogens der Papierversion entspricht, wurde mit der Formularfunktion geschützter Word 2003-Dokumente gearbeitet, in der der Wortlaut der Items nicht geändert, aber das markieren, korrigieren und Auslassen von Items in einer beliebigen Reihenfolge und in beliebigen Arbeitstempo bewusst möglich war. Da die Bearbeitung des digitalen Fragebogens nicht online erfolgte, wurden keine Unkosten für die digitale Bearbeitung erwartet und somit nicht erstattet. Der Fragebogen wurde als Download auf die Webseite <http://www.gerechtigkeitsforschung.de/Weisheit/Fragebogen.doc> im E-Mail-Anschreiben verlinkt. Diese Webseite wird von Prof. Dr. Schmitt, dem Erstbetreuer der Arbeit, seit Jahren für psychologische Forschungsarbeiten verwendet. Eine Rücksendung des bearbeiteten Dateiformulars erfolgte an die E-Mail-Adresse: Weisheit-2006@web.de, die speziell zu diesem Zweck erstellt wurde. Lediglich 8 Probanden mit sehr häufigem beruflichem Gebrauch eines Computers wählten die computergestützte Bearbeitung des digitalen Fragebogenäquivalentes.

6.4.3 Aufbau des Fragebogenheftes

Das Fragebogenheft ist eine Zusammenstellung der einzelnen Fragebögen und ermöglicht die Planung der zweiten, telefonischen, Erhebungsphase. Die Reihenfolge der Fragebögen im Testheft richtet sich nach dem jeweiligen Antwortformat der Tests, deren Länge sowie des zu messenden Konstruktes, siehe dazu Anhang C. Die methodische Beschreibung der verwendeten Instrumente erfolgte in Abschnitt 6.3.

Zu Anfang wurden die Weisheitstests 3D-WS und SAWS und die Skalen UK und PW gesetzt, die ein geschlossenes Antwortformat aufweisen. Im Fragebogenheft, Anhang C, gehören die Items 39 zum 3d-WS, Items 40 bis 53 erfassen UK, die darauf folgenden Items 54 bis 67 indizieren PW, während die Items 68 bis 107 den SAWS bilden.

Als Übergang zu den offenen Antwortformaten des TWR und AWR wurde der Satzergänzungstest WUSCT von Loevinger zwischengeschaltet. Da der TWR eine Replikation

eigener Erfahrungen fordert, ist er von der Anforderung niedriger einzustufen als der AWR, der es erfordert sich in eine fiktive Lage zu versetzen und das eigene Verhalten und Empfinden in dieser zu beschreiben. Mit der gewählten Reihenfolge, WUSCT vor TWR und AWR, steigert sich die geforderte Selbstoffenbarung somit sukzessive.

Wieder ein geschlossenes Antwortformat haben die sehr kurzen Instrumente PWS und PF, die sie daher für das Ende eines Fragebogens prädestinieren. Auf der letzten Seite erfolgte die Bestimmung eines individuellen dreistufigen Versuchspersonencodes und die Abfrage soziodemographischer Daten sowie die zeitliche Planung des Telefonates seitens der Teilnehmer, die ihre Wunschtermine angeben konnten.

Als Bearbeitungszeiträume des schriftlichen Materials wurde den Teilnehmern jeweils ein Zeitraum von 4 bis 6 Wochen zugestanden. Die Rekrutierung der Kontrollgruppe und der Psychotherapeuten sowie deren Datenerhebung wurden erst nach Beendigung der Erhebung der Nominierengruppe möglich. Insgesamt erstreckte sich die Stichprobenziehung und Datenerhebung über den Zeitraum von Juni 2006 bis Juli 2007.

6.4.4 Mündliche Erhebung

Das Telefoninterview dient der Erhebung der Leistungsmessung Allgemeiner Intelligenz. Der Zeitpunkt der Durchführung des jeweiligen Telefoninterviews wurde von den Teilnehmern nach eigenen Wünschen über das Fragebogenheft bestimmt. Jedes Telefonat wurde mittels eines Telefonleitfadens im Anhang D standardisiert durchgeführt, der die Instruktionen, die Reihenfolge der zu bearbeitenden Aufgaben sowie Möglichkeiten der Rücksprache über die Erhebung mit der durchführenden Diplomandin strukturierte.

Nach Zustimmung zur weiteren Teilnahme erfolgte die Testung der verbalen Fähigkeiten über den Wortschatztest des HAWIE-R, der Allgemeine Intelligenz widerspiegelt. Das standardisierte Vorgehen folgte dem Testmanual (Tewes, 1991) und dem Leitfaden im Anhang G. Das Vorgehen und die verwendeten Items des Wortschatztests wurden in Unterkapitel 6.3.9 beschrieben. Die Antworten der befragten Person wurden auf dem jeweiligen Dokumentationsbogen (Anhang D.2) festgehalten und anhand des Auswertungsmanuals bewertet. Der Endscore wurde den Teilnehmern nicht mitgeteilt. Die durchschnittliche Bearbeitungszeit über alle drei Teilnehmergruppen lag bei zirka 5 – 8 Minuten.

Im Anschluss an den Erhebungsteil des Telefonates, konnten fehlende soziodemographische Angaben im Fragebogen des Teilnehmers durch Nachfrage ergänzt werden. Bei Nichtbearbeitung des TWR oder AWR wurde Erläuterungen dieses Verhaltens notiert.

6.5 Methodische Umsetzung der statistischen Auswertung

Sämtliche statistische Analysen dieser Studie wurden mit SPSS Version 11.5 berechnet. Die Darstellung der verwendeten methodischen Verfahren folgt nicht chronologischen, sondern inhaltlichen Gesichtspunkten. Aufgrund der allgemein methodischen Ausrichtung dieses Abschnittes, wird die bisherige Indizierung von empirischen Variablen im Vergleich zu theoretischen Konstrukten ausgesetzt.

6.5.1 Interne Konsistenz und Interraterreliabilität

Die Reliabilität der Testverfahren dieser Studie wird über die Interne Konsistenz mittels Cronbachs α beurteilt. Cronbachs α ist das geeignete Maß für intervallskalierte und dichotome Testverfahren in einer querschnittlichen Erhebung. Dies trifft nach Kapitel 6.3 auf fast alle verwendeten Tests zu. Für die Gesamtscores der Maße TWR und AWR kann keine Angabe zur Internen Konsistenz errechnet werden. Sehr gute Werte Interner Konsistenz liegen über .90, gute Werte über .80, zufrieden stellende Werte erreichen .70. Ergeben sich jedoch niedrigere Werte, so sind die Ergebnisse anschließender Untersuchungen in ihrer Aussage nicht eindeutig, da eine mangelnde Reliabilität auf Messfehler hindeutet. Messfehler reduzieren die Genauigkeit einer empirischen Aussage oder Vorhersage. Systematische Messfehler beeinflussen darüber hinaus die inhaltliche Bedeutung der zu prüfenden Fragestellung, weshalb diese ihrer naturgemäß entweder vermieden oder kontrolliert werden sollten.

Zur Beurteilung der Interraterreliabilität bei den Testverfahren TWR, AWR wird Cohens κ bestimmt, wobei zufrieden stellende Werte .40 - .75 erreichen. Aufgrund ihrer Beurteilungsskalen werden diese Tests als intervallskaliert klassifiziert. Würde kein Intervallskalenniveau bei diesen Tests zu Grunde gelegt, müsste entweder Kendalls' Tau oder Spearmans' Rho als Koeffizient zur Güte der Beurteilerübereinstimmung gewählt werden.

6.5.2 Korrelationskoeffizienten

Die Wahl eines Korrelationskoeffizienten ist abhängig von den Skalenniveau der beteiligten Variablen. Werden Variablen korreliert, die Intervallskalenniveau aufweisen, erfolgt eine Produkt-Moment-Korrelation. Liegt das Skalenniveau eines der Merkmale dagegen im ordinalen Bereich wird Kendalls Tau bestimmt. Liegen zwei Variablen mit dichotomen Skalenniveau vor, ist der Phi-Koeffizient das passende Zusammenhangsmaß. Ein Punkt-biserialer Korrelationskoeffizient dagegen, bildet Zusammenhänge zwischen dichotomen und Intervallskalierten Variablen in dieser Studie ab. Die Annahme für Bestimmungen von Produkt-Moment-Korrelationsverhältnissen ist die bivariate Normalverteilung der intervallskalierten Variablen.

Eine Prüfung dieser Voraussetzung weist methodische Schwierigkeiten auf, weshalb im Allgemeinen lediglich eine Prüfung der Normalverteilung der einzelnen Variablen erfolgt (Bortz, 1999; Nachtigall & Wirtz, 2002). Der Signifikanztest von Korrelationen gilt als äußerst robust gegenüber der Verletzung dieser Verteilungsannahme (Nachtigall & Wirtz, 2002).

Zur Prüfung eines von 0 abweichenden Korrelationsverhältnisses über Fisher Z-Transformationen der zu testenden Korrelation werden Vergleiche mit dem kritischen Wert entsprechend des gewählten Signifikanzniveaus und der Stichprobengröße angestellt. Die Fisher Z-transformierten Korrelationskoeffizienten werden der Tabelle M von Diehl und Arbinger (2001) entnommen. Die Hypothesentestung erfolgt einseitig auf einen Niveau von $\alpha = .05$, dem ein kritischer Z-Wert von 1,645 entspricht. Bei einer gruppenspezifischen Fragestellung stehen maximal 38 Versuchspersonen ohne fehlende Werte zur Verfügung, während gruppenübergreifend maximal 114 Teilnehmer analysierbar sind. Einschätzungen der Effektstärke von Korrelationen ergeben über r . Sollte in den Voranalysen eine derart bedeutsame Korrelation zu einer Testvariable vorliegen, so wird der Effekt der Störvariablen für folgende Analysen über Semipartialkorrelation aus der Testvariablen auspartialisiert.

Unterschiedliche Vorgehensweisen zum Vergleich von Korrelationsverhältnissen und Partialkorrelationskoeffizienten werden in Diehl und Arbinger beschrieben (2001) und mittels dortiger Tabellen L und M bestimmt. Zur Signifikanzprüfung von Partialkorrelationen (Diehl & Arbinger, 2001), wird der errechnete Partialkorrelationskoeffizient mit einem kritischen Korrelationsniveau verglichen (Tabelle L), welches die Stichprobengröße berücksichtigt. Die Effektstärke beim Vergleich von Korrelationskoeffizienten ist q . Wenn das Korrelationsverhältnis zweier Variablen mit einer dritten verglichen wird, findet das Verfahren von Williams (1959; nach Diehl & Arbinger, 2001) Anwendung. Dies setzt mindestens 20 Versuchspersonen und bivariate Normalverteilungen der beteiligten Variablen voraus. Über die Formeln zur Bestimmung der Prüfgröße T_W (ebd.) erfolgt der Test gegen das nach Stichprobengröße, Freiheitsgraden und Alphafehlerniveau bestimmte T_{KRIT} aus Tabelle E (Diehl & Arbinger, 2001).

Die Bildung des multiplen Korrelationskoeffizienten erfolgt über dessen Regression. Für jeden der fünf Weisheitstests wird eine Regression des Testwertes durch die restlichen vier Tests über die Gesamtstichprobe durchgeführt. Bivariate Normalverteilung der Testwerte, sowie deren Arrayverteilungen unter Homoscedastizität ist Voraussetzung für die Durchführung von Regressionsanalysen. Verletzungen dieser Voraussetzungen führen zur Erhöhung der α und β -Fehler-Wahrscheinlichkeiten. Nach Bortz (1999) erübrigt sich die Prüfung dieser Voraussetzung, falls mehr als 40 Personen und weniger als 10 Prädiktoren vorliegen. Aus theoretischer Sicht ist eine Größenordnung von $R > .70$ eines multiplen Korrelationskoeffi-

zienten gut, während in der Praxis Werte $R > .60$ erwartet werden. In Testbatterien können jedoch viel geringere multiple Korrelationskoeffizienten praktischen Nutzen beweisen, wenn sie eine eigenständige Facette des zu messenden Konstruktes erfassen. Die Effektstärke multipler Korrelationen ist f^2 .

6.5.3 Mittelwertsvergleiche

Auch für die Mittelwertsvergleiche der vorliegenden Studie sind die Skalenniveaus der zu vergleichenden Variablen sowie deren Verteilungseigenschaften bestimmend für die Wahl der methodischen Umsetzung des Vergleichs. Da keine Hypothesen vorliegen, die Unterschiede zwischen den Teilnehmergruppen auf mehreren Tests gleichzeitig vorhersagen, wozu eine mehrfaktorielle multivariate Varianzanalyse insbesondere hinsichtlich der größeren Teststärke das Verfahren der Wahl wäre (Backhaus, 1999), werden nur die Einzelvergleiche geprüft.

Ob ein Auswertungsdesign für unabhängige oder Abhängige Stichproben Anwendung findet, ist in der vorliegenden Studie derart zu entscheiden, ob die zu prüfende Variable mit den Matchingkriterien zur Bildung der parallelisierten Gruppen zusammenhängt oder nicht. Wird ein positiver Zusammenhang zwischen diesen Variablen angenommen, hat die Verwendung von abhängigen Designs Vorteile in der Power des Testverfahrens (Diehl & Arbinger, 2001). Durch eine positive Korrelation der Messwertreihen verringert sich der Standardfehler der Prüfgrößen, was zu einer Verringerung des Betafehlerniveaus führt. Bei negativer Korrelation steigt der Standardfehler der Prüfgrößen der abhängigen Analyse auf ein Niveau über dem der unabhängigen Analysemethoden. Je höher die Korrelation ist, desto höher ist die Wahrscheinlichkeit die Nullhypothese abzulehnen. Zu beachten ist, dass bei Ausbleiben einer signifikanten Korrelation die Power des T-Tests für abhängige Stichproben nicht in dem Maße gehoben wird, dass die Einbußen der Halbierung der Freiheitsgrade gegenüber dem Design für unabhängige Stichproben wettgemacht werden. Anhand der Korrelation kann nachträglich erkannt werden, ob die Verwendung von abhängigen Analyseverfahren sinnvoll war. Bei signifikant negativer Korrelation und Beibehaltung der Nullhypothese empfehlen Diehl und Arbinger (2001) auf Analysemethoden für unabhängige Stichproben umzuschwenken.

Aufgrund dieser Überlegungen wird in der vorliegenden Studie folgendes, konservatives Vorgehen gewählt. Können über den Zusammenhang zwischen den Matchingkriterien und der zu testenden Variable nur Vermutungen angestellt werden, empfiehlt sich daher, eine Analyse über abhängige Stichprobendesigns zu rechnen und die Korrelation der Variablen zu inspizieren. Wird die Nullhypothese bei fehlender positiver Korrelation beibehalten, kann die zusätzliche Verwendung von unabhängigen Analysemethoden Aufschluss darüber liefern, ob

die unterschiedliche Anzahl der Freiheitsgrade zu einer Verwerfung der Hypothese führt.

Vergleiche von Mittelwerten intervallskaliertter Variablen abhängiger Gruppen ergeben über T-Test für abhängige Stichproben. Die Beurteilung der Effektstärke erfolgt über d' (Diehl & Arbinger, 2001). Bei kleinen Stichproben wird Wilcoxon's Vorzeichen Rangtest für abhängige Gruppen durchgeführt, da eine Aussage über die Normalverteilung der Populationsmitteldifferenzen in diesem Falle nicht zu treffen ist. Speziell bei nicht normalverteilten Populationsmitteldifferenzen ist die Power dieses Tests mindestens so gut wie die eines T-Tests. Mit steigender Stichprobengröße normalisiert sich die zu prüfende Differenzverteilung.

Thematisiert eine Hypothese Mittelwertsunterschiede intervallskaliertter Variablen unabhängiger Stichproben, ergeht die Analyse über einen W-Test (Diehl & Arbinger, 2001). Dieser ist dem T-Test im Hinblick auf die Robustheit überlegen. Ein T-Test für unabhängige Stichproben setzt normalverteilte Populationen und Varianzheterogenität voraus. Der W-Test ist gegenüber Verletzungen der Voraussetzung normalverteilter Populationswerte bei annähernd gleichen n und symmetrischen Populationen robust. Hier ist die Effektgröße d . Um einer Alphafehlerkumulierung entgegen zu wirken, erfolgt für jede zu testende Fragestellung eine Minderungskorrektur des Alphaniveaus nach Bonferroni.

Der Friedmantest dient Vergleich von Mittelwerten ordinalskaliertter, nichtparametrischer Variablen aus mehr als zwei abhängigen Stichproben. Bei der Verwerfung der Nullhypothese bedeutet dies, dass mindestens zwischen zwei der getesteten Gruppen ein signifikanter Kontrast vorherrschen muss (Diehl & Arbinger, 2001).

Zum Test von Häufigkeitsunterschieden zwischen nominalskalierten Variablen verschiedener Gruppen wird Fishers exakter Test bestimmt. Aufgrund zu geringer Zelhäufigkeiten sind die Voraussetzungen einer Chi²-Teststatistik hier nicht erfüllt.

6.6 Statistische Hypothesen

Der folgende Abschnitt präsentiert die zu testenden statistischen Hypothesen zu den bereits theoretisch hergeleiteten Hypothesen in identischer Reihenfolge. Die Darstellung bemüht sich um Sparsamkeit, weshalb einzelne Hypothesen bestmöglich zusammengefasst werden.

6.6.1 Statistische Hypothesen der Inneren Validität

Für die Prüfung von Hypothese 1 werden multiple Korrelationskoeffizienten der einzelnen Weisheitstestverfahren 3d-WS, SAWS, PWS, TWR und AWR berechnet. R_y sollte sich im Falle der Gültigkeit der Hypothese nicht von $R = .60$ signifikant unterscheiden, bei einem

einseitig getesteten Alphafehlerniveau von 5%.

$$H_0: R_y < .60$$

$$H_1: R_y = .60$$

Hypothese 2 und 3 beziehen sich auf Vergleiche von Korrelationskoeffizienten. Hypothese 2 testet die Korrelation zwischen TWR und PWS einseitig gegen das kritische Niveau von $r = .30$ auf dem 5% Alphafehlerniveau.

$$H_0: r_{PWS,TWR} = .30$$

$$H_1: r_{PWS,TWR} < .30$$

Hypothese 3 ist erfüllt, wenn die Korrelation zwischen SAWS und 3d-WS signifikant höher ist als alle anderen bivariaten Korrelationen dieser beiden Tests zu den Tests PWS, TWR und AWR. Zur Vereinfachung wird die Korrelation von 3d-WS und SAWS lediglich einseitig gegen die höchste Korrelationen (r_{xy}) dieser sechs bivariaten Korrelationen: $r_{SAWS,PWS}$, $r_{SAWS,TWR}$, $r_{SAWS,AWR}$, $r_{3d_WS,PWS}$, $r_{3d_WS,TWR}$ und $r_{3d_WS,AWR}$ getestet.

$$H_0: r_{SAWS,3d-WS} = r_{xy}$$

$$H_1: r_{SAWS,3d-WS} > r_{xy}$$

6.6.2 Statistische Hypothesen der Äußeren Validität

Die Hypothesen 4 bis 6 beziehen sich auf Mittelwertsunterschiede der einzelnen abhängigen Teilnehmergruppen auf den einzelnen Weisheitstests 3d-WS, SAWS, PWS, TWR und AWR, siehe Tabelle 6.2. Dazu werden für jeden Mittelwertsvergleich W-Tests berechnet. Es handelt sich hierbei nicht um zusammenfassende Hypothesen, weshalb $\alpha = .05$ verwendet wird. Sollten die Wertereihen der Gruppen Nominierte und Psychotherapeuten nicht miteinander korrelieren, es sich also um unabhängige Gruppen handeln, wird mit der Methode für abhängige Stichproben keine Überschätzung des wahren Wertes erfolgen.

Tabelle 6.2. Statistische Hypothesen der Äußeren Validierung.

Hypothese 4	Hypothese 5	Hypothese 6
$H_0: M_{NOM} = M_{KG}$	$H_0: M_{NOM} = M_{PSY}$	$H_0: M_{PSY} = M_{KG}$
$H_1: M_{NOM} > M_{KG}$	$H_1: M_{NOM} > M_{PSY}$	$H_1: M_{PSY} > M_{KG}$

6.6.3 Statistische Hypothesen der Konvergenten Validität

Die Hypothesen 7 bis 9 prüfen bivariate Korrelationsverhältnisse der Weisheitstestverfahren (3d-WS, SAWS, PWS, TWR und AWR) mit allen Persönlichkeitsreifetests (WUSCT, PF und PW) gegen das Niveau von $r = .30$ für die Gesamtgruppe. Übersteigen die Korrelationen dieses Niveau, ist von einer bedeutsamen konvergenten Beziehung auszugehen. Es wird hier einseitig auf dem Alphafehlerniveau von 5% für jede bivariate Beziehung getestet.

$$H_0: r_{\text{Weisheitstest, Persönlichkeitsreifetest}} < .30$$

$$H_1: r_{\text{Weisheitstest, Persönlichkeitsreifetest}} = .30$$

6.6.4 Statistische Hypothesen der Diskriminanten Validität

Zur Absicherung des Konstruktes werden diskriminante Zusammenhänge der Hypothesen 10 bis 13 geprüft, die genau in entgegengesetzter Größenordnung zu denen der konvergenten Hypothesen 7, 8 und 9 stehen. Die Diskriminanten Konstrukte sind dabei jeweils Adaptivität (UK), Verbale Intelligenz, Alter und Geschlecht. Da hier ebenfalls keine zusammengesetzten Hypothesen getestet werden, erfolgt die Prüfung der bivariaten Zusammenhänge auf dem Alphafehlerniveau von 5% für jeden Weisheitstest. Für die Variable Geschlecht sind hier punkt-biseriale Korrelationskoeffizienten zu bestimmen, während die anderen Variablen über Produkt-Moment-Korrelationen geprüft werden.

$$H_0: r_{\text{Weisheitstest, Diskriminantes Konstrukt}} = .30$$

$$H_1: r_{\text{Weisheitstest, Diskriminantes Konstrukt}} < .30$$

6.7 Voranalysen zu den Testverfahren

Hier folgen die Darstellung und Analyse der Mittelwerte, der Interraterreliabilitäten und Internen Konsistenzen der Testverfahren, die Faktorenbildung, sowie die Analyse von Einflüssen der Variablen Antwortlänge und Verbaler Intelligenz auf die Testverfahren mit offenem Antwortformat. Sich daraus ergebende Konsequenzen können die Auswertung der zentralen Hypothesen dieser Validierungsstudie wesentlich beeinflussen.

6.7.1 Mittelwerte der Testverfahren

Zur Prüfung der Hypothesen wird im Ergebnisteil vermehrt auf die Mittelwerte der einzelnen Teilnehmergruppen auf den Weisheitstests zurückgegriffen. In Tabelle 6.7.1 werden daher

Mittelwerte, Standardabweichungen und Standardfehler präsentiert. Tabelle F.2 des Anhangs komplettiert die Darstellung um andere Teilnehmergruppierungen. Die Analyse dieser Daten erfolgt im Ergebnisteil. Der Maximalwert der Skala SAWS ist 240, in der PWS ist er 17, im TWR 5 und im AWR 20. Die Faktorbildung der 3d-WS Skala wird in 6.7.4 beschrieben.

Tabelle. 6.3 Mittelwerte der Teilnehmergruppen auf den Weisheitstestverfahren.

Gruppierung	<i>M</i>	<i>N</i>	<i>SD</i>	<i>SE</i>
3DWS				
Gesamt	,00	114	,88	,08
NOM	,10	38	,84	,14
KG	-,03	38	,90	,15
PSY	-,08	38	,90	,15
SAWS				
Gesamt	171,07	113	20,90	1,97
NOM	169,30	37	17,69	2,91
KG	169,89	37	25,74	4,23
PSY	173,41	37	19,06	3,13
PWS				
Gesamt	13,95	111	2,22	,21
NOM	14,06	36	2,50	,42
KG	14,00	36	1,90	,32
PSY	13,64	36	2,28	,38
TWR				
Gesamt	3,74	102	,75	,07
NOM	3,57	30	,82	,15
KG	3,77	30	,77	,14
PSY	3,97	31	,71	,13
AWR				
Gesamt	9,82	114	2,87	,27
NOM	9,21	38	2,77	,45
KG	9,84	38	2,80	,45
PSY	10,39	38	2,97	,48

Anmerkungen. Der 3d-WS wird als Faktor aus drei Subskalen gebildet, daher ist der Mittelwert 0.

6.7.2 Reliabilitätsanalysen der Testverfahren

Zur Bestimmung der Beurteilerübereinstimmung wurde die Interraterreliabilität für den Gesamtscore TWR und den Summenscore des AWR ermittelt. Eine gute Übereinstimmung in Cohens Kappa wurde für beide Testverfahren berechnet (Werte in analoger Reihenfolge, $\kappa = .63$, näherungsweise $T^2 = 8,75$, $p < .01$; $\kappa = .72$ näherungsweise $T^2 = 16,74$; $p < .01$).

Die Reliabilität der Testverfahren wurde über die Interne Konsistenz bestimmt. Gute Interne Konsistenzen haben SAWS ($\alpha = .90$; $n = 104$), PW ($\alpha = .89$; $n = 112$), UK ($\alpha = .87$; $n = 111$) und 3d-WS ($\alpha = .81$; $n = 108$). Zufriedenstellende Werte erreichen die Tests WUSCT ($\alpha = .78$; $n = 106$), PF ($\alpha = .76$; $n = 111$) und der Wortschatztest des HAWIE-R ($\alpha = .76$; $n = 105$). Nicht zufriedenstellend ist die Reliabilität des PWS mit $\alpha = .61$ ($n = 91$). Als Post hoc gewählte Itemzusammenstellung ist das niedrige Alpha-Niveau des PWS einsehlich, und erlaubt weitere Analysen dieses Tests.

6.7.3 Anpassung der Auswertungs- und Bewertungskategorien von TWR & AWR

Die offenen Antworten in TWR und AWR erreichen nur in wenigen Fällen die von Helson und Srivastava (2002) geforderte Mindestwortanzahl von 70 für gültige Antworten. Im TWR sind dies 42 Teilnehmer, im AWR 29 und in beiden Tests 22 Personen. Um möglichst viele Daten auswerten zu können, bieten sich verschiedene Strategien an, die hinsichtlich der Verteilungseigenschaften der Testwerte beider Tests ausgewählt und vorgestellt werden.

Die Bewertungskriterien des TWR warf für die Analyse der hier vorliegenden Anwendung ein Problem auf. Es gibt in den gegebenen Antworten einige Fälle, bei denen aus der Art der Antwort entweder eine klare Nichtzuschreibung von Weisheit auf die eigene Person getroffen wird, oder die Beantwortung der Frage dies nahe legt durch – oder nein. Das Bewertungsschema verdeutlicht, dass keine der 5 Kategorien einen solchen Fall abdeckt. Die betreffenden 12 Fälle wurden bei weiteren Analysen des Tests ausgeschlossen.

Die Werteverteilung des TWR ($M = 3,73$; $SD = ,77$) hat eine Rechtsschiefe $.25$ ($S_E = .24$) und einen negativen Exzess mit $-.81$ ($S_E = .47$). Die Korrelation zu Antwortlänge im TWR mit $r = .51^{††}$ ($p < .01$) ist in einseitiger Testung signifikant und erklärt im bivariaten Fall 25% Varianz der TWR-Statistik. Da die Mindestantwortlänge nicht von allen Teilnehmern erreicht wurde ($M = 73,80$; $SD = 67,69$) und der Range der Wortanzahl in TWR mit 425 Worten groß ist, kann die Antwortlänge einen Effekt auf die Bewertung des Tests haben. Der Charakter dieses Einflusses ist jedoch nicht als systematische Störvariable zu bezeichnen, sondern als wahre Varianz die den Testwert beeinflusst. In den folgenden Analysen wird daher mit der

Gesamtbewertung des TWR ohne zusätzliche Kriterien gearbeitet.

Die Werteverteilung der vier Dimensionen des AWR über die Gesamtgruppe zeigt starke Bodeneffekte für die Dimensionen Emotionales Verständnis, Moralische Komplexität und Kognitive Differenziertheit (in analoger Reihenfolge, Schiefe = 0,90; 1,45; 0,71 (*SD* jeweils .23), siehe Tabelle 6.4. Die Dimension Prozedurales Wissen hat die geringste Linkschiefe mit 0,17. Analysiert wird nach Helson und Srivastava (2002) der Summenscore dieser vier Dimensionen. Die Schiefe mit 0,50 (*SD* = ,23) dieses Summenscores ($M = 9,74$, $SD = 2,87$) zeigt weniger ausgeprägte Bodeneffekte in der Gesamtgruppe als die drei kritischen Dimensionen.

Tabelle 6.4 Verteilungswerte des Summenscores und der Dimensionen des AWR in Gesamtgruppe, nach dem Kriterium der Wortanzahl = 70 und des Mehrdimensionalen Antwortstils.

AWR Dimensionen	<i>M (SE; SD)</i>	<i>Schiefe (SE)</i>	<i>Exzess(SE)</i>	<i>r</i> Antwortlänge
Gesamt (N = 114)				
Emotion	2,11 (.13; 1,36)	,90 (.23)	-,46 (.45)	.53 ^{††}
Moral	1,74 (.11; 1,17)	1,45 (.23)	,87 (.45)	.43 ^{††}
Kognition	2,22 (.12; 1,32)	,71 (.23)	-,70 (.45)	.56 ^{††}
Prozedural	3,46 (.10; 1,02)	,17 (.23)	-,89 (.45)	.33 ^{††}
Summenscore	9,74 (.27; 2,87)	,50 (.23)	-,34(.45)	.64 ^{††}
Wortanzahl = 70 (<i>n</i> = 29)				
Emotion	3,38 (.27; 1,43)	-,34 (.43)	-1,14 (.85)	.17
Moral	2,66 (.27; 1,47)	,29 (.43)	-1,37 (.85)	.10
Kognition	3,17 (.21; 1,14)	,21 (.43)	-,16 (.85)	.49 ^{††}
Prozedural	3,79 (.17; 0,90)	,13 (.43)	-1,20 (.85)	.34
Summenscore	12,62 (.45; 2,43)	,17 (.43)	-,27 (.85)	.37
Mehrdimensionalität (<i>n</i> = 91)				
Emotion	2,40 (.15; 1,39)	,58 (.25)	-,89 (.50)	.46 ^{††}
Moral	1,92 (.13; 1,24)	1,11 (.25)	,03 (.50)	.36 ^{††}
Kognition	2,52 (.14; 1,32)	,39 (.25)	-,94 (.50)	.47 ^{††}
Prozedural	3,49 (.11; 1,00)	,15 (.25)	-,77 (.50)	.33 ^{††}
Summenscore	10,45 (.28; 2,71)	,41 (.25)	-,38 (.50)	.58 ^{††}

Anmerkungen. Die Abkürzungen der Unterskalen des AWR bedeuten: Emotion= emotionales Verständnis, Moral = Moralische Komplexität, Kognition = Kognitive Differenziertheit, Prozedural = Prozedurales Wissen. Die Dimensionen des AWR werden von 1 -5 in der Ausprägung bewertet. Die letzte Spalte gibt jeweils die Korrelationen zu Antwortlänge wieder.

^{††} $p = .01$, [†] $p = .05$ (einseitig).

Die einseitig positiven Korrelationen zwischen der Antwortlänge und den AWR-Dimensionen in analoger Reihenfolge liegen bei .53, .43, .56 und .33 und für den Gesamtsummenscore bei .64, die alle auf dem 1% Niveau signifikant sind. Auch hier ist Antwortlänge als systematischer Faktor wahrer Varianz anstatt Fehlervarianz zu beurteilen. Die Stichprobenwerte der Antwortlänge im AWR von $M = 55,98$ ($SD = 47,030$) und $Range = 290$ liegen unter denen des TWR. Wird die Höhe der Korrelationen mit den Beurteilungen des AWR verglichen, zeigt sich, dass die kurzen Antworten vor allem Aspekte der Prozeduralen Dimension thematisieren, während bei längeren Antworten auch andere Dimensionen angesprochen werden.

Die Verwendung der Mindestwortzahl von 70 verringert daher nicht nur die Schiefe der Dimensionen, sondern hebt gleichzeitig deren Mittelwerte und senkt den Zusammenhang mit der Wortanzahl, was Tabelle 6.3 zeigt. Der Summenscore über diese 29 Personen zeigt dies ebenso ($M = 12,62$; $SD = 2,43$). Einzig die Korrelation zwischen Kognitiver Differenziertheit und Antwortlänge bleibt signifikant.

Um den Bodeneffekt der drei kritischen Dimensionen zu verringern, könnten alternativ die Personen ausgeschlossen werden, die auf diesen Dimensionen geringste Bewertung erzielen, also einzig auf der Prozeduralen Dimension scoren. Die verbleibenden 91 Teilnehmer haben einen mehrdimensionalen Antwortstil, siehe Tabelle 6.3. Die Mittelwerte und die Schiefe dieses Kriteriums sind vergleichbar aber besser als die der Gesamtpopulation.

6.7.4 Faktorbildung

Für den Weisheitstest 3d-WS (Ardelt, 2003) sowie des Gesamtweisheitstests von Helson und Srivastava (2002), der hier als HS eingeführt wird, muss ein Faktor über die zugrunde liegenden Dimensionen bzw. einzelnen Weisheitstests gebildet werden. Dazu wird der erste Faktor per Hauptachsenanalyse der Korrelationsmatrix extrahiert. Die Bestimmung der Faktorwerte der Dimensionen ergeht über deren Regression.

Entsprechend den Ergebnissen von Ardelt (2003) entfällt auf die Reflektive Unterskala des 3d-WS der größte Faktorwert mit .67, während die Faktorenwerte der beiden anderen Unterskalen niedrig sind (.20 und .17). Das Kaiser–Meyer–Olkin-Kriterium beurteilt den Faktor des 3d-WS als mäßig (Bühner, 2006), siehe Tabelle 6.4.

Als nächstes wurden die Faktorenvarianten des Weisheitsfaktors nach Helson und Srivastava gebildet. Die HS-Varianten, HS_{GES} , HS_{MD} und HS_{WORT} aus den Tests PWS, TWR und der entsprechenden AWR-Variante, gesamt ($_{GES}$), Mehrdimensionalität ($_{MD}$) und Wortanzahl ($_{WORT}$) zusammen.

Tabelle 6.5 Faktorenbildung 3d-WS-Faktor und HS-Faktor-Versionen.

Faktoren	n	Faktorwerte der Facetten			COV	Varianz- aufklärung	KMO	Bartlett-Test		
		1.	2.	3.				χ^2	df	p
3d-WS	114	,67	,20	,17	,77	55,71 %	,63	55,71	3	< .01
HS _{GES}	102	,96	,03	,00	,94	39,22 %	,51	25,34	3	< .01
HS _{MD}	83	1,01	-,01	-,03	,78	39,75 %	,47	19,62	3	< .01
HS _{WORT}	28	1,01	-,02	-,03	,80	39,46 %	,48	5,66	3	.13

Anmerkungen. KMO ist die Abkürzung für Kaiser-Meyer-Olkin-Kriterium. Die HS-Faktoren greifen auf die AWR-Varianten zurück. HS_{GES} steht für Gesamtstichprobe, HS_{MD} für Mehrdimensionalität und HS_{WORT} für Erfüllen der Mindestwortanzahl ≥ 70 . Die 3d-WS Facetten sind erstens die Reflektive, zweitens die Kognitive und drittens die Affektive Facette. Die HS-Facetten sind jeweils erstens TWR, zweitens AWR und drittens PWS. $p < .01$ steht hier für hohe Signifikanz des Bartlett-Tests.

Die Analyse der Faktorenversionen des HS belegt deren sehr hohe Übereinstimmung. Zusammenfassend kann gesagt werden, dass alle HS-Faktoren durch den TWR gebildet werden, was die Faktorwerte des TWR (.96, 1,01 und 1,01) verdeutlichen. Die Varianzaufklärung der drei Varianten liegt jeweils bei etwa 39,5 %, siehe Tabelle 6.5. Entsprechend des Kaiser-Meyer-Olkin-Kriteriums, erreicht lediglich HS_{GES} knapp das auswertbare Niveau von über .50. Die Faktorwerte der HS-Faktoren replizieren nicht die Ergebnisse der Studie von Helson und Srivastava (2002). In jener Studie erreichte der AWR die höchste Faktorladung zum gemeinsamen Faktor HS ($\lambda_{AWR} = .80$, $\lambda_{TWR} = .63$ und $\lambda_{PWR} = .59$). Die Faktorladungen des HS_{GES} in der vorliegenden Studie sind dagegen im Einzelnen $\lambda_{TWR} = .97$; $\lambda_{AWR} = .46$ und $\lambda_{PWS} = .17$. Somit sind Ergebnisse des HS aus jener Studie des HS_{GES} in dieser Studie nicht zu erwarten. Da HS_{GES} den TWR repräsentiert, wäre eine Beschreibung der Ergebnisse für HS_{GES} eine Redundanz zu den entsprechenden Ergebnissen des TWR. Aus diesem Grund wird auf eine Ergebnispräsentation des Faktors HS im Weiteren verzichtet.

6.7.5 Einflüsse der Antwortlänge auf WUSCT

Für den Satzergänzungstest der Ego-Entwicklung wird eine leichte Beziehung zwischen den Testwert und der Gesamtzahl der verwendeten Wörter im Test angenommen. In dieser Studie übersteigt deren Korrelation mit $r = .61$ ($p < .01$; $R^2 = .37$) diese Größenordnung.

6.7.6 Einflüsse der Verbalen Intelligenz auf TWR, AWR und WUSCT

Die Verbale Intelligenz der Teilnehmer kann einen Einfluss auf deren Antwortqualität in den offenen Antwortformaten des TWR, AWR und des WUSCT haben. Es werden daher die biva-

riaten Korrelationen zwischen diesen Testverfahren und der gemessenen Verbalen Intelligenz über den Wortschatztest des HAWIE-R ermittelt. Tabelle 6.6 belegt, dass keinerlei Korrelation signifikant wird. Für die höchste, nicht signifikante Korrelation, belegt die Varianzaufklärung von 3%, dass der Einfluss von Intelligenz auf Ego-Entwicklung vernachlässigbar ist. Auf eine Bereinigung der Tests für die Effekte der Verbalen Intelligenz wird daher verzichtet.

Tabelle 6.6 Bivariate Korrelation zwischen Verbaler Intelligenz und WUSCT, TWR, und AWR.

	WUSCT	TWR	AWR _{GES}
HAWIE-R	.17 n.s	-.10 n.s	-.08 n.s
<i>n</i>	106	97	106

Anmerkungen. Die Abkürzung AWR_{GES} steht für Gesamtstichprobe.
^{††}*p* > .01, [†]*p* > .05 (einseitig).

6.8 Voranalyse der Teilnehmergruppen

Neben den Matchingprofilvariablen der Nominierten und deren Kontrollpersonen Alter, Geschlecht, Beruf und Bildungsniveau werden die deskriptiven Variablen Berufstätigkeit, Gesamtdauer der Schul- und Ausbildungszeit, der Familienstand und die Anzahl der Kinder beschrieben und verglichen. Zusätzlich erfolgen Vergleiche der Gruppen hinsichtlich der Durchführung eigener Psychotherapeutischer Behandlungen und der Verbalen Intelligenz.

