



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

**„Antworttendenzen und Verfälschungsverhalten in
Persönlichkeitsfragebögen:
Erfassung von Antworttendenzen innerhalb verschiedener
Verfälschungs-Gruppen mittels multidimensionaler IRT-Modelle“**

verfasst von

Victoria Grothe, Bakk.rer.soc.oec.

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag.rer.nat.)

Wien, 2013

Studienkennzahl lt. Studienblatt:
Studienrichtung lt. Studienblatt:
Betreut von:

A 298
Diplomstudium Psychologie
Mag^a. Dr. Lale Khorramdel

Danksagung

Zunächst möchte ich die Gelegenheit nutzen und mich bei meiner Betreuerin, Mag^a. Dr. Lale Khorramdel, bedanken. Sie hat einen entscheidenden Beitrag zu meinem Verständnis zum hier behandelten Thema geliefert und mir auf dem Weg zur fertigen Diplomarbeit stets mit Rat und Tat zur Seite gestanden.

Des Weiteren möchte ich mich bei meinen Eltern bedanken, die mir solange ich denken kann, aber insbesondere in den letzten drei Jahren jede Form von Unterstützung haben zukommen lassen und auf deren Hilfe ich mich zu jedem Zeitpunkt mehr als verlassen konnte. Ein wirklich gutes Gefühl.

Ausserdem gilt mein Dank Philipp, meinem Ehemann. Er hat mich mehr als einmal durch seinen klaren Blick und seine Fähigkeit, Dinge logisch zu hinterfragen aus dem Dschungel der Diplomarbeit gerettet und mir wieder festen Boden unter den Füßen verschafft.

Auch meiner Tochter Alma möchte ich danken. Denn auch wenn Sie dem Gelingen von Studium und Diplomarbeit nicht nur zuträglich war, so hat Sie mir doch immer wieder den Blick für das Wesentliche geschärft.

Wien, im September 2013

Abstract – English

The present study investigates the relation between faking behavior (as socially desirable response behavior) and trait-unrelated response styles (Extreme Response Style, ERS; Middle Response Style, MRS) when administering personality questionnaires in the context of job applications. Before the three experimental groups ($n_1=127$, $n_2=126$, $n_3=132$) consisting of law and economics students received an online personality questionnaire (consisting of BIP, NEO-FFI, Social Desirability Scale) they obtained different faking instructions. The first experimental group was instructed to show strong faking behavior. The second experimental group was expected to show medium faking behavior and the third experimental group was expected to show honest response behavior. The intended faking effect could be proven for one questionnaire scale (Social Desirability Scale). By using a new IRT approach it was investigated whether for this scale an extreme response style (ERS) or a middle response style (MRS) could be observed. In order to carry out this investigation the original responses of the test takers on a 5-point Likert Scale were decomposed into binary so-called pseudo items. Multi-dimensional IRT models were used in order to examine whether the response styles could be unidimensionally displayed across the original questionnaire scale. Whereas MRS could be displayed unidimensionally across all three experimental groups, for ERS this was the case only for the third experimental group. A relation between response styles and stereotypical faking behavior could therefore only be proven for MRS. The analysis of the remaining 19 scales revealed important insights into modelling response styles by using pseudo items d, e and m.

The present study firstly reveals helpful insights for future studies carrying out investigations on the relation between faking behavior and the existence of response styles. Secondly, it provides information for the scaling of personality questionnaires by allowing to correct for response styles and thereby receiving undistorted results.

Key words: Item Response Theory (IRT), unidimensional IRT models, multidimensional IRT models, faking behavior, social desirability, faking instruction, response style, response process, rating scale

Abstract –Deutsch

Die vorliegende Arbeit dient der Untersuchung des Zusammenhangs zwischen Verfälschungsverhalten (im Sinne sozial erwünschten Antwortverhaltens) und Iteminhalts-unabhängiger Antworttendenzen (Tendenz zu Extremurteilen, ERS; Tendenz zur Mitte, MRS) im Rahmen der Vorgabe von Persönlichkeitsfragebögen im Bewerbungskontext. Drei Versuchsgruppen ($n_1=127$, $n_2=126$, $n_3=132$) bestehend aus Studenten der Rechts- und Wirtschaftswissenschaften erhielten vor Bearbeitung eines Online-Fragebogens (bestehend aus BIP, NEO-FFI, Soziale-Erwünschtheits-Skala) verschiedene Verfälschungsinstruktionen. Während die erste Versuchsgruppe mit dem Ziel instruiert wurde, möglichst starkes Verfälschungsverhalten zu zeigen, wurde für die zweite Versuchsgruppe mittelstarkes Verfälschen und für die dritte Versuchsgruppe ehrliches Antwortverhalten erwartet. Der beabsichtigte Verfälschungseffekt konnte dabei lediglich für eine Fragebogenskala (Soziale-Erwünschtheits-Skala) nachgewiesen werden. Für diese Skala wurde mittels eines neuen IRT-Ansatzes geprüft, ob eine Tendenz zu Extremantworten (Extreme Response Style: ERS) oder eine Tendenz zur Mitte (Midpoint Response Style: MRS) vorliegt. Dabei wurden die ursprünglichen Antworten der Testpersonen auf einer 5-kategoriellen Ratingskala in binäre sog. Pseudo-Items zerlegt und es wurde mittels multidimensionaler IRT-Modelle geprüft, ob sich Antworttendenzen eindimensional über die Fragebogenskala hinweg abbilden lässt. Dabei konnte MRS in allen drei Versuchsgruppen, ERS jedoch nur in VG 3 als eigenständige Dimension abgebildet werden. Ein Zusammenhang zwischen Antworttendenzen und stereotypem Verfälschungsverhalten liess sich somit nur für die Tendenz zur Mitte (MRS) nachweisen. Die Analyse der übrigen 19 Skalen ergab aufschlussreiche Erkenntnisse zu den zur Modellierung der Antworttendenzen verwendeten Pseudo-Items d, e und m.

Die vorliegende Arbeit liefert (1.) hilfreiche Erkenntnisse für zukünftige Studien zur Überprüfung eines Zusammenhangs von Verfälschungsverhalten und dem Vorliegen von Antworttendenzen und (2.) Hinweise für die Skalierung von Persönlichkeitsfragebögen mit dem Ziel Antworttendenzen zu korrigieren und weitgehend unverzerrte Ergebnisse zu erhalten.

Inhaltsverzeichnis

I. EINLEITUNG	1
II. THEORETISCHER TEIL.....	5
1 Persönlichkeitsverfahren in der Personalauswahl	5
2 Kritische Betrachtung von Persönlichkeitsfragebögen.....	6
3 Response Bias in Persönlichkeitsfragebögen.....	9
3.1 Inhaltliche Verfälschung – Response Sets.....	10
3.1.1 Verfälschen als stereotypes vs. individuell variierendes Antwortverhalten.....	12
3.1.2 Methoden zur Erhebung von Verfälschungsverhalten	18
3.1.3 Verfälschungsinstruktionen zur Erzeugung und Untersuchung von Verfälschungsverhalten	20
3.1.4 Methoden zur Vermeidung sowie Berücksichtigung von Verfälschungsverhalten	22
3.2 Formale Verfälschung – Antworttendenzen (Response Styles)	25
3.2.1 Extreme Antworttendenz (ERS) und Tendenz zur Mitte (MRS)	28
3.2.2 Erhebung, Vermeidung und Berücksichtigung von ERS und MRS..	33
4 Zusammenhang von Antworttendenzen und stereotypem Verfälschungsverhalten.....	35
5 (Multidimensionale) Item-Response-Theorie-Modelle zur Untersuchung von Antworttendenzen	37
III. EMPIRISCHER TEIL.....	42
6 Ziel der Untersuchung	42
6.1 Hypothesen.....	42
7 Methode.....	45
7.1 Untersuchungsdesign	45
7.2 Erhebungsinstrumente	53
7.2.1 Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP)	54
7.2.2 NEO Fünf Faktoren Inventar (NEO-FFI)	58
7.2.3 Soziale-Erwünschtheits-Skala-17 (SES-17)	62
7.3 Durchführung der Untersuchung	63
7.4 Stichprobe	63
8 Ergebnisse zum Effekt der Verfälschungsinstruktionen.....	70
8.1 Ergebnisse zum subjektiven Effekt der Verfälschungsinstruktionen	70
8.1.1 Durchführung der Analyse und Darstellung der Ergebnisse	70
8.1.2 Interpretation der Ergebnisse	73
8.2 Ergebnisse zum tatsächlichen Verfälschungsverhalten	73
8.2.1 Durchführung der Analyse und Darstellung der Ergebnisse	74

8.2.2	Interpretation der Ergebnisse	80
9	Ergebnisse zu IRT- Analysen der Antworttendenzen	81
9.1	Ergebnisse zur Skala Soziale Erwünschtheit.....	85
9.1.1	Durchführung der Analyse.....	85
9.1.2	Darstellung der Ergebnisse	87
9.1.3	Interpretation der Ergebnisse	89
9.2	Ergebnisse zu den restlichen 19 Skalen.....	91
9.2.1	Durchführung der Analysen.....	91
9.2.2	Darstellung der Ergebnisse	92
9.2.3	Interpretation der Ergebnisse	109
10	Diskussion.....	117
11	Zusammenfassung.....	122
	Literaturverzeichnis.....	126
IV.	ANHANG	134
1	Anhang zum BIP	135
1.1	Skalen-Interkorrelationsmatrix BIP	135
2	Online-Fragebogen	136
2.1	Online-Fragebogen.....	136
3	Anhänge zur Analyse des Verfälschungsverhaltens	164
3.1	Skaleninterkorrelationsmatrizen (VG)	165
3.2	Differenzen der Skaleninterkorrelationen (VG)	174
4	Anhänge zur Analyse der Antworttendenzen mittels IRT-Modellen.....	181
4.1	Beispiel- Syntax der IRT-Analyse mittels mdlm-Software	182
4.2	Beispiel-Output der IRT-Analyse mittels mdlm-Software	185
4.3	Modellpassung der Skala Soziale Erwünschtheit (GSP)	208
4.4	Beobachtete Interkorrelations-Matrizen zu Modellvergleich I (d-Items, GSP).....	208
4.5	Marginale IRT-basierte Reliabilitäten der Faktoren zur Skala Soziale Erwünschtheit (VG und GSP)	210
4.1	Geschätzte Mittelwerte und Standardabweichungen der Faktoren der Skala Soziale Erwünschtheit (GSP)	211
4.2	Marginale IRT-basierte Reliabilitäten der Faktoren der übrigen 19 Skalen auf Basis der einzelnen VG (I-III).....	212
_Toc366527201		

Abbildungsverzeichnis

Abbildung 1: Ursachen und Einflussfaktoren von Antworttendenzen	29
Abbildung 2: Multipler Entscheidungsprozess - Entscheidungsbaum	39
Abbildung 3: Verteilung der Studienfächer	65
Abbildung 4: Geschlechterverteilung	65
Abbildung 5: Verteilung der Nationalitäten	69
Abbildung 6: Veranschaulichung der Mittelwerte Soziale Erwünschtheit	80

Tabellenverzeichnis

Tabelle 1: Response Bias - Antworttendenzen und Response Sets	10
Tabelle 2: Erhebungsmethoden von Verfälschungsverhalten	18
Tabelle 3: Startseite Online-Fragebogen	47
Tabelle 4: Verfälschungsinstruktionen	48
Tabelle 5: Interne Konsistenz BIP	57
Tabelle 6: Interne Konsistenz NEO-FFI	61
Tabelle 7: Altersverteilung	64
Tabelle 8: Verteilung der Semesteranzahl	66
Tabelle 9: Verteilung des höchsten Ausbildungsabschlusses	67
Tabelle 10: Verteilung der Berufserfahrung	68
Tabelle 11: Mittelwert und Standardabweichung zum subjektiven Effekt der Verfälschungsinstruktionen	70
Tabelle 12: Box-M-Test zur Überprüfung der Homogenität der Varianzen	71
Tabelle 13: T-Test für unabhängige Stichproben bei ungleichen Varianzen	72
Tabelle 14: Mittelwerte und Standardabweichungen aller 20 Skalen	74
Tabelle 15: Box-M-Test zur Überprüfung der Homogenität der Varianzen der 20 Dimensionen	76
Tabelle 16: Effekt des Dimensions-Ausschlusses auf Box-M-Test	77
Tabelle 17: T-Test für unabhängige Stichproben bei ungleichen Varianzen	77
Tabelle 18: Kodierung der Pseudo-Items d, e und m	85
Tabelle 19: Modellberechnungen Skala Soziale Erwünschtheit (VG und GSP)	86
Tabelle 20: Modellvergleiche der Skala Soziale Erwünschtheit (VG) auf Basis aller Pseudo-Items	87
Tabelle 21: Geschätzte Mittelwerte, Standardabweichungen und die marginalen Reliabilitäten zur Skala Soziale Erwünschtheit (VG) für die Modelle B und C	88
Tabelle 22: Modellvergleiche I-IV	91
Tabelle 23: Modellvergleiche I-III (GSP) auf Basis der einzelnen Pseudo-Items	93
Tabelle 24: Geschätzte IRT-basierte Mittelwerte, Standardabweichungen und die marginalen Reliabilitäten für Modelle A und B (I-III)	95
Tabelle 25: Modellvergleiche IV (GSP) unter Einbeziehung aller Skalen und Pseudo-Items	97
Tabelle 26: Geschätzte Mittelwerte, Standardabweichungen und die marginalen Reliabilitäten für Modelle A und C (IV)	98
Tabelle 27: Modellvergleiche I-III (AV) unter Betrachtung der einzelnen Pseudo- Items	100
Tabelle 28: Modellvergleiche I-III des Bereichs BO (BIP) unter Betrachtung der einzelnen Pseudo-Items	101

Tabelle 29: Modellvergleiche I-III des Bereichs PK (BIP) unter Betrachtung der einzelnen Pseudo-Items	103
Tabelle 30: Modellvergleiche I-III des Bereichs SK (BIP) unter Betrachtung der einzelnen Pseudo-Items	104
Tabelle 31: Modellvergleiche I-III auf Basis des NEO-FFI (NEO) unter Betrachtung der einzelnen Pseudo-Items	105
Tabelle 32: Geschätzte Skaleninterkorrelationen auf Basis der d- und Original-Items für die Bereiche AV, BO und PK (BIP) aus Modellvergleich I (GSP).....	106
Tabelle 33: Geschätzte Skaleninterkorrelationen auf Basis der d- und Original-Items für den Bereich SK (BIP) und NEO (NEO-FFI) aus Modellvergleich I (GSP)	108

I. EINLEITUNG

Die vorliegende Arbeit dient der Untersuchung des Zusammenhangs zwischen Verfälschungsverhalten (*Response Sets*: sozial erwünschtes Antwortverhalten, *Impression Management*) und inhaltsunabhängigen Antworttendenzen (*Response Styles*) in Persönlichkeitsfragebögen im Rahmen der Bewerberauswahl. Vor dem Hintergrund der starken Bedeutung der optimalen Bewerberauswahl und des steigenden Einsatzes von Persönlichkeitsverfahren im Rahmen dieser, ist der Aspekt der Verfälschbarkeit bzw. Verfälschung von Persönlichkeitsfragebögen sowohl ökonomisch als auch ethisch problematisch. Ein weiteres Problem der Persönlichkeitsdiagnostik mittels Persönlichkeitsfragebögen ist, dass beim Einsatz von Ratingskalen Antworttendenzen auftreten können, welche das Ergebnis ebenfalls verzerren und einen fairen Vergleich von Personen erschweren bzw. verhindern. Aktuelle Studien zu Antworttendenzen zeigen, dass diese mittels eines neuen IRT-Ansatzes als eindimensionale Faktoren modelliert und erfasst, sowie (zumindest) auf Gruppenebene korrigiert werden können (Böckenholt, 2012; Meiser & Böckenholt, 2011; Khorramdel & von Davier, under review; von Davier & Khorramdel, in press).

Die hier vorliegende Arbeit geht nun folgender Frage nach: gibt es einen Zusammenhang zwischen der extremen Antworttendenz (ERS) und starkem Verfälschungsverhalten bzw. zwischen mittelstarkem Verfälschungsverhalten und der Antworttendenz zur Mitte (MRS) Ziegler (2009) äusserte diesbezüglich bereits die Überlegung, dass Verfälschungsverhalten und bestimmte Antworttendenzen möglicherweise äquivalent seien:

„The extent to which (...) response styles relate to slight and extreme faking should be examined (...). Otherwise, we might be talking of the same phenomena using different labels.” (Ziegler, 2009, S. 34). Auch van Hooft und Born (2012) kommen auf Basis ihrer Studie zu der zu untersuchenden Vermutung, dass ein Zusammenhang zwischen Verfälschungsverhalten und der Verwendung extremer Antworttendenzen bestehen könnte.

Für diesen Zusammenhang könnte die Annahme sprechen, dass es Faking-Schemata zu geben scheint (Dalen et al., 2001; Mahar et al., 1995; Holden & Hibbs, 1995; Holden et al., 1992), d.h. stereotype Annahmen von Bewerbern

darüber, wie z.B. ein optimaler Bewerber hinsichtlich seiner Persönlichkeitsmerkmale ausgestattet sein sollte. Aus dieser Annahme leitet sich die Vermutung ab, dass Antworten von Personen, welche ihre Antworten in einem Persönlichkeitsfragebogen verfälschen, eher stereotyp ausfallen, ähnlich wie bei Antworttendenzen, welche auch über verschiedene Fragebogeninhalte bzw. –dimensionen hinweg stabil zeigen (Baumgartner & Steenkamp, 2001; Weijters et al., 2010c). Ist dies der Fall, müsste nicht nur eine Antworttendenz sondern auch eine stereotype Verfälschungs-Tendenz eindimensional über mehrdimensionale Fragebogeninhalte hinweg modellierbar sein, d.h. sich über verschiedene Fragebogeninhalte hinweg konstant zeigen.

Gestaltet sich Verfälschungsverhalten jedoch komplexer als dies mittels Antworttendenzen darstellbar ist und entspricht nicht (zumindest nicht für alle Personen) einem stereotypen Antwortmuster, sollte eine eindimensionale Modellierung über mehrdimensionale Fragebogenskalen hinweg nicht möglich sein. Hierfür würde sprechen, dass gezeigt werden konnte, dass Personen unterschiedlich (stark) verfälschen bzw. unterschiedliche Verfälschungsstile zeigen (Zickar et al., 2004; Ellingson et al., 2001; McFarland & Ryan, 2000). Auch variierende intrapersonale Modelle zum Verfälschungsverhalten (z.B. Ellingson & McFarland, 2011; Griffith et al., 2011; Tett & Simonet, 2011; Sackett, 2011; Kuncel et al., 2011) lassen ein komplexeres Zusammenspiel mehrerer Variablen vermuten.

Möglich wäre jedoch, dass die Verfälschungsstile sich zwar in der Stärke der Verfälschung unterscheiden (Rosse et al., 1998), innerhalb der unterschiedlichen Stile jedoch konsistente, stereotype Antwortmuster gezeigt werden.

Um diese Annahmen zu untersuchen, sollen in der hier vorliegenden Studie drei verschiedene Verfälschungs- bzw. Antwortstile auf deren Zusammenhang mit Antworttendenzen untersucht werden. In einem ersten Schritt wurden Studierende in drei Versuchsgruppen unterteilt und erhielten unterschiedliche Verfälschungsinstruktionen: die erste Versuchsgruppe wurde instruiert, sich eine Bewerbungssituation vorzustellen, in der es um eine höchst wünschenswerte Position geht. Dabei wurden die Testpersonen aufgefordert, sich bei der Bearbeitung eines Persönlichkeitsfragebogens auf eine Art und Weise darzustellen, welche die Wahrscheinlichkeit diese Position zu erhalten, erhöhen würde. Die Testpersonen der zweiten Versuchsgruppe wurden instruiert, sich

eine Bewerbungssituation vorzustellen, im Rahmen derer sie sich unsicher darüber seien, ob sie die Position wirklich haben wollen. Weitere Instruktionen hinsichtlich der Bearbeitung des Persönlichkeitsfragebogens waren nicht gegeben. Testpersonen der Versuchsgruppe 3 wurden aufgefordert sich vorzustellen, sie hätten sich nach Hochschulabschluss freiwillig für eine Berufsberatung entschieden, im Rahmen derer ein Persönlichkeitsfragebogen vorgegeben würde. Die Ergebnisse wären nur dann hilfreich, wenn die Testpersonen die Fragen ehrlich beantworteten. Somit wurden die Testpersonen der Versuchsgruppe 3 aufgefordert, ehrlich zu antworten. Mittels dieser drei variierenden Situationsvorgaben (Instruktionen) sollte sichergestellt werden, dass in bestimmte Richtungen und in bestimmter Stärke (stark, mittelstark, gar nicht) verfälscht wird. In einem zweiten Schritt wurde mittels des neuen IRT-Ansatzes untersucht, ob 1. über alle drei Versuchsgruppen hinweg und 2. innerhalb der einzelnen Versuchsgruppen Antworttendenzen nachgewiesen werden können, wobei der Fokus auf der Erfassung von ERS (in Zusammenhang mit starkem Verfälschungsverhalten) und MRS (in Zusammenhang mit mittelstarkem Verfälschungsverhalten) lag.

In den folgenden Abschnitten der Arbeit wird zunächst eine theoretische Wissensgrundlage gelegt. Nach einer kurzen Einführung zu Persönlichkeitsverfahren in der Bewerberauswahl folgen Informationen zur Kritik am Einsatz von Persönlichkeitsverfahren (Abschnitt 2). Abschnitt 3 behandelt sodann den Themenbereich *Response Bias* und dessen Bestandteile *Response Sets* (Verfälschungsverhalten; Abschnitt 3.1) und *Response Styles* (Antworttendenzen; Abschnitt 3.2), sowie bereits untersuchte Methoden zum Umgang mit derartigem Antwortverhalten.

Auf Basis dieses Hintergrundwissens folgen sodann Überlegungen zum Zusammenhang zwischen Verfälschungsverhalten und Antworttendenzen (Abschnitt 4) unter der Annahme stereotypen, jedoch quantitativ variierenden Verfälschungsverhaltens (starkes, mittelstarkes, kein Verfälschungsverhalten).

Den Abschluss des theoretischen Teils dieser Arbeit (Abschnitt 5) bildet die Darstellung und Beschreibung von bereits eingesetzten IRT-Methoden zur Untersuchung von Antworttendenzen und die Erläuterung der in der vorliegenden Arbeit angewandten Untersuchungsmethode mittels eines neuen IRT-Ansatzes nach Böckenholt (2012) und dessen Erweiterung auf multidimensionale IRT-

Modelle (Khorramdel & von Davier, under review; von Davier & Khorramdel, in press).

Der dritte Teil (Abschnitt III) der vorliegenden Arbeit dient der empirischen Untersuchung dieses Zusammenhangs. Nach Darlegung des Untersuchungsziels und der Hypothesen (Abschnitt 6) werden Vorgehensweise sowie verwendete Verfahren näher beschrieben (Abschnitt 7). Es folgt die Analyse, Darstellung und Interpretation der empirischen Ergebnisse (Abschnitte 8 und 9). Abschliessend werden die Ergebnisse und Vorgehensweisen der Studie kritisch betrachtet und diskutiert (Abschnitt 10).

II. THEORETISCHER TEIL

1 Persönlichkeitsverfahren in der Personalauswahl

Nach Kauffeld und Grohmann (2011) bezeichnet der Begriff Personalauswahl „die Zuweisung von Bewerbern zu Stellen in der Organisation durch Rekrutierung und Einsatz von Auswahltechniken, die auf einer Anforderungsanalyse beruhen und der Identifizierung des am besten geeigneten Bewerbers dienen.“ (Kauffeld & Grohmann, 2011, S. 94). Mittels Durchführung eines Personalauswahl-Prozesses soll dabei das Risiko einer Fehlentscheidung für die Besetzung einer offenen Stelle minimiert werden.

Psychologisch-diagnostische Persönlichkeitsverfahren finden im Rahmen der Personalauswahl, insbesondere im deutschsprachigen Raum wenig Anwendung. In lediglich 20% der von Schuler und Kollegen (2007) untersuchten deutschen Unternehmen werden Persönlichkeitsverfahren als Unterstützung bei der Personalentscheidung durchgeführt. Ein ähnliches Bild zeigt sich für US-Amerikanische Unternehmen, von denen laut einer Studie der Society for Human Resource Management (2011) 18% Persönlichkeitsverfahren im Rahmen der Bewerberauswahl vorgeben. Den häufigsten Einsatz finden die Persönlichkeitsverfahren dabei zur Beurteilung von Kandidaten für Berufspositionen im mittleren (Society for Human Resource Management, 2011; Schuler et al., 2007) und oberen (Schuler et al., 2007) Management.

Schuler et al. (2007) konnten jedoch feststellen, dass eine Entwicklung in Richtung des Einsatzes valider psychologisch-diagnostischer Persönlichkeitsverfahren zu verzeichnen sei. Dieser Trend wird nur dann Bestand haben, wenn sich die Nützlichkeit des Einsatzes von Persönlichkeitsverfahren erhöht und Faktoren, welche diesem Ziel abträglich sind, identifiziert und Persönlichkeitsverfahren darauf aufbauend adaptiert werden. Innerhalb der vorliegenden Arbeit werden verschiedene dieser Aspekte diskutiert. Auf Basis dessen ist es sodann das Ziel, ebenfalls einen Beitrag zur Steigerung der Nützlichkeit von Persönlichkeitsverfahren zu leisten.

2 Kritische Betrachtung von Persönlichkeitsfragebögen

Persönlichkeitsverfahren werden häufig in Form von Persönlichkeitsfragebögen vorgegeben, welche „auf die (reflektierte) Selbsteinschätzung einer Tp (eig. Anm.: Testperson) bezüglich ihrer Eigenschaften und typischen Erlebens- und Verhaltensweisen“ (Kubinger, 2009, S. 164) abzielen. Dabei werden den Teilnehmern verschiedene persönlichkeitsbezogene Fragen oder Aussagen vorgelegt, zu denen er/sie jeweils den Grad der Zustimmung oder Ablehnung angeben soll. Der Einsatz ist dann nützlich und sinnvoll, wenn die Verfahrensergebnisse welche ein einzelner Teilnehmer erzielt, eine präzise Einschätzung der Ausprägung einer oder mehrerer untersuchter Persönlichkeitseigenschaften dieser Person darstellt und wenn diese Einschätzung der Ausprägung einer Persönlichkeitseigenschaft einen relevanten Zusammenhang zum zukünftigen Verhalten dieser Person darstellt (Ziegler et al. 2012).

Die Vorzüge sowie Nachteile psychologisch-diagnostischer Persönlichkeitsverfahren sind Bestandteil zahlreicher wissenschaftlicher Arbeiten, wobei es häufig um die mangelnde Validität derartiger Verfahren bzw. den daraus erwachsenden Konsequenzen geht (z.B. Dilchert et al., 2006; Arthur et al., 2001; Hough et al., 1990; Ones et al., 2005; Hough & Oswald, 2009; Morgeson et al., 2007; Murphy, 2005; Hogan, 2005). Da das Gütekriterium der Validität Hauptbestandteil dieser kritischen Auseinandersetzungen ist, wird es im Folgenden erläutert.

Das Gütekriterium der Validität (oder Gültigkeit) kann in vier Unterkonzepte untergliedert werden (American Psychological Association, 1954):

Inhaltliche Validität

„Von inhaltlicher Gültigkeit eines Tests ist dann zu sprechen, wenn dieser selbst, quasi definitionsgemäß, das optimale Kriterium des interessierenden Merkmals darstellt.“ (Kubinger, 2009, S. 55). Die inhaltliche Validität kann laut Kubinger (2009) am leichtesten mittels Experten-Rating, also der Befragung von Experten zum zu validierenden Konstrukt erzielt werden.

Konstruktvalidität

„Erfüllt ein Test nicht nur einen pragmatischen Anspruch, sondern gewisse theoretische bzw. theoriegeleitete Vorstellungen in Bezug auf irgendein

„Konstrukt“, dann ist ihm Konstruktvalidität zuzusprechen.“ (Kubinger, 2009, S. 57). Die Konstruktvalidität kann nach Campbell und Fiske (1959) in diskriminante und konvergente Validität unterteilt werden. Konvergente Validität ist dann erreicht, wenn ein Test im gleichen Faktor lädt wie ein konstruktnaher Test. Im Gegensatz dazu ist diskriminante Validität gegeben, wenn ein Test in einem Faktor eines konstruktnahen Verfahrens nicht lädt (Kubinger, 2009).

Kriteriumsvalidität (Übereinstimmungsvalidität und prognostische Validität)

Die Kriteriumsvalidität ist den zuvor genannten Arten der Validität nach Kubinger (2009) überlegen, da sie als einzige der drei Validitätskonzepte statistische Kennzahlen liefert und zwar dadurch, dass Sie „eine bestimmte als relevant angesehene Variable (sog. „Aussenkriterium“) [...] mit dem interessierenden Merkmal korreliert.“ (Kubinger, 2009, S. 64). Dabei kann die Kriteriumsvalidität, je nach Aussenkriterium gruppiert werden: „Die sog. Übereinstimmungsvalidität bezieht sich auf die Korrelation mit einem anderen Test, der (angeblich) dasselbe Konstrukt erfasst.“ (Kubinger, 2009, S. 64). Im Gegensatz dazu wird die prognostische Validität durch die Korrelation des untersuchten Verfahrens mit einem, in der Zukunft liegenden sowie für das Verfahren relevanten, Aussenkriterium erhoben (Kubinger, 2009). Die prognostische Validität spielt für den Einsatz von Persönlichkeitsfragebögen in der Personalauswahl eine tragende Rolle, da Persönlichkeitsfragebögen genau zu diesem Zweck, nämlich der möglichst treffsicheren Prognose eingesetzt werden. Selbstverständlich führt eine falsche Prognose aufgrund niedriger prognostischer Validität zu einer möglichen Fehlbesetzung die es zu vermeiden gilt. In der wissenschaftlichen Literatur wird der Begriff prognostische Validität gelegentlich unklar verwendet, indem er mit dem Begriff der Kriteriumsvalidität gleichgesetzt wird (z.B. Dilchert et al., 2006).

Wie eingangs bereits erwähnt, gehen die Meinungen hinsichtlich der Validität der vorhandenen, häufig Big-Five-basierten Persönlichkeitsinventare auseinander. Barrick & Mount (1991) wiesen Anfang der 90er Jahre nach, dass die Dimension Gewissenhaftigkeit einen guten Prädiktor für diverse berufliche Leistungen darstellt. Auch die Dimension Emotionale Stabilität (bzw. Emotionale Labilität/Neurotizismus) sei in der Lage, die umfassende berufliche Leistung vorherzusagen (Barrick, 2005). Auf die Studie von Barrick und Mount (1991)

folgten eine Vielzahl wissenschaftlicher Studien, welche ähnliche Ergebnisse zu Kriteriums- und Konstruktvalidität von Persönlichkeitsverfahren im beruflichen Kontext hervorbrachten (z.B. Hurtz & Donovan, 2000; Hogan & Holland, 2003; Ones et al., 2005) und die Berechtigung des Einsatzes von Persönlichkeitsverfahren in der Personalauswahl verfestigten (Hogan, 2005; Hough & Oswald, 2009). Dabei konnten die Prädiktorleistungen der Dimensionen Emotionale Stabilität und Gewissenhaftigkeit für die allgemeine berufliche Leistung bestätigt werden, wobei die Ergebnisse hinsichtlich der restlichen drei Faktoren (Extraversion, Offenheit für Erfahrungen, Verträglichkeit) als Prädiktoren lediglich für spezifische Berufs-Teilleistungen oder Berufsbilder bestätigt werden konnten (Barrick, 2005; Hough & Oswald, 2005).

Murphy und Dzieweczynski (2005) bemängeln auf Basis von Metaanalysen im Gegensatz zu den genannten Ergebnissen, die geringen Validitätswerte der Big-Five basierten Persönlichkeitsverfahren und die damit einhergehende geringe Relevanz der Persönlichkeitsverfahren für die erfolgreiche Personalauswahl. Ziegler und Bühner (2009) konnten eine verminderte Konstruktvalidität mittels Feststellung erhöhter Interkorrelationen zwischen Big-Five Dimensionen auf Grund von Verfälschungsverhalten nachweisen. Auch Dilchert et al. (2006) konnten den geringen Nutzen von Persönlichkeitsfragebögen zur Vorhersage beruflicher Leistungen betreffend die Kriteriums- bzw. prognostische Validität der eingesetzten Verfahren nachweisen (Hough et al., 1990; Mueller-Hanson et al., 2003).

Hogan (2005) setzt dem entgegen, dass die Metaanalysen welche geringe Validitätswerte der Big-Five Inventare nachweisen, methodisch mangelhaft durchgeführt würden. So würden im Rahmen der Metaanalysen solche Konstrukte welche zwar ähnlich benannt, jedoch inhaltlich unterschiedlich seien gemeinsam ausgewertet. Weiter würden Ergebnisse aus Verfahren welche hohe Validitätswerte aufweisen, mit Ergebnissen aus gering validierten Verfahren gemeinsam ausgewertet (Hogan, 2005). Hough und Oswald (2005) befinden, dass im Rahmen der durchgeführten Validitätsstudien und Metaanalysen insbesondere darauf zu achten sei, dass das gewählte Konstrukt (z.B. Verträglichkeit) und das darauf abgestimmte Prädiktorenkonstrukt (z.B. Berufsleistungen unter Teamarbeit) sorgfältig aufeinander abgestimmt werden müssten, da ein Mangel an Abstimmung durch fehlende Passung der Konstrukte

zu geringen Validitätswerten führe. Ein Aspekt, dem sowohl Kritiker (Murphy & Dziewieczynski, 2005; Morgeson et al., 2007; Douglas et al., 1996) als auch Befürworter (Hogan, 2005; Ones et al., 2005; Hough & Oswald, 2005) des Einsatzes von Persönlichkeitsfragebögen zustimmen, ist die Tatsache, dass Gütekriterien wie Validität und Reliabilität bei der Auswahl von Persönlichkeitsverfahren zur Personalauswahl in Organisationen häufig nicht berücksichtigt werden, was in der Folge bedeutet, dass qualitativ minderwertige Verfahren zum Einsatz kommen, welche wenig überraschend nicht in der Lage sind die gewünschte Vorhersageleistung zu erbringen.

Ein weiterer Kritikpunkt am Einsatz von Persönlichkeitsfragebögen in der Personalauswahl und möglicherweise auch Ursache für geringe Validitätswerte ist die Tatsache, dass Persönlichkeitsfragebögen anfällig für Verfälschung sind. Da der Aspekt des Verfälschens von Persönlichkeitsfragebögen einen zentralen Punkt der vorliegenden Arbeit darstellt, ist er im Folgenden detailliert erläutert.

3 Response Bias in Persönlichkeitsfragebögen

Ein wissenschaftlicher Oberbegriff, der die zentralen Themen der vorliegenden Arbeit beinhaltet, ist der des *Response Bias* (systematische Antwortverzerrung). Dieser Begriff beinhaltet sowohl die formelle Antwortverzerrung (auch Antworttendenzen, *Response Style*) als auch die Item-inhaltsabhängige Antwortverzerrung (*Response Set*) (z.B. Bardwell & Dimsdale, 2001; van Herk et al., 2004). Während die verschiedenen Formen der inhaltsunabhängigen Antworttendenzen (*Response Style*) klar voneinander abgegrenzt werden, ist diese Begriffsabgrenzung im Falle der inhaltsabhängigen Antwortverzerrung (*Response Set*) bei Betrachtung der wissenschaftlichen Literatur nicht immer gegeben. Den zahlreichen wissenschaftlichen Artikeln und Büchern, welche sich mit dem Thema der inhaltlichen Antwortverzerrung auseinander setzen, liegen oftmals unterschiedliche Begriffsverwendungen zugrunde (z.B. Reeder & Ryan, 2012; Smith & McDaniel, 2012; Ellingson, 2012). Diese unterschiedlichen Definitionen verwenden in der Regel einen oder mehrere Begrifflichkeiten aus Tabelle 1, I.

Tabelle 1 dient somit der Auflistung der wichtigsten inhaltsunabhängigen Antworttendenzen (II) sowie der Darstellung all jener Begriffe, welche im

Zusammenhang mit *Response Sets* verwendet werden (I), um diese in der Folge inhaltlich voneinander abzugrenzen und eine einheitliche Definition zur weiteren Verwendung in der vorliegenden Arbeit zu erarbeiten. Während die obere Zeile der Tabelle 1 (*Response Sets*) im nächsten Abschnitt (Abschnitt 3.1) erläutert wird, finden sich nähere Erläuterungen zur unteren Zeile (Antworttendenzen) in Abschnitt 3.2.

Tabelle 1: Response Bias - Antworttendenzen und Response Sets

I Response Sets <u>Inhaltsabhängig</u>	Verfälschen <i>Faking</i> , z.B. Reeder & Ryan (2012) Soziale Erwünschtheit <i>Social Desirability</i> , z.B. Paulhus (1984) Eindruckskontrolle <i>Impression Management</i> , z.B. Paulhus (1984), Turnley & Bolino, (2001) Absichtliche Antwortverfälschung <i>Intentional Response Distortion</i> , z.B. Smith & McDaniel (2012) Verstellung <i>Dissimulation</i> , z.B. Ellingson, 2012 Selbsttäuschung <i>Self-Deception</i> , z.B. Paulhus (1984)
II Antworttendenzen¹ <i>Response Styles</i> <u>Formal</u> , <u>Inhaltsunabhängig</u>	Extreme Antworttendenz (ERS) Tendenz zur Mitte (MRS) Antwortbereichs-Tendenz (RR) Netto-Zustimmungstendenz (NARS) Ja-Sage-Tendenz (ARS) Nein-Sage-Tendenz (DARS) Willkürliche Antworttendenz (NCR) Milde Antworttendenz (MLRS)

eigene Darstellung

3.1 Inhaltliche Verfälschung – Response Sets

Der Begriff der inhaltlichen Verfälschung bezieht sich auf die in Tabelle 1 angeführten Begrifflichkeiten (I). Dabei stehen meist die Begriffe Verfälschung (*Faking*) und sozial erwünschtes Antwortverhalten (*Socially Desirable*

¹Die angeführten Antworttendenzen beziehen sich auf Vaerenbergh & Thomas (2012).

Responding) im Vordergrund, wobei der Begriff der Verfälschung häufig als Oberbegriff für die weiteren angeführten Ausdrücke verwendet wird. Eine bedeutende wissenschaftliche Arbeit, welche sich mit dem Begriff des sozial erwünschten Antwortverhaltens (s.Tabelle 1, I) auseinander setzt, ist die Arbeit von Delroy Paulhus (1984). Paulhus (1984) erarbeitete ein Zwei-Komponenten Modell der sozialen Erwünschtheit. Dieses besteht zum einen aus *Impression Management* (Kontrolle des Eindrucks), verstanden als bewusstes Kontrollieren des entstehenden Eindrucks bei anderen Personen (Paulhus, 1984; Turnley & Bolino, 2001). Die zweite Komponente bildet die *Self-Deception* (Selbsttäuschung), ein Vorgang bei dem der Antwortende selbst von seinem übermäßig positiven Selbstbericht überzeugt ist und somit unbewusst verfälscht (Paulhus, 1984). Paulhus dazu: "(...) socially desirable responding centers on distinguishing self-deception, where the respondent actually believes his or her positive self-reports, (...) from impression management, where the respondent consciously dissembles." (Paulhus, 1984, S.599).

In vielen Studien werden die Begriffe sozial erwünschtes Antwortverhalten, *Impression Management* und Verfälschen oder *Faking* synonym verwendet oder zumindest begrifflich nicht klar abgegrenzt. So definieren beispielsweise Reeder und Ryan (2012) den Begriff Verfälschen als sozial erwünschtes Antwortverhalten. McFarland und Ryan (2000) hingegen verwenden in ihrer Studie den Begriff *Impression Management* synonym mit Verfälschen. Ein weiterer Begriff, welcher im gleichen Zusammenhang vorkommt, ist der der *Intentional Response Distortion* (vorsätzliche Antwortverzerrung). Auch dieser Begriff wird synonym mit Verfälschen, sozial erwünschtem Antwortverhalten und *Impression Management* verwendet. Smith und McDaniel (2012) definieren Verfälschen beispielsweise als motiviertes *Impression Management* durch die Verwendung vorsätzlicher Antwortverzerrung (*Intentional Response Distortion*). Auch *Dissimulation* (Verstellung) kommt im Zusammenhang mit den oben erläuterten Begriffen vor, unterliegt der gleichen Problematik der mangelnden Abgrenzung (z.B. Ellingson, 2012) und wird häufig als Synonym für den Begriff Verfälschen bzw. *Faking* verwendet (z.B. Ellingson, 2012), beschreibt im Gegensatz zum sozial erwünschten Antwortverhalten jedoch keine festgelegte Richtung der Verfälschung (positives vs. negatives Verfälschen).

MacCann et al. (2012) versuchten, die vielfach unterschiedlichen Definitionen für den Begriff des Verfälschens unter Betrachtung der wissenschaftlichen Literatur auf einen gemeinsamen Nenner herunterzubrechen und kommen auf Basis dessen zu den folgenden Grundannahmen (MacCann et al., 2012):

- Verfälschen ist eher eine Verhaltensweise als ein Persönlichkeitsmerkmal.
- Dem Verfälschen muss eine Motivation bzw. eine Zielorientierung zugrunde liegen.
- Verfälschungsverhalten hinterlässt einen inakkuraten oder optimierten Eindruck.
- Verfälschen beinhaltet Interaktionen zwischen Personen (Persönlichkeitsmerkmalen) und situationsabhängigen Faktoren (Umwelteinflüssen).
- Die Definition von Verfälschen wird mehrheitlich auf bewusstem und absichtlichem Verhalten abgestellt.

Diese Definitionsbestandteile grenzen den Begriff Verfälschen von unabsichtlichem und unbewussten Antwortverzerrungen (i.S. *Self-Deception*) ab. **Verfälschen (*Faking*) ist demnach ein bewusstes, situationsabhängiges Verhaltensmuster mit dem Ziel, einen bestimmten, vorgetäuschten Eindruck zu hinterlassen** (MacCann et al., 2012). Diese Definition betrachtet somit *Impression Management* als Bestandteil von Verfälschen und schliesst unbewusstes Verhalten (z.B. *Self-Deception* nach Paulhus, 1984) als Definitionsbestandteil des Begriffs Verfälschen aus. Die Definitionskomponente des „bestimmten“ Eindrucks bezieht sich wohl auf einen sozial erwünschten Eindruck. Die Verwendung des Ausdrucks „mit dem Ziel“ impliziert weiters eine Zielorientierung. Somit beinhaltet die obige Definition all jene Komponenten, die im Kontext des Bewerbungsprozesses, Bewerber-Antwortverhalten und somit der vorliegenden Arbeit von Relevanz sind. In der vorliegenden Arbeit wird daher die Definition von MacCann et al. (2012) für den Begriff des Verfälschens bzw. *Faking* aufgegriffen und im Weiteren verwendet.

3.1.1 Verfälschen als stereotypes vs. individuell variierendes Antwortverhalten

Personen in Bewerberauswahlsituationen haben in der Regel den Anreiz, in vorgegebenen Persönlichkeitsverfahren ein Ergebnis zu erzielen, welches die

Wahrscheinlichkeit, eine angestrebte Arbeitsstelle tatsächlich zu erhalten, erhöht (Dilchert et al., 2005). Donovan et al. (2003) befragten diesbezüglich Testpersonen zur Einschätzung ihres Antwortverhaltens. 50% der befragten Testpersonen gaben an, ihre positiven Eigenschaften übertrieben dargestellt zu haben. Mehr als 60% der Testpersonen bestätigten, ihre schlechten Seiten in abgeschwächter Form dargestellt zu haben und 15% der Testpersonen gaben zu, Angaben gemacht zu haben die vollständig gelogen seien (Donovan et al., 2003).

Die Differenz zwischen der wahren Persönlichkeit eines Bewerbers bzw. den wahren Antwortwerten (*true score*) und den beobachteten Ergebniswerten lässt sich demnach, neben dem Einfluss des Messfehlers, auf die Verfälschung durch den Kandidaten zurückführen (Dilchert et al., 2005). Der Ergebniswert erhebt bei starker Verfälschung dann nicht mehr die Ausprägung eben dieser Dimension, sondern auch das Ausmaß der Verfälschung (Ziegler et al., 2012) und hat somit einen Einfluss auf die Konstruktvalidität (s. Abschnitt 2) eines Verfahrens.

Hinsichtlich der Frage, wie sich Verfälschungsverhalten auf Basis von Fragebogenergebnissen äussert, sind dabei verschiedene Szenarien vorstellbar. Im Folgenden werden zunächst solche Ergebnisse erläutert die auf stereotypes, jedoch quantitativ variierendes Verfälschungsverhalten hindeuten. Dem liegt die Annahme zugrunde, dass Testpersonen ein einheitliches, stereotypes Bild jener (Persönlichkeits-)Merkmale haben, welche im Allgemeinen als wünschenswert bzw. sozial erwünscht gelten. Im Anschluss daran werden solche Ergebnisse und Modelle erläutert, die von einem komplexeren Verfälschungsverhalten, vor dem Hintergrund individuell variierender Annahmen über sozial erwünschte (Persönlichkeits-) Merkmale ausgehen.

Visweswaran und Ones (1999) konnten zeigen, dass (Test-) Personen nach Erhalt einer Verfälschungsinstruktion² in der Lage sind, alle fünf Faktoren eines Big-Five-Persönlichkeitsinventars sowohl in verbessernde (*faking good*) als auch in verschlechternde (*faking bad*) Richtung zu beeinflussen.

²Eine Verfälschungsinstruktion ist ein Verfahrens-Bearbeitungshinweis, welcher den Teilnehmer auffordert, das (Persönlichkeits-)Verfahren auf eine bestimmte Art und Weise zu beantworten, somit zu verfälschen. Im Zusammenhang dieser Arbeit geht es dabei um die Aufforderung sich selbst mittels bewusst gewählter Antwortkategorien auf eine Art und Weise darzustellen, die die Wahrscheinlichkeit eine Arbeitsstelle zu erhalten, erhöht. Ergänzend dazu s. Abschnitt 3.1.3, S. 28

Holden und Hibbs (1995) sowie Holden, Kroner, Fekken und Popham (1992) legten im Rahmen ihrer Studien mittels Erhebung von Item-Latenzzeiten dar, dass Personen wenn Sie dazu aufgefordert werden, gemäss bestehender stereotyper Schemata zu *faking good* oder *faking bad* auf Items aus Persönlichkeitsfragebögen reagieren. Dalen et al. (2001) untersuchten, inwiefern die Herausgabe dezidierter Informationen zur Arbeitsstelle das Verfälschungsverhalten der Bewerber beeinflusste und konnten diesbezüglich keine Effekte feststellen. Aus diesen Ergebnissen leiten auch Dalen und Kollegen (2001) die Annahme über ein stereotypes Verfälschungsmuster ab. So scheinen sowohl wahre Berufsanwärter als auch Studienteilnehmer, welche eine Verfälschungsinstruktion in Bezug auf eine Bewerbungssituation erhalten, vorgegebene Persönlichkeitsfragebögen in Bezug auf bestimmte oder als relevant erachtete Dimensionen zu verfälschen (z.B. Galic et al., 2012; Furnham, 1990). Nach Rosse et al. (1998) kann der Grad an Verfälschungsverhalten (in Bezug auf die als allgemein sozial erwünscht erachteten Dimensionen) zwischen den verfälschenden Testpersonen variieren. Im Rahmen ihrer Studie lag der Variierungsgrad an Verfälschungsverhalten zwischen einer bis drei Standardabweichungen jenseits des Mittelwertes (Rosse et al., 1998). Auch die Ergebnisse einer Studie, durchgeführt von Zickar und Robie (1999), legen über variierende Standardabweichungen der Testpersonen einer zur Verfälschung angehaltenen Versuchsgruppe im Vergleich mit einer ehrlich antwortenden Kontrollgruppe nahe, dass verfälschende Personen quantitativ variierendes Verfälschungsverhalten an den Tag legen. Diese Ergebnisse legen die Vermutung nahe, dass Verfälschungsverhalten tatsächlich als stereotypes Antwortverhalten, im Sinne einheitlicher Annahmen über sozial erwünschte (Persönlichkeits-) Merkmale verstanden werden kann.

Im nun folgenden Abschnitt werden solche Erkenntnisse erläutert, die Verfälschungsverhalten unter der Annahme individuell variierender Überzeugungen zu sozial erwünschten (Persönlichkeits-) Merkmalen, nahelegen. So überprüften Ellingson et al. (2001) im Rahmen einer Studie, inwieweit Verfälschungsverhalten zur Veränderung der einem Persönlichkeitsverfahren zugrundeliegenden Faktorenstruktur führt, konnten dies jedoch nicht bestätigen. Die Tatsache dass das Verfälschungsverhalten die Faktorenstruktur des Verfahrens nicht beeinflusst, stellt die soeben erläuterte Annahme über

Verfälschungsverhalten als eindimensionalen Faktor in Frage und spricht somit für ein Verfälschungsverhalten als komplexes und qualitativ variierendes Antwortmuster. McFarland und Ryan (2000) konnten weiters zeigen, dass die Art und Weise des Verfälschungsverhaltens mit der tatsächlichen Ausprägung bestimmter Persönlichkeitsmerkmale (Gewissenhaftigkeit, Neurotizismus) zusammenhängt und somit individuell variiert. Basierend auf diesen Erkenntnissen entwickelten McFarland und Ryan (2000) ein Modell zur Erklärung variierenden Verfälschungsverhalten, welches die Faktoren Verfälschungsfähigkeit (*ability to fake*), Überzeugungen zum Verfälschungsverhalten (*beliefs toward faking*), Verfälschungsgelegenheit (*opportunity to fake*) sowie situationale Faktoren (persönliche Relevanz der Arbeitsstelle, Aussprechen von Verfälschungswarnungen bzw. Einsatz von Lügenskalen) berücksichtigt. Zickar et al. (2004) konnten zeigen, dass Verfälschungsverhalten unter gleichen Bedingungen hinsichtlich der verfälschten Dimensionen variieren kann und führen diesen Umstand auf verschiedene Faktoren bspw. die Fähigkeit zu Verfälschungsverhalten, die Fähigkeit zur Selbstbeobachtung, motivationale Faktoren, den Willen zur Ehrlichkeit sowie Überzeugungen zur Nachweisbarkeit des eigenen Verfälschungsverhaltens zurück. König et al. (2012) versuchten im Rahmen einer qualitativen Studie zu ergründen, welche Faktoren einen Einfluss auf das Verfälschungsverhalten haben und konnten mittels Interview-Technik feststellen, dass das Antwortverhalten der Testpersonen innerhalb verschiedener Verfälschungsgruppen insbesondere von Überzeugungen zu folgenden Faktoren beeinflusst wird: die Konsistenz der Antworten innerhalb einer Dimension wird überprüft und die Testperson so als Verfälscher entlarvt; die Häufigkeit der Wahl der mittleren sowie extremen Antworttendenz wird beachtet und hinterlässt einen bestimmten Eindruck, wobei die Wahl der mittleren Antwortkategorie als Zeichen für Mittelmässigkeit, die Wahl der extremen Antwortkategorie als Zeichen für ausgeprägte Dickköpfigkeit gedeutet werden kann; die Ergebnisse des Verfahrens werden mit einem im vorhinein festgelegten idealen Kandidatenprofil abgeglichen, wobei Persönlichkeitsmerkmale auch zu gut ausgeprägt sein könnten (bspw. könnte sich zu starke Extraversion abträglich auswirken). Variierende Annahmen der Testpersonen zu diesen Faktoren liefern eine Basis für stark variierendes Verfälschungsverhalten hinsichtlich der erfragten

Konstrukte. Auch Mueller-Hanson et al. (2006) konnten qualitativ variierendes Verfälschungsverhalten von Testpersonen unter Verfälschungsinstruktionen feststellen. Die entscheidenden Einflussfaktoren auf die Verfälschungsabsicht und somit das Verfälschungsverhalten lagen nach Mueller-Hanson et al. (2006) in der Wahrnehmung der Situation (Überzeugungen zur Relevanz von Verfälschen, wahrgenommene Verhaltenskontrolle, Überzeugungen zu subjektiven Normen), der Gewissenhaftigkeitsausprägung und der emotionalen Stabilität der Teilnehmenden.