6.8.1 Alter

Die 114 Personen der Gesamtstichprobe sind im Mittel 52,22 Jahre alt ($SD = 13,99$). Diese teilen sich in drei Gruppen mit jeweils 21 Männern und 17 Frauen ($n = 38$). Die Altersspanne der Nominierten (24 - 85 Jahre), deren Kontrollpersonen (25 - 85 Jahre) und der Psychotherapeuten (26 - 86 Jahre) sind nahezu identisch. Bei einem bonferronikorrigierten Alphafehlerniveau von $\alpha_v = 0,0167$ unterscheiden sich das Alter der Kontrollgruppe und der Psychotherapeuten nicht von dem der Nominierten. Tabelle 6.7 präsentiert die Mittelwerte ($M_N = 54,43$; $M_K = 54,47$; $M_P = 48,00$), Standardabweichungen ($SD_N = 14,84$; $SD_K = 14,63$; $SD_P = 11,50$), und Teststatistiken für abhängige Gruppen in identischer Reihenfolge.

Tabelle 6.7 Altersvergleiche der Teilnehmergruppen.

Gruppe	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i> (2-seitig)
Gesamt ^a	52,25	13,94			
Nominierte ^b	55,13	14,84			
Kontrollgruppe ^b	54,13	14,63	0,01	74	.99 n.s.
Psychotherapeuten ^b	48,00	11,50 ^c	2,06	74	.04 n.s.

Anmerkungen. ^aN = 114. ^bn = 38.
 $\alpha_v < 0.0167$

6.8.2 Beruf

Die Gruppe der Psychotherapeuten hat keine Varianz auf der Variable Beruf. Für das Matching der gepaarten Gruppen ist die Berufsbezeichnung eine wichtige Variable und zeigt in den gewonnenen Paarungen eine gute Übereinstimmung. Werden die Berufe numerisch kodiert und die gepaarten Gruppen daraufhin im Wilcoxon-Test verglichen, sind diese weder signifikant unterschiedlich, noch deutet die exakte Wahrscheinlichkeit mit $p = .003$ auf eine Zufallsverteilung hin. Bei 34 Paarungen stimmt der Beruf perfekt überein. In den anderen Fällen wurde ein möglichst eng verwandtes Berufsfeld ausgewählt, was Tabelle F.1 im Anhang verdeutlicht.

6.8.3 Bildungsniveau und Berufstätigkeit

27 Psychotherapeuten haben ein Diplom, die übrigen sind darüber hinaus promoviert. Bei den gepaarten Gruppen haben jedoch 18 Matchingpaare keinen Hochschulabschluss. Daher erzielt die homogene Gruppe berufstätiger Psychotherapeuten in den Variablen Bildungsniveau, Anzahl der Schul- und Ausbildungsjahre sowie Berufstätigkeit höhere Werte als die beiden gepaarten Gruppen, die sich nicht unterscheiden.

Für Bildungsniveau wird der Friedman-Test über alle Gruppen signifikant, $\chi^2 = 111,00$ ($df = 1$, $n = 114$, $p < .01$). Die gepaarten Gruppen unterscheiden sich nicht im entsprechenden Wilcoxon-Test $Z = -7,47$ (bei $p < .01$). Beim Vergleich der Schul- und Ausbildungsjahre ($M_{NOM} = 17,84$; $M_{KG} = 17,89$; $M_{PSY} = 24,35$; $SD_{NOM} = 4,41$; $SD_{KG} = 4,61$; $SD_{PSY} = 4,80$) wird eine ANOVA für abhängige Stichproben signifikant, ($F = 34,70$; $df_1 = 2$, $n = 38$ $p < .01$; partielles $\varepsilon^2 = .66$). Die Gruppenunterschiede in der Berufstätigkeit sind signifikant (Fishers exakter Test = 18,38; $p < .01$; $CC = ,35$; $p = .02$). Tabelle 6.8 illustriert den Berufstätigkeitsstatus der Teilnehmergruppen in Prozent. Während über 20% der gepaarten Gruppen bereits nicht mehr erwerbstätig sind, trifft dies auf keinen der Psychotherapeuten zu.

Tabelle 6.8 Berufstätigkeitsstatus der Teilnehmergruppen in Prozent.

Gruppe	Berufstätigkeit			
	berufstätig	in Ausbildung	in Rente	nicht berufstätig
NOM	65,7 %	2,5 %	26,3 %	5,0 %
KG	71,1 %	5,3 %	21,1 %	2,5 %
PSY	89,5 %	10,5 %	0%	0%

Anmerkungen. Die Prozentzahlen basieren auf 38 Teilnehmern je Gruppe, 63 Männern und 51 Frauen.

6.8.4 Verbale Intelligenz

Da es signifikante Bildungsvorteile der Psychotherapeuten gibt, zeigen sich diese auch in Gruppenunterschieden Verbaler Intelligenz. Die Mittelwertsunterschiede der Gruppen ($M_{NOM} = 28,66$; $M_{KG} = 27,77$; $M_{PSY} = 30,11$; $SD_{NOM} = 1,89$; $SD_{KG} = 2,93$; $SD_{PSY} = 0,82$) werden bei signifikanten Levene-Test (31,90; $df_1 = 2$; $df_2 = 103$; $p > .01$) im Welch-Test hoch signifikant, $W = 17,08$ ($df_1 = 2$; $df_2 = 54,79$; $p > .01$). Die Psychotherapeuten erzielen signifikant höhere Intelligenzwerte als die sich nicht unterscheidenden gepaarten Gruppen. Die Korrelation zwischen gemessener Verbaler Intelligenz und Bildungsniveau für die 106 Probanden ohne fehlende Daten ist $\tau = .586^{**}$ ($p > .01$).

6.8.5 Psychotherapeutische Behandlung

Eine eigene Psychotherapeutische Behandlung haben 2 Nominierte und 5 Kontrollgruppenmitglieder durchgeführt. 14 Psychotherapeuten geben dies ebenfalls an, jedoch wird aus Anmerkungen heraus klar, dass dazu von einigen Probanden Lehranalysen zählten, die Teil einer psychotherapeutischen Ausbildung sind. Offensichtlich zählten andere Psychotherapeuten ihre Lehranalyse nicht zu einer psychotherapeutischen Behandlung. Die Daten erfassen also kein homogenes Konstrukt. Aus diesem Grunde ist die Verwendung des Items als mögliche Kontrollvariable wenig sinnvoll und wird daher außer Acht gelassen.

6.8.6 Familienstand

Die Prozentverteilung der Daten zu Familienstand und Elternschaft wird in Tabelle 6.9 für die Teilnehmergruppen sowie für Männer und Frauen separat präsentiert. Es ergeben sich keine signifikanten Unterschiede für die Gruppen oder Geschlechter. Dies bedeutet aber nicht, dass eine Gleichverteilung in den Abstufungen der Variablen Familienstand und Elternschaft vorliegt, da entsprechende Fishers exakte Tests jeweils hoch signifikante Verteilungsunterschiede belegten ($\chi^2 = 61,09$, $df = 4$, $p < .01$; $\chi^2 = 15,47$, $df = 1$, $p < .01$). Auch die An-

zahl von Kindern mit dem Range 0 bis 6 stellte sich erwartungsgemäß als nicht gleichverteilt heraus ($\chi^2 = 61,37$, $df = 5$, $p < .01$).

Table 6.9 Familienstand der Teilnehmer in Prozent.

	Gruppenzugehörigkeit			Geschlecht	
	NOM	KG	PSY	Männer	Frauen
Familienstand					
Ledig	21,0 %	10,5 %	18,4 %	12,7 %	21,6 %
Feste Partnerschaft	10,5 %	15,8 %	23,7 %	17,5 %	15,7 %
Verheiratet	52,6 %	50,0 %	42,1 %	57,1 %	37,3 %
Geschieden	7,9 %	15,8 %	13,2 %	9,5 %	15,7 %
Verwitwet	7,9 %	7,9 %	2,6 %	3,2 %	9,8 %
Elternschaft	63,2 %	79,0 %	63,2 %	71,4 %	64,7 %

Anmerkungen. Die Prozentzahlen basieren auf 38 Teilnehmern je Gruppe, 63 Männern und 51 Frauen.

7 Ergebnisse der Validierungsstudie

In diesem Kapitel erfolgt die Prüfung der Hypothesen, sowie der weiteren Analysen der Validierungsstudie. Im Folgenden kennzeichnen die Indices NOM für Nominierte, KG für Kontrollgruppe und PSY für die Psychotherapeuten die Gruppenzugehörigkeit der Messwerte. Die Ergebnisdarstellung wird im nächsten Kapitel Diskussion erörtert.

7.1 Kriteriumsvalidierung anhand der Inneren Validität

Hypothese 1: Der multiple Korrelationskoeffizient der einzelnen Weisheitstestverfahren erreicht einen Wert von .60.

Die Größenordnung der Interkorrelation der Weisheitstests wird zur Prüfung der Inneren Validität bestimmt. Tabelle 7.1 beinhaltet die über Regression bestimmten Werte des multiplen Korrelationskoeffizienten.

Tabelle 7.1. Multiple Korrelationskoeffizienten der Weisheitstestverfahren.

	<i>R</i>	<i>R</i> ²	<i>R</i> ² <i>korr</i>	<i>f</i> ²	<i>Z</i> (<i>r</i> = .60)	<i>F</i> (<i>df1,df2</i>)	<i>p</i>	<i>Sign. Tests</i>
3d-WS	.49	.24	.20	0,31	- 1,47	7,42 (4, 86)	<.01	SAWS, PWS
SAWS	.45	.20	.17	0,25	- 2,06 [†]	6,13 (4, 86)	<.01	3d-WS
PWS	.32	.10	.06	0,11	- 3,57 ^{††}	2,69 (4, 86)	.04	3d-WS
TWR	.51	.26	.22	0,34	- 1,29	8,21 (4, 86)	<.01	AWR
AWR _{GES}	.45	.21	.17	0,26	- 2,06 [†]	6,18 (4, 86)	<.01	TWR

Anmerkungen. *R* indiziert den Multiplen Korrelationskoeffizienten, *R*²*korr* dessen korrigierte Anpassung für die angenommene Grundgesamtheit. *f*² ist die Effektgröße. *Z* indiziert den Abstand von einem Korrelationsniveau von *r* = .60 in Fisher-Z-transformierten Werten. *df1* und *df2* indizieren die Freiheitsgrade des F-Tests, *p* dessen Wahrscheinlichkeit. In der letzten Spalte werden alle signifikanten Tests der multiplen Regression für den jeweiligen Test benannt. Listenweiser Fallausschluss: *n* = 101.

^{††} *Z*_{KRIT} = - 2,326, [†] *Z*_{KRIT} = - 1,645.

Alle multiplen Korrelationskoeffizienten liegen unter *R* = .60. Werden die Koeffizienten mit dem Niveau von .60 verglichen, so sind die Tests 3d-WS und TWR nicht signifikant darunter. Die Hypothese 1 zur Inneren Validität der Testverfahren wird daher für 3d-WS und TWR beibehalten. Um diese zu falsifizieren, muss *R* = .47 unterschritten werden, was auf die Tests SAWS (*R* = .45), PWS (*R* = .32) und AWR_{GES} (*R* = .45) zutrifft.

Der TWR hat die größte Kriteriumsvalidität der vorliegenden Tests. Die Varianzaufklärung durch die anderen Testverfahren liegt bei $R^2 = .26$. Der 3d-WS mit dem zweithöchsten multiplen Korrelationskoeffizienten ($R = .49$) wird als einziger Test durch den signifikanten Einfluss von zwei Tests bestimmt, SAWS und PWS. Angesichts dieses Befundes, dass vier von fünf Tests nur den signifikanten Einfluss eines Tests bei der Multiplen Korrelation aufweisen, kann nicht behauptet werden, dass alle Testverfahren übereinstimmende Merkmalsvarianz aufklären. Es sind verschiedene Konstrukte oder getrennte Facetten, die von den hier zu testenden Weisheitstests erfasst werden.

Effektstärken zwischen $f^2 = 0,15$ und $0,35$ haben mittlere Größe. Alle Tests außer der kleinen Effektgröße des PWS ($f^2 = .11$) erreichen dieses Niveau. Die Effektgröße des TWR liegt nur knapp unter dem Niveau eines großen Effektes. Die Durbin Watson Statistik zeigt leicht positive Autokorrelationen der Residuen für 3D-WS, SAWS und TWR, negative für PWS und AWR_{GES} . Die Differenzen zum Ideal von 2,00 sind gering. Die Kollinearitätsdiagnosen und Residuenanalysen offenbaren für alle Tests keine auffälligen Werte.

Wenige Unterschiede ergeben sich bei alternativer Analyse des AWR_{MD} anstatt des AWR_{GES} . Der wichtigste Unterschied ist jener des multiplen Korrelationskoeffizienten des SAWS ($R = .48$; $F = 5,78$ (4, 83); $p < .01$), der in diesem Niveau nicht signifikant von $R = .60$ unterscheidbar ist. Soll dieselbe Analyse mit dem AWR_{WORT} erfolgen, so muss bedacht werden, dass bei 28 Personen die Voraussetzung zur bivariaten Normalverteilung der Testwerte überprüft werden müsste, wozu leider geeignete Verfahren fehlen. Folgende Ergebnisse sind daher mit Vorsicht zu deuten. Bei multipler Korrelation mit dem AWR_{WORT} verlieren PWS, TWR und AWR ihre Signifikanz aufgrund der kleinen Stichprobe; während die multiple Korrelation des SAWS ($R = .59$; $F = 3,02$ (4, 28); $p = .04$) und 3d-WS ($R = .59$; $F = 3,04$ (4, 28); $p = .04$) dem in der Hypothese geforderten Niveau von $R = .60$ ($Z = -0,08$) nahe kommen.

Tabelle 7.2 Bivariate Korrelationen der Weisheitstestverfahren der Gesamtgruppe.

	3d-WS	SAWS	PWS	TWR	AWR_{GES}
3d-WS		>.01	>.01	.04	.10
SAWS	.42 ^{††}		.13	.01	.30
PWS	.28 ^{††}	.11		.03	.25
TWR	.17 [†]	.23 [†]	.19 [†]		>.01
AWR_{GES}	.13	.05	.07	.44 ^{††}	

Anmerkungen. Listenweiser Fallausschluss: $n = 101$.

^{††} $p > .01$, [†] $p > .05$ (einseitig).

Tabelle 7.2 zeigt die bivariaten Korrelationen zwischen sämtlichen Weisheitstestverfahren. Diese Beziehungen gehen zwar über die Hypothese hinaus, untermauern aber deren Aussage. Alle Korrelationen zwischen TWR und den anderen Tests werden signifikant. Der AWR_{GES} korreliert am höchsten mit dem TWR ($r = .44$; $p < .01$), während die anderen Korrelationen lediglich auf dem 5% Niveau signifikant sind. Die einzigen anderen signifikanten Korrelationen bestehen zwischen dem 3d-WS und SAWS sowie PWS. Besonders hoch ist dabei die Korrelation zu SAWS ($r = .42$; $p < .01$).

Hypothese 2: Die Testverfahren PWS und TWR sind nicht im mittleren oder hohen Maße miteinander positiv korreliert.

Hypothese 2 formuliert die Annahme der Unkorreliertheit von TWR und PWS als Replikation von Wink und Helson (1997). Tabelle 2 dokumentiert die bivariate Korrelation mit $r = .19$ ($p = .03$) als signifikant von Null unterscheidbar, aber nicht von $r = .30$ ($Z = -0,06$). Die Messung Praktischer Weisheit im Selbstbericht eines dichotomen Fragebogens ist somit signifikant mit dem Selbstbericht von Transzendentaler Weisheit korreliert. Auch von $r = .11$, dem Korrelationsniveau der Studie von Wink und Helson, ist die Korrelation dieser Studie nicht zu unterscheiden ($Z = 0,41$). Es kann somit weder ausgeschlossen werden, dass die Korrelation konvergenter Natur ist, noch dass $r = .11$ das wahre Korrelationsniveau beider Tests ist.

Hypothese 3: Die Testverfahren SAWS und 3d-WS korrelieren signifikant höher miteinander als mit den drei Weisheitstests PWS, TWR und AWR.

Hypothese 3 verkörpert einen aus der Kategorisierung der Weisheitstestverfahren abgeleiteten Zusammenhangsunterschied, siehe Kapitel 2.3. Die höchste bivariate Korrelation von 3d-WS oder SAWS zu einem der anderen Weisheitstests ist jene zwischen 3d-WS und PWS ($r = .28$, $p < .01$), siehe Tabelle 7.2. Diese ist signifikant niedriger als die Korrelation von 3d-WS und SAWS ($r = .42$, $p < .01$), weshalb diese Hypothese nach dem Verfahren von Williams (1959) weiterhin beibehalten wird ($T_W = 4,033 > T_{KRIT} = 1,661$). Die nach der Kategorisierung vorhergesagte Nähe der Testverfahren 3d-WS und SAWS hinsichtlich ihrer Übereinstimmung des Gegenstands, der Orientierung oder der Modalität gegenüber den anderen Tests TWR, AWR, und PWS, die sich mindestens in einer dieser Kategorien von 3d-WS und SAWS unterscheiden, hat somit einer empirischen Prüfung stand gehalten.

7.2 Kriteriumsvalidierung anhand der Äußeren Validität

Hypothese 4: Die Nominierten haben ein signifikant höheres Mittel auf den jeweiligen Weisheitstestverfahren als die Kontrollgruppe.

Die Gruppenmittelwerte werden durch T-Tests für abhängige Gruppen miteinander verglichen. Alle Mittelwertsunterschiede zwischen den Nominierten und der Kontrollgruppe widersprechen der Hypothese, siehe Tabelle 7.3. Die negativen Teststatistiken stehen sogar die Überlegenheit der Kontrollgruppe. Hierzu werden die von SPSS bestimmten zweiseitigen Wahrscheinlichkeitswerte nach deren Höhe und der Richtung des Mittelwertsunterschieds interpretiert. Die Effektstärken belegen bei AWR_{GES} ($M_{NOM} = 9,21$, $M_{KG} = 9,84$, $d' = ,22$) einen kleinen Effekt und verfehlen diesen bei TWR ($M_{NOM} = 3,57$, $M_{KG} = 3,77$, $d' = ,19$) knapp, wozu $d' = 0,20$ hätte erreicht werden müssen. In Konsequenz ist Hypothese 4 für jeden Weisheitstest, 3d-WS, SAWS, PWS, TWR und AWR_{GES} , zu verwerfen. Tabelle 6.3 dokumentiert die Mittelwerte der Weisheitstests für die Teilnehmergruppen.

Aufgrund der fehlenden signifikant positiven Korrelation der Messwertreihen beider Teilnehmergruppen, erfolgt die Absicherung dieser Befunde mittels W-Tests für unabhängige Gruppen. Es ergeben sich hier ebenfalls keine signifikanten Unterschiede. Auf eine Darstellung dieser Teststatistiken wird daher verzichtet.

Tabelle 7.3 Vergleich der Mittelwertsdifferenzen der Nominierten und Kontrollgruppe auf den einzelnen Weisheitstestverfahren.

	<i>n</i>	M_D (SE)	<i>SD</i>	$r_{NOM,KG}$	$t_{NOM,KG}$	<i>df</i>	<i>p</i> (2-seitig)	<i>d'</i>
3d-WS	38	,13 (.21)	1,30	-.12 n.s.	,61	37	.27 n.s	0,09
SAWS	38	-,59 (5,04)	30,65	.04 n.s.	-,12	37	.46 n.s	0,02
PWS	36	,06 (.53)	3,19	-.04 n.s.	,11	35	.46 n.s	0,02
TWR	30	-,20 (.20)	1,10	.05 n.s.	-1,00	29	.17 n.s	0,19
AWR_{GES}	38	-,63 (.54)	3,34	.28 n.s.	-1,17	37	.13 n.s	0,22

Anmerkungen. Negative Mittelwertsdifferenzen und T-Werte indizieren höhere Werte der Kontrollgruppe im Vergleich zu den Nominierten. *d'* ist die Effektgröße.

***p* < .01, **p* < .05 (zweiseitig).

Hypothese 5: Die Psychotherapeuten haben ein signifikant höheres Mittel auf den einzelnen Weisheitstestverfahren als die Kontrollgruppe.

Tabelle 7.4 Vergleich der Mittelwertsdifferenzen der Psychotherapeuten und Kontrollgruppe auf den einzelnen Weisheitstestverfahren.

	<i>n</i>	<i>M_D</i> (<i>SE</i>)	<i>SD</i>	<i>r_{PSY,KG}</i>	<i>t_{PSY,KG}</i>	<i>df</i>	<i>p</i> (2-seitig)	<i>d'</i>
3d-WS	38	,05 (.20)	1,25	.03 n.s.	,25	37	.80 n.s.	0,04
SAWS	37	-3,51 (4,68)	28,45	.22 n.s.	-,75	36	.46 n.s.	0,14
PWS	36	,31 (.47)	2,82	.12 n.s.	,65	35	.52 n.s.	0,01
TWR	30	,27 (.19)	1,05	-.05 n.s.	1,39	29	.17 n.s.	0,26
AWR _{GES}	38	,55 (.62)	3,80	.13 n.s.	,90	37	.38 n.s.	0,16

Anmerkungen. Negative Mittelwertsdifferenzen und T-Werte Werte indizieren höhere Werte der Kontrollgruppe im Vergleich zu den Nominierten. *d'* ist die Effektgröße.

***p* < .01, **p* < .05 (zweiseitig).

Hypothese 5 erforscht die Überlegenheit in den Mittelwertsdifferenzen der einzelnen Weisheitstests zwischen Psychotherapeuten und der Kontrollgruppe. Tabelle 7.4 präsentiert die statistischen Kennwerte für diese Hypothese. Analog zu Hypothese 4 kann die hypothetische Überlegenheit zwischen den Psychotherapeuten und der Kontrollgruppe nicht statistisch signifikant belegt werden. Lediglich ein kleiner Effekt ($M_{PSY} = 3,97$, $M_{KG} = 3,77$, $d' = ,26$) im TWR zu Gunsten der Psychotherapeuten liegt nahe, der die statistische Signifikanz verfehlt. Tabelle 6.3 präsentiert die Mittelwerte je Gruppe für die Weisheitstests. Hypothese 5 ist daher für jeden Test, 3d-WS, SAWS, PWS, TWR und AWR_{GES} falsifiziert. Auch hier liefert eine nachgeschobene Absicherung der Ergebnisse über Testverfahren für unabhängige Stichproben keine Unterschiede.

Hypothese 6: Die Nominierten haben ein signifikant höheres Mittel auf den jeweiligen Weisheitstestverfahren als die Psychotherapeuten.

Abgeleitet von den nicht signifikanten Effektstärken der beiden vorangegangenen Hypothesen, ist ein Beleg der Hypothese 6, der Dominanz der Weisheit der Nominierten über die klinischen Psychotherapeuten, nicht zu erwarten. Die statistische Prüfung bestätigt diesen Schluss, siehe Tabelle 7.5. In keinem Test dieser Studie konnten die Nominierten signifikant höhere Werte als die Psychotherapeuten erreichen. Hypothese 6 ist daher für alle Testverfahren

ren falsifiziert. Tabelle 7.5 präsentiert ebenso die aufschlussreichen Teststatistiken, die für zweiseitige Signifikanzprüfung durch das Statistikprogramm SPSS erstellt wurden. Es stellte sich entgegen der Hypothese heraus, dass Psychotherapeuten signifikant bessere Werte im TWR erreichen als die Nominierten ($t = -2,19$, $df = 30$, $p = .04$). Dieser Mittelwertsunterschied ist mit $M_{NOM} = 3,57$ und $M_{PSY} = 3,97$ ($d' = ,36$) als kleiner Effekt einzustufen. Auch die Mittelwerte des AWR_{GES} legen eine Dominanz der Psychotherapeuten mit kleinem Effekt ($M_{NOM} = 9,21$, $M_{PSY} = 10,39$, $d' = ,29$) nahe, der jedoch nicht statistische Signifikanz erreicht. Eine Übersicht der Mittelwerte je Gruppe liefert Tabelle 6.3.

Werden zur Absicherung dieser Hypothese ebenfalls W-Tests für unabhängige Stichproben durchgeführt, so stimmen die Ergebnisse mit den für abhängige Stichproben berichteten überein. Die Überlegenheit der Psychotherapeuten gegenüber den Nominierten im TWR wird als einzige statistisch signifikant ($W = 2,25$; $df = 66,47$; $p = ,028$). Im Falle des 3d-WS, SAWS und PWS kann nicht zwischen den beiden Teilnehmergruppen unterschieden werden. Somit ist Hypothese 6 für alle Tests 3d-WS, SAWS, PWS, TWR und AWR_{GES} falsifiziert.

Tabelle 7.5 Vergleich der Mittelwertsdifferenzen der Nominierten und Psychotherapeuten auf den einzelnen Weisheitstestverfahren.

	<i>n</i>	$M_D (S_E)$	<i>SD</i>	$r_{NOM,PSY}$	$t_{NOM,PSY}$	<i>df</i>	<i>p</i> (2-seitig)	<i>d'</i>
3d-WS	38	,18 (.22)	1,36	-.22 n.s.	,82	37	.42 n.s.	0,12
SAWS	37	-4,11 (4,38)	26,62	-.05 n.s.	,80	36	.35 n.s.	0,16
PWS	36	,44 (3,31)	3,31	.05 n.s.	,81	35	.43 n.s.	0,14
TWR	31	-.45 (.21)	1,15	.14 n.s.	-2,19	30	.04	0,36
AWR_{GES}	38	-1,18 (.61)	3,77	.14 n.s.	-1,94	37	.06 n.s.	0,29

Anmerkungen. Negative Mittelwertsdifferenzen und T-Werte Werte indizieren höhere Werte der Kontrollgruppe im Vergleich zu den Nominierten. *d'* ist die Effektgröße.

** $p < .01$, * $p < .05$ (zweiseitig).

In Summe muss über die Hypothesen der Äußeren Validierung gesagt werden, dass alle falsifiziert wurden, da es keinen einzigen hypothesenkonformen Gruppenunterschied auf nur einem einzelnen Weisheitstest in den vorliegenden Analysen gefunden werden konnte. Die Stichprobe kann demnach für die Tests 3d-WS, SAWS, PWS und AWR_{GES} anstatt durch drei abhängige Teilnehmergruppen, die sich hinsichtlich ihrer Weisheit unterscheiden, durch eine Teilnehmergruppe repräsentiert werden. Bei Verwendung der abhängigen oder unabhängigen Auswertungsdesigns stellten sich keine Unterschiede ein, weshalb bei späteren Studien auf die Anwendung abhängiger Auswertungsdesigns für diese Fragestellungen verzichtet wird.

7.3 Konstruktvalidierung anhand der Konvergenten Validität

Auch bei der Konstruktvalidierung über die Konvergente Validität zu verwandten Konstrukten erfolgen einzelne Vergleiche, anstatt die Testung einer summativen Hypothese. Die bivariaten Korrelationen werden in den Tabellen 7.6 und 7.7 präsentiert.

Hypothese 7: Ego-Entwicklung korreliert moderat positiv mit den einzelnen Weisheitstestverfahren.

Zunächst wird auf ausgewählte Aspekte der deskriptiven Verteilung der WUSCT-Werte eingegangen. Es finden sich generell hohe Werte der Ego-Entwicklung in den Teilnehmergruppen. Die sechste, Gewissenhafte Stufe, die von 85,84 % der Teilnehmer erreicht oder überschritten wird ($M_{NOM} = 6,54$; $M_{KG} = 6,26$; $M_{PSY} = 6,87$), liegt über den Durchschnitt der Allgemeinbevölkerung (Loevinger, 1976). Die beiden oberen Stufen 8, Autonomie, und 9, Integration, die Weisheit verkörpern sollen, erreichen 14,16 % ($NOM = 4$, $KG = 2$ und $PSY = 10$).

Wie die erste Datenspalte der Tabelle 7.6 präsentiert, werden die Korrelationen von Ego-Entwicklung zu allen Tests, außer dem PWS signifikant. Die Beziehungen von WUSCT zu 3d-WS, SAWS, TWR und AWR_{GES} können nicht von einer konvergenten Korrelation unterschieden werden. Einzig die Korrelation zwischen PWS und WUSCT ist signifikant unter $r = .30$ ($Z_{PWS} = 3,06$). Die höchste Korrelation erreicht der AWR_{GES} mit .25 und teilt somit 5% Varianz mit WUSCT. Hypothese 7 ist somit für 3d-WS, SAWS, TWR und AWR_{GES} gestützt.

Tabelle 7.6 Bivariate Korrelationen der Weisheitstestverfahren mit Ego-Entwicklung.

	WUSCT		WUSCT ohne Antwortlänge		Antwortlänge	
	r	$Z(r=.30)$	r	$Z(r=.30)$	r	$Z(r=.30)$
3d-WS	.25 [†]	-0,54	.24 [†]	-0,70	.16	-1,48
SAWS	.23 [†]	-0,75	.19 [†]	-1,22	.14	-1,67 [†]
PWS	.01	-3,06 ^{††}	-.04	-3,47 ^{††}	.16	-1,48
TWR	.34 ^{††}	0,04	.19 [†]	-1,13	.27 [†]	-0,33
AWR_{GES}	.42 ^{††}	1,37	.25 ^{††}	-0,55	.27 [†]	-0,33

Anmerkungen. Die erste Spalte enthält die Korrelation zwischen WUSCT und den Tests, die zweite enthält die um Antwortlänge bereinigten Semipartialkorrelationen des WUSCT, die dritte Spalte die Korrelation der Antwortlänge im WUSCT mit den einzelnen Testverfahren. Z indiziert den Abstand von einem Korrelationsniveau von $r = .30$ in Fisher-Z-transformierten Werten. Listenweiser Fallausschluss: $n = 101$.

^{††} $p > .01$, [†] $p > .05$ (einseitig); ^{††} $Z_{KRIT} = -2,326$, [†] $Z_{KRIT} = -1,645$.

Als mögliche Einflussvariable zwischen der Korrelation des WUSCT und den Weisheitstests ist die Antwortlänge, also die Summe der verwendeten Worte, zu berücksichtigen. Zur Aufklärung ihres Einflusses stellt Tabelle 7.6 die Korrelationen der Weisheitstests zum WUSCT mit und ohne den Einfluss der Antwortlänge, sowie der Antwortlänge ansich vor.

Zunächst werden die Kernpunkte der ersten beiden Datenspalten vorgestellt. Die Semipartialkorrelationen des um Antwortlänge bereinigten WUSCT und der Tests 3d-WS, SAWS, TWR und AWR_{GES} unterscheiden sich signifikant von Null. Der Fischer Z-Vergleich der Korrelationen WUSCT mit und WUSCT ohne Antwortlänge der Weisheitstests wirf für 3d-WS ($Z = 0,18$), SAWS ($Z = 0,45$) und TWR ($Z = 1,56$) nicht signifikant, während der Unterschied bei AWR_{GES} ($Z = 1,91$) signifikant ist. Zusammenfassend, wird die Antwortlänge aus dem WUSCT herauspartialisiert, so sinkt einzig die Korrelation zu AWR signifikant.

Werden die in der dritten Spalte dargestellten Korrelationen der Antwortlänge des WUSCT und den Weisheitstestverfahren betrachtet, so zeigen sich zwei Gruppierungen. Während keine signifikanten Korrelationen zu den Testverfahren mit geschlossenen Antwortformaten vorliegen, korrelieren die Testwerte von TWR und AWR_{GES}, beides Tests mit offenem Antwortformat, signifikant mit der Antwortlänge im WUSCT. Diese beiden Korrelationen haben gleiche Höhe ($r = .27, p = .02$). Antwortlänge im WUSCT ist nur für TWR und AWR_{GES} bedeutsam. Offenbar besitzen diese drei Tests eine weitere Varianzquelle, für die Antwortlänge ein Indikator ist. Im Diskussionsteil findet hierzu eine Erörterung statt.

Hypothese 8: Psychologisches Feingefühl korreliert moderat positiv mit den einzelnen Weisheitstestverfahren.

Tabelle 7.7 Bivariate Korrelationen der Weisheitstestverfahren mit PF und PW.

	PF		PW	
	<i>r</i>	<i>Z</i> (<i>r</i> = .30)	<i>r</i>	<i>Z</i> (<i>r</i> = .30)
3d-WS	.48 ^{††}	2,08 [†]	.45 ^{††}	1,77 [†]
SAWS	.39 ^{††}	,013	.63 ^{††}	4,19 ^{††}
PWS	.28 ^{††}	0,20	.05	- 2,60 ^{††}
TWR	.02	- 2,91 ^{††}	.14	- 1,68 [†]
AWR _{GES}	-.03	- 3,38 ^{††}	.05	- 2,55 ^{††}

Anmerkungen. *Z* indiziert den Abstand von einem Korrelationsniveau von $r = .30$ in Fisher-Z-transformierten Werten. Listenweiser Fallausschluss: $n = 101$

^{††} $p > .01$, [†] $p > .05$ (einseitig); ^{††} $Z_{KRIT} = - 2,326$, [†] $Z_{KRIT} = - 1,645$; (einseitig).

Hypothese 8 ist für das Konstrukt Psychologisches Feingefühl, gemessen über den IPIP, erfüllt für drei Testverfahren, den 3d-WS, SAWS und PWS. Insbesondere der 3d-WS korreliert mit PF signifikant über dem Niveau von .30 ($r = .48$; $p < .01$, $Z_{3d-WS} = 2,08$). Für TWR und AWR zeigen sich klare diskriminante Beziehungen aufgrund der insignifikanten Korrelationen ($r_{TWR} = .02$; $p = .438$; $r_{AWR} = -.03$; $p = .381$), siehe Tabelle 7.7.

Hypothese 9: Persönliches Wachstum korreliert moderat positiv mit den einzelnen Weisheitstestverfahren.

Die einzelnen Hypothesen zur Hypothese 9, der konvergenten Beziehung zwischen Persönlichem Wachstum (PW) und den Weisheitstests, sind für die Testverfahren 3d-WS ($r = .45$, $Z = 1,77$, $p < .01$) und SAWS ($r = .63$; $Z = 4,19$; $p < .01$) signifikant. Die Korrelation des SAWS liegt signifikant über dem Niveau von $r = .30$, siehe Tabelle 7.7. Für die Testverfahren PWS, TWR und AWR_{GES} wird Hypothese 8 falsifiziert.

Ergänzend werden die bivariaten Korrelationen der Persönlichkeitsreife-tests untereinander angegeben. Die beiden Tests PF und PW korrelieren in moderater Höhe ($r = .51$; $p < .01$). Die Korrelation zwischen WUSCT und PW ist signifikant ($r = .19$; $p = .03$), während die Beziehung zu PF insignifikant ist ($r = .13$; $p = .11$). Die Semipartialkorrelationen des um Antwortlänge bereinigten WUSCT zu den theoretisch verwandten Tests PW ($r = .11$; $p = .14$) und PF ($r = .11$; $p = .13$) sind sehr gering. Da das kritische Niveau der Semipartialkorrelation bei $r = 0,17$ ($n = 101$, $df = 98$, einseitiges Alpha von 0,05 %) liegt, sind beide Korrelationen somit insignifikant von Null und daher diskriminanter Natur (jeweils $Z = -1,98$).

Die zur Validierung verwendeten Persönlichkeitsreife-tests zeigen jedoch ebenfalls kein übereinstimmend konvergentes Muster. PW und PF korrelieren signifikant und in eindeutig moderater Größenordnung zueinander, die Korrelationen von WUSCT zu den beiden anderen Konstrukten sind jedoch sehr klein und werden nur zwischen WUSCT und PW signifikant. Die gewünschte konvergente Beziehung zwischen allen drei Verfahren fehlt somit. Daher ist belegt, dass WUSCT ein anderes Konstrukt erfasst als von den Tests PW und PF gemeinsam abgedeckt wird.

Abschließend wird für die Testverfahren 3d-WS und SAWS jede der drei Hypothesen bestätigt. Für den PWS trifft nur Hypothese 7 zur Konvergenten Validierung des Psychologischen Feingefühls zu. Die hypothetischen Beziehungen des PWS nach Hypothese 6 und 8 müssen als falsifiziert eingestuft werden. Bei den beiden Tests TWR und AWR_{GES} kann nur Hypothese 6 zur Konvergenten Beziehung zu WUSCT beibehalten werden. Die Hypothesen 7

und 8 werden für diese Tests falsifiziert. Bei Analyse von AWR_{MD} und AWR_{WORT} zeigen sich die gleichen Ergebnisse wie für AWR_{GES} .

7.4 Konstruktvalidierung anhand der Diskriminanten Validierung

Dieser Abschnitt prüft, ob die Konstrukte Adaptivität, indiziert durch UK, Verbale Intelligenz, Alter und Geschlecht keine Korrelationen zu den einzelnen Weisheitstestverfahren haben, die die Größenordnung von mehr als 10% gemeinsame Varianz nicht überschreiten. Auch die Hypothesen 10 bis 13 werden einseitig für jeden einzelnen Test geprüft.

Hypothese 10: Die Korrelation von Adaptivität und den einzelnen Weisheitstestverfahren liegt im diskriminanten Bereich.

Die Korrelation der Weisheitstests zu Umweltkontrolle (UK) als Maß für Adaptivität schildert Tabelle 7.8. Einzig für TWR wird die Hypothese der diskriminanten Beziehung zu UK beibehalten. Die Korrelationen des SAWS, PWS und AWR_{GES} zum UK können nicht klar in den diskriminanten Bereich eingeordnet werden, da sich ihr Niveau nicht von $r = .30$ unterscheidet. Also gilt die Hypothese für diese Tests falsifiziert. Die Korrelation zwischen UK und 3d-WS ist nicht divergent, sondern konvergent ($r = .60$; $p < .01$; $Z = 3,62$) und sogar signifikant höher als die übrigen Korrelationen. Im dazugehörigen Vergleich ist die Korrelation zu 3d-WS ($Z = 0,68$) signifikant höher als jene zu PWS ($Z = 0,30$), was eine nahezu mittlere Effektstärke von Fisher-Z-Transformierten-Korrelationskoeffizienten zeigt ($q = 0,39$).

Tabelle 7.8 Bivariate Korrelationen der Diskriminanten Validierung von UK und HAWIE-R.

	UK		HAWIE-R	
	r	$Z (r = .30)$	r	$Z (r = .30)$
3d-WS	.60 ^{††}	3,62 ^{††}	.14	- 1,67 [†]
SAWS	.24 ^{††}	- 0,61	.13	- 1,78 [†]
PWS	.29 ^{††}	- 0,10	-.01	- 3,06 ^{††}
TWR	.049	- 2,49 ^{††}	-.09	- 3,83 ^{††}
AWR_{GES}	-.074	- 3,66 ^{††}	-.03	- 3,21 ^{††}

Anmerkungen. Z indiziert den Abstand von einem Korrelationsniveau von $r = .30$ in Fisher-Z-transformierten Werten. Listenweiser Fallausschluss: $n = 94$.

^{††} $p > .01$, [†] $p > .05$ (einseitig); ^{††} $Z_{KRIT} = 2,326$, [†] $Z_{KRIT} = 1,645$ (einseitig).

Die Korrelation zwischen PW und UK, den Untertests des PWB von Ryff (1989; Ryff & Keyes, 1995), liegt im erwarteten niedrigen Bereich ($r = .32$, $p < .01$). Daher kann nicht von einem bedeutsamen gemeinsamen Anteil beider Tests ausgegangen werden, der die Korrelation von UK und Weisheit erklären könnte. Zur Prüfung dieser Annahme wird aus den UK die mit PW geteilte Testvarianz auspartialisiert. Die Semipartialkorrelationen werden noch immer signifikant für den 3d-WS ($r = .48$; $p < .01$) und PWS ($r = .37$; $p < .01$). Trotz des Absinkens der Korrelation zwischen PW und 3d-WS sind die Korrelationen weiterhin konvergent.

Der Zusammenhang zwischen 3d-WS und UK ist im Hinblick auf die Argumentation von Staudinger, Dörner und Mickler (2005) interessant, da sie den 3d-WS nicht als Weisheitstest klassifizieren. Die bivariate Beziehung der beiden Tests liegt bei $r = .60$ ($p < .01$) für die Gesamtgruppe. Werden die drei Teilnehmergruppen einzeln hinsichtlich dieser Korrelation verglichen, so unterscheiden sich ihre Korrelationskoeffizienten in Fisher-Z-Transformierten Werten jedoch alle signifikant. Es besteht für die Weisheitsnominierten ein sehr hoher Zusammenhang ($r = .75$; $p < .01$), während für die Kontrollgruppe ($r = .47$; $p < .01$) und die Psychotherapeuten ($r = .59$; $p < .01$) eine signifikante, moderate Beziehung vorliegt. Diese Unterschiede spiegeln sich in mittleren und großen Effektstärken ($Z_{NOM} = 0,97$; $Z_{KG} = 0,51$; $Z_{PSY} = 0,68$), einzig der Effekt zwischen Kontrollgruppe und Psychotherapeuten ist nur klein.

Hypothese 11. Die Korrelation von Verbaler Intelligenz und den einzelnen Weisheitstestverfahren liegt im diskriminanten Bereich.

Hypothese 11 prüft den diskriminanten Zusammenhang zwischen Verbaler Intelligenz, gemessen über den Wortschatztest des HAWIE-R, und den Weisheitstests. Tabelle 7.8 zeichnet die höchste Korrelation zwischen 3d-WS und Verbaler Intelligenz ($r = .14$, $p = .10$, $Z = -1.67$) als diskriminant aus. Der HAWIE-R korreliert diskriminant mit den Weisheitstestverfahren, da die Anteile gemeinsamer Varianz 10% nicht übersteigen. Diese Ergebnisse legen nahe, dass Psychometrische Intelligenz nicht mit Weisheit korreliert, da der Wortschatztest des HAWIE-R aufgrund seiner hohen Korrelation zum Gesamtwert des HAWIE-R auch als Indikator für Allgemeine Intelligenz gesehen werden kann.

Hypothese 12. Die Korrelation zwischen dem Alter der Teilnehmer und den einzelnen Weisheitstestverfahren liegt im diskriminanten Bereich.

Da sich weder Altersmittelwerte noch Testmittelwerte der einzelnen Teilnehmergruppen un-

terscheiden, ist Hypothese 12 aus den bisherigen Ergebnissen wahrscheinlich. In Tabelle 7.9 zeigen sich signifikante Korrelationen der Testverfahren zu Alter, die sich nicht signifikant von $r = .30$ unterscheiden lassen für 3d-WS, SAWS und PWS. Um dieses Niveau zu unterschreiten, müsste die jeweilige Korrelation unter $r = .15$ liegen. Die höchste bivariate Korrelation zwischen Alter und dem SAWS ($r = .21$; $p = .01$) steht für 4,5% gemeinsame Varianz.

Die Korrelationen zwischen Alter und 3d-WS ist diskriminant, ($Z = -1,96$) siehe Tabelle 7.9. Die drei Teilnehmer über 80 Jahren werden in einer alternativen Analyse ausgeschlossen, da diese Altersgruppe negative Beziehungen zwischen Weisheit und Alter in bisherigen Studien zeigte, siehe Kapitel 4. Dass unter dieser Auswahl die Korrelationen sinken, widerspricht den Erwartungen an die Beziehung von Alter und Weisheit. Jedoch erlauben die 3 Personen der vorliegenden Studie keine Verallgemeinerungen auf die Gesamtbevölkerung oder Vergleichbarkeit mit den Probanden der Berliner Studien. Die Korrelation des PWS ($r = .16$) und des SAWS ($r = .17$; $p < .05$, $Z = -1,43$) bleiben weiterhin uneindeutig in der Zuordnung zu konvergenten oder diskriminanten Beziehungen. Die Hypothese des diskriminanten Zusammenhangs von Alter und den Weisheitstests ist daher für 3d-WS, SAWS, und PWS nicht haltbar. Die Hypothese bleibt jedoch für TWR und die AWR-Varianten erhalten.

Tabelle 7.9 Korrelation zwischen Alter und den Weisheitstestverfahren.

	Alter _{GESAMT}			Alter < 80		
	r	$Z(r = .30)$	n	r	$Z(r = .30)$	n
3d-WS	.20 [†]	-1,00	94	.12	-1,96 [†]	98
SAWS	.21 [†]	-0,91	94	.17 [†]	-1,34	98
PWS	.15 [†]	-1,52	94	.16	-1,88 [†]	98
TWR	.02	-1,34	94	.01	-2,91 ^{††}	98
AWR _{GESAMT}	-.05	-3,43 ^{††}	94	-.13	-4,26 ^{††}	98
AWR _{MD}	-.03	-3,15 ^{††}	88	-.14	-4,13 ^{††}	89
AWR _{WORT}	.27	-0,02	28	.08	-1,16	28

Anmerkungen. Z indiziert den Abstand von einem Korrelationsniveau von $r = .30$ in Fisher- Z -transformierten Werten.

^{††} $p > .01$, [†] $p > .05$ (einseitig); ^{††} $Z_{KRIT} = 2,326$, [†] $Z_{KRIT} = 1,645$ (einseitig).

Hypothese 13. Die Korrelation zwischen dem Geschlecht der Teilnehmer und den einzelnen Weisheitstests liegt im diskriminanten Bereich.

Bei der Hypothese 13 zu Geschlechtsunterschieden der Weisheitstests, Tabelle 7.10, zeigen sich nicht die erwarteten Ergebnisse. Während 3d-WS und SAWS keine Geschlechtsunterschiede aufweisen, zeigt sich PWS ($r_{pb} = .25, p < .01$) und TWR ($r_{pb} = .27, p < .01$) signifikant korreliert. Die weiblichen Teilnehmer schneiden in diesen Tests besser ab als die männlichen. Die Größenordnung dieser Korrelationen von TWR und PWS können nicht eindeutig dem konvergenten oder diskriminanten Bereich zugeordnet werden ($Z_{PWS} = -0,55$; $Z_{TWR} = -0,29$), jedoch sind über die kleinen Werte der Effektgröße η^2 erkennbar, dass kein wesentlicher Varianzanteil durch die Geschlechtsunterschiede aufgeklärt wird. Der Anteil an geteilter Varianz erreicht für TWR als höchsten Wert nur den Bereich von 8%.

Tabelle 7.10 Punkt-Biseriale Korrelation der Weisheitstests zu Geschlecht.

	3d-WS	SAWS	PWS	TWR	AWR
r_{pb}	.04	.08	.25**	.27**	.11
η^2	.00	.01	.06	.08	.01
$Z (r = .30)$	-2,84	-2,41	-0,55	-,29	-2,13

Anmerkungen. Positive Korrelationen stehen für Vorteile der Frauen. Z indiziert den Abstand von einem Korrelationsniveau von $r = .30$ in Fisher- Z -transformierten Werten. Listenweiser Fallausschluss $n = 101$.

** $p < .01$, * $p < .05$ (zweiseitig); $^{\dagger\dagger} Z_{KRIT} = -2,326$, $^{\dagger} Z_{KRIT} = -1,645$ (einseitig).

Zusammenfassend sind viele der einzelnen Unterhypothesen der Diskriminanten Konstruktvalidierung erfüllt worden. Insbesondere Intelligenz zeigt keine bedeutsame Korrelation zu Weisheit. Die Skala Umweltkontrolle als Indikator der Adaptivität zeigt sich dagegen mit Weisheit in bivariater Testung mehr als lediglich diskriminant korreliert. Leichte Altersunterschiede ließen sich für die Tests 3d-WS, SAWS und PWS belegen, die jedoch nicht mehr als 5% der Testvarianz erklären. Auch die Analyse der Geschlechtsunterschiede ist meist hypothesenkonform. Einzig für die Tests TWR und PWS zeigen sich signifikante Vorteile der Frauen. Auch hier liegt der maximal geteilte Varianzanteil mit 8% auf niedrigem Niveau.

7.5 Weitere Analysen

Als Exkurs, wird zunächst eine alternative Analyseverfahren des SAWS aus dessen Erstversion (Webster, 2003, 2006) auf die vorliegende Stichprobe übertragen. Das damalige Cut-Off-Verfahren zog die Grenze bei $\frac{5}{6}$ der maximalen Punktzahl. Übertragen auf die neue Itemzahl überschreiten lediglich 1 Nominierter, 2 Kontrollgruppenmitglieder und 3 Psychotherapeuten

den Grenzwert. Diese Verteilungsunterschiede sind insignifikant (*Fishers exakter Test* = 6,88, $p = .38$). Der Anteil von 5,3% Weiser Personen in der Gesamtgruppe ist gering im Vergleich zum Wert der Websterstudie. Der Gesamtmittelwertes von knapp $\frac{3}{4}$ der Maximalpunktzahl liegt relativ hoch aber eben nicht im Niveau von Weisheit, siehe Anhang F.2.

Um die Zusammenhänge der vorliegenden Teilstichproben näher verstehen zu können, als es die Kernhypothesen erlauben, werden weitere Analysen präsentiert. Die Analysen werden zu Gunsten der Verständlichkeit in angemessener Detailtiefe besprochen.

WA 1: Für jede Teilnehmergruppe werden die Interkorrelationen der Weisheitstests auf bivariater Ebene miteinander verglichen.

Werden alle fünf Weisheitstests nach gruppenspezifischen Korrelationsunterschieden analysiert, zeigen sich solche zwischen den Teilnehmergruppen, siehe Tabellen 7.11 sowie 7.12. Die signifikanten Korrelationen in einer Gruppe werden mit denen der anderen Teilnehmergruppen verglichen.

Tabelle 7.11 Interkorrelationen der Weisheitstests der Nominierten und der Kontrollgruppe.

	3DWS	SAWS	PWS	TWR	AWR _{SUM}
3DWS		.63 ^{††}	.35 [†]	.11	.17
SAWS	.16		.23	.14	.21
PWS	.10	-.01		.15	.16
TWR	-.10	.23	.01		.53 ^{††}
AWR _{SUM}	.42 ^{††}	.04	.04	.22	

Anmerkungen. Das obere Dreieck beinhaltet die Korrelationen der 34 Nominierten, das untere Dreieck die Korrelationen der 33 Kontrollgruppenmitglieder nach listenweisem Fallausschluss.

^{††} $p < .01$, [†] $p < .05$ (einseitig).

Tabelle 7.12 Interkorrelationen der Weisheitstests der Psychotherapeuten.

	3DWS	SAWS	PWS	TWR	AWR _{SUM}
3DWS		.56 ^{††}	.45 [†]	.31 [†]	.22
SAWS			.24	.20	-.065
PWS				.36 [†]	.11
TWR					.51 ^{††}
AWR _{GES}					

Anmerkungen. Das obere Dreieck beinhaltet die 32 Psychotherapeuten nach listenweisem Fallausschluss.

^{††} $p < .01$, [†] $p < .05$ (einseitig).

Die Korrelation zwischen 3d-WS und SAWS für Nominierte ($r = .63$; $p < .01$; $Z = 0,74$) liegt deutlich über der der Kontrollpersonen ($r = .16$; $p = .19$; $Z = 0,16$), während die Korrelation der Psychotherapeuten ($r = .56$; $p < .01$; $Z = 0,63$) eine mittlere Größenordnung annimmt. Die Unterschiede zwischen Nominierten und Kontrollpersonen werden in Fisher-Z-Transformationen signifikant und zeigen einem großen Effekt $q = 0,58$; der Unterschied zwischen Kontrollgruppe und Psychotherapeuten hat einen mittleren Effekt $q = 0,47$.

Der PWS korreliert signifikant mit dem 3d-WS für Psychotherapeuten zu $.45$ und für Nominierte zu $.35$ (jeweils $p < .05$), während für die gematchte Kontrollgruppe keine signifikante Beziehung zwischen den Tests vorliegt. Diese Diskrepanzen werden signifikant und belegen einen mittleren Effekt für die Psychotherapeuten $q = 0,39$ und einen kleinen Effekt für die Nominierten ($q = 0,27$).

Ähnlich, wenn auch höher, zeigt sich das Korrelationsmuster zwischen TWR und AWR_{GES} . Auch hier korrelieren für die Nominierten und die Psychotherapeuten beide Tests signifikant in mittlerer Höhe ($r = .53$ und $.51$), während in der Kontrollgruppe keine signifikante Korrelation vorliegt. Die Unterschiede zur Kontrollgruppe liegen an der Grenze zu einem mittleren Effekt ($q_{NOM} = .36$; $q_{PSY} = .34$). Einzig innerhalb der Kontrollgruppe korreliert der AWR_{GES} mit dem 3d-WS in moderater Höhe. Nur diese Korrelation ist in der Kontrollgruppe signifikant. In der Gruppe der Psychotherapeuten korreliert der TWR mit dem 3d-WS ($r = .31$, $p = .04$) und dem PWS ($r = .36$, $p = .02$) signifikant. Diese Korrelationen sind nur für die Psychotherapeuten signifikant.

Zur genaueren Analyse werden nun alle verwendeten Persönlichkeitstests aus der Beziehung der Weisheitstests herausgerechnet. Werden die Einflüsse über Partialkorrelationen auspartialisiert, korreliert einzig in der Nominierengruppe 3d-WS mit SAWS signifikant ($r = .46$, $p < .01$, $n = 25$).

Bei den Nominierten und den Psychotherapeuten bleiben auch nach Auspartialisierung der PW, der UK und des Wortschatztests die Korrelationen zwischen AWR_{GES} und TWR erhalten. Für die Gruppe der Nominierten zeigt sich die höchste signifikante Korrelation ($r = .56$; $p < .01$; $n = 25$). Für die Psychotherapeuten ergibt sich $r = .51$ ($p < .01$; $n = 24$). Außer dieser Korrelation ist TWR mit PWS ($r = .50$; $p < .01$; $n = 24$) signifikant positiv korreliert, während sich für die Korrelationsresiduen aus SAWS und AWR_{GES} eine signifikant negative Beziehung zeigt ($r = .38$; $p = .03$; $n = 24$). Die Korrelationen in der Kontrollgruppe umfassen die Bestätigung der Beziehung zwischen 3d-WS und AWR_{GES} ($r = .33$; $p < .05$; $n = 25$), sowie die nun signifikante Beziehung zwischen TWR und SAWS ($r = .35$; $p = .036$; $n = 25$). Die übrigen Weisheitstests korrelieren weiterhin nicht miteinander.

Tabelle 7.13 Multiple Korrelationskoeffizienten der Weisheitstestverfahren.

	<i>R</i>	<i>R</i> ²	<i>R</i> ² <i>korr</i>	<i>f</i> ²	<i>Z</i> (<i>r</i> = .60)	<i>F</i> (<i>df</i> ₁ , <i>df</i> ₂)	<i>p</i>	<i>Sign. Tests</i>
3d-WS _{NOM}	.67	.44	.37	0,80	0,56	5,76 (4, 29)	< .01	SAWS
3d-WS _{KG}	.14	.02	-.12	,02	- 2,76 ^{††}	,14 (4, 29)	.96	-
3d-WS _{PSY}	.70	.49	.42	,96	0,88	6,70 (4, 28)	< .01	SAWS
SAWS _{NOM}	.64	.41	.33	,69	0,32	4,99 (4, 29)	< .01	3d-WS
SAWS _{KG}	.31	.10	-.03	,11	- 1,86 [†]	,77 (4, 29)	.55	-
SAWS _{PSY}	.63	.40	.32	,67	0,27	4,69 (4, 28)	< .01	3d-WS
TWR _{NOM}	.54	.29	.19	1,16	- 0,48	2,92 (4, 29)	.04	AWR
TWR _{KG}	.41	.17	.17	,69	- 1,32	1,99 (4, 29)	.24	-
TWR _{PSY}	.61	.38	.29	1,59	0,11	4,24 (4, 28)	< .01	AWR
AWR _{NOM}	.55	.30	.21	1,22	- 0,38	3,14 (4, 29)	.03	TWR
AWR _{KG}	.26	.07	-.06	,35	- 2,14 [†]	,53 (4, 29)	.72	-
AWR _{PSY}	.56	.32	.22	1,28	- 0,29	3,24 (4, 28)	.03	TWR
PWS _{NOM}	.37	.14	.02	,16	- 1,55	1,16 (4, 29)	.35	-
PWS _{KG}	.26	.07	-.06	,07	- 2,12 [†]	,54 (4, 29)	.71	-
PWS _{PSY}	.45	.20	.09	,25	-1,07	1,77 (4, 28)	.16	-

Anmerkungen. *R* indiziert den Multiplen Korrelationskoeffizienten, *R*²*korr* dessen korrigierte Anpassung für die angenommene Grundgesamtheit. *f*² ist die Effektgröße. *Z* indiziert den Abstand von einem Korrelationsniveau von *r* = .60 in Fisher-*Z*-transformierten Werten. *df*₁ und *df*₂ indizieren die Freiheitsgrade des F-Tests, *p* dessen Wahrscheinlichkeit. In der letzten Spalte werden alle signifikanten Tests der multiplen Regression für den jeweiligen Test benannt. Den Daten liegen jeweils 34 Nominierte, 34 Kontrollpersonen und 33 Psychotherapeuten zugrunde.

^{††} *Z*_{KRIT} = - 2,326, [†] *Z*_{KRIT} = - 1,645.

Werden die multiplen Korrelationskoeffizienten für jede Teilnehmergruppe getrennt berechnet, so unterschreitet die Teilnehmerzahl 40 Personen, da es Daten zu 34 Nominierten, 34 Kontrollpersonen und 33 Psychotherapeuten zu allen Weisheitstests gibt, siehe Tabelle 7.13. Diese Unterschreitung erfordert nach Bortz (1999) die Prüfung der bivariaten Normalverteilung aller Testwerte gegeneinander. Da es für diese Analyse keine Verfahren gibt, wird auf diese Absicherung verzichtet. Ohne diese Prämisse testen zu können, sind die Ergebnisse weniger vertrauenswürdig als für die Gesamtgruppe.

Die Analysen zeigen für den PWS keinen signifikanten multiplen Korrelationskoeffizienten. Für die andern vier Tests erscheinen die Ergebnisse sehr übereinstimmend. Bei keinem Test wird der multiple Korrelationskoeffizient in der Kontrollgruppe signifikant. Dagegen werden in allen übrigen Tests die Koeffizienten für Nominierte und Psychotherapeuten signifikant und zeigen eine vergleichbare Größenordnung. Die Höhe dieser Korrelationskoeffizienten

fizienten erreicht das in Hypothese 1 geforderte Niveau, bzw. sie sind nicht von diesem zu unterscheiden. Der größte Unterschied liegt zwischen dem Wert $R = .61$ ($p < .01$) der *Psychotherapeuten* und $R = .54$ ($p = .04$) der Nominierten für den TWR. Dieser Befund spricht für die differentielle Validität der Testverfahren. Für die Psychotherapeuten und die Nominierten nehmen die Validitätskoeffizienten die erwarteten Werte an, während für die Kontrollgruppe diese Testverfahren nicht in diesem Maße kovariieren.

WA 2. Bivariate Korrelationen der Testverfahren zur konvergenten und diskriminanten Validierung werden für die drei Teilnehmergruppen berichtet.

Zu dieser Partialkorrelationskoeffizientenanalyse passt die Fragestellung, ob es Gruppenunterschiede auch auf den Testverfahren zur konvergenten und diskriminanten Validierung gibt. Grundsätzlich müssen die Testverfahren der konvergenten (PW, PF, WUSCT) und diskriminanten (UK, HAWIE) Validierung ihre theoretische Trennung ebenfalls auf empirischer Ebene belegen. Die Tabellen 7.14 und 7.15 präsentieren die einzelnen bivariaten Korrelationen für die Gesamtgruppe als auch für die einzelnen Teilnehmergruppen getrennt. Werden nun die Nominierten mit der Kontrollgruppe verglichen, siehe Tabelle 7.14, zeigen diese eine höhere Korrelation zwischen PF zu PW sowie zu UK als deren gematchte Kontrollpersonen. Die höchste Korrelation der Gesamtgruppe erreichen UK mit PF, was aufgrund der zugrunde liegenden theoretischen Konzepte nicht erwartet wurde. Diese beiden Skalen UK und PF korrelieren in jeder Gruppe höher miteinander als alle übrigen bivariaten Paarungen. Hier offenbart sich gemeinsame Konstruktvarianz, die noch zu erklären ist.

Tabelle 7.14 Interkorrelationen der konvergenten und diskriminanten Testverfahren der Validierungsstudie für die Nominierten und die Kontrollgruppe.

	WUSCT	PF	PW	UK	HAWIE R
WUSCT		.25	-.03	.14	-.23
PF	-.01		.55 ^{††}	.63 ^{††}	-.20
PW	-.08	.39 ^{††}		.33 [†]	-.09
UK	-.15	.46 ^{††}	.31 [†]		.01
HAWIE-R	.25	.05	.08	.02	

Anmerkungen. Das obere Dreieck beinhaltet die Korrelationen der 35 Nominierten; das untere Dreieck die Korrelationen der 35 Kontrollgruppemitglieder nach listenweisem Fallausschluss.

^{††} $p < .01$, [†] $p < .05$ (einseitig).

Tabelle 7.15 Interkorrelationen der konvergenten und diskriminanten Testverfahren der Validierungsstudie für die Psychotherapeuten und für die Gesamtgruppe.

	WUSCT	PF	PW	UK	HAWIE R
WUSCT		.23	.24	.10	.30 ^{††}
PF	.15		.42 ^{††}	.55 ^{††}	.48 ^{††}
PW	.05	.44 ^{††}		.39 ^{††}	.19
UK	.01	.52 ^{††}	.32 ^{††}		.33 [†]
HAWIE-R	.19 [†]	.04	.09	.02	

Anmerkungen. Das obere Dreieck beinhaltet die Korrelation der 36 Psychotherapeuten; das untere Dreieck zeigt die Korrelationen der 106 Personen der Gesamtgruppe nach listenweisem Fallausschluss.

^{††} $p < .01$, [†] $p < .05$ (einseitig).

WA 3: Für die Erfolgstypen und die Suchenden werden die Mittelwerte der Weisheitstestverfahren verglichen.

Es folgt die Erörterung der Studienergebnisse von Helson und Srivastava (2001) anhand der empirischen Daten dieser Studie. Die Skalen UK und PW werden zur Bildung der ebendort beobachteten Typen herangezogen. Beide sechsstufigen Skalen sind rechtsschief, was Tabelle 7.16 anhand der deskriptiven Statistiken darstellt. Beide Verteilungen zeigen hohe Werte zentraler Tendenz und ein hohes Minimum, welches etwa einen halben Skalenpunkt unter der rechnerischen Mitte der Skalierung liegt. Als Konsequenz müssen der Mehrheit der Teilnehmer hohe Werte in Umweltkontrolle und Persönlichkeitsreife zugesprochen werden.

Nach Mediansplittung befinden sich 38 Erfolgstypen, 20 Personen des Typus Suchende in der Gesamtgruppe. Diese Verteilung weicht signifikant ($p = .025$) von der anteilmäßigen Gleichverteilung beider Typen in der Gesamtstichprobe ab. Bei den Nominierten sind es 13 Erfolgstypen und 9 Suchende; in der Kontrollgruppe finden sich 14 Erfolgstypen und 4 Suchende, während die Psychotherapeuten 11 Erfolgstypen und 7 Suchende haben, siehe Abbildung 1. Diese Verteilungsunterschiede zwischen Gruppen und beiden Typen sind nicht signifikant (*Fishers exakter Test*: 1,77, $p = .48$).

Tabelle 7.16 Deskriptive Statistiken für Umweltkontrolle und Persönliches Wachstum.

	<i>M (SE)</i>	<i>Mdn</i>	<i>SD</i>	<i>Schiefe (SE)</i>	<i>Exzess (SE)</i>	<i>MIN</i>	<i>MAX</i>
UK	4,81 (.06)	4,86	,60	-,65 (.23)	,72 (.45)	2,79	6,00
PW	4,99 (.06)	5,07	,64	-,70 (.23)	,27 (.45)	3,14	6,00

Anmerkungen. N = 114. SE steht jeweils für den Standardfehler der Schätzung des Kennwertes.

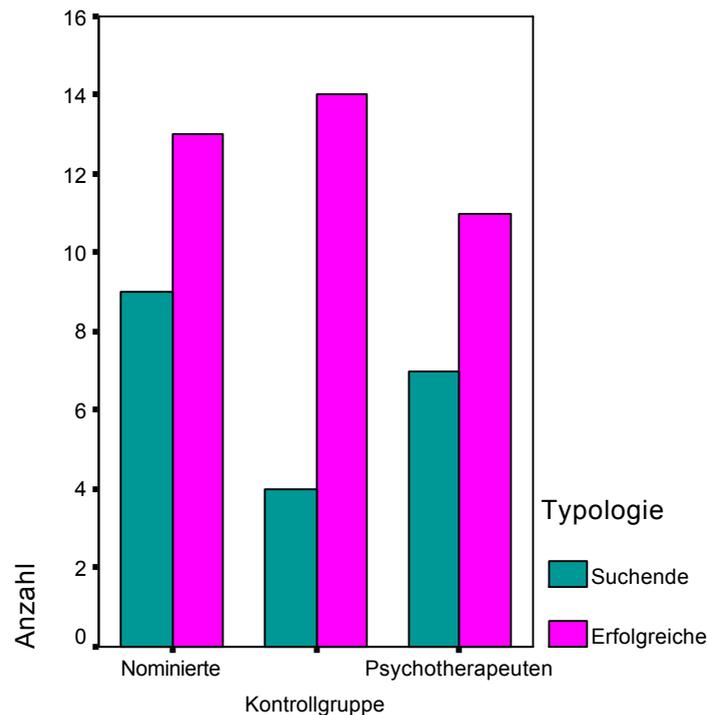


Abbildung 7.1 Verteilung der Typen Suchenden und Erfolgreiche in den Teilnehmergruppen.

Werden die Mittelwerte der Tests für diese beiden abhängigen Gruppen gegen das 5% Alphafehlerniveau geprüft, präsentieren sich einzig Unterschiede auf dem PWS, siehe Tabelle 7.17. Hier haben die Erfolgstypen signifikant höhere Weisheitswerte als die Suchenden ($M_D = 1,65$; $W = 2,45$, $df = 29,17$; $p = .02$). In den empirischen Daten liegen die Mittelwerte der Erfolgstypen in jedem Test über denen der Suchenden, sie werden jedoch nur im PWS signifikant.

Tabelle 7.17 W-Tests zu Mittelwertsunterschieden der Erfolgstypen und Suchenden in den Weisheitstestverfahren.

	n	M_D	SE	W	df	p	d'
3d-WS	58	,26	,23	1,17	37,90	.25 n.s	0,08
SAWS	58	2,5	5,31	,48	36,67	.64 n.s	0,07
PWS	57	1,65	,67	2,45	29,17	.02	0,52
TWR	54	,07	,25	,27	27,76	.79 n.s	0,09
AWR _{GESAMT}	58	,41	,85	,49	36,26	.63 n.s	0,14

Anmerkungen. Eine positive Differenz steht für höhere Mittelwerte der Erfolgstypen. SE steht für Standardfehler der Mittleren Differenzwerte M_D .

WA 4: Für die Erfolgstypen und die Suchenden werden die Interkorrelationen der Weisheitstests auf bivariater Ebene miteinander verglichen.

Abschließend werden die Korrelationsunterschiede beider Typen auf den Weisheitstestverfahren analysiert und in Tabelle 7.18 verkörpert. Die einzigen signifikanten Korrelationen für beide Typen sind jene von TWR zu AWR_{GES} , wobei die Korrelation der Erfolgstypen ($r = .28, p = .05$) lediglich leicht, die der Suchenden ($r = .70, p < .01$) dagegen hoch ist. Nur jeweils eine weitere signifikante Korrelation existiert in beiden Typen. In der Gruppe der Erfolgstypen ist dies die leichte Korrelation zwischen AWR_{GES} und 3d-WS ($r = .32, p = .03$). Dagegen ist die einzig weitere signifikante Korrelation für die Suchenden jene zwischen 3d-WS und SAWS ($r = .48, p = .03$) mittlerer Größe.

Anhand dieser Unterschiede in Mittelwerten und Korrelationen kann die Unterschiedlichkeit der beiden Typen in Hinsicht auf die Weisheitstests als belegt gelten. In den Mittelwerten aller Tests erreichen die Erfolgstypen jeweils ein höheres Niveau, der Unterschied wird jedoch nur für PWS statistisch signifikant.

Insbesondere geben jedoch die korrelativen Unterschiede Aufschluss über die Charakteristik der Typen. Für die Erfolgstypen korrelieren zwei der Weisheitstests signifikant. Das AWR hängt mit der 3d-WS und mit dem TWR zusammen, deren Größenordnung lässt ein konvergentes Niveau vermuten. Für den Typus der Suchenden dagegen ist die Korrelation zwischen TWR und AWR_{GES} wirklich hoch. Die zweite signifikante Korrelation für die Suchenden ist jene zwischen 3d-WS und SAWS.

Tabelle 7.18. Interkorrelationen der Weisheitstests der Erfolgstypen und Suchenden.

	3d-WS	SAWS	PWS	TWR	AWR_{GES}
3d-WS		.26	.23	.01	.32 [†]
SAWS	.48 [†]		.15	-.01	-.02
PWS	.23	-.00		-.06	-.08
TWR	-.02	.19	.27		.28 [†]
AWR_{GES}	.06	.32	.06	.70 ^{††}	

Anmerkungen. Das obere Dreieck beinhaltet die Korrelationen der 36 Erfolgstypen, das untere Dreieck die Korrelationen der 17 Suchenden nach listenweisem Fallausschluss.

^{††} $p < .01$, [†] $p < .05$ (einseitig).

8 Diskussion

Die Aufteilung dieses Unterkapitels orientiert sich an den Validierungszielen der Kriteriums- und Konstruktvalidierung der Studie. Sie integriert die Ergebnisse der Weiteren Analysen an entsprechender Stelle der Fragestellungen der Kernhypothesen. Zunächst werden die zentralen Befunde der Studie diskutiert, bevor auf die Planung, Stichprobe, Durchführung- und Auswertungsmethoden eingegangen wird.

8.1 Kriteriumsvalidierung anhand der Inneren Validität

Die Innere Validität der Testverfahren 3d-WS, SAWS, PWS, TWR und AWR wird über deren geteilte Varianz bestimmt. Diese Analysen beziehen sich einerseits auf die gleichzeitige Betrachtung aller Testverfahren zur Bildung von multiplen Validitätskoeffizienten als auch auf spezielle bivariate Vergleiche zwischen den einzelnen Testverfahren, die theoriegeleitete Unterschiede über das Ausmaß der geteilten Varianz prüfen.

8.1.1 Multiple Validitätskoeffizienten

Die Innere Validität als Teilbereich der Kriteriumsvalidität eines Fragebogens kann über einen multiplen Validitätskoeffizienten zu anderen Testverfahren ermittelt werden, die das zu erfassende Konstrukt ebenfalls messen (Lienert & Raatz, 1998). In der vorliegenden Studie wird dieser Koeffizient parallel für die Tests 3d-WS, SAWS, PWS, TWR und AWR erhoben. Beeinflusst ein gemeinsames, übergeordnetes Weisheitskonstrukt alle Tests, sollten sich idealerweise jeweils mittlere Validitätskoeffizienten bestimmen lassen.

Hypothese 1 zur Bestimmung der multiplen Validitätskoeffizienten zeigte, dass alle fünf zu validierenden Weisheitstestverfahren signifikante multiple Korrelationskoeffizienten erreichen konnten. Die Höhe der Koeffizienten ergibt folgendes Bild. Die Koeffizienten des 3d-WS und TWR sind nicht von dem in der Hypothese geforderten Niveau unterscheidbar. SAWS und TWR zeigen moderate Interkorrelationen, während die Höhe des Validitätskoeffizienten der PWS nicht zufrieden stellend ist.

Da, außer im Falle des 3d-WS, die Tests jeweils nur durch einen Test signifikante erklärt werden, gibt es kein gemeinsames übergeordnetes Konstrukt Weisheit, das einheitlich erfasst wird. Dies erlaubt zwei Erklärungen. Entweder wurden unterschiedliche Facetten des Konstruktes operationalisiert oder mindestens einem Testverfahren fehlt es an Validität (Lie-

ner & Raatz, 1998). Hier scheint vor allem die erstere Variante plausibel, da die Testverfahren unterschiedliche Definitionen von Weisheit operationalisieren. Verschiedene Erklärungsmodelle für Facettierungen der Weisheit wurden beispielsweise in Tabelle 1 des Abschnitts 3.4 systematisiert. Die Kategorisierung der Weisheitstestverfahren trennt zwischen der Modalität der Weisheit, (Prozess, Produkt und Persönlichkeit), der Orientierung der Weisheit analytisch oder synthetisch) und des Gegenstands der Weisheit (Selbstbezogen oder Allgemein). Diese konstruktbezogenen Unterschiede erklären die Ergebnisse, wobei die jeweiligen methodischen Unterschiede nicht unwesentlich sind.

Die PWS ist, als einzige, eine analytisch orientierte Skala. Während sie vollkommen unterschiedlich als das TWR verortet wird, teilt sie sich mit AWR den Gegenstandsbereich der Allgemeinen Weisheit und mit SAWS sowie 3d-WS die Modalität der Persönlichkeit. Diese Randstellung der analytischen PWS spiegelt sich offenbar im niedrigen multiplen Korrelationskoeffizient.

Nach der Kategorisierung sind 3d-WS und SAWS identisch verortet, während TWR und AWR sich lediglich in ihrem Gegenstandsbereich unterscheiden. Zusätzlich ist in beiden Paaren die methodische Herangehensweise gleich, entweder mehrstufige Persönlichkeitsinventare oder narrative Berichte. Diese Zweiteilung bestätigen die multiplen und bivariaten Korrelationen. Die erste Gruppe umfasst TWR und AWR, die zweite Gruppe 3d-WS, SAWS und PWS. Zwischen den Gruppen variiert die Modalität der Weisheitskonstrukte. Die erste Gruppe erfasst weise Produkte, die zweite Gruppe Persönlichkeitsfaktoren der Weisheit.

Diese Befunde gelten für die Gesamtgruppe. Das Bild der multiplen Validitätskoeffizienten in den Teilnehmergruppen kann jedoch differieren. Da unterschiedliche Kriterien für die Validierung eines Tests herangezogen werden können, ist nach Lienert und Raatz (1998) nicht nur von einer Validität zu sprechen. Ein Test der gegenüber mehreren Kriterien validiert wurde, hat somit mehrere Validitäten, die sich in ihrer Höhe unterscheiden können. Ein solcher Befund wird als differentielle Validität bezeichnet. Die drei Teilnehmergruppen zur Äußeren Validierung sind derartige Kriterien, für die differentielle Validität besteht.

Während es in der Kontrollgruppe keinerlei signifikante multiple Korrelationen gibt, offenbaren Nominierte und Psychotherapeuten höhere multiple Korrelationen als die Gesamtgruppe. Die multiplen Validitätskoeffizienten beider Gruppen sind nicht signifikant von dem erwarteten Niveau aus Hypothese 1 zu unterscheiden. Die Testverfahren kommen für die Nominierten und die Psychotherapeuten zu vergleichbaren Ergebnissen, was von validen Testverfahren erwartet wird (Häcker, Leutner, & Amelang, 1998). Dies bedeutet in vereinfachter Form, dass die Übereinstimmung zwischen hohen und hohen oder niedrigen und nied-

rigen Werten in den einzelnen Weisheitstests für einen Teilnehmer der Nominiertengruppe oder Psychotherapeuten wesentlich höher ist als für die gemachten Kontrollpersonen.

Betont sei, dass es sich bei Weisheit um ein theoriebelegt breites Konstrukt unterschiedlicher Facetten handelt. Der Zusammenhang dieser Facetten, statistisch deren geteilte Varianz, ist auf das Vorhandensein von Weisheit zurückführbar. Ohne die Verbindung durch Weisheit kovariieren die einzelnen Facetten nicht miteinander. Es ist daher ein der Validität der Testverfahren zuträglicher Befund, dass die Kriterien der Äußeren Validierung, Nominierung und Beruf Psychotherapeut, diesen Zusammenhang belegen, während für die Kontrollgruppe kein Zusammenhang herrscht.

8.1.2 Bivariate Korrelationen

Die Bestätigung der Hypothesen 2 und 3 stützen die aus bisherigen Studien und der Kategorisierung abstrahierten Annahmen zur Enge der Beziehung verschiedener Weisheitstests. Der Befund der Hypothese 2 zur Unkorreliertheit der Testverfahren PWS und TWR repliziert die Ergebnisse der Studie von Wink und Helson (1997). Dieses Ergebnis stützt somit auch die theoretische Trennung von Transzendenter und Praktischer Weisheit aus Abschnitt 3.3.1. Die Hypothese 3, dass die unterschiedliche Modalität der Weisheitstestverfahren über deren korrelative Nähe bestimmt, konnte bestätigt werden. 3d-WS und SAWS teilen untereinander mehr gemeinsame Varianz als mit einem der Testverfahren TWR, AWR oder PWS. Ein Methodeneinfluss kann hier nicht ausgeschlossen werden.

Generell können methodische Effekte nicht ausgeschlossen werden. Ohne jedoch auf die Ergebnisse der konvergenten Validierung vorzugreifen, kann der ebenfalls als möglicher Weisheitstest diskutierte WUSCT (Staudinger et al., 2005) zu einer ersten Klärung betrachtet werden. Dieses projektive Verfahren der Satzergänzung steht der Erhebungsmethode des Selbstberichtes von TWR und AWR näher als den Persönlichkeitsfragebögen. Diese Trennung findet sich jedoch nicht in den Korrelationskoeffizienten, was gegen die Bedeutung eines derartigen Methodeneffektes spricht. Das vorgestellte Klassifikationsschema liefert auch hier eine stimmige Erklärung, da sich der WUSCT die Synthetische Orientierung mit diesen Tests teilt. Mit der analytischen Skala PWS ist er nicht signifikant korreliert.