Darüberhinaus existieren weitere komplexe intrapersonale Modelle zur Erklärung variierenden Verfälschungsverhaltens: Tett und Simonet (2011) stellen diesbezüglich ein sog. Multisaturations-Modell vor und differenzieren zwischen den von Paulhus (1984) (s. Abschnitt 3.1) geprägten Begriffen *Impression Management* und *Self-Deception*. Im Falle des *Impression Management* nennen sie als Einflussfaktoren Risikofreude, Machiavellismus, geringe Integrität und Ehrgeiz. Bezogen auf die *Self-Deception* nennen sie die Faktoren Anerkennungsbedürfnis, Idealismus bzgl. der abgefragten Eigenschaft und Narzissmus. Zusätzlich gehen sie davon aus, dass die Gelegenheit, Fähigkeit und Motivation zum Verfälschen einen Einfluss auf das Vorkommen von *Impression Management* und *Self-Deception* haben, die dann schlussendlich in Kombination die Differenz zwischen wahren und beobachteten Ausprägungen ausmachen (Tett & Simonet, 2011).

Ellingson und McFarland (2011) stellen ebenfalls ein grundlegendes Konzept zum Verfälschungsvorgang dar. *Valence* (Wertigkeit) steht dabei für den Wert des antizipierten Verhaltens hinsichtlich der zu erwartenden Befriedigung, *Instrumentality* (Zweckdienlichkeit) stellt das Ausmaß dar, zu dem ein bestimmtes (in diesem Fall Verfahrens-) Ergebnis bei der Erreichung eines Ziels, beispielsweise dem Erhalt einer Arbeitsstelle behilflich sein kann und *Expectancy* (Erwartungswert) bezieht sich auf die Überzeugung einer Person, über die (hier: Verfälschungs-) Fähigkeit zu verfügen, das gewünschte Ziel bzw. bestimmte, erwünschte Merkmalsausprägungen zu erreichen (Ellingson & McFarland, 2011). Weitere Faktoren aus denen sich dann das Verfälschungsverhalten ergibt, stellen die Verfälschungsmotivation (beeinflusst von situationalen Faktoren), die tatsächliche Fähigkeit zum Verfälschen sowie intellektuelle Fähigkeiten dar (Ellingson & McFarland, 2011).

Die *Trait Contract Classification Theory* (Griffith, Peterson & Zickar, 2011) stellt Verfälschungsverhalten anhand vierer Kategorien dar: *Self-Presentation* (Selbstpräsentation), *Exaggeration* (Übertreiben), *Reactive Responding* (Reaktives Antwortverhalten) und *Fraudulent Responding* (betrügerisches Antwortverhalten).

Die Faktoren, welche dann schlussendlich zu einem oder mehreren dieser vier Formen der Verfälschung führen, speisen sich aus verschiedenen Quellen:

*Self-Monitoring*³ (Eigenüberwachung), *Self-Esteem* (Selbstwertgefühl), *Narcissism* (Narzissmus), *Impulsivity* (Impulsivität), *Integrity/Honesty-Humility*⁴ (Integrität/Ehrlichkeits-Demut) und *Locus of control*⁵ (Kontrollort bzw. Attribuierungsstil).

Diese Modelle legen ebenfalls die Vermutung nahe, dass Verfälschungsverhalten in nicht stereotyper sondern individuell variierender Form auftritt. Kritisch an einigen dieser Modelle (Tett & Simonet, 2011; Ellingson & McFarland, 2011; Griffith et al., 2011) ist jedoch anzumerken, dass sie keiner empirischen Überprüfung unterzogen wurden.

Somit scheint die Erkenntnislage zur Frage der Art des Verfälschungsverhaltens (stereotyp vs. individuell variierend) zum jetzigen Zeitpunkt nicht eindeutig geklärt zu sein. Die vorliegende Arbeit stellt jedoch auf die geäußerte Vermutung des qualitativ stereotypen, womöglich jedoch quantitativ variierenden Verfälschungsverhaltens ab und überprüft in diesem Sinne die Annahme eines Zusammenhangs zwischen derartigem Verfälschungsverhalten und dem Vorliegen von Antworttendenzen (ERS und MRS).

³*Self-Monitoring* bezeichnet die stetige Kontrolle der Präsentation der eigenen Person nach außen und den Versuch, das eigene Verhalten an die Erwartungen der Außenwelt anzupassen (Griffith et al., 2011)

⁴*Honesty-Humility* als Persönlichkeitsmerkmal ist charakterisiert durch Begriffe wie Ehrlichkeit, Fairness, Aufrichtigkeit, fehlender Eingebildetheit, fehlendem Ausbeutungswillen und fehlender Habgier (Lee et al., 2005 zitiert nach Griffith et al., 2011).

⁵*Locus of Control* bezeichnet das Ausmaß in welchem Personen Ereignisse auf externe (Glück oder Schicksal) bzw. interne Quellen (eigene Fähigkeit oder eigene Schuld) zurückführen (Rotter, 1990 zit. nach Griffith et al., 2011).

3.1.2 Methoden zur Erhebung von Verfälschungsverhalten

Um die Art und Ausprägung von Verfälschungsverhalten zu erheben, haben sich verschiedene, unterschiedlich komplexe Methoden etabliert. Burns und Christiansen (2011) verwenden ein Drei-Faktoren-Modell zur Darstellung der aus der wissenschaftlichen Literatur extrahierten Methoden zur Messung von Verfälschungsverhalten. Dabei geht es um die Frage ob (1.) die Erhebungsmethode das einmalige vs. wiederholte Erheben von Daten benötigt und ob (2.) zur Erhebung von Verfälschung die nicht-kognitiven Erhebungsinstrumente, also Persönlichkeitsverfahren (internal) vs. zusätzliche Erhebungsinstrumente (external) herangezogen werden. Der dritte Faktor bezieht sich auf die (3.) Erscheinungsform des Verfälschungseffekts in Form eines Einflusses auf die Konstruktvalidität oder einer Verschiebung des Mittelwertes (Burns & Christiansen, 2011). Tabelle 2 enthält die verschiedenen Erhebungsmethoden und die dazugehörigen Ausprägungen der drei oben genannten Faktoren (in Anlehnung an Burns und Christiansen, 2011).

Tabelle 2: Erhebungsmethoden von Verfälschungsverhalten

	Erhebungsmethode	Wiederholte vs. Einmalige Erhebung	Internales vs. Externales Erhebungsinstrument	Erscheinungsform des Verfälschungseffekts
I.	Score-Abweichung	Wiederholt	Internal	Mittelwertsverschiebung
II.	Residualisierte Score- Abweichung	Wiederholt	Internal	Mittelwertsverschiebung
III.	Validitäts- und Soziale Erwünschtheitsskalen	Einmalig	External	Mittelwertsverschiebung
IV.	Schein-Items	Einmalig	External	Mittelwertsverschiebung
V.	Spezifische Item- Antwort	Einmalig	Internal	Mittelwertsverschiebung
VI.	Prozentuale Antwortkonsistenz	Wiederholt	Internal	Konstruktvalidität
VII.	Within-Subject- Korrelation	Wiederholt	Internal	Konstruktvalidität
VIII.	Within-Subject- Differenzvarianzen	Wiederholt	Internal	Konstruktvalidität
IX.	Kovarianz-Index	Einmalig	Internal	Konstruktvalidität

Nach Burns und Christiansen (2011) ist die Erhebung des Effekts der Verfälschung mittels Messung der Mittelwertsverschiebung (s. Tabelle 2, rechte Spalte) die robusteste und auch am häufigsten angewandte Form der Messung, welche beispielsweise auch im Rahmen der viel zitierten Metaanalyse von Viswesvaran und Ones (1999) Anwendung fand. In der Regel werden den Testpersonen Persönlichkeitsverfahren unter ehrlicher vs. Verfälschungsbedingung (mittels Verfälschungsinstruktionen, s. Abschnitt 3.1.3, S.20) im Rahmen einer einmaligen Erhebung (s. zweite Spalte) vorgegeben, wie es auch im Rahmen der vorliegenden Arbeit geschehen ist. Oder aber es werden tatsächlichen Bewerbern Persönlichkeitsfragebögen innerhalb des Bewerbungsprozesses und nach begunnenem Arbeitsverhältnis, also mittels wiederholter Erhebung, vorgegeben (Burns & Christiansen, 2011). Die Mittelwertsverschiebung wird dabei entweder für das Persönlichkeitsverfahren (internales Erhebungsinstrument) oder aber mittels externalem Erhebungsinstrument, beispielsweise einer sozialen Erwünschtheits- bzw. Verfälschungsskala erhoben. Eine detaillierte Erläuterung der einzelnen Methoden findet sich bei Burns und Christiansen (2011).

Jedoch stellen die von Burns und Christiansen (2011) angeführten Methoden zur Identifizierung von Verfälschungsverhalten nicht die vollständige Breite an zur Verfügung stehenden Methoden dar. So wurde beispielsweise bereits die Reaktionszeit pro Item (Shoss & Strube, 2011; van Hooft & Born, 2012) als Maß für Verfälschungsverhalten herangezogen, stellte sich jedoch als hinfällig dar (Shoss & Strube, 2011). Van Hooft und Born (2012) erhoben die Augenbewegungen der Testpersonen unter Bearbeitung von Fragen unter ehrlichen sowie Verfälschungsbedingungen, welches vielversprechende Ergebnisse lieferte (dazu s. Abschnitt 4).

Kompliziertere Methoden, beispielsweise die Erhebung von Verfälschungsverhalten mittels Modellen der Item-Response-Theorie (z.B. Robie & Zickar, 2001) oder Struktur-Gleichungs-Modellen (Ziegler & Bühner, 2009) werden ebenfalls zur Feststellung von Verfälschungsverhalten angewendet und liefern vielversprechende Ergebnisse. Auf die Verwendung von IRT-Modellen zur Erhebung von Verfälschungsverhalten wird zu einem späteren Zeitpunkt (Abschnitt 5) jedoch näher eingegangen, da die Verwendung derartiger IRT-

Modelle im gegebenen Zusammenhang einen essentiellen Bestandteil der vorliegenden Arbeit darstellen.

3.1.3 Verfälschungsinstruktionen zur Erzeugung und Untersuchung von Verfälschungsverhalten

Eine verbreitete Methode, die Effekte von Verfälschungsverhalten auf Persönlichkeitsverfahren zu erheben, ist die Vorgabe von Persönlichkeitsverfahren unter Verfälschungsbedingungen (Raymark & Tafero, 2009). Diese Verfälschungsbedingung wird dabei mittels Verwendung einer sog. Verfälschungsinstruktion (*Faking Instruction*) erzeugt. Die Instruktion kann auf mündliche oder schriftliche Weise übermittelt werden und fordert VerfahrensteilnehmerInnen in der Regel dazu auf, sich so zu verhalten oder Items so zu beantworten, wie sie es in einer bestimmten (z.B. Bewerbungs-) Situation tun würden (z.B. in Raymark & Tafero, 2009; Röhner et al., 2011; Pauls & Crost, 2005b; Nguyen et al., 2005; Widhiarso, 2011). Dabei variiert der Grad an Spezifikation dieser Instruktion. So kann eine derartige Instruktion zum reinen besser Darstellen auffordern oder aber den Teilnehmenden instruieren, sich so darzustellen, dass die Chancen eine bestimmte Stelle (bspw. BuchhalterIn, s. Raymark & Tafero, 2009; Pauls & Crost, 2005b; Nguyen et al., 2005; Widhiarso, 2011) zu erhalten, erhöht werden. Einige Studien fordern die Testpersonen auf bestimmte Eigenschaftsausprägungen (z.B. hohen Ehrgeiz, Röhner et al., 2011) vorzutäuschen. Des Weiteren kann auch die Vehemenz der Aufforderung variieren. Eine abgeschwächte Version verwenden beispielsweise Raymark und Tafero (2009, S.93): *“Please answer the questions (...) in such a way to make the most favorable impression possible.”* Ziegler et al. (2007) formulieren eine Verfälschungsinstruktion mittels der Angabe dass der Job erstrebenswert sei und man solle sein Bestes geben. Eine tatsächliche Täuschungs-Aufforderung verwenden Röhner et al. (2011) („*fake (...) to maximise the chances of being offered the job*“, Röhner et al., S.467). Um den Verfälschungseffekt zu maximieren ergänzen Nguyen et al. (2005) die Verfälschungsinstruktion mit dem Hinweis auf die Anonymität und den lediglich wissenschaftlichen Verbleib der Daten, in der Hoffnung, dass den Testpersonen die verfälschungshemmende Angst vor Konsequenzen genommen wird. Douglas et al. (1996) intensiviert die

Verfälschungsinstruktion folgendermassen: „*we are not interested in your true opinion, rather we would like you to identify the most socially desirable response for each question*“ (Douglas et al., 1996, S.3). Die hier erwähnten wissenschaftlichen Studien führten ihre Untersuchungen bzw. die Vorgabe der Instruktionen dabei ausschliesslich im Rahmen tatsächlicher face-to-face-Situationen durch (z.B. Ziegler et al., 2007; Widhiarso, 2011; Nguyen, 2005; Douglas et al., 1996; Raymark & Tafero, 2009).

Abschliessend ist jedenfalls festzustellen, dass die Variation der Verfälschungsinstruktionen womöglich auch zu unterschiedlichem Verfälschungsverhalten führt. Raymark und Tafero (2009) konnten diesbezüglich am Beispiel einer *faking good* – Verfälschungsinstruktion und einer *faking accountant* (BuchhalterIn)-Verfälschungsinstruktion ⁶ feststellen, dass die Ergebnisse unter diesen beiden Instruktionen insofern unterschiedlich waren, als dass solche Berufsbild-Facetten, welche für die Tätigkeit eines Buchhalters negativ bewertet werden (z.B. Offenheit für Tätigkeiten, Geselligkeit und Erlebnissuche) unter der *faking accountant*- Bedingung geringere Ausprägungen als unter einer Ehrlichkeitsbedingung⁷ erzielten und in einer reinen *faking good*-Bedingung höhere Werte erzielten, als dies in der Ehrlichkeitsbedingung der Fall war (Raymark und Tafero, 2009). Dieses Beispiel zeigt, dass abhängig von der Art der Verfälschungsinstruktion unterschiedliche Effekte zu erwarten sind. Weiters könnte dies eine Begründung für die Differenz zwischen den Effekten von Verfälschungsinstruktionen und Effekten realer Verfälschungsvorgängen in Bewerbungssituationen sein (Raymark & Tafero, 2009). Pauls und Crost (2005) bestätigen diese Ergebnisse insofern, als dass sie ebenfalls feststellen konnten, dass spezifische berufsbezogene Verfälschungsinstruktionen wesentlich differenziertere Mittelwertsprofile erzeugten, als dies der Fall für allgemeine *faking good*- Instruktionen war, die zu einheitlicher Erhöhung der Big Five-Mittelwerte führten.

Weitere Einflussquellen auf die Effekte der verschiedenen Methoden sind im Folgenden angeführt:

⁶Dabei wurden die Testpersonen gebeten, ein Persönlichkeitsverfahren so zu beantworten, als würden sie sich für eine Stelle als BuchhalterIn bewerben.

⁷Dabei wird der Teilnehmer konkret aufgefordert, ehrlich auf die Items zu antworten.

So zeigen sich Within-Subject-Design⁸-Effekte von Verfälschungsinstruktionen grösser als jene Effekte, welche unter Verwendung von Between-Subject-Designs auftreten (Raymark & Tafero, 2009). Weiters scheint es, dass wahre Bewerber wesentlich taktischer und überlegter beim Verfälschen vorgehen, als solche, die mittels Verfälschungsinstruktion aufgefordert werden zu verfälschen (Raymark & Tafero, 2009). Ausserdem muss angenommen werden, dass Verfälschungsinstruktionen andere kognitive Prozesse hervorrufen, als dies der Fall ist für eine wahre Bewerbungssituation (Raymark & Tafero, 2009). Darüberhinaus führen fehlende Konsequenzen in künstlichen Situationen zu höheren Verfälschungswerten, da Testpersonen im Rahmen der Vorgabe von Verfälschungs-Instruktionen oftmals von der Aufgabe befreit sind, einen möglichst plausiblen Eindruck zu hinterlassen (Steffen, 2004).

Die im Rahmen der vorliegenden Studie eingesetzten Verfälschungsinstruktionen werden zu einem späteren Zeitpunkt (Abschnitt 7.1) näher erläutert.

3.1.4 Methoden zur Vermeidung sowie Berücksichtigung von Verfälschungsverhalten

Im Folgenden sind nun einige wiederholt verwendete Methoden beschrieben, welche dem Ziel dienen, das Auftreten von Verfälschungsverhalten zu verhindern.

1. *Forced-Choice Items*⁹. Christiansen, Edelstein und Fleming (1998) (zitiert nach Jackson et al., 2000) sowie Jackson et al. (2000) und Kanning und Kuhne (2006) konnten zeigen, dass die Verwendung von Forced-Choice-Formaten zu geringerem Verfälschungsverhalten führt und die Vorgabe von Single-Choice-Items zu einer mehr als drei mal so hohen Mittelwertsdifferenz zwischen Standard- und Verfälschungsinstruktionsgruppe führt, wie dies der Fall bei Verwendung von Forced-Choice-Items ist (Jackson et al., 2000). Heggstad et al. (2006) konnten einen derartigen Effekt mittels Forced-Choice-Items zumindest für eine Analyse auf Gruppenniveau bestätigen.

⁸*Within-Subject-Design* bezeichnet das wiederholte Vorgeben von Verfahren an ein und die selbe Person, während *Between-Subject-Designs* die jeweils einmalige Vorgabe von Verfahren an verschiedene Personen beinhaltet.

⁹Items, bei denen eine Testperson gezwungen wird, sich zwischen zwei oder mehreren vorgegebenen, inhaltlich ähnlichen Items eindeutig für die Zustimmung zu einem der vorgegebenen Items zu entscheiden (Jackson et al., 2000).

2. *Warnungen und Ehrlichkeitsaufforderungen*¹⁰ (Fan et al., 2012; McFarland, 2003). McFarland (2003) konnte zeigen, dass der Einsatz von Warnungen nicht nur zur Minderung von Verfälschungsverhalten führt, sondern auch die Korrelationen zwischen den unterschiedlichen Dimensionen verringert, was auf die Minderung des gemeinsamen Einflusses auf die verschiedenen Faktoren, nämlich das Verfälschungsverfahren, hindeutet. Fan et al. (2012) zeigten, dass potenzielle Verfälscher (im Gegensatz zu ehrlichen Testpersonen) ihre Antworten nach Erhalt einer Warnung tatsächlich in die nicht-sozial erwünschte Richtung abmilderten. Khorramdel und Kubinger (2006) hingegen, konnten in ihrer Studie keinen derartigen Effekt durch die Vorgabe von Warnungen erzielen. Somit ist die Wirkung dieser Methode umstritten.

3. *Intransparente oder subtile Items (intransparent, subtle Items)* sind solche Items, aus denen für die Testpersonen nicht direkt ersichtlich ist, welche Beantwortung nun die sozial erwünschtere Antwort darstellt (Honkaniemi, 2012) bzw. welcher zugrundeliegende Faktor mittels des gebotenen Items erhoben werden soll (Hough, 1990). Hough (1990) führt jedoch an, dass solche intransparenten Items oftmals geringere Validitätswerte aufweisen, als dies für transparente Items der Fall ist und weiter, dass die Effektivität solcher Items in Hinblick auf das angestrebte Ziel, nämlich geringeres Verfälschungsverhalten, nicht empirisch belegt ist (Hough, 1990). Thiemann (2006) ergänzt, dass die Wirkungsweise solcher intransparenter Items mit steigender Intransparenz bzw. sinkender Augenscheinvalidität verbessert würde, betont in diesem Zusammenhang jedoch die Problematik der fehlenden Bewerberakzeptanz im Falle geringer Augenscheinvalidität der verwendeten Verfahren (Thiemann, 2006).

4. *Implicit Association Tests (IAT, Implizite Assoziationstests)* basieren auf der Annahme, dass Personen auf stark miteinander assoziierte Begriffe schneller reagieren, als auf zwei lediglich schwach assoziierte Begriffe für welche die Reaktionszeit somit langsamer ausfallen würde (Steffens, 2004).¹¹ Steffens

¹⁰Dabei wird Teilnehmern beispielsweise mitgeteilt, dass Verfälschung der Antworten negative Konsequenzen mit sich brächten, dass sie ihre Antworten zu einem späteren Zeitpunkt rechtfertigen müssten oder dass ein verstecktes Verfahren Verfälschungsverhalten aufdecken würde (Honkaniemi, 2012).

¹¹Beispielsweise werden im Rahmen eines IAT Wortlisten vorgegeben. Wenn eine Testperson den auf dieser Liste vorhandenen Begriff Gewissenhaftigkeit dann schneller dem Begriff ‚Selbst‘

(2004) konnte nachweisen, dass IATs tatsächlich weniger anfällig für Verfälschungsverhalten zu sein scheinen. Jedoch ist anzumerken, dass mit steigender IAT-Testerfahrung und vor allem Wissen um die Erhebung und Relevanz der Reaktionszeit in diesem Verfahren, auch die Verfälschungsanfälligkeit steigt (Steffens, 2004; Röhner et al., 2011; Filipkowski, 2012).

Mueller-Hanson et al. (2006) schlagen auf Basis der Ergebnisse einer von Ihnen durchgeführten Studie vor, das Verfälschungsverhalten durch die Veränderung oder Beeinflussung der Test-Situationswahrnehmung und/ oder persönliche Überzeugungen der Testpersonen bezüglich des Verfahrens oder des Verfahrenskontexts zu verringern.

Ziegler et al. (2007) empfehlen auf Basis ihrer Befunde den Einsatz objektiver Persönlichkeitsverfahren zur Verminderung auftretenden Verfälschungsverhaltens. Auf Basis der Studien von Komar et al. (2010) und Khorramdel und Kubinger (2006) zeigt der Einsatz von Zeitbeschränkungen der Bearbeitungszeit in Kombination mit bestimmten Antwortformaten (dichotom und analog nach Khorramdel & Kubinger, 2006) ebenfalls Wirkung hinsichtlich des verminderten Auftretens von Verfälschungsverhalten.

Weiters existieren solche Methoden, die die Berücksichtigung von Verfälschungsverhalten im Rahmen der Auswertung von empirischen Daten ermöglichen (Reeder & Ryan, 2012). Dabei sind insbesondere die residualisierte Ausprägungsanpassung (Korrektur von Daten auf Basis der Ergebnisse aus Soziale-Erwünschtheits-Skalen; Reeder & Ryan, 2012), der Ausschluss verfälschungsbehafteter Fälle (Reeder & Ryan, 2012; Goffin & Christiansen, 2003) und das wiederholte, verifizierende Testen¹² (Reeder & Ryan, 2012) zu nennen. Insbesondere hinsichtlich eines negativen Einflusses auf die (Konstrukt-) Validität sind diese Methoden jedoch kritisch zu hinterfragen (Ones et al., 1996; Reeder & Ryan, 2012). Die Verwendung Soziale-Erwünschtheits-Skalen zur Korrektur der von Verfälschungsverhalten beeinflussten Daten ist dabei insofern kritisch, als dass die Korrelationskoeffizienten zwischen Verfälschungsverhalten

zuordnet, als einer gegenteiligen Kategorie, dann lässt dies den Rückschluss zu, dass die Testperson sich selbst mit dem Begriff der Gewissenhaftigkeit assoziiert (Filipkowski, 2012).

¹²Testpersonen, welchen auf Basis von Soziale- Erwünschtheit-Skalen Verfälschungsverhalten attestiert wurde, erhalten eine Rückmeldung über die Identifizierung des Verfälschungsverhaltens mit der Bitte um Wiederholung des gleichen Verfahrens (Reeder & Ryan, 2012).

und den Ergebnissen der Soziale-Erwünschtheits-Skalen je nach Autor stark variieren: nach Christiansen (2007, zitiert nach Reeder & Ryan, 2007) liegen diese bei 0,32 bzw. 0,12 (Griffith & Peterson, 2008) und nach Ones (2005) und Viswesvaran und Ones (1999) bei 0,95.

Während nun der in Tabelle 1 (S. 10) als *Response Set* gekennzeichnete Bereich hinreichend erläutert wurde, befasst sich der nun folgende Abschnitt mit den ebenfalls in Tabelle 1 angeführten Antworttendenzen (*Response Styles*), die sich nunmehr auf die formale, Iteminhalts-unabhängige Ebene beziehen.

3.2 Formale Verfälschung – Antworttendenzen (Response Styles)

Eine häufig in Persönlichkeitsfragebögen verwendete Rating-Skala ist die Likert-Skala. Diese von Rensis Likert im Jahr 1932 erstmals entwickelte Skala, enthält typischerweise fünf, im Zustimmungsgrad variierende Antwortmöglichkeiten (Edmondson, 2005). Diese Variation wird dabei häufig zusätzlich mittels Beschriftung der Antwortmöglichkeiten (z.B.: sehr unzutreffend, eher unzutreffend, weder/noch, eher zutreffend, sehr zutreffend) vorgenommen. Die mittlere Antwortskala kann dabei entweder als weitere Abstufung (z.B. „weder/noch“) oder aber als Ausdruck von Unentschiedenheit (z.B. „weiss nicht“) beschrieben und beschriftet werden (Khorramdel & von Davier, under review). Der Vorteil des Einsatzes derartiger Rating-Skalen basiert auf der Möglichkeit interindividueller Vergleiche, geringem Konstruktionsaufwand der Verfahren und der Vereinfachung des Antwortprozesses für den Beantwortenden, indem dieser sich ausschliesslich auf eine Form der Einschätzungsmethode konzentrieren muss (vgl. Khorramdel & von Davier, under review). Jedoch bringt der Einsatz von Ratingskalen auch Nachteile in Form unterschiedlicher Interpretation der Distanzen zwischen den Einschätzungsstufen oder Beschriftungen mit sich (vgl. Khorramdel & von Davier, under review). Weiter ist fraglich, ob einer dem Item angemessener Zustimmungsbereich vorliegt (vgl. Khorramdel & von Davier, under review).

Dem Einsatz von Rating-Skalen liegt darüber hinaus die Annahme zugrunde, dass die antwortende Person ihre Antwort einzig und allein vom Inhalt der Frage abhängig macht. Tatsächlich hängen die Antworten jedoch auch von inhaltsunabhängigen Faktoren ab (Baumgartner & Steenkamp, 2001).

Antworttendenzen (*Response Styles*) welche unter der Verwendung von Rating-Skalen auftreten, sind solche inhaltsunabhängigen Faktoren. Sie stellen die systematische Tendenz dar, auf eine bestimmte, vom Inhalt eines Items unabhängige Art und Weise auf ein Item zu antworten (Vaerenbergh & Thomas, 2012). Damit beeinflussen sie je nachdem welche Antworttendenz vorliegt, den Messfehler (*systematic error*) innerhalb der Messung, univariate Datenverteilungen (über den Einfluss auf Mittelwert und Varianz) sowie multivariate Datenverteilungen (über den Einfluss auf Korrelationen) (Vaerenbergh & Thomas, 2012). Die Verwendung von Antworttendenzen führt in der Folge nicht nur zu Inflation oder Deflation der ermittelten Skalenwerte, sondern als Konsequenz aus dieser Tatsache auch zur möglicherweise fälschlichen Beeinflussung von Entscheidungen, welche auf diesen Skalenwerten basieren (Baumgartner & Steenkamp, 2001). Auch die Ausprägung der Korrelationen zwischen verschiedenen Dimensionen eines Verfahrens, wird durch das Vorliegen von Antworttendenzen beeinflusst (Baumgartner & Steenkamp, 2001). Diese Angaben könnten zu der Annahme verleiten, dass die Vorgabe von Persönlichkeitsverfahren weniger der Erhebung der tatsächlichen Merkmalsausprägung einer Person dient, als vielmehr der Erhebung unterschiedlicher Verwendungsarten der Antwortkategorien (Harzing, 2006). Bezüglich der Einflussgrößen existieren zum jetzigen Zeitpunkt verschiedene Angaben. Weijters, Geuens und Schillewaert (2010c) konnten zeigen, dass Antworttendenzen eine zeitlich stabile Komponente aufweisen, welche unter anderem durch demographische Angaben der Testpersonen beeinflusst wird (Alter, Geschlecht und Ausbildungsniveau). Sie betonen jedoch, dass ein grosser, bisher unerklärter Erklärungsanteil ausserhalb der demographischen Daten läge. Weitere Einflussgrößen betreffen die Anzahl an Antwortkategorien sowie die Art der Beschriftung dieser (Weijters, Cabooter & Schillewaert, 2010a; Kieruj & Moors, 2010) und die Datenerhebungsmethode bezogen auf Paper-Pencil-Verfahren, Telefonerhebungen und Onlineerhebungen (Weijters, Schillewaert & Geuens, 2008). Auch die Kulturabhängigkeit von Antworttendenzen (z.B. Harzing, 2006; Khorramdel & von Davier, under review; Vaerenbergh & Thomas, 2012) wissenschaftlichen gewinnt vor dem Hintergrund zunehmender Globalisierung an Relevanz. Die Ergebnisse diesbezüglicher wissenschaftlicher Studien zeigen, dass die Herkunft bzw. das kulturelle Umfeld

einer Person, einen Einfluss auf die Nutzung von Antwortkategorien zu haben scheinen.

Im folgenden werden nun die in Tabelle 1 (S. 10) bereits aufgelisteten Antworttendenzen näher erläutert.

Antwortbereichstendenz (*Response Range, RR*) beschreibt die Tendenz, einen entweder engen oder zu breiten Bereich an Antwortkategorien um die Mittelkategorie zu wählen (Vaerenbergh & Thomas, 2012; Baumgartner & Steenkamp, 2001).

Netto-Zustimmungstendenz (*Net Acquiescence Response Style/Directional Bias, NARS*) beschreibt die Tendenz, unabhängig vom Item-Inhalt eher Zustimmung als Ablehnung zu äussern (Vaerenbergh & Thomas, 2012; Baumgartner & Steenkamp, 2001).

Ja-Sage-Tendenz/Zustimmungstendenz (*Acquiescence Response Style, ARS*) beschreibt die Zustimmung zu Items, unabhängig vom Item-Inhalt (Vaerenbergh & Thomas, 2012; Baumgartner & Steenkamp, 2001).

Nein-Sage-Tendenz/Ablehnungstendenz (*Disacquiescence Response Style, DARS*) beschreibt die Wahl der ablehnenden Antwortkategorie, unabhängig vom Item-Inhalt (Vaerenbergh & Thomas, 2012; Baumgartner & Steenkamp, 2001).

Willkürliche Antworttendenz (*Noncontingent Responding, NCR*) beschreibt die Tendenz auf zufällige, willkürliche Art und Weise Items zu beantworten (Vaerenbergh & Thomas, 2012; Baumgartner & Steenkamp, 2001).

Milde Antworttendenz (*Mild Response Style, MLRS*) bezeichnet das Vermeiden der höchsten und niedrigsten Antwortkategorien und ist das Gegenstück zur nun folgenden extremen Antworttendenz (Vaerenbergh & Thomas, 2012; Baumgartner & Steenkamp, 2001).

Extreme Antworttendenz (*Extreme Response Style, ERS*) bezeichnet die Verwendung der ersten und letzten Antwortkategorien, unabhängig vom Iteminhalt und stellt das Gegenstück zur MLRS (s. oberhalb) dar (Vaerenbergh & Thomas, 2012; Baumgartner & Steenkamp, 2001).

Tendenz zur Mitte (*Mid-Point Response Style, MRS*) beschreibt die Tendenz, unabhängig vom Iteminhalt die Mittelkategorie zu wählen (Vaerenbergh & Thomas, 2012; Baumgartner & Steenkamp, 2001).

Da die extreme Antworttendenz (ERS) sowie die Tendenz zur Mitte (MRS) essentieller Bestandteil der vorliegenden Arbeit sind, wird auf diese im Weiteren näher eingegangen.

3.2.1 *Extreme Antworttendenz (ERS) und Tendenz zur Mitte (MRS)*

Weijters (2006) (zit. nach Vaerenbergh & Thomas, 2012) differenziert die Ursachen des Auftretens von Antworttendenzen mittels zweier Kriterien. Innerhalb des ersten Kriteriums, dem *Respondent Level* (Befragten-Ebene), wird die Antworttendenz als Charakteristikum des Befragten angesehen. Innerhalb des zweiten Kriteriums, dem *Stimulus Level* (Stimulus-Ebene), wird das Befragungsinstrument als Ursache etwaiger Antworttendenzen angesehen. Abbildung 1 folgt in Anlehnung an Weijters (2006) dieser Aufteilung.

Abbildung 1: Ursachen und Einflussfaktoren von Antworttendenzen

ANTWORTTENDENZ	KONSEQUENZ	URSACHENEBENE	EINFLUSSQUELLE	AUSWIRKUNG	LITERATUR
ERS Extreme Antworttendenz	Überhöht oder verringert Mittelwerte und Varianzen ¹ Verringert das Ausmaß an multivariaten Relationen ²	<i>Stimulus Level</i>	Format/Anzahl der Antwortkategorien	- Keine Auswirkung auf das Auftreten von ERS bei Variation von 5 bis 11 Antwortkategorien¹ - Höhere Anzahl an Antwortkategorien verringert Auftreten von ERS² - Mittlere Antwortkategorie verringert ERS²	¹ Kieruj & Moors, 2013 ² Weijters et al., 2010a
			Benennung der Antwortkategorien	Benennung aller Antwortkategorien verringert ERS³	³ Weijters et al., 2010a
			Datenerhebungsmethode	Auftreten von ERS: Online < Telefon < Paper&Pencil⁴	⁴ Weijters et al., 2008
	^{1,2} Vaerenbergh & Thomas, 2012	<i>Respondent Level</i>	Themenrelevanz	Je relevanter ein Thema für den Befragten, desto stärkeres Auftreten von ERS⁵	⁵ Gibbons et al., 1999
			Personenangaben	- Höheres Alter → stärkere ERS^{6,7} - Mittleres Alter → geringere ERS; Jüngere und Ältere → stärkere ERS⁸	⁶ Kieruj & Moors, 2013 ⁷ Greenleaf, 1992 ⁸ De Jong et al., 2008
				Weibliches Geschlecht → stärkere ERS^{9,10}	⁹ Moors, 2012 ¹⁰ De Jong et al., 2008
				Geringere Ausbildung → stärkere ERS¹¹	¹¹ Greenleaf, 1992
				Geringeres Einkommen → stärkere ERS¹²	¹² Greenleaf, 1992
			Persönlichkeit*	- Extravertierte, starke Meinungs-Vertreter → stärkere ERS¹³ - Self-Regulatory Focus²⁶: Promotion Focus → stärkere ERS¹⁴	¹³ Kieruj & Moors, 2013 ¹⁴ Cabooter, 2010

* Die Einflüsse der Persönlichkeit auf die Verwendung von Antworttendenzen ist insofern kritisch zu betrachten, als dass viele dieser persönlichkeitsbezogenen Ergebnisse selbst mittels Vorgabe von Persönlichkeitsverfahren erhoben werden, welche auch Gegenstand der Verwendung von Antworttendenzen sind (Vaerenbergh & Thomas, 2012).

			Stabilität	ERS ist zeitlich stabil ^{15, 16, 17}	¹⁵ Greenleaf, 1992 ¹⁶ Kieruj & Moors, 2013 ¹⁷ Weijters et al., 2010c
MRS Tendenz zur Mitte	Nähert Mittelwert an Mittelkategorie an u. verringert die Varianz ¹⁹	<i>Stimulus Level</i>	Anzahl der Antwortkategorien	Stärkeres Auftreten von MRS ab 9-10 Antwortkategorien ¹⁸	¹⁸ Kieruj & Moors, 2013
			Datenerhebungsmethode	Auftreten von MRS: Telefon < Paper&Pencil < Online ²⁰	²⁰ Weijters et al., 2008
	Erhöht Ausmaß multivariater Relationen ¹⁹ ¹⁹ Vaerenbergh & Thomas, 2012	<i>Respondent Level</i>	Personenangaben	- Höheres Alter → Stärkere MRS ²¹ - Geringere Ausbildung → Stärkere MRS ²² - Weibliches Geschlecht → stärkere MRS ²³	^{21, 22} Weijters et al., 2010c ²³ Harzing, 2006
			Persönlichkeit*	Self-Regulatory Focus²⁶: Prevention Focus → stärkere MRS ²⁴	²⁴ Cabooter, 2010
			Stabilität	MRS ist zeitlich stabil ²⁵	²⁵ Weijters et al., 2010c

eigene Darstellung

²⁶ **Self-Regulatory Focus-Theorie:** Diese Theorie geht von zwei unterschiedlichen, die menschlichen Handlungs- und Wahrnehmungsweisen beeinflussenden und situationsabhängigen Orientierungs-Foki aus: im *Prevention Focus* stehen Sicherheitsbedürfnisse im Vordergrund. Dabei werden zu verfolgende Ziele als Verpflichtungen empfunden und die Handlungsabsicht ist die Vermeidung von unerwünschten Zuständen oder Fehlern. Im *Promotion Focus* stehen hingegen Hoffnungen und Weiterentwicklung im Vordergrund. Die Handlungsabsicht besteht in der Erreichung von Zielen und Nutzung von Chancen (Higgins, 1997).

Als Einflussquellen für das Auftreten der extremen Antworttendenz (ERS) werden auf *Stimulus-Ebene* hinsichtlich der Antwortkategorien, wie aus Abbildung 1 ersichtlich, das Format, die Anzahl und die Benennung der Skalen-Stufen genannt (Kieruj & Moors, 2013; Weijters et al., 2010a). Dabei führt eine höhere Anzahl an Antwortkategorien, das Einführen einer mittleren Antwortkategorie und die vollständige Benennung der Antwortkategorien zu einem geringerem Ausmaß an ERS (Weijters et al., 2010a). Weiter beeinflusst auch die Datenerhebungsmethode das Auftreten von ERS insofern, als dass nach Weijters et al. (2008) ERS in Paper & Pencil-, Telefon-, und Onlineerhebungen in absteigender Reihenfolge abnimmt. Somit empfiehlt es sich zur Vermeidung von ERS Paper & Pencil-Erhebungen zu umgehen. Hinsichtlich der *Befragten-Ebene* konnten Gibbons et al. (1999) feststellen, dass je höher die Relevanz des erfragten Inhalts, desto stärker tendieren Befragte zur Verwendung der extremen Antwortkategorien. Darüberhinaus scheinen auch verschiedene biographische Merkmale einen Einfluss zu haben. Das Alter der Befragten hat nach Greenleaf (1992), Kieruj und Moors (2013) und DeJong et al. (2008) insofern einen Einfluss, als dass sowohl ältere als auch jüngere Testpersonen, im Gegensatz zu Testpersonen der dazwischen angesiedelten Altersgruppen, verstärkt auf die Verwendung der extremen Antwortkategorien zurückgreifen. Das Geschlecht (weibliche Befragte tendieren zu stärker ausgeprägter ERS; Moors et al., 2012; DeJong et al., 2008), sowie die Ausbildung (geringerer Ausbildungsstatus führt zu stärkerer ERS; Greenleaf, 1992) und das Einkommen (geringeres Einkommen führt zu erhöhter ERS; Greenleaf, 1992) scheinen ebenfalls Einflussquellen auf Befragten-Ebene darzustellen. Darüberhinaus gibt es Persönlichkeitsfaktoren, welche auf das Auftreten von ERS förderlich wirken. Dies bezieht sich auf extravertierte und starke Meinungsvertreter (Kieruj & Moors, 2013) und Menschen mit sog. *Promotion Focus* (Cabooter, 2010; zur Erklärung s. Fussnote 26, S. 30). Verschiedene Autoren (Greenleaf, 1992; Kieruj & Morris, 2011; Weijters et al., 2010c) konnten darüberhinaus die zeitliche also intrapersonale Stabilität der Verwendung extremer Antworttendenzen feststellen, was auf einen Bezug zwischen Persönlichkeit und Antwortstil eines Befragten hindeutet.

Die wissenschaftlichen Ergebnisse zur mittleren Antworttendenz oder Tendenz zur Mitte (MRS) lassen sich anhand ähnlicher Kriterien darstellen. So hat auf *Stimulus-Ebene* erneut die Anzahl an Antwortkategorien einen Einfluss insofern,

als dass nach Kieruj und Moors (2013) ab dem Einsatz von neun Antwortkategorien, MRS verstärkt auftritt. Selbstverständlich muss zum Auftreten von MRS eine Mittelkategorie vorhanden sein. Verzichtet man auf eine Mittelkategorie, mit dem Zweck das Auftreten von MRS zu umgehen, muss jedoch beachtet werden, dass dies nach Weijters et al. (2010a) zum Anstieg von ERS führt. Des Weiteren hat auch die Datenerhebungsmethode einen Einfluss. Dabei tritt MRS im Rahmen von Online-, Paper & Pencil-, sowie Telefonumfragen in absteigender Reihenfolge in geringerem Ausmaß auf (Weijters et al., 2008). Wenn man sich die Angaben bezüglich der Datenerhebungsmethode für ERS anschaut, so stellt man fest, dass die günstigsten Erhebungsmethoden bezüglich der Vermeidung der beiden Antworttendenzen unterschiedlich ausfallen (Online-Erhebung (ERS) vs. Telefon (MRS)). So kann die Vermeidung der einen Antworttendenz mittels entsprechender Erhebungsmethoden-Wahl womöglich zur Begünstigung der anderen Antworttendenz führen. Auf *Befragten-Ebene* zeigen sich erneut Ausbildung (je geringer die Ausbildung, desto stärkere MRS; Weijters et al., 2010c), Alter (höheres Alter führt zu stärkerer MRS; Weijters et al., 2010c) und Geschlecht (weibliches Geschlecht tendiert zu stärkerer MRS; Harzing, 2006) der Befragten als Einflussquelle. Cabooter (2010) konnte hinsichtlich der Persönlichkeit als Einflussfaktor feststellen, dass Personen mit stärkerem *Prevention-Focus* (zur Erklärung s. Fussnote 26, S. 30) eher zur Verwendung der mittleren Antwortkategorie neigen, vermutlich da die mittlere Antwortkategorie die Flucht vor einer klaren Meinungskundgabe und damit verbundenen Konsequenzen darstellt. Und wie schon die extreme Antworttendenz, zeigt sich auch die mittlere Antworttendenz als zeitlich stabil und somit womöglich verankert in der Persönlichkeit (Weijters et al., 2010c).

Die Feststellung dieser Einflussquellen dient in erster Linie dazu, die Korrektur von mittels Antworttendenzen verzerrten Datensätzen und die Vermeidung von Antworttendenzen möglich zu machen. Diese Aspekte werden im folgenden Abschnitt 3.2.2 näher beleuchtet.

3.2.2 Erhebung, Vermeidung und Berücksichtigung von ERS und MRS

Vaerenbergh und Thomas (2012) zählen die gängigsten Methoden zur Erhebung und Berücksichtigung von Antworttendenzen bei der Ergebnisinterpretation auf:

1. *Zählmethode*. Hierbei werden die Anzahl der extremen Antworten (im Sinne von ERS) und/oder mittlere Antworten (im Sinne von MRS) im gesamten Fragebogen gezählt. Sodann wird der Prozentsatz der entsprechenden Antworttendenzen in Bezug gesetzt, zu den Antworten welche nicht der jeweiligen Antworttendenz entsprechen (Reynolds & Smith, 2010). Diese Methode benötigt heterogene, inhaltlich voneinander unabhängige Items (Morren et al., 2011; Vaerenbergh & Thomas, 2012). Überaus problematisch ist allerdings die Tatsache, dass bei diesem Ansatz jener Anteil der Antworten, welcher nicht durch die Verwendung von Antworttendenzen, sondern tatsächlich inhaltsabhängig zustande kommt, ignoriert wird (de Beuckelaer et al., 2010), so dass sich vor diesem Hintergrund zahlreiche Ansätze entwickelt haben, die diese Problematik zu umgehen versuchen und im Folgenden erläutert werden.

2. *Latent Class Regression Analysis*. Diese Methode, u.a. eingesetzt von Moors (2010), dient der Identifikation latenter Klassen von Personen, welche über die unterschiedliche Verwendung der Antwortkategorien im Sinne von Antworttendenzen in Gruppen zusammengefasst werden können (Moors, 2010). Kritisch zu betrachten ist nach Vaerenbergh und Thomas (2012) und Moors (2010) jedoch die Tatsache, dass sowohl Fachkenntnisse zum Einsatz dieser Methode, als auch spezielle Software zur Durchführung der Modellberechnungen von Nöten sind.

3. *Latent Class Confirmatory Factor Analyse (LCFA)*. Die LCFA kann zur Erhebung von ERS eingesetzt werden. Sie wird über die Modellierung der Antworttendenzen als latente Faktoren, neben den eigentlichen inhaltlichen Faktoren durchgeführt (Moors, 2012; Morren et al., 2011). Dabei laden alle inhaltlichen Faktoren auf den Antworttendenz-Faktoren. Über die erhöhte Modellpassung wird sodann das Vorliegen von Antworttendenzen überprüft. Vaerenbergh und Thomas (2012) merken auch hinsichtlich dieser Methode kritisch an, dass sowohl Fachkenntnisse zum Einsatz dieser Methode als auch spezielle Software zur Durchführung der Modellberechnungen von Nöten sind.

4. *Representative Indicators for Response Styles – Methode* (RIRS; *Repräsentative Antworttendenz-Indikatoren*). Hierbei werden einige unkorrelierte, heterogene Items in den bestehenden Fragebogen aufgenommen und deren gewichtete Antworttendenz-Indikatoren berechnet (de Beuckelaer et al., 2010; Vaerenbergh & Thomas, 2012). Der dahinterstehende Gedanke ist der, dass diese zusätzlichen heterogenen Items den Konflikt der Konfundierung inhaltsabhängiger Antworten und inhaltsunabhängiger Antworttendenzen bereinigen. Diese Methode kann sowohl zur Aufdeckung als auch zur Berücksichtigung von ERS und MRS bei der Ergebnisinterpretation eingesetzt werden (Vaerenbergh & Thomas, 2012). Nachteilig an dieser Methode ist laut Vaerenbergh und Thomas (2012) die Tatsache, dass zusätzliche Items in bestehende Verfahren eingefügt werden müssen.

5. *Item-Response-Theorie (IRT)*. Die Anwendung der Item-Response-Theorie führt zur Modellierung der Lösungswahrscheinlichkeit eines bestimmten Items, welche eine Funktion der zugrundeliegenden latenten Variablen bildet (Vaerenbergh & Thomas, 2012). Der Vorteil der Verwendung von Item-Response-Modellen ist, dass diese sowohl zwischen der tatsächlichen Merkmalsausprägung einer Testperson und der Verwendung von Antworttendenzen unabhängig von der tatsächlichen Merkmalsausprägung differenzieren, als auch psychometrische Itemcharakteristiken in Hinblick auf die Verwendung von Antworttendenzen berücksichtigen (Bolt & Johnson, 2009).

Da derartige Modelle einen relevanten Punkt der vorliegenden Arbeit bilden, ist diesen ein eigener Abschnitt (Abschnitt 5) gewidmet.

Hinsichtlich der Vermeidung des Auftretens oder des Einflusses von ERS und MRS lassen sich weiter aus Abbildung 1 einige Empfehlungen ableiten. Dabei sind die Auswirkungen auf Stimulus-Ebene eher bei der Fragebogenentwicklung und -konstruktion zu berücksichtigen und dienen der Vermeidung von Antworttendenzen an der Ursprungsquelle. Auswirkungen auf Befragten-Ebene werden bei der Auswertung oder der Interpretation der ausgewerteten Daten berücksichtigt, somit nicht an der Quelle des Ursprungs verhindert.

Legt man beispielsweise die wissenschaftlichen Erkenntnisse von Weijters et al. (2010a) zugrunde, so empfiehlt es sich die Anzahl an Antwortkategorien in Abhängigkeit der Art der Befragung auf mindestens fünf bis sieben Antwortkategorien festzulegen und diese vollständig zu benennen. Die hohe

Anzahl an Antwortkategorien fördert nach Kieruj und Moors (2013) das Auftreten von MRS, jedoch ist dies erst ab neun Antwortkategorien der Fall. Darüberhinaus sollten die Antwortkategorien zur Verringerung des Auftretens von ERS eine Mittelkategorie beinhalten, welche in der Folge selbstverständlich das Auftreten von MRS bedingt (Weijters et al., 2010a). Weijters et al. (2008) konnten zeigen, dass ERS in Paper & Pencil-Erhebungen häufiger auftritt, als in Telefon-Befragungen und das Auftreten bei Onlineerhebungen am geringsten ist. MRS tritt im Gegensatz dazu am häufigsten bei Onlineerhebungen, gefolgt von Paper&Pencil-Erhebungen auf. Um MRS zu vermeiden eignet sich somit insbesondere die Telefonumfrage (Weijters et al., 2008).

Abschliessend ist zu sagen, dass das Auftreten von Antworttendenzen auch mittels diesbezüglicher Optimierung eines Fragebogens nicht zu vermeiden ist. Aus Abbildung 1 ist anhand der Vielfältigkeit der Einflussgrössen die Komplexität der Thematik abzulesen. Die wissenschaftliche Erkenntnislage zum jetzigen Zeitpunkt ist somit zu gering, als dass sichere Methoden zum Umgang mit Antworttendenzen abgeleitet werden könnten. Es scheint somit sinnvoller, Erkenntnisse zu den Mechanismen und Ursachen für das Auftreten von Antwortverhalten und Antworttendenzen weiter zu vertiefen und auf Basis dessen die Erhebungsmethoden anzupassen und die Auswertung sowie Interpretation von Daten diesbezüglich zu optimieren. Ein vielversprechender Ansatz ist hierbei die Verwendung von multidimensionalen IRT-Modellen, welche in Abschnitt 5 näher erläutert wird.