8.1.3 Zusammenfassung der Befunde zur Inneren Validität

Zusammenfassend betrachtet erreichen die Ergebnisse zur Inneren Validität nahezu die erwarteten Größenordnungen. Angesichts der unterschiedlichen Definitionen der Weisheit, aus denen unterschiedliche Facetten jedes Tests resultieren, erscheint die Höhe der multiplen Validi-

tätskoeffizienten stimmig. Es kann nicht von einem gemeinsamen übergeordneten Konstrukt *Weisheit* ausgegangen werden, da zwei von einander getrennte Varianzquellen in den Tests wieder zu finden sind, die anhand des Kategoriensystems aus Kapitel 2.3.3 als Modalitätsunterschiede erkannt werden können. Die beiden Tests TWR und AWR zur Messung weiser Produkte erfordern offenbarte Introspektion, während die Fragebogen 3d-WS, SAWS und PWS Selbstzuschreibung von Persönlichkeitseigenschaften verlangen, die nur im Falle hoher Ausprägung aller Dimensionen oder Items, auf Weisheit hindeuten. Die Ergebnisse zur Beziehung von verschiedenen Weisheitstestverfahren untereinander sind hypothesenkonform. Die Befunde zeigen die Bedeutung des vorgestellten Kategoriensystems der Weisheitstests.

Interessanterweise interagieren die Kriterien der Äußeren Validierung mit der Fragestellung der Inneren Validität in Form einer differentiellen Validität. Die Weisheitstests erreichen für Nominierte und Psychotherapeuten höhere multiple Validitätskoeffizienten als für die Kontrollgruppe. Dies spricht sowohl für die Validität der Äußeren Kriterien als auch für die Innere Validität der Testverfahren. Der nächste Abschnitt thematisiert die Prüfung der Äußeren Validität anhand der angenommenen Unterschiedshypothesen.

8.2 Kriteriumsvalidierung anhand der Äußeren Validität

Die Äußere Validität steht für die Übereinstimmung eines möglichst objektiven, äußeren Merkmals mit den Ergebnissen der zu validierenden Testverfahren. Wegener fasst die Probleme der Kriteriumsvalidität wie folgt zusammen: „Es gibt sehr häufig keine hinreichend genau gemessene Kriteriumsvariable für die Validierung einer Messung, und sofern es sie doch gibt, ist fraglich, worin eigentlich der Anlaß für die neue Messung besteht (Wegener, 1983, 95-96, nach Schnell, Hill & Esser, 1999, S, 150).“ Für diese Studie werden das Kriterium der Fremdeingeschätzten Weisheit und das Kriterium Beruf Psychotherapeut verwendet.

Es handelt sich bei den Hypothesen 4 bis 6 streng genommen nicht um die Replikation von Studienergebnissen, sondern um die Generalisierung der Befunde der Allgemeinen Weisheit auf andere Tests (Baltes et al., 1995). Es war anzunehmen, dass die Nominierten höhere Mittelwerte in einem Weisheitstest erreichen als deren gematchte Kontrollgruppe. Genauso sollten sich Psychotherapeuten aufgrund ihrer beruflichen Expertise gegenüber der Kontrollgruppe durchsetzen. Die Hypothesen konnten jedoch sämtlich nicht bestätigt werden. Keines der Testverfahren zeigte hypothesenkonforme Mittelwertsunterschiede.

Auffällig ist jedoch, dass die adaptierte Version (Helson & Srivastava, 2001, 2002) jener Allgemeinen Weisheitsaufgabe von Baltes et al. in dieser Studie keine Gruppenunter-

schiede indizierte, während gerade in dieser Aufgabe die Gruppenunterschiede in den Interviewstudien deutlich wurden. Die fehlende Übereinstimmung könnte an der Adaption des Testverfahrens liegen. Eine Betrachtung über die Beschränkungen der durchgeführten Testanwendung ergeht im nächsten Abschnitt. Die hohe Bedeutung von beruflicher Spezialisierung für das Allgemeine Weisheitsparadigma, belegt in Studien von Smith et al. (1994) und Staudinger et al. (1998), konnte sich in den hier geprüften Weisheitstests nicht durchsetzen.

Entgegen der Vorhersage zeigte sich eine Überlegenheit der Psychotherapeuten in Transzendentaler Weisheit gegenüber den Weisheitsnominierten. Da der TWR nach einem Beispiel persönlich erlangter Weisheit der Teilnehmer fragt, und die Nominierten gerade wegen ihrer wahrgenommene Weisheit nominiert wurden, muss diskutiert werden, warum sie in diesem Test die niedrigsten Mittelwerte erzielten. Im Grunde kann über die Antwort in diesen Test eine mögliche Rechtfertigung der Nominierung durch den Nominierten geliefert werden. Widerstände gegen die Nominierung könnten daher an dieser Stelle sichtbar werden. Im späteren Abschnitt 8.6.3 werden einzelne Begründungen von Nominierten wiedergegeben, die die Beantwortung dieser Frage völlig ausgelassen haben. Die Diskussion dieser Gründe lässt obige Begründung der niedrigen Werte im TWR plausibel erscheinen.

Die Ursache der fehlenden Überlegenheit der Psychotherapeuten oder der Nominierten kann möglicherweise in der Definition von Weisheit der Testverfahren gefunden werden. Wenn diese Teilnehmergruppen nicht überlegen sind, die Äußere Validität von Nominierung und Beruf Psychotherapeut aber als gegeben angenommen wird, so muss gefolgert werden, dass die Tests nicht dieselbe Definition von Weisheit widerspiegeln, die den Äußeren Kriterien zu Grunde liegt. Aufgrund der sehr unterschiedlichen Definitionen von Weisheit in den fünf Tests ist es möglich, dass von jedem Test eine oder mehrere Facetten der Weisheit erfasst werden, doch andere wiederum nicht abgedeckt werden. Demnach besteht weiterer Testentwicklungsbedarf. Einige Anregungen in diese Richtung werden in Abschnitt 8.7 über Impulse für die weitere Forschung thematisiert.

Eine zentrale alternative Erklärung für die fehlenden Weisheitsunterschiede der Teilnehmergruppen ist jene, dass keine der drei Gruppen als weise zu beurteilen ist. Die Cut-Off-Wert Analyse des SAWS legt diesen Schluss nahe. Die Teilnehmergruppen der Studie zeichneten sich somit zwar durch hohe Persönlichkeitsreife aus, siehe 8.3, Weisheit erreichen sie jedoch nicht, weshalb auch keine bedeutsamen Gruppenunterschiede zu finden sind.

Als Zusammenfassung der Ergebnisse der Äußeren Validität bietet sich daher folgende Sichtweise an. Da die drei Teilnehmergruppen im Mittel gleiche Weisheitswerte erreichen, konnten die zur Äußeren Validierung gewählten Kriterien der Nominierung und des Berufs-

bildes Psychotherapeut ihre Bedeutung für die Validierung der Weisheitstestverfahren nicht hypothesenkonform bestätigen. In Anbetracht der differentiellen Kriteriumsvalidität muss anerkannt werden, dass jedoch die bereits diskutierten Unterschiede für eine gewisse Validität der Kriterien sprechen. Die Höhe der Gruppenmittelwerte in den Weisheitstests lässt jedoch schließen, dass das Weisheitsniveau der Teilnehmer, insbesondere der Nominerten hätte höher sein können, siehe Anhang F.2. Entweder ist die Validität der Nominierung anzuzweifeln, oder die Übereinstimmung von Laiennominierungen der Weisheit zu schriftlichen Erhebungsverfahren ist mangelhaft. Da alle Tests gleiche Ergebnisse liefern, beginnt die Fremdwahrnehmung von Weisheit vermutlich auf einem niedrigeren Niveau, welches auch von Kontrollgruppenmitgliedern erreicht werden konnte.

8.3 Konstruktvalidierung anhand der Konvergenten Validität

Die Konstruktvalidierung prüft die Verortung eines Konstruktes in den Rahmen wissenschaftlich bereits anerkannter Konstrukte. Auch auf theoretischer Ebene verwandte Konstrukte sollten in konvergenter Größenordnung miteinander korrelieren. Die Konvergente Validierung der *Weisheitstests* ergeht in den Hypothesen 7 bis 9 zu den Persönlichkeitsreifeindikatoren Psychologisches Feingefühl, Persönliches Wachstum und Ego-Entwicklung.

Um zu einem allgemeinen Urteil der konvergenten Validität über die drei Tests PF, PW und WUSCT kommen zu können, müssen diese ebenso in konvergenter Beziehung zueinander stehen. Diese Annahme konnte jedoch nicht bestätigt werden. Der WUSCT korreliert nicht in entsprechender Höhe mit den beiden anderen Tests. Neben den methodischen Unterschieden zwischen einem projektiven Verfahren und Persönlichkeitsfragebögen sind inhaltliche Differenzen der Testverfahren anzunehmen. Nicht auszuschließen ist, dass bei dem durchschnittlich sehr hohen Niveau der Teilnehmer auf dem WUSCT die Beziehungen zu PF oder PW nicht höher zu erwarten sind. Andererseits sind inhaltliche Differenzen der Konzepte möglicherweise in der kognitiven Ausrichtung des WUSCT zu finden, siehe Kapitel 4.2.2.3, 4.2.3.2 und 4.2.3.3, die von PF oder PW nicht abgedeckt wird. Dies belegt, dass kein gemeinsames Konstrukt von Persönlichkeitsreife für die vorliegende Stichprobe durch die drei dazu erwähnten Testverfahren erfasst werden konnte. Wenn es sich also um zwei inhaltlich oder methodisch getrennte Aspekte von Persönlichkeitsreife handelt, die von diesen drei Testverfahren erfasst werden, müssten sich die Beziehungen von PF und PW zu den Weisheitstests ähneln, aber von denen des WUSCT unterscheiden. Diese Annahme erscheint plausibel aufgrund der beobachteten Daten, die im Folgenden zunächst getrennt berichtet werden.

8.3.1 Psychologisches Feingefühl und Persönliches Wachstum als Indikatoren der Persönlichkeitsreife

Die angenommenen Beziehungen zu Psychologischem Feingefühl und Persönlichem Wachstum stellten sich für die Testverfahren 3d-WS und SAWS erwartungsgemäß ein. Der PWS ist einzig mit dem PF signifikant korreliert. Zu TWR und AWR konnte weder PF noch PW eine konvergente Beziehung belegen. Auch bei Einzelanalyse der drei Teilnehmergruppen gibt es keine signifikante Beziehung zwischen TWR oder AWR zu entweder PF oder PW.

Die Konvergente Beziehung von PW zu den beiden Weisheitstests 3d-WS und SAWS tritt erwartungsgemäß ein (Ardelt, 2003; Helson & Srivastava, 2001; Mickler, 2004). Entgegen der Studie zu Selbstbezogener Weisheit und PF (Mickler, 2004), die eine fehlende konvergente Beziehung belegte, zeigte sich die Wahl eines alternativen Fragebogens für PF in dieser Studie effektiv. Im Grunde sind somit die konvergenten Korrelation zwischen Allgemeiner Weisheit und PF aus den Studien von Pasupathi und Staudinger (2001) und Staudinger et al. (1997) auf die anderen Weisheitstests 3d-WS, SAWS und PWS übertragen, repliziert.

Dass für die Testverfahren TWR und AWR keine Konvergente Validität zu Persönlichkeitsreife nachgewiesen werden konnte, widerspricht den theoretischen Annahmen. Gerade Persönliches Wachstum sollte mit dem TWR korrelieren. Erklärbar sind diese fehlenden Beziehungen durch die mangelhafte Beantwortung der Frage durch die Teilnehmer. Wären längere und somit detailliertere Antworten gegeben worden, wäre eine adäquate Beurteilung in absoluter und relativer Bewertung leichter. Da die durchschnittlichen Antwortlängen der Teilnehmer sehr gering waren und sich hohe Korrelationen zwischen Antwortlänge und Gesamtbewertung im TWR oder AWR zeigten, muss davon ausgegangen werden, dass diese Tests im Grunde die eigentlichen Konstrukte nur verzerrt messen konnten. Es erscheint an dieser Stelle am plausibelsten, dass die Bereitschaft, ausführlich auf die Fragen zu antworten, etwas aus der eigenen Entwicklung im TWR preiszugeben, oder die Auseinandersetzung mit einer solchen existenziellen Situation im AWR gründlich zu beschreiben, mit der Selbstoffenbarungstendenz der Teilnehmer zusammenhängt. In diesem Sinne kann Micklers (2004) Haltung, kurze Antworten als generelle Antwortverweigerung anzusehen, nachvollzogen werden. Die Verwendung von TWR und AWR mit der gegebenen Instruktion, umrahmt von geschlossenen Persönlichkeitsfragebögen, führte demnach nicht dazu, dass die Probanden ihre Gedanken zu den Themen der Fragestellungen ausführlich zu Papier brachten. Insgesamt ist das über TWR und AWR somit gemessene Konstrukt der Selbstoffenbarung von Transzendenter Reife und Allgemeiner Weisheit in den nicht konvergenten Korrelationen zu Persönlichem Wachstum und Psychologischen Feingefühl stimmig repräsentiert.

8.3.2 Ego-Entwicklung als Indikator der Persönlichkeitsreife

Der WUSCT als Maß der Ego-Entwicklung zeigte andere Korrelationsmuster als PW und PF. Das Maß korreliert signifikant mit 3d-WS, SAWS, TWR und AWR. Die Größenordnung dieser Beziehung kann weder klar dem konvergenten noch dem diskriminanten Bereich zugeordnet werden. Der WUSCT korreliert nicht signifikant mit PWS. In anderen Studien zeigte sich WUSCT moderat zu Selbstbezogener Weisheit (Mickler, 2004), Allgemeiner Weisheit (Pasupathi et al. 2001) und dem latenten Weisheitskonstrukt HS von Helson und Srivastava (2001) korreliert, das sich aus TWR, AWR und PWR zusammensetzt.

Die Zusammenhangsmuster wurden bereits im Abschnitt 8.1.3 anhand des Kategoriensystems der Weisheitstests erklärt. Der WUSCT erfasst ein Synthetisches Persönlichkeitsbild, wie es auch von SAWS, 3d-WS, TWR oder AWR abgedeckt wird. Die geringen Korrelationen sind möglicherweise auf die generell hohe Ego-Entwicklung der Teilnehmer zurückführbar. Alle drei Teilnehmergruppen erzielen vergleichbare und höhere Werte als die Normalbevölkerung im WUSCT. Sie befinden sich mehrheitlich auf Gewissenhaften, Individualistischen oder den beiden oberen Stufen der Autonomen oder Integrierten Persönlichkeit.

Nicht auszuschließen aber unwahrscheinlich ist, dass die Testverfahren TWR und AWR mit WUSCT einzig aufgrund von Methodeneffekten oder Störvariablen korrelieren. Die Intelligenzvariable nach HAWIE-R bietet sich aufgrund der Unkorreliertheit zu den Tests nicht als gemeinsame Varianzquelle an. Alternativ kann von unterschiedlich starker Motivation oder Selbstoffenbarungstendenz der Teilnehmer zur Bearbeitung der offenen Antwortformate WUSCT, TWR und AWR ausgegangen werden, die zu unterschiedlich guter Leistung in den Tests, je nach Motivation, führte.

In Zusammenschau der vorliegenden Befunde ist die Deklaration des WUSCT als Maß der Selbstbezogenen Weisheit (Staudinger et al., 2005) oder als Prozessmessung von Weisheit (Hartmann, 2004) nicht gesichert. Die dazu erwarteten höheren Korrelationen zwischen WUSCT und den Weisheitstests müssen entweder auf mangelnde Validität der Tests oder auf Besonderheiten der Stichprobe zurückgeführt werden. Wesentlich ist, dass alle drei Teilnehmergruppen vergleichbare und hohe Mittelwerte im WUSCT erreichen, ihnen somit eine hohe Persönlichkeitsreife nach Definition des WUSCT zugesprochen werden kann. Weitere Erforschung zur Stellung des WUSCT entweder als Weisheitstest oder als konvergentes Maß sollte an einer größeren Stichprobe erfolgen. Insbesondere aufgrund der standardisierten Administration, des ökonomischen Umfangs, der ausführlichen Validierung und Normierung und des projektiven Charakters des Tests, ist der WUSCT ein interessantes Verfahren.

8.3.3 Zusammenfassung der Befunde der Konvergenten Konstruktvalidierung

Das Fazit der Konvergenten Validierung ist, dass 3D-WS und SAWS die für einen Weisheitstest erwarteten Befunde mehrheitlich zeigen. TWR und AWR belegen keine konvergente Beziehung zu den Persönlichkeitsreifeindikatoren, da sie durch Effekte der unterschiedlichen Selbstoffenbarung, nur bedingt das eigentliche Konstrukt der Transzendenten oder Allgemeinen Weisheit präsentieren. Somit ergeht der Rat, beide Tests entweder über mündliche Erhebung zu erfassen, oder einerseits das Commitment der Teilnehmer zu steigern und andererseits durch veränderte Instruktion gezielter das gewünschte Antwortverhalten zu fördern. Impulse in diese Richtung werden in Abschnitt 3.5.3 gegeben. Das PWS einzig mit PF korreliert, und somit eine gewisse Persönlichkeitsreife beweist, die jedoch nicht auf Persönliches Wachstum ausgerichtet ist, belegt die Unkorreliertheit zu PW und WUSCT. Dies ist stimmig zur praktischen, analytischen Ausrichtung der PWS.

8.4 Konstruktvalidierung anhand der Diskriminanten Validität

Die Konstruktvalidität eines neu zu etablierenden Konstruktes zeigt sich auch in seiner Eigenständigkeit gegenüber anderen in der Forschungswelt anerkannten Konstrukten. Eine solche Prüfung erwartet geringe korrelative Beziehungen theoretisch nicht verwandter Konstrukte. Die Hypothesenprüfungen 10 bis 13 zur diskriminanten Validität der Weisheitstests zu Adaptivität, Intelligenz, Alter und Geschlecht zeigen allerdings uneinheitliche Ergebnisse und werden somit in unterschiedlichen Abschnitten diskutiert.

8.4.1 Umweltkontrolle als Indikator der Adaptivität

Hypothesenkonform korrelieren TWR und AWR nicht mit Umweltkontrolle. Die Adaptivität, gemessen über Umweltkontrolle, zeigt entgegen den Erwartungen sehr hohe Korrelationen zu 3d-WS, SAWS und PWS. Diese hohen Korrelationen müssen bedeuten, dass die Weisen der Studie auch die Anforderungen ihrer Umwelt bewältigen, bzw. diese zur eigenen Zufriedenheit erfüllen. Möglicherweise sind die relativ hohen Korrelationen zu UK auch in dessen niedriger Schwierigkeit zu finden, da die Mittelwerte im UK außerordentlich hoch liegen.

Es ist festzuhalten, dass die Fragebögen zur Weisheitsmessung eine größere Nähe zu Umweltkontrolle und somit Adaptivität haben, als dies aus theoretischer Sicht anzunehmen war. Die Befunde sprechen daher dafür, dass Umweltkontrolle dieses Niveaus indikativ für Weisheit ist. Ob dieser Befund als fehlende diskriminante Validität der Testverfahren zu deuten ist, sollte anhand von Replikationsstudien mit einer schwierigeren Umweltkontrollskala

geklärt werden. Zumindest die Möglichkeit eines Methodeneffektes entkräftet die Ergebnisse der Weiteren Analysen. Würde ein Methodeneffekt das Kovariieren der Variablen determinieren, so müssten sich dieselben Zusammenhangsmuster bei jeder der drei Teilnehmergruppen finden lassen. Die kritischen positiven Korrelationen des UK zu 3d-WS, SAWS und PWS treten nur für die Teilnehmergruppen der Nominierten und Psychotherapeuten auf. In der Kontrollgruppe sind Korrelation von UK zu SAWS und PWS nicht signifikant, während eine leicht negative Beziehung von AWR zu UK signifikant ist.

Eingehend auf die Adaptivitätskritik von Staudinger et al. (2005) an Ardelts 3d-WS muss dieser stattgegeben werden. Die korrelative Beziehung zwischen 3d-WS und UK ist in allen drei Teilnehmergruppen hoch signifikant. Bei spezieller Betrachtung der Nominierten reicht die Korrelation sogar in Größenordnungen einer hohen Korrelation. Die Korrelation in Ardelts Studie (2003) zu einer anderen Skala der Umweltkontrolle ist von dem hier vorliegenden Korrelationsniveau nicht zu unterscheiden.

8.4.2 Zusammenfassung der Befunde zur Typologie der Suchenden und der Erfolgstypen

An dieser Stelle eingehend auf die Ergebnisse der Weiteren Analysen zur Typologie von Helson und Srivastava (2001), erscheinen die hohen Korrelationen der UK in einem neuen Licht. Durch Kombination von PW und UK zu vier Typen erwachsenen Reifeverhaltens und spezieller Analyse der Erfolgstypen (hohe Werte auf beiden Skalen) und dem Typus der Suchenden (hohe PW-Werte und niedrige UK-Werte), fanden sich in besagter Studie wichtige Unterschiede hinsichtlich der Weisheit dieser Teilnehmer. Die Suchenden erreichten höhere Mittelwerte auf dem latenten Weisheitskonstrukt der Forscher, bestimmt durch die Tests TWR, AWR und PWR und zeigten für Weisheit stimmigere korrelative Beziehungen zu weiteren Konstrukten als die Erfolgstypen. Auch in der vorliegenden Arbeit zeigen sich Unterschiede in den Mittelwerten und korrelativen Beziehungen dieser beiden Typen.

Die Verteilung von Suchenden und Erfolgstypen ist in den Teilnehmergruppen gleich. Somit scheinen die Kriterien der Äußeren Validierung, zumindest für die hier getesteten Versuchspersonen, unabhängig von diesen Typen erwachsenen Verhaltens zu sein. Auch für Männer und Frauen ergeben sich Gleichverteilungen der Typen.

Bei Betrachtung der Mittelwerte der Gesamtgruppe liegen die der Erfolgstypen auf allen Tests höher, und erreichen im PWS statistische Signifikanz. Gerade die Betonung adaptiver Eigenschaften der PWS, nüchtern denkend, reif, vernünftig, leichtsinnig (umgepolt), legt die Dominanz des Typus der Erfolgreichen nahe.

Diese Befunde scheinen den Ergebnissen von Helson und Srivastava (2001) zu widersprechen, deren Suchende auf dem latenten Faktor HS aus AWR, TWR und PWS die höheren Werte erreichten. Da dieser Faktor für die vorliegende Studie nicht gebildet werden konnte, ist eine genaue Replikation nicht möglich. Werden jedoch die Interkorrelationen zwischen TWR und AWR inspiziert, können Schlüsse gezogen werden. Angemerkt sei, dass PWS und TWR unkorreliert sind. Für die Suchenden korrelieren die Tests in hohem Maße, während sie für die Erfolgstypen nicht zusammenhängen. Somit ist anzunehmen, dass auch in dieser Studie die Suchenden die höheren Faktorwerte auf dem latenten Faktor HS von Helson und Srivastava erzielen könnten. Dieser Befund scheint unabhängig von den Mittelwerten dieser Gruppen auf den Weisheitstests zu sein.

Die Korrelation von TWR und AWR erklärend, bedarf es bei beiden Tests der offenbarten Introspektion. Diese stellt für den Typus der Suchenden einen wichtigen Betrag zum Ziel der Steigerung des eigenen Persönlichkeitswachstums dar. Dieses Ziel verfolgen Erfolgstypen mit wesentlich weniger Anstrengung als ihre gesellschaftlichen Aufgaben und Ziele (Helson & Srivastava (2001)). Bei Analyse der Gesamtgruppe kann sich die negative Korrelation von UK zu TWR und AWR in der Teilgruppe der Suchenden nicht durchsetzen.

Mittlere Korrelationen zeigen die Suchenden zwischen 3d-WS und SAWS. Für die Erfolgstypen korrelieren diese Tests nicht, dagegen sind PWS und 3d-WS signifikant korreliert. Der 3d-WS ist folglich für beide Teilnehmertypen bedeutsam. Für die Gruppe der Erfolgreichen stehen vermutlich die adaptiven Elemente der Skalen miteinander in Beziehung.

Die Unterschiede für die Gruppe der Suchenden können in den drei Facetten des 3d-WS begründet sein. Da der 3d-WS als Faktor analysiert wird, können vergleichbare Faktorwerte aus unterschiedlichen Profilverläufen ergehen. Die beiden Typen könnten daher unterschiedliche Profilverläufe auf den drei Dimensionen im 3d-WS aufweisen. Da alle drei Testverfahren reflektive und kognitive Aspekte abdeckt, bietet sich hier geradezu die affektive Dimension zur Erklärung an.

Da die PWS keine emotionalen Aspekte erfasst, könnte die gemeinsame Varianz von SAWS und 3d-WS von diesen erklärt werden. Die Emotionsregulationsskala des SAWS erfasst die Kontrolle der eigenen Gefühle mit Betonung auf die Minderung negativer Emotionen. Zieht man die Studie von Helson und Srivastava (2001) heran, so zeichnen sich die Erfolgstypen gerade durch ein solches Muster der Emotionsoptimierung aus. Jedoch besteht die Korrelation zwischen SAWS und 3d-WS nicht für diese, sondern für die Suchenden. Der 3d-WS erfasst in seiner affektiven Dimension Mitgefühl und Zuneigung in sehr hoher Ausprägung gegenüber allen Menschen. Eine derart grundsätzlich respektvolle, emotionale Wert-

schätzung aller Menschen hängt nicht unmittelbar mit Emotionsregulationskompetenz zusammen. Werden die Korrelationen zwischen den beiden Unterskalen inspiziert, korrelieren diese für die Suchenden sehr hoch ($r = .80, p < .00$), während eine moderate Korrelation für die Erfolgstypen existiert ($r = .37, p = .01$). Besonders interessant ist, dass für die Nominierten ($r_{NOM} = .72, p < .00$) ebenfalls eine viel höhere Korrelation zwischen beiden Skalen existiert als für die anderen beiden Teilnehmergruppen ($r_{KG} = .38, p = .01$; $r_{PSY} = .33, p = .01$). Die Vorteile der Suchenden und Nominierten liegen jeweils im Bereich eines großen Effektes ($q_{S,E} = .70$; $q_{NOM,KG} = .51$). Anhand dieser Korrelationen erscheint plausibel, dass beide emotionalen Komponenten für Weisheit wichtig sind. Weisheit verlangt folglich die Wertschätzung der menschlichen Natur gemeinsam mit der Fähigkeit zur Emotionsregulation.

8.4.3 Der Wortschatztest als Indikator der Intelligenz

Werden die anderen Fragestellungen der diskriminanten Validität betrachtet, ergeben sich weniger komplexe Zusammenhänge. Die hypothesenkonforme fehlende Beziehung zwischen psychometrisch erfasster kristalliner Intelligenz, über den Wortschatztest des HAWIE-R, wird für alle Testverfahren belegt. Offenbar sind die Weisheitstestverfahren nicht durch die gemessene verbale Kompetenz der Teilnehmer beeinflusst. Dies gilt insbesondere für TWR und AWR mit offenen Antwortformaten. Dieses Ergebnis bestätigt die Haltung der Weisheitsforscher, dass psychometrische Intelligenz nicht mit Weisheit gleichzusetzen ist (Ardelt, 2000b, 2003, 2004b; Clayton, 1982, Staudinger et al., 1997; Staudinger et al. 1998).

8.4.4 Alter

Die gefundenen Alterseffekte für 3d-WS und SAWS sind angesichts der theoretischen und empirischen Befundlage unerwartet eingetreten. In dieser Studie präsentieren die Testverfahren leicht positive Beziehungen zwischen Alter und Weisheit, deren Höhe nicht klar außerhalb des konvergenten Bereiches liegt. Dies kann auf einen unglücklichen Selektionseffekt zurückgeführt werden, der darin begründet ist, dass vor allem solche älteren Teilnehmer an der Studie partizipierten, die ein persönliches Interesse am Thema Weisheit und Persönlichkeitswachstum haben. Genau für diese Personen lässt sich aber aus theoretischen Gesichtspunkten erwarten, dass diese im Laufe ihres Lebens einen stetigen Zuwachs ihrer eigenen Weisheit erzielen können (Ardelt, 1997). Insbesondere die 3 Teilnehmer über 80 Jahren beeinflussen die Korrelationskoeffizienten bedeutsam, siehe Kapitel 7. 4. Werden sie nicht mit-analysiert, sind die Beziehungen zu 3d-WS und PWS insignifikant, während der SAWS nur knapp signifikantes Niveau erreicht. Es ist daher möglich, dass diese älteren Personen Zuge-

winne in Weisheit über den Verlauf ihres Erwachsenenalters erreichten, die nicht der Allgemeinbevölkerung entsprechen und ebenso wenig von hirnrorganischen Abbauprozessen beeinträchtigt wurden, die von einigen Weisheitsforschern (Ardelt, 2003; Lindenberger & Baltes, 1997) diskutiert werden.

8.4.5 Geschlecht

Für 3d-WS, SAWS und AWR zeigten sich hypothesenkonform eindeutig diskriminante Beziehungen zur Geschlechtsvariablen. Dagegen können die signifikanten Vorteile der Frauen in den Tests TWR und PWS nicht eindeutig konvergenter oder diskriminanter Größenordnung zugeordnet werden. Die Größenordnung dieses Unterschiedes ist nach η^2 jedoch gering, weshalb auf eine detaillierte Analyse dieses Befundes verzichtet wird.

Die hier signifikanten Tests PWS und TWR wurden bisher nur an Frauen erhoben. Die Überlegenheit der Frauen in 3d-WS und SAWS der Studien von Ardelt (2003) und Webster (2003) konnte auch in dieser Studie nicht repliziert werden. Die damalige Erklärung beider Autoren zur Überlegenheit der Frauen war, diese seien den Männern in den emotionsbetonten Unterskalen bzw. Dimensionen überlegen. Eine statistische Prüfung dieser Behauptung fehlte allerdings genauso wie deren Replikation. In Studien zur Allgemeinen Weisheit konnten keine Geschlechtsunterschiede gefunden werden (Staudinger & Baltes, 1996).

8.4.6 Zusammenfassung der Befunde der Diskriminanten Konstruktvalidierung

TWR und AWR sind hypothesenkonform diskriminant mit jedem der Konstrukte korreliert. Entgegen den Erwartungen zeigen sich konvergente Beziehungen von 3d-WS, SAWS und PWS zu Umweltkontrolle dieser Schwierigkeit. Zu beachten ist, dass die Mehrheit der Teilnehmer sehr hohe Werte in Umweltkontrolle erreichte. Um diesen Deckeneffekt bei Replikationsstudien zu vermeiden, bietet es sich an, ein schwierigeres Instrument zu verwenden. An dieser Stelle muss Umweltkontrolle als konvergentes Konstrukt zu Weisheit nach 3d-WS und SAWS angesehen werden. Die Befunde der Diskriminanten Validierung der Testverfahren zu Intelligenz über den Wortschatztest des HAWIE-Rs bestimmt, bestätigen die diskriminante Beziehung zu Weisheit. Für Alter zeigen sich leicht positive Beziehungen zu 3d-WS, SAWS und PWS. Diese bleiben nach Ausschluss der über 80jährigen einzigen für SAWS signifikant. Die korrelative Beziehung kann jedoch nicht eindeutig von konvergenter Größenordnung ausgeschlossen werden. Entweder sind Selektionseffekte oder das besondere Interesse für Persönlichkeitswachstum der Teilnehmer für diesen Befund verantwortlich. Die leichten Vor-

teile der Frauen bei PWS und TWR sind auf theoretischer Ebene nicht zu erklären, für 3d-WS, SAWS und AWR zeigen sich keine Unterschiede.

8.5 Schlüsse auf die Inhaltsvalidität der Weisheitstestverfahren

Die Thematik der Inhaltsvalidität gehört nicht zu den empirischen Forschungsfragen dieser Diplomarbeit. Wird diese Arbeit dennoch zur Erörterung der Inhaltsvalidität der Weisheitstests betrachtet, so sind folgende Erkenntnisse abzuleiten. Keiner der hier validierten Testverfahren scheint Weisheit in der vollen Komplexität dieses Konstruktes abzudecken. Im Hinblick auf die Synthetische Weisheit, spiegeln die Testverfahren verschiedene Facetten dieser wieder. Es fehlt somit an einem Test, der alle dieser Facetten vereinen kann. Auf der Suche nach Facetten der Weisheit lässt sich anhand der Unterskalen, Dimensionen oder Bewertungskategorien der Tests auf theoretischer Ebene folgender Vorschlag festmachen.

Weisheit integriert als emotionale Facette unbedingte Wertschätzung gegenüber Anderen, in Form von Mitgefühl und Verständnis (3d-WS, AWR), und die Fähigkeit zur Emotionsregulation (SAWS). Reflektion oder Introspektion und tiefere Einsicht (TWR, 3d-WS, SAWS, PWS) sind der zentrale Bestandteil, um Persönlichkeitswachstum und Weisheit zu erlangen. Um diesen Prozess zu unterstützen und dessen Ergebnisse zum Beispiel in Form von weisen Handlungen (AWR) anzuwenden, bedarf es einer gewissen kognitiven Facette der Weisheit (3d-WS, SAWS, PWS, AWR), die das Bedürfnis und die Kompetenz zum Verständnis von komplexen Vorgängen beschreibt. Darüber hinaus sind eine besondere Wertorientierung zur Maximierung des Allgemeinwohls, am ehesten durch die emotionale Wertschätzung des menschlichen Lebens an sich (3d-WS, AWR) und Moralbewertungskategorie des AWR repräsentiert, sowie Offenheit und Humor (SAWS) als Persönlichkeitseigenschaften notwendig, um eine weise Person abzubilden. Die Bedeutung der Werteorientierung betont vor allem Sternberg (1990) auf theoretischer Ebene, während sie bei der Validierung der Selbstbezogenen Weisheit empirische Bedeutsamkeit erlangte (Mickler, 2004).

8.6 Begrenzungen der Studie

Die Begrenzungen der vorliegenden Studie können in verschiedene Bereiche aufgliedert werden. Zunächst werden das Studiendesign, die Stichprobenziehung und deren Zusammensetzung, die verwendeten Testverfahren und schließlich die methodische Auswertung kritisch diskutiert. Im nächsten Abschnitt erfolgt eine Sammlung von Impulsen für weitere Studien zum Bereich der Weisheitstestvalidierung. Eine Ausdehnung auf weitere Fragestellungen in

Zusammenhang mit Weisheit würde den Rahmen dieser Studie sprengen.

8.6.1 Begrenzungen des Studiendesigns

Die Konstrukt- und Kriteriumsvalidierung von Fragebogenverfahren zur Weisheitsmessung wurde in dieser Studie querschnittlich an drei verschiedenen Personengruppen durchgeführt. Die gleichzeitige Validierung mehrerer Testverfahren hat Vorteile und Nachteile gegenüber getrennt durchgeführten Validierungen. Vorteile einer derart vergleichenden Studie liegen in der Möglichkeit der gegenseitigen Konstruktvalidierung der Fragebögen bei Analyse einer einzigen Stichprobe. Dieses, das Nebengütekriterium der Ökonomie berücksichtigende Vorgehen, birgt jedoch Gefahr, falls die Mehrheit der Testverfahren nicht valide sind oder Verzerrungen in der gezogenen Stichprobe vorliegen. Unter diesen Bedingungen können valide Testverfahren als nicht valide eingestuft werden. Das Gegenteil läge vor, wenn ein nicht valides Instrument in einer verzerrten Stichprobe anhand weiterer nicht für das eigentliche Konstrukt valider Testverfahren als valide eingeschätzt würde.

Die Durchführung einer Kreuzvalidierung anhand einer größeren Stichprobe, wäre ein Verfahren, um die Risiken der Stichprobenverzerrung zu Minimieren. Eine Ausdehnung des Rekrutierungszeitraums bis ein ausreichender Stichprobenumfang für eine Kreuzvalidierung vorgelegen hätte, hätte den zeitlichen Rahmen einer Diplomarbeit sprengen. Dies liegt insbesondere an der Seltenheit von hoher Weisheit. Aus diesem Grund war ebenso eine durch Poweranalyse festgelegte Stichprobengrößenbestimmung nicht indiziert. In Kauf zu nehmen war daher, mögliche Effekte aufgrund der zu kleinen Stichproben nicht zu erkennen.

Aufgrund des Umfangs des Fragebogens wurde ebenso auf Maßnahmen zur Absicherung der Befunde durch Einflüsse von Methodeneffekten verzichtet. Ein solches Vorgehen hätte nach der Multitrait-Multimethode Methode (Campbell & Fiske, 1959) vor allem eine umfangreichere Erhebung der Konstrukte zur diskriminanten Validität erfordert. In der vorliegenden Studie wären daher alternative Erhebungsverfahren für Intelligenz oder Umweltkontrolle zu erheben gewesen.

Eine längsschnittliche Erhebung würde in ihrer Übereinstimmung von Ergebnissen über Messzeitpunkte für die Stabilität der Befunde im Sinne der Reliabilität sprechen. Darüber hinaus ist die interne Validität von längsschnittlichen Untersuchungen jener von querschnittlichen Designs überlegen. Da Weisheit als sehr stabile Persönlichkeitseigenschaft gilt, ist die Absicherung der Validierungsergebnisse über die Zeit von nachrangiger Bedeutung gegenüber zentraleren Fragen der Kriteriums- und Konstruktvalidität. In Anlehnung an Solomons Viergruppenplan, hätte sich in der vorliegenden Studie die Erhebung eines weiteren

Messzeitpunktes angeboten. Da jedoch die Erhebung anonym und sukzessive erfolgte, ist eine wiederholte Erhebung der Daten nach einem festen zeitlichen Intervall nicht möglich.

Bezüglich der Externen Validität sind quasiexperimentelle Designs, wie die vorliegende querschnittliche Studie, reinen Laborstudien überlegen. In dieser Studie führen die Probanden die Erhebung selbst angeleitet innerhalb ihres natürlichen Tagesablaufes und ihrer gewohnten Umgebung durch. Nachteile einer solchen Erhebung liegen jedoch in der Internen Validität, die für Erfassung von Ergebnissen unter Kontrolle von bekannten und Konstanthaltung von möglichen Störeinflüssen steht, was besonders unter Laborbedingungen möglich ist.

Bei den quasiexperimentellen Designs wird zur Steigerung der Internen Validität von Bortz und Döring (2006) empfohlen, Kontrollgruppen derart zu erheben, dass möglichst viele Einflussfaktoren über die Experimentalgruppe und die Kontrollgruppe gleichverteilt sind. Dieses Verfahren erfordert die Bildung von so genannten Matched Samples oder gepaarten Stichproben. In der vorliegenden Studie wurde daher zur Nominierengruppe eine gematchte Kontrollgruppe rekrutiert, deren Teilnehmer nach Alter, Geschlecht, Beruf und Bildungsniveau jeweils zu einem Nominierten passten. Somit wurden diese potentiellen Störvariablen und solche, die mit diesen kovariieren, auf versuchsplanerischer Ebene konstant gehalten. Im folgenden Abschnitt werden die einzelnen Teilnehmerstichproben in deren Rekrutierung und Zusammensetzung kritisch diskutiert.

Da die Fragebogen in selbstadministrativer Weise erhoben wurden, ist jedoch auch nicht auszuschließen dass eine unbestimmte Anzahl von Störeinflüssen auf die Bearbeitung der Fragebogen einwirkte. Es wird angenommen, dass sich die Einflüsse in der Summe der Teilnehmer in ihrer Wirkrichtung aufheben. Um nur einen möglichen Einflussfaktor exemplarisch zu nennen, könnte die Bearbeitung des Fragebogens durch die Interaktion mit einer weiteren Person beeinflusst worden sein, mit der beispielsweise über Tests oder Items gesprochen wurde. Um die Bedeutung eines solchen Effektes zu verdeutlichen, sei auf die Studien des Allgemeinen Weisheitsparadigmas verwiesen (Smith et al., 1994; Staudinger, 1996), in denen reale oder imaginierte Interaktion mit anderen, zu einer Erhöhung des Weisheitsgehaltes der Antwort führte.

Zur Absicherung der Ergebnisse gegen Einflüsse Sozialer Erwünschtheit, wurden einerseits Laiennominierungen zur Gewinnung des Außenkriteriums, objektive Testverfahren, multimodale Erhebung, genauso wie Testverfahren verwendet, deren Unabhängigkeit von Sozialer Erwünschtheit bereits bewiesen werden konnte (Ardelt, 2003). Als einzige weitere Option hätte ein Maß zur Bestimmung sozialer Erwünschtheit in den Fragebogen integriert werden können. Aufgrund der Diskussion von Kapitel 4.2.3.5 wurde von einer solchen Maß-

nahme abgesehen.

Nicht auszuschließen ist, dass die gefundenen empirischen Ergebnisse der vorliegenden Studie nicht auf andere Stichproben übertragbar sind, also nicht repliziert werden können. Diese mögliche Singularität der Ergebnisse kann durch mögliche verzerrende Störeinflüsse in der vorliegenden Studie verursacht worden sein, die beispielsweise in der Zusammensetzung der Stichprobe oder deren Rekrutierung begründet sind. Wären jene möglichen Störeinflüsse in einer größeren Replikationsstichprobe kontrollierbar, wären daher andere Ergebnisse zu erwarten. Im folgenden Abschnitt wird die Stichprobenziehung in Rekrutierung und Zusammensetzung diskutiert.

8.6.2 Diskussion der Teilnehmergruppen bezüglich der Rekrutierung, Zusammensetzung und Ergebnisgeneralisierung

Die vorliegende Studie wurde für die Äußere Kriteriumsvalidierung bei drei verschiedenen Teilnehmergruppen durchgeführt. Das Kriterium einer Laiennominierung, positiv beurteilte Fremdeingeschätzte Weisheit, und die berufliche Expertise in der Bearbeitung menschlicher Probleme, angezeigt durch den Beruf Psychotherapeut, wurden gewählt, da es für Weisheit keine direkteren Indikatoren zur Äußeren Validierung gibt. Die Ergebnisse der Äußeren Validierung über die höhere Weisheit dieser Teilnehmer gegenüber einer zu den Nominierten nach Alter, Geschlecht, Beruf und Bildungsniveau gemachten Kontrollgruppe widersprechen den Hypothesen. Dies liegt vermutlich an der zufällig überdurchschnittlich Ego-Entwickelten Kontrollgruppe, kann jedoch auch durch fehlende Validität der Testverfahren oder Äußeren Kriterien zurückgeführt werden. Die erteilten Nominierungen könnten aufgrund eines Weisheitsniveaus erfolgt sein, dass auch von den Psychotherapeuten und der Kontrollgruppe erreicht wurde. Alternativ könnte die relativ geringe Stichprobengröße der Teilnehmergruppen die möglicherweise existierenden kleinen Effekte verdeckt haben. Die folgende Diskussion soll die Rolle der Teilnehmer als Ursache fehlender Äußerer Validität der Studie erhellen.

Zunächst die Nominierungen zur Bildung der Gruppe der Nominierten betrachtend, muss darauf hingewiesen werden, dass durch die mehrheitliche Ansprache von studentischen Nominatoren eine besondere Gruppe von möglichen Nominierten begünstigte, nämlich solche, die akademischen Gesellschaftsteilen angehören oder diesen nahe stehen. Da des Weiteren ein Schneeballverfahren über die Internetkontakte einiger Kommilitonen der Diplomandin erfolgte, kann keine Aussage über die Zusammensetzung der abschließend angesprochenen Gruppe der potentiellen Nominatoren erfolgen. In Anbetracht der direkten Ansprache von einer Vielzahl von Studenten, ist deren Anteil an erfolgten Nominierungen jedoch klein.

Die Qualität der Nominierungen kann von unterschiedlichen Motivationen der Nominatoren negativ beeinflusst werden, die empirisch nicht kontrollierbar sind. Wurde die persönliche Aufforderung zur Nominierung eines Bekannten als unangenehm empfunden und daher vermieden, ist dies unabhängig von der Qualifikation der Bekannten für diese Untersuchung. Diese Haltung verringert die Gesamtzahl von Nominierungen geeigneter und nicht geeigneter Personen. Andererseits kann der Wunsch der Unterstützung der Studie dazu führen, dass Personen nominiert werden, die nur bedingt dem Kriterium wahrgenommener Weisheit entsprechen. Diese Haltung vergrößert die Gesamtzahl von Nominierungen geeigneter und nicht geeigneter Personen für die Gruppe der Weisheitsnominierten.

Werden die Studien (Orwoll & Perlmutter, 1990) zu der typischen Nominiertenpersönlichkeit betrachtet, so stimmen die Überzahl der männlichen Nominierten mit einem relativ hohen gesellschaftlichen Status, abstrahiert an deren Bildungsniveau und Beruf, mit den Nominierten in dieser Studie gut überein. Wie bereits in Abschnitt 2.2.2 diskutiert, wird die Wahrscheinlichkeit einer Nominierung auch durch den Bekanntheitsgrad der betreffenden Person, unabhängig von ihrer Weisheit, beeinflusst.

Die Bereitschaft der Nominierten zur Teilnahme wurde nach deren Aussagen durch das Interesse am Thema geleitet. Die Mehrzahl der angesprochenen Nominierten wies jedoch die Zuschreibung von Weisheit im Telefongespräch zurück. Einige Nominierte verweigerten die Teilnahme aus diesem Grund, andere beurteilten die wissenschaftliche Herangehensweise zur Studie der Weisheit als dem Konstrukt widersprechend und nahmen daher nicht teil.

Die nach Alter, Geschlecht, Beruf und Bildungsniveau zu den Nominierten passenden Kontrollpersonen stellten als ihre Teilnahmemotivation ebenso besonderes Interesse am Thema in den Vordergrund, neben dem Wunsch die wissenschaftliche Forschung zu unterstützen. Anhang B.2 präsentiert die kontrollgruppenspezifische Aufklärung über die Teilnehmergruppen. Die Aufklärung über ihre Funktion in der Stichprobe, als Kontrollperson zu einer als weise nominierten Person, könnte die Motivation der Selbstdarstellung beeinflusst haben. Eine Kontrollperson zu sein, legt eine Wettbewerbssituation nahe, in der die Überlegenheit eines Fremden geprüft und bestätigt werden soll. Eine solche Einschätzung der Situation kann Ansporn zur Erbringung einer sehr hohen Leistung sein, was hier dem Ziel einer besonders weisen Selbstdarstellung entspräche. Läge eine derartige Motivation vor, könnte diese die Messung der wahren Weisheitswerte verfälschen, falls es den Personen gelingt, sich wirklich weiser zu präsentieren als sie es sind. Für die leistungsbezogenen Testverfahren und den WUSCT als projektives Maß sind solche Verfälschungen unwahrscheinlicher als für die Persönlichkeitsfragebögen. Da in allen Testverfahren vergleichbare Befunde festgestellt wurden,

ist ein bedeutsamer Einfluss einer solchen Verzerrung nicht erkennbar.

In Anbetracht der Ergebnisse der Ego-Entwicklung ist die Kontrollgruppe in ihrer Persönlichkeitsentwicklung der Normalbevölkerung überlegen. Einschränkend müssen die fehlenden multiplen Korrelationen zwischen den Weisheitstestverfahren für Kontrollpersonen und die von den Nominierten und Psychotherapeuten abweichenden bivariaten Korrelationsmuster so gedeutet werden, dass die einzelnen von diesen Tests abgedeckten Weisheitsfacetten nicht zusammenhängen und dass die Kontrollgruppe Personen vereint, deren Profilverläufe auf den einzelnen Testverfahren sich gegenseitig aufheben. Von dieser Heterogenität sind die Mittelwerte auf den einzelnen Skalen nicht betroffen, weshalb die Hypothesen der Äußeren Validierung nicht bestätigt werden konnten. Bei der Kontrollgruppe ist insgesamt eher von hoher Persönlichkeitsreife anstatt von Weisheit zu sprechen.

Welche Faktoren ursächlich für die hohe Persönlichkeitsreife der Kontrollpersonen sind, ist zu diskutieren. Entweder bilden die Matchingprofile wichtige soziodemographische Kennzeichen von Personen mit hoher Persönlichkeitsreife ab, oder es handelt sich um eine über das besondere Interesse am Thema der Studie selektierte oder aber um eine zufällige Stichprobe. Vermutlich spielen in diesem Falle alle drei Faktoren eine Rolle. Da die Profile typische Nominierungsprofile für Weisheit darstellen, impliziert, dass damit Personen identifiziert werden könnten, die hohe Werte in Weisheitstests erreichen. Bei der Rekrutierung der Teilnehmer für die Kontrollgruppe wurde von einigen angesprochenen Personen verkündet, sich bereits seit einiger Zeit mit Themen der Weisheit, ob philosophischer oder religiöser Natur, auseinander zu setzen. Es bleibt darüber hinaus wahrscheinlich, dass durch Zufall ähnlich wie Personen in allen drei Teilnehmergruppen zu finden sind, die auf persönlicher Ebene ein vergleichbares Niveau erreicht haben.

Eine Absicherung dieser Möglichkeit wäre nur dann möglich gewesen, wenn ein und der derselbe vorab festgelegte Bekanntenkreis für Nominierte und Kontrollpersonen die Einteilung der Personen auf beide Gruppen vorgenommen hätte. Leider hätte dies eine Vorabselektion der Nominierten und der Kontrollpersonen bedeutet. Besonders aufgrund der ethischen Einwände gegen eine Fremdeinschätzung fehlender Weisheit muss generell von einem derartigen Verfahren abgesehen werden. Im vorliegenden Fall ist nicht prüfbar, ob eine Person des Bekanntenkreises der Kontrollpersonen diese für ihre Weisheit nominiert hätte. Die Trennung in diese Teilnehmergruppen erfolgte somit lediglich daher, dass die Existenz eines Merkmals, Fremdeingeschätzte Weisheit, für eine der beiden Gruppen geprüft und belegt wurde, während die andere, die Kontrollgruppe, nicht für dieses Kriterium geprüft werden konnte. Die Ergebnisse der Nominierten auf die Population aller Personen zu generalisieren,

die eine positiv fremdeingeschätzte Weisheit von ihren Bekannten attestiert bekommen, könnte somit einige oder sogar alle Teilnehmer der vorliegenden Kontrollgruppe umschließen. Entsprechend erlaubt die Generalisierung über die Eigenschaften der Matchingprofile auf die Population von Personen mit identischen Kriterienausprägungen des Alters, Geschlechts, Berufes und Bildungsniveaus wie die Nominierten der Studie, ebenfalls keine Trennung zwischen Nominierten und Kontrollpersonen. Da sich jedoch bereits vorgestellte Befunde differentieller Validität dieser Kriterien zu den Weisheitstests offenbarten, müssen zumindest im Durchschnitt die Teilnehmer der Nominierengruppe eher die Breite der Facetten der Weisheit repräsentieren als dies bei den Kontrollpersonen der Fall ist.

Die Rekrutierung von klinischen Psychotherapeuten für die Stichprobe erfolgte über eine Liste kassenärztlich zugelassener und niedergelassener Psychotherapeuten aus der Vorderpfalz und dem Rhein-Neckarraum. Es wurden rein psychotherapeutisch arbeitende Probanden geworben sowie jene mit weiteren ärztlichen Fachgebieten. Auch die Variable Berufserfahrung deckt das volle Spektrum ab. Die aus dieser Zusammensetzung resultierenden Expertiseunterschiede waren in Studien von Baltes et al (1995) nicht bedeutsam, weshalb auf eine diesbezügliche Beschränkung der Teilnahmekriterien verzichtet wurde. Die Ergebnisse der Studie können daher auf den Berufszweig der Psychotherapeuten generalisiert werden.

Aufgrund der fehlenden Überlegenheit der Psychotherapeuten gegenüber der Kontrollgruppe, muss die Existenz eines von Baltes et al. (1995) diskutierten Artefaktes durch deren psychologische Vorbildung, welche ihre Testleistung positiv beeinflussen kann, ausgeschlossen werden. Die Vorteile von Psychotherapeuten sollten insbesondere in der Fähigkeit zur Introspektion liegen, welche in höheren Werten der Tests TWR und AWR resultieren könnten, die jedoch nicht nachzuweisen waren.

8.6.3 Diskussion der Weisheitstestverfahren

Die vorliegende Studie hat zur Aufgabe, die Weisheitstests 3d-WS, SAWS, PWS, TWR und AWR zu validieren. Die Begrenzungen dieser Tests werden in diesem Abschnitt nacheinander diskutiert, während eine spezielle Diskussion der Testverfahren zur Konvergenz und Diskriminanten Validierung von nachgeordneter Relevanz ist und daher entfällt.

Der Gesamtwert des 3d-WS wird in der Originalstudie von Ardel (2003) über die Bildung eines latenten Faktors ermittelt. Hierbei konnte der Ladungskoeffizient der Reflektiven Dimension frei variieren, während dieser für die beiden anderen Dimensionen gleichgesetzt wurde. Da es keine Kausalhypothesen zu prüfen galt, konnte auf latente Strukturanalysen verzichtet werden und der Faktor 3d-WS konnte ohne Festlegung von Ladungskoeffizienten über

eine Hauptachsenanalyse gebildet werden. Trotzdem entsprechen die Faktorladungen der drei Dimensionen dem oben genannten Muster, siehe Kapitel 6.7.3, wenn auch in sehr niedriger Größenordnung. Daher beurteilt das Kaiser-Meyer-Olkin Kriterium die Daten zur Analyse des 3d-WS Faktors lediglich als mäßig (Bühner, 2006). Gerade da einzig hohe Ausprägungen auf allen Dimensionen Weisheit indiziert, wäre der Anteil geteilter Varianz bei guter Kriteriumsvalidität höher zu erwarten. Die Nähe zu Adaptivität muss weiter erforscht werden. Da Ardelts (1997, 2000a) 3d-WS hoch mit Lebenszufriedenheit korreliert und diese ein Indikator der Adaptivität ist, ist dieser Befund jedoch teilweise nachvollziehbar. Möglicherweise sind andere Aspekte der Adaptivität adäquatere Diskriminante Konstrukte, als Umweltkontrolle, besonders auf diesem Schwierigkeitsniveau. Das hier verwendete Instrument UK erfasst, ob und wie sehr die Personen die alltäglichen Anforderungen meistern. Hier könnte die Verwendung eines Inventars sinnvoll sein, das Sozialen Erfolg auf einem hohen Niveau operationalisiert.

Die fünf Dimensionen des SAWS werden als Summenscore aufaddiert zur Weisheitsmessung herangezogen. Die Erhebung und Analyse der Daten verlief problemlos. Die hier erstellte Übersetzung der Skala ins Deutsche wurde jedoch nicht speziell auf ihre Übereinstimmung mit der amerikanischen Faktorenstruktur getestet. Diese Analyse sollte an einer größeren Stichprobe erfolgen. In der Studie von Webster (2003) wurde mit einem Cut-off-Wert gearbeitet. Wird dieser auf die neue Itemzahl übertragen und in der vorliegenden Stichprobe betrachtet, so erreichen lediglich etwa 3% der Teilnehmer dieses Niveau.

Die Bildung des gemeinsamen Faktors HS der Weisheit nach Helson und Srivastava (2001, 2002) konnte in dieser Studie nicht erfolgen, da es sich aufgrund der niedrigen Interkorrelation der drei Verfahren nach dem Kaiser-Meyer-Olkin-Kriterium nicht um eine analytisch wertige Faktorenbildung gehandelt hätte. Bei exemplarischer Bildung des Faktors mit unterschiedlichen AWR-Varianten zeigte sich ein konsistentes Bild. Einzig der TWR wurde wirklich von dem latenten Weisheitsfaktor repräsentiert. Aus diesem Grund wurde in der vorliegenden Studie auf die Betrachtung des latenten Faktors über alle drei Testverfahren verzichtet. Die Ergebnisse des TWR können näherungsweise für die Beziehungen des latenten Faktors über die drei Testverfahren bezeichnet werden. Die vorliegende Arbeit konnte somit die Studienergebnisse von Helson und Srivastava (2001, 2002) zu einem latenten Faktor über die Weisheitskomponenten der Transzendenten, Allgemeinen und Praktischen Weisheit nicht replizieren. Im Folgenden werden somit die drei Tests, die Praktische Weisheitsskala, und die beiden Ratings der Transzendenten und Allgemeinem Weisheit einzeln diskutiert.

Die PWS zeigt die geringste Interne Konsistenz aller Testverfahren und gleichzeitig

die meisten fehlenden Angaben. Aufgrund ihres Entwicklungsprozesses sind zumindest die niedrigen Werte der Internen Konsistenz verständlich. Die hohe Anzahl fehlender Daten ist auf das dichotome Antwortformat zurückzuführen, das nach Lienert und Raatz (1998) zu besonderen Widerständen in der Bearbeitung führt, da es den Teilnehmern erschwert wird, sich differenziert darstellen zu können.

Die PWS nimmt darüber hinaus als analytisch ausgerichtetes Instrument eine Sonderstellung gegenüber den vier anderen Testverfahren ein. Möglicherweise ist somit der niedrige multiple Validitätskoeffizient erklärbar. Die Ergebnisse der Konvergenten Validierung für diese Skala erreichten lediglich uneindeutiges Niveau. Somit ist PWS nicht sicher im nomologischen Netzwerk der bereits etablierten Testverfahren zu verorten. Es muss daher die Empfehlung ergehen, Praktische Weisheit über alternative Erhebungsverfahren wie Rating (PWR) zu erheben (Helson & Srivastava, 2001, 2002). Gerade diese Skala wird im Hinblick auf Verzerrungstendenzen im Folgenden diskutiert.

Da in der ursprünglichen Längsschnittstudie die Vollversion des ACL (Gough & Heilbrun, 1983) erhoben wurde, ist laut Wink & Helson (1997) nicht damit zu rechnen, dass eine Verzerrung der 17 der PWS-Items durch soziale Erwünschtheit auftrat. Zu dem späteren Erhebungszeitpunkt wurde das Konstrukt der Praktischen Weisheit über ein Interviewverfahren, dem PWR, bestimmt, das allerdings aufgrund fehlender zeitlicher und finanzieller Ressourcen nicht in der vorliegenden Studie durchgeführt werden konnte.

Um den Einfluss von verzerrtem Antwortverhalten bei lediglicher Präsentation der 17 Items gering zu halten, wurde das Einstiegsitem ehrlich vor die PWS gesetzt. Mit ehrlich soll das wahrheitsgemäße Beantworten von Items ins Bewusstsein der Probanden gerückt werden. Ob diese Maßnahme Wirkung gezeigt hat, ist nicht überprüfbar. Möglich wäre sogar ein gegenteiliger Effekt, so dass gerade durch „ehrllich“ als erwünschte Verhaltensweise, die folgenden Items ebenso in erwünschter Ausprägungsrichtung beantwortet wurden.

Das Item reif wurde bei der Vortestung des Fragebogens von zwei der Probanden aus Gründen negativ wahrgenommenen Konnotation des Begriffes abgelehnt. Es wurde angenommen, dass das Alter der Probanden von 32 und 35 für diese Einschätzung ursächlich ist und ältere Teilnehmergruppen den Begriff nicht als negativ belegt wahrnehmen.

Das Item weise der PWS sollte der Definition nach die größte Indikation für das zu validierende Konstrukt aufweisen. Jedoch kann eine Selbstzuschreibung dieser Eigenschaft der gesellschaftlichen Norm widersprechend wahrgenommen werden, da wahre Weisheit mit Bescheidenheit und der Einsicht, nicht weise zu sein, einhergeht. Eine weitere Hürde für die Selbstzuschreibung von Weisheit kann das dichotome Antwortformat sein, welches keinen

Spielraum für eine individuelle Einschätzung erlaubt (Lienert & Raatz, 1998). In Konsequenz wurde das Item weise von knapp 33% mit ja beantwortet, es fehlen jedoch die Angaben von 12 Teilnehmern. Die Gruppen unterscheiden sich in der Beantwortung dieses Items nicht.

Der TWR fragt nach eigenen Erfahrungen, die zu einer Steigerungen der persönlichen Weisheit beitragen. Den Richtlinien der Studie von Helson und Srivastava dürften nur jene Antworten ausgewertet werden, die mindestens 70 Worte umfassen. Da in der vorliegenden Studie der Mittelwert der Antwortlänge jedoch unter diesem Niveau liegt, was den Ausschluss der Mehrzahl der Teilnehmer zur Folge hätte, musste von diesem Selektionskriterium abgesehen werden. Dieses Kriterium stammt aus den Interviewstudien der Allgemeinen Weisheit. Von Mickler (2004) wird berichtet, bei der Validierung von Selbstbezogener Weisheit nur solche Antworten ausgewertet zu haben, die über 200 Worte aufwiesen. Einzig bei einem Teilnehmer hätte der Wert unter diesem Niveau gelegen. Mickler argumentiert, dass solche kurze Antworten einer Verweigerung der Bearbeitung der Fragestellung gleichkämen.

Möglicherweise sind derartig kurze Antworten in der Mills Studie nicht aufgetreten, da die Teilnehmer dieser Längsschnittstudie mit der Bearbeitung solcher Aufgaben vertrauter sind und der Studie gegenüber höheres Commitment aufweisen. Werden Studienergebnisse von Bluck und Glück (2004) konsultiert, wäre zu erwarten, dass jeder Teilnehmer mindestens eine Situation erinnern kann, in der er nach persönlichen Kriterien weise reagierte und zu diesen Ereignissen eine detaillierte Beschreibungen geben könnte. Jedoch kann die fehlende Übereinstimmung mit diesen Studien auch auf folgende Kriterien zurückgeführt werden. Neben einer gewissen Verpflichtung durch die monetäre Vergütung der Teilnahme erfolgte die Erhebung im Interview. In einer Interviewsituation liegt ein höherer Aufforderungscharakter durch Ermunterungen des Interviewers vor als bei einer schriftlichen Bearbeitung.

Im Vortest erreichten die Antworten im TWR im Durchschnitt 164 Worte, im AWR 116 Worte, wobei keine Antwort unter 70 Worten lag. Aufgrund der erreichten Mindestlänge, wurde die Problematik zu kurzer Antwortlängen für die Tests TWR und AWR nicht abgesehen.

Während des Rücklaufs der Fragebögen wurde erkannt, dass es einige Teilnehmer gibt, die eine Beantwortung vollkommen verweigern. Im Zuge der mündlichen Erhebung erklärten sich jene Teilnehmer auf Nachfrage in folgender Weise. Einige Probanden würden sich selbst nie als weise bezeichnen, weshalb sie konsequent nicht antworten. Andere Probanden sehen den weisen Akt darin, nicht zu antworten. Erikson (1964, nach Hartmann 2005) gibt ebenfalls zu bedenken, dass kein wirklich weiser Mensch jemals von sich sagen würde, er sei weise. Die Beantwortung dieser Frage ist für einige zu persönlich und wird daher vermie-

den. Von anderen wird bei genauerer Evaluation einer erinnerten Situation mit Weisheitsbezug beschlossen, dass diese nicht den eigenen Ansprüchen an Weisheit genügt, und möglicherweise den wissenschaftlichen Maßstäben nicht gerecht wird. Eine weitere Erklärung für das Fehlen einer Antwort kann aus der Absicht entstehen, eine einzelne Erfahrung nicht über zu betonen und andere Erfahrungen dagegen nicht erwähnen zu können. Das Fehlen oder die Verwendung von pauschalen Antworten kann auch aus mangelnder Bereitschaft zur Selbstoffenbarung zu Stande kommen, obwohl die Person möglicherweise eine geeignete Situation berichten könnte. Bei anderen Probanden liegt eine persönlich sehr genaue Vorstellung von Weisheit vor, die mit dem restlichen Fragebogen nicht übereinstimmt, eine Verweigerung der Bearbeitung des TWR wurde als Protest gegen die restlichen Fragebogen genutzt. Es mussten 12 Probanden daher bei der Analyse des TWR ausgeschlossen werden, da sie entweder gar nicht antworteten, mit nein, Erfahrungen oder Ich halte mich nicht für weise antworteten. Dies betraf 4 Nominierte, 5 Kontrollpersonen und 3 Psychotherapeuten.

Das Scoring des TWR anhand des Bewertungsschemas zeigte, dass kaum Antworten auf den beiden niedrigsten Bewertungsniveaus gescort wurden. Die Kriterien der „Verkümmerten Persönlichkeit“ konnten nicht in den gegebenen Antworten entdeckt werden. Antworten mit oberflächlicher Bedeutungsebene des Begriffes Weisheit wurden dagegen mit einer niedrigen Wertung bedacht.

Zur Verbesserung der schriftlichen Anwendung dieses Tests müsste die Instruktion dazu auffordern, die eigenen Gedanken möglichst präzise aufzuschreiben. Da vor allem der transzendente Aspekt der Antworten interessiert, könnte es möglicherweise für schriftliche Erhebungen sinnvoll sein, Übertragungen dieses Einsichtsbeispiels in andere Bereiche zu liefern. Das Auswertungsmanual muss vor allem für das Scoring von niedrigen Bewertungseinheiten detaillierter ausgestaltet werden. Insbesondere fehlt dem Manual eine Bewertungsanweisung für begründete Antwortverweigerungen oder ausweichende Antworten.

Die Begrenzungen des AWR liegen in ähnlichen Bereichen wie jene des TWR. Die aus der Interviewmessung von Weisheit adaptierte Fragestellung der Allgemeinen Weisheit, bleibt undifferenziert. Stellen sie sich vor, sie erhalten einen Anruf eines guten Freundes. Dieser erzählte ihnen er könne nicht so weiter machen und er wolle sich das Leben nehmen. Was würden Sie denken oder tun? Dies verleitet zu undifferenzierten Antworten.

Inbesondere aufgrund der schriftlichen Erhebung zeigen sich Schwächen des Instrumentes. Da die Antwortenlängen im AWR in der Mehrheit ebenso wenig ausreichendes Niveau erreichten, erzeugt die adäquate Bewertung dieser Antworten auch hier Probleme. Da das Scoringmanual der amerikanischen Studie von Helson und Srivastava zum AWR führt

einzig die oberen Skaleneinheiten der vier Kategorien aus. Beispiele für ein Scoring fehlen völlig. Es ergab sich daher der Bedarf, das Bewertungsmanual auszudifferenzieren. Dies erfolgte mit besonderer Beachtung der Intervallskalierung der Gesamtskala (Anhang E.2).

Das Scoring blieb an Folgenden Stellen problematisch. Eine Antwort kann emotionales Verständnis ausdrücken ohne dies zu verbalisieren. Solche impliziten Botschaften wurden nicht bewertet, da nicht mit Sicherheit bei anderen Antworten auf ein Fehlen der nicht explizit verbalisierten Emotionalen Komponente geschlossen werden konnte. Die Kognitive Differenziertheit einiger Personen zeigte sich eher in der formalen Erarbeitung von Lösungsvorschlägen als in deren exemplarischer Diskussion und Lösung. Solche Antworten wurden als indikativ für das Konstrukt Kognitive Differenziertheit gewertet und daher entsprechend gescored. Ein strengeres Vorgehen erging bei der prozeduralen Unterskala. Eine Antwort, die die Notwendigkeit der Berücksichtigung fremder Hilfsquellen, wie beispielsweise Notarzt, Rettungswagen, Polizei, nicht berücksichtigten, klammern die unbestimmte Brisanz des Anrufes aus. Dies entspricht einer niedrigeren Ausprägung in Handlungswissen, auch wenn ein ansonsten elaboriertes Hilfsangebot erfolgt.

Die Antworten der Probanden legten den Schwerpunkt auf Handlungen in der betreffenden Situation, während die kognitive Auseinandersetzung mit der Situation wenig Beachtung findet. Da sich jedoch drei der Kategorien mit der kognitiven Auseinandersetzung befassten, Kognitive Differenziertheit, Moralische Komplexität und Emotionales Verständnis, zeigten diese drei starke Bodeneffekte während die Prozedurale Wissen Kategorie weniger schief verteilt ist. Dies liegt in der schriftlichen Erhebung begründet. Im Interview werden sämtliche während 20 Minuten geäußerten Gedanken protokolliert und bewertet. Beim Niederschreiben ist damit zu rechnen, dass vergleichbare Gedankengänge in den Teilnehmern vorgehen, jedoch werden diese nicht komplett niedergeschrieben. Es ist demnach wahrscheinlicher, vage Rahmenbedingungen, Handlungsmöglichkeiten und Schlussfolgerungen zu verbalisieren, als diese aus Gründen der Konsistenz niederzuschreiben.

Als Fazit für TWR und AWR muss gefolgert werden, dass eine Erweiterung der Instruktion zur Steigerung der Antwortlänge für TWR und AWR entwickelt und bei der Fragestellung präsentiert werden sollte. Beispielhaft könnte folgende Instruktion verwendet werden: Bitte geben Sie ihre Gedanken möglichst detailliert wieder. Falls dies nicht möglich ist, sollte von einer schriftlichen Administration dieser Tests abgesehen werden.

8.6.4 Begrenzungen des Methodischen Vorgehens

Das methodische Vorgehen der Studie kritisch zu beleuchten ist sinnvoll, um die Einschät-

zung der empirischen Ergebnisse vorzunehmen. In der vorliegenden Studie sind daher vor allem folgende Diskussionspunkte angezeigt: Eignung der Stichprobengröße, Verwendung von abhängigen Analysemethoden für die Teilnehmergruppen Nominierte, Kontrollgruppe und Psychotherapeuten, und die Bildung der beiden Faktoren für 3d-WS, sowie die drei Testverfahren TWR, AWR und PWS.

Dass aufgrund der geringen Stichprobengröße von insgesamt 114 Teilnehmern möglicherweise Mittelwertsunterschiede oder Korrelationen nicht in ihrer Größenordnung voneinander trennbar sind, ist nicht auszuschließen. Insbesondere für die Fragestellungen der Konstruktvalidierung, für die Korrelationskoeffizienten von dem kritischen Niveau von .30 unterscheidbar sein sollten, sind größere Stichproben vorteilig. Für diese Studie mussten Korrelationen, vereinfacht gesagt, entweder weniger als halb oder doppelt so groß wie die kritische Korrelation von .30 sein, um sich von dieser signifikant unterscheiden zu lassen.

In der vorliegenden Studie wurden zur Äußeren Validierung abhängige Gruppenstatistiken für die Analyse von Unterschieden zwischen Nominierten, Kontrollgruppe und Psychotherapeuten verwendet. Da es sich angesichts der empirischen Ergebnisse zur Allgemeinen Weisheit bei den Matchingkriterien Alter und Geschlecht, wofür alle drei Teilnehmergruppen parallelisiert sind, nicht um mit dem Zielkonstrukt Weisheit zusammenhängende Variablen handelt, hätte eine Analyse für unabhängige Stichproben durchgeführt werden können. Genauso liegt die Befundlage für die Kriterien der Bildung. Sollte es keine Korrelation zwischen dem Kriterium und den zu testenden Variablen geben, so wäre eine Verwendung eines Designs für abhängige Stichproben nicht durch Powervorteile begünstigt. Dass an den gegebenen Stellen trotzdem Analysen für abhängige Stichproben erfolgten, stellt eine konservative Herangehensweise dar. Die ungeprüfte Übertragung der Ergebnisse von einem Interviewverfahren auf die Fragebögen dieser Studie, die neben den methodischen Unterschied unterschiedliche Definitionen des Konstruktes Weisheit vertreten, welche in Konsequenz zur Verwendung von unabhängigen Analysemethoden geführt hätte, wäre unwissenschaftlich.

Die Verwendung abhängiger Gruppen unter Inspektion der Korrelation zwischen den Messwertreihen erlaubt festzustellen, ob dieses Vorgehen sinnvoll war. Da keine bedeutsamen positiven oder negativen Korrelationen vorlagen, wurde zur Absicherung der Beibehaltung der Nullhypothese nach dem abhängigen Design eine Analyse über unabhängige Gruppen nachgeschoben. Das elegantere Vorgehen wäre gewesen, einzig die Korrelation zu inspizieren und darauf die Prüfgrößen für unabhängige Tests zu berichten. Da jedoch beide Analysemethoden, ob für unabhängige oder abhängige Gruppen, zu insignifikanten Ergebnissen führten, wurde auf diese Form der Ergebnisdokumentation verzichtet.

Zur Auswertung der dreidimensionalen Weisheitsskala sowie zur Bildung eines latenten Faktors HS für die drei Tests der Praktischen, Transzendenten und Allgemeinen Weisheit muss jeweils eine Faktorenanalyse erfolgen. In den Originalstudien wurden diese Analysen über latente Strukturgleichungsmodelle mit LISREL (Ardelt, 2003) oder SAS Proc Mixed (Helson & Srivastava, 2001, 2002) bestimmt, da Kausalhypothesen und Wachstumskurven zu modellieren und analysieren waren. Da dies in der vorliegenden Studie nicht zutrifft, wurde auf eine Verwendung einer solchen Auswertungsmethode verzichtet. Vorteilhaft bei jener Form der Analyse ist, dass lediglich drei manifeste Indikatoren genügen, um eine latente Variable zu bestimmen, wozu die drei Unterskalen des 3d-WS und die drei Tests TWR, AWR und PWS ausreichen.

Alternativ wurde die Extraktion eines Faktors über SPSS durchgeführt. Mittels Hauptachsenanalyse wurde der Faktor der höchsten Varianzaufklärung extrahiert, die im gegebenen Falle die gleichen Ergebnisse wie eine Maximum-Likelihood Schätzung erzielt. Hier findet die Hauptkomponentenanalyse keine Anwendung, da es sich nicht primär um eine Varianzreduktion der zu Grunde liegenden Items und Tests handelt, sondern inhaltliche Annahmen über eine ursächliche Beziehung der gemeinsamen Varianz vorliegen. Ebenso deckt sich die Annahme der Hauptachsenanalyse eines nichtgeteilten, eigenständigen Varianzanteils der einzelnen Dimensionen oder Testverfahren mit der theoretischen Grundlage dieser Testverfahren. Eine Hauptkomponentenanalyse würde der Existenz einer solchen eigenständigen Varianzquelle keine Rechnung tragen.

Da das Kaiser-Meyer-Olkin-Kriterium für den Faktor HS keine Faktorenextraktion empfahl (Bühner, 2006) und sich der latente Faktor über TWR, AWR und PWS lediglich aus dem TWR signifikant determinierte, wurde entschieden, auf die Bildung dieses Faktors zu verzichten. Die fehlende Übereinstimmung bewies sich ebenso bei alternativer Modellierung des Faktors HS durch AMOS. Das Modell des Faktors HS war nicht signifikant und TWR offenbarte eine hohe negative Fehlervarianz.

Einschränkend muss anerkannt werden, dass die Voraussetzungen für ein faktorenanalytisches Vorgehen aufgrund der Itemzahl 3d-WS nicht erfüllt waren. So lautet ein Vorschlag von Guadagnoli und Velicer (1998, nach Bortz, 1999) mindestens 10 oder mehr Variablen für einen Faktor in der Planungsphase zu veranschlagen, um mit einem n von 150 zu rechnen. Für jeden bedeutsamen Faktor sollen mindestens 4 Ladungen über .60 aufzuweisen sein, um auf Einschränkungen des Stichprobenumfangs verzichten zu können. Diesem entspräche es, wenn 10 bis 12 Variablen mit einer Ladung von .40 einen Faktor bilden. Dies sei nach Stevens (2002, nach Bortz, 1999) generell die untere Grenze für Faktorenladungen, die bei der Inter-

pretation eines Faktors berücksichtigt werden sollten. Ansonsten, bei vielen sehr geringen Faktorladungen seien mindestens 300 Teilnehmer zu analysieren, um die Notwendigkeit von Ergebnisreplikationen zu umgehen. Keine dieser Voraussetzungen wird von 3d-WS erfüllt. Bei alternativer Verwendung der einzelnen 39 Items des 3d-WS zeigen sich dagegen gleiche Faktorenwerte und Gütekriterien wie bei Verwendung der Unterskalen, obwohl die Itemzahl nun eher den Voraussetzungen von Guadagnoli und Velicer entsprachen.

8.7 Ausblick

Da das psychologische Forschungsfeld der Weisheit bisher zu wenigen empirischen Studien führte, ist eine Vielzahl von Fragestellungen bisher unbearbeitet. Im folgenden Ausblick auf weitere interessante Fragestellungen ergeht daher eine Beschränkung auf den Kontext der Validierung von Weisheitstests. Ein Versuch, einen Überblick über anders gelagerte Fragestellungen zu geben, sprengt den Rahmen dieser Validierungsstudie.

Die Ergebnisse dieser Studie können als erste Validierungsstudie über mehrere Weisheitstestverfahren bei einer Stichprobe von 114 Teilnehmern Belege über die Zusammenhangsstruktur der Weisheitstests und deren Verortung in nomologischen Netzwerk von Konstrukten der Persönlichkeitsreife und kognitiver Kompetenz auf empirischer Ebene liefern. Die Schlüsse aus den empirischen Daten weisen auf Trennungen zwischen den Weisheitstestverfahren hin, die auf theoretischer Ebene die Trennung zwischen Synthetischer und Analytischer Ausrichtung und zwischen Produktmessungen und Persönlichkeitsbeschreibungen einer weisen Person identifiziert werden können. Insgesamt wird deutlich, dass keiner der kurzen schriftlichen Testverfahren das komplexe Konstrukt Weisheit umfassen kann. Die varianzstärksten Testverfahren der Studie TWR und 3d-WS spiegeln sehr unterschiedliche Konzepte der Weisheit und der adäquaten Messung dieser. Es bleibt also in weiteren Studien zur Inhaltsvalidität zu klären, wie entweder ein übergreifendes Konstrukt der Weisheit definiert und operationalisiert werden soll, oder ob eine Aufteilung verschiedener Weisheitskonzepte die Lösung der Wahl zur adäquaten Beschreibung dieses psychologischen Phänomens sind.

Falls der erste Weg weiter beschritten wird, bietet es sich an, eine vergleichende Validierung dieser Fragebogenverfahren zu den Interviewverfahren, Allgemeine Weisheit und Selbstbezogene Weisheit, anzustreben. Eine solche Herangehensweise offenbart einerseits die methodischen Einflüsse und Effekte der Erhebung, und andererseits die inhaltliche Nähe der Definitionen der Weisheit über deren Operationalisierungen. Eine weitere Option der Kriteriumsvalidierung bietet der empirische Vergleich von Testverfahren mit der Weisheitsentwick-

lungsskala (Brown & Grene, 2006). Da das Modell dieses Tests über Grounded Theory aus einer riesigen Collegestichprobe gewonnen wurde, stellt diese der Weisheit der Großen Zahl folgend das Laienmodell der Weisheit dar, zumindest jenes von Collegestudenten.