4 Zusammenhang von Antworttendenzen und stereotypem Verfälschungsverhalten

Der folgende Abschnitt dient der Erläuterung eines möglichen Zusammenhangs zwischen Antworttendenzen und Verfälschungsverhalten. Ein entsprechender Zusammenhang wurde kaum untersucht und konnte bisher auch nicht nachgewiesen werden. Ziegler (2009) weist mit Verweis auf Rost et al. (1999) jedoch auf einen möglichen Zusammenhang zwischen der Verwendung extremer Antworttendenzen (ERS) und Verfälschungsverhalten hin: „The extent to which these response styles relate to slight and extreme faking should be examined

using large samples. Otherwise, we might be talking of the same phenomena using different labels.” (Ziegler, 2009, S. 34).

Diese Aussage von Ziegler (2009) beschreibt im Kern das Vorhaben der vorliegenden Arbeit, nämlich festzustellen, ob ein Zusammenhang zwischen bestimmten Antworttendenzen (ERS und MRS) und dem Auftreten von Verfälschungsverhalten besteht.

Obwohl zu der Frage nach dem Zusammenhang zwischen Antworttendenzen und Verfälschungsverhalten kaum wissenschaftliche Literatur vorliegt, liefern van Hooft und Born (2012) einen Hinweis, der die zu prüfende Annahme über einen derartigen Zusammenhang zulässt: van Hooft und Born (2012) versuchten mittels der Verfolgung von Augenbewegungen der teilnehmenden Personen, Informationen über die kognitiven Vorgänge während der Bearbeitung eines Persönlichkeitsfragebogens zu erheben. Sie ließen die Testpersonen ein Persönlichkeits- sowie ein Integritätsverfahren, jeweils unter Verfälschungsinstruktion und unter Ehrlichkeitsaufforderung bearbeiten. Dabei stellten sie fest, dass die Testpersonen unter Verfälschungsinstruktion die beiden extremen Antwortkategorien sowohl länger, als auch schneller nach Lesen des Items fixierten. Gleichzeitig wurden unter der Ehrlichkeitsaufforderungs-Bedingung längere Fixationszeiten für die mittlere Kategorie und geringere Fixationszeiten für die extremen Antwortkategorien festgestellt (van Hooft & Born, 2012). Daraus schließen van Hooft und Born, dass Verfälschungsverhalten zur verstärkten Wahl extremer Antwortkategorien führen bzw. Verfälschungsverhalten mit der Verwendung extremer Antworttendenzen gleichzusetzen sein könnte.

Der Annahme dass Verfälschungsverhalten und Antworttendenzen möglicherweise ein und dasselbe Phänomen darstellen bzw. ein entsprechender Zusammenhang besteht, muss die Hypothese zugrunde gelegt werden, dass Verfälschungsverhalten ein stereotypes (über verschiedene Fragebogeninhalte hinweg konstantes) Antwortverhalten darstellt. Um die Annahme zu prüfen, ob es einen Zusammenhang zwischen Verfälschungsverhalten und Antworttendenzen gibt wurde im Rahmen der vorliegenden Studie unter Vorgabe unterschiedlicher (Verfälschungs-) Instruktionen versucht, bestimmte Arten von Verfälschungsverhalten auf Seiten der Testpersonen zu induzieren, um dann

mittels multidimensionaler IRT-Modelle untersuchen zu können, ob bestimmte Antworttendenzen in den Daten vorliegen.

Der nun folgende Abschnitt enthält zum besseren Verständnis eine Einführung zur Verwendung von IRT-Modellen. Darüberhinaus wird im letzten Teil des Abschnitts auf den in der vorliegenden Arbeit verwendeten Ansatz unter Verwendung multidimensionale IRT-Modell (MIRT) eingegangen.

5 (Multidimensionale) Item-Response-Theorie-Modelle zur Untersuchung von Antworttendenzen

Zur Verwendung von aus der Item-Response-Theorie (IRT) hervorgegangenen Modellen zur Untersuchung und Messung von Antworttendenzen existieren bereits zahlreiche wissenschaftliche Untersuchungen (z.B. Austin et al., 2006; Johnson & Bolt, 2010; Bolt & Newton, 2011; Böckenholt, 2012; Bolt & Johnson, 2009, DeJong et al., 2008). Die Verwendung von IRT-Modellen bietet gegenüber anderen Methoden zur Messung von Antworttendenzen erhebliche Vorteile. So bietet die Verwendung von IRT-Modellen die Möglichkeit, im Gegensatz zu in Abschnitt 3.2.2 erläuterten Methoden sowohl zwischen der tatsächlichen Merkmalsausprägung einer Testperson und der Verwendung von Antworttendenzen unabhängig von der tatsächlichen Merkmalsausprägung zu differenzieren, als auch psychometrische Itemcharakteristiken in Hinblick auf die Verwendung von Antworttendenzen zu berücksichtigen (Bolt & Johnson, 2009).

DeJong et al. (2008) zeigen beispielsweise, dass ein auf der IRT basierendes Latent-Trait-Modell mit der Differenzierung in Item- und Personenparameter ein sinnvolles Instrument zur Feststellung von Antworttendenzen darstellt, da nach DeJong et al. (2008) Antworttendenzen genau durch die Interaktion von persönlicher Disposition und Itemcharakteristik zustande kommen. Dabei wird die zugrundeliegende Eigenschaft (*latent trait*) durch die personenabhängig variierende Nutzung extremer Antwortkategorien repräsentiert. Des Weiteren erlaubt das Modell solche Items mittels Itemdiskriminanzparameter zu differenzieren, die sich besser oder schlechter für die Identifikation von ERS eignen (DeJong et al., 2008). Somit lassen sich aus dem Modell direkte

Implikationen zur Verfahrenskonstruktion sowie Antwortverzerrungsbereinigung ableiten.

Bolt und Johnson (2009) präsentieren ein multidimensionales IRT-Modell, welches eine Antworttendenz (ERS) als explizite statistische Dimension behandelt. Die Antworttendenz als Merkmalsausprägung beeinflusst die Attraktivität der verschiedenen Antwortkategorien und somit die Frequenz mit welcher eine bestimmte Antwortkategorie gewählt wird (Bolt & Johnson, 2009). Das Modell stellt Itemantworten als Funktion aus Merkmalsausprägung und Antworttendenzen dar (Bolt & Johnson, 2009), aus denen sich die Schätzung der eigentlichen Merkmalsausprägung und ERS ableiten lässt. Bolt und Johnson (2009) zeigen mit ihrem Ansatz ausserdem eine Möglichkeit, Antworttendenzen als gruppenabhängigen Indikator (*Differential Item Functioning*, DIF) zu anderen gruppenabhängigen Indikatoren, nämlich der Itemcharakteristik, abzugrenzen.

Austin et al. (2006) verwenden ein Mixture-Rasch-Modell zur Analyse von durch die unterschiedliche Nutzung der Antwortkategorien charakterisierten Sub-Gruppen oder latenten Klasse mit unterschiedlichen Personen- und Itemparametern (statt Individuen). Sie zeigten, dass für drei der fünf Dimensionen des Fünf-Faktoren-Modells (Neurotizismus, Verträglichkeit, Gewissenhaftigkeit) zwei Subgruppen im Sinne der Verwendung von Antwortkategorien bzw. Antworttendenzen (ERS vs. MRS) existieren.

Der in der vorliegenden Arbeit verwendete Ansatz aus der Item-Response-Theorie basiert auf einem Ansatz von Böckenholt (2012), erweitert durch Khorramdel und von Davier (under review). Den Kern dieses Modells stellt dabei die Zerlegung des Antwortprozesses in multiple sequentielle Antwortsubprozesse dar. D.h. es wird kein singulärer Antwortprozess angenommen, sondern mehrere sequentielle Entscheidungen, welche die finale Antwortentscheidung für eine der Antwortkategorien einer Ratingskala oder Likert-Skala bedingen. Um diesen multiplen Entscheidungsprozess darzustellen, verwendet Böckenholt (2012) einen Entscheidungsbaum. Dieser Entscheidungsbaum ist in Anlehnung an Khorramdel und von Davier (under review) für eine 5-stufige Likert-Skala im Folgenden dargestellt und beschreibt drei mögliche Antwortsubprozesse:

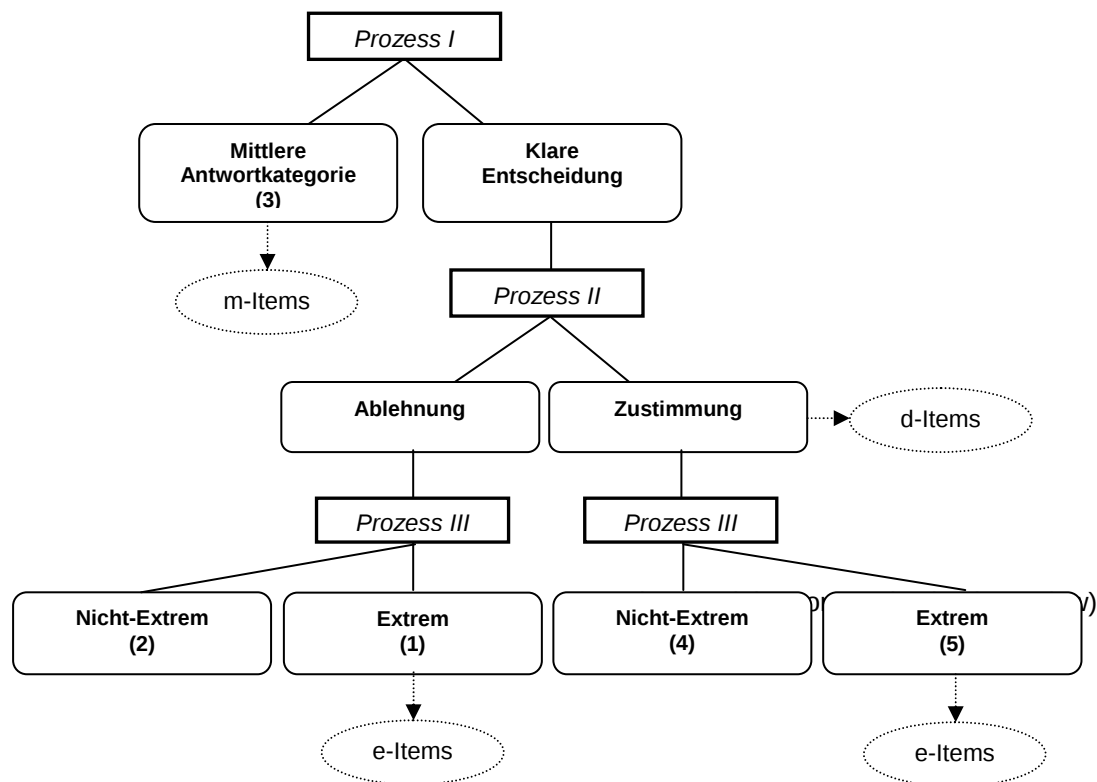


Abbildung 2: Multipler Entscheidungsprozess - Entscheidungsbaum

Der in Abbildung 2 dargestellte *Prozess I* beinhaltet zunächst eine klare Entscheidung für eine Zustimmung oder Ablehnung im Gegensatz zu einer Entscheidung für die Mittelkategorie. Dabei bezeichnet (3) die Position der gewählten Antwortkategorie innerhalb der Likertskala (1-2-3-4-5). Wird eine klare Entscheidung getroffen, so steht innerhalb des *Prozesses II* die Richtungsentscheidung hinsichtlich des Items an. Es muss also grundsätzlich Zustimmung oder Ablehnung beschlossen werden. Ein weiterer möglicher Entscheidungsprozess, *Prozess III*, beinhaltet die Intensität der Entscheidung (extrem vs. nicht-extrem).

Khorramdel und von Davier (under review) modellieren die verschiedenen Antwortprozesse in Anlehnung an Böckenholt (2012) mit sog. Pseudoitems, um auf Basis dieser Pseudoitems eine Modellschätzung vornehmen zu können. Diese Pseudoitems sind in Abbildung 2 mittels gestrichelter Ovale dargestellt. Pseudo-Item m (midpoint) repräsentiert Antworten, bei denen die mittlere Antwortkategorie gewählt wurde und Pseudo-Item e (extreme) repräsentiert die Wahl der extremen Antwortkategorie in positiver oder negativer Richtung (1 oder

5 auf der Likert-Skala). Pseudo-Item d (direction) repräsentiert extreme und nicht extreme Antworten in positiver Richtung bzw. Antworten, die für eine positive Ladung auf dem jeweils zu messenden Konstrukt stehen.

Während der Ansatz von Böckenholt (2012) sich lediglich auf die Analyse eines Fragebogens mit nur einer Skala bezieht und Antworttendenzen als mehrdimensionale Faktoren modelliert, erweitern Khorramdel und von Davier (under review) diesen Ansatz auf die Anwendung eines multidimensionalen Fragebogens mit fünf Skalen und modellieren Antworttendenzen nicht nur als mehrdimensionale, sondern auch als eindimensionale Faktoren. Ziel ihrer Arbeit war, mittels Pseudoitems zu prüfen, ob Antworttendenzen zur Mitte oder zu Extremurteilen (MRS oder ERS) vorliegen und sich von tatsächlich konstruktbezogenen Antworten differenzieren lassen, sowie einen Score für die einzelnen Fragebogenskalen zu finden, welcher von diesen Antworttendenzen nicht bzw. weniger verzerrt ist. Die fünf Fragebogendimensionen stellen die Dimensionen des Five-Factor-Modells (FFM) nach McCrae und Costa (1987) dar: Verträglichkeit, Emotionale Stabilität, Gewissenhaftigkeit, Extraversion und Offenheit. Um zu untersuchen, inwiefern die Pseudo-Items in der Lage sind, zwischen Antworttendenzen und merkmalsbezogenen Antworten zu differenzieren, wurden eindimensionale IRT-Modelle mit multidimensionalen IRT-Modellen (MIRT) auf Basis des 2PL-Modells verglichen.

Auf Basis der oben genannten Pseudo-Items (e,d und m) stellen Khorramdel und von Davier (under review) folgende Hypothesen auf: wenn extreme und mittlere Antworttendenzen in den Daten vorliegen, dann lassen sich diese über die Pseudo-Items e und m messen und jeweils eindimensional bzw. konsistent über die fünf Dimensionen des FFM hinweg abbilden, während Pseudo-Items d, unbeeinflusst (oder kaum beeinflusst) von diesen beiden Antworttendenzen, die Big Five Dimensionen erhebt.

Die Ergebnisse von Khorramdel und von Davier (under review) zeigen, dass sich die extreme Antworttendenz mittels Pseudoitems und über alle Fragebogendimensionen hinweg eindimensional erfassen lässt, während das für Pseudo-Items m nicht zutrifft. Da die Interkorrelationen der fünf Persönlichkeitsdimensionen sowohl auf Basis der Pseudo-Items e als auch auf Basis der Pseudo-Items m hoch ausfallen, wird für Pseudo-Items m die Vermutung aufgestellt, dass diese beides erfassen, die mittlere Antworttendenz

sowie konstruktrelevante Antworten zu den fünf Persönlichkeitsdimensionen. Zusätzlich konnten Khorramdel und von Davier (under review) zeigen, dass Pseudo-Item d die fünf Persönlichkeitsdimensionen adäquat erfasst und zu geringeren Skaleninterkorrelationen führt, als die Verwendung der fünf Antwortkategorien (gem. Likert-Skala, wie im Fragebogen vorgesehen).

Ziel der vorliegenden Arbeit ist nun mittels dieses Ansatzes den zu diesem Zweck eigens erhobenen Datensatz auf Antworttendenzen zu prüfen und einen möglichen Zusammenhang zu (instruiertem) Verfälschungsverhalten zu untersuchen.

III. EMPIRISCHER TEIL

6 Ziel der Untersuchung

Das Ziel der vorliegenden Arbeit ist es, einen möglichen Zusammenhang zwischen Verfälschungsverhalten und der extremen und mittleren Antworttendenz in Persönlichkeitsfragebögen zu untersuchen. Dazu wurde ein entsprechendes Untersuchungs-Design (Between-Subject-Design¹³) aufgestellt, wobei in Anlehnung an unterschiedliche Situationsbedingungen drei Arten von Verfälschungsverhalten (hohe, unsichere und geringe Faking-Motivation bzw. ehrliche Beantwortung) mittels Verwendung von Verfälschungs-Instruktionen (s. dazu Abschnitt 3.1.3, S.20) erzeugt werden sollten. Das Experiment wurde an einer Stichprobe aus Studenten der Rechts- und Wirtschaftswissenschaften unterschiedlicher Universitäten sowie Fachhochschulen in Österreich und Deutschland durchgeführt. Die Instruktionen sind im folgenden Abschnitt 6.1 gemeinsam mit den darauf aufbauenden Hypothesen beschrieben.

6.1 Hypothesen

Über die drei unterschiedlichen Verfälschungs-Instruktionen sollten drei Arten von Verfälschungsverhalten hervorgerufen werden, deren Zusammenhang mit Antworttendenzen in Persönlichkeitsfragebögen (unter Verwendung einer Ratingskala) untersucht werden sollte. Für jedes Verfälschungsverhalten wurde vorab eine Hypothese aufgestellt, inwiefern dieses mit Antworttendenzen bzw. mit welcher Art von Antworttendenz zusammenhängen könnte. Die Instruktionen und darauffolgende Persönlichkeitsfragebögen wurden den drei Versuchsgruppen (VG) von Studierenden im Rahmen eines Onlinefragebogens vorgegeben.

Die erste Instruktion (VG 1: hohe Verfälschungs-Motivation) stellte eine Bewerberauswahlsituation dar, wobei der Testperson mitgeteilt wurde, dass es um eine für sie äusserst erstrebenswerte Position ginge und sie den

¹³Ein *Between-Subject-Design* beinhaltet die jeweils einmalige Vorgabe von Verfahren an verschiedene Personen, im Gegensatz zu einem *Within-Subject-Design*, welches die mehrmalige Vorgabe an den selben Personenkreis bezeichnet.

Persönlichkeitsfragebogen im Rahmen eines Assessment-Centers beantworten würde. Die Testperson wurde aufgefordert, sich beim Beantworten der Fragebogenitems auf eine Art darzustellen, die die Wahrscheinlichkeit die angebotene Position zu erhalten, erhöhen würde. Ziel war dabei das Induzieren starken Verfälschungsverhaltens, begründet durch eine hohe Motivation, die Stelle erhalten zu wollen.

Die Situation stellte sich im Rahmen der zweiten Instruktion (VG 2: unsichere Verfälschungs-Motivation) ähnlich dar, jedoch wurde den Testpersonen hier mitgeteilt, sie seien sich unsicher ob sie die Stelle haben wollten. Ziel war es dabei, eine unsichere Verfälschungsmotivation herbeizuführen, die in der Folge zu geringerem Verfälschen im Vergleich zu VG 1 führen würde.

Instruktion 3 (VG 3: geringe Verfälschungs-Motivation bzw. ehrliches Antwortverhalten) stellte eine freiwillige Berufsberatungssituation dar, im Rahmen derer ein Persönlichkeitsfragebogen vorgegeben würde, um so persönliche Stärken und Schwächen in Hinblick auf eine zukünftige Berufstätigkeit feststellen zu können. Dabei wurde der Testperson mitgeteilt, dass ehrliches Antwortverhalten in dieser Situation den grössten Nutzen bringen würde. Vor diesem Hintergrund wurde erwartet, dass die Testpersonen eine geringe Verfälschungsmotivation aufweisen würden, was in der Folge zu eher ehrlichem Antwortverhalten statt Verfälschungsverhalten führen würde.

Es wird vermutet, dass starkes Verfälschungsverhalten einen Zusammenhang mit extremen Antworttendenzen (vgl. Ziegler, 2009; van Hooft & Born, 2012) aufzeigt, während geringeres Verfälschungsverhalten eher einen Zusammenhang mit der Antworttendenz zur Mitte aufweisen könnte. Im Rahmen der vorliegenden Arbeit werden daher die folgenden Hypothesen untersucht:

Hypothese 1a

H_{(0) 1a}: Unter Vorgabe der **1. Instruktion (VG 1)** zeigt sich kein erhöhtes Ausmaß an Verfälschungsverhalten (sozial erwünschtes Antwortverhalten, *Impression Management*) bei der Beantwortung der Fragebogenitems, verglichen mit der 2. und 3. Instruktion.

H_{(1) 1a}: Unter Vorgabe der **1. Instruktion (VG 1)** zeigt sich ein signifikant erhöhtes Ausmaß an Verfälschungsverhalten (sozial erwünschtes Antwortverhalten, *Impression Management*) bei der Beantwortung der

Fragebogenitems, verglichen mit der 3. Instruktion, sowie ein erhöhtes Ausmaß an Verfälschungsverhalten verglichen mit der 2. Instruktion.

Hypothese 1b

H_{(0) 1b}: Unter Vorgabe der **1. Instruktion (VG 1)** bzw. bei starkem Verfälschungsverhalten, tritt keine vermehrte Antworttendenz zu Extremurteilen (ERS) auf.

H_{(1) 1b}: Unter Vorgabe der **1. Instruktion (VG 1)** bzw. bei starkem Verfälschungsverhalten, tritt vermehrt die Antworttendenz zu Extremurteilen (ERS) auf.

Hypothese 2a

H_{(0) 2a}: Unter Vorgabe der **2. Instruktion (VG 2)** zeigt sich kein erhöhtes Ausmaß an Verfälschungsverhalten (sozial erwünschtes Antwortverhalten, *Impression Management*) bei der Beantwortung der Fragebogenitems verglichen mit der 3. Instruktion.

H_{(1) 2a}: Unter Vorgabe der **2. Instruktion (VG 2)** zeigt sich ein signifikant erhöhtes Ausmaß an Verfälschungsverhalten (sozial erwünschtes Antwortverhalten, *Impression Management*) bei der Beantwortung der Fragebogenitems verglichen mit der 3. Instruktion, jedoch geringeres Verfälschungsverhalten im Vergleich zur 1. Instruktion.

Hypothese 2b

H_{(0) 2b}: Unter Vorgabe der **2. Instruktion (VG 2)** bzw. bei geringerem Verfälschungsverhalten oder Unsicherheit, tritt keine vermehrte Antworttendenz zur Mitte (MRS) auf.

H_{(1) 2b}: Unter Vorgabe der **2. Instruktion (VG 2)** bzw. bei geringerem Verfälschungsverhalten oder Unsicherheit, tritt vermehrt die Antworttendenz zur Mitte (MRS) auf.

Hypothese 3

H_{(0) 3}: Unter Vorgabe der **3. Instruktion (VG 3)** bzw. bei ehrlicher Beantwortung des Fragebogens, treten vermehrt Antworttendenzen (ERS oder MRS) auf.

H_{(1) 3}: Unter Vorgabe der **3. Instruktion (VG 3)** bzw. bei ehrlicher Beantwortung des Fragebogens, treten keine bzw. kaum Antworttendenzen (ERS oder MRS) auf.

Die Überprüfung der Hypothesen erfolgte im Rahmen eines zweistufigen Prozesses: nach der Datenerhebung mittels Online-Fragebogen unter Vorgabe der drei Verfälschungsinstruktionen wurde in einem ersten Schritt überprüft, ob die Verfälschungsinstruktionen zu dem jeweils hypothetisierten Verfälschungsverhalten geführt haben. In einem zweiten Schritt wurden die Daten der drei Versuchsgruppen mittels IRT-Analysen auf Antworttendenzen untersucht (ERS, MRS).

In den nun folgenden Abschnitten finden sich detaillierte Angaben zu Untersuchungsplan (Abschnitt 7.1), Erhebungsinstrumenten (Abschnitt 7.2), Durchführung (Abschnitt 7.3) sowie der teilnehmenden Stichprobe (Abschnitt 7.4). Abschnitte 8 und 9 beinhalten so dann die Ergebnisauswertung der erhobenen Daten.

7 Methode

7.1 Untersuchungsdesign

Die Vorbereitung der wissenschaftlichen Untersuchung innerhalb der vorliegenden Arbeit untergliederte sich in die folgenden Stufen:

- a. Erstellung der Fragebogen-Instruktionen und der Rückmeldung**
- b. Programmierung des Online-Fragebogens bestehend aus den drei Verfahren BIP, NEO-FFI und SES-17**
- c. Pilottestung von Instruktionen und Onlinefragebogen**

Sodann erfolgte die Durchführung der Studie.

Obige Schritte sind im folgenden erläutert:

a. Erstellung der Fragebogen-Instruktionen und der Rückmeldung

Trotz berechtigter Kritik am Einsatz von Verfälschungsinstruktionen in künstlichen Situationen (s. Abschnitt 3.1.3, S. 20), wurden in der vorliegenden Studie Verfälschungsinstruktionen eingesetzt. Zunächst hat dies pragmatische Gründe: die Verfügbarkeit reeller Bewerber für derartige wissenschaftliche Untersuchungen ist begrenzt. Dies liegt zum einen daran, dass Unternehmen sowohl aus rechtlichen als auch aus ethischen Gründen externen Personen ungern Zugang zum Bewerbungsprozess gewähren oder diese in den Bewerbungsprozess eingreifen lassen. Darüber hinaus könnte sich ein derartiger Eingriff wohl auch auf die Reaktion des Bewerbers auf das Unternehmen in eine nicht gewünschte negative Richtung auswirken.

Über diese pragmatischen Gründe hinaus gibt es jedoch weitere Gründe, die die Qualität der erhobenen Daten beeinflussen: so lässt die Verwendung von Verfälschungsinstruktionen unter künstlichen Bedingungen die Kontrolle der Art der Verfälschung zu, was bei Testung tatsächlicher Bewerber nicht möglich ist. Des Weiteren kann eine Untersuchung an einer homogenen Stichprobe (z.B. wie in der vorliegenden Arbeit ausschliesslich Studierende der Rechts- und Wirtschaftswissenschaften) durchgeführt werden, als dies für eine reelle Bewerbungssituation der Fall wäre. Weiters ist die Anzahl an Bewerbern innerhalb eines Unternehmens meist nicht ausreichend groß, um eine wie im vorliegenden Fall auf Basis der Item-Response-Theorie durchgeführte Studie zuzulassen. Eine Verbreitung der Datenerhebung auf mehrere Unternehmen würde hingegen sowohl die Homogenität der Stichprobe, als auch die Machbarkeit der Studie gefährden.

Im folgenden findet sich zunächst die Startseite des Online-Fragebogens, welche sich für alle drei Versuchsgruppen gleich gestaltete. Danach folgt die Darstellung der Instruktionstexte für VG 1, 2 und 3.

Startseite für VG 1, 2 und 3

Wissen Sie, welchen Eindruck Sie bei Persönlichkeitsverfahren in Bewerbungssituationen hinterlassen?

Nein? Dann nehmen Sie an dieser Studie teil!

Denn Sie erhalten eine **detaillierte Rückmeldung ihrer Ergebnisse**, wie Sie in einer realen Bewerbungssituation dem potenziellen Arbeitgeber vorliegen würde. Diese Rückmeldung liefert wertvolle Informationen über die individuellen Ausprägungen grundlegender Persönlichkeitseigenschaften.

Der Fragebogen dient somit zur Vorbereitung und Übung für künftige Assessment-Center!

Es lohnt sich also **20-30 Minuten** ihrer Zeit in die Bearbeitung des Fragebogens zu investieren!

Die Studie wird im Rahmen einer wissenschaftlichen Arbeit (Diplomarbeit) der Fakultät für Psychologie an der Universität Wien durchgeführt und behandelt das Thema **Persönlichkeitsverfahren im [Assessment-Center](#)**.

Für Rückfragen können Sie sich jederzeit an Ihre **Ansprechperson Frau Victoria Albertz** wenden (a0542736@unet.univie.ac.at).

Selbstverständlich werden Ihre Angaben vertraulich behandelt, so dass Ihre Anonymität zu jedem Zeitpunkt gewährleistet ist.

Tabelle 3: Startseite Online-Fragebogen

Obige Darstellung enthält jene Information, welche die Testpersonen der Studie (unabhängig von ihrer Gruppenzugehörigkeit zu VG 1, 2 oder 3) nach Betätigung des Links zum Fragebogen erhielten.

1. Instruktion (VG 1)

Stellen Sie sich bitte vor, Sie haben sich für eine Stelle beworben! Das Unternehmen hat Sie nun zu einem Kennenlern-Gespräch eingeladen und möchte auch eine Auswahltestung (Assessment-Center) mit Ihnen machen. Bei Ankunft im Unternehmen müssen Sie feststellen, dass Sie nicht der einzige Bewerber für die Stelle sind. Der Stellenanzeige konnten Sie bereits einige Informationen zur Stelle entnehmen. Diese machten einen sehr positiven Eindruck auf Sie. Während des Kennenlern-Gesprächs hat sich dieser positive erste Eindruck der Stelle vollkommen bestätigt. **Sie sind überzeugt, dass die Stelle genau ihren Vorstellungen entspricht und nehmen sich vor, ihr Bestes zu geben um den Job zu bekommen. Sie sind daher bemüht, im Assessment-Center Angaben zu machen, die ihre Chancen auf die Stelle erhöhen.**

Nach dem das Kennenlern-Gespräch beendet ist, werden Sie für das Assessment-Center gemeinsam mit einigen Mitbewerbern in einen Raum geführt. Dort wird Ihnen ein Platz am Computer zugewiesen, mit der Bitte einen Fragebogen auszufüllen. Auf dem Bildschirm erscheint die folgende Instruktion:

2. Instruktion (VG 2)

Stellen Sie sich bitte vor, Sie haben sich für eine Stelle beworben! Das Unternehmen hat Sie nun zu einem Kennenlern-Gespräch eingeladen und möchte auch eine Auswahltestung (Assessment-Center) mit Ihnen machen. Bei Ankunft im Unternehmen müssen Sie feststellen, dass Sie nicht der einzige Bewerber für die Stelle sind. **Während des Kennenlern-Gesprächs wird Ihnen klar, dass die Stelle nicht hundert-prozentig Ihren Vorstellungen entspricht. Sie finden die Stelle zwar nicht uninteressant, sind sich aber nicht sicher, ob sie den Job überhaupt haben möchten.**

Nach dem das Kennenlern-Gespräch beendet ist, werden Sie für das Assessment-Center gemeinsam mit einigen Mitbewerbern in einen Raum geführt. Dort wird Ihnen ein Platz am Computer zugewiesen, mit der Bitte einen Fragebogen auszufüllen. Auf dem Bildschirm erscheint die folgende Instruktion:

3. Instruktion (Kontrollbedingung, VG 3)

Stellen Sie sich vor, Sie wollen nach erfolgreichem Studienabschluss eine professionelle Berufsberatung in Anspruch nehmen. Im Rahmen dieser Berufsberatung soll ein Persönlichkeitsprofil Aufschluss darüber geben, welche persönlichkeitsbezogenen Stärken (Soft Skills) Sie in Hinblick auf eine zukünftige Berufstätigkeit mitbringen. Dieses Persönlichkeitsprofil wird auf Basis eines von Ihnen ausgefüllten Persönlichkeitsfragebogens erstellt. **Ihnen ist klar, dass nur eine ehrliche Beantwortung des Fragebogens ein realistisches Bild Ihrer berufsbezogenen Persönlichkeit vermitteln kann. Nur dann sind die Ergebnisse, die Ihnen rückgemeldet werden für Sie hilfreich. Sie sind daher sehr bemüht, so ehrlich und realistisch wie möglich zu antworten.**

Nachdem Ihnen ein Platz am Computer zugewiesen worden ist, werden Sie nun gebeten, den Persönlichkeitsfragebogen auszufüllen. Auf dem Bildschirm erscheint die folgende Instruktion:

Tabelle 4: Verfälschungsinstruktionen

Im folgenden werden die Hauptkriterien der Instruktionstexte dargestellt:

während die Ausgangssituation der ersten beiden Instruktionstexte (VG 1 und 2) durch eine Bewerbungssituation gekennzeichnet war, stellte die dritte Instruktion (VG 3) eine Berufsberatungssituation dar. Die Suggestion einer Bewerbungssituation sollte dabei einen Rahmen schaffen, in welchem Verfälschungsverhalten wahrscheinlich scheint, im Gegensatz zur Berufsberatungssituation, welche als einzige Konsequenz das Wissen über die eigenen berufsbezogenen Stärken und Schwächen birgt und somit ein möglichst authentisches Antwortverhalten fördern soll bzw. als Kontrollbedingung dienen soll.

Der in den ersten beiden Instruktionen (VG 1 und 2) enthaltene Zusatz, „Bei Ankunft im Unternehmen müssen Sie feststellen, dass Sie nicht der einzige

Bewerber für die Stelle sind“, dient dazu, die Verfälschungsmotivation der Testpersonen zu erhöhen.

Die in der ersten und zweiten Instruktion nun folgenden Abschnitte unterscheiden sich insofern, als dass die erste Instruktion dem Ziel dient, stärkeres Verfälschungsverhalten hervorzurufen als dies der Fall für die zweite Instruktion ist, indem der Testperson suggeriert wird, er/sie sei von der Stelle überzeugt und versuche daher, im Rahmen des Assessment-Center „sein Bestes zu geben um den Job zu bekommen“. „Sie sind daher bemüht, im Assessment-Center Angaben zu machen, die ihre Chance auf die Stelle erhöhen“ entspricht dabei einer Täuschungsaufforderung.

Die zweite Instruktion hingegen dient ebenfalls dem Ziel Verfälschungsverhalten hervorzurufen, jedoch in abgeschwächter Form im Vergleich zu der ersten Bedingung: „Während des Kennenlern-Gesprächs wird Ihnen klar, dass die Stelle nicht hundertprozentig Ihren Vorstellungen entspricht. Sie finden die Stelle zwar nicht uninteressant, sind sich aber nicht sicher, ob sie den Job überhaupt haben möchten“. Diese Instruktion verzichtet im Gegensatz zur ersten Instruktion gänzlich auf eine Aufforderung an die Testperson, sich auf eine bestimmte Art und Weise zu verhalten oder das Antwortverhalten dementsprechend anzupassen.

Die dritte Instruktion enthält im Gegensatz zur zweiten Instruktion erneut eine Aufforderung zu einem bestimmten Antwortverhalten: „Ihnen ist klar, dass nur eine ehrliche Beantwortung des Fragebogens ein realistisches Bild Ihrer berufsbezogenen Persönlichkeit vermitteln kann. (...) Sie sind daher sehr bemüht, so ehrlich und realistisch wie möglich zu antworten“. Ziel ist dabei die Vermeidung von Verfälschungsverhalten seitens der Testpersonen.

Die Testpersonen wurden während des Verfahrens an die jeweilig variierende Situation erinnert um die Verfälschungs- bzw. Ehrlichkeitsmotivation aufrecht zu erhalten. Unter der ersten Instruktion lauteten die erscheinenden Hinweise beispielsweise folgendermaßen: „Vergessen Sie nicht, dass Sie in einer Bewerbungssituation sind! Sie geben ihr Bestes, um den Job zu bekommen!“. Innerhalb der zweiten Bedingung fiel dieser Hinweis entsprechend aus: „Vergessen Sie nicht, dass Sie in einer Bewerbungssituation sind!“. Die dritte Bedingung enthielt beispielsweise eine Ehrlichkeitsaufforderung: „Vergessen Sie nicht, dass Sie nur dann (...) profitieren können, wenn Sie die Fragen ehrlich

beantworten!“. Des Weiteren erhielten die Testpersonen jeweils drei Meldungen zum Fortschritt der Fragebogenbearbeitung (z.B. „Sie haben bereits mehr als die Hälfte der Items beantwortet!“). Der vollständige Fragebogen findet sich im Anhang (S. 136).

Um die Attraktivität der Studienteilnahme zu erhöhen bzw. einen Anreiz zum vollständigen Ausfüllen des Fragebogens zu schaffen, wurde den Studienteilnehmern eine individualisierte Rückmeldung ihrer im BIP erzielten Ergebnisse angekündigt und unmittelbar nach Abschluss des Fragebogens ausgegeben. Die Rückmeldung war insofern individualisiert, als dass die 14 Ergebnisse der Dimensionen aus dem Bochumer Inventar zur berufsbezogenen Persönlichkeit (BIP) auf eine Skala von 0 bis 10 Punkten umgerechnet und dargestellt wurden. Zusätzlich wurden zu jeder Dimension Beschreibungen ausgegeben, welche sich auf Personen mit hohen sowie niedrigen Ausprägungen bezogen. So hatte jede Testperson die Möglichkeit, sich selbst anhand der ausgegebenen Punktzahl diesen Beschreibungen zuzuordnen. Eine Muster-Rückmeldung findet sich im Anhang (S. 156).

b. Programmierung des Online-Fragebogens

Die Erstellung des Online-Fragebogens erfolgte über das Online-Fragebogen-Erstellungs-Tool Onlineumfragen.com GmbH (<https://www.onlineumfragen.com/>). Der Persönlichkeitsfragebogen wurde über 16 Online-Seiten präsentiert, wobei die Testpersonen mittels zu Beginn ausgegebenem Passwort die Möglichkeit erhielten, die Beantwortung des Fragebogens zu einem späteren Zeitpunkt fortzuführen. D.h. es waren selbstgewählte Pausen möglich.

Die Befragung wurde mittels Verwendung dreier unterschiedlicher Links zu jeweils einer der drei Bedingungen durchgeführt.

c. Pilottestung von Instruktionen und Onlinefragebogen

Um möglichen Fehlerquellen und Missverständnissen vorzubeugen, wurde vor Beginn der tatsächlichen Datenerhebung eine Pilottestung (Vorabtestung) durchgeführt. Im Rahmen dieser Pilottestung erhielten 15 Personen (5 Personen pro Bedingung) aus dem näheren Umfeld der Autorin, einen Link zu dem jeweiligen Fragebogen per E-Mail zugesandt. Nach 2 Wochen wurden diese

Personen um eine mündliche Rückmeldung bezüglich des Fragebogens gebeten. Dabei wurden die folgenden Fragen erörtert:

1. Was hat die Instruktionsangabe (Berufsberatung bzw. Bewerbungssituation) bei der Beantwortung des Fragebogen bewirkt? Warst du ehrlich, oder nicht?
2. Wie verständlich war diese Instruktion?
3. Welche Angaben hätten geholfen, um die Berufsberatungs- bzw. Bewerbungssituation besser nachzuempfinden?
4. Hattest du Vorstellungen darüber, wie die optimalen Antworten auf die Fragen aussehen würden?
5. Wie wäre der Umgang mit den Fragen gewesen, wenn du keine Berufserfahrung hättest? Hättest du Schwierigkeiten die Fragen zu beantworten?
6. Meinst du, dass Studenten der Wirtschafts- und Rechtswissenschaften den Begriff „Assessment-Center“ kennen? Was verstehen Sie darunter?
7. Gab es in der Abfolge der Fragen Auffälligkeiten? Hat sich für dich merklich der Fragenstil geändert?
8. Wie lange hast du für den Fragebogen gebraucht?
9. Hättest du, und wenn ja wo, den Fragebogen unter normalen Umständen abgebrochen?
10. Wie hat dir die Rückmeldung gefallen? Was würdest du anders/besser machen?

Die Rückmeldung von 11 Teilnehmern (6 Wirtschaftswissenschaftsstudierende, 2 Rechtswissenschaftsstudierende, 3 Psychologie-Studierende) konnte sodann gesichtet und die Fragebögen entsprechend modifiziert werden. Die Teilnehmer gaben bezüglich der dargestellten Fragen die folgenden Rückmeldungen:

ad 1. Die Teilnehmer der Pilottestung verhielten sich hinsichtlich der ersten sowie dritten Instruktion nach eigenen Angaben gemäß der Instruktion. Innerhalb der zweiten Bedingung gaben die Teilnehmer an, sich ebenfalls besser dargestellt zu haben, da Sie sich unabhängig von der persönlichen Bewertung einer Stelle immer so darstellen würden, dass die Chance auf die Stelle erhöht wird, um sich auf diesem Wege alle Optionen offen zu halten. Somit wurden für die Hauptstudie keinerlei Änderungen an den Instruktionen mehr vorgenommen.

ad 2. Es ergaben sich keinerlei Verständnisschwierigkeiten, so dass die Instruktionen diesbezüglich nicht verändert werden mussten.

ad 3/4. Diese Fragen führten zu unterschiedlichen Ergebnissen. Die Mehrheit der Befragten hatte eine Vorstellung darüber wie die idealen Antworten aussähen, andere jedoch gaben an dazu eine Stellenbeschreibung zu benötigen. Für die Hauptstudie wurde jedoch von der Vorgabe einer Stellenbeschreibung abgesehen, da diese insofern einen unkontrollierten Einflussfaktor darstellen würde, als dass die Teilnehmer eine beschriebene Position bzw. das Anforderungsprofil als unterschiedlich ansprechend empfinden würden und dies eventuell in verändertem Verfälschungsverhalten hätte resultieren können.

ad 5. Die Berufserfahrung war nach Ansicht der Teilnehmer irrelevant für die Beantwortung der Fragen. Somit wurden für die Hauptstudie auch solche Teilnehmer akzeptiert, die bereits über Berufserfahrung verfügten.

ad 6. Die Teilnehmer waren geschlossen der Meinung dass der Begriff „Assessment-Center“ der Zielstichprobe geläufig sei, so dass der Begriff im Rahmen der Hauptstudie angewendet werden konnte.

ad 7. Die Teilnehmer konnten keine Auffälligkeiten oder Veränderungen bemerken. Diese Tatsache führte zur Beibehaltung des Fragebogaufbaus (BIP, NEO-FFI und SES-17 nacheinander vorgegeben).

ad 8. Die Teilnehmer gaben Zeiten zwischen 20 und 40 Minuten an. Diese Angabe war insofern hilfreich, als dass den Teilnehmern der Hauptstudie bereits im Voraus die ungefähre Bearbeitungsdauer mitgeteilt werden konnte, in der Hoffnung so die Abbruchquote zu verringern.

ad 9. Einige Teilnehmer gaben an, dass Sie den Fragebogen abbrechen würden. Es sei denn, Sie könnten aus der Beantwortung bzw. der Rückmeldung persönlichen Nutzen, beispielsweise für einen laufenden Bewerbungsprozess, ziehen. Um die Abbruchquote möglichst gering zu halten wurden vor Durchführung der Hauptstudie zusätzlich motivierende Smileys am Seitenkopf einiger Seiten eingefügt. Ausserdem wurden an entsprechender Stelle die Hinweise „Sie haben bereits x von 287 Fragen beantwortet!“ , „Sie haben bereits mehr als die Hälfte der Fragen beantwortet“ und „Sie haben es fast geschafft! Dies ist die letzte Seite!“ hinzugefügt.

ad 10. Die Teilnehmer empfanden die Rückmeldung als informativ und brachten keinerlei Veränderungs- oder Verbesserungsvorschläge an, so dass diese in ihrer Form beibehalten werden konnte (s. Anhang S. 156).

7.2 Erhebungsinstrumente

Der den Testpersonen im Rahmen der vorliegenden Arbeit vorgegebene Fragebogen setzte sich aus Items von zwei Persönlichkeitsfragebögen und einer Soziale-Erwünschtheits-Skala zusammen: Bochumer Inventar zur berufsbezogenen Persönlichkeit (BIP, 210 Items, Hossiep & Paschen, 2003; erläutert in Abschnitt 7.2.1), NEO Fünf Faktoren Inventar (NEO-FFI, 60 Items, Borkenau und Ostendorf, 1993; erläutert in Abschnitt 7.2.2) und Soziale-Erwünschtheits-Skala-17 (SES-17; 17 Items; Stöber, 1999; erläutert in Abschnitt 7.2.3). Insgesamt bearbeiteten die Testpersonen somit einen Gesamtfragebogen mit 287 Items. Zusätzlich erhielten die Testpersonen nach Beantwortung des eigentlichen Fragebogens drei weitere Fragen, um so den subjektiven Einfluss der Verfälschungsinstruktionen auf die Testpersonen erheben zu können: (1.) wie leicht es Ihnen fiel, sich in die beschriebene Situation hineinzusetzen, (2.) ob Sie sich bemüht hätten, einen besonders guten Eindruck zu hinterlassen und (3.) ob Sie sich bemüht hätten, ehrlich zu sein. Die Reihenfolge der Fragebogenvorgabe gestaltete sich folgendermaßen: zunächst erschienen die 210 Items des BIP, danach die 60 Items des NEO-FFI gefolgt von den 17 Items der SES-17. Die Reihenfolge der BIP-, NEO-FFI- und Soziale-Erwünschtheitsskala -Items wurde dabei jeweils, wie in den Original-Fragebögen vorgesehen, beibehalten. Da die drei Verfahren ein jeweils unterschiedliches Antwortformat aufweisen, wurde das Antwortformat für die vorliegende Studie vereinheitlicht: alle 287 Items wurden mittels einer fünfstufigen Likert-Skala (trifft völlig zu – trifft eher zu – neutral – trifft eher nicht zu – trifft nicht zu) vorgegeben. In den folgenden Abschnitten finden sich nähere Beschreibungen der drei angewandten Erhebungsinstrumente.

7.2.1 Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP)

Das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP) dient der „standardisierten Erfassung des Selbstbildes eines Testkandidaten in Hinblick auf relevante persönlichkeitsorientierte Beschreibungsdimensionen aus dem Berufsleben“ (Manual BIP, S. 14, Hossiep & Paschen, 2003). Das Verfahren basiert nicht auf dem Fünf-Faktoren-Modell, beinhaltet jedoch Teilaspekte der fünf Faktoren (Hossiep et al., 2003). Die Erfassung insbesondere berufsbezogener Persönlichkeitsmerkmale erleichtert im Falle einer Bewerbungssituation die Feststellung der erforderlichen Persönlichkeitsmerkmale, da das BIP die beruflichen Tätigkeiten welche mit den entsprechenden Persönlichkeitsmerkmalen einhergehen genau beschreibt.

Der Anwendungsbereich bezieht sich somit auf Beratungsgespräche oder Platzierungsentscheidungen im berufsbezogenen Bereich, in welchen der BIP beispielsweise als „Basis für eine tiefergehende Exploration“ (Hossiep und Paschen, 2003, S.14) eingesetzt werden kann. Einen weiteren Anwendungsbereich stellt die Durchführung des BIP als Grundlage „zur Vorbereitung von Coaching- oder Trainingsmaßnahmen“ (Hossiep und Paschen, 2003, S.14) dar.

Da die Polung der Items (positiv oder negativ in Bezug auf das interessierende Merkmal) einen Einfluss auf die Verwendung von Antwortkategorien (z.B. Vermeidung des *Acquiescence Response Style*, ARS durch die Verwendung sowohl positiv als auch negativ formulierter Items) haben kann (Weijters & Baumgartner, 2012; Herche & Engelland, 1996) wurde diese im Rahmen der vorliegenden Studie mittels Umformulierung der betroffenen Items (94 von 210) vereinheitlicht. D.h. alle negativ gepolten Items wurden so formuliert, dass sie in Bezug auf das zu messende Merkmal eine positive Polung haben. Im folgenden findet sich ein Beispiel zur Umpolung der Items:

Original-Item (Dimension Belastbarkeit): „Wenn ich extrem hart arbeiten muss, gerate ich bei zusätzlichen Schwierigkeiten aus dem Gleichgewicht.“

Umformuliertes Item: „Wenn ich extrem hart arbeiten muss, gerate ich auch bei zusätzlichen Schwierigkeiten nicht aus dem Gleichgewicht.“

Das BIP enthält 210 Items welche auf vier Bereiche und 14, den Bereichen zugeordneten Dimensionen, aufgeteilt sind. Im folgenden sind die vier Bereiche mit den dazugehörigen Dimensionen nach Hossiep et al. (2003) dargestellt. Diese Darstellungen dienen gleichzeitig als Grundlage zur Interpretation der Dimensionen. Detailliertere Interpretationshinweise finden sich in der Muster-Rückmeldung im Anhang (S. 156).

Arbeitsverhalten (AV)

Gewissenhaftigkeit entspricht laut Manual teilweise der Dimension Gewissenhaftigkeit des Big-Five-Modells (s. Abschnitt 7.2.2) und misst im wesentlichen Sorgfalt und Präzision. Beispiel-Item: „Bevor ich mich Freizeitaktivitäten zuwende, erledige ich alle anstehenden Arbeiten.“

Flexibilität wird verstanden als „die Bereitschaft und Fähigkeit, sich auf veränderliche berufliche Bedingungen und wechselnde Situationen einzustellen“. Die Dimension weist laut Manual Überschneidungen zur Dimension Offenheit für Erfahrungen des Big-Five-Modells auf. Beispiel-Item: „Ich empfinde es als Herausforderung, wenn ich mit unvorhergesehenen Situationen konfrontiert werde.“

Handlungsorientierung, als zeitlich schnelles Aufnehmen von Aktivitäten und Umsetzung von Absichten. Beispiel-Item: „Meine Zeiteinteilung gelingt mir nicht so, dass ich meine Aufgaben rechtzeitig erledige.“

Berufliche Orientierung (BO)

Leistungsmotivation, als „Bereitschaft zur Auseinandersetzung mit einem hohen Gütemaßstab.“ Beispiel-Item: „Ich empfinde Genugtuung dabei, mit meinen Kräften bis an meine Grenzen zu gehen.“

Gestaltungsmotivation, als die Bereitschaft, „andere zu einem Tun zu veranlassen, das sie ohne diese Einflussnahme nicht unbedingt zeigen würden.“ Beispiel-Item: „Manche Kollegen denken, ich dränge zu stark auf Veränderungen, wenn ich mit ihnen zusammenarbeite.“

Führungsmotivation, als Motivation, direkten Einfluss auf soziale Vorgänge zu nehmen. Beispiel-Item: „Es stellt mich zufrieden, wenn ich andere beeinflussen kann.“

Soziale Kompetenz (SK)

Sensitivität, als „Fähigkeit zur Wahrnehmung auch schwacher Signale im zwischenmenschlichen Bereich“. Beispiel-Item: „Ich finde auch zu sehr schwierigen Personen einen guten Draht.“

Kontaktfähigkeit, als Fähigkeit zum Aufbau zwischenmenschlicher Beziehungen und der Nutzung beruflicher Kontakte. Diese Dimension weist laut Manual Überschneidungen mit der Dimension Extraversion des Big-Five-Modells auf. Beispiel-Item: „Ich nutze einen grossen Teil meiner Freizeit, um Kontakte zu pflegen.“

Soziabilität erfasst „eine im sozialen Umgang freundliche und rücksichtsvolle Grundhaltung“ und weist laut Manual Ähnlichkeiten mit der Skala Verträglichkeit des Big-Five-Modells auf. Beispiel-Item: „In Gruppendiskussionen bin meist ich es, der ausgleichend wirkt.“

Teamorientierung erfasst das „Interesse und die Bereitschaft, die eigene Position zugunsten der Zusammenarbeit in einem Team zurückzunehmen“. Beispiel-Item: „Wenn ich etwas plane, überlege ich zunächst, wer noch bei dem Projekt mitarbeiten könnte.“

Durchsetzungsstärke erfasst, inwiefern eine Person bereit ist „eigene Auffassungen auch dann weiter zu verfolgen, wenn sich Schwierigkeiten ergeben, die durch andere Personen verursacht sind“. Beispiel-Item: „Bei der Arbeit im Team setze ich mich nachhaltig dafür ein, andere von meinem Standpunkt zu überzeugen.“

Psychische Konstitution (PK)

Emotionale Stabilität misst die Veranlagung zu einer emotional ausgeglichenen Grundhaltung, welche die angemessene Kontrolle der eigenen emotionalen Reaktion erlaubt. Diese Skala weist laut Manual deutliche Parallelen mit der Skala Neurotizismus aus dem Big-Five-Modell auf. Beispiel-Item: „Ich nehme des Leben im Allgemeinen leicht.“

Belastbarkeit erhebt das Ausmaß, zu dem sich eine Person als robust und widerstandsfähig bezeichnet. Beispiel-Item: „Ich kann problemlos viele Stunden ohne Pause arbeiten.“

Selbstbewusstsein wird verstanden als Bereitschaft, eigene Vorstellungen und Verhaltensweisen zu vertreten, auch wenn diese womöglich auf Widerspruch und

Missgunst stoßen könnten. Beispiel-Item: „Ich besitze Eigenschaften, in denen ich den meisten Menschen überlegen bin.“

Ein wichtiger Maßstab um die Qualität eines Verfahrens zu beurteilen sind die Gütekriterien, welche für das BIP im folgenden dargestellt sind:

Interne Konsistenz

Die Mehrheit, nämlich 12 der Dimensionen aus dem BIP, weisen gute bis sehr gute (> 0.8) Interne-Konsistenz-Maße auf, während zwei der angeführten Dimensionen (Gestaltungsmotivation und Soziabilität) lediglich zufriedenstellende Werte (>0.7) aufweisen. In der rechten Spalte (Tabelle 5) finden sich die auf Basis der vorliegenden Studie ermittelten Cronbach α Werte (Reliabilitätskoeffizient). Die Ergebnisse sprechen insgesamt für die inhaltliche Homogenität der Dimensionen und eine gute bis zufriedenstellende Messgenauigkeit (Hossiep et al., 2003).

Tabelle 5: Interne Konsistenz BIP

Skala	BIP- Normstichprobe (N=9139)	Gesamtstichprobe, aktuelle Studie (N=385)
	Cronbach α	Cronbach α
<i>Leistungsmotivation</i>	0,81	0,84
<i>Gestaltungsmotivation</i>	0,74	0,76
<i>Führungsmotivation</i>	0,87	0,89
<i>Gewissenhaftigkeit</i>	0,83	0,89
<i>Flexibilität</i>	0,87	0,85
<i>Handlungsorientierung</i>	0,85	0,85
<i>Sensitivität</i>	0,85	0,84
<i>Kontaktfähigkeit</i>	0,90	0,90
<i>Soziabilität</i>	0,76	0,80
<i>Teamorientierung</i>	0,89	0,90
<i>Durchsetzungsstärke</i>	0,85	0,89
<i>Emotionale Stabilität</i>	0,89	0,87
<i>Belastbarkeit</i>	0,91	0,89
<i>Selbstbewusstsein</i>	0,85	0,88

Hossiep et al., 2003

vorliegende Studie

Retest – und Split-Half-Reliabilität

Während die Split-Half Reliabilitäten sich in einem guten Bereich bewegen (0,72-0,91), weisen die Retest-Reliabilitäten teilweise geringere Werte auf (0,69-0,80). Dobrietz (2001; zit. nach Hossiep et al., 2003) teilte zur Erfassung der Stabilität der Reliabilitätswerte eine Stichprobe in drei unterschiedlich lang andauernde Zeitcluster und konnte darauf basierend feststellen, dass die Dimensionen als langfristig stabil anzusehen sind.

Kriteriumsvalidität – Berufliches Entgelt und die hierarchische Position

Hinsichtlich der Bedeutung der Gesamtskala für das berufliche Entgelt ergibt sich ein $R^2=0,15$. Somit sind 15% der Unterschiede im Entgelt auf die unterschiedlichen Ergebnisse der Testpersonen in den einzelnen Dimensionen zurückzuführen (Hossiep et al., 2003). Hinsichtlich der hierarchischen Position liegt ein R^2 von 0,16 vor (Hossiep et al., 2003). Die Zusammenhänge zwischen den Dimensionen des BIP und den Aussenkriterien berufliches Entgelt (0,10-0,31 für 12 Dimensionen; 0,08 und 0,09 für Sensitivität und Soziabilität) sowie hierarchische Position (0,10-0,37 für 11 Dimensionen; <0,10 für Gewissenhaftigkeit, Sensitivität und Teamorientierung) können als zufriedenstellend gewertet werden (Hossiep et al., 2003).

Ein kritischer Aspekt des BIP besteht jedoch in den teilweise hoch ausgeprägten Skaleninterkorrelationen (Hossiep et al., 2003), welche auf die Tatsache zurückzuführen sind, dass das zugrundeliegende Persönlichkeitsmodell mehr als die fünf mittels Big-Five-Modell extrahierten Faktoren annimmt. Hossiep et al. (2003) betonen zwar, dass diese Interkorrelationen inhaltlich plausibel sind da sie jene Zusammenhänge repräsentieren, welche die Skalen im Alltagsverständnis auch tatsächlich aufweisen, jedoch stellen sie sich aus testtheoretischer Sicht problematisch dar. Eine vollständige Auflistung der Interkorrelationen findet sich im Anhang (S. 135).

7.2.2 NEO Fünf Faktoren Inventar (NEO-FFI)

Im folgenden findet sich eine Beschreibung des NEO Fünf Faktoren Inventar nach Costa und McCrae (NEO-FFI) (Borkenau & Ostendorf, 1993).

Das NEO-FFI basiert auf den fünf Faktoren des Fünf-Faktoren-Modells nach Costa & McCrae und erhebt diese fünf Faktoren anhand von 60 Items. Der Einsatz des Verfahrens empfiehlt sich nach Borkenau und Ostendorf (1993) insbesondere zur „groben aber vollständigen Erfassung der Bereiche individueller Unterschiede“ (Manual NEO-FFI, S. 8). Somit eignet sich das Verfahren beispielsweise für den Einsatz im Rahmen von Forschungsprojekten oder als Breitbandverfahren (*Screening*-Verfahren). Wie bereits im Falle des BIP, wurden auch für das NEO-FFI negativ gepolte Items für die vorliegende Studie so umformuliert, dass eine starke Zustimmung (Antwortkategorie „trifft völlig zu“) tatsächlich einer Bestätigung der abgefragten Dimension entspricht. Im folgenden findet sich ein entsprechendes Beispiel:

Original-Item (Dimension Neurotizismus): „Ich bin leicht beunruhigt.“

Umformuliertes Item: „Ich bin nicht leicht beunruhigt.“

Es wurden insgesamt 27 der 60 Items des NEO-FFI umformuliert.

Die fünf Faktoren des NEO-FFI sind im folgenden dargestellt.

Neurotizismus

Die Skala Neurotizismus erfasst die Ausprägung der emotionalen Stabilität oder Labilität einer Person (Borkenau & Ostendorf, 1993). Personen mit einer hohen Ausprägung in dieser Dimension berichten, dass sie leicht aus dem Gleichgewicht zu bringen seien, im Gegensatz zu Personen mit niedriger Ausprägung in dieser Dimension, welche sich selbst als ruhig, ausgeglichen und sorgenfrei beschreiben. Beispiel-Item: „Wenn ich unter starkem Stress stehe, fühle ich mich manchmal, als ob ich zusammenbräche.“

Extraversion

Personen mit hohen Werten in dieser Dimension beschreiben sich selbst als gesellig, aktiv, gesprächig und optimistisch. Beispiel-Item: „Ich habe gerne viele Leute um mich herum.“

Offenheit für Erfahrungen

Personen, welche hohe Ausprägungen in dieser Dimension vorweisen, beschreiben sich selbst als wissbegierig, kreativ und unabhängig in ihrem Urteilsvermögen. Beispiel-Item: „Ich habe oft Spass daran, mit Theorien oder abstrakten Ideen zu spielen.“

Verträglichkeit

Personen mit hohen Werten in dieser Dimension beschreiben sich selbst als altruistisch, mitfühlend und wohlwollend. Sie zeigen sich oftmals kooperativ, nachgiebig und harmoniebedürftig. Beispiel-Item: „Ich würde lieber mit anderen zusammenarbeiten, als mit Ihnen zu wetteifern.“

Gewissenhaftigkeit

Personen mit hoher Ausprägung in dieser Dimension bezeichnen sich selbst als ordentlich, zuverlässig, penibel, ehrgeizig und systematisch im Gegensatz zu Personen mit einer niedrigen Ausprägung in dieser Dimension, welche sich selbst eher als nachlässig und gleichgültig beschreiben. Beispiel-Item: „Ich versuche alle mir übertragenen Aufgaben sehr gewissenhaft zu erledigen.“

Die Normierungsstichprobe des NEO-FFI besteht aus 2112 Testpersonen (Borkenau & Ostendorf, 1993). Im folgenden werden die wichtigsten Gütekriterien zum NEO-FFI dargestellt und erläutert:

Interne Konsistenz

Während die Skalen Gewissenhaftigkeit, Neurotizismus sowie Extraversion auf eine gute interne Konsistenz der Items hinweisen (0,80-0,85), liegt das Cronbach's α (Reliabilitätskoeffizient) mit 0,71 für die Dimensionen Offenheit für Erfahrungen und Verträglichkeit im mittleren, jedoch akzeptablen Bereich (Tabelle 6). Diese Ergebnisse konnten von Kanning & Holling (2001) sowie Lüdtker et al. (2004) (s. mittlere Spalte, Tabelle 6) bestätigt werden. Ergänzend sind in der rechten Spalte (Tabelle 6) entsprechende Cronbach α Werte auf Basis der vorliegenden Studie enthalten. Die Ergebnisse ($>0,70$) lassen auf inhaltlich homogene und genau messende Dimensionen schließen.

Tabelle 6: Interne Konsistenz NEO-FFI

Skala	Normstichprobe des NEO-FFI (N=2112)	Gesamtstichprobe, vorliegende Studie (N=385)	
Cronbach α			
Neurotizismus	0,85	0,81	0,88
Extraversion	0,80	0,77	0,79
Offenheit für Erfahrungen	0,71	0,73	0,80
Verträglichkeit	0,71	0,73	0,78
Gewissenhaftigkeit	0,85	0,83	0,87
	Borkenau & Ostendorf, 1993	Lüdtke et al., 2004	vorliegende Studie

Retest-Reliabilität

Hinsichtlich der Retest-Reliabilitäten weisen die Dimensionen Neurotizismus, Extraversion und Gewissenhaftigkeit starke Zusammenhänge auf (0,80-0,81), während die Dimensionen Offenheit für Erfahrungen sowie Gewissenhaftigkeit einen mittleren Zusammenhang (0,76 bzw. 0,65) abbilden. Somit weisen die Skalen eine zufriedenstellende bis gute Zuverlässigkeit auf.

Konstruktvalidität

Die Konstruktvalidität des NEO-FFI wurde von Borkenau und Ostendorf (1993) mittels der Befragung dreier Bekannter einer jeden Testperson erhoben. Borkenau und Ostendorf (1993) erklären die teils geringen Korrelationswerte (0,23-0,45) durch die Tatsache, dass die beurteilenden Bekannten aus organisatorischen Gründen nicht immer nahe Bekannte waren und verweisen auf die von Ostendorf (1990) (zitiert nach Borkenau & Ostendorf, 1993) durchgeführte Studie, bei der mittels gleicher Methode jedoch unter Teilnahme ausschliesslich naher Bekannter erheblich zufriedenstellendere Korrelationen (0,43-0,61) festgestellt werden konnten.

Insgesamt erweist sich das NEO-FFI als teststatistisch hochwertiges Instrument (Kanning & Holling, 2001), welches durch die Überprüfung und darauf basierende Bestätigung der Faktorenstruktur, sowie mindestens zufriedenstellende

Reliabilitäten bestätigt werden kann (Lüdtke et al., 2004; Kanning & Holling, 2001).

7.2.3 Soziale-Erwünschtheits-Skala-17 (SES-17)

Die in der vorliegenden Arbeit vorgegebene Soziale-Erwünschtheits-Skala-17 (SES-17, Stöber 1999) entspricht denen in Abschnitt 3.1.2 (S.18) erwähnten Validitäts- und sozialen Erwünschtheits-Skalen. Derartigen Verfahren liegt die Annahme stereotypen Verfälschungsverhaltens (wie dargelegt in Abschnitt 3.1.1) zugrunde. Auf Basis dieser Annahme wird vermutet, dass verfälschende Personen mittels Vorgabe dieses Verfahrens identifiziert werden können. Im vorliegenden Fall dient der Einsatz der SES-17-Skala eben diesem Zweck, nämlich der Identifikation variierenden Verfälschungsverhaltens zwischen den drei Versuchsgruppen.

Die SES-17 ist eine auf Basis der Marlowe-Crowne-Skala¹⁴ entwickelte Skala und umfasst 17 Items.

Im folgenden finden sich drei Beispiel-Items:

„Manchmal werfe ich Müll einfach in die Landschaft oder auf die Strasse.“

„Ich habe schon einmal jemanden ausgenutzt oder über Ohr gehauen.“

„Ich lästere gelegentlich über andere hinter deren Rücken.“

Hinsichtlich der **Gütekriterien** gelten folgende Angaben nach Stöber (1999):

<u>Interne Konsistenz</u>	Cronbach α =0,72-0,75
<u>Konvergente Validität</u> (mit SDS-CM ¹⁵ Skala)	r =0,67-0,74
<u>Retest-Reliabilität</u>	r_{xx} = 0,82

Die interne Konsistenz der Items der SES-17 Skala in der vorliegenden Studie liegt bei einem Cronbach α von 0,77 (N=385).

Somit kann bei der SES-17 von einer ökonomischen Sozialen-Erwünschtheits-Skala mit mindestens zufriedenstellenden Gütekriterien ausgegangen werden.

¹⁴Social Desirability-Scale (Crowne & Marlow, 1960) nach Stöber (1999)

7.3 Durchführung der Untersuchung

Die Datenerhebung erstreckte sich über einen knapp dreimonatigen Zeitraum vom 18.09.2012 bis zum 11.12.2012. Die Akquirierung der teilnehmenden Studenten fand über verschiedenste Wege statt: als informelle Akquirierungswege wurden das soziale Netzwerk Facebook (privates FB-Netzwerk, Facebook-Seiten deutscher und österreichischer Rechts- und Wirtschaftswissenschaftlicher Fakultäten) sowie verschiedene universitäre rechts- und wirtschaftswissenschaftliche Foren genutzt. Des Weiteren wurden insgesamt vier, an Universitäten als Lehrpersonal tätige Personen aus dem privaten Umfeld der Autorin der vorliegenden Arbeit gebeten, die Studie im Rahmen ihrer Lehrveranstaltung kurz vorzustellen und den LV-Teilnehmern den Web-Link zum Fragebogen zukommen zu lassen. Darüberhinaus erhielten 68 rechts- bzw. wirtschaftswissenschaftliche Dekanate sowie 71 Career Center deutscher und österreichischer Universitäten entsprechende Teilnahme-Anfragen.

Die Reihenfolge der Item-Vorgabe war (nach Erhalt der variierenden Instruktionen, s. Abschnitt 7.1) unter allen drei Bedingungen gleichbleibend (BIP, NEO-FFI, SES-17). Die durchschnittliche Bearbeitungsdauer der Testpersonen der drei Versuchsgruppen lag zwischen 37 und 38 Minuten.

7.4 Stichprobe

Die Gesamtstichprobe (N=385) setzt sich aus den drei Versuchsgruppen (VG) zusammen, welche den Versuchsbedingungen „hohe Verfälschungs-Motivation“ (VG1, n=127), „unsichere Verfälschungs-Motivation“ (VG 2, n=126) und „geringe Verfälschungs-Motivation“ bzw. „ehrliche Beantwortung“ (VG3, n=132) und den entsprechenden Verfälschungsinstruktionen (s. Abschnitt 7.1) zugeordnet sind.

Die folgende Tabelle dient zur Darstellung der Altersverteilung innerhalb der Versuchsgruppen.

Tabelle 7: Altersverteilung

VG 1 <i>hohe Verfälschungs- Motivation</i>			VG 2 <i>unsichere Verfälschungs- Motivation</i>			VG 3 <i>geringe Verfälschungs- Motivation</i>		
Alter	n	%	Alter	n	%	Alter	n	%
18-21	19	15,0	18-21	23	18,2	18-21	36	27,2
22-25	59	46,4	22-25	65	51,6	22-25	66	50,0
26-30	39	30,7	26-30	28	22,2	26-30	24	18,1
>30	10	7,90	>30	10	8,00	>30	6	4,70
Σ	127	100,0	Σ	126	100	Σ	132	100
$\bar{x}=25,08$	$\sigma = 3,77$		$\bar{x}= 25,05$	$\sigma = 5,38$		$\bar{x}= 24,01$	$\sigma = 4,37$	

Die Altersverteilung innerhalb der Versuchsgruppen stellt sich wie aus Tabelle 7 ersichtlich, für die drei Versuchsgruppen ähnlich dar. So wird in etwa die Hälfte der jeweiligen Versuchsgruppe durch Testpersonen zwischen 22 und 25 Jahren gestellt. Die Gruppe der Testpersonen zwischen 18 und 21 Jahren macht in VG 3 einen grösseren Anteil aus als in VG 1 und VG 2. Diese Differenz schlägt sich in der Folge in einem niedrigeren Anteil der 26 bis 30-jährigen Testpersonen innerhalb der VG 3 nieder.

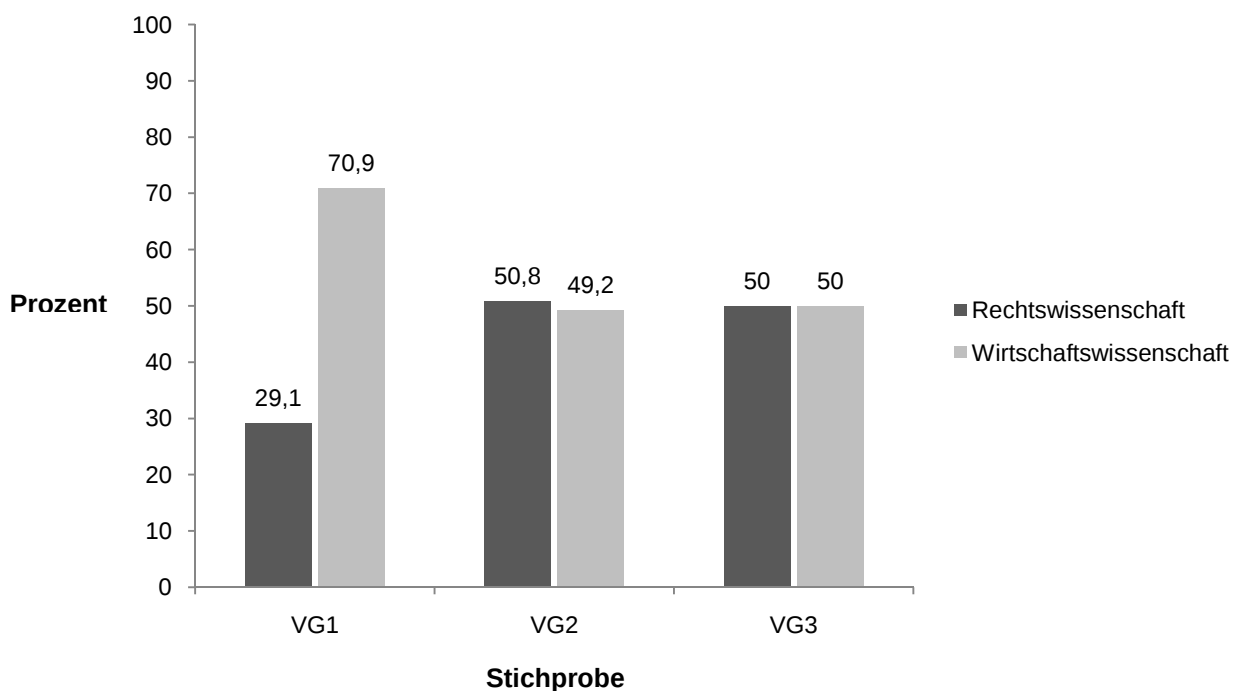


Abbildung 3: Verteilung der Studienfächer

Hinsichtlich der Verteilung der belegten Studienfächer der Testpersonen gibt es ebenfalls eine Differenz zwischen den Versuchsgruppen, wie in Abbildung 3 ersichtlich. Während die VG 2 und VG 3 mit jeweils ca. 50% durch Wirtschafts- und Rechtswissenschaftler dargestellt werden, weist VG 1 diesbezüglich eine ungleiche Verteilung auf. So wurden 70,9% der Testpersonen durch Wirtschaftswissenschaftler, jedoch nur 30% der Teilnehmenden durch Studenten der Rechtswissenschaft repräsentiert.

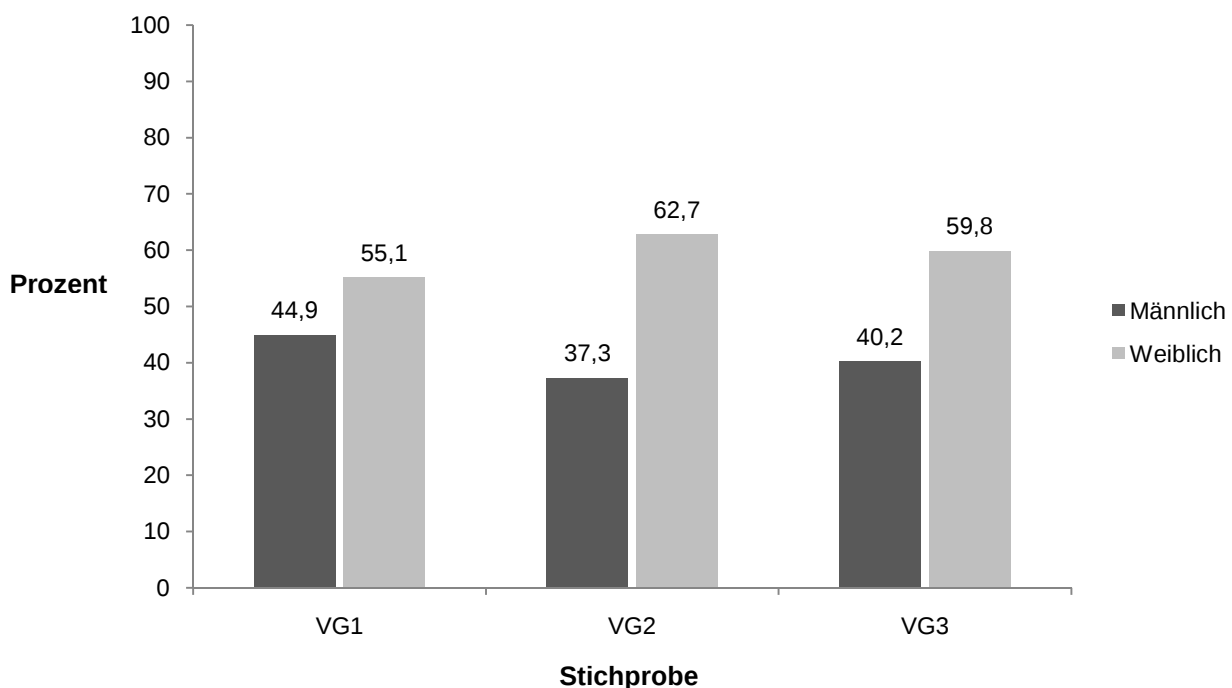


Abbildung 4: Geschlechterverteilung

Hinsichtlich der Geschlechterverteilung zeigt sich, wie aus Abbildung 4 ersichtlich, über die drei Versuchsgruppen ein ähnliches Bild geprägt von einem erhöhten Anteil weiblicher Testpersonen, der bei ca. 60% liegt.

Tabelle 8: Verteilung der Semesteranzahl

VG 1 <i>hohe Verfälschungs- Motivation</i>			VG 2 <i>unsichere Verfälschungs- Motivation</i>			VG 3 <i>geringe Verfälschungs- Motivation</i>		
Anzahl Semester	<i>n</i>	<i>%</i>	Anzahl Semester	<i>n</i>	<i>%</i>	Anzahl Semester	<i>n</i>	<i>%</i>
1-3	13	10,2	1-3	11	8,7	1-3	26	19,7
4-6	34	26,8	4-6	29	23,0	4-6	29	22,0
6-10	47	37,0	6-10	62	49,2	6-10	56	42,4
10-14	25	19,7	10-14	20	15,9	10-14	17	12,9
14-18	2	1,6	14-18	2	1,6	14-18	1	0,8
>18	1	0,8	>18	0	0,0	>18	1	0,8
Σ	122	96,1	Σ	124	98,4	Σ	130	98,5

Hinsichtlich der Verteilung der Semesteranzahl der teilnehmenden Studierenden lassen sich bezüglich der in Tabelle 8 gemachten Angaben folgende Aussagen tätigen: grundsätzlich stellt sich die Verteilung zwischen den drei Versuchsgruppen als sehr ähnlich dar. Der grösste Anteil an Testpersonen pro VG hatte zum Zeitpunkt der Studienteilnahme zwischen 6 und 10 Fachsemestern absolviert (37%, 49,2% und 42,4% entsprechend VG 1, 2 und 3). Gefolgt von solchen Testpersonen, die zum Teilnahmezeitpunkt zwischen 4 und 6 (26,8%, 23% und 22%) und zwischen 1 und 3 Fachsemestern absolviert hatten (10,2%, 8,7% und 19,7%).

Tabelle 9: Verteilung des höchsten Ausbildungsabschlusses

VG 1 <i>hohe Verfälschungs- Motivation</i>			VG 2 <i>unsichere Verfälschungs- Motivation</i>			VG 3 <i>geringe Verfälschungs- Motivation</i>		
Ausbildung	n	%	Ausbildung	n	%	Ausbildung	n	%
Lehre/ Pflichtschule	2	1,6	Lehre/ Pflichtschule	1	0,8	Lehre/ Pflichtschule	0	0,0
Realschule	1	0,8	Realschule	1	0,8	Realschule	0	0,0
Matura/ Abitur	73	57,5	Matura/ Abitur	70	55,6	Matura/ Abitur	78	59,1
Universität/ FH	47	37,0	Universität/ FH	50	39,7	Universität/ FH	48	36,4
Sonstige	1	0,8	Sonstige	1	0,8	Sonstige	4	3,0
Σ	124	97,6	Σ	123	97,6	Σ	130	98,5

Die Verteilung der höchsten Ausbildungsabschlüsse, dargestellt in Tabelle 9, gestaltet sich über die drei Versuchsgruppen hinweg ebenfalls sehr ähnlich. So verfügen zwischen 55,6% und 59,1% der Testpersonen über Matura bzw. Abitur als höchste abgeschlossene Ausbildung. Zwischen 36,4% und 39,7% der Testpersonen verfügte zum Zeitpunkt der Studienteilnahme bereits über einen Universitäts- oder FH-Abschluss. Auch hier stellt sich die Verteilung der Studierenden über die drei VG als sehr ähnlich und somit günstig dar.

Tabelle 10: Verteilung der Berufserfahrung

VG 1 <i>hohe Verfälschungs- Motivation</i>			VG 2 <i>unsichere Verfälschungs- Motivation</i>			VG 3 <i>geringe Verfälschungs- Motivation</i>		
Berufs- erfahrung	n	%	Berufs- erfahrung	n	%	Berufs- erfahrung	n	%
Keine	13	10,2	keine	7	5,6	keine	11	8,3
Nebenjob	21	16,5	Nebenjob	32	25,4	Nebenjob	38	28,8
Praktikum	25	19,7	Praktikum	31	24,6	Praktikum	28	21,2
Teilzeit	20	15,7	Teilzeit	21	16,7	Teilzeit	27	20,5
Vollzeit	43	33,9	Vollzeit	34	27,0	Vollzeit	27	20,5
Σ	124	96,0	Σ	125	99,3	Σ	131	99,3

Die Verteilung der Berufserfahrung zwischen den VG ist ein Aspekt, welcher Unterschiede zwischen den Versuchsgruppen aufweist. Tabelle 10 stellt diese Unterschiede dar, insofern als dass in VG 1 der Anteil solcher Testpersonen mit Berufserfahrung in Form eines Nebenjobs zugunsten der Testpersonen ohne jegliche Berufserfahrung im Gegensatz zu VG 2 und VG 3 zwischen 5% und 8% niedriger ausfällt. Weiters auffallend ist, dass der Anteil solcher Testpersonen, welche bereits in einem Vollzeit-Arbeitsverhältnis arbeiten oder gearbeitet haben um ca. 7% bis 14% grösser ausfällt, als dies in VG 2 und VG 3 der Fall ist. Diese Unterschiede könnten in der Folge insofern einen Einfluss auf die Ergebnisse haben, als dass eine Vollzeit-Anstellung im Regelfall mit einem fundierteren Bewerbungsprozess einhergeht, als dies für beispielsweise Nebentätigkeiten der Fall ist.

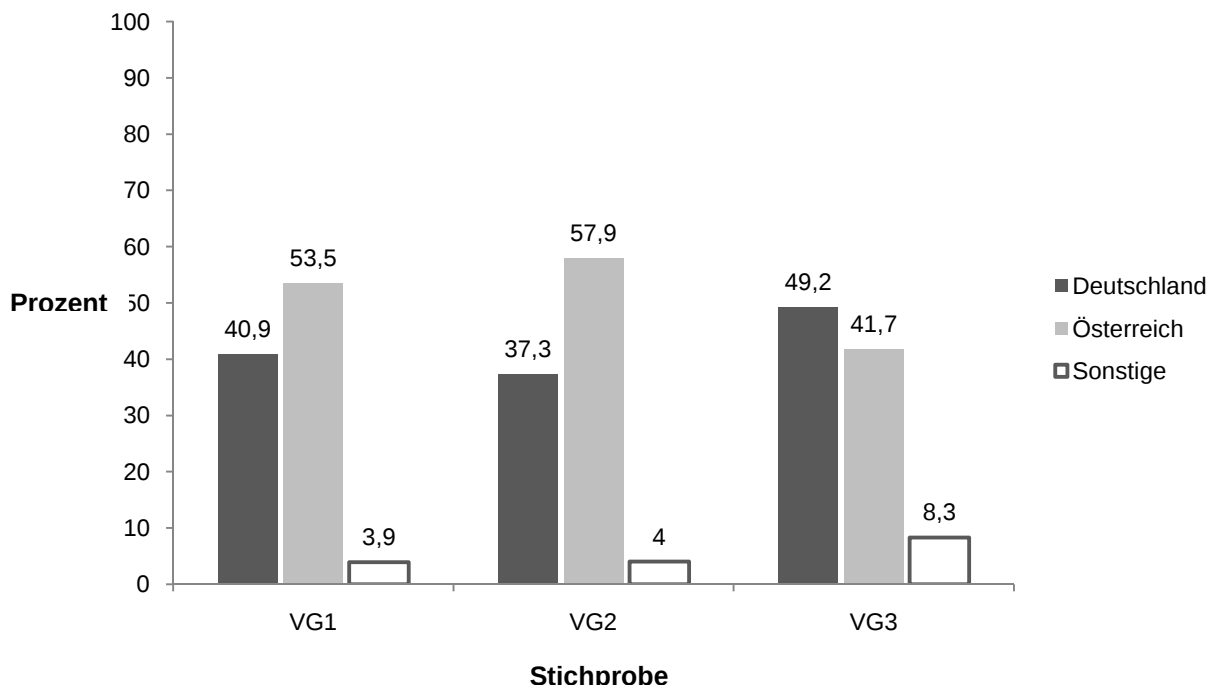


Abbildung 5: Verteilung der Nationalitäten

Obige Abbildung 5 stellt die Verteilung der Nationalitäten innerhalb der Versuchsgruppen dar. Auch diesbezüglich stellen sich die Nationalitäten der Testpersonen zwischen den Versuchsgruppen vorteilhafterweise als sehr ähnlich dar. Während VG 1 und 2 über einen grösseren Anteil an österreichischen Testpersonen verfügen, ist der Anteil deutscher Testpersonen in VG 3 grösser als der Anteil österreichischer Testpersonen. Der Anteil an Testpersonen anderer Nationalitäten ist verhältnismässig gering, auch wenn er in VG 3 etwas höher ausfällt, als dies für die VG 1 und VG 2 der Fall ist.

Hinsichtlich der Teilnahme-Quoten können folgende Angaben gemacht werden: 604 Personen (VG 1), 590 Personen (VG 2) und 498 Personen (VG 3) betätigten den Link zum jeweiligen Fragebogen. Davon beantworteten 226 Personen, 233 Personen bzw. 227 Personen mindestens eine Frage. Der Rücklauf vollständig beantworteter Fragebögen betrug 140 (VG 1), 144 (VG 2) und 155 (VG 3) Fragebögen, wovon einige Fälle (14 Personen der VG 1, 13 Personen aus VG 2 und 13 Personen aus VG 3) aus der Analyse ausgeschlossen werden mussten, da Sie weder einem rechts- noch wirtschaftswissenschaftlichen Studiengang angehörten und die Stichprobe möglichst homogen gehalten werden sollte.

Hinsichtlich des Abbruchpunktes ist zu sagen, dass die Mehrheit der abgebrochenen Befragungen während der Bearbeitung der ersten drei Fragen-Seiten (zwischen Beantwortung des Items 1 und 75 von 287 Items) abgebrochen wurden.

8 Ergebnisse zum Effekt der Verfälschungsinstruktionen

8.1 Ergebnisse zum subjektiven Effekt der Verfälschungsinstruktionen

Um eruieren zu können wie die Testpersonen der drei Versuchsgruppen die Effektivität der Instruktionen und ihre Reaktion auf die Verfälschungsinstruktionen einschätzen, wurden den Testpersonen nach Beendigung des Fragebogens zusätzlich drei Fragen zum Effekt der Verfälschungsinstruktion gestellt. Dabei sollten sie (1.) einschätzen, wie leicht es Ihnen fiel, sich in die beschriebene Situation hineinzusetzen, (2.) ob Sie sich bemüht hätten, einen besonders guten Eindruck zu hinterlassen und (3.) ob Sie sich bemüht hätten, ehrlich zu sein. Für die statistischen Analysen wurde das Programmpaket *SPSS für Windows* (Version 20) herangezogen.

8.1.1 Durchführung der Analyse und Darstellung der Ergebnisse

Zunächst wurden die Mittelwerte und Standardabweichungen der drei Zusatzfragen betrachtet (siehe Tabelle 11).

Tabelle 11: Mittelwert und Standardabweichung zum subjektiven Effekt der Verfälschungsinstruktionen

AV		Versuchsgruppe 1	Versuchsgruppe 2	Versuchsgruppe 3
<i>Leichtes Hineinversetzen</i>	Mittelwert	4,10	4,05	4,14
	Std.Abweichung	0,82	0,88	0,80
<i>Guter Eindruck</i>	Mittelwert	1,81	1,85	1,39
	Std.Abweichung	0,93	0,93	0,59
<i>Ehrlich sein</i>	Mittelwert	4,47	4,45	4,75
	Std.Abweichung	0,82	0,65	0,45

Die grauen Markierungen kennzeichnen die vergleichsweise höchsten Mittelwerte zwischen den Versuchsgruppen je Dimension.

Während die Testpersonen der VG 3 den höchsten Mittelwert (graue Markierung) beim Bemühen ehrlich zu antworten, sowie hinsichtlich der Leichtigkeit, sich in

die Situation zu versetzen, erzielten, erreichte VG 2 hinsichtlich des Bemühens, einen guten Eindruck zu hinterlassen, den vergleichsweise höchsten Mittelwert.

Es zeigt sich, dass die Differenzen der Mittelwerte der VG 1 und 2 in allen drei Dimensionen wesentlich geringer ausfallen, als dies für die entsprechenden Differenzen zu VG 3 der Fall ist.

Um überprüfen zu können, welchen subjektiven Effekt die Verfälschungsinstruktionen auf das Verfälschungsverhalten der Testpersonen hatte, bzw. inwiefern signifikante Mittelwertsunterschiede zwischen den drei Versuchsgruppen hinsichtlich der Beantwortung der drei Fragen vorliegen, war die Durchführung einer multivariaten Varianzanalyse (MANOVA) beabsichtigt.

Die Überprüfung der Voraussetzung der Homogenität der Varianz-Kovarianz-Matrix wurde für die drei Dimensionen (i.S. der Zusatzfragen) mittels Box-M-Test durchgeführt. Der Box-M-Test dient dabei der Überprüfung der Nullhypothese, dass die beobachteten Varianz-Kovarianz-Matrizen der abhängigen Variablen über die Gruppen gleich sind (Rasch & Kubinger, 2006). Bei einem Signifikanzniveau von $\alpha=0,05$ weist der Box-M-Test ein signifikantes Ergebnis mit einem p-Wert von 0,000 auf (siehe Tabelle 12).

Tabelle 12: Box-M-Test zur Überprüfung der Homogenität der Varianzen

Box-M-Test	129,328
<i>F</i>	10,651
<i>df1</i>	12
<i>df2</i>	667681,526
<i>Sig.</i>	0,000

Die graue Markierung bezeichnet jenen Wert, der bei einem Signifikanzniveau von $\alpha=0,05$ signifikant ist.

Somit ist die Annahme homogener Varianzen verletzt. Da die Signifikanzwerte des Levene-Tests auf Gleichheit der Fehlervarianzen die Verletzung der Annahme homogener Varianzen für zwei der drei Dimensionen (0,001 für *Guter Eindruck* und 0,000 für *Ehrlich sein*) bestätigten, wurde von der Durchführung einer multivariaten Varianzanalyse aufgrund nicht erfüllter Voraussetzungen abgesehen. Es wurde stattdessen der gegenüber ungleichen Varianzen robuste T-Test bei unabhängigen Stichproben unter Annahme ungleicher Varianzen durchgeführt (Welch-Test). Aufgrund der Kumulierung des Risikos 1. Art bei

Durchführung mehrerer Signifikanztests innerhalb derselben Fragestellung (Kubinger et al., 2011), wurde das Signifikanzniveau gemäß der Bonferroni-Methode angepasst. Dies führt für die folgend dargestellten Signifikanzwerte zu einem $\alpha_{\text{Vergleich}} = \alpha/m = 0,05/3 \approx 0,0167$ (m entspricht der Anzahl durchgeführter Tests, in diesem Fall $m=3$).

Die Ergebnisse zum paarweisen T-Test bei unabhängigen Stichproben unter Annahme ungleicher Varianzen finden sich in Tabelle 13.

Tabelle 13: T-Test für unabhängige Stichproben bei ungleichen Varianzen

Abhängige Variable	Levene-Test		T-Test für MW-Gleichheit		
	<i>F</i>	<i>Sig.</i>	<i>T</i>	<i>df</i>	<i>Sig. (2-seitig)</i>
VG 1 vs. VG 2					
<i>Leichtes Hineinversetzen</i>	0,05	0,83	0,45	244,73	0,66
<i>Guter Eindruck</i>	0,01	0,94	-0,34	245,99	0,73
<i>Ehrlich sein</i>	1,38	0,24	0,25	232,30	0,80
VG 2 vs. VG 3					
<i>Leichtes Hineinversetzen</i>	0,17	0,68	-0,85	247,23	0,397
<i>Guter Eindruck</i>	12,86	0,00	4,72	206,43	0,00
<i>Ehrlich sein</i>	35,86	0,00	-4,22	220,34	0,00
VG 1 vs. VG 3					
<i>Leichtes Hineinversetzen</i>	0,04	0,84	-0,41	250,87	0,683
<i>Guter Eindruck</i>	13,48	0,00	4,29	205,62	0,00
<i>Ehrlich sein</i>	27,76	0,00	-3,26	187,72	0,001

Die graue Markierung bezeichnet jene Werte, die bei einem Signifikanzniveau von $\alpha=0,0167$ signifikant sind.

Unter Berücksichtigung des mittels Bonferroni-Adjustierung angepassten Signifikanzniveaus von $\alpha=0,0167$ kann von einem signifikanten Mittelwertsunterschied innerhalb der Versuchsgruppen 2 und 3 sowie 1 und 3 in den Skalen *Guter Eindruck* sowie *Ehrlich sein* (graue Markierung in obiger Tabelle) ausgegangen werden.

8.1.2 Interpretation der Ergebnisse

Unter Betrachtung der Ergebnisse zur Skala *Guter Eindruck*, scheinen die Testpersonen der Versuchsgruppe 3 weniger stark bemüht gewesen zu sein einen guten Eindruck mittels Beantwortung der Fragen zu vermitteln, als die Testpersonen der VG 1 und 2. Dies kann als Hinweis gedeutet werden, dass zumindest aus subjektiver Sicht der Testpersonen, die Instruktionen ihrem Zweck zum Induzieren ehrlichen Antwortverhaltens innerhalb der VG 3 bzw. stärkerem Verfälschungsverhalten in VG 1 und 2, in gewünschter Art und Weise gedient haben.

Aus den Ergebnissen zur Skala *Ehrlich sein* kann geschlossen werden, dass die Testpersonen der VG 3, gemäß der Aufforderung zum ehrlichen Antwortverhalten innerhalb des Instruktionstextes, ihrer eigenen Einschätzung nach stärker bemüht waren, die Fragen ehrlich zu beantworten, als dies für die Testpersonen der VG 1 und 2 der Fall war. Dies lässt ebenfalls den Rückschluss auf den gewünschten subjektiven Effekt der Instruktionen zu.

Die Tatsache, dass kein signifikanter Unterschied zwischen den drei Versuchsgruppen in der Skala *Leichtes Hineinversetzen* bestand, stellt einen wünschenswerten Umstand dar, da die Fähigkeit, sich in die beschriebene Situation Hineinzuversetzen somit als Einflussvariable auf das Verfälschungsverhalten ausgeschlossen werden kann.

Da es sich bei diesen Ergebnissen lediglich um die subjektive Einschätzung des eigenen Antwortverhaltens seitens der Testpersonen der jeweiligen Versuchsgruppe handelt, finden sich im folgenden die Ergebnisse zum tatsächlich, objektiv beobachtbaren Verfälschungsverhalten der Testpersonen.

8.2 Ergebnisse zum tatsächlichen Verfälschungsverhalten

Zur Überprüfung der unter Abschnitt 6.1 (S.42) dargestellten Hypothesen 1a, 2a und 3a wurde das Programmpaket *SPSS für Windows* (Version 20) herangezogen. Vor Durchführung der eigentlichen Analysen wurden zunächst fehlende Werte durch den pro Person und Dimension ermittelten Mittelwert ersetzt, negativ gepolte Items der Skala *Soziale Erwünschtheit* umgepolzt und die zur weiteren Analyse benötigten Skalenrohwerte für die insgesamt 20

Dimensionen (14 Dimensionen des BIP, fünf Dimensionen des NEO-FFI sowie eine Soziale Erwünschtheits-Dimension) ermittelt.

8.2.1 Durchführung der Analyse und Darstellung der Ergebnisse

Im folgenden findet sich zunächst eine Tabelle zur Darstellung der Mittelwerte und Standardabweichungen nach Versuchsgruppe und Skala. Die grauen Markierungen bezeichnen jene Versuchsgruppe, für die der Mittelwert in der jeweiligen Fragebogenskala am höchsten ausgeprägt ist. Dabei entspricht eine höhere Mittelwertsausprägung einer stärkeren Ausprägung der jeweiligen Skala.

Tabelle 14: Mittelwerte und Standardabweichungen aller 20 Skalen

Bereich	Abhängige Variable		VG 1	VG 2	VG 3
AV	BIP Flexibilität	Mittelwert	35,06	35,80	37,24
		Std.Abw.	7,90	8,46	8,65
AV	BIP Gewissenhaftigkeit	Mittelwert	30,60	30,63	32,91
		Std.Abw.	8,40	8,94	10,53
AV	BIP Handlungsoptimierung	Mittelwert	29,90	28,99	31,98
		Std.Abw.	6,90	8,13	8,40
BO	BIP Führungsmotivation	Mittelwert	35,06	35,80	37,24
		Std.Abw.	7,90	8,46	8,65
BO	BIP Gestaltungsmotivation	Mittelwert	31,61	32,17	32,00
		Std.Abw.	6,23	6,25	6,41
BO	BIP Leistungsmotivation	Mittelwert	31,29	30,75	32,24
		Std.Abw.	7,68	7,56	7,73
PK	BIP Belastbarkeit	Mittelwert	28,46	27,72	30,68
		Std.Abw.	7,58	8,19	8,16
PK	BIP Emotionale Stabilität	Mittelwert	42,98	42,42	44,66
		Std.Abw.	9,74	10,00	9,05
PK	BIP Selbstbewusstsein	Mittelwert	43,27	43,51	44,72
		Std.Abw.	9,55	10,04	11,21
SK	BIP Durchsetzungsstärke	Mittelwert	29,99	30,19	30,34
		Std.Abw.	6,45	7,39	7,95
SK	BIP Kontaktfähigkeit	Mittelwert	38,27	38,23	37,27
		Std.Abw.	10,72	10,79	10,70
SK	BIP Sensitivität	Mittelwert	24,23	24,04	25,06
		Std.Abw.	6,02	5,63	6,37
SK	BIP Soziabilität	Mittelwert	33,14	33,08	34,92
		Std.Abw.	7,18	6,38	7,95
SK	BIP Teamorientierung	Mittelwert	32,55	32,77	35,38
		Std.Abw.	7,96	8,02	8,99
NEO	NEO-FFI Extraversion	Mittelwert	25,27	23,92	25,02
		Std.Abw.	6,65	6,03	5,63

NEO	<i>NEO-FFI Gewissenhaftigkeit</i>	Mittelwert	21,49	21,37	23,25
		Std.Abw.	5,75	6,32	7,04
NEO	<i>NEO-FFI Neurotizismus*</i>	Mittelwert	45,37	45,04	42,64
		Std.Abw.	8,14	8,59	8,10
NEO	<i>NEO-FFI Offenheit</i>	Mittelwert	27,80	27,70	27,75
		Std.Abw.	7,53	7,50	8,04
NEO	<i>NEO-FFI Verträglichkeit</i>	Mittelwert	27,50	26,89	28,19
		Std.Abw.	6,49	6,13	6,72
SES	<i>Soziale Erwünschtheit</i>	Mittelwert	67,38	66,86	66,04
		Std.Abw.	8,64	8,04	8,30

Die grauen Markierungen kennzeichnen die vergleichsweise höchsten Mittelwerte zwischen den Versuchsgruppen je Dimension, bzw. jene Versuchsgruppe, die im Vergleich mit den anderen beiden Versuchsgruppen das am stärksten sozial erwünschte Antwortverhalten aufweist. Die Skala Neurotizismus (*) erhebt entgegengesetzt der sozialen Erwünschtheit, somit muss hier der geringste Mittelwert als jener bezeichnet werden, der in höchstem Ausmaß sozial erwünscht ausfällt.

Aus obiger Tabelle ist bereits ersichtlich, dass die Mittelwerte der VG 3 in 14 von 20 Dimensionen über jenen der anderen beiden Versuchsgruppen liegen. In fünf der Dimensionen (Kontaktfähigkeit (BIP), Extraversion (NEO-FFI), Neurotizismus (NEO-FFI), Offenheit (NEO-FFI), Soziale Erwünschtheit (SES-17)) liegen die Mittelwerte der VG 1 (stärkste Verfälschungsinstruktion) über denen der anderen beiden Versuchsgruppen. Für die Skala Neurotizismus (NEO-FFI) entspricht eine höhere Ausprägung einem geringeren Maß sozial erwünschten Antwortverhaltens.

Zur Überprüfung der Frage nach dem Einfluss der unterschiedlichen Instruktionen auf das Antwortverhalten im Fragebogen und damit auf die Ausprägung in den 20 Fragebogenskalen, war die Durchführung einer multivariaten Variananalyse (MANOVA) beabsichtigt. Die Überprüfung der Voraussetzungen einer MANOVA wurden für die 20 vorliegenden Dimensionen mittels Box-M-Test durchgeführt.

Tabelle 15: Box-M-Test zur Überprüfung der Homogenität der Varianzen der 20 Dimensionen

Box-M-Test	572,212
<i>F</i>	1,264
<i>df1</i>	420
<i>df2</i>	378954,795
<i>Sig,</i>	0,000

Die graue Markierung bezeichnet jenen Wert, der bei einem Signifikanzniveau von $\alpha=0,05$ signifikant ist.

Bei einem Signifikanzniveau von $\alpha=0,05$ weist der Box-M-Test ein signifikantes Ergebnis mit einem p-Wert von 0,000 auf (siehe Tabelle 15). Somit ist die Annahme homogener Varianzen verletzt.

Da die Signifikanzwerte des Levene-Tests auf Gleichheit der Fehlervarianzen die Verletzung der Annahme homogener Varianzen lediglich hinsichtlich einer Dimension (0,025 für Handlungsoptimierung, BIP) bestätigten, wurde versucht, mittels Ausschlusses jener Skalen, welche die heterogenen Varianzen bzw. Kovarianzen verursachten, ein nicht-signifikantes Ergebnis des Box-M-Test zu erzielen. Dabei wurden Interkorrelationsmatrizen (Spearman-Rang-Korrelationen) je Versuchsgruppe erstellt (Anhang, S.165). Es folgte die Ermittlung der Differenzen der jeweiligen Skaleninterkorrelationen zwischen den drei Versuchsgruppen (Differenzen zwischen VG 1 und 2, 2 und 3 sowie 1 und 3). Aus den drei daraus resultierenden Differenztabellen (Anhang, S.174) wurden nun jene Fragebogenskalen, welche auffallend hohe Differenzwerte aufwiesen, gemäß Kubinger et al. (2011) extrahiert, nach Differenzhöhe sowie -häufigkeit gerangreicht und aus dem Box-M-Test in dieser Reihenfolge ausgeschlossen. Die folgende Tabelle zeigt die Ausschlussreihenfolge der auffälligen Dimensionen, sowie den aus dem Ausschluss der jeweiligen Dimension folgenden Signifikanz-Wert des Box-M-Tests.

Tabelle 16: Effekt des Dimensions-Ausschlusses auf Box-M-Test

Dimension(in Ausschlussreihenfolge)	Sig. (Box-M-Test)
<i>Soziale Erwünschtheit</i>	0,000
<i>BIP Soziabilität</i>	0,001
<i>BIP Gestaltungsmotivation</i>	0,012
<i>BIP Handlungsoptimierung</i>	0,010
<i>NEO-FFI Offenheit für Erfahrungen</i>	0,011

Die graue Markierung bezeichnet jene Werte, die bei einem Signifikanzniveau von $\alpha=0,05$ signifikant sind.

Da auch der Ausschluss dieser fünf Skalen, wie in Tabelle 16 ersichtlich, nicht zum gewünschten Ergebnis eines nicht-signifikanten Box-M-Test führte, wurde die Durchführung einer MANOVA für alle Dimensionen verworfen.

Um den Effekt der Instruktionen auf die Fragebogenskalen überprüfen zu können, wurde in der Folge der gegenüber ungleichen Varianzen robuste T-Test bei unabhängigen Stichproben unter Annahme ungleicher Varianzen (Welch-Test) durchgeführt. Aufgrund der Kumulierung des Risikos 1. Art bei Durchführung mehrerer Signifikanztests innerhalb derselben Fragestellung (Kubinger et al., 2011), wurde das Signifikanzniveau gemäß der Bonferroni-Methode angepasst. Dies führt für die folgend dargestellten Signifikanzwerte zu einem $\alpha_{\text{Vergleich}} = \alpha/m = 0,05/20 = 0,0025$ (m entspricht der Anzahl durchgeführter Tests; in diesem Fall $m=20$). Die Ergebnisse des paarweisen T-Test für unabhängige Stichproben bei ungleichen Varianzen finden sich in der folgenden Tabelle 17.