Hinweise auf die Form eines inhaltsvaliden Weisheitstests, der die unterschiedlichen Aspekte der hier verwendeten Testverfahren vereint, können aus den Ergebnissen dieser Studie gezogen werden. Es erscheint bei schriftlicher Erfassung von Weisheit notwendig, deren vielfältige Facetten im Instrument abzubilden. Anhaltspunkte für Zusammenhänge einzelner Facetten ergeben sich aus den Korrelationsmustern der Nominierten und dem Typus der Suchenden (Helson & Srivastava, 2001).

Die Emotionale Komponente der Weisheit betrachtend, wird ersichtlich, dass sowohl Emotionsregulationskompetenz als auch tiefe emotionale Wertschätzung gegenüber den Menschen gemeinsam auftreten müssen, um auf Weisheit hinzudeuten. Beide Kompetenzen können durch hohe Persönlichkeitsentwicklung, durch die Bewältigung eigener Konflikte, erreicht werden. Diese emotionale Reife ist zumindest teilweise durch Introspektion bedingt, die in den Weisheitstestverfahren (3d-WS, SAWS, PWS, TWR) in Form von Reflektion beschrieben wird, die eigenen Entwicklungsschritte ermöglicht. Dieser Zusammenhang wird beispielhaft für den Typus der Suchenden deutlich. Hohe Weisheit, die sich von einer durchschnittlichen Entwicklung unterscheidet, bleibt jedoch nicht auf episodische Erfahrungen verhaftet, sondern erfordert die Transzendenz dieser Einsichten.

Um an dieser Stelle exemplarisch auf den TWR einzugehen, dessen Ziel offenbar die Messung genau dieses Aspektes ist, müssten folgende Veränderungen erwogen werden. Neben einer instruktionellen Aufforderung die eigenen Gedanken möglichst detailliert wiederzugeben, müsste die Instruktion klarer Transzendenz der geschilderten Einsicht über verschiedene Lebensbereiche und Episoden fordern. Darüber hinaus fehlt es der TWR-Bewertungsskala an Codierungsmöglichkeiten von Personen, die die Fragen aus unterschiedlichen Gründen nicht beantworten wollen.

Wie diese Einsichten und die emotionale Reife in Handlungen angewendet werden und somit zu messen sind, bedarf weiterer theoretischer Vorarbeit. Möglicherweise stellt die Praktische Weisheitsskala einen Versuch dar, diese Komponente weiser Handlungen sozusagen in ihrer Entstehung auf kognitiver Ebene zu operationalisieren. Die übrigen Testverfahren integrieren ebenso kognitiv anmutende Konstrukte wie Kognitive Differenziertheit und Prozedurales Wissen (beide AWR), eine kognitive Unterskala im (3d-WS) oder Konstrukte wie Offenheit und Humor (SAWS). Dass verbale Intelligenz für diesen Zwischenschritt zur Anwendung nicht in Frage kommt, belegen die Studienergebnisse. Auch fluide Intelligenzin-

strumente konnten in bisherigen Studien zur Allgemeinen Weisheit ihre Bedeutung im Zuge der Weisheitsmessung nicht belegen (Staudinger et al., 1997, 1998). Intelligenz muss hier eher im Rahmen einer Befähigung gedeutet werden, die die Form ihrer Anwendung nicht determiniert. In der Bearbeitung dieser hier abzuleitenden Fragestellungen, des kognitiven Korrelates von Weisheit und der Erforschung weiser Handlungen liegen die Anforderungen an eine Weiterentwicklung der Validierung in der Weisheitsforschung. Das Konzept der kognitiven Denkstile von Sternberg (1988, 1997) sollte in diesem Zusammenhang weiter erforscht werden, siehe Kapitel 4.2.2.3. Kognitive Denkstile schaffen eine Verbindung zwischen kognitiven Prozessen und deren inhaltlicher Orientierung bei Erkenntnis- und Bewertungsprozessen. Die Fragestellung, welche Bewertungsprozesse Entscheidungen derart beeinflussen, dass weise Entscheidungen und Handlungen wahrscheinlicher werden, ist weiter zu erforschen.

Zu diesem Zwecke können Studien anhand von Typologien wie Helsons und Srivastavas Konzept (2001) hilfreich sein. Zumindest die Bestimmung von geeigneten Teilnehmern einer Studie fällt aus praktischer Sicht über ein derartiges Verfahren leichter als über Laien-nominierungen von Weisheit. Die Bedeutung von Nominierungsverfahren ist jedoch ungebrochen, wenn speziell Personen mit hohen Weisheitswerten in einer Studie analysiert werden sollen. Die ersatzweise Rekrutierung von Psychotherapeuten, oder einem bestimmten Profil entsprechende Personen führen, wie in dieser Arbeit, vor allem auf korrelativer Ebene nicht zu vergleichbaren Ergebnissen.

Wird Weisheit hingegen in unterschiedliche Forschungsfelder aufgeteilt, so empfiehlt sich, die oben vorgenommenen Trennungen beizubehalten: die Emotionale Dimension, die Reflektive Dimension und die Forschungsfelder der Kognitiven Korrelate der Weisheit. Darüber hinaus erscheint es ebenso für Fragestellungen der Äußeren Validierung unablässig, weise Handlungen und Werteorientierung zu erforschen. Hierbei kann auf Arbeiten von Sternberg (2004a, 2005) zurückgegriffen werden, die einerseits pädagogische Elemente zur Förderung von Weisheit enthält und andererseits das Konzept der Foolishness erforscht, welche das Fehlen von Weisheit im Sinne von typischen Fehlschlüssen systematisiert. Diese Forschungsfragen wären für die zukünftige Weisheitsforschung wichtige und vielversprechende Vorhaben.

9 Zusammenfassung

Diese Diplomarbeit dient der Konstrukt- und Kriteriumsvalidierung folgender schriftlicher Weisheitstestverfahren: 3d-WS (Ardelt, 2003, 2006), SAWS (Webster, 2003, 2007) PWS (Wink & Helson, 1997), TWR (Helson & Srivastava, 2001, 2002) und AWR (Helson & Srivastava, 2002). Die querschnittliche Validierungsstudie besteht aus je 38 Personen in drei ad hoc Stichproben. Die Probanden nahmen ohne Entlohnung an der schriftlichen und telefonischen Erhebung teil.

Kriterien der Stichprobenziehung

Da es keine eindeutigen objektiven Indikatoren von Weisheit gibt, wurde auf folgende zwei Kriterien zur Bestimmung von Teilnehmergruppen zurückgegriffen.

1. Über die Fragestellung - Kennen Sie eine Person in Ihrem Bekanntenkreis, die Sie als weise bezeichnen würden? - eine so genannte Laiennominierung, wurden Personen für die Studie vorgeschlagen. Zu diesen Nominierten wurden Kontrollpersonen nach Alter, Geschlecht, Beruf und Bildungsniveau für die Kontrollgruppe gematcht.

2. Über den Beruf Psychotherapeut, da die Expertise von Psychotherapeuten bereits in anderen Validierungsstudien zu Weisheit deutlich wurde (Baltes et al., 1995; Helson & Srivastava, 2002). Die Psychotherapeutengruppe wurde nach Alter und Geschlecht zu den anderen Gruppen gematcht.

Kriteriumsvalidierung über die äußere Validität

Die angenommene Dominanz der Nominierten und der Psychotherapeuten gegenüber der Kontrollgruppe bei Analyse der T-Tests für abhängige Gruppen konnte nicht bestätigt werden. Die Interkorrelationen der einzelnen Testverfahren in den Gruppen der Nominierten und Psychotherapeuten sind jedoch wesentlich höher als in der Kontrollgruppe. Diese differentielle Kriteriumsvalidität stützt die theoretische Sichtweise, dass Weisheit ein breites Konstrukt unterschiedlicher Facetten ist, die diese durch ihr gemeinsames Auftreten indizieren.

Kriteriumsvalidierung über die innere Validität

Zur Kategorisierung der Weisheitstests dienen deren Modalität (Produkt oder Persönlichkeit), Orientierung (analytisch oder synthetisch), und dem Gegenstand (Selbstbezogen oder Allgemein). Übereinstimmend mit dem eingeführten Klassifikationsschema zeigen sich die Korrelati-

onsmuster der Tests. Die multiplen Validitätskoeffizienten (Lienert & Raatz, 1998) sind für TWR ($R = .51$) und 3d-WS ($R = .49$) hypothesenkonform nicht von $R = .60$ unterscheidbar. Die multiplen Korrelationskoeffizienten von SAWS ($R = .45$) und AWR ($R = .45$) sind moderat, lediglich PWS ($R = .32$) hat nur wenig geteilte Varianz.

Konvergente Validierung

Die Weisheitstests sollten eine konvergente, moderate Korrelation zu Persönlichkeitsreife (Ryff, 1989), Ego-Entwicklung (Hy & Loevinger, 1996) und Psychologischem Feingefühl (Goldberg, et al., 2006) haben. Für 3D-WS und SAWS trafen die erwarteten Befunde ein. Die PWS bildet hier eine Ausnahme. Die theoretische Ausrichtung von PWS legt nahe, dass es wenig Überschneidung mit Persönlichem Wachstum und Ego-Entwicklung gibt, was durch die Daten gestützt ($r = .05$, n.s.; $r = .01$, n.s.) wird. Die Instrumente TWR und AWR zeigen entgegen der Hypothese keine signifikante Beziehung zu den Konstrukten Persönlichkeitsreife und Psychologisches Feingefühl.

Diskriminante Validierung

Die Weisheitstests sollten eine diskriminante Beziehung zu Verbaler Intelligenz, Alter, Geschlecht und Umweltkontrolle als Indikator der Adaptivität beweisen. Die Befunde der fünf Weisheitstestverfahren zu Intelligenz, über den Wortschatztest des HAWIE-R (Tewes, 1991) telefonisch erhoben, sind hypothesenkonform. Die höchste Korrelation erreicht hier 3d-WS ($r = .14$, $p = .10$, $Z = -1.67$).

In der Studie zeigte sich ein leicht positiver Alterstrend bei 3d-WS, SAWS, und PWS. Die höchste Korrelation erreicht SAWS mit $r = .21$ ($p = .01$). Bei TWR und AWR dagegen, gibt es erwartungsgemäß keine Alterstrends.

Weibliche Teilnehmer schnitten im den Tests PWS ($r_{pb} = .25$, $p < .01$) und TWR ($r_{pb} = .27$, $p < .01$) leicht besser ab. Weitere Geschlechtseffekte finden sich nicht.

SAWS, TWR, AWR und PWS zeigen hypothesenkonform Werte, die den erwarteten diskriminanten Zusammenhang zu Umweltkontrolle (Ryff, 1989) nicht ausschließen. Dagegen bestimmen die Befunde des 3d-WS ($r = .60$; $p < .01$) Umweltkontrolle als konvergentes Konstrukt zu Weisheit. Aufgrund eines unerwartet hohen Deckeneffektes der Skala Umweltkontrolle, ist die Aussagekraft der Befunde eingeschränkt.

Einschränkung der Weisheitstests

Die Einschränkung der Ergebnisaussagekraft von TWR und AWR ist den kurzen Antworten

der Proband geschuldet. Die durchschnittliche Antwortlänge auf beiden Tests lag unter dem Grenzwert von 70 Worten, die zur Analyse von Helson und Srivastava empfohlen wurden. Der Detailreichtum eigener Introspektion variiert mit der Selbstoffenbarungstendenz, und daher auch deutlich mit der Antwortlänge. Eine Anpassung des Beurteilungsschemas konnte daher diesen Effekt nicht völlig kompensieren. Die Testwerte präsentieren daher nur bedingt die eigentlichen Konstrukte der Transzendenten oder Allgemeinen Weisheit. Die Bereitschaft Selbstoffenbarung ist ein wesentlicher Faktor bei der Erhebung. Bei der weiteren Verwendung dieser Tests ist darauf zu achten, entweder eine höhere Antwortlänge zu fordern, oder die Selbstoffenbarung der Teilnehmer zu kontrollieren.

Generalisierung

Da sich für Nominierte und Psychotherapeuten erwartungsgemäße Validitätskoeffizienten ergeben, wird angenommen, dass diese Ergebnisse auf deren Population verallgemeinerbar sind. Es ist also möglich, anhand der Kriterien der Äußeren Validierung Personenpopulationen zu identifizieren, die sich im Durchschnitt durch ein gemeinsames Kovariieren der einzelnen Weisheitstests auszeichnen, wenn auch nicht auf ausgesprochen hohem Niveau.

Eine Generalisierung fällt für die Kontrollgruppe schwerer, da die Nominierten und die Kontrollpersonen den typischen Profilen von Weisheitsnominierten entsprechen. Anhand der Matchingkriterien von Alter, Geschlecht, Beruf und Bildungsniveau der Nominierten-Gruppe, sind Personengruppen mit hoher Persönlichkeitsreife identifizierbar, für die die Befunde dieser Studie generalisiert werden können. Diese Profile identifizieren besonders Ego-Entwickelte Personengruppen, da sich keine Mittelwertsdifferenz auf dem WUSCT abbildet. Aufgrund der festgestellten Differentiellen Validität der Testverfahren jedoch, kann zwischen Kontrollpersonen und den Nominierten unterschieden werden. Die Kontrollgruppe erzielte sich gegenseitig ausgleichende Profilverläufe in den einzelnen Weisheitstestverfahren, während die der Nominierten gemeinsam variierten.

10 Literaturverzeichnis

- Achenbaum, W. A. (2004). Wisdom's Vision of Relations. *Human Development*, 47, 300-303.
- Achenbaum, W. A. & Orwoll, L. (1991). Becoming wise: A psycho-gerontological interpretation of the Book of Job. *International Journal of Aging and Human Development*, 32, 21-39.
- Allemand, M., Zimprich, D. & Hertzog, C. (2007). Cross-Sectional Age Differences and Longitudinal Age Changes of Personality in Middle Adulthood and Old Age. *Journal of Personality*, 75, 2, 323-358.
- Ardelt, M. (1997). Wisdom and life satisfaction in old age. *Journal of Gerontology: Series B: Psychological Sciences & Social Sciences*, 52B, 1, 15-27.
- Ardelt, M. (1998). Social crisis and individual growth: The long-term effects of the Great Depression. *Journal of Aging Studies*, 12, 3, 291-314.
- Ardelt, M. (2000a). Antecedents and effects of wisdom in old age: A longitudinal perspective on aging well. *Research on Aging*, 22, 4, 360-394.
- Ardelt, M. (2000b). Intellectual versus wisdom-related knowledge: The case for a different kind of learning in the later years of life. *Educational Gerontology*, 26, 8, 771-789.
- Ardelt, M. (2003). Empirical assessment of a three-dimensional wisdom scale. *Research on Aging*, 25, 3, 275-324.
- Ardelt, M. (2004a). Wisdom as Expert Knowledge System: A Critical Review of a Contemporary Operationalization of an Ancient Concept. *Human Development*, 47, 5, 257-285.
- Ardelt, M. (2004b). Where Can Wisdom Be Found?. A Reply to the Commentaries of Baltes and Kunzmann, Sternberg and Achenbaum. *Human Development*, 47, 5, 304-307.
- Ardelt, M. (2006). 3-dimensionale Weisheitsskala. *Geo* [online]. Verfügbar unter: <http://www.geo.de/GEO/interaktiv/tests/50033.html> [23. Mai 2006].
- Arlin, P. K. (1990). Wisdom: The Art of Problem Finding. In R. J. Sternberg (Eds.), *Wisdom: Its Nature, Origins, and Development* (pp. 230-243). Cambridge: Cambridge University Press.

- Argyle, M. (1987). *The psychology of happiness*. London: Methuen.
- Assmann, A. (1991). *Weisheit: Archäologie der literarischen Kommunikation III*. München Wilhelm Fink Verlag.
- Baltes, P. B. & Kunzmann, U. (2004). The Two Faces of Wisdom: Wisdom as General Theory of Knowledge and Judgment about Excellence in Mind and Virtue vs. Wisdom as Everyday Realization in People and Products. *Human Development*, 47, 290-299.
- Baltes, P. B. & Smith, J. (1990). The psychology of wisdom and its ontogenesis. In R. J. Sternberg (Eds.), *Wisdom: Its nature, origins, and development* (pp. 87-120). Cambridge: Cambridge University Press
- Baltes, P. B. & Staudinger, U. M. (2000). Wisdom: A metaheuristic (pragmatic) to orchestrate mind and virtue toward excellence. *American Psychologist*, 55, 122-136.
- Baltes, P.B., Staudinger, U. M., Maercker, A. & Smith J. (1995). People nominated as wise: A comparative study of wisdom-related knowledge, *Psychology and Aging*, 10, 155-166.
- Bauer, J. J., McAdams, D. P. & Sakaeda A. R. (2005). Interpreting the Good Life: Growth Memories in the Lives of Mature and Happy People. *Journal of Personality and Social Psychology*, 88, 1, 203-237.
- Birren, J. E. & Fischer, L.M. (1990). Conceptualizing wisdom: the primacy of affect-cognition relations. In R. J. Sternberg (Eds.), *Wisdom, its nature, origins, and development*, (pp. 317-332). Cambridge: Cambridge University Press.
- Birren, J. E. & Svensson, C. A. (2005). Wisdom in History. In R. J. Sternberg & J. Jordan (Eds.), *A Handbook of wisdom*, (pp.3-31). Cambridge: Cambridge University Press.
- Blanchard-Fields, F., Brannan, J. R. & Camp, C. J. (1987). Alternative conceptions of wisdom: An onion-peeling exercise. *Educational Gerontology*, 13, 497-502.
- Bluck, S. & Glück, J. (2004). Making Things Better and Learning a Lesson: Experiencing Wisdom across the Lifespan. *Journal of Personality*, 72, 3, 543-572.
- Bluck, S. & Glück, J. (2005). From the inside out. In R. J. Sternberg & J. Jordan (Eds.), *A Handbook of wisdom*, (pp. 84-109). Cambridge: Cambridge University Press
- Borkenau, P. & Ostendorf, F. (1989). Untersuchungen zum Fünf-Faktoren-Modell der Persönlichkeit und seiner diagnostischen Erfassung. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 10, 239-251.

- Borkenau, P. & Ostendorf, F. (1992). Social desirability scales as moderator and suppressor variables. *European Journal of Personality*, 6, 199–214.
- Bortz J. & Döring N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. (4. überarbeitete Auflage). Heidelberg: Springer.
- Brown, S. C. (2004). Learning across campus: How college facilitates the development of wisdom. *Journal of College Student Development*, 45, 134-148.
- Brown, S. C. & Greene J. A. (2006). The wisdom Development Scale: Translating the Conceptual to the Concrete. *Journal of College Student Development*, 47, 1, 1-19.
- Brown, S. C., Stevens, R. A., Trojano, P. E. & Schneider M. K. (2002). Exploring complex phenomena: Grounded theory in student affairs research. *Journal of College Student Development*, 43, 173-183.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Person Studium.
- Bursik, K. (1991). Adaption to divorce and ego development in adult women. *Journal of Personality and Social Psychology*, 60, 300-306.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multi-trait-multimethod matrix. *Psychological Bulletin*. 56, 81-105.
- Clayton, V. P. (1982). Wisdom and Intelligence: The Nature and Function of Knowledge in the Later Years. *International Journal of Aging and Development* 15, 315–23.
- Clayton, V. P. & Birren, J. E. (1980). The development of wisdom across the lifespan: A reexamination of an ancient topic. In P. B. Baltes & O. G. Brim (Eds.), *Life-span development and behavior* (Vol. 3, pp. 103-135). San Diego, Academic Press.
- Cronbach, L. J. (1971). Test validation. In R.L. Thorndike (Eds.), *Educational measurement* (Vol. 2, pp. 443-507). Washington, DC: American Council on Education.
- Csikszentmihalyi, M. & Nakamura, J. (2005). The Role of Emotions in the development of wisdom. In R. J. Sternberg & J. Jordan (Eds.), *A Handbook of Wisdom* (pp. 220-242). Cambridge: Cambridge University Press.
- Csikszentmihalyi, M. & Rathunde (1990). The psychology of wisdom: An evolutionary interpretation. In R. J. Sternberg (Eds.), *Wisdom, its nature, origins and development*, (pp. 22-51). Cambridge: Cambridge University Press.

- Cohn, L. D. & Westenber, P.M. (2004). Intelligence and Maturity: Meta-Analytic Evidence for the Incremental and Discriminant Validity of Loevinger's Measure of Ego Development. *Journal of Personality and Social Psychology*, 86, 5, 760-772.
- Costa, P. T., Zonderman, A. B., & McCrae, R. R. (1991). Personality, defense, coping, and adaptation in older adulthood. In E. M. Cummings, A. L. Greene & K. H. Karraker (Eds.), *Life-span developmental psychology: Perspectives on stress and coping* (pp. 277-293). Hillsdale: Erlbaum.
- Denney, N. W., Dew, J. R. & Kroupa, S. L. (1995). Perceptions of wisdom: What is wisdom and who has it? *Journal of Adult Development*, 2, 37-47.
- Diener, E., Sandvik, E., Pavot, W. & Gallagher, D. (1991). Response artifacts in the measurement of subjective well-being. *Social Indicators Research*, 24, 35-56.
- Dittmann-Kohli, F. (1984). Weisheit als mögliches Ergebnis der Intelligenzentwicklung im Erwachsenenalter. *Sprache und Kognition*, 2, 112-132.
- Drewes, M. J. & Westenber, M. (2001). The Impact of Modified Instructions on Ego-Level Scores: A Psychometric Hazard or Indication of Optimal Ego-Level? *Journal of Personality Assessment*, 76, 2, 229-249.
- Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology*, 37, 90-93.
- Egloff, B. & Schmuckle S. C. (2002). Predictive Validity of the Impact Association Test for Assessing Anxiety. *Journal of Personality and Social Psychology*, 82, 6, 1441-1455.
- Ellis, A. (1977). *Die rational-emotive Therapie*. München: Verlag J. Pfeiffer.
- Fisseni, H.-J. (2003). *Persönlichkeitspsychologie*. Göttingen: Hogrefe.
- Fromm, E. (1954). *Psychoanalyse und Ethik*. Stuttgart: Diana.
- Gardener, G. (1993). *Frames of mind: A theory of multiple Intelligences*. London: Fontana Press.
- Glück, J. & Bluck, S. (2003). Jedes Alter hat seine Weisheit. *Psychologie Heute*, 4, 36-39.
- Glück, J., Bluck, S., Baron, J. & McAdams, D. P. (2005). The wisdom of experience: Autobiographical narratives across adulthood. *International Journal of Behavioral Development*, 29, 3, 197- 208.

- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R. & Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40, 84-96.
- Gough, H. G. (1964). *The California Psychological Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Gough, H. G. & Heilbrun, A. B. (1983). *The Adjective Check List manual*. Palo Alto, CA: Consulting Psychologists Press.
- Grigorenko, E. L. & Sternberg, R. J. (1995). Thinking Styles. In D. H. Saklofske und M. Zeidner (Eds.), *International Handbook of Personality and Intelligence*, (pp. 205-229). New York: Plenum Press.
- Häcker, H., Leutner, D. & Amelang, M. (1998). Standards für pädagogisches und psychologisches Testen. *Supplementum 1/1998 der Diagnostica und der Zeitschrift für Differentielle und Diagnostische Psychologie*. Göttingen, Bern: Hogrefe und Hans Huber.
- Härder, M. & Härder S. (1998). Neure Entwicklungen bei der Delphimehtode. Literaturbericht II. *ZUMA-Arbeitsbericht 98,05*. Mannheim:
- Hartmann, C. (2004). Perspective[wisdom]. In C. Peterson & M. E. P. Seligman, (Eds.), *Character strengths and virtues: a handbook and classification*. (pp. 181-196). Oxford: Oxford University Press.
- Helson, R. & Roberts, B. W. (1994). Ego development and personality change in adulthood. *Journal of Personality and Social Psychology*, 66, 911-920.
- Helson, R. & Srivastava, S. (2001). Three Paths of Adult Development: Conservators, Seekers and Achievers. *Journal of Personality and Social Psychology*, 80, 6, 955-1010.
- Helson, R. & Srivastava, S. (2002). Creative and Wise People: Similarities, Differences and How they Develop. *Personality and Social Psychology Bulletin*, 28, 10, 1430-1440.
- Helson, R. & Wink, P (1987). Two Conceptions of Maturity Examined in the Findings of a Longitudinal Study. *Journal of Personality and Social Psychology*, 53, 3, 531-541.
- Hershey, D. A. & Farrell, A. H. (1997). Perceptions of wisdom associated with selected occupations and personality characteristics. *Current Psychology: Developmental, Learning, Personality, Social*, 16, 115-130.
- Holliday, S. G. & Chandler, M. H. (1986). *Wisdom: Explorations in adult competence*. New York: Karger.

- Holt, R. R. (1980). Loevinger's Measure of Ego Development: Reliability and National Norms for Male and Female Short Forms. *Journal of Personality and Social Psychology*, 39, 5, 909-920.
- Jason, L. A., Reichler, A., Kinge, C., Madsen, D., Chamacho, J. & Marchese W. (2001). The measurement of wisdom: A preliminary effort. *Journal of Community Psychology*, 29, 585-598.
- Jordan, J. (2005). The Quest for wisdom in Adulthood: A Psychological Perspective . In R. J. Sternberg & J. Jordan (Eds.), *Handbook of wisdom*, (p 160-188). Cambridge: Cambridge University Press.
- Jung, C. G. (1930). *Psychologische Typen*. Zürich: Rascher.
- Kekes, J. (1983). Wisdom. *American Philosophical Quarterly*, 20, 277–286.
- Kitchener, K. S. (1983). Changing conceptions of the philosophy of science and the foundations of development psychology. *Human Development*, 8, 1-30.
- Kitchener, K. S. & Brenner, H. G. (1990). Wisdom and Reflective Judgment: Knowing in the Face of Uncertainty. In R. J. Sternberg (Eds.), *Wisdom: Its Nature, Origins, and Development*, (pp. 212–29). Cambridge: Cambridge University Press.
- Klemp, G. O. Jr. & McClelland, D. C. (1988). What characterizes intelligent functioning among senior managers? In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 31-50). New York: Cambridge University Press.
- Kunzmann, U. & Baltes, P. B. (2003) Wisdom-related knowledge: Affective, motivational, and interpersonal correlates. *Personality and Social Psychology Bulletin*, 29, 1104-1119.
- Kunzmann, U. & Baltes, P. B. (2005). The Psychology of wisdom. In R. J. Sternbergs & J. Jordan (Eds.), *A Handbook of wisdom*, (pp. 110-135). Cambridge, Cambridge University Press.
- Kunzmann, U., Stange, & Jordan, J. (2005). Positive Affectivity and Lifestyle in Adulthood: Do You Do What You Feel? *Personality and Social Psychology Bulletin*, 31, 4, 574-588

- Kramer, D. A. (1990). Conceptualizing Wisdom: The Primacy of Affect-Cognition Relations. In R. J. Sternberg (Eds.), *Wisdom: Its Nature, Origins, and Development*, (pp. 279–313). Cambridge: Cambridge University Press.
- Kramer, D. A. (2000). Wisdom as a Classical Source of Human Strength: Conceptualization and Empirical Inquiry. *Journal of Social and Clinical Psychology* 19, 83–101.
- Labouvie-Vief, G. (1990). Wisdom as integrated thought: Historical and developmental perspectives. In R. J. Sternberg (Eds.), *Wisdom: Its nature, origins, and development* (pp. 52-86). New York: Cambridge University Press.
- Labouvie-Vief, G. (2003). Dynamic integration: Affect, cognition, and the self in adulthood. *Current Directions in Psychological Science*, 12, 201-206.
- Labouvie-Vief, G., DeVoe, M. & Bulka, D. (1989). Speaking about feelings: Conceptions of emotion across the life span. *Psychology and Aging*, 4, 425-437.
- Labouvie-Vief, G. & Diehl, M. (1998). The role of ego development in the adult self. In P. M. Westenberg, A. Blasi & L. D. Cohn (Eds.), *Personality development: Theoretical, empirical, and clinical investigations of Loevinger's conception of ego development* (pp. 219–235). Mahwah: Lawrence Erlbaum Associates, Inc.
- Labouvie-Vief G. & Medler, M. (2002). Affect Optimization and Affect Complexity: Modes and Styles of Regulation in Adulthood. *Psychology and Aging*, 17, 4, 571-588.
- Levitt, H. M. (1999). The Development of Wisdom: An Analysis of Tibetan Buddhist Experience. *Journal of Humanistic Psychology*, 39, 86–105.
- Li, S. C. & Kunzmann, U. (2004) Research on Intelligence in German-Speaking Countries. In R. J. Sternberg (Eds.), *International Handbook of Intelligence*. (pp. 135-169). Cambridge: Cambridge University Press.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse*. (6. überarbeitete Auflage). Weinheim: Beltz.
- Lindenberger, U. & Baltes, P. B. (1997). Intellectual functioning in old and very old age: Cross sectional results from the Berlin Aging Study. *Psychology and Aging*, 12, 410-432.
- Loevinger, J. (1976). *Ego development: Conception and theory*. San Francisco: Jossey-Bass.
- Loevinger, J. (1998). *Technical Foundations for Measuring Ego Development*. Mahwah, Lawrence Erlbaum Associates.

- Lönnqvist, J.-E., Paunonen S., Tuulio-Henriksson A., Lönnqvist J. & Verkasalo M.. (2007). Substance and style in Socially Desirable Responding. *Journal of Personality*, 75, 2, 291-322.
- Lorr, M. & Manning, T. T. (1978). Measurement of ego development by sentence completion and Personality test. *Journal of Clinical Psychology*, 34, 354-360.
- Mackintosh, N. J. (1998). *IQ and Human Intelligence*. Oxford University Press: Oxford.
- Maercker, A., Böhmig-Krumhaar, S. & Staudinger, U. M. (1998). Existentielle Konfrontation als Zugang zu weisheitsbezogenem Wissen und Urteilen: Eine Untersuchung von Weisheitsnominierten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 30, 2-12.
- Manners, J. & Durkin, K. (2001). A Critical Review of the Validity of Ego Development Theory and Its Measurement. *Journal of Personality Assessment*, 77, 3, 541-567.
- Maslow, A. H. (1999). *Motivation und Persönlichkeit*. Reinbek bei Hamburg: Rowohlt.
- McCrae, R.R. & Costa, P. T. (1980a). Openness to Experience and Ego Level in Loevinger's Sentence Completion Test: Dispositional Contributions to Developmental Models of Personality. *Journal of Personality and Social Psychology*, 39, 6, 1179-1190.
- McCrae, R.R. & Costa, P. T. (1980b). Psychological maturity and subjective well-being: Toward a new synthesis. *Developmental Psychology*, 19, 243-248.
- McCrae, R. R. & Costa, P. T., Jr. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology*, 51, 882-888.
- McCrae, R. R. & Costa, P. T. (1997). Conceptions and correlates of openness to experience. In R. Hogan, S. Briggs & J. Johnson (Eds.), *Handbook of personality psychology* (pp. 825-847). San Diego: Academic Press.
- McCrae, R.R. & John, O.P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60, 175-215.
- Meacham, J. A. (1983). Wisdom and the Context Of Knowledge: Knowing That One Doesn't Know. In D. Kuhns und J. A. Meacham (Eds.), *On the Development Of Developmental Psychology*, (pp. 111-134). Basel: Karger.
- Meacham, J. A. (1990). The loss of wisdom. In R. J. Sternberg (Eds.), *Wisdom, its nature, origins and development*, (pp. 181-211). Cambridge: Cambridge University Press.

- Michel, L. & Conrad, W. (1982). Theoretische Grundlagen psychometrischer Tests. In K-J Groffmann & L. Michel (Hrsg.), *Enzyklopädie der Psychologie, Serie II Grundlagen psychologischer Diagnostik, Band 1 Psychologische Diagnostik* (S. 1-129). Göttingen: Hogrefe.
- Mickler, C. (2004). Selbstbezogene Weisheit Ein Instrument zur Messung von Persönlichkeitreife – Entwicklung, Validierung, Alterseffekte. Dissertation. Bremen: Jacobs Center for Lifelong Learning and Institutional Development.
- Newman, D. L., Tellegen, A. & Bouchard, T.J. (1998). Individual differences in adult ego development: Sources of influence in twins reared apart. *Journal of Personality and Social Psychology*, 74, 4, 985-995.
- Orwoll, L. & Achenbaum, W. A. (1993). Gender and the development of wisdom. *Human Development*, 36, 276-296.
- Orwoll, L. & Perlmutter, M. (1990). The study of wise persons: Integrating a personality perspective. In R.J. Sternberg (Eds.), *Wisdom: Its nature, origins, and development* (pp. 160-177). Cambridge: Cambridge University Press.
- Paulhus, D. L. (1991). Measurement and control of response biases. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measurement of personality and social psychological attitudes* (pp. 17-59). San Diego: Academic press.
- Paulhus, D.L., & John, O.P. (1998). Egoistic and moralistic bias in self-perceptions: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, 66, 2024-1060.
- Pascual-Leone, J. (1990). An Essay on Wisdom: Toward Organismic Processes That Make It Possible. In R. J. Sternberg (Eds.), *Wisdom: Its nature, origins, and development* (pp. 244–278). Cambridge: Cambridge University Press.
- Pasupathi, M. & Staudinger U. M. (2001). Do advanced moral reasoners also show wisdom? Linking moral reasoning and wisdom-related knowledge and judgement. *International Journal of Behavioral Development*, 25, 401-415.
- Pasupathi, M., Staudinger U.M. & Baltes, P. B. (2001). Seeds of wisdom: Adolescents' knowledge and judgement about difficult life problems. *Developmental Psychology*, 37, 351-361.

- Peterson, C. & Seligman, M.E. (2004). *Character strengths and virtues*. New York: Oxford University Press.
- Piedmont, R. L., McCrea, R. C., Riemann, R. & Angleitner, A. (2000). On the Invalidity of Validity Scales: Evidence From Self-Reports and Observer Ratings in Volunteer Samples. *Journal of Personality and Social Psychology*, 78, 3, 582- 593.
- Rathunde, K. (1995). Wisdom and Abiding Interest: Interviews with Three Noted Historians in Later Life. *Journal of Adult Development* 2, 159–72.
- Reber, A. S., Walkenfeld, F. F., & Hernstadt, R. (1991). Implicit and explicit learning: individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 5, 888-96.
- Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, 38, 1, 119- 125.
- Richardson, M. J. & Pasupathi, M. (2005) Young and Growing Wiser. In R. J. Sternberg & J. Jordan (Eds.), *Handbook of wisdom* (pp. 139-159). Cambridge: Cambridge University Press.
- Rogers, C. R. (1961). *On becoming a person*. Boston: Houghton Mifflin.
- Rowe, G. & Wright G. (1999). The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting*, 15, 353-375.
- Ryan R. M., Deci, E. L. (2001). On happiness and human potentials. A review of research of hedonistic and eudaimonic well-being. *Annual Review of Psychology*, 52, 141-166.
- Ryff, C. D. (1989). Happiness is everything or is it? Exploration on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57, 6, 1069-1081.
- Ryff, C. D. & Keyes, C. L. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, 69, 4, 719-727.
- Saum-Aldehoff, T. (2001). „Weisheit ist leicht zu erkennen –aber schwer zu erreichen“. Ein Gespräch mit Ursula Staudinger. *Psychologie Heute*, 10, S, 22-27.
- Schippan, B., Baumann, K., Linden, M. (2004). Weisheitstherapie – kognitive Therapie der posttraumatischen Verbitterungsstörung. *Verhaltenstherapie*, 14, 282-293.
- Schmutte, P. S. & Ryff, C. D. (1997). Personality and Well-Being: Reexamining Methods and Meanings. *Journal of Personality and Social Psychology*, 73, 3, 549-559.

- Schnell R., Hill P. B. & Esser, E. (1999). *Methoden der empirischen Sozialforschung*. (6. überarbeitete Auflage). München: Oldenburg.
- Seligman, M. E. P. & Csikszentmihalyi, M. (2000). Positive Psychology: An Introduction. *American Psychologist*, 55, 1, 5-14.
- Smith, J. & Baltes, P. B. (1990). A study of wisdom-related knowledge: Age/cohort differences in responses to life planning problems. *Developmental Psychology*, 26, 494-505.
- Smith, J., Staudinger U. M. & Baltes, P. B. (1994). Occupational settings facilitating wisdom-related knowledge: The sample of clinical psychologists. *Journal of Consulting and Clinical Psychology*, 62, 989-999.
- Sowarka, D. (1989). Weisheit und weise Personen: Common-Sense-Konzepte älterer Menschen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 21, 87-109.
- Starke-Perschke, S., Bliesener, T., Dannecker, M., Dlugosch, G., Flor, D., Gstalter, H., Jäger, R. S., Kalb, P. E., Mangold, R., Pritzel, M., Richter-Appelt, H., Rothenburg, C., Schmidbauer, W., Sigusch, V., Spieß, E. & Wenninger, G. (2001). *Der Brockhaus Psychologie. Fühlen, Denken und Verhalten verstehen*. Mannheim: Brockhaus.
- Staudinger, U. M. (1996). Wisdom and the social-interactive foundation of the mind. In P. B. Baltes & U. M. Staudinger (Eds.), *Interactive minds: Life-span perspectives on the social foundation of cognition* (pp. 276-315). New York: Cambridge University Press.
- Staudinger, U. M. (1999). Older and wiser? Integrating results on the relationship between age and wisdom-related performance. *International Journal of Behavioral Development*, 23, 641-664.
- Staudinger, U. M. (2005). Weisheit, Lebens- und Selbsteinsicht. In H. Weber & T. Rammsayer (Hrsg.), *Handbuch der Persönlichkeitspsychologie und Differentiellen Psychologie*. (S. 342-349). Göttingen: Hogrefe.
- Staudinger, U. M. & Baltes, P.B. (1996a). Interactive minds: A facilitative setting for wisdom-related performance? *Journal of Personality and Social Psychology*, 71, 746-762.
- Staudinger, U. M. & Baltes, P. B. (1996b). Weisheit als Gegenstand psychologischer Forschung. *Psychologische Rundschau*, 47, 57-77.