Tabelle 17: T-Test für unabhängige Stichproben bei ungleichen Varianzen

Tabelle 27.1 T-Test für unabhängige Stichproben bei ungleichen Varianzen						
Abhängige Variable		Levene-Test		T-Test für MW-Gleichheit		
VG 1 vs. VG 2						
BIP		F	Sig.	T	df	Sig. (2-seitig)
AV	Flexibilität	0,11	0,74	0,-72	249,58	0,48
AV	Gewissenhaftigkeit	0,18	0,68	-0,03	249,75	0,98
AV	Handlungsoptimierung	3,03	0,08	0,96	243,91	0,34
BO	Führungsmotivation	0,11	0,74	-0,72	249,58	0,48
BO	Gestaltungsmotivation	0,12	0,73	-0,72	250,97	0,47

BO	<i>Leistungsmotivation</i>	0,03	0,86	0,56	250,98	0,58
PK	<i>Belastbarkeit</i>	0,69	0,41	0,76	249,21	0,45
PK	<i>Emotionale Stabilität</i>	0,45	0,50	0,45	250,70	0,65
PK	<i>Selbstbewusstsein</i>	0,57	0,45	-0,20	250,15	0,84
SK	<i>Durchsetzungsstärke</i>	1,80	0,18	-0,23	245,96	0,82
SK	<i>Kontaktfähigkeit</i>	0,00	0,97	0,04	250,95	0,97
SK	<i>Sensitivität</i>	0,54	0,46	0,26	250,13	0,80
SK	<i>Soziabilität</i>	1,08	0,30	0,07	248,08	0,94
SK	<i>Teamorientierung</i>	0,10	0,75	-0,22	250,94	0,83
NEO-FFI						
NEO	<i>Extraversion</i>	0,52	0,47	1,69	248,96	0,09
NEO	<i>Gewissenhaftigkeit</i>	1,39	0,24	0,16	248,38	0,87
NEO	<i>Neurotizismus</i>	1,20	0,28	0,32	250,05	0,75
NEO	<i>Offenheit</i>	0,14	0,71	0,10	251,00	0,92
NEO	<i>Verträglichkeit</i>	0,03	0,88	0,77	250,39	0,44
SES	<i>Soziale Erwünschtheit</i>	0,03	0,87	0,49	250,02	0,62
VG 2 vs. VG 3						
BIP		<i>F</i>	<i>Sig.</i>	<i>T</i>	<i>df</i>	<i>Sig. (2-seitig)</i>
AV	<i>Flexibilität</i>	0,43	0,51	-1,36	255,85	0,18
AV	<i>Gewissenhaftigkeit</i>	3,15	0,08	-1,88	252,61	0,06
AV	<i>Handlungsoptimierung</i>	0,82	0,37	-2,91	255,95	0,004
BO	<i>Führungsmotivation</i>	0,43	0,51	-1,36	255,85	0,18
BO	<i>Gestaltungsmotivation</i>	0,34	0,56	0,22	255,88	0,82
BO	<i>Leistungsmotivation</i>	0,34	0,56	-1,54	256,00	0,12
PK	<i>Belastbarkeit</i>	0,14	0,71	-2,92	255,36	0,004
PK	<i>Emotionale Stabilität</i>	1,83	0,18	-1,89	250,65	0,06
PK	<i>Selbstbewusstsein</i>	1,72	0,19	-0,91	255,00	0,36
SK	<i>Durchsetzungsstärke</i>	0,90	0,34	-0,15	255,83	0,88
SK	<i>Kontaktfähigkeit</i>	0,00	0,98	0,72	255,22	0,48
SK	<i>Sensitivität</i>	2,35	0,13	-1,36	254,50	0,17

SK	<i>Soziabilität</i>	3,61	0,06	-2,06	248,83	0,04
SK	<i>Teamorientierung</i>	3,62	0,06	-2,45	254,84	0,02
NEO-FFI						
NEO	<i>Extraversion</i>	1,19	0,28	-1,51	252,68	0,13
NEO	<i>Gewissenhaftigkeit</i>	1,30	0,26	-2,70	255,09	0,02
NEO	<i>Neurotizismus</i>	1,62	0,20	2,31	253,16	0,02
NEO	<i>Offenheit</i>	0,04	0,84	-0,04	255,86	0,97
NEO	<i>Verträglichkeit</i>	0,63	0,43	-1,62	255,46	0,11
SES	<i>Soziale Erwünschtheit</i>	0,19	0,66	2,91	254,70	0,004
VG 1 vs. VG 3						
BIP		<i>F</i>	<i>Sig.</i>	<i>T</i>	<i>df</i>	<i>Sig. (2-seitig)</i>
AV	<i>Flexibilität</i>	1,06	0,30	-2,12	256,33	0,04
AV	<i>Gewissenhaftigkeit</i>	5,02	0,03	-1,96	248,54	0,05
AV	<i>Handlungsoptimierung</i>	7,80	0,01	-2,19	250,87	0,03
BO	<i>Führungsmotivation</i>	1,06	0,30	-2,12	256,33	0,04
BO	<i>Gestaltungsmotivation</i>	0,06	0,81	-0,49	256,97	0,62
BO	<i>Leistungsmotivation</i>	0,55	0,46	-0,98	256,98	0,33
PK	<i>Belastbarkeit</i>	0,18	0,67	-2,27	256,69	0,02
PK	<i>Emotionale Stabilität</i>	0,30	0,53	-1,44	253,81	0,15
PK	<i>Selbstbewusstsein</i>	4,21	0,04	-1,12	253,32	0,26
SK	<i>Durchsetzungsstärke</i>	5,46	0,02	-0,39	249,95	0,70
SK	<i>Kontaktfähigkeit</i>	0,005	0,94	0,76	256,56	0,45
SK	<i>Sensitivität</i>	0,61	0,44	-1,08	256,92	0,28
SK	<i>Soziabilität</i>	0,77	0,38	-1,89	255,98	0,06
SK	<i>Teamorientierung</i>	2,72	0,10	-2,68	255,25	0,008
NEO-FFI						
NEO	<i>Extraversion</i>	3,03	0,08	0,32	246,77	0,75
NEO	<i>Gewissenhaftigkeit</i>	5,27	0,02	-2,21	250,50	0,03
NEO	<i>Neurotizismus</i>	0,03	0,86	2,71	256,49	0,007
NEO	<i>Offenheit</i>	0,28	0,60	0,05	256,82	0,96

NEO	Verträglichkeit	0,36	0,55	-0,83	257,00	0,41
SES	Soziale Erwünschtheit	0,33	0,57	3,32	252,50	0,001

Die graue Markierung bezeichnet jene Werte, die bei einem Signifikanzniveau von $\alpha=0,0025$ signifikant sind.

Unter Berücksichtigung des mittels Bonferroni-Adjustierung angepassten Signifikanzniveaus von $\alpha=0,0025$ kann von einem signifikanten Mittelwertsunterschied lediglich zwischen Versuchsgruppe 1 und 3 in der Skala Soziale Erwünschtheit (graue Markierung, siehe Tabelle 17) ausgegangen werden.

8.2.2 Interpretation der Ergebnisse

Um auf obigen Ergebnissen aufbauend Aussagen bezüglich der aufgestellten Hypothesen (s. Abschnitt 6.1, S. 42) zu treffen, muss überprüft werden ob diese Mittelwertsunterschiede in die der Alternativhypothese entsprechende Richtung ausschlagen. Die angegebenen Mittelwerte aus Tabelle 14 enthalten diesbezüglich einen Hinweis, der im folgenden grafisch veranschaulicht wird.

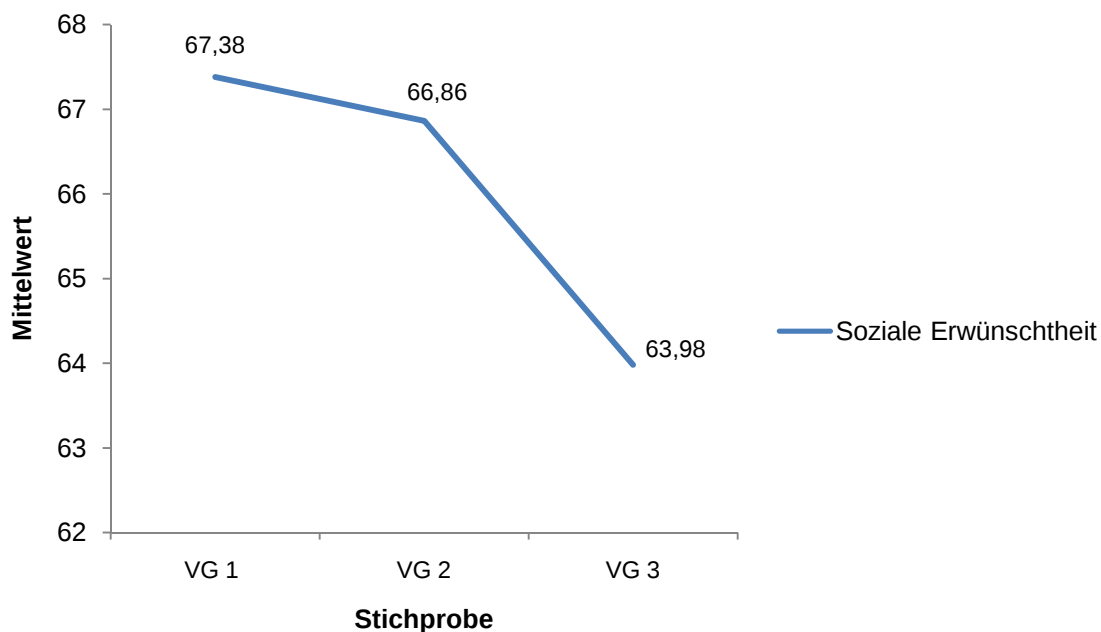


Abbildung 6: Veranschaulichung der Mittelwerte Soziale Erwünschtheit

Abbildung 6 zeigt, dass der Mittelwert der dritten Versuchsgruppe unterhalb derer, der Versuchsgruppe 1 (und 2) liegt. Somit hat VG 1 (und 2), gemäß der

verwendeten Skala, ein stärkeres sozial erwünschtes Verfälschungsverhalten gezeigt, als dies für VG 3 der Fall ist.

Es folgt die Annahme eines statistisch bedeutsamen Effekts der Instruktionen auf die Skala Soziale Erwünschtheit zwischen den Versuchsgruppen 1 und 3 in die erwartete Richtung.

Folglich wird die **Alternativhypothese $H_{(1)1a}$** für die **Versuchsgruppe 1** und die Skala **Soziale Erwünschtheit** angenommen:

Unter Vorgabe der 1. Instruktion (VG 1) zeigt sich ein signifikant erhöhtes Ausmaß an Verfälschungsverhalten (sozial erwünschtes Antwortverhalten, *Impression Management*) bei der Beantwortung der Fragebogenitems, verglichen mit der 3. Instruktion, sowie ein erhöhtes Ausmaß an Verfälschungsverhalten verglichen mit der 2. Instruktion.

Weiters wird die **Nullhypothese $H_{(0)2a}$** für **Versuchsgruppe 2** für die Skala **Soziale Erwünschtheit** beibehalten:

Unter Vorgabe der 2. Instruktion (VG 2) zeigt sich kein signifikant erhöhtes Ausmaß an Verfälschungsverhalten (sozial erwünschtes Antwortverhalten, *Impression Management*) bei der Beantwortung der Fragebogenitems, verglichen mit der 3. Instruktion.

Hinsichtlich der **19 übrigen Skalen**, werden die in Abschnitt 6.1 formulierten **Null-Hypothesen $H_{(0)1a}$ und $H_{(0)2a}$** angenommen.

Abschliessend ist zu sagen, dass die vorliegenden Ergebnisse hinsichtlich der Unterschiede im Verfälschungsverhalten zwischen den drei Versuchsgruppen eine weitere Untersuchung zum Zusammenhang von Verfälschungsverhalten und der Verwendung von Antworttendenzen ausschliesslich für die Skala Soziale Erwünschtheit ermöglichen. Hinsichtlich der übrigen Skalen konnten zumindest hinsichtlich des Verfälschungsverhaltens (gem. Abschnitt 3.1.1) keine signifikanten Unterschiede zwischen den drei Versuchsgruppen festgestellt werden.

9 Ergebnisse zu IRT- Analysen der Antworttendenzen

Wie weiter oben ausgeführt, macht eine Analyse zum Zusammenhang von Verfälschungsverhalten und Antworttendenzen ausschliesslich für die Skala

Soziale Erwünschtheit Sinn. Für die übrigen 19 Skalen kann zu den entsprechenden Hypothesen 1b, 2b und 3 (Abschnitt 6.1, S.42) keine Stellung bezogen werden. Es werden für diese Skalen dennoch IRT-Analysen durchgeführt um feststellen zu können, inwiefern Antworttendenzen mittels der entsprechenden Pseudo-Items erhoben werden können und wie die Modellpassungs-Güte auf Basis der Pseudo-Item-Kodierung für die berechneten Modelle gelingt.

Im Rahmen der vorliegenden Arbeit wurden sämtliche IRT-Modelle auf Basis des eindimensionalen dichotomen logistischen Testmodells von Rasch (Rasch-Modell, 1 PL-Modell, Kubinger et al., 2011) geschätzt. Das Rasch-Modell postuliert, dass die Lösungswahrscheinlichkeit eines Items i für Person v (oder im Falle von Fragebogenitems: die Wahrscheinlichkeit zugunsten eines Persönlichkeitsmerkmals zu antworten) von zwei Parametern abhängig ist. Nämlich der Itemschwierigkeit σ_i (oder der Schwierigkeit dem Item zuzustimmen) und dem Fähigkeitsparameter (bzw. dem Merkmalsausprägungs-Parameter) der Person, ξ_v . Darauf aufbauend kann der Zusammenhang für die Lösungswahrscheinlichkeit des Items i folgendermaßen dargestellt werden:

$$P("1" | \xi_v, \sigma_i) = \frac{e^{\xi_v - \sigma_i}}{1 + e^{\xi_v - \sigma_i}}$$

In Bezug auf die in der vorliegenden Arbeit betrachteten Antworttendenzen, wird ξ_v als die personenabhängige Tendenz eine gewisse Antworttendenz zu verwenden betrachtet. Der Parameter σ_i stellt die Attraktivitätsausprägung eines Items dar, mit der es eine Antworttendenz bei einer Person hervorruft. $P("1" | \xi_v, \sigma_i)$ beschreibt somit die Wahrscheinlichkeit des Vorliegens einer Antworttendenz für Item i , gezeigt von Person v . Je grösser ξ_v und je kleiner σ_i , desto grösser ist somit die Wahrscheinlichkeit des Vorliegens einer Antworttendenz.

Zur Überprüfung der unter Abschnitt 6.1 (S.42) dargestellten Hypothesen 1b, 2b und 3 in Bezug auf die Skala Soziale Erwünschtheit, sowie zur Durchführung der Hypothesen-unabhängigen Antworttendenz-Analysen für die 19 übrigen Skalen wurden eindimensionale und mehrdimensionale Rasch-Modelle berechnet. Dafür wurde die Software *mdltm* (Software for Multidimensional Discrete Latent Trait Models; von Davier, 2005) herangezogen, welche Marginale Maximum-

Likelihood-Schätzungen (MML) über die Verwendung des EM-Algorithmus (E steht für Erwartungswerte, M für Maximierung) durchführt. *mdltm* ermöglicht die Schätzung von Generellen Diagnostik Modellen (GDM), Log-Linearen Kognitiven Diagnostik Modellen (LCDM), sowie Mischverteilungs-Modellen und hierarchischen GDM.

Für die IRT Modellvergleiche wurden die Informationsmaße Akaike Informationskriterium (Akaike Information Criterion, AIC) sowie das Bayes` Informationskriterium (Bayesian Information Criterion, BIC) herangezogen. Diese Informationsmaße „beziehen sich auf die Likelihoodfunktion der Daten unter der getroffenen Modellannahme“ (Kubinger et al., 2011, S. 504).

$$\begin{aligned} \text{AIC} &= 2k - 2 \ln L(y | \hat{\theta}) \\ \text{BIC} &= k \ln n - 2 \ln L(y | \hat{\theta}) \end{aligned}$$

$L(y | \hat{\theta})$ stellt dabei die entsprechende Likelihood dar, y steht für die beobachteten Daten, n stellt die Untersuchungseinheiten dar und $\hat{\theta}$ steht schliesslich für die k Modellparameter. Nach Kubinger et al. (2011) stellen die Informationskriterien jenen Informationsverlust dar, der dadurch zustande kommt, dass die beobachteten Daten lediglich über die jeweiligen Modelle repräsentiert werden und somit nicht der in der Realität existierenden Komplexität gerecht werden können. Daraus folgt, dass jenes Modell, welches den geringeren Informationskriteriums-Wert vorweisen kann (folglich geringeren Informationsverlust mit sich bringt), als besser an die Realität angepasst gilt. Die Werte aus AIC und BIC sind somit lediglich relativ, also im Vergleich auf Basis der selben Datengrundlage, nicht jedoch absolut zu interpretieren.

Im Falle uneindeutiger Aussagen hinsichtlich dieser beiden Kriterien wurden vom Modell abweichende Items sukzessive anhand der jeweiligen Itemfits (RMSEA-Werte) ausgeschlossen, um eine bessere Modellanpassung zu erzielen (s. Erläuterungen zu Tabelle 23).

Aufbereitung der Daten – Kodierung von Pseudo-Items

Die IRT-Analysen wurden für BIP, NEO-FFI und die Soziale Erwünschtheits-Skala getrennt vorgenommen um differenzierte inhaltliche Aussagen tätigen zu

können. Aufgrund der (für IRT-Analysen eher geringen) Stichprobengröße und um die Rechenleistung für die Software zumutbar zu halten, wurden die IRT-Analysen des BIP zusätzlich getrennt für die 4 Bereiche des BIP (Arbeitsverhalten: AV; Berufliche Orientierung: BO; Psychische Konstitution: PK; Soziale Kompetenz: SK), sowie den Skalen des NEO-FFI (NEO) vorgenommen. Zur Generierung der binären Pseudo-Items (für jedes Item werden nur die Werte 0 und 1 vergeben) zur Erfassung von ERS, MRS sowie konstrukt-relevanten (von Antworttendenzen weitgehend unverzerrten) Antworten, wurden die ursprünglichen Itemantworten (von 1 bis 5) wie bei Khorramdel und von Davier (under review) umkodiert.

Auf Basis des in Abbildung 2 (Abschnitt 5, S. 39) dargestellten Entscheidungsbaumes wurde der ursprüngliche Datensatz (Ratings gemäß der Likert-Skala von 1 bis 5) in dreifacher Weise (je Pseudo-Item ein Datensatz) umkodiert. Dabei repräsentiert jeder der drei neu geschaffenen Datensätze eine der aus Abbildung 2 ersichtlichen Antworttendenzen: (1.) Antworten bei denen die mittlere Antwortkategorie gewählt wurde (Pseudo-Item m), (2.) Antworten, bei denen eine positive oder negative Extremantwort (1. oder 5. Antwortkategorie der 5-stufigen Likert-Skala) gewählt wurde (Pseudo-Item e), sowie (3.) nicht-extreme und extreme Antworten (4. oder 5. Kategorie einer fünfstufigen Likert-Skala, Pseudo-Item d). Die Repräsentation der ursprünglichen Item-Antworten gemäß dieser Methodik (Zerteilung des Gesamtantwortprozesses in einzelne sequentielle Sub-Antwortprozesse) erlaubt in der Folge eine Modellierung des Antwortprozesses sowohl hinsichtlich der Verwendung von Antworttendenzen mittels m- und e-Items (als Repräsentanten von ERS und MRS) als auch hinsichtlich der inhaltsbezogenen Antwort mittels der Pseudo-Items d. Die Kodierung der ursprünglichen Items in diese drei Datensätze (Pseudo-Items d, e und m) ist in Anlehnung an Khorramdel und von Davier (under review) in der folgenden Tabelle illustriert.

Tabelle 18: Kodierung der Pseudo-Items d, e und m

ursprüngl. Item-Antwort (5-stufige Likert- Skala)	Pseudo-Item d	Pseudo-Item e	Pseudo-Item m	Wahrscheinlichkeit f. die entspr. Antwortkategorie
1	0	1	0	P(1)
2	0	0	0	P(2)
3	9	9	1	P(3)
4	1	0	0	P(4)
5	1	1	0	P(5)
Fehlender Wert	9	9	9	

Dabei wird die Wahl der mittleren Antwortkategorie (3 auf Likert-Skala) im Rahmen der Kodierung von e- und d-Items nicht mit einer 0, sondern mit einem Code für fehlende Werte (hier 9) kodiert, da keine implizite Abhängigkeit zwischen den drei Pseudo-Items angenommen wird (vgl. Khorramdel & von Davier, under review). Fehlende Werte werden wie bei Khorramdel und von Davier (under review) ebenfalls mit 9 kodiert.

Gemäß der unterschiedlichen Ergebnisse hinsichtlich des Verfälschungsverhaltens wird im folgenden zunächst auf die Ergebnisse zur Skala Soziale Erwünschtheit eingegangen (Abschnitt 9.1, Ergebnisse zur Skala Soziale Erwünschtheit), um einen möglichen Zusammenhang zwischen Verfälschungsverhalten und dem Auftreten von Antworttendenzen zu untersuchen.

Sodann folgen die Ergebnisse zum Vorliegen möglicher Antworttendenzen hinsichtlich der restlichen 19 Skalen (Abschnitt 9.2, Ergebnisse zu den restlichen 19 Skalen).

9.1 Ergebnisse zur Skala Soziale Erwünschtheit

9.1.1 Durchführung der Analyse

Um ERS und MRS auf Basis der Pseudo-Items (e und m) zu modellieren und von konstrukt-relevanten Antworten (Pseudo-Items d) zu differenzieren, wurden eindimensionale und multidimensionale Rasch-Modelle getrennt für die drei

Versuchsgruppen (VG) und für die Gesamtstichprobe (GSP) berechnet und verglichen.

Da die SES-Skala nur eine Dimension darstellt ist eine Analyse auf Pseudo-Item-Ebene (vgl. Khorramdel & von Davier, under Review; von Davier & Khorramdel, in print) mittels eindimensionaler Pseudo-Item-Faktoren nicht möglich. Es können lediglich IRT-Modelle mit multidimensionalen Pseudo-Item-Faktoren unter Einbeziehung aller Pseudo-Item-Skalen berechnet werden. Eine Veranschaulichung der berechneten Modelle für die SES-Skala findet sich in Tabelle 19.

Tabelle 19: Modellberechnungen Skala Soziale Erwünschtheit (VG und GSP)

Modelle			
A 1-dimensional	B 3-dimensional	C 2-dimensional	D 2-dimensional
e-Items, m-Items und d-Items werden einem gemeinsamen Faktor zugeordnet, der für die SES-Skala steht.	e-Items, m-Items und d-Items werden jeweils einem Faktor zugeordnet: ERS-Faktor (e-Dimension), MRS-Faktor (m-Dimension), Trait-Faktor (d-Dimension)	d- und e-Items werden einem gemeinsamen Trait-Faktor (d&e-Dimension), die m-Items einem MRS-Faktor (m-Dimension) zugeordnet	d- und m-Items werden einem gemeinsamen Trait-Faktor (d&m-Dimension), die e-Items einem ERS-Faktor (e-dimension) zugeordnet

Modell A stellt die Hypothese dar, dass sämtliche Pseudo-Items die eigentlich zu messende Skala (SES) erfassen und keine Antworttendenzen vorliegen. Modell B trägt der Idee Rechnung, dass sowohl extreme (ERS) als auch mittlere Antworttendenzen (MRS) neben der Erfassung der eigentlichen SES-Skala (dem Trait; d-Items) vorliegen. Modell C verfolgt die Hypothese, dass neben der eigentlichen SES-Skala (erfasst durch d-Items und e-Items) ein MRS vorliegt (m-Items), aber kein ERS, während Modell D auf der Hypothese basiert, dass neben der SES-Skala (erfasst durch d-Items und m-Items) ein ERS (e-Items) vorliegt, aber kein MRS.

9.1.2 Darstellung der Ergebnisse

Um den Zusammenhang zwischen dem Verfälschungsverhalten und dem Vorliegen von Antworttendenzen überprüfen zu können, liegt der Fokus hier auf einem differenzierten Ergebnisvergleich der drei Versuchsgruppen hinsichtlich der Modellpassungsgüte, so dass die Ergebnisse der Gesamtstichprobe lediglich zur Ergänzung im Anhang zu finden sind (ab S. 208).

Tabelle 20 beinhaltet nun die Ergebnisse des AIC- und BIC-Kriteriums für die drei Versuchsgruppen; je niedriger das AIC- bzw. BIC-Kriterium, desto besser die Modellpassung.

Tabelle 20: Modellvergleiche der Skala Soziale Erwünschtheit (VG) auf Basis aller Pseudo-Items

Soziale Erwünschtheit SES-Skala	Kriterium	Modell A 1- dimensional SES-Skala	Modell B 3- dimensional Pseudo-Item Skalen	Modell C 2- dimensional d-&e-Skala, m-Skala	Modell D 2- dimensional d-&m-Skala, e-Skala
VERSUCHSGRUPPE 1					
d-, e-, m-Items	<i>AIC-Index</i>	5192,05	5023,09	5010,01	5072,99
	<i>BIC-Index</i>	5348,48	5216,50	5183,50	5246,48
VERSUCHSGRUPPE 2					
d-, e-, m-Items	<i>AIC-Index</i>	5255,94	5026,02	5003,71	5159,38
	<i>BIC-Index</i>	5411,93	5218,89	5176,72	5332,39
VERSUCHSGRUPPE 3					
d-, e-, m-Items	<i>AIC-Index</i>	5866,83	5702,48	5731,63	5759,10
	<i>BIC-Index</i>	6025,38	5898,51	5907,48	5934,95

Die grauen Markierungen kennzeichnen die vergleichsweise geringsten AIC/BIC -Werte der berechneten Modelle innerhalb eines Bereichs, die folglich die höchste Modellanpassungsgüte aufweisen.

Wie anhand der grauen Markierungen ersichtlich, sind die Ergebnisse zur Modellpassung nicht für alle Versuchsgruppen gleich. Während das 2-dimensionale Modell C (Pseudo-Items d und e bilden den inhaltsabhängigen Trait-Faktor, Pseudo-Items m den MRS-Faktor) für die Versuchsgruppen 1 und 2 die beste Passung aufweist, zeigt in Versuchsgruppe 3 das 3-dimensionalen Modell B (drei Faktoren, ein Faktor je Pseudo-Item Gruppe) die beste Passung. Zur näheren Betrachtung der angeführten Ergebnisse findet sich im folgenden eine Tabelle der geschätzten Mittelwerte sowie Standardabweichungen der Faktoren aus den Modellen B und C, da dies jene Modelle mit der jeweils höchsten Modellpassung gemäß Tabelle 20 sind. Des Weiteren sind die

marginalen IRT-basierten Reliabilitäts-Kennwerte (Sireci et al., 1991) der verschiedenen Faktoren innerhalb der einzelnen Modelle für die m- Dimension (für VG 1, 2 und 3; 3. Zeile) sowie e-Dimension (für VG 3) angeführt.

Tabelle 21: Geschätzte Mittelwerte, Standardabweichungen und die marginalen Reliabilitäten zur Skala Soziale Erwünschtheit (VG) für die Modelle B und C

Soziale Erwünschtheit SES-Skala	Kriterium	Modell B	Modell C
		3-dimensional Pseudo-Item Skalen	2-dimensional d-&e-Skala, m-Skala
VERSUCHSGRUPPE 1			
d-Dimension	MW	2,956	
	Std.Abw.	1,183	
e-Dimension	MW	0,008	
	Std.Abw.	1,27	
m-Dimension	MW	-0,720	-0,829
	Std.Abw.	1,30	1,20
	Reliabilität	0,651	0,630
d&e-Dimension	MW		1,404
	Std.Abw.		1,00
VERSUCHSGRUPPE 2			
d-Dimension	MW	2,691	
	Std.Abw.	1,50	
e-Dimension	MW	0,091	
	Std.Abw.	1,18	
m-Dimension	MW	0,038	-0,326
	Std.Abw.	1,05	0,91
	Reliabilität	0,619	0,615
d&e-Dimension	MW		0,961
	Std.Abw.		1,05
VERSUCHSGRUPPE 3			
d-Dimension	MW	1,940	
	Std.Abw.	1,03	
e-Dimension	MW	0,080	
	Std.Abw.	1,04	
	Reliabilität	0,740	
m-Dimension	MW	-0,109	-0,235
	Std.Abw.	0,88	0,77
	Reliabilität	0,537	0,511
d&e-Dimension	MW		0,680
	Std.Abw.		0,76

Die grauen Markierungen kennzeichnen die MW, Std.Abw. und Reliabilitäten jener Faktoren, welche den jeweiligen Modellen mit der höchsten Anpassungsüte gemäß AIC- und BIC-Werte entnommen sind.

9.1.3 Interpretation der Ergebnisse

Hinsichtlich des Vorliegens von Antworttendenzen lässt sich aus den dargestellten Ergebnissen nun folgendes ableiten: im Falle des stärksten Verfälschungsverhaltens (VG 1) gemäß der mittels SPSS durchgeführten Analysen in der SES-Skala, lässt sich die Pseudo-Item-Dimension e (ERS) nicht einzeln abbilden und muss somit mit der Pseudo-Item-Dimension d als gemeinsame, womöglich Iteminhalts-abhängige Dimension angesehen werden. Zumindest scheint es so, als würde Pseudo-Item e nicht ausschließlich ERS, sondern auch inhaltsrelevante Antworten beinhalten. Gleiches gilt auch für die Versuchsgruppe 2. Die Annahme eines Zusammenhangs zwischen Verfälschungsverhalten und der Verwendung der extremen Antworttendenz (ERS) kann somit nicht bestätigt werden. Eine eindimensionale Darstellung *beider* Antworttendenzen MRS und ERS lässt sich ausschliesslich für Versuchsgruppe 3 feststellen, welche auf Basis der im Vorfeld getätigten Analysen (s. Abschnitt 8.2) im Vergleich zu VG 1 und 2 in geringerem Maße Verfälschungsverhalten (i.S. sozialer Erwünschtheit) zeigt.

Die Pseudo-Item-Dimension m (m-Items, MRS) lässt sich für alle drei Versuchsgruppen 1, 2 und 3 abbilden. Die höhere Ausprägung des geschätzten Mittelwertes der m-Dimension für Versuchsgruppe 2 im Vergleich zu VG 1 im 2-dimensionalen Modell B (VG 1: -0,829; VG 2: -0,326) könnte vor dem Hintergrund der vorgegebenen Instruktion eine Reaktion auf die Unsicherheit der Testpersonen der VG 2, mittels Wahl der entsprechend unsicheren, nämlich mittleren Antwortkategorie darstellen. Somit kann von einem Zusammenhang zwischen Verfälschungsverhalten und der Verwendung der mittleren Antwortkategorie im Sinne einer Tendenz zur Mitte (MRS) in der hier vorliegenden Studie ausgegangen werden.

Für alle drei Versuchsgruppen ist weiters festzustellen, dass das eindimensionale Modell A die geringste Modellpassung (höchster AIC- sowie BIC-Wert i.V. zu Modellen B, C und D) aufweist und somit jedenfalls keine eindimensionale Skala im Sinne des zu messenden Konstrukts (Soziale Erwünschtheit) vorliegt.

Die marginalen Reliabilitäten der m-Dimension für VG 1 und 2 reichen von 0,615 bis 0,651. Für die VG 3 ergibt sich für die e-Dimension des 3-dimensionalen

Modells ein Reliabilitätswert von 0,740, jedoch geringere Reliabilitätswerte (0,537 und 0,511) für die m-Dimension des 3- bzw. 2-dimensionalen Modells. Ergänzende (marginale) Reliabilitätskennwerte zu den übrigen Faktoren finden sich im Anhang (ab S. 210).

Die Ergebnisse zeigen, dass in **Versuchsgruppe 1** keine vermehrte **Tendenz zu Extremurteilen (ERS)** auftritt, was zur Beibehaltung der **Nullhypothese $H_{(0)1b}$** führt.

Des Weiteren wird für die **Versuchsgruppe 2** hinsichtlich des Vorliegens einer **Tendenz zur Mitte (MRS)** die **Alternativhypothese $H_{(1)2b}$** angenommen, d.h. es tritt eine vermehrte Antworttendenz zur Mitte (MRS) auf.

Hinsichtlich der Ergebnisse zu **Versuchsgruppe 3** wird die **Nullhypothese $H_{(0)3}$** angenommen, d.h. für diese Versuchsgruppe können sowohl **extreme als mittlere Antworttendenzen (ERS oder MRS)** nachgewiesen werden.

Abschliessend ist zum Zusammenhang von Verfälschungsverhalten (i.S. sozialer Erwünschtheit) und dem Vorliegen von Antworttendenzen festzuhalten, dass ein entsprechender Zusammenhang für eine Tendenz zu Extremurteilen (ERS) nicht festzustellen ist. Die Tatsache, dass die Pseudo-Item-Dimension e lediglich im Falle des am geringsten ausgeprägten sozial erwünschten Antwortverhaltens als alleinige Dimension extrahiert werden konnte, bekräftigt den fehlenden Zusammenhang zu positivem Verfälschungsverhalten im vorliegenden Fall.

Das Vorliegen einer Antworttendenz zur Mitte (MRS) in beiden Verfälschungsgruppen (VG 1 und 2) lässt den Rückschluss auf einen Zusammenhang zwischen MRS und Verfälschungsverhalten zu. Dabei scheint ein (jedoch nicht-signifikant) schwächer ausgeprägtes Verfälschungsverhalten der VG 2 unter unsicherer Verfälschungsmotivation im Vergleich zu VG 1 (starke Verfälschungsmotivation) mit einer erhöhten Verwendung der mittleren Antwortkategorie (MRS) gemäß geschätztem IRT-basierten Mittelwert einherzugehen.

Die Ergebnisse zeigen, dass die Antworttendenz MRS (sowie teilweise auch ERS) als eigene Dimensionen abgebildet werden kann und die Fragebogenskala SES (unter Einbeziehung aller Antwortkategorien) keine eindimensionale Skala im Sinne des zu messenden Konstrukts darstellt.

9.2 Ergebnisse zu den restlichen 19 Skalen

Auf die Untersuchung der formulierten Hypothesen 1b, 2b und 3 im Falle von 19 der 20 untersuchten Dimensionen muss auf Grund nicht nachweisbaren Verfälschungsverhaltens verzichtet werden. Es wird jedoch mittels IRT Analysen überprüft, inwiefern die erhobenen Daten mittels der Pseudo-Item-Dimensionen d, e und m repräsentiert werden können bzw. inwiefern vom Vorliegen von einer Tendenz zu Extremwerten (ERS, mittels Pseudo-Item e) bzw. einer Tendenz zur mittleren Antwortkategorie (MRS mittels Pseudo-Item m) auszugehen ist. Des Weiteren soll überprüft werden, ob die Messung der Fragebogenskalen mittels Pseudo-Item d, jener mittels Original-Items (Likert-Skala) überlegen ist.

9.2.1 Durchführung der Analysen

Die folgende Tabelle dient der Veranschaulichung der berechneten Modelle:

Tabelle 22: Modellvergleiche I-IV

Modellvergleich	Pseudo-Items	Modelle		
für GESAMTSTICHPROBE sowie VG 1, 2 und 3				
I	d-Items	A <i>1-dimensional</i> d-Dimension	B <i>Mehrdimensional</i> Fragebogenskalen	
II	e-Items	A <i>1-dimensional</i> e-Dimension	B <i>Mehrdimensional</i> Fragebogenskalen	
III	m-Items	A <i>1-dimensional</i> m-Dimension	B <i>Mehrdimensional</i> Fragebogenskalen	
GESAMTSTICHPROBE				
IV	d-, e-, m-Items	A <i>3-dimensional</i> d-, e-, m-Dimension	B <i>Mehrdimensional</i> Fragebogenskalen	C <i>Mehrdimensional</i> Fragebogenskalen, e-, m-Dimension

Modellvergleiche I bis III werden sowohl für die Gesamtstichprobe, als auch für die einzelnen Versuchsgruppen durchgeführt. Modellvergleich IV wird ausschliesslich für die Gesamtstichprobe durchgeführt, da dieser umfassendere Modellvergleich IV auf Basis der für IRT-Analysen verhältnismässig kleinen

Stichprobengröße der einzelnen Versuchsgruppen keine genaue Modellschätzung erlauben würde.

Modellvergleich I überprüft die Annahme, dass Pseudo-Item d tatsächlich die eigentlichen Fragebogenskalen (Modell B), im Gegensatz zu einer einzelnen womöglich inhaltsunabhängigen Dimension (Modell A) erhebt.

Modellvergleich II überprüft inwiefern Pseudo-Item e zur modellhaften Repräsentation der Daten bzw. zur Erhebung der inhaltsunabhängigen Antworttendenz zu Extremurteilen (ERS, Modell A) im Gegensatz zu den eigentlichen Fragebogenskalen (Modell B) geeignet ist.

Modellvergleich III überprüft inwiefern Pseudo-Item m zur modellhaften Repräsentation der Daten bzw. zur Erhebung der inhaltsunabhängigen mittleren Antworttendenz (MRS, Modell A) im Gegensatz zu den eigentlichen Fragebogenskalen (Modell B) geeignet ist.

Modellvergleich IV überprüft mittels Integration aller Pseudo-Items d, e und m inwiefern sowohl extreme (ERS, erfasst durch e-Items) als auch mittlere Antworttendenzen (MRS, erfasst durch m-Items) neben der Erfassung der eigentlichen Fragebogen-Skalen (d-Items) erfasst werden können (Modell C). Das Modell B trägt dabei der Idee Rechnung, dass sämtliche Pseudo-Items (d, e und m) die eigentlich zu messenden Fragebogenskalen erfassen und Modell A erfasst neben der extremen (ERS, erfasst durch e-Items) als auch mittleren Antworttendenzen (MRS, erfasst durch m-Items) eine zusätzliche womöglich inhaltsunabhängige Dimension auf Basis der d-Items.

Die angeführten Modellvergleiche I bis IV werden getrennt für die vier Bereiche des BIP (AV, BO, PK, SK) sowie den NEO-FFI (NEO) berechnet.

9.2.2 Darstellung der Ergebnisse

Im folgenden finden sich die Ergebnisse zu:

- a) Modellpassung auf Basis der Gesamtstichprobe (Modellvergleiche I-III und IV)**
- b) Modellpassung auf Basis der VG 1, 2 und 3 (Modellvergleiche I-III)**
- c) Ergebnisse zur Messgüte des Pseudo-Item d Basis der GSP**

Im Anschluss daran findet sich die Interpretation der Ergebnisse aus diesen drei Bereichen a), b) und c).

a) Ergebnisse zur Modellpassung auf Basis der Gesamtstichprobe (GSP)

Modellvergleiche I-III für GSP

Tabelle 23 veranschaulicht die Ergebnisse der Modellvergleiche I bis III (Tabelle 22) für die Gesamtstichprobe (GSP). Dabei sind jeweils jene Felder grau markiert, deren AIC- und BIC-Werte auf die im Vergleich zwischen Modell A und B bessere Modellanpassung hindeuten.

Tabelle 23: Modellvergleiche I-III (GSP) auf Basis der einzelnen Pseudo-Items			
Arbeitsverhalten (AV) Gewissenhaftigkeit Flexibilität Handlungsoptimierung	Kriterium	A	B
		1-dimensional	3-dimensional
		d-/e-/m-Dimension	Fragebogen-Skalen
d-Items	<i>AIC-Index</i>	10989,41	9845,51
	<i>BIC-Index</i>	11551,14	10561,52
e-Items	<i>AIC-Index</i>	14913,57	14711,58
	<i>BIC-Index</i>	15475,29	15427,59
m-Items	<i>AIC-Index</i>	15172,58	15246,53
	<i>BIC-Index</i>	15734,31	15962,54
Beruf. Orientierung (BO) Leistungsmotivation Gestaltungsmotivation Führungsmotivation	Kriterium	A	B
		1-dimensional	3-dimensional
		d-/e-/m-Dimension	Fragebogen-Skalen
d-Items	<i>AIC-Index</i>	10000,48	9841,83
	<i>BIC-Index</i>	10550,34	10534,1
e-Items	<i>AIC-Index</i>	13420,97	13628,27
	<i>BIC-Index</i>	13970,83	14332,41
m-Items	<i>AIC-Index</i>	16721,99	16905,94
	<i>BIC-Index</i>	17271,85	17610,08
Psych. Konstitution (PK) Emotionale Stabilität Belastbarkeit Selbstbewusstsein	Kriterium	A	B
		1-dimensional	3-dimensional
		d-/e-/m-Dimension	Fragebogen-Skalen
d-Items	<i>AIC-Index</i>	11382,08	11226,56
	<i>BIC-Index</i>	11979,41	11978,17
e-Items	<i>AIC-Index</i>	15035,70	15061,01
	<i>BIC-Index</i>	15633,04	15812,62
m-Items	<i>AIC-Index</i>	16876,47	16885,26
	<i>BIC-Index</i>	17473,80	17636,86
Soz. Kompetenz (SK) Sensitivität Kontaktfähigkeit	Kriterium	A	B
		1-dimensional	5-dimensional

Soziabilität Teamorientierung Durchsetzungstärke		d-/e-/m-Dimension	Fragebogen- Skalen
d-Items	<i>AIC-Index</i>	14823,30	13829,11
	<i>BIC-Index</i>	15639,59	14996,08
e-Items	<i>AIC-Index</i>	21640,93	21797,09
	<i>BIC-Index</i>	22511,21	22964,06
m-Items	<i>AIC-Index</i>	26461,19	27293,06
	<i>BIC-Index</i>	27331,47	28460,03
NEO-FFI (NEO)		A	B
Neurotizismus		1-dimensional	5-dimensional
Extraversion			
Offenheit			
Verträglichkeit			
Gewissenhaftigkeit			
d-Items	<i>AIC-Index</i>	15046,38	13770,59
	<i>BIC-Index</i>	15821,72	14842,63
e-Items	<i>AIC-Index</i>	22099,19	22219,49
	<i>BIC-Index</i>	22874,54	23291,53
m-Items	<i>AIC-Index</i>	20751,18	21566,46
	<i>BIC-Index</i>	21526,52	22638,49

Die angegebenen Ergebnisse beziehen sich auf die mit **I bis III** gekennzeichneten Modellvergleiche aus Tabelle 22. Die grauen Markierungen kennzeichnen die vergleichsweise geringsten AIC/BIC -Werte der berechneten Modelle innerhalb eines Bereichs.

Wie der grauen Markierung (Tabelle 23) zu entnehmen ist, zeigt sich für die BIP-Bereiche BO, PK, SK sowie den NEO-FFI ein einheitliches Ergebnis: während hinsichtlich des d-Items das 3-dimensionale, auf den Fragebogenskalen basierende Modell B die geringeren AIC- und BIC-Werte aufweist, zeigen sich für die e- sowie m-Items geringere AIC- und BIC-Werte für das eindimensionale Modell A. Lediglich der Bereich Arbeitsverhalten (AV) des BIP weist hinsichtlich der d- als auch e-Items geringere AIC- und BIC-Werte für das 3-dimensionale Modell (B) auf, während die AIC- und BIC-Werte auf Basis des Pseudo-Item m, erneut die geringeren Werte für das 1-dimensionale Modell A aufzeigen.

In der folgenden Tabelle finden sich die geschätzten IRT- basierten Mittelwerte, Standardabweichungen sowie marginalen IRT-basierten Reliabilitäten jener Modelle, welche die höchste Passungsgüte gemäß Tabelle 23 aufweisen, getrennt für die einzelnen Versuchsgruppen 1, 2 und 3.

Tabelle 24: Geschätzte IRT-basierte Mittelwerte, Standardabweichungen und die marginalen Reliabilitäten für Modelle A und B (I-III)

Bereich	Item	Dimension	Modell A			Modell B			
			1-dimensional			3-dimensional			
			VG 1	VG 2	VG 3	VG 1	VG 2	VG 3	
Arbeits- verhalten (AV)	d-Items	Flexibilität				-1,569	-0,376	-0,144	
						1,58	1,681	1,82	
						0,75	0,76	0,79	
		Gewissenhaftigkeit				-1,003	-0,846	-0,450	
						2,12	2,09	2,28	
						0,69	0,75	0,82	
		Handlungsoptimierung				-0,895	-1,073	-0,271	
						1,91	2,08	2,02	
						0,69	0,73	0,76	
	e-Items	Flexibilität				-1,364	-0,126	-0,648	
						1,18	1,50	1,51	
						0,68	0,74	0,73	
		Gewissenhaftigkeit				-0,186	-0,224	-0,168	
						1,37	1,59	1,55	
						0,76	0,80	0,80	
Handlungsoptimierung					-0,072	-0,207	-0,224		
					1,05	1,54	1,53		
					0,70	0,82	0,79		
m-Items	m-Dimension		-1,578	-0,295	-0,138				
			0,78	0,86	0,72				
			0,78	0,76	0,72				
		Dimension	Modell A			Modell B			
			1-dimensional			3-dimensional			
Berufl. Orien- tierung (BO)	d-Items	Führungsmotivation				-1,720	-1,095	-0,231	
						1,75	2,07	2,09	
						0,77	0,79	0,82	
		Gestaltungsmotivation				-0,240	-0,307	-0,125	
						1,60	1,70	1,40	
						0,68	0,73	0,68	
		Leistungsmotivation				-0,406	-0,820	-0,206	
						1,82	2,10	1,82	
						0,71	0,71	0,73	
	e-Items	e-Dimension		-1,081	-0,331	-0,328			
				1,08	1,15	1,01			
				0,86	0,85	0,84			
	m-Items	m-Dimension		-1,282	-0,115	-0,166			
				0,71	0,63	0,86			
				0,76	0,70	0,80			
		Dimension	Modell A			Modell B			
			1-dimensional			3-dimensional			
Psych. Konstitution (PK)	d-Items	Belastbarkeit				-2,836	-1,430	-0,650	
						1,79	2,37	2,37	
						0,73	0,73	0,76	
		Emot. Stabilität				-0,447	-0,275	-0,064	
						1,77	1,99	1,61	
						0,77	0,83	0,80	
		Selbstbewusstsein				-0,433	-0,339	-0,091	
						1,60	1,82	1,87	
						0,78	0,82	0,85	

		-1,179	-0,254	-0,320		
	e-Items	<i>e-Dimension</i>	1,05	1,09	1,02	
			0,85	0,85	0,86	
	m-Items	<i>m-Dimension</i>	-1,532	-0,147	-0,199	
			0,93	0,71	0,88	
			0,84	0,74	0,80	
		Dimension	Modell A	Modell B		
			1-dimensional	5-dimensional		
Soziale Kompetenz (SK) MW Std.Abw. Reliabilität		<i>Durchsetzungstärke</i>		-2,388	-1,328	-1,035
				2,06	2,59	2,52
				0,71	0,78	0,77
		<i>Kontaktfähigkeit</i>		-0,929	-0,604	-0,851
				2,27	2,37	2,23
				0,80	0,81	0,81
	d-Items	<i>Sensitivität</i>		-0,760	-0,808	-0,568
				2,53	2,84	2,32
				0,57	0,55	0,61
		<i>Soziabilität</i>		-0,485	-0,193	-0,398
				1,74	1,83	1,80
				0,67	0,68	0,74
		<i>Teamorientierung</i>		-0,497	-0,757	-0,090
				2,93	2,65	2,95
				0,78	0,75	0,83
	e-Items	<i>e-Dimension</i>	-1,187	-0,337	-0,286	
			1,08	1,06	1,14	
			0,90	0,88	0,90	
	m-Items	<i>m-Dimension</i>	-1,331	-0,153	-0,167	
			0,71	0,83	0,71	
			0,86	0,86	0,83	
		Dimension	Modell A	Modell B		
			1-dimensional	5-dimensional		
NEO-FFI (NEO) MW Std.Abw. Reliabilität		<i>Extraversion</i>		-2,375	-0,161	-0,264
				1,77	2,03	1,76
				0,69	0,68	0,67
		<i>Gewissenhaftigkeit</i>		-0,165	-0,921	-0,420
				2,18	2,11	1,54
				0,71	0,64	0,62
	d-Items	<i>Neurotizismus</i>		-1,268	-2,156	-0,882
				3,18	3,13	3,26
				0,52	0,53	0,68
		<i>Offenheit</i>		-0,927	-0,722	-0,172
				2,67	2,73	2,14
				0,67	0,72	0,76
		<i>Verträglichkeit</i>		-0,029	-0,222	-0,113
				2,00	1,98	1,86
				0,75	0,73	0,75
	e-Items	<i>e-Dimension</i>	-0,602	-0,127	-0,219	
			1,00	1,11	0,97	
			0,90	0,91	0,89	
	m-Items	<i>m-Dimension</i>	-1,666	-0,194	-0,175	
			0,75	0,85	0,67	
			0,83	0,83	0,75	

Hinsichtlich der geschätzten Mittelwerte (Tabelle 24) fällt auf, dass die geschätzten Mittelwerte der d-Items innerhalb des Modells B, in nahezu allen Fällen die geringste Ausprägung für VG 1 oder 2 aufweisen. Hinsichtlich der e-Items bzw. e-Dimension (Modell A) in den Bereichen BO, SK, PK und NEO zeigen sich die jeweils kleinsten geschätzten Mittelwerte für die VG 1. Die Mittelwerts-Differenzen der VG 1 zu den geschätzten Mittelwerten der VG 2 und 3 sind teilweise hoch ausgeprägt, während die Differenz der Mittelwerte zwischen VG 2 und 3 sehr gering ausfällt. Hinsichtlich der m-Items bzw. m-Dimension (Modell A) zeigen sich die kleinsten Mittelwerte in allen Fällen für VG 1. Die marginalen Reliabilitätswerte der Modelle A bewegen sich konstant oberhalb von 0,70. Die Reliabilitätswerte der Modelle B liegen insbesondere in den Bereichen SK (BIP) und NEO (NEO-FFI) teilweise darunter, mit Werten unterhalb von 0,60.

Modellvergleich IV für GSP

Wie aus Tabelle 25 ersichtlich, liegen die AIC- und BIC-Werte in den 5- bzw. 7-dimensionalen Modellen (Modell C, s. graue Markierung) für die Bereiche AV, SK und NEO jeweils unter jenen der Modelle A oder B. Für die Bereiche BO und PK zeigen sich hingegen die geringsten AIC- und BIC-Werte für die 3-dimensionalen Modelle A, welche eine Modellstruktur auf Basis der Pseudo-Item-Dimensionen (d, e und m) repräsentieren.