- Staudinger, U. M., Dörner, J. & Mickler M. (2005). Wisdom and Personality. In R.J. Sternberg & J. Jordan (Eds.), *A Handbook of wisdom*. (pp. 191-219). Cambridge: Cambridge University Press.
- Staudinger, U. M., Lopez, D. F. & Baltes, P. B. (1997). The Psychometric Location of Wisdom-Related Performance: Intelligence, Personality, and More? *Personality and Social Psychology Bulletin*, 23, 11, 1200-1214.
- Staudinger, U. M., Maciel, A. G., Smith, J. & Baltes, P. B. (1998). What predicts wisdom-related performance? A first look at Personality, Intelligence and Facilitative Experimental Contexts. *European Journal of Personality*, 12, 1-17.
- Staudinger, U. M. & Pasupathi, M. (2003). Correlates of wisdom-related performance in adolescence and adulthood: Age-graded differences in "paths" toward desirable development. *Journal of Research on Adolescence*, 13, 239-268.
- Staudinger, U. M., Smith, J. & Baltes, P. B. (1992). Wisdom-related knowledge in a life-review task: Age differences and the role of professional specialization. *Psychology and Aging*, 7, 271-281.
- Staudinger, U. M., Smith, J., & Baltes, P. B. (1994). *Handbuch zur Erfassung von weisheitsbezogenem Wissen. Materialien aus der Bildungsforschung. Nr. 46*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Stern, W. (1912). *Die psychologischen Methoden der Intelligenzprüfung und deren Anwendung an Schulkindern*. Leipzig: Barth.
- Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology*, 49, 607-627.
- Sternberg, R. J. (1988). *The triarchic mind: A new theory of human intelligence*. New York: Viking.
- Sternberg, R. J. (1990). Wisdom and its relations to intelligence and creativity. In R. J. Sternberg (Eds.), *Wisdom: Its nature, origins, and development* (pp. 142-149). Cambridge: Cambridge University Press.
- Sternberg, R. J. (1997). *Thinking Styles*. Cambridge: Cambridge University Press.
- Sternberg, R. J. (1998). A balance theory of wisdom. *Review of General Psychology*, 2, (4), 347-365.

- Sternberg, R. J. (2004a). What is Wisdom and how you can develop it? *The Annals of the American academy*, 591, 164-174.
- Sternberg, R. J. (2004b). Words to the wise about wisdom. A commentary of Ardelt's Critique of Baltes. *Human Development*, 47, 286-289.
- Sternberg, R. J. (2005). Foolishness. In R.J. Sternberg und J. Jordans (Eds.) *A Handbook of Wisdom*. (pp. 331-352). Cambridge: Cambridge University Press.
- Sternberg R. J., Wagner, R. K., Williams W. M. & Horvath J. A. (1995). Testing common sense. *American Psychologist*. 50, 912-926.
- Stöber, (2001). The Social Desirability Scale-17. Convergent Validity, Diskriminant Validity, and Relationship with Age. *European Journal of Psychological Assessment*, 17, 3, 222-232.
- Takahashi, M. (2000). Toward a culturally inclusive understanding of wisdom: Historical roots in the East and West. *International Journal of Aging and Human Development*, 51, 3, 217-230.
- Takahashi, M. & Bordia P. (2000). The concept of wisdom: A cross-cultural comparison. *International Journal of Psychology*, 35, 1, 1-9.
- Takahashi, M. & Overton, W. F. (2002). Wisdom: A culturally inclusive developmental perspective. *Interantional Journal of Behavioral Development*, 26, 267-277.
- Takahashi, M. & Overton, W. F. (2005). Cultural Foundations of wisdom. In R. J. Sternberg & J. Jordan (Eds.), *A Handbook of Wisdom* (pp. 32-60). Cambridge: Cambridge University Press.
- Taranto, M.A. (1989). Facets of Wisdom: A Theoretical Synthesis. *International Journal of Aging and Human Development*, 29, 1-21.
- Tewes, U. (1991). *HAWIE-R. Hamburg-Wechsler Intelligenztest für Erwachsene. Revision 1991. Handbuch und Testanweisung*. Bern: Huber.
- Tügel, H. (2006). Die Wurzeln der Weisheit. *Geo*, 3, 144-168.
- Webster, J. D. (2003). An exploratory analysis of a self-assessed wisdom scale. *Journal of Adult Development*, 10, 13-22.
- Webster, J. D. (2007). Measuring the character strength of wisdom. *International Journal of Aging and Human Development*, 65, (2), 163-183.

-
- Westenberg, P. M. & Block, J. (1993). Ego development and individual differences in personality. *Journal of Personality and Social Psychology*, 65, 792-800.
- Wink, P. & Helson, R. (1987). Two Conceptions of Maturity Examined in the Findings of a Longitudinal Study. *Journal of Personality and Social Psychology*, 53, 3, 531-541.
- Wink, P. & Helson, R. (1997). Practical and Transcendent Wisdom: Their Nature and Some Longitudinal Findings. *Journal of Adult Development*, 4, 1, 1-14.
- Vaillant, G. E. (1977). *Adaption to life*. Boston. Little, Brown.

11 Tabellenverzeichnis

<i>Tabelle 3.1</i> Kategorisierung ausgewählter Weisheitstestverfahren.	37
<i>Tabelle 6.1.</i> Verwendete Testverfahren.	69
<i>Tabelle 6.2.</i> Statistische Hypothesen der Äußeren Validierung.	81
<i>Tabelle 6.3.</i> Mittelwerte der Teilnehmergruppen auf den Weisheitstestverfahren	83
<i>Tabelle 6.4</i> Verteilungswerte des Summenscores und der Dimensionen des AWR in Gesamtgruppe, nach dem Kriterium der Wortanzahl = 70 und des Mehr- dimensionalen Antwortstils	85
<i>Tabelle 6.5</i> Faktorenbildung 3d-WS-Faktor und HS-Faktor-Versionen.	87
<i>Tabelle 6.6</i> Korrelation zwischen Verbaler Intelligenz und WUSCT, TWR, und AWR.	88
<i>Tabelle 6.7</i> Altersvergleiche der Teilnehmergruppen.	89
<i>Tabelle 6.8</i> Berufstätigkeitsstatus der Teilnehmergruppen in Prozent.	90
<i>Tabelle 6.9</i> Familienstand der Teilnehmer in Prozent.	91
<i>Tabelle 7.1.</i> Multiple Korrelationskoeffizienten der Weisheitstestverfahren.	92
<i>Tabelle 7.2</i> Bivariate Korrelationen der Weisheitstestverfahren der Gesamtgruppe.	93
<i>Tabelle 7.3</i> Vergleich der Mittelwerte der Nominierten und Kontrollgruppe auf den einzelnen Weisheitstestverfahren.	95
<i>Tabelle 7.4</i> Vergleich der Mittelwerte der Psychotherapeuten und Kontrollgruppe auf den einzelnen Weisheitstestverfahren.	96
<i>Tabelle 7.5</i> Vergleich der Mittelwerte der Nominierten und Psychotherapeuten auf den einzelnen Weisheitstestverfahren.	97
<i>Tabelle 7.6</i> Bivariate Korrelationen der Weisheitstestverfahren mit Ego-Entwicklung.	98
<i>Tabelle 7.7</i> Bivariate Korrelationen der Weisheitstestverfahren mit PF und PW.	99
<i>Tabelle 7.8</i> Bivariate Korrelationen der Diskriminanten Validierung von UK und HAWIE-R.	101
<i>Tabelle 7.9</i> Korrelation zwischen Alter und den Weisheitstestverfahren.	103
<i>Tabelle 7.10</i> Punkt-Biseriale Korrelation der Weisheitstests zu Geschlecht.	104
<i>Tabelle 7.11</i> Interkorrelationen der Weisheitstests der Nominierten und der Kontroll- gruppe.	105
<i>Tabelle 7.12</i> Interkorrelationen der Weisheitstests der Psychotherapeuten.	105
<i>Tabelle 7.13</i> Multiple Korrelationskoeffizienten der Weisheitstestverfahren.	107
<i>Tabelle 7.14</i> Interkorrelationen der konvergenten und diskriminanten Testverfahren der Validierungsstudie für die Nominierten und die Kontrollgruppe.	108

<i>Tabelle 7.15</i> Interkorrelationen der konvergenten und diskriminanten Testverfahren der Validierungsstudie für die Psychotherapeuten und für die Gesamtgruppe.	109
<i>Tabelle 7.16</i> Deskriptive Statistiken für Umweltkontrolle und Persönliches Wachstum.	109
<i>Tabelle 7.17</i> W-Tests zu Mittelwertsunterschieden der Erfolgstypen und Suchenden in den Weisheitstestverfahren.	110
<i>Tabelle 7.18.</i> Interkorrelationen der Weisheitstests der Erfolgstypen und Suchenden.	110

Tabellen zur weiteren Übersicht der statistischen Auswertung

<i>Tabelle F.1</i> Matchingprofile der Paarungen Nominierte & Kontrollperson.	181
<i>Tabelle. F.2</i> Mittelwerte verschiedener Teilnehmergruppen auf den Weisheitstestverfahren	184
<i>Tabelle F.3.1</i> Korrelationen der Weisheitstests mit den Konstrukten der Konvergenten und Diskriminanten Validierung in der Gesamtgruppe.	186
<i>Tabelle. F.3.2</i> Korrelationen der Weisheitstests mit den Konstrukten der Konvergenten und Diskriminanten Validierung in der Nominierengruppe.	186
<i>Tabelle F.3.3</i> Korrelationen der Weisheitstests mit den Konstrukten der Konvergenten und Diskriminanten Validierung in der Kontrollgruppe.	187
<i>Tabelle F.3.4</i> Korrelationen der Weisheitstests mit den Konstrukten der Konvergenten und Diskriminanten Validierung in der Psychotherapeutengruppe.	187
<i>Tabelle F.3.5</i> Korrelationen der Weisheitstests mit den Konstrukten der Konvergenten und Diskriminanten Validierung für den Typus der Suchenden.	188
<i>Tabelle F.3.6</i> Korrelationen der Weisheitstests mit den Konstrukten der Konvergenten und Diskriminanten Validierung für den Typus der Erfolgreichen.	188

12 Abbildungsverzeichnis

<i>Abbildung 4.1.</i> Typologie der Persönlichkeitsreife nach Helson und Srivastava (2001).	52
<i>Abbildung 7.1</i> Verteilung der Typen Suchenden und Erfolgreiche in den Teilnehmergruppen.	109

13 Anhang

A Falblatt Weisheitsnominierungsverfahren

Dies ist eine um 50 % verkleinerte Darstellung der Vorder- und Rückseite des Falblattes.

Wie nehmen Sie am Nominierungsverfahren teil?

Ohne **freiwillige Teilnahme** ist ein Vorschreiten der Forschung vor allem im Bereich der Psychologie ausgeschlossen. Für meine Diplomarbeit bin ich auf Ihre freundliche Unterstützung angewiesen und Ihnen im Voraus dafür sehr dankbar.

Die Teilnahme an der Studie dient der wissenschaftlichen Forschung, daher wird keine finanzielle Entlohnung für die Teilnahme geleistet.

☞ **Vertraulichkeit der Daten:**

Ich versichere **absolute Geheimhaltung** über Ihre Teilnahme. Die von Ihnen nominierten Personen werden für Dritte nicht mit Ihnen in Verbindung zu bringen sein und die gewonnenen Daten werden natürlich streng vertraulich behandelt.

☞ **Kontaktadresse:**

Von einer beliebigen Email-Adresse können Sie mir eine Zusage über die Teilnahme an der Studie zusenden. Die weitere Korrespondenz, erste und zweite Befragungsrunde, wird einzig über diese Email-Adresse erfolgen.

☞ **E-Mail:** weisheit-2006@web.de

Durchführungszeitraum: Juni 2006




Nominierungsverfahren zur Validierung von 3 Weisheitsmaßen

- Einladung -

Telefon: 06341/141261
E-Mail: hoffmann.marina@web.de



Marina Hoffmann
Godramsteinstr. 8
76829 Landau in der Pfalz

Telefon: 06341/141261
E-Mail: hoffmann.marina@web.de

Betreuung:
Prof. Dr. Manfred Schmitt
Universität Koblenz-Landau, Campus Landau
Forstr. 7
76829 Landau in der Pfalz

Telefon: 06341/280-495
Telefax: 06341/280-490
E-Mail: schmittm@uni-landau.de

Weisheitsnominierungsverfahren



Weisheit ist leicht zu erkennen, aber schwer zu erlangen

Diesem Sprichwort folgend, kann die **Allgemeingültigkeit** von verschiedenen Weisheitstestverfahren durch ein Nominierungsverfahren überprüft werden.

Studienleitung:

☞ **Diplomarbeitsstudie an der Universität Koblenz-Landau**

Dieses Falblatt informiert Sie über ein Nominierungsverfahren das Teil einer Diplomarbeit an der Universität Koblenz-Landau ist. Diese Untersuchung wird von **Marina Hoffmann** unter der Betreuung von **Prof. Dr. Manfred Schmitt** durchgeführt. Sollten Sie weitere Fragen zur Studie oder über das Nominierungsverfahren haben, können Sie gerne Kontakt zu uns aufnehmen.

Das Schneeballverfahren:

Ein solches Verfahren bedient sich der **Netzwerke der persönlichen Beziehungen** der Teilnehmer. Über diese Methode kann eine Vielzahl von Personen zu den möglichen Nominierten gezählt werden.

Die Studie:
Vergleich und Validierung dreier Weisheitsmessverfahren

Fast alle Verfahren zur Weisheitsmessung in der Psychologie gründen sich auf unterschiedliche Ansätze. In diesem jungen Forschungsgebiet sind vergleichende Studien notwendig.

☞ **Weisheitsmessverfahren**

In meiner Diplomarbeit werde ich zwei **Fragebogenverfahren** sowie eine Kombination aus **Fragebogen- und Testelementen** in Ihrer Identifikation von Weisheit empirisch untersuchen.

☞ **Validierung**

Ein Validierungsverfahren erfordert die **Absicherung der inhaltlichen Bedeutung** der Testverfahren in Ihrer Beziehung zu verschiedenen theoretisch relevanten Konstrukten auf **statistischer Basis**.

☞ **Identifikation von weisen Personen**

Durch das Nominierungsverfahren wird eine Gruppe von Weisen bestimmt, die mit einer Kontrollgruppe verglichen wird.

Ihr individuelles Verständnis von Weisheit ist entscheidend.

Möglicherweise können Sie mit Ihrem Konzept von Weisheit eine oder mehrere Personen verbinden, die sie gut kennen. Dies können Männer und Frauen jeder Altersgruppe oder Sozialschicht sein.

Die Nominierung:

☞ **Sie nominieren Personen, die :**

- sie **gut kennen** und **einschätzen** können
- Ihrem **Verständnis von Weisheit** entsprechen
- nach Ihrer Einschätzung **grundsätzlich bereit** sind, an einer eher umfangreichen psychologischen Studie **teilzunehmen**.

☞ **Die Weisheitsnominierten**

Bitte geben Sie bei der Nominierung an, wie Kontakt zu der jeweiligen Person hergestellt werden kann.

Die Nominierten werden daraufhin von der Nominierung in Kenntnis gesetzt und gebeten, an der Studie teilzunehmen.

B Anschreiben

B.1 Computervermitteltes Anschreiben für die Nominiertengruppe



Ansprechpartnerin:
Marina Hoffmann
Godramsteinerstr. 8
76829 Landau in der Pfalz

Telefon: 06341/707769
E-Mail: hoffmann.marina@web.de

Betreuung:
Prof. Dr. Manfred Schmitt
Universität Koblenz-Landau,
Campus Landau
Im Fort 7
76829 Landau in der Pfalz
Telefon: 06341/280-495
Telefax: 06341/280-490

<Datum>

Vergleichende Studie zu Weisheitsfragebögen

Sehr geehrte/r <Name>,

Dieses Schreiben informiert Sie über eine wissenschaftliche Studie an der Universität Koblenz-Landau, für die Sie als Teilnehmer vorgeschlagen wurden.

Das Forschungsgebiet:

In der Psychologie ist das Forschungsgebiet der Weisheit noch recht jung. Erste empirische Forschung wurde über Laientheorien zu Weisheit durchgeführt, bevor der Versuch unternommen wurde, Weisheit psychometrisch messbar zu machen. Eine Anwendung in klinischen Bereich wurde durch die Weisheitstherapie bei Anpassungsstörungen entwickelt.

Die Studie:

In dieser Studie werden erstmals verschiedene Fragebögen mehrerer Forschergruppen zur Erfassung von individueller Weisheit miteinander verglichen. Ziel der Studie ist es, Aussagen über den Nutzen und Validität der Fragebögen treffen zu können. Des Weiteren werden systematische Zusammenhänge zu anderen verwandten Konstrukten untersucht.

Ich möchte Sie herzlich bitten, Teilnehmer in dieser Studie zu werden.

Wie Sie für diese Studie nominiert wurden:

In einer Befragung zum Thema Weisheit wurde verschiedenen Personen folgende Frage gestellt: „Kennen Sie jemand, der Ihrem persönlichen Verständnis nach einer weisen Person entspricht?“ Sie wurden als Teilnehmer für diese Untersuchung vorgeschlagen.

Auch falls Ihnen das mit der Nominierung verbundene Lob unangemessen erscheint, so bitte ich Sie trotzdem, an dieser Forschungsarbeit teilzunehmen. Durch Ihre Teilnahme können Sie einen Beitrag zur Weiterentwicklung und Verbesserung der Weisheitsforschung leisten, der in diesem Ansatz noch völlig neu ist. Zu Ihrer Information erhalten Sie nach Abschluss der Studie einen Überblick über die dank ihrer Hilfe gewonnenen Erkenntnisse.

Umfang der Studie:

Die Studie besteht aus 2 Teilen:

Fragebogenheft (Bearbeitungszeit: ~ 1 Stunde)

- kann zur Bearbeitung unter folgendem Link heruntergeladen werden:

<http://www.gerechtigkeitsforschung.de/Weisheit/Fragebogen.doc>

- oder wird Ihnen per Post zugestellt, falls Sie mir Ihre Postadresse zukommen lassen.

Rücksendung des bearbeiteten Fragebogens bitte bis Ende <Zeitraum von 6 Wochen ab dem Versendedatum> entweder per Rücksendeumschlag oder als Anhang von beliebiger Emailadresse an <spezielle Emailadresse> .

Telefonat (Dauer: ~ 15 min)

Auf der letzten Seite des Fragebogens geben Sie eine Telefonnummer und Ihren Wunschzeitpunkt für ein weiterführendes Gespräch an. Bitte bedenken Sie bei der Terminfestlegung, dass Sie während des Gespräches nicht gestört werden sollten. Es entstehen keine Kosten für Sie.

Es entstehen keine Kosten für Sie.

Studienleitung:

Prof. Dr. Manfred Schmitt, Lehrstuhl für Diagnostik und Persönlichkeitspsychologie an der Universität Koblenz-Landau und cand. Psych. Marina Hoffmann.

Datenschutz:

Ihre Kontaktdaten werden streng vertraulich behandelt und Dritten nicht zur Verfügung gestellt. Die Untersuchungsdaten werden einzig für den oben angegebenen Forschungszweck verwendet und die Datenanalyse erfolgt anonymisiert. Die Anonymisierung Ihrer Daten wird von Ihnen persönlich auf der letzten Seite des Fragebogens durchgeführt und im späteren Telefonat erfragt. Sie müssen daher weder auf dem frankierten Rückumschlag noch am Telefon Ihren Namen nennen.

Eine individualisierte Ergebnismeldung an einzelne Teilnehmer ist nicht möglich.

Ihre Einstellung gegenüber der Erhebung ist wichtig:

Es steht außer Frage, dass Sie bei der Bearbeitung der Fragebögen Antwortalternativen erkennen können, die ihrem Verständnis von Weisheit am nächsten kommen. Eine derartig positive Einschätzung einzelner Antwortalternativen darf Ihr Antwortverhalten bitte nicht beeinflussen. Ich möchte Sie an dieser Stelle darauf hinweisen, dass die Erhebung nicht das Ziel hat, ihr persönliches Verständnis eines weisen Antwortverhaltens zu messen, sondern einzig der Beurteilung der Fragebögen an sich nachstrebt. Aus diesem Grund bitte ich Sie, nach bestem Wissen und Gewissen, die auf Sie persönlich zutreffende Antwort auf die Fragen zu geben, ungeachtet des vermeintlichen Weisheitsgehaltes Ihrer Antwort.

Sollten Sie sich für die Teilnahme entschieden haben, so ist es für die sichere Interpretation der Ergebnisse von Bedeutung, dass Sie versuchen, alle Fragen offen und ehrlich zu beantworten. Fehlende Daten stellen für die Interpretierbarkeit der Ergebnisse ein großes Problem dar.

Vorab besten Dank für Ihre Unterstützung!“

Mit freundlichen Grüßen

Marina Hoffmann

B.2 Anschreiben – spezifisches für Kontrollgruppe

Der folgende Text ersetzt den ersten Abschnitt in dem Brief für Nominierte.

Sehr geehrte/r XY,

dieses Schreiben informiert Sie über eine wissenschaftliche Studie an der Universität Koblenz-Landau, für die Sie als Teilnehmer aufgrund Ihres Werdegangs in Frage kommen. Sie entsprechen aufgrund Ihres Geschlechts, Alters und ihres Berufes einer speziell als „weise“ benannten Person, die bereits an dieser Studie teilgenommen hat.

Der folgende Text ersetzt den dritten Abschnitt in dem Brief für Nominierte, der erklärt, wie diese Personen für diese Studie nominiert wurden.

Warum ist ihre Teilnahme wichtig?

Die erste Gruppe dieser Untersuchung bilden Personen, die speziell als weise nominiert wurden. Die zweite Gruppe, der Sie angehören würden, ähnelt der ersten nach Geschlecht, Alter und Berufsprofil. Da die Weisheitsforschung nicht allein theoretischen Definitionen genügen möchte, sondern auch das Allgemeinverständnis von Weisheit in diesem Forschungsbereich ein wichtiges Kriterium ist, wurde die erste Teilnehmergruppe nach dem Allgemeinverständnis von Weisheit ausgewählt. Da weise Personen rar sind, ist für die statistische Auswertung ein größerer Personenkreis notwendig. Aus diesem Grund, wird eine zweite Gruppe von Teilnehmern, die große Ähnlichkeit zu der ersten hat, gebildet. Falls Sie sich nicht zur Teilnahme bereit erklären, beginnt die Suche nach einer Person mit ihrem Geschlecht, Alter und Berufsprofil neu. Auf Ihre freundliche Unterstützung meines Forschungsprojektes bin ich daher angewiesen. Sollten Sie nicht teilnehmen, jedoch eine weitere passende Person kennen, die Ihrer Meinung nach an dieser Studie teilnehmen würde, so wenden Sie sich bitte telefonisch oder per Email an uns, damit ich Kontakt zu dieser Person aufbauen kann. Findet sich kein passender Ersatz, ist die statistische Auswertung der gesamten Studie in ihrer Aussagekraft eingeschränkt und es könnte lediglich ein Teil der gewünschten Fragestellungen untersucht werden.

B.3 Anschreiben - spezifisches für Psychotherapeuten

Der folgende Text ersetzt den dritten Abschnitt in den Brief für Nominierte, der erklärt, wie diese Personen für diese Studie nominiert wurden.

Klinische Psychologen & Psychotherapeuten gesucht:

Dank Forschungsergebnissen des Max-Planck-Institutes für Bildungsforschung in Berlin und Studien der Berkeley University in Kalifornien wurde festgestellt, dass klinische Psychologen & Psychotherapeuten eine interessante Teilnehmergruppe darstellen. Insgesamt wird es drei Gruppen in dieser Untersuchung geben, eine Gruppe mit „Weisen“, eine gematchte Kontrollgruppe und die Gruppe der Klinischen Psychologen & Psychotherapeuten.

Ich möchte Sie herzlich bitten, Teilnehmer in der Gruppe der Klinischen Psychologen & Psychotherapeuten dieser Studie zu werden.

Falls sich nur wenige Klinische Psychologen & Psychotherapeuten zur Teilnahme bereit erklären, ist die statistische Auswertung der gesamten Studie in ihrer Aussagekraft sehr eingeschränkt und es kann lediglich ein Teil der gewünschten Fragestellungen untersucht werden. Auf Ihre freundliche Unterstützung meines Forschungsprojektes bin ich daher angewiesen.

C Das Fragebogenheft

Bei der Bearbeitung der folgenden Fragebögen interessiert es, welche Antwort Sie persönlich besonders gut widerspiegelt. Daher gibt es generell keine richtigen oder falschen Antworten, sondern nur solche, die entweder mehr oder weniger gut auf Sie zutreffen.

Suchen Sie bitte diejenige heraus, die Sie für sich am treffendsten halten.

Dieser Fragebogen hat ein Antwortformat mit fünf Abstufungen. Versuchen Sie, die gesamte Breite des Antwortformates auszunutzen.

	Trifft ganz bestimmt für mich zu	Trifft meistens für mich zu	Trifft teilweise für mich zu	Trifft kaum für mich zu	Trifft nicht für mich zu
1. Ich versuche, jede Seite einer Meinungsverschiedenheit zu betrachten, bevor ich eine Entscheidung treffe.					
2. Wenn ich hilfsbedürftige Menschen sehe, versuche ich ihnen in irgendeiner Weise zu helfen.					
3. Wenn ich mich über jemanden aufrege, versuche ich mich für eine Zeit in dessen Lage zu versetzen.					
4. Es gibt gewisse Menschen, die ich so wenig leiden kann, dass ich mich innerlich freue, wenn sie erwischt werden und für das was sie getan haben, bestraft werden.					
5. Ich versuche immer alle Seiten eines Problems zu betrachten.					
6. Manchmal empfinde ich echtes Mitgefühl für alle Menschen.					
7. Ich versuche Situationen, wo die Wahrscheinlichkeit besteht, tiefgründig über etwas nachzudenken zu müssen, vorauszusehen und zu vermeiden.					
8. Wenn ich darüber nachdenke, was mir schon passiert ist, kann ich nur Verbitterung empfinden.					
9. Ich habe oft einen anderen Menschen nicht getröstet, wenn er oder sie es gebraucht hätte.					
10. Ein Problem interessiert mich nicht, wenn ich denke, dass es keine Lösung gibt.					
11. Ich werde entweder sehr ärgerlich oder sehr niedergeschlagen, wenn Sachen schief gehen.					
12. Wenn andere Leute Probleme haben, dann ist das ihre Sache.					
13. Oft verstehe ich das Verhalten von anderen nicht.					
14. Manchmal bin ich emotional so aufgebracht, dass ich unfähig werde, alle Lösungen für meine Probleme zu sehen.					
15. Wenn Leute sich mit mir unterhalten, wünsche ich mir manchmal, dass sie weggehen würden.					
16. Ich nehme Dinge lieber hin, als zu versuchen, zu verstehen, warum sie so gekommen sind.					

	Trifft ganz bestimmt für mich zu	Trifft meistens für mich zu	Trifft teilweise für mich zu	Trifft kaum für mich zu	Trifft nicht für mich zu
17. Wenn ich mit einem Problem nicht zu Recht komme, versuche ich als erstes, mir einen Überblick über die Situation zu verschaffen und alle relevanten Informationen zu berücksichtigen.					
18. Ich mag es nicht, die Sorgen anderer Leute anzuhören.					
19. Ich zögere dabei, wichtige Entscheidungen zu treffen, nachdem ich über sie nachgedacht habe.					
20. Bevor ich jemand kritisiere, versuche ich mir vorzustellen, wie ich mich an seiner Stelle fühlen würde.					
21. Ich ärgere mich leicht über Leute, die sich mit mir Streiten.					
22. Ich fühle mich betrogen, wenn ich darauf zurückblicke, was mir alles widerfahren ist.					
23. Es ist mir durchaus recht, wenn ich nur die Antwort auf ein Problem kenne, anstatt die Gründe für die Antwort zu verstehen.					
24. Ich finde es manchmal schwierig, die Dinge aus der Sicht einer anderen Person zu sehen.					
Hier ändert sich das Antwortformat →	Stimme sehr zu	Stimme zu	Weder noch	Stimme nicht zu	Stimme überhaupt nicht zu
25. In unserer komplizierten Welt muss man sich auf vertrauenswürdige Autoritäten oder Experten verlassen, um herauszufinden, was los ist.					
26. Unglückliche Menschen, die sich nur selbst bemitleiden, verärgern mich.					
27. Das Leben ist im Grunde immer gleich.					
28. Menschen nehmen die Gefühle und das Empfinden von Tieren zu ernst.					
29. Man kann fast alle Menschen als ehrlich oder unehrlich einteilen.					
30. Es würde mir sehr viel besser gehen, wenn sich meine aktuelle Situation ändern würde.					
31. Es gibt für alles nur einen richtigen Weg.					
32. Ich weiß, dass es Menschen gibt, die ich niemals mögen werde.					
33. Es ist besser, nicht zu viel über Dinge zu wissen, die nicht verändert werden können.					

	Stimme sehr zu	Stimme zu	Weder noch	Stimme nicht zu	Stimme ü- berhaupt nicht zu
34. Bei mir gehen Sachen oft ohne eigene Schuld schief.					
35. Was ich nicht weiß, macht mich nicht heiß.					
36. Ich komme mit jeder Art von Menschen gut aus.					
37. Entweder kennt jemand die Antwort auf eine Frage, oder er kennt sie nicht					
38. Es ist eigentlich nicht mein Problem, wenn andere in Schwierigkeiten sind und Hilfe brauchen					
39. Menschen sind entweder gut oder schlecht.					

Gehen Sie bei dem folgenden Fragebogen genauso vor, aber beachten Sie, dass es sich hier um ein sechsstufiges Antwortformat handelt.

Sie müssen sich also für oder gegen die Aussage entscheiden und versuchen Sie alle Antwortabstufungen auszuschöpfen.

	Trifft sehr auf mich zu	Trifft auf mich zu	Trifft eher auf mich zu	Trifft eher nicht auf mich zu	Trifft nicht auf mich zu	Trifft überhaupt nicht auf mich zu
40. Im Großen und Ganzen habe ich das Gefühl, dass ich mein Leben recht gut im Griff habe.						
41. Oft erdrückt mich der Alltag mit seinen Anforderungen.						
42. Ich passe nicht besonders gut zu den Menschen in meiner Umgebung, irgendwie gehöre ich nicht dazu.						
43. Ich erledige meine vielen alltäglichen Aufgaben und Pflichten ganz gut.						
44. Oft habe ich das Gefühl, meine Pflichten wachsen mir über den Kopf.						
45. Wenn ich mit meinen Lebensumständen nicht zufrieden wäre, dann würde ich Maßnahmen ergreifen, um das zu ändern.						
46. Die Regelung meiner Finanzen und meiner persönlichen Angelegenheiten gelingt mir im Großen und Ganzen gut.						
47. Es belastet mich, dass ich bei dem, was täglich zu tun ist, nicht hinterher komme.						
48. Ich verstehe es, meine Zeit so einzuteilen, dass ich auch wirklich alles schaffe, was getan werden muss.						

	Trifft sehr auf mich zu	Trifft auf mich zu	Trifft eher auf mich zu	Trifft eher nicht auf mich zu	Trifft nicht auf mich zu	Trifft überhaupt nicht auf mich zu
49. Ich habe jeden Tag eine Menge zu tun, aber es befriedigt mich auch, dass ich alles schaffe.						
50. Schon bei dem Versuch, meinen Tag zu planen, verliere ich oft den Mut, weil ich sowieso nicht alles schaffe.						
51. Ich habe es geschafft, die Beschäftigungen zu finden und die Beziehungen aufzubauen, die für mich wichtig und richtig sind.						
52. Es fällt mir schwer mein Leben so zu gestalten, dass ich damit zufrieden bin.						
53. Es ist mir gelungen, mein Zuhause und meine Art zu Leben so ganz nach meinen Vorstellungen zu gestalten.						
54. Ich bin nicht an Aktivitäten interessiert, die meinen Horizont erweitern.						
55. Ich habe das Gefühl, dass ich mit der Zeit immer mehr über mich lerne.						
56. Ich gehöre zu denen, die gerne etwas Neues ausprobieren.						
57. Ich habe keine Lust irgendetwas Neues zu probieren, ich finde mein Leben so, wie es läuft, ganz gut.						
58. Ich denke es ist wichtig, immer wieder neue Erfahrungen zu machen, die in Frage stellen, was man über sich und die Welt denkt.						
59. Wenn ich es recht bedenke, habe ich mich in den letzten Jahren nicht besonders weiterentwickelt.						
60. Meiner Ansicht nach können Menschen jedes Alters sich weiterentwickeln.						
61. Mit der Zeit habe ich eine Menge Lebenserkenntnis gewonnen und bin dadurch als Mensch viel stärker und leistungsfähiger geworden.						
62. Ich habe das Gefühl, dass ich mich im Laufe der Zeit persönlich sehr viel weiter entwickelt habe.						
63. Ich bin kein Freund von neuen Situationen, in denen ich meine altvertrauten Gewohnheiten ändern muss.						
64. Für mich ist das Leben ein ständiger Lern- und Entwicklungsprozess.						
65. Wenn ich mir so ansehe, wie sich meine Ansichten im Lauf der Jahre verändert haben und reifer geworden sind, dann bin ich recht zufrieden.						
66. Ich habe es schon lange aufgegeben, mein Leben wesentlich verändern oder verbessern zu wollen.						
67. Da ist schon was Wahres dran, wenn man sagt: „Was Hänschen nicht lernt, lernt Hans nimmermehr.“						

	Trifft sehr auf mich zu	Trifft auf mich zu	Trifft eher auf mich zu	Trifft eher nicht auf mich zu	Trifft nicht auf mich zu	Trifft überhaupt nicht auf mich zu
68. Ich habe viele schmerzhaftere Ereignisse in meinem Leben überwunden.						
69. Es fällt mir leicht meine Gefühle der aktuellen Situation anzupassen.						
70. Oft denke ich über Zusammenhänge zwischen meiner Vergangenheit und der Gegenwart nach.						
71. Ich kann über persönliche Peinlichkeiten kichern.						
72. Ich lese gerne Bücher, welche mich dazu herausfordern Sachverhalte anders zu denken.						
73. Ich musste viele wichtige Lebensentscheidungen treffen.						
74. Gefühle überwältigen mich nicht, wenn ich persönliche Entscheidungen treffe.						
75. Ich denke oft über meine persönliche Vergangenheit nach.						
76. Sogar in schwierigen Lebenssituationen kann es erheitende Elemente geben.						
77. Ich genieße es, eine Vielfalt von Musikstilen neben meiner liebsten Art zu hören.						
78. Ich habe mich während meines Lebens mit vielen verschiedenen Arten von Menschen befasst.						
79. Ich bin auf meine eigenen Gefühle eingestellt.						
80. Ich schwelge recht häufig in Erinnerungen.						
81. Ich versuche und finde eine lustige Seite, wenn ich einen bedeutenden Übergang in meinem Leben zu bewältigen habe.						
82. Ich genieße es, eine breite Vielfalt von Gerichten aus verschiedenen Kulturkreisen zu versuchen.						
83. Ich habe viele moralische Zwangslagen erlebt.						
84. Ich bin sehr gut darin, meine emotionalen Zustände zu lesen.						
85. Zurückblicken auf meine Vergangenheit hilft mir Perspektiven für meine aktuellen Belange zu gewinnen.						
86. Man kann mich sehr leicht zum Lachen bringen.						
87. Ich suche oft nach neuen Dingen zum Ausprobieren.						
88. Ich habe schon viel von der negativen Seite des Lebens gesehen (Unehrlichkeit, Scheinheiligkeit).						

	Trifft sehr auf mich zu	Trifft auf mich zu	Trifft eher auf mich zu	Trifft eher nicht auf mich zu	Trifft nicht auf mich zu	Trifft überhaupt nicht auf mich zu
89. Ich kann meine Emotionen frei ausdrücken, ohne mich zu fühlen als könnte ich die Kontrolle verlieren.						
90. Ich erinnere mich oft an frühere Zeiten in meinem Leben, um zu sehen, wie ich mich seither geändert habe.						
91. An diesem Punkt meines Lebens, finde ich es leicht über meine Fehler zu lachen.						
92. Kontroverse Kunstwerke spielen eine wichtige und wertvolle Rolle in der Gesellschaft.						
93. Ich habe viele schwierige Lebensübergänge durchlebt.						
94. Ich bin gut darin, meine feinen Emotionen in mir zu identifizieren.						
95. Mich an meine früheren Tage zu erinnern, hilft mir, Einblicke in wichtige Lebensangelegenheiten zu gewinnen.						
96. Ich benutze oft Humor, um andere zu beruhigen.						
97. Ich mag es unter Menschen zu sein, deren Ansichten sich sehr stark von meinen unterscheiden.						
98. Persönlich habe ich entdeckt, dass man nicht immer von dem Einband auf den Inhalt eines Buches schließen kann.						
99. Ich kann meine Gefühle regulieren falls es die Situation erfordert.						
100. Ich stelle oft fest, dass Erinnerungen aus meiner Vergangenheit wichtige Bewältigungsressourcen sein können.						
101. Jetzt glaube ich, die kleinen Ironien des Lebens richtig schätzen zu können.						
102. Ich bin sehr neugierig auf andere religiöse und/oder philosophische Glaubensrichtungen.						
103. Von anderen habe ich wertvolle Lebenslehren gelernt.						
104. Es scheint, ich hätte ein Talent dafür, die Gefühle anderer zu lesen.						
105. Das Wiedererleben vergangener Leistungen in der Erinnerung erhöht meine Zuversicht für heute.						
106. Ich kann mich über mich selbst lustig machen um andere zu ermutigen.						
107. Ich habe mich oft über das Leben und das was es jenseits davon gibt gewundert.						

**Vervollständigen Sie bitte die folgenden Sätze möglichst spontan.
Es gibt keine richtigen oder falschen Antworten.**

1. Wenn sich ein Kind nicht an Gruppenaktivitäten anschließen möchte...

2. Kinder erziehen...

3. Wenn ich kritisiert werde...

4. Die Aufgabe eines Mannes...

5. Mit anderen Menschen zusammen zu sein...

6. Was ich an mir mag, ist...

7. Meine Mutter und ich...

8. Was mich in Schwierigkeiten bringt, ist...

9. Erziehung...

**Vervollständigen Sie bitte die folgenden Sätze möglichst spontan.
Es gibt keine richtigen oder falschen Antworten.**

10. Wenn Menschen hilflos sind...

11. Frauen haben Glück, weil...

12. Ein guter Vater...

13. Ein Mädchen hat das Recht...

14. Wenn sie über Sex gesprochen haben, habe ich...

15. Eine Ehefrau sollte...

16. Es tut mir leid...

17. Ein Mann fühlt sich gut, wenn...

18. Regeln sind...

**Stellen Sie sich vor, Sie erhalten einen Anruf eines guten Freundes.
Dieser erzählt Ihnen, er könne nicht so weiter machen und er wolle sich das Leben nehmen.
Was würden Sie denken und tun?**

Viele Leute hoffen, mit steigendem Alter weiser zu werden.

Würden Sie ein Beispiel von Weisheit geben, die Sie erworben haben und beschreiben, wie es dazu gekommen ist?

Kreuzen Sie im Folgenden an, ob die Beschreibungen und die Aussagen auf Sie zutreffen oder nicht.

ehrllich	Ja	Nein	tolerant	Ja	Nein	unreif	Ja	Nein
unparteiisch	Ja	Nein	reflektiert	Ja	Nein	weise	Ja	Nein
gedankenvoll	Ja	Nein	realistisch	Ja	Nein	intelligent	Ja	Nein
verständnisvoll	Ja	Nein	leichtsinnig	Ja	Nein	intolerant	Ja	Nein
vielseitig interessiert	Ja	Nein	reif	Ja	Nein	vernünftig	Ja	Nein
nüchtern denkend	Ja	Nein	oberflächlich	Ja	Nein	einsichtig	Ja	Nein

Hier das bekannte sechsstufige Antwortformat →	Trifft sehr auf mich zu	Trifft auf mich zu	Trifft eher auf mich zu	Trifft eher nicht auf mich zu	Trifft nicht auf mich zu	Trifft überhaupt nicht auf mich zu
108. Ich kann mit guten Lösungen aufwarten.						
109. Ich bin leicht verletzbar.						
110. Ich bin in gar nichts hervorragend.						
111. Ich weiß, wie ich mein Wissen anwende.						
112. Ich komme leicht durcheinander.						
113. Ich fühle mich bei Rückschlägen erdrückt.						
114. Ich bin voller Ideen.						
115. Ich stelle meine Fähigkeit in Frage, meine Arbeit ordentlich zu erledigen.						
116. Ich weiß, dass ich kein besonderer Mensch bin.						
117. Ich fühle mich leicht angegriffen.						