Tabelle 25: Modellvergleiche IV (GSP) unter Einbeziehung aller Skalen und Pseudo-Items

Dimensionen	Kriterium	A	B	C
		3-dimensional d-, e-, m- Dimension	3-dimensional Fragebogen- Skalen	5-dimensional Fragebogenskalen, e-, m-Dimension
Arbeitsverhalten (AV) Flexibilität Gewissenhaftigkeit Handlungsoptimierung	<i>AIC-Index</i>	40885,32	43561,19	40157,26
	<i>BIC-Index</i>	42589,17	45265,04	42003,42
Beruf. Orientierung (BO) Führungsmotivation Gestaltungsmotivation Leistungsmotivation	<i>AIC-Index</i>	40121,21	43216,27	40605,79
	<i>BIC-Index</i>	41789,48	44884,54	42416,38
Psych. Konstitution (PK) Belastbarkeit Emotionale Stabilität Selbstbewusstsein	<i>AIC-Index</i>	43158,42	47131,93	43613,43
	<i>BIC-Index</i>	44696,00	48942,52	45566,33
Dimensionen	Kriterium	A 3-dimensional	B 5-dimensional	C 7-dimensional

		d-, e-, m-Dimension	Fragebogen-Skalen	Fragebogenskalen, e-, m-Dimension
Soz. Kompetenz(SK)				
Durchsetzungstärke	<i>AIC-Index</i>	62771,34	68485,63	62303,95
Kontaktfähigkeit				
Sensitivität				
Soziabilität	<i>BIC-Index</i>	65400,25	71256,86	65134,47
Teamorientierung				
NEO-FFI (NEO)				
Extraversion	<i>AIC-Index</i>	57867,29	60792,89	56907,00
Gewissenhaftigkeit				
Neurotizismus				
Offenheit	<i>BIC-Index</i>	60211,56	63279,48	59452,89
Verträglichkeit				

Die angegebenen Ergebnisse beziehen sich auf die mit **IV** gekennzeichneten Modellvergleiche aus Tabelle 22. Die grauen Markierungen kennzeichnen die vergleichsweise geringsten AIC/BIC/-Werte der berechneten Modelle innerhalb eines Bereichs.

In der folgenden Tabelle finden sich die geschätzten IRT-basierten Mittelwerte, Standardabweichungen sowie marginalen IRT-basierten Reliabilitäten jener Modelle, welche die höchste Passungsgüte gemäß Tabelle 25 aufweisen, getrennt für die einzelnen Versuchsgruppen 1, 2 und 3.

Tabelle 26: Geschätzte Mittelwerte, Standardabweichungen und die marginalen Reliabilitäten für Modelle A und C (IV)

Bereich	Dimension	A			C		
		3-dimensional			5-dimensional		
		VG 1	VG 2	VG 3	VG 1	VG 2	VG 3
Arbeits- verhalten (AV)	e-Dimension				0,050	-0,122	-0,239
					0,93	1,23	1,18
					0,83	0,89	0,88
	m-Dimension				-0,413	-0,096	0,052
					0,77	0,88	0,79
					0,77	0,79	0,77
	Flexibilität				-1,729	-0,793	-0,162
					1,73	1,74	1,94
					0,74	0,75	0,80
	Gewissenhaftigkeit				-0,901	-0,267	-0,264
					2,56	2,58	2,62
					0,74	0,75	0,81
	Handlungsoptimierung				-0,474	-0,685	0,268
					2,36	2,52	2,33
					0,71	0,73	0,77
Berufl. Orientierung (BO)	d-Dimension	-1,566	-0,433	-0,084			
		1,25	1,45	1,37			
		0,84	0,86	0,86			
	e-Dimension	-0,242	-0,223	-0,378			
		1,16	1,24	1,15			
		0,84	0,86	0,84			
	m-Dimension	-0,191	-0,094	-0,087			
		0,80	0,75	0,90			
		0,78	0,75	0,81			
	Psych. Konstitution (PK)	d-Dimension	-1,386	-0,408	0,005		
1,31			1,57	1,41			
0,86			0,90	0,90			

MW Std.Abw. Reliabilität	e-Dimension	-0,282	-0,228	-0,457	
		1,03	1,09	1,06	
	m-Dimension	0,86	0,88	0,86	
		0,377	0,101	0,048	
		0,88	0,70	0,85	
		0,83	0,77	0,82	
		A	C		
		3-dimensional	7-dimensional		
Soz. Kompetenz(SK)	e-Dimension		0,140	-0,791	-0,102
			1,24	1,14	1,33
			0,93	0,91	0,92
	m-Dimension		-0,203	0,238	-0,312
			0,81	0,89	0,85
			0,83	0,88	0,88
	Durchsetzungsstärke		-2,342	-1,228	-0,138
			2,14	2,65	2,49
			0,73	0,80	0,78
	Kontaktfähigkeit		-0,932	-0,927	-1,301
			2,36	2,39	2,23
			0,80	0,81	0,83
	Sensitivität		-1,015	-0,945	-0,786
			2,44	2,68	2,31
			0,58	0,55	0,62
	Soziabilität		-0,553	-0,319	-0,202
			2,02	2,08	2,02
			0,70	0,69	0,72
	Teamorientierung		-0,565	-0,958	-0,332
			2,84	2,56	2,80
			0,80	0,78	0,85
NEO-FFI (NEO)	e-Dimension		0,622	0,195	-0,328
			0,92	1,02	0,90
			0,90	0,92	0,88
	m-Dimension		-0,596	0,380	0,217
			0,63	0,75	0,70
			0,76	0,81	0,78
	Extraversion		-2,310	-0,949	-0,248
			1,99	2,21	1,58
			0,71	0,65	0,62
	Gewissenhaftigkeit		-0,785	-2,463	-0,385
			3,81	3,19	3,67
			0,56	0,53	0,67
	Neurotizismus		1,426	0,403	-0,161
			2,58	2,93	2,235
			0,69	0,73	0,76
	Offenheit		0,507	0,239	0,538
			2,07	2,08	1,95
			0,76	0,74	0,76
	Verträglichkeit		-0,476	0,019	-0,147
			2,00	2,12	1,78
			0,71	0,69	0,67

Bei Betrachtung von Tabelle 26 zeigt sich, dass die auf dem e-Item basierenden Dimensionen der Modelle A (für BO und PK) sowie C (für AV, SK und NEO) nahezu durchgängig die höchsten geschätzten Mittelwerte für VG 1 und 2 im Vergleich mit VG 3 aufweisen (Ausnahme SK: höchster MW für VG 1 und 3). Die

geschätzten Mittelwerte auf Basis der m-Items für die Bereiche BO und PK (Modell A) sowie AV und SK (Modell C) gestalten sich in abfallender Reihenfolge von VG 3 zu VG 2 und VG 1. Die Verwendung der d-Items zur Erhebung der Fragebogenskalen in Modell C zeigt in 10 von den abgebildeten 13 Fragebogen-Dimensionen (aus AV, PK und NEO) den höchsten Mittelwert für die VG 3 im Vergleich zu VG 1 und 2 (Ausnahme: SK Kontaktfähigkeit, NEO Neurotizismus, NEO Verträglichkeit).

Hinsichtlich der marginalen IRT-basierten Reliabilitäten zeigen sich in nahezu allen Fällen Reliabilitätswerte grösser als 0,70. Lediglich für die mittels Pseudo-Item d erhobenen Dimensionen Sensitivität aus dem BIP (0,55-0,62) sowie Extraversion (0,62-0,65), Gewissenhaftigkeit (0,53-0,67), Neurotizismus (0,69) und Verträglichkeit (0,67-0,69) aus dem NEO-FFI bewegen sich die Reliabilitätswerte unter 0,70.

b) Ergebnisse zur Modellpassung auf Basis der VG 1, 2 und 3

Im folgenden finden sich nun die Ergebnisse der Modellvergleiche I bis III (Tabelle 22). Grundsätzlich kann man von einer Entsprechung der Ergebnisse auf Basis der Gesamtstichprobe (s. Tabelle 23), mit jenen der Versuchsgruppen ausgehen. Es stellen sich jedoch einige Abweichungen dar, die im Weiteren näher erläutert werden.

Modellvergleiche I-III für VG 1, 2 und 3

Tabelle 27 stellt die Ergebnisse zum Bereich Arbeitsverhalten (AV) des BIP dar.

Tabelle 27: Modellvergleiche I-III (AV) unter Betrachtung der einzelnen Pseudo-Items

Arbeitsverhalten (AV)		A	B
Gewissenhaftigkeit	Kriterium	1-dimensional	3-dimensional
Flexibilität		d-/e-/m-Dimension	Fragebogen-
Handlungsoptimierung			Skalen
VG 1			
d-Items	AIC-Index	3367,55	3039,0
	BIC-Index	3498,74	3207,27
e-Items	AIC-Index	34,19	60,50
	BIC-Index	0,19	0,50
m-Items	AIC-Index	5047,72	5095,35
	BIC-Index	5178,92	5263,62
VG 2			
d-Items	AIC-Index	3403,65	3157,45
	BIC-Index	3534,48	3325,26

<u>e-Items</u>	<i>AIC-Index</i>	4938,87	4870,02
	<i>BIC-Index</i>	5069,70	5037,83
nach Ausschluss von 6 von 42 Items	<i>AIC-Index</i>	4199,05	4209,51
	<i>BIC-Index</i>	4312,82	4360,25
m-Items	<i>AIC-Index</i>	4794,25	4803,21
	<i>BIC-Index</i>	4925,08	4971,02
VG 3			
d-Items	<i>AIC-Index</i>	4256,78	3708,46
	<i>BIC-Index</i>	4389,74	3878,99
<u>e-Items</u>	<i>AIC-Index</i>	5025,69	4980,48
	<i>BIC-Index</i>	5158,65	5151,01
nach Ausschluss von 11 von 42 Items	<i>AIC-Index</i>	3654,68	3658,97
	<i>BIC-Index</i>	3755,84	3797,71
m-Items	<i>AIC-Index</i>	5353,98	5356,92
	<i>BIC-Index</i>	5486,94	5527,45

Die unterstrichenen Zellinhalte bezeichnen solche Modellvergleiche, deren Passung entgegen der Erwartungen lag. Die grauen Markierungen kennzeichnen die vergleichsweise geringsten AIC/BIC-Werte der berechneten Modelle innerhalb eines Bereichs.

Für die d-Items des Bereich AV (Tabelle 27) zeigt sich die 3-dimensionale Repräsentation (Modell B) mittels eigentlicher Fragebogenskalen in allen drei Versuchsgruppen als jene mit den geringsten AIC- und BIC-Werten im Vergleich zu Modell A, während die Repräsentation der m-Items mittels eindimensionalen Modells A (m-Dimension, MRS) geringere AIC- und BIC-Werte aufweist, als dies für Modell B der Fall ist. Hinsichtlich der e-Items (e-Dimension, ERS) konnten für VG 2 und VG 3 erst nach Ausschluss¹⁶ von sechs (VG 2) bzw. 11 (VG 3) Items geringere AIC- und BIC-Werte des 1-dimensionalen Modells A im Vergleich zu Modell B erzielt werden (unterstrichen). Dies entspricht 11% (VG 2) bzw. 26% (VG 3) der Items.

Die in Tabelle 28 dargestellten AIC- und BIC-Werte betreffen den Bereich BO (BIP).

Tabelle 28: Modellvergleiche I-III des Bereichs BO (BIP) unter Betrachtung der einzelnen Pseudo-Items

Berufl. Orientierung (BO)		A	B
Leistungsmotivation	Kriterium	1-dimensional	3-dimensional
Gestaltungsmotivation		d-/e-/m-Dimension	Fragebogen-Skalen
Führungsmotivation			

¹⁶Der Ausschluss der Items geschieht hierbei gemäß Analyse des RMSEA-Wertes (Root Mean Square Error Approximation, dtsh.: Approximationsdiskrepanz-Wurzel); Erläuterung s. Fussnote 17

VERSUCHSGRUPPE 1			
d-Items	<i>AIC-Index</i>	3263,58	3251,39
	<i>BIC-Index</i>	3391,92	3416,81
nach Ausschluss von 9 von 41 Items	<i>AIC-Index</i>	2517,57	2479,11
	<i>BIC-Index</i>	2620,24	2618,86
e-Items	<i>AIC-Index</i>	4510,70	4582,72
	<i>BIC-Index</i>	4639,04	4748,13
m-Items	<i>AIC-Index</i>	5483,54	5549,54
	<i>BIC-Index</i>	5611,88	5714,96
VERSUCHSGRUPPE 2			
d-Items	<i>AIC-Index</i>	3068,35	3049,22
	<i>BIC-Index</i>	3196,34	3214,18
nach Ausschluss von 1 von 41 Items	<i>AIC-Index</i>	2990,95	2932,23
	<i>BIC-Index</i>	3116,09	3094,35
e-Items	<i>AIC-Index</i>	4398,18	4456,79
	<i>BIC-Index</i>	4526,17	4621,76
m-Items	<i>AIC-Index</i>	5538,58	5598,07
	<i>BIC-Index</i>	5666,56	5763,03
VERSUCHSGRUPPE 3			
d-Items	<i>AIC-Index</i>	3700,61	3683,94
	<i>BIC-Index</i>	3830,68	3851,58
nach Ausschluss von 3 von 41 Items	<i>AIC-Index</i>	3313,26	3274,44
	<i>BIC-Index</i>	3434,66	3433,4
e-Items	<i>AIC-Index</i>	4593,35	4691,60
	<i>BIC-Index</i>	4723,42	4859,24
m-Items	<i>AIC-Index</i>	5742,03	5794,19
	<i>BIC-Index</i>	5872,09	5961,84

Die kursiv geschriebenen Zellinhalte bezeichnen solche Modellvergleiche, im Rahmen derer keine eindeutige Aussage auf Basis des AIC- und BIC-Kriteriums getätigt werden konnte. Die grauen Markierungen kennzeichnen die vergleichsweise geringsten AIC/BIC-Werte der berechneten Modelle innerhalb eines Bereichs.

Gemäß obiger Tabelle 28 zeigt sich für den Bereich BO (BIP) hinsichtlich der Modelle auf Basis der e- sowie m-Items für alle drei Versuchsgruppen, dass hier die 1-dimensionalen auf ERS sowie MRS basierenden Modelle A die geringeren AIC- und BIC-Werte aufweisen. Für die d-Items lässt sich je nach Betrachtung des AIC- oder BIC-Wertes entweder Modell A oder B als jenes mit den geringeren Werten ausmachen. Nach Ausschluss von neun (VG 1, 22%), einem (VG 2, 2%) und drei (VG 3, 7%) von 41 Item(s) auf Basis des RMSEA-Wertes (s. Fussnote 17, S.110) zeigt sich jedoch für alle drei Versuchsgruppen ein geringerer AIC- und BIC-Wert für das 3-dimensionale, die Fragebogen-Skalen repräsentierende Modell B.

In der folgenden Tabelle 29 sind die Ergebnisse zum Bereich PK des BIP dargestellt.

Tabelle 29: Modellvergleiche I-III des Bereichs PK (BIP) unter Betrachtung der einzelnen Pseudo-Items

Psych. Konstitution (PK) Emotionale Stabilität Belastbarkeit Selbstbewusstsein	Kriterium	A 1-dimensional d-/e-/m-Dimension	B 3-dimensional Fragebogen- Skalen
VERSUCHSGRUPPE 1			
d-Items	AIC-Index	3655,81	3636,14
	BIC-Index	3795,55	3812,96
nach Ausschluss von 2 von 45 Items	AIC-Index	3364,18	3317,62
	BIC-Index	3495,37	3485,89
e-Items	AIC-Index	4949,86	4973,84
	BIC-Index	5089,61	5150,67
m-Items	AIC-Index	5474,23	5524,12
	BIC-Index	5613,98	5700,95
VERSUCHSGRUPPE 2			
d-Items	AIC-Index	3526,90	3489,83
	BIC-Index	3666,27	3666,17
e-Items	AIC-Index	5033,71	5022,81
	BIC-Index	5173,07	5199,16
nach Ausschluss von 4 von 45 Items	AIC-Index	4477,23	4477,31
	BIC-Index	4605,22	4642,27
m-Items	AIC-Index	5587,87	5572,93
	BIC-Index	5727,23	5749,27
nach Ausschluss von 6 von 45 Items	AIC-Index	4869,73	4870,52
	BIC-Index	4992,03	5029,79
VERSUCHSGRUPPE 3			
d-Items	AIC-Index	4245,69	4177,99
	BIC-Index	4387,32	4357,2
e-Items	AIC-Index	5146,94	5158,93
	BIC-Index	5288,56	5338,13
m-Items	AIC-Index	5838,52	5875,99
	BIC-Index	5980,15	6037,19

Die kursiv geschriebenen Zellinhalte bezeichnen solche Modellvergleiche, im Rahmen derer keine eindeutige Aussage auf Basis des AIC- und BIC-Kriteriums getätigt werden konnte. Die grauen Markierungen kennzeichnen die vergleichsweise geringsten AIC/BIC-Werte der berechneten Modelle innerhalb eines Bereichs.

In Tabelle 29 (Bereich PK, BIP) zeigen sich für die d-Items sowohl für die VG 1 nach Ausschluss zweier Items (4%) auf Basis des RMSEA-Wertes, als auch für die VG 2 und 3 die geringeren AIC- und BIC-Werte für das dreidimensionale Modell B, welches auf den eigentlichen Fragebogenskalen basiert. Für die e- und m-Items stellt sich für die VG 1 und 3 das 1-dimensionale Modell A als jenes, mit

den geringeren AIC- und BIC-Werten dar. Für die VG 2 zeigt sich das 1-dimensionale Modell A nach Ausschluss von vier (9%) e-Items bzw. sechs (13%) m-Items ebenfalls als jenes mit den geringeren AIC- und BIC-Werten.

Tabelle 30 dient der Darstellung der AIC- und BIC-Werte für den Bereich Soziale Kompetenz (SK) des BIP.

Tabelle 30: Modellvergleiche I-III des Bereichs SK (BIP) unter Betrachtung der einzelnen Pseudo-Items

Soz. Kompetenz(SK)		A	B
Kriterium		1-dimensional	5-dimensional
		d-/e-/m-Dimension	Fragebogen-Skalen
VERSUCHSGRUPPE 1			
d-Items	<i>AIC-Index</i>	4709,54	4481,87
	<i>BIC-Index</i>	4912,04	4755,67
e-Items	<i>AIC-Index</i>	7157,08	7252,34
	<i>BIC-Index</i>	7362,42	7528,99
m-Items	<i>AIC-Index</i>	8738,23	8967,88
	<i>BIC-Index</i>	8943,58	9244,53
VERSUCHSGRUPPE 2			
d-Items	<i>AIC-Index</i>	4677,28	4325,49
	<i>BIC-Index</i>	4882,06	4601,38
e-Items	<i>AIC-Index</i>	7048,88	7147,78
	<i>BIC-Index</i>	7253,66	7423,66
m-Items	<i>AIC-Index</i>	8725,87	8946,10
	<i>BIC-Index</i>	8930,65	9221,99
VERSUCHSGRUPPE 3			
d-Items	<i>AIC-Index</i>	5512,02	5065,23
	<i>BIC-Index</i>	5720,12	5345,59
e-Items	<i>AIC-Index</i>	7590,18	7760,19
	<i>BIC-Index</i>	7798,29	8040,55
m-Items	<i>AIC-Index</i>	9081,59	9167,52
	<i>BIC-Index</i>	9289,69	9447,88

Die grauen Markierungen kennzeichnen die vergleichsweise geringsten AIC/BIC-Werte der berechneten Modelle innerhalb eines Bereichs.

Für den Bereich Soziale Kompetenz (SK, Tabelle 30) zeigen sich hinsichtlich der d-Items für alle Versuchsgruppen 1, 2 und 3 geringere AIC- und BIC-Werte für das dreidimensionale Modell B (Fragebogenskalen) als es für Modell A der Fall ist. Für e- als auch m-Items zeigen sich in VG 1, 2 und 3 geringere AIC- und BIC-Werte für das die jeweilige Antworttendenz (ERS für e-Items und MRS für m-Items) repräsentierende Modell A im Vergleich zu Modell B.

Tabelle 31 enthält die Ergebnisse der AIC- und BIC-Werte zum NEO-FFI.

Tabelle 31: Modellvergleiche I-III auf Basis des NEO-FFI (NEO) unter Betrachtung der einzelnen Pseudo-Items

NEO-FFI (NEO) Neurotizismus Extraversion Offenheit Verträglichkeit Gewissenhaftigkeit	Kriterium	A 1-dimensional d-/e-/m-Dimension	B 5-dimensional Fragebogen- Skalen
VERSUCHSGRUPPE 1			
d-Items	AIC-Index	4735,04	4346,39
	BIC-Index	4911,87	4594,53
e-Items	AIC-Index	7327,61	7563,22
	BIC-Index	7510,14	7817,05
m-Items	AIC-Index	6814,34	7074,06
	BIC-Index	6996,87	7074,06
VERSUCHSGRUPPE 2			
d-Items	AIC-Index	4590,91	4157,74
	BIC-Index	4770,09	4408,03
e-Items	AIC-Index	7193,04	7220,28
	BIC-Index	7375,07	7473,42
m-Items	AIC-Index	6815,34	6931,72
	BIC-Index	6997,37	7184,86
VERSUCHSGRUPPE 3			
d-Items	AIC-Index	5776,45	5306,73
	BIC-Index	5961,44	5563,97
e-Items	AIC-Index	7697,04	7632,25
	BIC-Index	7882,02	7889,49
nach Ausschluss von 7 von 60 Items	AIC-Index	6765,59	6768,65
	BIC-Index	6930,35	7005,66
m-Items	AIC-Index	7178,21	7290,65
	BIC-Index	7363,19	7547,89

Die kursiv geschriebenen Zellinhalte bezeichnen solche Modellvergleiche, im Rahmen derer keine eindeutige Aussage auf Basis des AIC- und BIC-Kriteriums getätigt werden konnte. Die grauen Markierungen kennzeichnen die vergleichsweise geringsten AIC/BIC-Werte der berechneten Modelle innerhalb eines Bereichs.

Hier (Tabelle 31) zeigen sich für die d-Items der VG 1, 2 und 3 geringere AIC- und BIC- Werte für das 3-dimensionale Modell B auf Basis der eigentlichen Fragebogenskalen als für die 1-dimensionalen Modelle A. Hinsichtlich der e-Items weist das 1-dimensionale Modell B für VG 1 und 2 im Vergleich zu Modell A die geringeren AIC- und BIC-Werte auf. Um eine eindeutige Aussage zu den e-Items hinsichtlich VG 3 auf Basis der AIC- und BIC-Werte zu tätigen, müssen 7 von 60 e-Items (12%) ausgeschlossen werden. Sodann zeigt auch hier das 1-

dimensionale Modell A die im Vergleich zu B geringeren AIC- und BIC-Werte. Für das m-Item zeigen sich in allen drei Versuchsgruppen geringere AIC- und BIC-Werte für das die Tendenz zur mittleren Antwortkategorie (MRS) repräsentierende Modell A im Vergleich zu Modell B.

c) Ergebnisse zur Messgüte des Pseudo-Item d auf Basis der GSP

Tabelle 32 und Tabelle 33 enthalten die geschätzten Skalen-Interkorrelationen (paarweise Korrelationen der Skalen) aus dem IRT-Modellvergleich I (Tabelle 22) für die Gesamtstichprobe zum Vergleich der Messgüte von Pseudo-Items d und den Original-Items aus BIP und NEO-FFI auf Basis der Daten der Gesamtstichprobe. Der Untersuchung der Korrelationswerte liegt die Annahme zugrunde, dass ein höherer Korrelationskoeffizient eine stärkere inhaltliche Überlappung der erhobenen Dimensionen darstellt. Ziel eines Messinstruments ist es stets, so wenig Überlappung wie möglich zwischen den einzelnen Dimensionen zu erhalten bzw. möglichst unabhängige Dimensionen zu erheben. D.h. je geringer der jeweilige Korrelationskoeffizient, desto eher ist die Erfassung der entsprechenden Dimensionen als separate Konstrukte gerechtfertigt. Bezogen auf den vorliegenden Fall bedeutet dies, dass die ein oder andere Kodierungsmethode (Pseudo-Item d vs. Original-Items bzw. Likert-Skala) eine schärfere inhaltliche Trennung bei der Erhebung der zugrundeliegenden Konstrukte erlaubt. Die jeweiligen Korrelationskoeffizienten finden sich in Tabelle 32, wobei die grau markierten Werte jene Werte darstellen die im direkten Vergleich zwischen d- und Original-Items den geringeren Korrelationswert und folglich die höhere Messgüte (i.S. einer schärferen Trennung der Skalen) aufweisen.

Tabelle 32: Geschätzte Skaleninterkorrelationen auf Basis der d- und Original-Items für die Bereiche AV, BO und PK (BIP) aus Modellvergleich I (GSP)

d-Items			
Arbeitsverhalten (AV)	Flexibilität	Gewissenhaftigkeit	Handlungsoptimierung
Flexibilität	1,000		
Gewissenhaftigkeit	-0,164	1,000	
Handlungsoptimierung	0,164	0,470	1,000
Berufli. Orientierung (BO)	Führungsmotivation	Gestaltungsmotivation	Leistungsmotivation
Führungsmotivation	1,000		
Gestaltungsmotivation	0,602	1,000	
Leistungsmotivation	0,442	0,448	1,000

Psych. Konstitution (PK)	Belastbarkeit	Emotionale Stabilität	Selbstbewusstsein
Belastbarkeit	1,000		
Emotionale Stabilität	0,587	1,000	
Selbstbewusstsein	0,509	0,608	1,000

Original-Items (BIP)			
Arbeitsverhalten (AV)	Flexibilität	Gewissenhaftigkeit	Handlungsoptimierung
Flexibilität	1,000		
Gewissenhaftigkeit	0,159	1,000	
Handlungsoptimierung	0,403	0,585	1,000
Berufli. Orientierung (BO)	Führungsmotivation	Gestaltungsmotivation	Leistungsmotivation
Führungsmotivation	1,000		
Gestaltungsmotivation	0,620	1,000	
Leistungsmotivation	0,585	0,596	1,000
Psych. Konstitution (PK)	Belastbarkeit	Emotionale Stabilität	Selbstbewusstsein
Belastbarkeit	1,000		
Emotionale Stabilität	0,676	1,000	
Selbstbewusstsein	0,548	0,589	1,000

Die grau markierten Werte kennzeichnen jene Korrelationskoeffizient-Beträge, die im Rahmen des Vergleichs der Interkorrelationen zwischen d- und Original-Items geringer ausfallen.

Bei Betrachtung der Korrelationen des Bereichs Arbeitsverhalten (AV) sind die Korrelationen in zwei von drei Fällen für das Pseudo-Item d geringer (gemäß grauer Markierung; d- vs. Original-Item: 0,164 vs. 0,403 und 0,470 vs. 0,585) als für die Original-Items. Der Betrag des dritten Wertes (Korrelationen zwischen Flexibilität und Gewissenhaftigkeit) unterscheidet sich lediglich in sehr geringem Ausmaß (-0,164 für Pseudo-Item d und 0,159 für das Original-Item).

Der Vergleich von d- und Original-Items innerhalb des Bereichs Berufliche Orientierung (BO) weist für die Korrelationskoeffizienten der mittels Pseudo-Item d erhobenen Dimensionen in allen Fällen geringere Beträge auf (d- vs. Original-Items: 0,602 vs. 0,620, 0,442 vs. 0,585, 0,448 vs. 0,596).

Bei Betrachtung des Bereichs Psychische Konstitution (PK) ergeben sich in zwei von drei Fällen geringere Korrelationskoeffizienten für die Dimensionen deren Messung mittels Pseudo-Item d vorgenommen wurde (d- vs. Original-Item: 0,587 vs. 0,676 und 0,509 vs. 0,548). Der dritte Vergleich fällt zugunsten der Original-Items aus (0,608 vs. 0,589).

Im Bereich Soziale Kompetenz (SK) (Tabelle 33) des BIP zeigen alle der auf Basis des Pseudo-Item d ermittelten Korrelationskoeffizienten (s. graue Markierung der Korrelationskoeffizienten der d-Items, Bereich SK) einen

niedrigeren Betrag, als es für die auf Basis der Original-Items ermittelten Werte der Fall ist. Die Korrelationskoeffizienten der Dimensionen des NEO-FFI weisen in acht von zehn Fällen auf Basis der d-Items einen geringeren Betrag auf, als die entsprechenden Beträge auf Basis der Original-Items.

Tabelle 33: Geschätzte Skaleninterkorrelationen auf Basis der d- und Original-Items für den Bereich SK (BIP) und NEO (NEO-FFI) aus Modellvergleich I (GSP)

d-Items					
Soziale Kompetenz (SK)	Durchsetzungsstärke	Kontaktfähigkeit	Sensitivität	Soziabilität	Teamorientierung
Durchsetzungsstärke	1,00				
Kontaktfähigkeit	0,373	1,00			
Sensitivität	0,172	0,336	1,00		
Soziabilität	-0,126	0,236	0,325	1,00	
Teamorientierung	0,140	0,348	0,258	0,366	1,00
NEO-FFI (NEO)	Extraversion	Gewissenhaftigkeit	Neurotizismus	Offenheit	Verträglichkeit
Extraversion	1,00				
Gewissenhaftigkeit	0,060	1,00			
Neurotizismus	-0,220	-0,189	1,00		
Offenheit	0,114	-0,066	0,038	1,00	
Verträglichkeit	0,153	0,091	-0,102	0,105	1,00
Original-Items (BIP)					
Soziale Kompetenz (SK)	Durchsetzungsstärke	Kontaktfähigkeit	Sensitivität	Soziabilität	Teamorientierung
Durchsetzungsstärke	1,00				
Kontaktfähigkeit	0,405	1,00			
Sensitivität	0,369	0,516	1,00		
Soziabilität	0,148	0,381	0,516	1,00	
Teamorientierung	0,221	0,398	0,373	0,381	1,00
NEO-FFI (NEO)	Extraversion	Gewissenhaftigkeit	Neurotizismus	Offenheit	Verträglichkeit
Extraversion	1,00				
Gewissenhaftigkeit	0,280	1,00			
Neurotizismus	0,071	0,052	1,00		
Offenheit	0,366	0,223	0,162	1,00	
Verträglichkeit	0,369	0,360	0,130	0,357	1,00

Die grau markierten Werte kennzeichnen jene Korrelationskoeffizient-Beträge, die im Rahmen des Vergleichs der Interkorrelationen zwischen d- und Original-Items geringer ausfallen.

Die beobachteten bzw. empirischen (im Gegensatz zu den hier angeführten geschätzten) Interkorrelationswerte finden sich ergänzend im Anhang ab S.208.

9.2.3 Interpretation der Ergebnisse

Im folgenden findet sich die Interpretation der zuvor dargestellten Ergebnisse. Der Aufbau entspricht jenem, welcher bereits zu Beginn des Abschnitt 9.2.2 dargestellt wurde:

- a) Interpretation der Ergebnisse zur Modellpassung auf Basis der Gesamtstichprobe (Modellvergleiche I-III und IV)**
- b) Interpretation der Ergebnisse zur Modellpassung auf Basis der VG 1, 2 und 3 (Modellvergleiche I-III)**
- c) Interpretation der Ergebnisse zur Ergebnisse zur Messgüte des Pseudo-Item d Basis der GSP**

a) Interpretation der Ergebnisse zur Modellpassung auf Basis der GSP

Modellvergleiche I-III für GSP

Auf Basis der Tatsache, dass die Modellpassungen (Tabelle 23) ausser einer Abweichung (e-Items, Bereich AV) für die restlichen 14 Modellvergleiche innerhalb der Bereiche (AV, BO, SK, PK, NEO) gleich ausfallen zeigt sich, dass aus diesen Ergebnissen klare Aussagen hinsichtlich der Modell-Anpassungsgüte abgeleitet werden können. Für die mittels Pseudo-Item d berechneten Modelle weist das 3-dimensionale bzw. 5-dimensionale Modell B eine bessere Passung auf die Daten aus als das 1-dimensionale Modell, welches die Pseudo-Item-Dimension d repräsentiert. Daraus ist zu schliessen, dass die für Pseudo-Item d erarbeitete Kodierungsvariante die tatsächlichen inhaltsabhängigen Antworten auf die Items erhebt.

Bei Betrachtung der auf Basis des Pseudo-Items e berechneten Modelle zeigt sich, dass in vier der fünf Bereiche (BO, PK, SK und NEO-FFI) von einer besseren Anpassungsgüte des 1-dimensionalen Modells auszugehen ist. Dies deutet auf die Tatsache hin, dass Pseudo-Item e tatsächlich die extreme Antworttendenz (ERS) unabhängig vom Iteminhalt erhebt und die Betrachtung der Daten unter Annahme einer einzelnen Antworttendenz-Dimension (ERS) die Daten besser repräsentiert, als dies für eine Repräsentation mittels ursprünglicher Fragebogenskalen der Fall ist.

Hinsichtlich des abweichenden Ergebnisses der Pseudo-Items *e* in dem Bereich AV (BIP) wurde mittels des Ausschlusses solcher Items die einen geringen Itemfit bzw. hohen RMSEA-Wert ¹⁷ (*Root Mean Square Error Approximation*, Approximationsdiskrepanz-Wurzel) aufweisen untersucht, ob dies zur erhöhten Anpassungsgüte des 1-dimensionalen Modells führen würde. Im vorliegenden Fall lagen keinerlei Items mit einem RMSEA-Wert grösser 0,08 vor. Nichts desto trotz wurde versucht über einen Item-Ausschluss der Items mit den höchsten RMSEA-Werten eine verbesserte Modellpassung des 1-dimensionalen Modells zu erzielen. Dieses Vorhaben wurde nach Ausschluss von 50% (21 Items) der vorhandenen 42 Items abgebrochen, da die erhöhte Anpassungsgüte des 3-dimensionalen Modells bestehen blieb. Daraus ist zu schliessen, dass im Falle der Antworten des Bereichs AV (BIP) und den darin enthaltenen Dimensionen, keine Tendenzen zu extremen Antworten (ERS) vorliegen.

Bei Betrachtung der auf Pseudo-Item *m* basierenden Anpassungsgüte-Kriterien (Tabelle 23) der zu vergleichenden Modelle zeigt sich, dass die erhöhte Anpassungsgüte hinsichtlich aller untersuchten Bereiche (AV, BO, SK, PK, NEO) für das 1-dimensionale Modell angenommen werden kann. Daraus kann geschlossen werden, dass mittels Pseudo-Items *m* eine vom Fragebogeninhalt unabhängige mittlere Antworttendenz (MRS) gemessen werden kann.

Die dargestellten geschätzten Mittelwerte (Tabelle 24) erlauben differenzierte Aussagen zum Antwortverhalten zwischen den drei Versuchsgruppen (VG 1, 2 und 3). Versuchsgruppe 3 weist in nahezu allen Fragebogendimensionen die höchsten geschätzten Mittelwerte basierend auf dem *d*-Item (Modell B) im Vergleich zu VG 1 und 2 auf. Dies entspricht den Befunden aus den SPSS-Analysen (Tabelle 14) zum Verfälschungsverhalten, in denen die VG 3 in der Mehrheit der Fragebogendimensionen die höchsten empirischen Mittelwerte im Vergleich zu VG 1 und 2 aufwies.

Die geschätzten IRT-basierten Mittelwerte der *e*-Dimension (ERS) innerhalb des Modell A zeigen im Vergleich der drei Versuchsgruppen in den Bereichen BO, PK, SK und NEO die geringste Ausprägung für VG 1. Während die Differenzen

¹⁷Der RMSEA-Wert beschreibt die Diskrepanz eines Items bezogen auf die Modellapproximation. Je größer diese Diskrepanz (< 0,05 gute Modellpassung, > 0,08 schlechte Modellpassung nach Reinecke, 2005), desto schwerwiegender der Einfluss des jeweiligen Items auf die Modellpassung. Die Untergrenze (0,00) des RMSEA-Wert wird erreicht, wenn das Modell eine exakte Passung zu den Daten aufweist (Reinecke, 2005).

der Mittelwerte von VG 1 zu VG 2 und 3 grösser ausfallen ($\approx 0,70$), sind die Mittelwertsdifferenzen zwischen VG 2 und 3 deutlich geringer ($<0,10$). Daraus lässt sich schliessen, dass (1.) die Testpersonen der VG 1 das geringste Ausmaß extremen Antwortverhaltens (ERS) zeigen und (2.) dass Testpersonen der VG 2 und 3 ein ähnlich hohes Ausmaß extremen Antwortverhaltens (ERS) an den Tag legten.

Bei Betrachtung der geschätzten Mittelwerte der m-Items (MRS, Modell A) zeigt sich, dass in allen Bereichen (AV, BO, SK, PK, NEO) die geringste Ausprägung für VG 1 vorliegt und die Differenz der Mittelwerte zwischen VG 1 zu VG 2 und 3 wesentlich grösser ausfällt ($>1,0$) als die Differenz zwischen VG 2 und 3 ($<0,1$). Daraus ist zu schliessen, dass VG 1 im Gegensatz zu VG 2 und 3 eine wesentlich geringere Tendenz zur mittleren Antwortkategorie (MRS) an den Tag legt.

Die dargestellten marginalen IRT-basierten Reliabilitäten der Faktoren befinden sich mehrheitlich in einem akzeptablen Bereich ($>0,7$). Insbesondere für die eindimensionalen Modelle (A) liegen nahezu ausschliesslich gute Reliabilitätswerte ($0,7-0,8$) vor. Die marginalen Reliabilitäten der 3- und 5-dimensionalen Modelle B liegen tendenziell unterhalb derer der 1-dimensionalen Modelle. Geringere Reliabilitätswerte ($< 0,60$) wurden insbesondere für die d-Items der Bereiche SK (BIP) und den NEO-FFI ermittelt. Insgesamt befinden sich die Werte jedoch in einem akzeptablen Bereich, so dass von einer zufriedenstellenden Messgenauigkeit der Pseudo-Items d, e und m ausgegangen werden kann.

Abschliessend ist zu sagen, dass die Kodierung mittels Pseudo-Items zur Erhebung der inhaltsabhängigen Antworten auf die eigentlichen Fragebogenskalen (d-Items), sowie der inhaltsunabhängigen Antworttendenzen ERS (e-Items) und MRS (m-Items) geeignet ist. Weiters führt sowohl die Tendenz zu extremen (ERS) als auch mittleren (MRS) Antwortkategorie als eigene Dimensionen zur Verzerrung der eigentlichen (Fragebogen-) Dimensionsstruktur. Die geschätzten Mittelwerte zeigen weiters, dass die VG 1 (hohe Faking-Motivation) ein grundsätzlich geringeres Ausmaß extremer und mittlerer Antworttendenzen (ERS und MRS) zeigt, als dies für VG 2 und 3 (unsichere Faking-Motivation und ehrliches Antwortverhalten) der Fall ist,

während VG 2 und 3 ein ähnliches Antwortverhalten hinsichtlich ERS und MRS zeigen.

Modellvergleich IV für GSP

Die erhöhte Modellpassung der 5- bzw. 7-dimensionalen Modelle C für die Bereiche AV, SK und NEO lässt den Rückschluss zu, dass die Repräsentation der Fragebogendaten mittels der inhaltsunabhängigen Pseudo-Dimensionen e und m (als Abbildung von ERS und MRS) und der inhaltsabhängigen Pseudo-Dimension d (zur Erhebung der eigentlichen Fragebogenskalen) gemäß Modell C besser gelingt, als dies für ein ausschliesslich die eigentlichen Fragebogenskalen enthaltendes Modell der Fall ist (Modell B). Diese Ergebnisse bestätigen grösstenteils die Ergebnisse der Modellvergleiche I bis III.

Weiters bestätigt dies, dass für alle fünf Bereiche (AV, BO, SK, PK, NEO) Pseudo-Items e und m die jeweiligen inhaltsunabhängigen Antworttendenzen, nämlich ERS und MRS eindimensional erheben. Die erfolgreiche Verwendung des Pseudo-Item d zur Erhebung der inhaltsabhängigen Item-Antworten gemäß Fragebogenskalen (Modell C) kann jedoch nur für die Bereiche AV, PK und NEO bestätigt werden.

Da die geschätzten IRT-basierten Mittelwerte sowie Standardabweichungen (Tabelle 26) der einzelnen Dimensionen getrennt für die Versuchsgruppen vorliegen, kann ein differenzierteres Bild des Antwortverhaltens zwischen den einzelnen Versuchsgruppen gezeichnet werden. Hinsichtlich der inhaltsabhängigen d-Dimension liegen in 10 von 13 Fragebogendimensionen die höchsten Werte für VG 3 vor. Dieser Umstand korrespondiert mit den empirischen, mittels SPSS erhobenen Mittelwerten (Tabelle 14) insofern, als dass acht von jenen zehn Fragebogendimensionen auch hinsichtlich ihrer empirischen Mittelwerte höhere Ausprägungen in der VG 3 als VG 1 und 2 zeigen. Die Tatsache, dass die Mittelwerte der e-Dimension (e-Items) zur Erhebung einer Tendenz zu Extremurteilen (ERS) in vier der fünf Bereiche (Ausnahme: SK) für VG 1 oder 2 die geringsten geschätzten Mittelwerte aufweisen, deutet an, dass VG 1 und 2 tendenziell eine geringere Tendenz zu Extremurteilen (ERS) an den Tag legten als VG 3, welche in vier der fünf Bereiche den höchsten geschätzten Mittelwert aufweist. Gemeinsam mit der Tatsache, dass VG 3 in 15 von 19 Fragebogendimensionen den höchsten

empirischen Mittelwert (Tabelle 14) im Sinne (stereotypen) Verfälschungsverhalten an den Tag legt, kann als positiver Zusammenhang zwischen (stereotypem) Verfälschungshalten und der Antworttendenz zu Extremurteilen (ERS) gedeutet werden. In vier der fünf Bereiche (AV, BO, PK, SK) zeigen sich die höchsten geschätzten Mittelwerte der m-Dimension (m-Items) zur Erhebung mittlerer Antworttendenzen (MRS) für die VG 3. Gemeinsam mit den Ergebnissen zu den empirischen Mittelwerten der Fragebogendimensionen (Tabelle 14), bei welchen VG 3 in 15 von 19 Fragebogendimensionen die höchste Ausprägung erzielte, könnte somit, wie bereits bei den Ergebnissen zur Soziale-Erwünschtheits-Skala (Abschnitt 9.1), die Verwendung der mittleren Antwortkategorie (MRS) mit (stereotypem) Verfälschungsverhalten einhergehen.

Da die marginalen IRT-basierten Reliabilitäten (Tabelle 26) nahezu aller Dimensionen für die drei Versuchsgruppen oberhalb von 0,70 liegen, kann von einer zufriedenstellenden Messgenauigkeit der Pseudo-Item Dimensionen hinsichtlich einer Tendenz zur Mitte (für m-Items), einer Tendenz zu Extremurteilen (für e-Items) und der d-Items zur inhaltlichen Erhebung der eigentlichen Fragebogendimensionen ausgegangen werden.

Zusammenfassend zeigt sich, dass die Pseudo-Items e und m zur Erhebung der inhaltsunabhängigen Tendenz zu extremem und mittlerem Antwortverhalten (ERS und MRS) geeignet sind und die Darstellung mittels e- und m- Dimension die Daten besser repräsentiert, als die alleinige Repräsentation mittels der eigentlichen Fragebogenskalen. Auf Basis der geschätzten IRT-basierten Mittelwerte in Zusammenhang mit den empirischen Mittelwerten zum (stereotypen) Verfälschungsverhalten zeigen sich positive Zusammenhänge zwischen (stereotypem) Verfälschungsverhalten und sowohl extremer (ERS) als auch mittlerer (MRS) Antworttendenzen.

b) Interpretation der Ergebnisse zur Modellpassung auf Basis der VG 1, 2 und 3

Modellvergleiche I-III für VG 1, 2 und 3

Die Ergebnisse zu den Modellvergleichen I bis III auf Basis der einzelnen Versuchsgruppen 1, 2 und 3 lassen einen den vorhergehenden Analysen entsprechenden Trend feststellen. So zeichnet sich in allen fünf Bereichen (AV,

BO, PK, SK, NEO) die Tendenz ab, dass (1.) für die d-Items das 3- bzw. 5-dimensionale Modell B auf Basis der Fragebogenskalen die bessere Modellpassung aufweist. Dies entspricht der Annahme, dass die Kodierung der d-Items die Erhebung der inhaltsabhängigen Antworten erlaubt. Weiters zeigt sich (2.) dass unter Verwendung der e- bzw. m-Items, die jeweiligen 1-dimensionalen Modelle A im Sinne der Erhebung einer Antworttendenz zu Extremurteilen (e-Items, ERS) bzw. mittleren Antworttendenz (m-Items, MRS) tendenziell die bessere Modellpassung gemäß geringerem AIC- und BIC-Wert zeigen, verglichen mit den Modellen B (zur Erhebung der eigentlichen Fragebogenskalen) und dass entsprechende Antworttendenzen (ERS und MRS) vorliegen. Von dieser Tendenz kann eindeutig jedoch nur in sechs von insgesamt 15 Modellvergleichen gesprochen werden. In den übrigen neun Modellvergleichen zwischen Modell A und B mussten hauptsächlich d- und e- aber auch m-Items auf Basis der Itemfits (RMSEA-Werte) aus den jeweiligen Dimensionen ausgeschlossen werden um eine entsprechende Modellpassung zu erzielen. Der prozentuale Anteil an ausgeschlossenen Items (2, 4, 7 und 22% an d-Items; 9, 11, 12, 26% an e-Items, 13% an m-Items aus den jeweiligen Dimensionen der betroffenen Modellvergleiche) ist teilweise beträchtlich, so dass von der eindeutigen Passung des entsprechenden Modells nicht mehr ausgegangen werden kann. Bei Betrachtung der Zugehörigkeit der ausgeschlossenen Items zu den eigentlichen Fragebogendimensionen lässt sich kein Trend feststellen. Somit können die einzelnen Fragebogen-Items als Verursachung ausgeschlossen werden. Auch sind alle Versuchsgruppen 1, 2 und 3 in ähnlicher Weise vom Item-Ausschluss betroffen.

Die marginalen IRT-basierten Reliabilitäten finden sich ergänzend im Anhang ab S. 212.

Angeichts der geringen Stichprobengrößen der einzelnen Versuchsgruppen ($n_{VG1}=127$, $n_{VG2}=126$, $n_{VG3}=132$) erlauben die Ergebnisse auf Basis der Gesamtstichprobe ($N=385$) genauere Aussagen hinsichtlich der Interpretation, als jene auf Basis der getrennten Analyse der drei Versuchsgruppen.

c) Interpretation der Ergebnisse zur Messgüte des Pseudo-Item d auf Basis der GSP

Die in Tabelle 32 und Tabelle 33 dargestellten geschätzten Skaleninterkorrelationen auf Basis der d- und Original-Items für die Gesamtstichprobe (Modellvergleich I gem. Tabelle 22) bestätigen im Allgemeinen die Annahme, dass die Kodierung mittels Pseudo-Item d tatsächlich zu einer höheren Messgüte führt als die Original-Kodierung mittels fünfstufiger Likert-Skala, da die meisten Skaleninterkorrelationen auf Basis der d-items geringer ausfallen als auf Basis der Original-Items. Dabei zeigt sich, dass die Differenzbeträge zwischen den Korrelationskoeffizienten (d-Items vs. Original-Items) im Mittel jedoch nicht sehr hoch ausfallen. Angesichts der nachgewiesenen Verzerrung der Daten durch inhaltsunabhängige Antworttendenzen (ERS und MRS) bei Verwendung der ursprünglichen Kodierungsmethode, stellen die hier dargestellten Ergebnisse jedoch lediglich eine Ergänzung dar, welche durch die leicht geringeren Korrelationskoeffizienten der d-Items im Vergleich zu den Original-Items (Likert-Skala) erneut bestätigen, dass der Kodierung mittels Pseudo-Item d im Gegensatz zur ursprünglichen Kodierungsmethode der Vorzug zu geben ist.

10 Diskussion

Die vorliegende Arbeit diente der Untersuchung eines möglichen Zusammenhangs zwischen Verfälschungsverhalten in Persönlichkeitsfragebögen und dem Ausmaß der Verwendung extremer sowie mittlerer Antworttendenzen (ERS und MRS) im Sinne der Iteminhalts-unabhängigen Wahl der 1. und 5. (ERS) bzw. 3. (MRS) Antwortkategorie. Untersucht wurde dieser Zusammenhang im Rahmen eines Experiments, in welchem drei Versuchsgruppen (VG 1, 2 und 3) bestehend aus Studierenden der Studienrichtungen Rechts- und Wirtschaftswissenschaften, drei unterschiedliche Verfälschungsinstruktionen vorgegeben wurden, woraufhin sie zwei Persönlichkeitsfragebögen (BIP, NEO-FFI) sowie eine soziale Erwünschtheitsskala (SES) bearbeiteten. Da die Verfälschungsinstruktionen lediglich auf die SES-Skala einen Einfluss in die erwartete Richtung zeigten, wurde lediglich diese Skala für die Interpretation eines möglichen Zusammenhangs zwischen Verfälschungsverhalten und Antworttendenzen herangezogen. Ein entsprechender Zusammenhang in Bezug auf die SES-Skala konnte lediglich für die Antworttendenz zur Mitte (MRS) festgestellt werden. Hier zeigte sich, dass eine grundsätzlich vorhandene Verfälschungsmotivation (VG 1: hohe Verfälschungsmotivation; VG 2: unsichere Verfälschungsmotivation) mit einer stärker ausgeprägten Tendenz zur Mitte einhergeht, als dies für eine geringe Verfälschungsmotivation bzw. ehrliches Antwortverhalten (VG 3) der Fall ist. Es zeigte sich erwartungsgemäß (Hypothese 2b), dass die unsichere Verfälschungsmotivation (VG 2) dazu führte, dass die mittlere Antwortkategorie häufiger gewählt wurde bzw. eine stärkere Antworttendenz zur Mitte (MRS) vorlag als bei starker Verfälschungsmotivation (VG 1). Ein Zusammenhang zwischen Verfälschungsverhalten und der Antworttendenz zu Extremurteilen (ERS) ließ sich auf Basis der Skala Soziale Erwünschtheit nicht nachweisen.

Die Eignung der verwendeten Kodierungsmethode nach Böckenholt (2012) zur Erhebung der Antworttendenzen auf Basis der Pseudo-Items d, e und m kann grundsätzlich, wie schon von Khorramdel und von Davier (under review) bestätigt

werden. In 24 von 25 Modellvergleichen stellte sich MRS über die Verwendung von Pseudo-Items m als eindimensional abbildbare Dimension dar. In einem Modellvergleich konnte dies erst nach Ausschluss von 13% der Items festgestellt werden. Die Antworttendenz zu Extremurteilen (ERS mittels e-Items) liess sich in 20 von 25 Modellvergleichen eindimensional abbilden. Nach Ausschluss von 11%, 26%, 9% und 12% der Items ergab sich das gleiche Bild für weitere vier der 25 Modellvergleiche. Lediglich in einem Modellvergleich war die e-Dimension (ERS) nicht eindimensional abbildbar. Mittels der Modellierung von Antworttendenzen über Pseudo-Items war es so möglich aufzuzeigen, dass inhaltsunabhängige Antworttendenzen im Datensatz der aktuellen Studie die Messung der Persönlichkeitsmerkmale verzerren. Diese Verzerrung kann jedoch korrigiert werden, wenn man statt der vorgesehenen Kodierung aller fünf Antwortkategorien (Original-Items) nur jene Antwortkategorien für die Bildung eines Scores heranzieht, welche nicht durch ERS und MRS verzerrt sind, d.h. unter Verwendung der Pseudo-Items d. Für Pseudo-Items d zeigte sich in den IRT-Modellen, dass diese der erwarteten Dimensionalität der Fragebogenskalen gerecht werden: die mehrdimensionalen IRT-Modelle auf Basis der Pseudo-Items d, welche so viele Dimensionen wie Fragebogenskalen beinhalteten, beschrieben die Daten jeweils besser als ein eindimensionales Modell. Weiters zeigte sich, dass die Skaleninterkorrelationen auf Basis der d-items für die meisten Skalen geringer ausfallen, als Skaleninterkorrelationen auf Basis der Original-Items (Kodierung aller fünf Antwortkategorien). Diese Ergebnisse entsprechen denen von Khorramdel und von Davier (under review) und zeigen auf, dass Datensätze auf das Vorliegen von Antworttendenzen überprüft und gegebenenfalls korrigiert werden sollten um faire Gruppenvergleiche zu ermöglichen.

Die Tatsache, dass eine Tendenz zu Extremurteilen (ERS) im Rahmen der hier vorliegenden Studie nicht mit Verfälschungsverhalten in Zusammenhang gebracht werden konnte, lässt Raum für weitere Untersuchungen.

Eine mögliche Ursache für den fehlenden Zusammenhang könnte sein, dass es durch die Verfälschungsinstruktionen nicht (vollständig) gelungen ist, das interessierende Verfälschungsverhalten zu induzieren. Das erwartete Verfälschungsverhalten zeigte sich lediglich in einer einzigen Skala (SES). Der Annahme, dass es einen Zusammenhang zwischen Antworttendenzen und Verfälschungsverhalten gibt (bzw. diese beiden Begriffe ein und dasselbe

Phänomen darstellen) liegt die Hypothese zugrunde, dass Verfälschungsverhalten ein eher stereotypes Antwortverhalten darstellt (vgl. Holden & Hibbs, 1995; Holden, Kroner, Fekken, & Popham, 1992). Eine weitere Erklärung könnte daher sein, dass Verfälschungsverhalten nicht stereotyp, sondern wesentlich komplexer ist und individuell variiert (vgl. Ellingson et al., 2001; McFarland & Ryan, 2000; Zickar et al., 2004).

Des Weiteren unterliegt die vorliegende Arbeit einigen Limitationen.

So sind die dargestellten Ergebnisse nicht über alle Analysen und Bereiche hinweg konsistent. Es ergaben sich sowohl Belege dafür, dass Pseudo-Item d innerhalb einer Dimension, also inhaltsunabhängig abgebildet werden konnte, als auch dafür, dass Pseudo-Item e und m gemeinsam in einer Dimension mit Pseudo-Item d, somit inhaltsabhängig abgebildet werden konnten.

Weiters konnte der Zusammenhang zwischen dem Vorliegen von Antworttendenzen und stereotypem Verfälschungsverhalten lediglich auf Basis der Skala Soziale Erwünschtheit untersucht werden. Die dazu vorgegebene SES-17-Skala (Stöber, 1999) bezieht sich nicht auf Persönlichkeitsmerkmale, sondern auf bspw. triviale Begebenheiten des Alltags. Des Weiteren sind die Fragen so formuliert, dass sie eine hohe Augenscheinvalidität aufweisen bzw. durchschaubar sind (z.B. „Ich habe schon einmal jemanden ausgenutzt oder übers Ohr gehauen“ oder „Ich lästere gelegentlich über andere hinter deren Rücken“). Somit ist die Übertragbarkeit der Ergebnisse auf tatsächliche Persönlichkeitsfragebögen zu hinterfragen.

Die Tatsache, dass trotz entsprechender Verfälschungsinstruktion die Testpersonen der VG 1 im Gegensatz zu jenen der VG 2 und 3 kein stärkeres Verfälschungsverhalten an den Tag legten, gibt Anlass zur Vermutung, dass der Erhalt der Rückmeldung vor dem Hintergrund der Verfälschungsinstruktionen einen womöglich unerwünschten Effekt hatte. So könnte die Hoffnung auf eine realistische (im Gegensatz zu durch Verfälschungsverhalten beeinflussten) Rückmeldung der VG 1 zur verringerten Verfälschungsmotivation trotz Verfälschungsaufforderung geführt haben. Und auch die Tatsache, dass die VG 1 über den Erhalt einer Rückmeldung erfahren würde, wie sie im Rahmen eines Bewerbungsverfahrens tatsächlich abschneiden würden, hatte nicht den gewünschten (verfälschenden) Effekt auf das Antwortverhalten der VG 1. In Erwartung der Tatsache, dass die Testpersonen der VG 3 eine tatsächliche

Aussage über berufsbezogene Merkmale der eigenen Person erhalten würden, könnten weiters die Testpersonen der VG 3 die Fragen so beantwortet haben, dass Sie eine möglichst positive Rückmeldung, im Sinne einer möglichst positiven Selbstdarstellung der eigenen Person erhalten würden. Somit sollte die Ankündigung einer Rückmeldung als Anreiz zur Studienteilnahme kritisch hinterfragt werden, da sie womöglich einen weiteren unkontrollierbaren Einflussfaktor darstellt. Des Weiteren könnten die fehlenden Konsequenzen des Verfälschungsverhaltens innerhalb mittels Verfälschungsinstruktionen künstlich induzierter (Bewerbungs-) Situationen, einen Effekt auf die Art und Weise des Verfälschungsverhaltens haben (Steffen, 2004).

In Bezug auf die Stichprobe zeigten sich Unterschiede zwischen den drei Versuchsgruppen, welche hinsichtlich der Übertragbarkeit der Ergebnisse Beachtung finden sollten: während in Versuchsgruppe 1 und 2 ungefähr gleich viele 18- bis 21-Jährige vorhanden waren (15% bzw. 18%) lag der Prozentsatz der 18- bis 21-Jährigen in VG 3 bei 27%. Dies schlug sich sodann auch im Anteil der Studierenden mit einer Semesteranzahl zwischen einem und drei Semestern nieder (10%, 9% und 20% für VG 1, 2 und 3) sowie im Anteil jener, die bereits Vollzeit beschäftigt sind oder waren (34%, 27% und 21%). Die Tatsache, dass eine Person im Rahmen des Studiums oder der Vollzeit-Beschäftigung mit seinen berufsbezogenen Merkmalen konfrontiert wird, könnte sich somit auf die Beantwortung der Fragen im Sinne eines differenzierteren Persönlichkeitsbildes niederschlagen. Weiters lag der Anteil weiblicher und männlicher Testpersonen in allen Stichproben lag bei ca. 60:40.

Eine weitere Schwierigkeit bei der Interpretation der vorliegenden Ergebnisse ist die Tatsache, dass die ursprünglich als Kontrollstichprobe vorgesehene VG 3 entgegen der Erwartung kein ehrliches Antwortverhalten zeigte und somit auf eine Kontrollstichprobe verzichtet werden musste. Eine derartige Datengrundlage hätte womöglich einen klärenden Beitrag zu den derzeit vorliegenden Ergebnissen leisten können.

Des Weiteren wäre für IRT-Modellberechnungen eine grössere Stichprobe als die hier erhobene Stichprobe ($N=385$, bzw. $n_{VG1}=127$, $n_{VG2}=126$, $n_{VG3}=132$) zur Erzielung reliabler Parameterschätzungen wünschenswert gewesen.

Es ist weiters zu beachten, dass die Vermittlung der Verfälschungsinstruktionen aufgrund der fehlenden Kontrollierbarkeit der Online-Testsituation womöglich

nicht den gewünschten Einfluss auf vereinzelte Testpersonen hatte. Eine alternative Möglichkeit den Effekt von Verfälschungsinstruktionen zu kontrollieren, ist die Vorgabe eines konkreten Berufs-Anforderungsprofils (z.B. Pauls & Crost, 2005). Diese Vorgabe führt sowohl zu differenzierteren Mittelwertsprofilen (Pauls & Crost, 2005) als auch zu Ergebnissen, welche den Ergebnissen tatsächlicher Bewerbungssituationen ähnlicher sind (Raymark & Tafero, 2009). Durch die Vorgabe eines Anforderungsprofils lässt sich ausserdem eine sichere Aussage zum Vorliegen von Verfälschungsverhalten treffen, da von Beginn an klar ist in welchen Dimensionen (nämlich jenen, die hinsichtlich des Anforderungsprofils relevant sind) verfälschendes Antwortverhalten zu erwarten ist, so dass eine Untersuchung des Zusammenhangs zwischen Verfälschungsverhalten und dem Vorliegen von Antworttendenzen unter Vorgabe von Berufs-Anforderungsprofilen zu aufschlussreichen Erkenntnissen führen könnte. In der vorliegenden Studie wurde jedoch von der Vorgabe eines Anforderungsprofils abgesehen um der Tatsache vorzubeugen, dass die Testpersonen sich vom Anforderungsprofil womöglich nicht angesprochen fühlen und dies die Testbearbeitungsmotivation beeinflusst.

Der Versuch über Verfälschungsinstruktionen das Verfälschungsverhalten zu kontrollieren zeigt sich dennoch als vielversprechender Ansatz um gezielte Aussagen über einen möglichen Zusammenhang zu Antworttendenzen möglich zu machen. Zukünftige Studien sollten dabei die in der vorliegenden Studie getätigten Erfahrungen zum Einfluss von Verfälschungsinstruktionen berücksichtigen.

Der Erhalt einer schriftlichen Rückmeldung an die Testpersonen zeigte sich als Möglichkeit, die Testbearbeitungsmotivation der Testpersonen zu erhöhen bzw. zu beeinflussen.

Letztendlich konnte die Arbeit einen interessanten Zusammenhang zwischen Verfälschungsverhalten und einer Antworttendenz zur Mitte MRS aufzeigen und lieferte eine erneute positive Bestätigung über die Anwendung von Pseudo-Items zur Identifikation vorliegender Antworttendenzen innerhalb eines Datensatzes.

11 Zusammenfassung

Die vorliegende Arbeit diene der Untersuchung des Zusammenhangs zwischen stereotypem Verfälschungsverhalten im Sinne sozial erwünschten Antwortverhaltens und Iteminhalts-unabhängiger Antworttendenzen (Tendenz zu Extremurteilen, ERS; Tendenz zur Mitte, MRS) im Rahmen der Vorgabe von Persönlichkeitsfragebögen unter bestimmten Verfälschungsinstruktionen. Um einen möglichen Zusammenhang zwischen Verfälschungsverhalten und dem Vorliegen von extremen sowie mittleren Antworttendenzen überprüfen zu können, wurden drei Versuchsgruppen ($n_{VG1}=127$, $n_{VG2}=126$, $n_{VG3}=132$) bestehend aus Studierenden der Rechts- und Wirtschaftswissenschaften zunächst mittels verschiedener Verfälschungsinstruktionen auf unterschiedliche Art und Weise beeinflusst, um so ein starkes (VG 1), ein weniger ausgeprägtes (bzw. durch Unsicherheit geprägtes) Verfälschungsverhalten (VG 2), sowie geringes Verfälschungsverhalten bzw. ehrliches Antwortverhalten (VG 3) im Rahmen der Beantwortung des vorgegebenen Online-Fragebogens (bestehend aus Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung BIP, NEO-FFI und SES-17 zur Erhebung der Skala Soziale Erwünschtheit) zu induzieren. Das durch die Instruktionen intendierte Verfälschungsverhalten stellte sich dabei auf Basis der Ergebnisse des Welch-Tests lediglich für eine der 20 erhobenen Fragebogenskalen, die Skala Soziale Erwünschtheit, ein.

Für diese Skala konnte mittels multidimensionaler IRT-Modelle die Frage überprüft werden, ob Antworttendenzen (ERS und MRS) vorliegen und inwiefern das vorliegende Verfälschungsverhalten innerhalb der Versuchsgruppen mit dem Vorliegen von Antworttendenzen zusammenhängt. Dabei bestand die Vermutung, dass im Falle des stärksten Verfälschungsverhaltens bzw. der stärksten Verfälschungsmotivation (VG 1) insbesondere extreme Antworttendenzen auftreten würden. Für VG 2 (weniger starkes Verfälschungsverhalten bzw. unsichere Verfälschungsmotivation) wurde die Hypothese aufgestellt, dass das Antwortverhalten insbesondere mit mittleren Antworttendenzen einhergehen würde. Für Versuchsgruppe 3 (ehrliches Antwortverhalten) wurde vermutet, dass es hier zu keinem vorrangigen Vorliegen von Antworttendenzen kommen würde.

Zur Überprüfung dieser Annahmen wurden die Antworten der Testpersonen auf einer 5-kategoriellen Ratingskala (Likert-Skala) gemäß Khorramdel und von Davier (under review) zunächst in binäre sog. Pseudo-Items (d, e und m) zerlegt. Pseudo-Item d (zustimmende Antwortkategorien der Likert-Skala) diente der Iteminhalts-abhängigen Erhebung der eigentlichen Fragebogenskalen. Pseudo-Items e repräsentieren die extremen Antwortkategorien (1 und 5) und dienten der Erhebung der inhaltsunabhängigen extremen Antworttendenzen (ERS), während Pseudo-Items m (mittlere Antwortkategorie 3) die Repräsentation der inhaltsunabhängigen mittleren Antworttendenz (MRS) darstellen. Mittels dieser Pseudo-Items wurden sodann ein- und mehrdimensionale IRT-Modellvergleiche durchgeführt und überprüft, ob sich Antworttendenzen eindimensional über alle Fragebogenskalen abbilden lassen. Die Ergebnisse der Modellvergleiche gaben über Informationsverlust-Kriterien (AIC- und BIC-Kriterium) Aufschluss über die Frage, welches der verglichenen Modelle die tatsächliche Datenstruktur am besten repräsentierte und liessen somit Rückschlüsse über den Zusammenhang zwischen Verfälschungsverhalten und dem Vorliegen extremer und mittlerer Antworttendenzen zu.

Die Ergebnisse konnten die vermuteten Zusammenhänge zwischen stereotypem Verfälschungsverhalten und inhaltsunabhängigen Antworttendenzen jedoch nur teilweise bestätigen. So konnte ein Zusammenhang zwischen der inhaltsunabhängigen Verwendung der mittleren Antwortkategorie (MRS) und Verfälschungsverhalten für die SES Skala nachgewiesen werden. Dabei führte eine unsichere Verfälschungsmotivation (VG 2) erwartungsgemäß zur höheren Ausprägung von MRS, als dies für stärkeres Verfälschungsverhalten bzw. unter starker Verfälschungsmotivation (VG 1) der Fall war. Ein Zusammenhang zwischen einer Tendenz zu Extremurteilen (ERS) und Verfälschungsverhalten i.S. sozialer Erwünschtheit konnte für die SES Skala nicht nachgewiesen werden. In Bezug auf die restlichen 19 Fragebogenskalen (welche im Gegensatz zur SES Skala keine signifikanten Unterschiede zwischen den drei Versuchsgruppen aufwiesen) zeigte sich grundsätzlich, dass sich beide Antworttendenzen (ERS und MRS) eindimensional über alle Fragebogenskalen abbilden lassen, was die ausschliessliche Annahme inhaltlicher Fragebogendimensionen in Frage stellt. Es wurde ausserdem festgestellt, dass Pseudo-Item d bei der Erhebung der eigentlichen Fragebogenskalen im Vergleich zu den Original-Items (Likert-Skala)

eine höhere Messgüte zeigt: auf Basis der d-Items resultierten geringere Skaleninterkorrelationen als auf Basis der Original-Items, d.h. die inhaltlichen Konstrukte des Fragebogens lassen sich auf Basis der d-Items klarer voneinander abgrenzen.

Hinsichtlich der vorliegenden Ergebnisse gilt es zu beachten, dass die Ergebnisse zum Zusammenhang zwischen MRS und Verfälschungsverhalten lediglich auf Basis der Skala Soziale Erwünschtheit (SES, keine Persönlichkeits-Skala) gezeigt werden konnten. Des Weiteren stellen sich die für IRT-Analysen relativ geringe Stichprobengrösse ($N=385$, bzw. $n_{VG1}=127$, $n_{VG2}=126$, $n_{VG3}=132$), sowie die unkontrollierten Testbedingungen der webbasierten Testung als Kritikpunkte dar. Im Rahmen zukünftiger Studien sollten diese Faktoren beachtet werden. Auf Basis der vorliegenden Ergebnisse besteht die Vermutung, dass sich Verfälschungsverhalten weniger stereotyp darstellt, sondern auf einem eher komplexen Prozess mit multiplen Einflussfaktoren basiert. Diese Annahme sollte in zukünftigen Studien überprüft werden.

Literaturverzeichnis

Arthur, W., Woehr, D. & Graziano, W. (2001). Personality testing in employment settings: Problems and issues in the application of typical selection practices. *Personnel Review*, 30, 657 – 676.

Austin, E., Deary, I. & Egan, V. (2006). Individual differences in response scale use: Mixed Rasch modeling of responses to NEO-FFI Items. *Personality and Individual Differences*, 40, 1235-1245.

Bardwell, W. & Dimsdale, J. (2001). The Impact of Ethnicity and Response Bias on the Self-Report of Negative Affect. *Journal of Applied Biobehavioral Research*, 6, 27-38.

Barrick, M. & Mount, M. (1991). The Big Five Personality Dimensions and Job Performance: a Meta-Analysis. *Personnel Psychology*, 44, 1-26.

Baumgartner, H. & Steenkamp, J.-B. (2011). Response Styles in Marketing Research: A Cross-National Investigation. *Journal of Marketing Research*, 36, 143-156.

Bolt, D. & Johnson, T. (2009). Addressing Score Bias and Differential Item Functioning Due to Individual Differences in Response Style. *Applied Psychological Measurement*, 33, 335-352.

Bolt, D. & Newton, J. (2011). Multiscale Measurement of Extreme Response Style. *Educational and Psychological Measurement*, 71, 814-833.

Borkenau, P. & Ostendorf, F. (1993). *NEO-Fünf-Faktoren-Inventar (NEO-FFI) nach Costa und McCrae*. Göttingen: Hogrefe.

Böckenholt, U. (2012). Modelling Multiple Response Processes in Judgment and Choice. *Psychological Methods*, Advance online Publication.

Burnham, K.P. & Anderson, D. R. (2004). Understanding AIC and BIC in Model Selection, *Sociological Methods & Research*, 33, 261-304.

Burns, G. & Christiansen, N. (2011). Methods of Measuring Faking Behavior, *Human Performance*, 24, 358-372.

Cabooter, E. (2010). *The Impact of situational and dispositional variables on Response Styles with respect to attitude measures*, Unveröff. Diss., Universiteit Gent, Gent.

Campbell, D. & Fiske, D. (1959). Convergent and Discriminant Validation by the Multi-trait-Multimethod Matrix. *Psychological Bulletin*, 56, 81-105.

Dalen, L., Stanton, N. & Roberts, A. (2001). Faking personality questionnaires in personnel selection. *Journal of Management Development*, 20, 729 – 742.

DeJong, M., Steenkamp, J.-B., Fox, J.-P. & Baumgartner, H. (2008). Using Item Response Theory to Measure Extreme Response Styles in Marketing Research: A Global Investigation. *Journal of Marketing Research*, 45, 104-115.

- Development Dimensions International, DDI (2005). DDI 2005 Selection Forecast (online). URL: <http://ddiworld.com.au/global-offices/asia-pacific/australia/australia-press-room/new-report-highlights-cost-of-bad-hiring-decisions> (9.01.2013).
- Dilchert, S., Ones, D., Viswesvaran, C. & Deller, J. (2006). Response Distortion in Personality Measurement: Born to Deceive, yet Capable of Providing Valid Self-Assessment? *Psychology Science*, 48, 209-225.
- Donovan, J., Dwight, S. & Hurtz, G. (2003). An Assessment of the Prevalence, Severity, and Verifiability of Entry-Level Applicant Faking Using the Randomized Response Technique. *Human Performance*, 16, 81-106.
- Douglas, E., McDaniel, M. & Snell, A. (1996). The Validity of Non-Cognitive Measures Decays When Applicants Fake. Academy of Management Proceedings, Cincinnati.
- Edmondson, D. (2005). *Likert Scales: A History*. 12th Conference on Historical Analysis and Research in Marketing (CHARM), Long Beach.
- Ellingson, J. (2012). People Fake Only When They Need to Fake. In M. Ziegler, C. MacCann & R. Roberts (Hrsg.), *New Perspectives on Faking in Personality Assessment*. New York: Oxford University Press.
- Ellingson, J. & McFarland, L. (2011). Understanding Faking Behavior Through The Lens of Motivation: An Application of VIE Theory. *Human Performance*, 24, 322-337.
- Fan, J., Gao, D., Carroll, S. A., Lopez, F. J., Tian, T. S., & Meng, H. (2012). Testing the Efficacy of a New Procedure for Reducing Faking on Personality Tests Within Selection Contexts, *Journal of Applied Psychology*, Advance online Publication.
- Filipkowski, J. (2007). Impression Management across Applicant and Incumbant Contexts: The Effect on Job Performance. Unveröff. Diss., Wright State University, Dayton.
- Furnham, A. (1990). Faking Personality Questionnaires: Fabricating Different Profiles for Different Purposes. *Current Psychology: Research & Reviews*, 9, 46-55.
- Galic, Z., Jerneic, Z. & Kovacic, M. (2012). Do Applicants Fake Their Personality Questionnaire Responses and How Successful are Their Attempts? A Case of Military Pilot Cadet Selection. *International Journal of Selection and Assessment*, 20, 229-241.
- Gernand, L. & Fenske, N. (2009). *Understanding AIC and BIC in Model Selection*. Manuskript im Rahmen der Veranstaltungsreihe Modellwahl und Modelldiagnostik, Ludwig-Maximilians-Universität, München.
- Goffin, R. & Christiansen, N. (2003). Correcting Personality Test for Faking: A Review of Popular Personality Tests and an Initial Survey of Researchers. *International Journal of Selection and Assessment*, 11, 340-344.
- Gibbons, J., Zellner, J. & Rudek, D. (1999). Effects of Language and Meaningfulness on the Use of Extreme Response Style by Spanish-English Bilinguals. *Cross-Cultural Research*, 33, 369-381.
- Greenleaf, E. (1992). Measuring Extreme Response Style. *The Public Opinion Quarterly*, 56, 328-351.

Griffith, R. & Peterson, M. (2008). Failure of Social Desirability Measures to Capture Applicant Faking Behavior. *Industrial and Organizational Psychology, Perspectives on Science and Practice*, 1, 308-311.

Griffith, R., Lee, L., Peterson, M. & Zickar, M. (2011). First Dates and Little White Lies: A Trait Classification Theory of Applicant Faking Behavior. *Human Performance*, 24, 338-357.

Griffith, R. & Peterson, M. (2011). One Piece at a Time: The Puzzle of Applicant Faking and a Call for Theory. *Human Performance*, 24, 291-301.

Harzing, A.-W. (2006). Response Styles in Cross-National Survey Research. *International Journal of Cross Cultural Management*, 6, 243-266.

Heggestad, E., Morrison, M., Reeve, C. & McCloy, R. (2006). Forced-Choice Assessments of Personality for Selection: Evaluating Issues of Normative Assessment and Faking Resistance. *Journal of Applied Psychology*, 91, 9-24.

Herche, J. & Engelland, B. (1996). Reversed-polarity items and scale unidimensionality. *Journal of the Academy of Marketing Science*, 24, 366-374.

Higgins, E. (1997). Beyond Pleasure and Pain. *American Psychologist*, 52, 1280-1300.

Hogan, R. (2005). In Defense of Personality Measurement: New Wine for old Whiners. *Human Performance*, 18, 331-341.

Hogan, J. & Holland, B. (2003). Using Theory to Evaluate Personality and Job-Performance Relations: A Socioanalytic Perspective. *Journal of Applied Psychology*, 88, 100-112.

Honkaniemi, L. (2012). *Applicants in a Real-Life Selection Context*. Doctoral Dissertation, University of Jyväskylä, Jyväskylä.

Hossiep, R. & Paschen, M. (2003). *BIP Das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung*. Göttingen: Hogrefe.

Hough, L., Eaton, N., Dunnette, M., Kamp, J. & McCloy, R. (1990). Criterion-Related Validities of Personality Constructs and the Effect of Response Distortion on Those Validities. *Journal of Applied Psychology Monograph*, 75, 581-595.

Hough, L. & Oswald, F. (2005). They're Right, Well ... Mostly Right: Research Evidence and an Agenda to Rescue Personality Testing From 1960s Insights. *Human Performance*, 18, 373-387.

Hurtz, G. & Donovan, J. (2000). Personality and Job Performance: The Big Five Revisited. *Journal of Applied Psychology*, 85, 869-879.

Jackson, D., Wroblewski, V. & Ashton, M. (2000). The Impact of Faking in Employment Tests: Does Forced Choice Offer a Solution? *Human Performance*, 13, 371-388.

Johnson, T. & Bolt, D. (2010). On the Use of Factor-Analytic Multinomial Logit Item Response Models to Account for Individual Differences in Response Styles. *Journal of Educational and Behavioral Statistics*, 35, 92-114.

- Kanning, U. P. & Holling, H. (2001). Struktur, Reliabilität und Validität des NEO-FFI in einer Personalauswahlsituation. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 22, 239-247.
- Kanning, U. & Kuhne, S. (2006). Social desirability in a multimodal personnel selection test battery. *European Journal of Work and Organizational Psychology*, 15, 241-261.
- Kauffeld, S. & Grohmann, A. (2011). Personalauswahl. In S. Kauffeld (Hrsg.), *Arbeits- und Organisationspsychologie für Bachelor*. Heidelberg: Springer
- Khorramdel, L. & Kubinger, K. (2006). The effect of speediness on personality questionnaires: an experiment on applicants within a job recruiting procedure. *Psychology Science*, 48, 378-397.
- Khorramdel, L. & von Davier, M. (under review). Measuring Response Styles across the Big Five: A Multi-Scale Extension of an Approach using Multinomial Processing Trees.
- Kieruj, N. & Moors, G. (2013). Response Style Behavior: question format dependent or personal style? *Quality and Quantity*, 47, 193-211.
- Komar, S., Komar, J., Robie, C. & Tagger, S. (2010). Speeding Personality Measures to Reduce Faking. *Journal of Personnel Psychology*, 9, 126-137.
- König, C., Melchers, K., Kleinmann, M., Richter, G. & Klehe, U.-C. (2006). The Relationship Between the Ability to Identify Evaluation Criteria and Integrity Test Scores. *Psychology Science*, 48, 369-377.
- Krahé, B. & Herrmann, J. (2003). Verfälschungstendenzen im NEO-FFI: Eine experimentelle Überprüfung. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24, 105-117.
- Kubinger, K. (2009). *Psychologische Diagnostik*. Göttingen: Hogrefe.
- Kubinger, K., Rasch, D. & Yanagida, T. (2011). *Statistik in der Psychologie*. Göttingen: Hogrefe.
- Kuncel, N., Goldberg, L. & Kiger, T. (2011). A Plea for Process in Personality Prevarication, *Human Performance*, 24, 373-378.
- Lüdtke, O., Trautwein, U., Nagy, G. & Köller, O. (2004). Eine Validierungsstudie zum NEO-FFI in einer Stichprobe junger Erwachsener: Effekte des Itemformats, faktorielle Validität und Zusammenhänge mit Schulleistungsindikatoren. *Diagnostica*, 50, 134-144.
- MacCann, C., Ziegler, M. & Roberts, R. (2012). Faking in Personality Assessment: Reflections and Recommendations. In M. Ziegler, C. MacCann & R. Roberts (Hrsg.), *New Perspectives on Faking in Personality Assessment*. New York: Oxford University Press.
- McCoach, D.B. & Black, A.C. (2008). Evaluation of Model Fit and Adequacy. In O'Connell, A.A. & McCoach, D.B. (Hrsg.), *Multilevel Modeling of Educational Data* (S. 245-272). Charlotte: Information Age Publishing.
- McCrae, R. & Costa, P. (1987). Validation of the Five-Factor Model of Personality Across Instruments and Observers. *Journal of Personality and Social Psychology*, 52, 81-90.

- McFarland, L. (2003). Warning Against Faking on Personality Test: Effects on Applicant Reactions and Personality Test Scores. *International Journal of Selection and Assessment*, 11, 265-276.
- McFarland, L. & Ryan, A. (2000). Variance in Faking Across Noncognitive Measures. *Journal of Applied Psychology*, 85, 812-821.
- Morgeson, F., Campion, M., Dipboye, R., Hollenbeck, J., Murphy, K. & Schmitt, N. (2007). Reconsidering the Use of Personality Tests in Personnel Selection Contexts. *Personnel Psychology*, 60, 683-729.
- Morren, M., Gelissen, J. & Vermunt, J. (2011). Dealing with Extreme Response Style in Cross-cultural Research: A Restricted Latent Class Factor Analysis Approach. *Sociological Methodology*, 41, 13-47.
- Mueller-Hanson, R., Heggstad, E. & Thornton III, G. (2006). Individual Differences in Impression Management: an Exploration of the Psychological Processes Underlying Faking. *Psychology Science*, 3, 288-312.
- Murphy, K. & Dziewieczynski, J. (2005). Why Don't Measures of Broad Dimensions of Personality Perform Better As Predictors of Job Performance? *Human Performance*, 18, 343-357.
- Nachtwei, J. & Schermuly, C. (2009). Acht Mythen über Eignungstests. *Harvard Business Manager*, 4, 6-10.
- Nguyen, N., Biderman, M. & McDaniel, M. (2005). Effects of Response Instructions on Faking a Situational Judgment Test. *International Journal of Selection and Assessment*, 13, 250-260.
- Obermann, C. (2009). *Assessment Center - Entwicklung, Durchführung, Trends*. Wiesbaden: Gabler.
- Ones, D. (2005). Personality at Work: Raising Awareness and Correcting Misconceptions. *Human Performance*, 18, 389-404.
- Ones, D., Viswesvaran, C. & Reiss, A. (1996). Role of social desirability in personality testing for personnel selection: The redherring. *Journal of Applied Psychology*, 81, 660-679.
- Paulhus, D. (1984). Two-Component Models of Socially Desirable Responding. *Journal of Personality and Social Psychology*, 46, 598-609.
- Pauls, C. & Crost, N. (2005b). Effects of different instructional sets on the construct validity of the NEO-PI-R. *Personality and Individual Differences*, 39, 297-308.
- Rademacher, F. (2008). *Herausforderung Mitarbeiterbindung*. Studienarbeit, Berufsakademie Mannheim, München: GRIN Verlag.
- Rasch D. & Kubinger K. (2006). *Statistik für das Psychologiestudium*. München: Elsevier.
- Raymark, P. & Tafero, T. (2009). Individual Differences in the Ability to Fake on Personality Measures. *Human Performance*, 22, 86-103.
- Reeder, M. & Ryan, M. (2012). Methods for correcting of Faking. In M. Ziegler, C.

- MacCann & R. Roberts (Hrsg.), *New Perspectives on Faking in Personality Assessment*. New York: Oxford University Press.
- Reinecke, J. (2005). *Strukturgleichungsmodelle in den Sozialwissenschaften*. München: Oldenbourg.
- Rost, J., Carstensen, C. & von Davier, M. (1999). Sind die Big Five Rasch-skalierbar? *Diagnostica*, 45, 119-127.
- Röhner, J., Schröder-Abé, M. & Schütz, A. (2011). Exaggeration is Harder Than Understatement, but Practice Makes Perfect! *Experimental Psychology*, 58, 464-472.
- Ryan, A., McFarland, L., Baron, H. & Page, R. (1999). An International Look at Selection Practices: Nation and Culture as Explanations for Variability in Practice. *Personnel Psychology*, 52, 359-391.
- Sackett, P. (2011). Integrating and Prioritizing Theoretical Perspectives on Applicant Faking of Personality Measures, *Human Performance*, 24, 379-385.
- Schuler, H. & Frintrup, A. (2008). Persönlichkeit macht den Unterschied. *Bildung Aktuell*, 5, 8-10.
- Schuler, H., Hell, B., Trapmann, S., Schaar, H. & Boramir, I. (2007). Die Nutzung psychologischer Verfahren der externen Personalauswahl in deutschen Unternehmen. *Zeitschrift für Personalpsychologie*, 6, 60-70.
- Shoss, M. & Strube, M. (2011). How Do You Fake a Personality Test? An Investigation of Cognitive Models of Impression-Managed Responses. *Organizational Behavior and Human Decision Processes*, 116, 163-171.
- Sireci, S. G., Thissen, D., & Wainer, H. (1991). On the reliability of testlet-based tests. *Journal of Educational Measurement*, 28, 237-247.
- Smith, D. & McDaniel, M. (2012). Questioning Old Assumptions: Faking and the Personality-Performance Relationship. In M. Ziegler, C. MacCann & R. Roberts (Hrsg.), *New Perspectives on Faking in Personality Assessment*. New York: Oxford University Press.
- Society for Human Resource Management, SHRM (2011). SHRM Poll: Personality Tests for Hiring and Promotion of Employees (online). URL: <http://www.shrm.org/research/surveyfindings/articles/pages/shrmpollpersonalitytestsforthehiringandpromotionofemployees.aspx> (9.02.2013).
- Steffens, M. (2004). Is the Implicit Association Test Immune to Faking? *Experimental Psychology*, 51, 165-179.
- Stöber, J. (1999). Die Soziale-Erwünschtheits-Skala-17 (SES-17): Entwicklung und erste Befunde zu Reliabilität und Validität. *Diagnostica*, 45, 173-177.
- Tett, R. & Simonet, D. (2011). Faking in Personality Assessment: A "Multisaturation" Perspective on Faking as Performance. *Human Performance*, 24, 302-321.
- Thiemann, H. (2006). *Sozial erwünschte Antworttendenzen bei der Bewerberauswahl mit Persönlichkeitsfragebögen*. Unveröff. Dissertation, Ruhr-Universität Bochum, Bochum.

- Turnley, W. & Bolino, M. (2001). Achieving Desired Images While Avoiding Undesired Images: Exploring the Role of Self-Monitoring in Impression Management. *Journal of Applied Psychology*, 86, 351-360.
- Vaerenbergh, Y. & Thomas, T. (2012). Response Styles in Survey Research: A Literature Review of Antecedents, Consequences and Remedies. *International Journal of Public Opinion Research*, Advance online Publication, 19. Juni 2012.
- Van Herk, H., Poortinga, Y. & Verhallen, T. (2004). Response Styles in Rating Scales. *Cross-Cultural Psychology*, 35, 346-360.
- Van Hooft, E. & Born, M. (2012). Intentional Response Distortion on Personality Tests: Using Eye-Tracking to Understand Response Processes When Faking. *Journal of Applied Psychology*, 97, 301-316.
- Viswesvaran, C. & Ones, D. (1999). Meta-Analyses of Fakability Estimates: Implications for Personality Measurement. *Educational and Psychological Measurement*, 59, 197-210.
- von Davier, M. (2005). *A general diagnostic model applied to language testing data* (Research Report No. RR-05-16). Princeton, NJ: Educational Testing Service.
- Weijters, B., & Baumgartner, H. (2012). Misresponse to Reversed and Negated Items in Surveys: A Review. *Journal of Marketing Research*, 49, 737-747.
- Weijters, B., Cabooter, E. & Schillewaert, N. (2010a). The effect of rating scale format on response styles: The number of categories and response category labels. *International Journal of Research in Marketing*, 27, 236-247.
- Weijters, B., Geuens, M. & Schillewaert, N. (2010b). The Individual Consistency of Acquiescence and Extreme Response Style in Self-Report Questionnaires. *Applied Psychological Measurement*, 34, 105-121.
- Weijters, B., Geuens, M. & Schillewaert, N. (2010c). The Stability of Individual Response Styles. *Psychological Methods*, 15, 96-110.
- Weijters, B., Schillewaert, N. & Geuens, M. (2008). Assessing Response Styles Across Modes of Data Collection. *Journal of the Academy of Marketing Science*, 36, 409-422.
- Weuster, A. (2008). Grundlagen der Personalauswahl. In A. Weuster (Hrsg.) *Personalauswahl*. Wiesbaden: Gabler.
- Widhiarso, W. (2011). Typology of Ideal Personality: Applicant Perspective in Job Selection Setting, 1st International Conference on Information Systems for Business Competitiveness (ICISBC), Semarang.
- Ziegler, M. & Bühner, M. (2009). Modeling Socially Desirable Responding and Its Effects. *Educational and Psychological Measurement*, 69, 548-565.
- Ziegler, M., McCann, C. & Roberts, R. (2012). Faking: Knowns, Unknowns and Points of Contention. In M. Ziegler, C. MacCann & R. Roberts (Hrsg.), *New Perspectives on Faking in Personality Assessment*. New York: Oxford University Press.

Ziegler, M., Schmidt-Atzert, L., Bühner, M. & Krumm, S. (2007). Fakability of different measurement methods for achievement motivation: questionnaire, semi-projective, and objective. *Psychology Science*, 49, 291-307.

IV. ANHANG

1 Anhang zum BIP

1.1 Skalen-Interkorrelationsmatrix BIP

Bereiche		Berufliche Orientierung (BO)			Arbeitsverhalten (AV)			Soziale Kompetenz (SK)					Psychische Konstitution (PK)		
	Dimension	LM	GM	FM	Ge	FI	HO	Sen	Ko	Soz	TO	Du	ES	Bel	Sb
BO	Leistungsmotivation (LM)	1.00													
	Gestaltungsmotivation (GM)	.56	1.00												
	Führungsmotivation (FM)	.48	.59	1.00											
AV	Gewissenhaftigkeit (Ge)	.22	-.04	-.03	1.00										
	Flexibilität (FI)	.47	.53	.55	-.26	1.00									
	Handlungsorientierung (HO)	.34	.26	.40	.33	.38	1.00								
SK	Sensitivität (Sen)	.20	.15	.32	.07	.32	.37	1.00							
	Kontaktfähigkeit (Ko)	.23	.33	.50	-.07	.43	.32	.54	1.00						
	Soziabilität (Soz)	-.14	-.34	-.20	.06	-.04	.08	.39	.14	1.00					
	Teamorientierung (TO)	.06	.09	.28	-.11	.30	.19	.24	.34	.28	1.00				
	Durchsetzungsstärke (Du)	.43	.62	.75	.05	.47	.41	.29	.46	-.33	.16	1.00			
PK	Emotionale Stabilität (ES)	.21	.23	.41	-.07	.50	.46	.32	.37	.14	.23	.39	1.00		
	Belastbarkeit (Bel)	.52	.37	.50	.10	.60	.58	.36	.33	.11	.21	.43	.69	1.00	
	Selbstbewusstsein (Sb)	.35	.49	.66	-.03	.54	.47	.36	.55	-.14	.19	.67	.66	.58	1.00

Die grau markierten Zellen beinhalten die Interkorrelationen zwischen den innerhalb eines Bereichs (AV, BO, SK, PK) zusammengefassten Dimensionen. Die fett markierten Werte kennzeichnen jene Korrelationskoeffizienten, deren Betrag $\geq .50$ beträgt und somit als inhaltlich besonders relevant interpretiert werden kann.

2 Online-Fragebogen

2.1 Online-Fragebogen



Wissen Sie, welchen Eindruck Sie bei Persönlichkeitsverfahren in Bewerbungssituationen hinterlassen?

Nein? Dann nehmen Sie an dieser Studie teil!

Denn Sie erhalten eine **detaillierte Rückmeldung ihrer Ergebnisse**, wie Sie in einer realen Bewerbungssituation dem potenziellen Arbeitgeber vorliegen würde. Diese Rückmeldung liefert wertvolle Informationen über die individuellen Ausprägungen grundlegender Persönlichkeitseigenschaften.

Der Fragebogen dient somit zur Vorbereitung und Übung für künftige Assessment-Center! Es lohnt sich also **20-30 Minuten** ihrer Zeit in die Bearbeitung des Fragebogens zu investieren!

Die Studie wird im Rahmen einer wissenschaftlichen Arbeit (Diplomarbeit) der Fakultät für Psychologie an der Universität Wien durchgeführt und behandelt das Thema **Persönlichkeitsverfahren im Assessment-Center**.

[Assessment-Center](#).

Für Rückfragen können Sie sich jederzeit an Ihre **Ansprechperson Frau Victoria Albertz** wenden (a0542736@unet.univie.ac.at).

Selbstverständlich werden Ihre Angaben vertraulich behandelt, so dass Ihre Anonymität zu jedem Zeitpunkt gewährleistet ist.

weiter...

Die folgenden zwei Seiten erhielten lediglich die Testpersonen der **Versuchsgruppe 1**

Hinweis

Stellen Sie sich bitte vor, Sie haben sich für eine Stelle beworben!

Das Unternehmen hat Sie nun zu einem Kennenlern-Gespräch eingeladen und möchte auch eine Auswahltestung (Assessment-Center) mit Ihnen machen.
Bei Ankunft im Unternehmen müssen Sie feststellen, dass Sie nicht der einzige Bewerber für die Stelle sind.

Der Stellenanzeige konnten Sie bereits einige Informationen zur Stelle entnehmen. **Diese machten einen sehr positiven Eindruck auf Sie.**

Während des Kennenlern-Gesprächs hat sich dieser positive erste Eindruck der Stelle vollkommen bestätigt.

Sie sind überzeugt, dass die Stelle genau ihren Vorstellungen entspricht und nehmen sich vor, ihr Bestes zu geben um den Job zu bekommen.

Sie sind daher bemüht im Assessment-Center Angaben zu machen, die ihre Chancen auf die Stelle erhöhen.

Nach dem das Kennenlern-Gespräch beendet ist, werden Sie für das Assessment-Center gemeinsam mit einigen Mitbewerbern in einen Raum geführt. Dort wird Ihnen ein Platz am Computer zugewiesen, mit der Bitte einen Fragebogen auszufüllen.

Auf dem Bildschirm erscheint die folgende Instruktion:

(bitte "weiter.."-Button anklicken)

weiter...

Herzlich Willkommen im Assessment-Center!

Es freut uns dass Sie Interesse an unserem Unternehmen haben!

Im folgenden Fragebogen finden Sie **287 Aussagen** zur Erfassung von Eigenschaften und **Verhaltensweisen im Arbeitsleben**.

Zur Beurteilung der Aussagen steht Ihnen eine 5-fach abgestufte Skala zur Verfügung.
Bitte markieren Sie für jede Frage die für Sie **am ehesten zutreffende Antwort**.
Bedenken Sie, dass es dabei weder richtige noch falsche Antworten gibt!

Sollten Situationen beschrieben sein, die Sie persönlich noch nicht erlebt haben, dann schätzen Sie bitte ein, wie Sie sich in solch einer Situationen am ehesten verhalten würden.

Einige der Aussagen könnten Ihnen in der Formulierung ähnlich vorkommen. Bitte lassen Sie sich davon nicht irritieren!
Und nun viel Spaß bei der Bearbeitung der Fragen!



[weiter...](#)

Die folgenden zwei Seiten erhielten lediglich die Testpersonen der **Versuchsgruppe 2**

Hinweis

Stellen Sie sich bitte vor, Sie haben sich für eine Stelle beworben!

Das Unternehmen hat Sie nun zu einem Kennenlern-Gespräch eingeladen und möchte auch eine Auswahltestung (Assessment-Center) mit Ihnen machen.
Bei Ankunft im Unternehmen müssen Sie feststellen, dass Sie nicht der einzige Bewerber für die Stelle sind.

**Während des Kennenlern-Gesprächs wird Ihnen klar, dass die Stelle nicht hundertprozentig Ihren Vorstellungen entspricht.
Sie finden die Stelle zwar nicht uninteressant, sind sich aber nicht sicher, ob sie den Job überhaupt haben möchten.**

Nach dem das Kennenlern-Gespräch beendet ist, werden Sie für das Assessment-Center gemeinsam mit einigen Mitbewerbern in einen Raum geführt.
Dort wird Ihnen ein Platz am Computer zugewiesen, mit der Bitte einen Fragebogen auszufüllen.
Auf dem Bildschirm erscheint die folgende Instruktion:

(bitte "weiter.."-Button anklicken)

weiter...

Herzlich Willkommen im Assessment-Center!

Es freut uns dass Sie Interesse an unserem Unternehmen haben!

Im folgenden Fragebogen finden Sie **287 Aussagen** zur Erfassung von Eigenschaften und **Verhaltensweisen im Arbeitsleben**.

Zur Beurteilung der Aussagen steht Ihnen eine 5-fach abgestufte Skala zur Verfügung.
Bitte markieren Sie für jede Frage die für Sie **am ehesten zutreffende Antwort**.
Bedenken Sie, dass es dabei weder richtige noch falsche Antworten gibt!

Sollten Situationen beschrieben sein, die Sie persönlich noch nicht erlebt haben, dann
schätzen Sie bitte ein, wie Sie sich in solch einer Situationen am ehesten verhalten
würden.

**Einige der Aussagen könnten Ihnen in der Formulierung ähnlich vorkommen. Bitte
lassen Sie sich davon nicht irritieren!**
Und nun viel Spaß bei der Bearbeitung der Fragen!



weiter...

Die folgenden zwei Seiten erhielten lediglich die Testpersonen der **Versuchsgruppe3**

Hinweis

Stellen Sie sich vor, Sie wollen nach erfolgreichem Studienabschluss eine professionelle Berufsberatung in Anspruch nehmen.

Im Rahmen dieser Berufsberatung soll ein Persönlichkeitsprofil Aufschluss darüber geben, welche persönlichkeitsbezogenen Stärken (Soft Skills) Sie in Hinblick auf eine zukünftige Berufstätigkeit mitbringen.

Dieses Persönlichkeitsprofil wird auf Basis eines von Ihnen ausgefüllten Persönlichkeitsfragebogens erstellt.

Ihnen ist klar, dass nur eine ehrliche Beantwortung des Fragebogens ein realistisches Bild Ihrer berufsbezogenen Persönlichkeit vermitteln kann. Nur dann sind die Ergebnisse, die Ihnen rückgemeldet werden, für Sie hilfreich.

Sie sind daher sehr bemüht, so ehrlich und realistisch wie möglich zu antworten.

Nachdem Ihnen ein Platz am Computer zugewiesen worden ist, werden Sie nun gebeten, den Persönlichkeitsfragebogen auszufüllen.

Auf dem Bildschirm erscheint die folgende Instruktion:

(bitte "weiter.."-Button anklicken)

weiter...

Herzlich Willkommen zur Berufsberatung!

Es freut uns, dass Sie sich die Zeit nehmen den folgenden Fragebogen auszufüllen!

Im folgenden Fragebogen finden Sie **287 Aussagen** zur Erfassung von Eigenschaften und **Verhaltensweisen im Arbeitsleben**.

Zur Beurteilung der Aussagen steht Ihnen eine 5-fach abgestufte Skala zur Verfügung.
Bitte markieren Sie für jede Frage die für Sie *am ehesten zutreffende Antwort*.
Bedenken Sie, dass es dabei weder richtige noch falsche Antworten gibt!

Sollten Situationen beschrieben sein, die Sie persönlich noch nicht erlebt haben, dann schätzen Sie bitte ein, wie Sie sich in solch einer Situationen am ehesten verhalten würden.

Einige der Aussagen könnten Ihnen in der Formulierung ähnlich vorkommen. Bitte lassen Sie sich davon nicht irritieren!
Und nun viel Spaß bei der Bearbeitung der Fragen!



[weiter...](#)

Bitte bewerten Sie die folgenden Statements...	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
1. Ich bin jemand, der eher selten mit anderen aneinander gerät.	☐	☐	☐	☐	☐
2. Wenn ich auf fremde Personen treffe, finde ich ohne Schwierigkeiten ein Gesprächsthema.	☐	☐	☐	☐	☐
3. Meine Arbeit stellt mich vor allem dann zufrieden, wenn ich Sie durch die Arbeit mit anderen erledigen kann.	☐	☐	☐	☐	☐
4. Ich grübele nicht lange über persönliche Probleme nach.	☐	☐	☐	☐	☐
5. Wenn ich extrem hart arbeiten muss, gerate ich auch bei zusätzlichen Schwierigkeiten nicht aus dem Gleichgewicht.	☐	☐	☐	☐	☐
6. Ich bin schlagfertig.	☐	☐	☐	☐	☐
7. Ich empfinde Genugtuung dabei, mit meinen Kräften bis an meine Grenzen zu gehen.	☐	☐	☐	☐	☐
8. Es ist mir in meinem Leben gelungen, eine ganze Menge zu bewegen.	☐	☐	☐	☐	☐
9. Auch Gespräche, in denen ich Einfluss auf andere nehmen muss, machen mir nur wenig aus.	☐	☐	☐	☐	☐
10. Andere haben es schwer, in einer Auseinandersetzung mit mir die Oberhand zu gewinnen.	☐	☐	☐	☐	☐
11. Bevor ich mich Freizeitaktivitäten zuwende, erledige ich alle anstehenden Arbeiten.	☐	☐	☐	☐	☐
12. Es macht mir nichts aus mich mit Aufgaben zu beschäftigen, bei denen ich im Vorfeld nicht ganz abschätzen kann, was mich erwartet.	☐	☐	☐	☐	☐
13. Meine Zeiteinteilung gelingt mir so, dass ich meine Aufgaben rechtzeitig erledige.	☐	☐	☐	☐	☐
14. Ich treffe in fast allen Situationen instinktiv den richtigen Ton.	☐	☐	☐	☐	☐
15. Es fällt mir nicht schwer, mich den Vorstellungen anderer anzupassen.	☐	☐	☐	☐	☐
16. Ich bin gerne mit Menschen zusammen, auch wenn ich Sie nicht so gut kenne.	☐	☐	☐	☐	☐
17. Man erzielt oftmals die besten Arbeitsergebnisse, wenn man im Team arbeitet.	☐	☐	☐	☐	☐
18. Wenn mir etwas nicht gelingt, stört mich das nach kurzer Zeit nicht mehr.	☐	☐	☐	☐	☐
19. Ich bleibe gelassen, auch wenn vieles gleichzeitig auf mich einströmt.	☐	☐	☐	☐	☐
20. Nach einer Begegnung mache ich mir eher wenig Gedanken darüber, wie ich auf mein Gegenüber gewirkt habe.	☐	☐	☐	☐	☐
21. Wenn ich meinen Lebensunterhalt nicht finanzieren müsste, würde ich deutlich weniger arbeiten.	☐	☐	☐	☐	☐
22. Ich bin ausgesprochen ehrgeizig.	☐	☐	☐	☐	☐
23. Auf einen großen Gestaltungsspielraum kann ich nur sehr schlecht verzichten, auch wenn ich ansonsten mit meiner Tätigkeit sehr zufrieden bin.	☐	☐	☐	☐	☐
24. Mir machen Entscheidungen nichts aus, auch wenn sie den Handlungsspielraum anderer Menschen einschränken.	☐	☐	☐	☐	☐
25. Meine Kollegen meinen, dass ich häufig versuche, meine Vorstellungen durchzusetzen.	☐	☐	☐	☐	☐

weiter...