D Material für das Telefoninterview

D.1 Gesprächsleitfaden für das Telefoninterview

1) Begrüßung:

Guten Tag, hier spricht Marina Hoffmann aus Landau. Ich rufe Sie wegen der Studie zum Vergleich von Weisheitsfragebögen an, an der Sie teilnehmen.“

2) Gesprächsrahmen:

„Haben Sie in der nächsten **viertel Stunde ungestört Zeit?**“

Falls Ja → weiter - Falls Nein → neuen Termin vereinbaren

3) Abfrage des Versuchspersonencodes:

„Als erstes möchte ich Sie nach **ihrem Verschlüsselungscode fragen**, damit die Zuordnung der Daten auch einwandfrei funktioniert.“ *Vorlesen und vergleichen*

„Nennen Sie mir die ersten beiden Anfangsbuchstaben des **Vornamens ihrer Mutter** und die ersten beiden Anfangsbuchstaben des **Vornamens ihres Vaters**.

An welchem **Tag** wurden Sie geboren?

In welchem **Monat** wurden Sie geboren?“

4) Falls bei demographischen Daten Angaben fehlen, dann:

Als nächstes möchte ich (eine) **fehlende Angabe(n)** zu ihren persönlichen Daten im Fragebogen ergänzen, die sie wahrscheinlich übersehen haben ...

5) Redebedarf aufgrund der Fragebogenerhebung?

„Möchten Sie noch etwas zur bisherigen Erhebung sagen, oder haben Sie offene Fragen die ich Ihnen nun gerne beantworten kann?“ *Antworten*

6) Situation für Intelligenzmessung geeignet?

„**Geht es Ihnen gut?**“

Können Sie sich in den nächsten Minuten **in Ruhe konzentrieren?**“

Falls Ja → weiter - Falls Nein → neuen Termin vereinbaren

„Bei diesem Gespräch geht es um die Erhebung **von Richtwerten für ihre Intelligenz**. Dieser Teil der Studie dient der Erforschung der Stärke des Zusammenhangs von Intelligenz und Weisheit, bzw. ob es überhaupt einen Zusammenhang zwischen Intelligenz und

Weisheit gibt.“ *Rückfragen klären und Anonymität dieser Erhebung betonen, etwa: „Ihre Intelligenzwerte werden natürlich genauso streng vertraulich behandelt wie alle anderen Daten. Ihre Telefonnummer oder ihr Name wird bei der weiteren Auswertung nicht mit ihren Intelligenzwerten in Verbindung zu bringen sein.“*

Zustimmung zu diesen Tests erfragen: Falls Ja → weiter - Falls Nein → Ende

7) Hawie-R Wortschatztest

„Als nächstes möchte ich Sie also, wie in einem Quiz, nach der Bedeutung oder dem Sinn von Wörtern oder Begriffen fragen, Ja?“

(Anleitung aus HAWIE-R S. 56-57) „Ich möchte, dass Sie mir die Bedeutung einiger Wörter nennen. Beginnen wir mit dem Wort ‚Apfel‘? Was bedeutet ‚Apfel‘?“

Die Antwort des ersten Begriffes Apfel wird den Probanden als Lösungsbeispiel vorgegeben, etwa: „Der Begriff Apfel steht für eine bestimmte Obstsorte.“

1. Apfel	17. Parlament
2. Beginn	18. Mandant
3. schleunigst	19. Deformation
4. Knie	20. vulgär
5. kriechen	21. Prestige
6. Pelz	22. Resistenz
7. Mikroskop	23. Potential
8. anonym	24. vehement
9. Gewebe	25. adäquat
10. Kerbe	26. Nihilismus
11. Naht	27. Sequenz
12. Chance	28. pragmatisch
13. Kapsel	29. Determinante
14. Oase	30. Thorium
15. Beklemmung	31. konkordant
16. Andeutung	32. Geoid

(Wortliste des Wortschatztests von S. 57.)

Folge der Instruktion des Handbuches.

Notiere, ob eine korrekte Antwort auf die Frage gegeben wurde anhand der Lösungen des Handbuches auf dem Bogen zur Dokumentation des Telefoninterviews.

Beginne bei dem Begriff Oase. Falls dieser Begriff und die folgenden nicht gelöst werden, frage nach Begriffen einfacheren Begriffen auf der Liste.

„**Vielen Dank.** Somit haben wir das Ende der Studie erreicht und ich bedanke mich sehr für Ihre Teilnahme.“

Zusicherung einer allgemeinen Ergebnismrückmeldung, etwa: „Ich werde ihnen gerne zum Ende der Untersuchung einen Überblick über die gewonnenen Ergebnisse zusenden.“

Eventuelle Fragen beantworten und Gespräch in angenehmer Atmosphäre beenden.

D.2 Dokumentation des Telefoninterviews

Termin: _____ Telefonnummer: _____

Code: _____ Gruppe: _____

Fehlende Daten: _____

HAWIE-R

Oase	
Mandant	23
Prestige	24
Adäquat	25
Nihilismus	26
Sequenz	27
Pragmatisch	28
Determinante	29
Thorium	30
Konkordant	31
Geoid	32

E Bewertungskategorien für TWR und AWR

E.1 TWR

Um die Höchstwertung von **5 Punkten** zu erhalten, muss eine Aussage einsichtig und tiefgründig sein, Ichbezogenheit überwinden und folgende Kerneigenschaften von Weisheit zeigen: Erkennen der Komplexität und Grenzen des Wissens, Integration von Denken und Fühlen sowie philosophische und spirituelle Tiefe.

Eine Wertung mit **4 Punkten**, zeichnet sich durch einige dieser Kernpunkte aus, hat aber weniger klare transzendente Bezüge. Die Information darüber, wie Personen ihre Weisheit erlangten, hilft in einigen Fällen zu entscheiden, wie tief oder oberflächlich diese Erfahrung einzustufen ist.

Eine **3 Punkt**-Wertung erfolgt, wenn die Antwort ein Aspekt von Weisheit in einem eingeschränkten persönlichen Rahmen beschreibt, wie Toleranz, Nachdenklichkeit oder Eigenständigkeit. In einigen Fällen erhalten Antworten diese Wertung, die einen philosophischen Standpunkt vertreten, bei denen angezweifelt wird, dass er dem jeweiligen Teilnehmer zuzuschreiben ist oder schlicht zu oberflächlich beschrieben wurde.

2 Punkte erhalten Antworten, die verkümmertes Persönlichkeitswachstum offenbaren und oberflächliche, zynische oder verzweifelte Einstellungen vertreten.

Lediglich **1 Punkt** erhalten Aussagen, die einen extremen Mangel an Selbsttranszendenz oder Ego-Entwicklung aufweisen, wie bitterer Egozentrismus. Des Weiteren erhalten Antworten einen 1 Punkt, wenn eine klare Unfähigkeit im Umgang mit der Frage festgestellt wird.

E.2 AWR

Emotionales Verständnis

Es gibt **5 Punkte** in „**Emotionalem Verständnis**“ für eine emotional offene Antwort in der Verständnis, Mitgefühl und Respekt für die Lage des fiktiven Freundes hoch ausgeprägt sind. **4 Punkte** erhält eine Antwort, die die Kriterien für eine 5er Wertung nur knapp verfehlt. **3 Punkte** belegen ein unelaboriertes Maß an Mitgefühl und emotionalem Verständnis. **2 Punkte** liegen vor, wenn nur eine dieser Formen gezeigt wird. Lediglich **1 Punkt** erhalten Antworten, die nahezu rein rationale Einstellungen repräsentieren.

Moralische Komplexität

Es gibt **5 Punkte** in „**Moralischer Komplexität**“ wenn diskutiert wird, ob Menschen das Recht haben unter bestimmten Bedingungen Selbstmord zu begehen, oder welche Form der Hilfe man einer solchen Person geben darf. Die Diskussion sollte Kognition und Affekt integrieren. **4 Punkte** werden vergeben, wenn nicht das notwendige Niveau für eine höhere Wertung vorliegt. **3 Punkte** erzielt eine klare unelaborierte moralische Aussage. **2 Punkte** erhält eine Antwort, wenn eine moralische Komponente mitschwingt. Eine **1-Punkt-Antwort** tangiert die Moralität eines Suizids oder Hilfestellung zu dieser nicht. Dies tritt beispielsweise dann ein, wenn eine schwere körperliche Krankheit als möglicher Grund für das Verhalten genannt wird, moralische Aspekte jedoch nicht ausgeführt werden. In diesem Falle ist schwere körperliche Krankheit auf dem Kriterium „**Kognitiver Differenzierung**“ zu scoren.

Kognitive Differenziertheit

Es gibt **5 Punkte** in „**Kognitiver Differenziertheit**“ für das Diskutieren verschiedener Kontexte eines Selbstmordes, (wie beispielsweise Depression, Diagnose einer tödlichen Krankheit, Drogenmissbrauch, Verlust, Krisen, etc.) mit Verknüpfung adäquater Reaktionsweisen des Teilnehmers. **4 Punkte** erhält eine vom Niveau her schwächere Antwort. Eine **Mittlere Bewertung** erfolgt, wenn lediglich ein Grund exemplarisch diskutiert wird, oder eine elaborierte Nachfrage nach den Beweggründen angekündigt wird. **2 Punkte** entsprechen unelaborierten Nachfragen nach den Beweggründen. **1 Punkt** gibt es, wenn nicht nach Ursachen für den Anruf oder die Tat gefragt wird.

Prozedurales Wissen

Es gibt **5 Punkte** in „**Prozeduralem Wissen**“ für die Diskussion hilfreicher Gesprächsinhalte, verschiedener Handlungsalternativen (bspw. dem Freund zuhören), diverser Hilfestellen (Notruf der Polizei, Psychologen, Telefonseelsorge, etc.) und die Beurteilung der zeitlichen Dringlichkeit. **4 Punkte** werden erreicht, wenn die Nutzung fremder Hilfsinstanzen als Option in Erwägung gezogen wird. Es gibt **3 Punkte** für ein persönliches elaboriertes Hilfsangebot, das jedoch fremde Hilfe außer Acht lässt. **2 Punkte** erlangt ein unelaboriertes Hilfsangebot. **1 Punkt** wird erreicht, wenn kein Hilfsangebot erfolgt.

F Tabellen zur weiteren Übersicht der statistischen Auswertung

F.1 Matchingprofiltabellen

Tabelle F.1 Matchingprofile der Paarungen Nominierte & Kontrollperson.

Gruppe ^a	Geschlecht ^b	Alter	Beruf	Bildungsniveau ^c
N	M	34	Ingenieur	5
K	M	37	Ingenieur	5
N	M	38	Landwirtschaftstechniker	3
K	M	33	Landwirtschaftstechniker	3
N	M	40	Wirtschaftsingenieur	5
K	M	40	Wirtschaftsingenieur	5
N	M	42	Ingenieur	5
K	M	38	Ingenieur	5
N	M	40	Beamter	4
K	M	42	Beamter	4
N	M	41	Personalfachkaufmann	3
K	M	48	Industriekaufmann	3
N	M	48	Mediziner	5
K	M	47	Mediziner	6
N	M	48	Heilpraktiker	3
K	M	49	Heilpraktiker	3
N	M	52	Ingenieur	5
K	M	51	Ingenieur	5
N	M	52	Lehrer	7
K	M	58	Lehrer	7
N	M	58	Volkswirt	5
K	M	59	Volkswirt	5
N	M	59	Architekt	5
K	M	59	Architekt	5
N	M	59	Ingenieur	5
K	M	59	Ingenieur	5
N	M	62	Pfarrer	6
K	M	64	Pfarrer	6

Anmerkungen. ^a N indiziert den Nominierten, K die Kontrollperson eines Paares. ^b m = männlich, w = weiblich.

^c Das Bildungsniveau wird in 7 Abschlüssen erfasst: 1 = ohne Abschluss; 2 = Haupt-/Volksschule; 3 = mittlere Reife; 4 = Abitur /Fachabitur /Fachoberschule; 5 = Hochschule; 6 = Promotion; 7 = Habilitation.

Tabelle F.1 (Fortgesetzt). Matchingprofile der Paarungen Nominierte & Kontrollperson.

Gruppe ^a	Geschlecht ^b	Alter	Beruf ^c	Bildungsniveau ^d
N	M	63	Selbstständiger	6
K	M	60	Selbstständiger	6
N	M	64	Psychologe	5
K	M	64	Psychologe	6
N	M	64	Pensionär	6
K	M	67	Pensionär	2
N	M	65	Installateur	2
K	M	64	Installateur	4
N	M	69	Mathematiker	5
K	M	65	Mathematiker	5
N	M	72	Elektroinstallateur	2
K	M	74	Elektroinstallateur	4
N	M	73	Mathematiker	5
K	M	66	Mathematiker	5
N	W	24	Psych. Studentin	4
K	W	25	Psych. Studentin	4
N	W	26	Bürokauffrau	3
K	W	28	Bürokauffrau	4
N	W	34	Architektin	4
K	W	32	Bauzeichnerin	5
N	W	34	Heilpraktikerin	4
K	W	38	Heilpraktikerin	3
N	W	36	Bürokauffrau	4
K	W	40	Bürokauffrau	3
N	W	40	Betriebswirtin	5
K	W	37	Industriekauffrau	5
N	W	55	Physikerin	5
K	W	59	Physikerin	6
N	W	56	Psychologin	4
K	W	54	Psychologin	5
N	W	56	Psychologin	7
K	W	59	Psychologin	7

Anmerkungen. ^a N indiziert den Nominierten, K die Kontrollperson eines Paares. ^b m = männlich, w = weiblich.

^c Das Bildungsniveau wird in 7 Abschlüssen erfasst: 1 = ohne Abschluss; 2 = Haupt-/Volksschule; 3 = mittlere Reife; 4 = Abitur /Fachabitur /Fachoberschule; 5 = Hochschule; 6 = Promotion; 7 = Habilitation.

Tabelle F.1 (Fortgesetzt). Matchingprofile der Paarungen Nominierte & Kontrollperson.

Gruppe ^a	Geschlecht ^b	Alter	Beruf ^c	Bildungsniveau ^d
N	W	57	Systemanalytikerin	5
K	W	58	Lehrerin	5
N	W	64	Heilpraktikerin	2
K	W	68	Heilpraktikerin	3
N	W	68	Medizinerin	5
K	W	55	Medizinerin	5
N	W	70	Bankkauffrau	3
K	W	69	Bankkauffrau	3
N	W	70	Biotechnikerin	3
K	W	68	Biotechnikerin	4
N	W	72	Kauffrau	2
K	W	74	Kauffrau	2
N	W	75	Nonne	4
K	W	71	Nonne	2
N	W	85	Schriftstellerin	3
K	W	85	Schriftstellerin	5

Anmerkungen. ^a N indiziert den Nominierten, K die Kontrollperson eines Paares. ^b m = männlich, w = weiblich.

^c Das Bildungsniveau wird in 7 Abschlüssen erfasst: 1 = ohne Abschluss; 2 = Haupt-/Volksschule; 3 = mittlere Reife; 4 = Abitur /Fachabitur /Fachoberschule; 5 = Hochschule; 6 = Promotion; 7 = Habilitation.

F.2 Mittelwertstabellen der Weisheitstests

Tabelle. F.2 Mittelwerte verschiedener Teilnehmergruppen auf den Weisheitstestverfahren

Gruppierung	<i>M</i>	<i>N</i>	<i>SD</i>	<i>SE</i>
PWS				
Gesamt	13,95	111	2,22	,21
NOM	14,06	36	2,50	,42
KG	14,00	36	1,90	,32
PSY	13,64	36	2,28	,38
Männlich	13,42	59	2,34	,30
Weiblich	14,53	49	1,97	,28
Erfolgstyp	14,65	37	1,86	,31
Suchende	13,00	20	2,68	,60
TWR				
Gesamt	3,74	102	,75	,07
NOM	3,57	30	,82	,15
KG	3,77	30	,77	,14
PSY	3,97	31	,71	,13
Männlich	3,53	55	,77	,10
Weiblich	3,98	45	,72	,11
Erfolgstyp	3,89	37	,77	,13
Suchende	3,82	17	,88	,21
AWR				
Gesamt	9,82	114	2,87	,27
NOM	9,21	38	2,77	,45
KG	9,84	38	2,80	,45
PSY	10,39	38	2,97	,48
Männlich	9,39	62	2,78	,35
Weiblich	10,10	49	2,89	,41
Erfolgstyp	10,26	38	2,92	,47
Suchende	9,85	20	3,15	,70

Tabelle. F.2 (Fortgesetzt). Mittelwerte verschiedener Teilnehmergruppen auf den Weisheitstestverfahren

Gruppierung	<i>M</i>	<i>N</i>	<i>SD</i>	<i>SE</i>
3DWS				
Gesamt	,00	114	,88	,08
NOM	,10	38	,84	,14
KG	-,03	38	,90	,15
PSY	-,08	38	,90	,15
Männlich	-.00	62	,86	,11
Weiblich	,04	49	,88	,13
Erfolgstyp	,35	38	,81	,13
Suchende	,08	20	,83	,19
SAWS				
Gesamt	171,07	113	20,90	1,97
NOM	169,30	37	17,69	2,91
KG	169,89	37	25,74	4,23
PSY	173,41	37	19,06	3,13
Männlich	169,06	62	21,06	2,68
Weiblich	172,46	48	20,59	2,97
Erfolgstyp	176,50	38	18,44	2,99
Suchende	174,00	20	19,64	4,39

F.3 Korrelationstabellen der Weisheitstests

Tabelle F.3.1 Korrelationen der Weisheitstests mit den Konstrukten der Konvergenten und Diskriminanten Validierung in der Gesamtgruppe.

	3d-WS	SAWS	PWS	TWR	AWR
PW	<u>.47^{††}</u>	<u>.63^{††}</u>	.05	.16	.03
PF	<u>.49^{††}</u>	.38 ^{††}	.28 ^{††}	.00	-.03
WUSCT	.29 ^{††}	.23 [†]	.09	.31 ^{††}	.43 ^{††}
UK	<u>.60^{††}</u>	.24 ^{††}	.29 ^{††}	.05	-.07
Wortschatz	.14	.12	-.01	-.09	-.03
Alter	.19 [†]	.21 [†]	.14	.03	-.02
Geschlecht	-.05	.14	.34 ^{††}	.36 ^{††}	.14

Anmerkungen. Listenweiser Fallausschluss $n = 95$. Unterstrichen sind Korrelationen die signifikant über .30 liegen. Signifikant unter diesem Niveau korrelieren die fettgedruckten Korrelationen. Alle übrigen signifikanten Korrelationen unterscheiden sich nicht signifikant von diesem Niveau.

^{††} $p < .01$ (einseitig); [†] $p < .05$ (einseitig).

Tabelle F.3.2 Korrelationen der Weisheitstests mit den Konstrukten der Konvergenten und Diskriminanten Validierung in der Nominiertengruppe.

	3d-WS	SAWS	PWS	TWR	AWR
PW	.47 ^{††}	.49 ^{††}	.16	.24	.11
PF	<u>.63^{††}</u>	.54 ^{††}	.44 ^{††}	.12	.07
WUSCT	.27	.21	.28	.34 [†]	.33 [†]
UK	<u>.76^{††}</u>	.43 ^{††}	.27	.03	.14
Wortschatz	-.01	-.18	.05	-.42^{††}	-.10
Alter	.11	.24	.14	-.13	-.14
Geschlecht	-.01	.21	.24	.40 [†]	.34 [†]

Anmerkungen. Listenweiser Fallausschluss $n = 32$. Unterstrichene Korrelationen liegen signifikant über .30. Aufgrund der wenigen Teilnehmer, unterschreiten dieses Niveau nur jene Korrelationen, die fettgedruckt sind signifikant. Alle übrigen Korrelationen unterscheiden sich nicht signifikant von diesem Niveau.

^{††} $p < .01$ (einseitig); [†] $p < .05$ (einseitig).

Tabelle F.3.3 Korrelationen der Weisheitstests mit den Konstrukten der Konvergenten und Diskriminanten Validierung in der Kontrollgruppe.

	3d-WS	SAWS	PWS	TWR	AWR
PW	.39 ^{††}	<u>.75^{††}</u>	-.16	.04	-.09
PF	.43 ^{††}	.18	.13	-.17	-.24
WUSCT	.21	.19	.03	.16	.48 ^{††}
UK	.49 ^{††}	-.02	.09	.01	-.32[†]
Wortschatz	.28	.20	-.10	.00	-.21
Alter	.13	.07	.26	.38 [†]	.14
Geschlecht	-.10	.37 [†]	.24 [†]	.40 [†]	.34 [†]

Anmerkungen. Listenweiser Fallausschluss $n = 32$. Unterstrichene Korrelationen liegen signifikant über .30. Aufgrund der wenigen Teilnehmer, unterschreiten dieses Niveau nur jene Korrelationen, die fettgedruckt sind signifikant. Alle übrigen Korrelationen unterscheiden sich nicht signifikant von diesem Niveau.

^{††} $p < .01$ (einseitig); [†] $p < .05$ (einseitig).

Tabelle F.3.4 Korrelationen der Weisheitstests mit den Konstrukten der Konvergenten und Diskriminanten Validierung in der Psychotherapeutengruppe.

	3d-WS	SAWS	PWS	TWR	AWR
PW	<u>.63^{††}</u>	<u>.57^{††}</u>	.18	.26	.12
PF	.45 ^{††}	<u>.58^{††}</u>	.28	.18	.21
WUSCT	.39 [†]	.25	-.04	.36 [†]	.43 ^{††}
UK	<u>.58^{††}</u>	.42 ^{††}	<u>.56^{††}</u>	.11	.00
Wortschatz	.18	.30	.09	-.22	.26
Alter	.36	.45 ^{††}	-.01	-.03	.04
Geschlecht	.04	-.21	.40 [†]	.26	-.08

Anmerkungen. Listenweiser Fallausschluss $n = 31$. Unterstrichene Korrelationen liegen signifikant über .30. Aufgrund der wenigen Teilnehmer, unterschreiten dieses Niveau nur jene Korrelationen, die fettgedruckt sind signifikant. Alle übrigen Korrelationen unterscheiden sich nicht signifikant von diesem Niveau.

^{††} $p < .01$ (einseitig); [†] $p < .05$ (einseitig).

Tabelle F.3.5 Korrelationen der Weisheitstests mit den Konstrukten der Konvergenten und Diskriminanten Validierung für den Typus der Suchenden.

	3d-WS	SAWS	PWS	TWR	AWR
PW	.47 [†]	.45	-.08	.13	.14
PF	.42	.41	.04	-.27	-.05
WUSCT	.44	.26	.22	.58 [†]	.27
UK	.57 [†]	.18	-.00	-.63^{††}	-.52[†]
Wortschatz	.32	-.08	-.11	.01	.14
Alter	.04	.60 ^{††}	.05	.17	.16
Geschlecht	-.23	-.10	.60 [†]	.29	.27

Anmerkungen. Listenweiser Fallausschluss $n = 14$. Aufgrund der wenigen Teilnehmer, unterschreiten das Niveau von $r = .30$ nur jene Korrelationen, die fettgedruckt sind signifikant. Keine Korrelation übersteigt dieses Niveau signifikant. Alle übrigen Korrelationen unterscheiden sich nicht signifikant von diesem Niveau.

^{††} $p < .01$ (einseitig); [†] $p < .05$ (einseitig).

Tabelle F.3.6 Korrelationen der Weisheitstests mit den Konstrukten der Konvergenten und Diskriminanten Validierung für den Typus der Erfolgreichen.

	3d-WS	SAWS	PWS	TWR	AWR
PW	.36 [†]	<u>.55^{††}</u>	-.19	.10	-.02
PF	.40 ^{††}	.13	.41 ^{††}	-.06	-.06
WUSCT	.17	.08	-.07	.17	.49 ^{††}
UK	.10	.31 [†]	.20	-.04	-.22
Wortschatz	-.08	.05	.16	-.11	-.18
Alter	.34 [†]	.25	.11	.02	-.07
Geschlecht	-.16	.19	.23	.40 ^{††}	.33 [†]

Anmerkungen. Listenweiser Fallausschluss $n = 36$. Unterstrichene Korrelationen liegen signifikant über .30. Aufgrund der wenigen Teilnehmer, unterschreiten dieses Niveau nur jene Korrelationen, die fettgedruckt sind signifikant. Alle übrigen Korrelationen unterscheiden sich nicht signifikant von diesem Niveau.

^{††} $p < .01$ (einseitig); [†] $p < .05$ (einseitig).

G Rückmeldung für die Teilnehmer



Ansprechpartnerin:
Marina Hoffmann
Godramsteinerstr. 8
76829 Landau in der Pfalz

Telefon: 06341/707769
E-Mail: hoffmann.marina@web.de

Betreuung:
Prof. Dr. Manfred Schmitt
Universität Koblenz-Landau,
Campus Landau
Im Fort 7
76829 Landau in der Pfalz
Telefon: 06341/280-495
Telefax: 06341/280-490

<Datum>

Sehr geehrte/r Studienteilnehmer / Studienteilnehmerin,

Erinnern Sie sich noch an die Studie zu Weisheit, an der Sie teilgenommen haben?

Dieser Brief informiert Sie über die zur „Weisheitsmessung“ verwendeten Testverfahren des von Ihnen bearbeiteten Fragebogenheftes, darüber wie Weisheit in der Psychologischen Forschung definiert wird und welche Erkenntnisse anhand der Studienergebnisse für Ihre Teilnehmergruppe gewonnen wurden.

Die Fragestellung, ob Weisheit über eine schriftliche Erhebung zu erfassen ist, initiierte das Forschungsvorhaben. Fragebogenerhebungen stellen eine ökonomische Form der Datengewinnung dar. Ob sich ein Fragebogen jedoch zur Messung des betreffenden „Konstruktes“ eignet, ist eine Frage der Validität: „Misst der Test das Konstrukt, was er zu messen vorgibt?“. Im Gegensatz zu Eichungen von Testverfahren, die an mehreren Tausend Probanden erfolgen, bei denen „Normwerte“ für verschiedene Altersgruppen etc. bestimmt werden, bestimmen Validierungsstudien die Entsprechung der Testwerte zu:

- a) inhaltlichen, theoriegeleiteten Annahmen (Inhaltsvalidität)
- b) auf anderem Wege feststellbaren Merkmalen des betreffenden Konstruktes (Kriteriumsvalidität)
- c) theoretischen Annahmen über die Enge der Beziehungen (Korrelation) zu weiteren verwandten oder nicht verwandten Konstrukten (Konstruktvalidität)

Die Studie zur Validierung der Weisheitsfragebögen thematisierte vor allem die Punkte b) und c). Nach eingehender Literaturanalyse wurden fünf der insgesamt sieben bekannten schriftlichen Weisheitstests in dieser Studie verwendet.

- 1) 3-dimensionale Weisheitsskala (3d-WS- Ardelt, 2003) - eröffnete das Fragebogenheft.
- 2) Selbstadministrative Weisheitsskala (SAWS - Webster, 2003) - war vor der Satzergänzungsaufgabe im Fragebogenheft lokalisiert.
- 3) Praktische Weisheitsskala (PWS – Wink & Helson, 1997) – erfragte Ja - Nein Antworten auf der oberen Hälfte der vorletzten Seite.
- 4) Transzendentes Weisheitsrating (TWR – Wink & Helson, 1997) - bewertet die Einzelfrage nach einem Beispiel eigener Weisheit.
- 5) Allgemeines Weisheitsrating (AWR – Helson & Srivastava, 2001, 2002) – bewertet die Einzelfrage nach Gedanken und Reaktionen auf einen Anruf eines Freundes, der angab sich das Leben zu nehmen.

** Auf der letzten Seite finden Sie die detaillierten Angaben zu diesen Publikationen.*

Diese fünf Testverfahren werden in Form einer Tabelle einzelnen Teilbereichen zugeordnet, die insgesamt ein eher breites Verständnis der Weisheit widerspiegeln. Weisheit beinhaltet nach Meinung von Forschern und Laien viele Facetten, die wie folgt vereinfachend zusammengefasst werden können.

Weisheit bedarf der Reflektion der eignen Erfahrungen. Diese ermöglicht es, zu Einsichten zu kommen, die das eigene Verhalten erklären, und wiederum auf andere Situationen übertragen zu können. An dieser Stelle spielen Werte, philosophische und religiöse Grundsätze ebenfalls eine Rolle (Werteorientierung). In der Qualität und Tragweite dieser Übertragungen, entweder für sich oder für andere genutzt, wird Weisheit offenbart. Dazu ist eine gewisse geistige Befähigung (Kognitive Komponente der Weisheit) notwendig, um die Angemessenheit einer Übertragung und die daraus zu ziehenden Konsequenzen (Handlungskompetenz) richtig einzuschätzen. Bis zu dieser Stelle beschreiben die Ausführungen ein analytisches Modell der Weisheit. Wenn der weise Mensch als Ganzes (synthetisches Modell) betrachtet wird, so sind dessen Persönlichkeit und gefühlsmäßige Seite (Emotionale Komponente) mit zu berücksichtigen.

	3d-WS	SAWS	PWS	TWR	AWR
Reflektion	Reflektion	Reflektion & Lebens- erfahrung	Praktische Weisheit	Transzendenz eigener Erfah- rungen	
Kognitive Komponente	Kognitive Komponente				Kognitive Komponente
Werte- orientierung	Wert- schätzung				Moral
Handlungs- kompetenz					Handlungs- kompetenz
Persön- lichkeit		Offenheit & Humor			
Emotionale Komponente	Mitgefühl & Verständnis	Stimmungs- regulation			Mitgefühl & Verständnis

Es wird für Sie ersichtlich, dass keiner der Tests alle Teilbereiche abdeckt. Dies beruht auf dem jeweiligen Weisheitsmodell des Entwicklers. In Anbetracht der Aufgabe Weisheit zu erfassen, erscheint die Beschränkung auf einzelne Teilbereiche fraglich. Die empirischen Ergebnisse der Studie spiegeln das Bild obiger Tabelle jedoch nicht wider. Die größte Nähe zeigen TWR und AWR, sowie 3d-WS und SAWS. Die PWS zeigt wenig Übereinstimmung mit den anderen Tests. Nachfolgend werden einige Befunde diskutiert.

Die beiden Testverfahren TWR und AWR, die eine gewisse Selbstoffenbarung eigener Gedanken und Lebenserfahrungen verlangten, wurden von der Mehrheit der Teilnehmer eher knapp oder undetailliert beantwortet. Eine Auswertung anhand des ursprünglichen Manuals war aufgrund der Mindestantwortlänge von 70 Worten nicht möglich. Über eine Anpassung des Manuals konnten Einschätzungen zu den jeweiligen Bewertungskategorien vorgenommen werden. Im Sinne der Validität – Messen TWR und AWR das, was sie zu messen vorgeben? - muss für diese Studie die Einschränkung ergehen, dass die Gültigkeit der zu messenden Konstrukte abhängig von der Bereitschaft zur Selbstoffenbarung der Teilnehmer war. Als weitere Empfehlung sollte hier entweder die Instruktion der Testverfahren verbessert werden, um ausführlichere Antworten zu erhalten, oder es sollte von einer schriftlichen Erhebung abgesehen werden.

In Folge dieser Beschränkung entsprachen die Ergebnisse der Konstruktvalidierung nicht allen

theoretisch abgeleiteten Hypothesen, während die anderen drei Testverfahren hier mehr Hypothesenkonformität bewiesen. Zu dieser Konstruktvalidierung dienten als der Weisheit verwandte Konzepte:

- 1) Ego-Entwicklung (Loevinger, 1970) - Satzergänzungsaufgabe
- 2) Persönliches Wachstum (Ryff & Keyes, 1992) – zwischen 3d-WS und SAWS
- 3) Psychologisches Feingefühl (IPIP, 2006) – unterer Teil der vorletzten Seite.

Die Konstrukte, die keinen Einfluss auf Weisheit haben sollten, sind:

- 1) Umweltkontrolle (Ryff & Keyes, 1992) – zwischen 3d-WS und SAWS
- 2) Wortschatz als Indikator der Intelligenz (Tewes, 1995) - telefonisch erhoben
- 3) Selbsteinschätzung von Intelligenz (Rammstedt & Rammsayer, 2002) – telefonisch erhoben
- 4) Alter
- 5) Geschlecht

Die Hypothese der Unabhängigkeit von Weisheit und Umweltkontrolle konnte für die drei Tests 3d-WS, SAWS und PWS nicht bestätigt werden. Die Mehrheit der Teilnehmer erreichten hier sehr hohe Werte. Aus statistischer Sicht muss daher angenommen werden, dass der Test zu dieser Eigenschaft zu leicht war. Fazit der Daten ist, dass Umweltkontrolle und Weisheit häufig in gleicher Weise auftreten.

Für Intelligenz belegen die Daten, dass diese unabhängig von der Weisheit der Teilnehmer ist. Der einzige Unterschied zwischen Männern und Frauen in den Weisheitstests ist, dass Frauen bessere Werte auf der PWS erzielen. Die Größenordnung dieses Effektes nimmt statistisch betrachtet ein mittleres Niveau an. Die Analysen zu Weisheit und Alter zeigen einen leichten Effekt für die PWS. Somit schreiben sich die älteren Teilnehmer im Vergleich zu den jüngeren mehr Praktische Weisheit zu.

Für die Fragestellung der Kriteriumsvalidierung wurden drei unterschiedlichen Teilnehmergruppen rekrutiert. Als subjektives, von einer Einschätzung abhängiges Merkmal wurden im Bekanntenkreis der als Nominatoren angesprochenen Personen jene nominiert, die in deren Augen als „Weise“ bezeichnet werden können. Derartige und ähnliche Nominierungsverfahren sind in der Weisheitsforschung nicht selten. Zu den Nominierten wurde nach Alter, Geschlecht, Beruf und Bildungsniveau entsprechende Kontrollpersonen rekrutiert. Als objektives, da offen erkennbares Merkmal, wurde der Beruf Psychotherapeut gewählt. Diese Berufsgruppe zeigte in einigen Studien zu Allgemeiner Weisheit, Transzendenter Weisheit und Praktischer Weisheit, dass sie für Validierungsstudien interessant sind. Diese Berufsausbildung und Praxis fördert analytische Kompetenzen und Bewältigungsstrategien im Umgang mit diversen menschlichen Problemen. Die Gruppenmittelwerte der Weisheitstestverfahren für die drei Teilnehmergruppen zeigen einheitlich hohes Niveau. Die Annahmen werden gestützt durch die Einschätzung der Ego-Entwicklung (Satzergänzungstest) für die aus Testeichungen Normwerte vorliegen. Demnach sind 86% aller Teilnehmer mindestens auf dem als „Gewissenhaft“ deklarierten Niveau, welches eine Stufe über dem Durchschnitt der Allgemeinbevölkerung liegt. 55% der Teilnehmer erreichen mindestens das nächste, das „Individuelle“ Niveau. Darüber hinaus werden die in der Allgemeinbevölkerung sehr seltenen oberen Kategorien der „Autonomie“ und „Integrität“ von 15% der Teilnehmer erreicht. Somit ergeht an die Teilnehmer der Studie die Gratulation zu hoher Persönlichkeitsreife, die in einigen Fällen höchstens Niveau erreicht und Weisheit verkörpert.

Eine gewisse Unterschiedlichkeit zwischen den Gruppen ließ sich in Bezug auf die Übereinstimmung der Testverfahren feststellen. Für die Nominiertengruppe und die Psychotherapeuten-Gruppe zeigten sich die Weisheitsfragebogen übereinstimmender als für die Kontrollgruppe. Vereinfacht ausgedrückt, ist die Anzahl von Personen in der Nominiertengruppe und der Psychotherapeuten-Gruppe höher als in der Kontrollgruppe, für die folgendes Muster zutrifft: hohe Werte auf einem Test gehen mit hohen Werten auf den anderen Tests einher, entsprechend gehen niedrige Werte auf einem Test mit niedrigen Werten auf einem anderen Test einher. Die Kontrollgruppe zeigte mehr Profile mit gegensätzlichen Verläufen, die sich in Summe des Gruppenwertes aufhoben. Statistisch benannt, zeigen die Tests für die unterschiedlichen Gruppen „differentielle Validität“. Dies bedeutet, dass eine Prognose eines Weisheitswertes auf einem der fünf Tests über die Kenntnis eines oder mehrerer Testergebnisse der anderen Tests für die Nominierten und die Psychotherapeuten leichter fällt als für die Kontrollpersonen.

Aufgrund der theoretischen Sichtweise, dass alle Komponenten der Weisheit im hohen Maße vorhanden sein müssen, um einer Person Weisheit zu attestieren, sollte auf allen Testverfahren hohe Werte erreicht werden. Im Anbetracht der Differentiellen Validität, trifft dies somit eher auf Personen zu, die entweder das subjektive oder das objektive Merkmal der Kriteriumsvalidierung erfüllen, also entweder von einem ihrer Bekannten für ihre Weisheit nominiert wurden, oder Psychotherapeuten sind.

Abschließend möchte Ich nochmals herzlichst meine Dankbarkeit für Ihre freundliche Unterstützung meiner Diplomarbeit entgegenbringen. Durch Ihre Bereitschaft ein langes Fragebogenheft zu bearbeiten, telefonisch für ein Interview zur Verfügung zu stehen, und mir Ihre persönlichen Rückmeldungen oder Definitionen von Weisheit zu liefern, war diese Studie erst möglich. Insbesondere durch Ihre hohen Werte in Persönlichkeitsreife und Weisheit wurde die Qualität dieser Studie gesichert. Schließlich war die Zielsetzung, zu Prüfung, ob ein Weisheitstest das misst, was er zu messen vorgibt. Ich habe von Ihrer Mitteilungsbereitschaft nicht nur in Form dieser Arbeit profitiert und wünsche Ihnen daher, weiterhin wertvolle Erkenntnisse erlangen zu können.

- Ardelt, M. (2003). Empirical assessment of a three-dimensional wisdom scale. *Research on Aging*, 25, 3, 275-324.
- Helson, R. & Srivastava, S. (2001). Three Paths of Adult Development: Conservers, Seekers and Achievers. *Journal of Personality and Social Psychology*, 80, 6, 955-1010.
- Helson, R. & Srivastava, S. (2002). Creative and Wise People: Similarities, Differences and How they Develop. *Personality and Social Psychology Bulletin*, 28, 10, 1430-1440
- Wink, P. & Helson, R. (1997). Practical and Transcendent Wisdom: Their Nature and Some Longitudinal Findings. *Journal of Adult Development*, 4, 1, 1-14.
- Webster, J. D. (2003). An exploratory analysis of a self-assessed wisdom scale. *Journal of Adult Development*, 10, 13-22.

Mit freundlicher Hochachtung

Marina Hoffmann