	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
26. Ich plane lieber systematisch, anstatt spontanen Einfällen zu folgen.	☐	☐	☐	☐	☐
27. Wenn sich die Schwerpunkte meiner Tätigkeit völlig verändern, kann ich mich problemlos darauf einstellen.	☐	☐	☐	☐	☐
28. Ich zögere nicht mit der sofortigen Umsetzung von Entschlüssen.	☐	☐	☐	☐	☐
29. Ich bemerke nahezu immer, wenn sich jemand unwohl fühlt.	☐	☐	☐	☐	☐
30. Ich gehe mit anderen rücksichtsvoll um.	☐	☐	☐	☐	☐
31. Ich wäre froh, wenn es zu meinen beruflichen Aufgaben gehörte, immer wieder neue Personen anzusprechen.	☐	☐	☐	☐	☐
32. Mich stört es nicht, wenn ich mich bei meiner Tätigkeit mit anderen abstimmen muss.	☐	☐	☐	☐	☐
33. Wenn meine Arbeitsergebnisse nicht wie erhofft ausfallen, komme ich schnell darüber hinweg.	☐	☐	☐	☐	☐
34. Selbst in arbeitsintensiven Phasen führt die Beanspruchung bei mir zu keinerlei körperlichen Beschwerden.	☐	☐	☐	☐	☐
35. Ich vertrete meine Meinung auch dann sehr offen, wenn das Gesprächsklima darunter leidet.	☐	☐	☐	☐	☐
36. Meine Handlungen werden von anderen in der Regel richtig verstanden.	☐	☐	☐	☐	☐
37. Es macht mir wenig aus zu arbeiten, während andere ihren Freizeitaktivitäten nachgehen.	☐	☐	☐	☐	☐
38. Manche Kollegen denken, ich dränge zu stark auf Veränderungen, wenn ich mit ihnen zusammenarbeite.	☐	☐	☐	☐	☐
39. Es stellt mich zufrieden, wenn ich andere beeinflussen kann.	☐	☐	☐	☐	☐
40. Ich lasse mir nichts gefallen.	☐	☐	☐	☐	☐
41. Aufgaben, bei denen ich die Genauigkeit einer Arbeit kontrollieren muss, kommen meinen Fähigkeiten entgegen.	☐	☐	☐	☐	☐
42. Ich empfinde es als Herausforderung, wenn ich mit unvorhergesehenen Situationen konfrontiert werde.	☐	☐	☐	☐	☐
43. Die zielorientierte Analyse von Problemen, führt bei mir zu schnellen Handlungen.	☐	☐	☐	☐	☐
44. Ich kann mich auf die unterschiedlichsten Menschen sehr gut einstellen.	☐	☐	☐	☐	☐
45. Harmonie spielt für mich beim Erreichen meiner Ziele eine übergeordnete Rolle.	☐	☐	☐	☐	☐
46. Es fällt mir nicht schwer, fremde Personen anzusprechen, um mit Ihnen in ein Gespräch zu kommen.	☐	☐	☐	☐	☐
47. Ich ziehe es vor, im Team zu arbeiten.	☐	☐	☐	☐	☐
48. Wenn man mich kritisiert, denke ich nicht länger als nötig darüber nach.	☐	☐	☐	☐	☐
49. Ich verkrafte lang andauernde, hohe Belastungen besser als andere.	☐	☐	☐	☐	☐
50. Wenn andere hinter meinem Rücken über mich reden, lässt mich das kalt.	☐	☐	☐	☐	☐

weiter...

	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
51. Ich bin nur selten ohne Grund traurig.	☐	☐	☐	☐	☐
52. Es beschäftigt mich, wenn meine Kollegen größere Erfolge erzielen als ich.	☐	☐	☐	☐	☐
53. Ich bin mit meiner Leistung erst dann zufrieden, wenn ich die Erwartungen übertreffe.	☐	☐	☐	☐	☐
54. Wenn es darum geht, etwas Neues zu initiieren, bin ich in der Regel die richtige Person dafür.	☐	☐	☐	☐	☐
55. Ich wirke auf andere mitreißend.	☐	☐	☐	☐	☐
56. Es bereitet mir nur selten Schwierigkeiten, meine Ideen durchzusetzen.	☐	☐	☐	☐	☐
57. Ich interessiere mich für Aufgaben, die viel Ausdauer und Sorgfalt verlangen.	☐	☐	☐	☐	☐
58. Für mich ist es nicht sehr wichtig, den Ablauf eines Arbeitstages genau planen zu können.	☐	☐	☐	☐	☐
59. Wenn ich an einer dringenden Sache arbeite, unterbreche ich die Arbeit nur sehr selten, um etwas anderes zu tun.	☐	☐	☐	☐	☐
60. Ich finde auch zu sehr schwierigen Personen einen guten Draht.	☐	☐	☐	☐	☐
61. Mir macht es nichts aus, die selbe Sache mehrmals zu erklären.	☐	☐	☐	☐	☐
62. Wenn ich jemanden kennen lerne, bin ich recht schnell in der Lage, mich ihm gegenüber ungezwungen zu verhalten.	☐	☐	☐	☐	☐
63. Ich bin davon überzeugt, dass nahezu alle aktuellen Probleme nur im Team zu bewältigen sind.	☐	☐	☐	☐	☐
64. Wenn mir Fehler nachgewiesen werden, bin ich nur kurz betrübt.	☐	☐	☐	☐	☐
65. Auch wenn ich sehr hart arbeiten muss, bleibe ich gelassen.	☐	☐	☐	☐	☐
66. Ich besitze Eigenschaften, in denen ich den meisten Menschen überlegen bin.	☐	☐	☐	☐	☐
67. Es passiert kaum, dass Menschen meine Handlungen anders interpretieren als es meine Absicht war.	☐	☐	☐	☐	☐
68. Zu einem häufigen Wechsel des beruflichen Einsatzortes bin ich gern bereit.	☐	☐	☐	☐	☐
69. Ich stelle mich gern schwierigen Situationen, um festzustellen, wie gut ich bin.	☐	☐	☐	☐	☐
70. Manchmal muss man meinen Tatendrang bremsen.	☐	☐	☐	☐	☐
71. Eine Führungsaufgabe ist mir lieber als eine Spezialistentätigkeit.	☐	☐	☐	☐	☐
72. Ich kann andere dazu bringen, sich für meine Sache zu engagieren.	☐	☐	☐	☐	☐
73. Ich beschäftige mich gern mit Aufgaben, die große Präzision erfordern.	☐	☐	☐	☐	☐
74. Mir ist die klare Abgrenzung eines Aufgabenbereiches nicht so wichtig.	☐	☐	☐	☐	☐
75. Auch wenn ich ein Ziel nicht klar vor Augen habe, gelingt es mir strukturiert vorzugehen.	☐	☐	☐	☐	☐

weiter...

Vor Präsentation der nächsten Fragen (76. – 100.) erscheinen je nach Versuchsgruppe folgende drei Texte im Tabellenkopf:

Version 1 (Versuchsgruppe 1)

Vergessen Sie nicht, dass Sie in einer Bewerbungssituation sind! Sie geben ihr Bestes, um den Job zu bekommen!	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
--	---------------------	-------------------	---------	-------------------------	--------------------

Version 2 (Versuchsgruppe 2)

Vergessen Sie nicht, dass Sie in einer Bewerbungssituation sind!	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
--	---------------------	-------------------	---------	-------------------------	--------------------

Version 3 (Versuchsgruppe 3)

Vergessen Sie nicht, dass Sie nur dann von der Rückmeldung profitieren können, wenn Sie die Fragen ehrlich beantworten!	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
--	---------------------	-------------------	---------	-------------------------	--------------------

<i>Text in Abhängigkeit der Version gemäß vorheriger Seite</i>	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
76. Ich bemerke mit großer Sicherheit, wie sich mein Gegenüber fühlt.	☐	☐	☐	☐	☐
77. Ich versuche, mich besser anzupassen, wenn mein Verhalten nicht ankommt.	☐	☐	☐	☐	☐
78. Menschen, die in der Lage sind sehr schnell Kontakte zu knüpfen, irritieren mich nicht.	☐	☐	☐	☐	☐
79. Wenn man eine Aufgabe optimal erledigen will, ist die Teamarbeit unerlässlich.	☐	☐	☐	☐	☐
80. Mein Verhalten meinen Mitmenschen gegenüber verändert sich nur wenig, selbst wenn mich Probleme belasten.	☐	☐	☐	☐	☐
81. Ich kann problemlos viele Stunden ohne Pause arbeiten.	☐	☐	☐	☐	☐
82. Ich kann mit der Situation vor einer größeren Gruppe reden zu müssen sehr gut umgehen.	☐	☐	☐	☐	☐
83. Bei anderen hinterlasse ich in der Regel genau den Eindruck, den ich auch hinterlassen möchte.	☐	☐	☐	☐	☐
84. Ständige berufsbedingte Reisen machen mir nichts aus.	☐	☐	☐	☐	☐
85. Auch nach sehr guten Leistungen bemühe ich mich, noch besser zu werden.	☐	☐	☐	☐	☐
86. Ich kann mich nur schlecht zurücknehmen, wenn etwas nicht optimal läuft. Auch wenn ich nicht direkt betroffen bin.	☐	☐	☐	☐	☐
87. Ich strahle Autorität aus.	☐	☐	☐	☐	☐
88. Bei der Arbeit in einer Gruppe gelingt es mir problemlos, die anderen für meine Ideen zu gewinnen.	☐	☐	☐	☐	☐
89. Ich arbeite gerne mit Personen zusammen, die die Dinge eher genau erledigen wollen.	☐	☐	☐	☐	☐
90. Selbst wenn Aufgaben nicht ganz klar definiert sind, macht mir das nicht viel aus.	☐	☐	☐	☐	☐
91. Wenn ich zu Hause arbeite, fällt es mir leicht, pünktlich anzufangen.	☐	☐	☐	☐	☐
92. Ich trete eher selten in „Fettnäpfchen“.	☐	☐	☐	☐	☐
93. Selbst wenn die Dinge nicht ganz nach meiner Vorstellung laufen, bewahre ich die Ruhe.	☐	☐	☐	☐	☐
94. Ich schließe sehr schnell neue Bekanntschaften.	☐	☐	☐	☐	☐
95. Bei nahezu allen Aufgaben kann die Bearbeitung in Gruppen helfen.	☐	☐	☐	☐	☐
96. Selbst starke nervliche Belastungen werfen mich weniger aus der Bahn als die meisten anderen.	☐	☐	☐	☐	☐
97. Ich fühle mich den Anforderungen, die an mich gestellt werden, meistens gewachsen.	☐	☐	☐	☐	☐
98. Vor Begegnungen mit wichtigen Personen werde ich weniger nervös als andere.	☐	☐	☐	☐	☐
99. In schwierigen Zeiten bitte ich meine Kollegen und Freunde schon mal um Unterstützung.	☐	☐	☐	☐	☐
100. Gelegentlich vernachlässige ich durch das viele Arbeiten mein Privatleben.	☐	☐	☐	☐	☐

weiter...

	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
101. Es fällt mir schwerer als anderen, mich Gegebenheiten anzupassen, die ich für verbesserungswürdig halte.	☐	☐	☐	☐	☐
102. Ich bin sehr daran interessiert, eine leitende Position innezuhaben.	☐	☐	☐	☐	☐
103. Ich kann anderen gegenüber meine Meinung durchsetzen.	☐	☐	☐	☐	☐
104. Lieber gelte ich als unkreativ als dass man mir sagt, ich sei nicht zuverlässig und unorganisiert.	☐	☐	☐	☐	☐
105. Ich befasse mich gern mit Aufgaben, die schnelles Handeln verlangen.	☐	☐	☐	☐	☐
106. Selbst wenn viele Aufgaben gleichzeitig bearbeitet werden müssen, fällt es mir leicht die Dinge zu organisieren.	☐	☐	☐	☐	☐
107. Mir ist schnell klar, wie ich mich unbekannten Personen gegenüber verhalten sollte.	☐	☐	☐	☐	☐
108. Ich würde nicht offen zeigen, wenn ich Menschen nicht mag.	☐	☐	☐	☐	☐
109. Ich nutze einen großen Teil meiner Freizeit, um Kontakte zu pflegen.	☐	☐	☐	☐	☐
110. Meine Kollegen meinen, ich sei ein guter Teamworker.	☐	☐	☐	☐	☐
111. Auch wenn ich mir über wichtige Dinge Sorgen mache, bin ich in der Lage zielorientiert zu handeln.	☐	☐	☐	☐	☐
112. Im Vergleich zu anderen kann ich mir ungewöhnlich viel abverlangen, ohne dass ich mich verausgaben muss.	☐	☐	☐	☐	☐
113. Ich bin gelassen, wenn ich einer Gruppe von Leuten vorgestellt werde.	☐	☐	☐	☐	☐
114. Es ist mir sehr wichtig, zu den Besten zu gehören.	☐	☐	☐	☐	☐
115. Mir ist wichtig, dass sich mein berufliches Entgelt direkt an meine Leistungen knüpft.	☐	☐	☐	☐	☐
116. Es ist für mich wesentlich, durch meine Tätigkeit etwas Maßgebliches zu bewirken.	☐	☐	☐	☐	☐
117. Anderen Anweisungen zu geben, beeinträchtigt mein Wohlbefinden kaum.	☐	☐	☐	☐	☐
118. Es fällt mir leicht, andere von meiner Auffassung zu überzeugen.	☐	☐	☐	☐	☐
119. Ich nehme Dinge ganz genau.	☐	☐	☐	☐	☐
120. Mir macht es nichts aus von von meinem festen Tagesablauf abzurücken.	☐	☐	☐	☐	☐
121. Mit der Bearbeitung eines komplexen Problems zu beginnen, fällt mir meist nicht schwer.	☐	☐	☐	☐	☐
122. Es gelingt mir meist gut, mich auf Menschen einzustellen, die sehr unzugänglich sind.	☐	☐	☐	☐	☐
123. Man schätzt mich als eher warmherzig und wenig berechnend ein.	☐	☐	☐	☐	☐
124. Ich bin ein geselliger Mensch.	☐	☐	☐	☐	☐
125. Ich kann meine Fähigkeiten vor allem in der Zusammenarbeit mit anderen voll entfalten.	☐	☐	☐	☐	☐

weiter...

	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
126. In der Regel lasse ich mich nur schwer entmutigen.	☐	☐	☐	☐	☐
127. Selbst unter starkem Druck gelingt es mir, meine Gefühle in Zaum zu halten.	☐	☐	☐	☐	☐
128. Ich trete meist eher forsch auf, obwohl ein etwas zurückhaltenderes Auftreten manchmal von Vorteil wäre.	☐	☐	☐	☐	☐
129. Ich werde in der Regel so behandelt, wie es meinem Verhalten auch entspricht.	☐	☐	☐	☐	☐
130. Ich halte auch dann an Zielen fest, wenn sie schwer zu erreichen sind.	☐	☐	☐	☐	☐
131. Um eigene Wege zu beschreiten, verzichte ich auf den Rückhalt durch andere.	☐	☐	☐	☐	☐
132. In Besprechungen übernehme ich vielfach die Gesprächsführung, auch wenn ich nicht die Leitungsposition inne habe.	☐	☐	☐	☐	☐
133. Bei Auseinandersetzungen gewinne ich andere leicht für meine Position.	☐	☐	☐	☐	☐
134. Es lohnt sich viel Aufwand in eine Sache zu stecken, da man so auch häufig das hundertprozentige Ergebnis erzielt.	☐	☐	☐	☐	☐
135. Es ist mir angenehm, wenn bei einer Tätigkeit die Anforderungen häufig wechseln.	☐	☐	☐	☐	☐
136. Wenn ich etwas Unangenehmes zu erledigen habe, erledige ich es meist sofort.	☐	☐	☐	☐	☐
137. Ich merke sofort, wenn andere z.B. durch ironische Bemerkungen verletzt sind.	☐	☐	☐	☐	☐
138. Ich gelte als ein offener Mensch.	☐	☐	☐	☐	☐
139. Bei der Bearbeitung einer Aufgabe macht es mir nichts aus, die Hilfe anderer zu beanspruchen.	☐	☐	☐	☐	☐
140. Ich sehe vieles weniger dramatisch als andere.	☐	☐	☐	☐	☐
141. Auch gleichzeitigen Anforderungen von mehreren Seiten stehe ich meist gelassen gegenüber.	☐	☐	☐	☐	☐
142. Auch wenn mich jemand anderes ablehnt, nehme ich mir das nicht sehr zu Herzen.	☐	☐	☐	☐	☐
143. Mir liegt viel daran, bessere Ergebnisse als andere zu erzielen.	☐	☐	☐	☐	☐
144. Ich setze mir auch mal Ziele, deren Erreichung nicht hundertprozentig sicher ist.	☐	☐	☐	☐	☐
145. Ich kämpfe für meine Überzeugungen, auch wenn ich dafür Nachteile in Kauf nehmen muss.	☐	☐	☐	☐	☐
146. Ich kann mir ein erfülltes Berufsleben ohne die Wahrnehmung von Führungsverantwortung nicht vorstellen.	☐	☐	☐	☐	☐
147. Es macht mir wenig aus, mich bei anderen unbeliebt zu machen, um etwas Neues durchzusetzen.	☐	☐	☐	☐	☐
148. Ich bin Perfektionist.	☐	☐	☐	☐	☐
149. Mir macht es nichts aus Aufgaben zu übernehmen, bei denen sich mein Arbeitsumfeld ständig ändert.	☐	☐	☐	☐	☐
150. Wenn ich eine schwierige Aufgabe in Angriff nehme, kann ich mich alleinig auf diese konzentrieren, ohne dass mir andere Dinge durch den Kopf gehen.	☐	☐	☐	☐	☐

weiter...

Sie haben bereits mehr als die Hälfte der Fragen beantwortet! 😊	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
151. Ich finde auch in schwierigen Situationen das richtige Wort.	☐	☐	☐	☐	☐
152. In Gruppendiskussionen bin meist ich es, der ausgleichend wirkt.	☐	☐	☐	☐	☐
153. Ich kann besser auf Menschen zugehen als viele andere.	☐	☐	☐	☐	☐
154. Wenn ich etwas plane, überlege ich zunächst, wer noch bei dem Projekt	☐	☐	☐	☐	☐
155. Ich nehme das Leben im Allgemeinen eher locker.	☐	☐	☐	☐	☐
156. Auch wenn ich sehr viel zu tun habe, ändert sich an meiner Verhaltensweise gegenüber meiner Umwelt kaum etwas.	☐	☐	☐	☐	☐
157. Peinliche Situationen bringen mich weniger in Verlegenheit als die meisten anderen.	☐	☐	☐	☐	☐
158. Anderen gegenüber bin ich nur wenig misstrauisch.	☐	☐	☐	☐	☐
159. Ich bin unzufrieden, wenn ich mein Leistungspotenzial nicht voll ausgeschöpft habe.	☐	☐	☐	☐	☐
160. Andere orientieren sich an mir.	☐	☐	☐	☐	☐
161. Bei der Arbeit im Team setze ich mich nachhaltig dafür ein, andere von meinem Standpunkt zu überzeugen.	☐	☐	☐	☐	☐
162. Ich arbeite deutlich sorgfältiger als andere Personen.	☐	☐	☐	☐	☐
163. Wenn das meiste vorhersehbar ist, verliert die berufliche Tätigkeit für mich ihren Reiz.	☐	☐	☐	☐	☐
164. Was ich mir für den Tag vornehme, ist am Abend erledigt.	☐	☐	☐	☐	☐
165. Ich bin sehr sensibel für Veränderungen der Gesprächsatmosphäre.	☐	☐	☐	☐	☐
166. Wenn andere mir gegenüber nicht klar äußern, was sie wollen, macht mir das nicht viel aus.	☐	☐	☐	☐	☐
167. Meine Sicherheit im Umgang mit Menschen ist größer als die vieler anderer.	☐	☐	☐	☐	☐
168. Es ist Teil meines Arbeitsstils, vieles mit anderen zu diskutieren.	☐	☐	☐	☐	☐
169. Mich wirft so leicht nichts aus der Bahn.	☐	☐	☐	☐	☐
170. Ich kann mit einer hohen Arbeitsbelastung sehr gut umgehen.	☐	☐	☐	☐	☐
171. Ich bin selbstbewusst.	☐	☐	☐	☐	☐
172. Ich bin mit mir erst dann zufrieden, wenn ich außergewöhnliche Leistungen vollbringe.	☐	☐	☐	☐	☐
173. Für einige bin ich ein unbequemer Querdenker.	☐	☐	☐	☐	☐
174. Für mich sind fachliche Kompetenzen weniger wichtig als Führungsqualitäten.	☐	☐	☐	☐	☐
175. Es kommt vor, dass ich anderen gegenüber sehr dominant bin.	☐	☐	☐	☐	☐

weiter...

	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
176. Ich verwalte meine Unterlagen so, dass ich alles auf Anhieb wiederfinde.	☐	☐	☐	☐	☐
177. Es ist mir unangenehm, exakte Handlungsanweisungen zu erhalten.	☐	☐	☐	☐	☐
178. Es fällt mir nicht schwer, mit der Erledigung eiliger Sachen anzufangen.	☐	☐	☐	☐	☐
179. Ich bin meist in der Lage, im Gespräch den richtigen Ton zu treffen.	☐	☐	☐	☐	☐
180. Ich vermeide es, andere zu provozieren.	☐	☐	☐	☐	☐
181. Viele Menschen schätzen mich wegen meiner kontaktfreudigen Art.	☐	☐	☐	☐	☐
182. Wenn ich die Wahl habe, bearbeite ich Aufgabe lieber gemeinsam mit anderen.	☐	☐	☐	☐	☐
183. Ich habe ein ziemlich dickes Fell.	☐	☐	☐	☐	☐
184. Mich stört es nicht, einer Tätigkeit nachzugehen, bei der ich unter hohem Leistungsdruck stehe.	☐	☐	☐	☐	☐
185. Vor bedeutenden Ereignissen (z.B. wichtigen Gesprächen oder Präsentationen) habe ich mich gut im Griff.	☐	☐	☐	☐	☐
186. Ich spiele in Gedanken gern mit abstrakten Ideen.	☐	☐	☐	☐	☐
187. Es fällt mir nicht schwer, angebrachte Kritik auch zu kommunizieren.	☐	☐	☐	☐	☐
188. Ich plane meine Aktivitäten möglichst genau im Voraus.	☐	☐	☐	☐	☐
189. Ich fühle mich auch dann wohl, wenn die Dinge anders laufen als ich es gewohnt bin.	☐	☐	☐	☐	☐
190. Wenn ich viele Aufgaben zu erledigen habe, stelle ich schnell fest, womit ich am besten beginne.	☐	☐	☐	☐	☐
191. In der Regel kann ich sehr gut abschätzen, was andere von mir erwarten.	☐	☐	☐	☐	☐
192. Ich habe ein dichtes Netz an beruflichen Kontakten geknüpft.	☐	☐	☐	☐	☐
193. Wenn mir einige Misserfolge hintereinander widerfahren, bin ich trotzdem in der Lage mich wieder aufzuraffen und weiter zu machen.	☐	☐	☐	☐	☐
194. Es macht mir nichts aus, dass einige Leute mich ablehnen, wenn ich mich nicht nach ihnen richte.	☐	☐	☐	☐	☐
195. Mit anderen zu wetteifern bereitet mir Vergnügen.	☐	☐	☐	☐	☐
196. Ich bin bereit, zugunsten meines beruflichen Engagements Einschränkungen meines Privatlebens hinzunehmen.	☐	☐	☐	☐	☐
197. Ich trage gern die Verantwortung für wichtige Entscheidungen.	☐	☐	☐	☐	☐
198. Wenn ich vor völlig unerwarteten Situationen stehe, fühle ich mich richtig in meinem Element.	☐	☐	☐	☐	☐
199. Es fällt mir leicht, meine Prioritäten bei der Arbeit einzuhalten.	☐	☐	☐	☐	☐
200. Ich komme mit jedem gut klar.	☐	☐	☐	☐	☐

Sie haben bereits 200 von 287 Items beantwortet! Denken Sie daran, dass Sie nach Abschluss des Fragebogens ein Feedback erhalten. Es lohnt sich den Fragebogen auszufüllen!	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
201. Bei Schwierigkeiten wendet man sich häufig an mich, weil ich überall die richtigen Leute kenne.	☐	☐	☐	☐	☐
202. Vergangene Misserfolge belasten mich nicht mehr.	☐	☐	☐	☐	☐
203. Mir macht es nur wenig aus im Mittelpunkt der Aufmerksamkeit zu stehen.	☐	☐	☐	☐	☐
204. Mich reizen besonders Probleme, die sehr schwer zu lösen sind.	☐	☐	☐	☐	☐
205. In Situationen, in denen die Leitung einer Gruppe erforderlich ist, macht es mir nichts aus, mich dieser Aufgabe anzunehmen.	☐	☐	☐	☐	☐
206. Ich bin bei Kontrollen lieber penibel als großzügig.	☐	☐	☐	☐	☐
207. Mit mir kann man Konflikte konstruktiv beilegen.	☐	☐	☐	☐	☐
208. Nach der Begegnung mit wichtigen Personen bemühe ich mich, den Kontakt aufrecht zu erhalten.	☐	☐	☐	☐	☐
209. Ich bin frei von Ängsten.	☐	☐	☐	☐	☐
210. Spannungen mit Kollegen kann ich gut ertragen.	☐	☐	☐	☐	☐

weiter...

	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
211. Ich bin leicht beunruhigt.	☐	☐	☐	☐	☐
212. Ich habe gerne viele Leute um mich herum.	☐	☐	☐	☐	☐
213. Ich schwelge gerne in Tagträumereien.	☐	☐	☐	☐	☐
214. Ich versuche zu jedem, dem ich begegne, freundlich zu sein.	☐	☐	☐	☐	☐
215. Ich halte meine Sachen ordentlich und sauber.	☐	☐	☐	☐	☐
216. Ich fühle mich anderen oft unterlegen.	☐	☐	☐	☐	☐
217. Ich bin leicht zum Lachen zu bringen.	☐	☐	☐	☐	☐
218. Ich finde philosophische Diskussionen spannend.	☐	☐	☐	☐	☐
219. Ich bekomme nur selten Streit mit meiner Familie und meinen Kollegen.	☐	☐	☐	☐	☐
220. Ich kann mir meine Zeit recht gut einteilen, so dass ich meine Angelegenheiten rechtzeitig beende.	☐	☐	☐	☐	☐
221. Wenn ich unter starkem Stress stehe, fühle ich mich manchmal, als ob ich zusammenbräche.	☐	☐	☐	☐	☐
222. Ich halte mich eher für fröhlich.	☐	☐	☐	☐	☐
223. Mich begeistern Motive, die ich in der Kunst und in der Natur finde.	☐	☐	☐	☐	☐
224. Ich kann mir nicht vorstellen dass mich jemand für selbstsüchtig und selbstgefällig halten könnte.	☐	☐	☐	☐	☐
225. Ich bin ein sehr systematisch vorgehender Mensch.	☐	☐	☐	☐	☐
226. Ich fühle mich öfters einsam oder traurig.	☐	☐	☐	☐	☐
227. Ich unterhalte mich wirklich gerne mit anderen Menschen.	☐	☐	☐	☐	☐
228. Ich glaube, dass es für Schüler von Vorteil ist, wenn man sie Rednern zuhören lässt, die kontroverse Standpunkte vertreten.	☐	☐	☐	☐	☐
229. Ich würde lieber mit anderen zusammenarbeiten, als mit Ihnen zu wetteifern.	☐	☐	☐	☐	☐
230. Ich versuche, alle mir übertragenen Aufgaben sehr gewissenhaft zu erledigen.	☐	☐	☐	☐	☐
231. Ich fühle mich oft angespannt und nervös.	☐	☐	☐	☐	☐
232. Ich bin gerne im Zentrum des Geschehens.	☐	☐	☐	☐	☐
233. Poesie beeindruckt mich ab und zu sehr.	☐	☐	☐	☐	☐
234. Im Hinblick auf die Absichten anderer bin ich wenig zynisch und skeptisch.	☐	☐	☐	☐	☐
235. Ich habe eine Reihe von klaren Zielen und arbeite systematisch auf sie zu.	☐	☐	☐	☐	☐

weiter...

Vor Präsentation der nächsten Fragen (263.-260.) erscheinen je nach Versuchsgruppe folgende drei Texte im Tabellenkopf:

Version 1 (Versuchsgruppe 1)

Sie sind in einem Assessment-Center! Bedenken Sie, dass Sie die Stelle unbedingt haben wollen!	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
--	---------------------	-------------------	---------	-------------------------	--------------------

Version 2 (Versuchsgruppe 2)

Bedenken Sie, dass Sie diesen Fragebogen im Rahmen eines Assessment-Center ausfüllen!	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
--	---------------------	-------------------	---------	-------------------------	--------------------

Version 3 (Versuchsgruppe 3)

Bitte versuchen Sie die Fragen so ehrlich wie möglich zu beantworten!	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
--	---------------------	-------------------	---------	-------------------------	--------------------

<i>Text in Abhängigkeit der Version gemäß vorheriger Seite</i>	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
236. Manchmal fühle ich mich völlig wertlos.	☐	☐	☐	☐	☐
237. Ich ziehe es gewöhnlich vor, Dinge gemeinsam statt allein zu tun.	☐	☐	☐	☐	☐
238. Ich probiere oft neue und fremde Speisen aus.	☐	☐	☐	☐	☐
239. Ich glaube, dass die meisten Leute es gut mit einem meinen und nicht versuchen andere auszunutzen.	☐	☐	☐	☐	☐
240. Ich beginne sofort mit einer Arbeit.	☐	☐	☐	☐	☐
241. Ich empfinde manchmal Furcht oder Angst.	☐	☐	☐	☐	☐
242. Ich habe oft das Gefühl vor Energie überzuschäumen.	☐	☐	☐	☐	☐
243. Ich nehme die Stimmungen oder Gefühle, die verschiedene Umgebungen hervorrufen oft wahr.	☐	☐	☐	☐	☐
244. Die meisten Menschen, die ich kenne, mögen mich.	☐	☐	☐	☐	☐
245. Ich arbeite hart, um meine Ziele zu erreichen.	☐	☐	☐	☐	☐
246. Ich ärgere mich oft darüber, wie andere Leute mich behandeln.	☐	☐	☐	☐	☐
247. Ich bin ein fröhlicher, gut gelaunter Mensch.	☐	☐	☐	☐	☐
248. Ich glaube, dass wir bei ethischen Entscheidungen nicht so sehr auf die Ansichten unserer religiösen Autoritäten achten sollten.	☐	☐	☐	☐	☐
249. Die meisten Leute halten mich für warmherzig und wenig berechnend.	☐	☐	☐	☐	☐
250. Wenn ich eine Verpflichtung eingehe, so kann man sich auf mich bestimmt verlassen.	☐	☐	☐	☐	☐
251. Zu häufig bin ich entmutigt und will aufgeben, wenn etwas schief geht.	☐	☐	☐	☐	☐
252. Ich bin ein gut gelaunter Optimist.	☐	☐	☐	☐	☐
253. Wenn ich Literatur lese oder ein Kunstwerk betrachte, empfinde ich manchmal ein Frösteln oder eine Welle der Begeisterung.	☐	☐	☐	☐	☐
254. In Bezug auf meine Einstellungen bin ich nur selten nüchtern und unnachgiebig.	☐	☐	☐	☐	☐
255. In der Regel bin ich so verlässlich oder zuverlässig wie ich es sein sollte.	☐	☐	☐	☐	☐
256. Ich bin manchmal traurig oder deprimiert.	☐	☐	☐	☐	☐
257. Ich führe ein hektisches Leben.	☐	☐	☐	☐	☐
258. Es interessiert mich, über die Natur des Universums oder die Lage der Menschheit zu spekulieren.	☐	☐	☐	☐	☐
259. Ich versuche stets rücksichtsvoll und sensibel zu handeln.	☐	☐	☐	☐	☐
260. Ich bin eine tüchtige Person, die ihre Arbeit immer erledigt.	☐	☐	☐	☐	☐

weiter...

Sie haben es fast geschafft. Dies ist die letzte Seite! 😊	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
261. Ich fühle mich oft hilflos und wünsche mir eine Person, die meine Probleme löst.	☐	☐	☐	☐	☐
262. Ich bin ein sehr aktiver Mensch.	☐	☐	☐	☐	☐
263. Ich bin sehr wissbegierig.	☐	☐	☐	☐	☐
264. Auch wenn ich Menschen nicht mag, würde ich Ihnen das nicht offen zeigen.	☐	☐	☐	☐	☐
265. In meinem Leben herrscht Ordnung.	☐	☐	☐	☐	☐
266. Manchmal ist mir etwas so peinlich, dass ich mich am liebsten verstecken würde.	☐	☐	☐	☐	☐
267. Lieber würde ich eine Gruppe anführen, als allein meinen eigenen Weg zu gehen.	☐	☐	☐	☐	☐
268. Ich habe oft Spaß daran, mit Theorien oder abstrakten Ideen zu spielen.	☐	☐	☐	☐	☐
269. Auch wenn es mir um etwas wichtiges ginge, würde ich niemanden manipulieren.	☐	☐	☐	☐	☐
270. Bei allem, was ich tue, strebe ich nach Perfektion.	☐	☐	☐	☐	☐
271. Manchmal werfe ich Müll einfach in die Landschaft oder auf die Straße.	☐	☐	☐	☐	☐
272. Eigene Fehler gebe ich stets offen zu und ertrage gelassen etwaige negative Konsequenzen.	☐	☐	☐	☐	☐
273. Im Straßenverkehr nehme ich stets Rücksicht auf die anderen Verkehrsteilnehmer.	☐	☐	☐	☐	☐
274. Ich habe schon einmal Drogen (Tabletten, Haschisch oder "ähnliches") konsumiert.	☐	☐	☐	☐	☐
275. Ich akzeptiere alle anderen Meinungen, auch wenn sie mit meiner eigenen nicht übereinstimmen.	☐	☐	☐	☐	☐
276. Meine Wut oder schlechte Laune lasse ich hin und wieder an unschuldigen oder schwächeren Leuten aus.	☐	☐	☐	☐	☐
277. Ich habe schon einmal jemanden ausgenutzt oder übers Ohr gehauen.	☐	☐	☐	☐	☐
278. In einem Gespräch lasse ich den anderen stets ausreden und höre ihm aufmerksam zu.	☐	☐	☐	☐	☐
279. Ich zögere niemals, jemandem in einer Notlage beizustehen.	☐	☐	☐	☐	☐
280. Wenn ich etwas versprochen habe, halte ich es ohne Wenn und Aber.	☐	☐	☐	☐	☐
281. Ich lästere gelegentlich über andere hinter deren Rücken.	☐	☐	☐	☐	☐
282. Ich würde niemals auf Kosten der Allgemeinheit leben.	☐	☐	☐	☐	☐
283. Ich bleibe immer freundlich und zuvorkommend anderen Leuten gegenüber, auch wenn ich selbst gestreßt bin.	☐	☐	☐	☐	☐
284. Im Streit bleibe ich stets sachlich und objektiv.	☐	☐	☐	☐	☐
285. Ich habe schon einmal geliehene Sachen nicht zurückgegeben.	☐	☐	☐	☐	☐
286. Ich ernähre mich stets gesund.	☐	☐	☐	☐	☐
287. Manchmal helfe ich nur, weil ich eine Gegenleistung erwarte.	☐	☐	☐	☐	☐

weiter...

Herzlichen Dank für das vollständige Ausfüllen des Fragebogens!
Ihre Rückmeldung erhalten Sie mittels Link auf der letzten Seite.

Bitte gehen Sie nun zur nächsten Seite um letzte Angaben zu machen!



[weiter...](#)

Bitte vervollständigen Sie nun noch die folgenden Angaben:

Alter in Jahren:

Geschlecht:

Studiengang:

Studiendauer in Semestern:

Berufserfahrung

Höchste abgeschlossene Ausbildung:

Nationalität:

weiter...

Bevor Sie nun die Rückmeldung erhalten, bewerten Sie bitte noch die folgenden drei Aussagen zum Fragebogen:	trifft völlig zu	trifft eher zu	neutral	trifft eher nicht zu	trifft nicht zu
Mir fiel es sehr leicht, mich in die beschriebene Bewerbungssituation hinein zu versetzen.	☐	☐	☐	☐	☐
Ich habe mich beim Ausfüllen des Fragebogens eher bemüht, einen besonders guten Eindruck zu hinterlassen als ehrlich zu antworten.	☐	☐	☐	☐	☐
Ich habe mich beim Ausfüllen des Fragebogens eher bemüht, ehrlich zu antworten.	☐	☐	☐	☐	☐

weiter...



Vielen Dank für die Teilnahme an dieser Studie!

Sollten sich ihrerseits Fragen ergeben, so wenden Sie sich bitte an **Frau Victoria Albertz** (a0542736@unet.univie.ac.at).

Ihre Rückmeldung erhalten Sie indem Sie folgenden Link anklicken:

[Klicken Sie hier, um Ihre persönliche Auswertung anzusehen und/oder auszudrucken \(PDF\).](#)

[Alles beantwortet - Umfrage schliessen](#)

2.2 Online-Fragebogen - Rückmeldung

Ihre persönliche Auswertung

Leistungsmotivation

Sie haben x von 10 Punkten

Personen mit **hoher Ausprägung** in dieser Skala stellen enorme Anforderungen an die eigene Leistung und nehmen auch überdurchschnittliche Anstrengungen auf sich, um tatsächlich zu erreichen, was Sie sich vorgenommen haben. Es wird auch dann an Zielen festgehalten, wenn erkennbar ist, dass sich diese nur schwer erreichen lassen. Diese Personen werden in erheblichem Maße dadurch motiviert, dass ihnen die Möglichkeit eröffnet wird, außergewöhnliche berufliche Leistungen zu erbringen. Im Extremfall werden die eigenen Ziele fortwährend nach oben korrigiert, um die Herausforderung aufrecht zu halten. Sehr leistungsmotivierte Personen sind häufig mit ihren Ergebnissen nur kurzfristig zufrieden und streben nach einer ständigen Optimierung ihrer

Kompetenzen. Sie suchen im Allgemeinen solche beruflichen Kontexte auf, in denen viel Spielraum "nach oben hin" besteht und sich die Ergebnisse des eigenen Engagements möglichst genau nachvollziehen lassen. Diese Personen können im Berufsleben bisweilen durch folgendes Motto beschrieben werden: Höher, schneller, weiter.

Personen mit **niedrigen Ausprägungen** in dieser Skala messen dem beständigen Ausschöpfen der eigenen Leistungsreserven kein positives Gewicht bei. Sie könnten einige Aufgaben besser bewältigen, verfügen aber nicht über den Ehrgeiz, stets höchsten Ansprüchen zu genügen. Wenn absehbar ist, dass sich ein Ziel nur schwer erreichen lässt, wird nicht weiter daran festgehalten. Stattdessen erfolgt gegebenenfalls eine Neuorientierung oder Prioritätenverschiebung. Das Vollbringen außergewöhnlicher beruflicher Leistungen ist für diese Personen kein zentraler Anreiz. Im wirtschaftlichen Kontext werden sie bisweilen dadurch beschrieben, dass Ihnen der nötige "Drive" fehle.

Gestaltungsmotivation

Sie haben x von 10 Punkten

Für Personen mit **hohen Werten** auf dieser Skala ist es von großer Bedeutung, sich aktiv an einer Veränderung und Gestaltung ihrer Umgebung zu beteiligen. Dementsprechend wird eine berufliche Tätigkeit mit hohem Gestaltungsspielraum angestrebt. Ideen werden mit hohem Engagement verfolgt. Gestaltungsmotivierte Personen sind bereit, auch gegen deutliche Widerstände anzugehen, wenn sie eine Aufgabenstellung begeistert. Es ist sehr motivierend für Sie, wenn es ihnen gelingt, Missstände zu beseitigen, etwas Neues durchzusetzen und eigene Vorstellungen zu verwirklichen. Die Möglichkeit der Machtausübung stellt für diese Personen einen wichtigen Anreiz für berufliche Leistungen und großen beruflichen Einsatz dar.

Personen mit **niedrigen Skalenwerten** verspüren wenig inneren Antrieb, sich für Veränderungen und Personen mit niedrigen Skalenwerten verspüren wenig inneren Antrieb, sich für Veränderungen und Umstrukturierungen ihrer Umgebung zu engagieren. Ihre Stärke liegt eher in der Kontinuität und Wahrung des Bestehenden. Große Einflussmöglichkeiten auf Prozesse in ihrer Tätigkeit

stellen keinen bedeutsamen motivierenden Faktor dar. Es bereitet ihnen keine Schwierigkeiten, sich in bestehende Strukturen einzufügen und im Rahmen von vorgegebenen Grenzen zu agieren. Die Möglichkeit, auf das Umfeld einzuwirken und Prozesse in ihrem Sinne zu beeinflussen, übt wenig Anziehungskraft auf Sie aus. Macht ist für Sie ein wenig positiv besetzter Begriff und Spielräume zur Machtausübung werden nicht aktiv aufgesucht und ausgebaut.

Führungsmotivation

Sie haben x von 10 Punkten

Für Personen mit **hoher Ausprägung** dieser Dimension zählt es zu ihren beruflichen Zielsetzungen, die Tätigkeit anderer anzuleiten und zu koordinieren. Sie geben ohne zu zögern Anweisungen. In Gruppen sehen Sie sich gern in der Leitungsfunktion und genießen es, andere für ihre Auffassungen zu begeistern und für ihren Standpunkt zu gewinnen. Sie schreiben sich die für Führungskräfte typischen Merkmale zu, namentlich etwa andere zu begeistern oder Orientierung zu stiften. Sie erleben sich in der sozialen Einflussnahme als stark und kompetent. Sie erwarten, dass man Ihnen für gewöhnlich folgt.

Für Personen mit **niedrig ausgeprägten Werten** stellt die Tätigkeit andere Menschen zu führen, keinen nachhaltigen Anreiz dar. Fachliche Kompetenzen stehen dem gegenüber vermehrt im Mittelpunkt ihres Engagements. Es widerstrebt ihnen, in den Handlungsspielraum anderer einzugreifen oder sich in Gruppen um die Führung zu bemühen. Beim Erteilen von Anweisungen fühlen sie sich „nicht ganz wohl in ihrer Haut“. In Arbeitsgruppen werden persönliche Fähigkeiten eher zur Bewältigung von fachlich anspruchsvollen Aufgaben eingesetzt. Bei zahlreichen Personen mit niedrigen Werten auf dieser Skala geht das geringere Führungsmotiv mit einer hohen Wertschätzung für fachlich anspruchsvolle Aufgaben einher.

Gewissenhaftigkeit

Sie haben x von 10 Punkten

Personen mit **hohen Ausprägungen** in dieser Skala erweisen sich bei der Planung und Ausführung von Tätigkeiten höchst verlässlich. Sie versuchen, jede begonnene Aufgabe so gründlich und präzise wie möglich zu bearbeiten. Dabei ist ihnen das korrekte Einhalten von Vereinbarungen und Fristen sehr wichtig. Sorgfalt ist für sie von großer Bedeutung, auch wenn sie häufig viel Zeit und Energie investieren müssen, damit die Arbeitsergebnisse ihren persönlichen Standards genügen. Sie bleiben auch dann noch ausdauernd bei der Sache, wenn es auf Einzelheiten und Details ankommt. Bei der Zusammenarbeit mit anderen achten Sie ebenfalls auf deren Genauigkeit und Zuverlässigkeit. Sie vertiefen sich gern in Einzelheiten und fundierte Lösungen von Problemen.

Personen mit **niedrigen Skalenwerten** investieren nur ungern viel Aufwand in die sehr präzise Ausführung von Aufgaben. Wenn es auf Einzelheiten und Geduld ankommt, verlieren die Tätigkeiten für sie relativ schnell an Attraktivität. Sie schätzen statt dessen eher pragmatische Lösungen und sind der Ansicht, dass nicht jedes Arbeitsergebnis vollkommen sein muss. Spontaneität ist Ihnen wichtiger als das exakte Einhalten von Vereinbarungen und Fristen. Aufgaben,

die über lange Zeiträume hinweg mit hohen Anforderungen an Ausdauer Vereinbarungen und Fristen. Aufgaben, die über lange Zeiträume hinweg mit hohen Anforderungen an Ausdauer und Sorgfalt verbunden sind, werden weniger gern bearbeitet. Sie engagieren sich bevorzugt für Tätigkeiten, bei denen ohne langwierige Vorbereitungen gehandelt werden kann.

Flexibilität

Sie haben x von 10 Punkten

Personen mit **hoher Ausprägung** in dieser Dimension fällt es leicht, sich auf wechselnde Bedingungen einzustellen. Sie schätzen es sehr, mit immer neuen Herausforderungen konfrontiert zu werden. Unbekanntes verursacht bei ihnen keinerlei Unbehagen, denn sie verfügen über Improvisationstalent und sind stets offen für neue Erfahrungen. Sie werden vergleichsweise wenig dadurch beeinträchtigt, dass Aufgaben nicht klar definiert sind, da sie ein hohes Maß an Ungewissheit tolerieren können und neuen Situationen zuversichtlich begegnen. Personen mit **niedrigen Werten** in dieser Dimension fällt es eher schwer, sich auf wechselnde Bedingungen einzustellen. Da sie sich in einem stabilen Umfeld wohler fühlen, sagt ihnen eine ständige Konfrontation mit neuen Situationen nicht zu. In ihrer Tätigkeit bevorzugen sie definierte Aufgaben, die nicht mit einem hohen Maß an Uneindeutigkeit behaftet sind – bei Herausforderungen dieser Art arbeiten sie am effektivsten. Auf ihr Arbeitsumfeld wirken sie damit stabilisierend. Dagegen fühlen Sie sich bei beruflichen Aufgaben vielfach nicht wohl, die durch ein hohes Maß an Ungewissheit charakterisiert sind.

Handlungsoptimierung

Sie haben x von 10 Punkten

Personen mit **hohen Skalenwerten** gehen ihre Aufgaben außergewöhnlich rasch und zielorientiert an. Wenn eine Entscheidung für ein bestimmtes Vorgehen getroffen wurde, beginnen Sie unverzüglich mit der Umsetzung ihrer Vorstellungen. Dabei sind sie in der Lage, ihre Aufmerksamkeit auf die relevanten Aspekte zu richten. Sie werden als Menschen wahrgenommen, die sich bei Schwierigkeiten nicht lange den Kopf zerbrechen, sondern zügig handeln. Sie setzen sich Ziele, die sie ausdauernd verfolgen und verfügen dabei über ein hohes Maß an effektiver Selbstorganisation. Komplexe Aufgaben bewältigen Sie durch eine Gliederung in erreichbare Teilziele. Personen mit **niedrigen Werten** beschreiben sich als Menschen, die häufig unsicher hinsichtlich des optimalen Vorgehens bei der Bewältigung ihrer Aufgaben sind. Wenn sie eine Entscheidung für eine bestimmte Handlungsalternative getroffen haben, zögern sie eine gewisse Zeit bis sie mit der Umsetzung beginnen. Es fällt ihnen manchmal schwer, die Aufmerksamkeit auf die relevanten Aspekte zu richten. Besonders bei unangenehmen Aufgaben besteht die Tendenz, die Bearbeitung hinauszuschieben. Für wenig handlungsorientierte Menschen ist es wichtig, sich immer wieder überschaubare und erreichbare Teilziele zu setzen - damit wird Ihnen auch die Bewältigung von komplexen oder eher unangenehmen Tätigkeiten leichter gelingen.

Sensitivität

Sie haben x von 10 Punkten

Personen mit **hohen Werten** beschreiben sich als Menschen, die im Kontakt zu anderen über ein feines Gespür für deren Stimmungen verfügen. Sie finden schnell heraus wie diffizile Gesprächssituationen zu meistern sind. Dies ermöglicht ihnen, sich auch auf unzugängliche Menschen einzustellen, eine Vielzahl unterschiedlicher Situationen angemessen zu interpretieren und ihr Verhalten darauf abzustimmen.

Personen mit **niedrigen Werten** sind in einer Reihe sozialer Situationen eher unsicher in der Wahrnehmung der Befindlichkeit ihrer Gesprächspartner und der Einschätzung angemessenen Verhaltens. Ihnen ist nicht immer deutlich, wie ihre Handlungen von anderen aufgefasst und interpretiert werden. Ihnen ist nicht immer deutlich, wie ihre Handlungen von anderen aufgefasst und interpretiert werden.

Kontaktfähigkeit

Sie haben x von 10 Punkten

Personen mit **hohen Werten** bereitet es keinerlei Schwierigkeiten, auf andere zuzugehen, Kontakte mit Unbekannten herzustellen und ein Netz an persönlichen Beziehungen aufzubauen. Sie verfügen über große Sicherheit und Unbefangenheit im Umgang mit anderen Menschen. Sowohl im beruflichen wie auch im privaten Umfeld sind sie offen für den persönlichen Austausch und schätzen es, mit zahlreichen Personen in Verbindung zu stehen. Ihre zahlreichen Kontakte ermöglichen es Ihnen, bei vielfältigen Problemen einen geeigneten Ansprechpartner zu finden.

Personen mit **niedrigen Skalenwerten** verhalten sich in sozialen Situationen eher zurückhaltend und gehen seltener von sich aus auf andere Menschen zu. Es dauert in der Regel eine Weile, bis sie neue Bekanntschaften knüpfen. Dies kann einerseits darin begründet sein, dass ihnen ein kleiner Kreis von engen Freunden und Bekannten genügt. Eine zweite Facette liegt in einer gewissen Unsicherheit, so dass sie auf andere reserviert und zurückhaltend wirken. Neben dem Zugehen auf andere Menschen fällt diesen Personen auch der Aufbau persönlicher Netzwerke, auf die Sie bei Problemen zurückgreifen können, nicht leicht.

Soziabilität

Sie haben x von 10 Punkten

Für Personen mit **hohen Werten** ist es bedeutsam, von anderen als freundlich und rücksichtsvoll wahrgenommen zu werden. Sie schätzen ein harmonisches Miteinander und bemühen sich, ausgleichend und integrierend auf ihr Umfeld einzuwirken. Wenn sie andere Menschen kritisieren müssen, so tun sie dies vorzugsweise indirekt, weil sie vermeiden möchten ihren Gesprächspartner vor den Kopf zu stoßen. Es fällt Ihnen schwer, Unangenehmes offen anzusprechen, wenn Sie dadurch Ärger oder Missmut auf sich ziehen könnten. Sie treten

anderen Menschen in der Regel mit Wohlwollen gegenüber und werden daher als unterstützend und liebenswürdig erlebt.

Personen mit **niedrigen Werten** gehören nicht zu den Menschen, die überall beliebt sein möchten. Es ist für Sie wenig bedeutsam, von anderen als angenehm im Umgang und stets rücksichtsvoll wahrgenommen zu werden. Kritik sprechen sie offen aus und nehmen dabei in Kauf, andere vor den Kopf zu stoßen. Sie tolerieren es, Missmut auf sich zu ziehen, denn für sie führt häufig nicht Harmonie, sondern Auseinandersetzung zum Ziel. Auf andere Menschen wirken sie markant oder provokant. Man weiß von Ihnen, dass Sie ehrlich aussprechen was Sie denken und nicht zu den Personen gehören, die anderen schmeicheln um Sympathien zu erlangen.

Teamorientierung

Sie haben x von 10 Punkten

Personen mit **hoch ausgeprägten Werten** sind in hohem Maße kooperationsbereit und bewerten die Zusammenarbeit mit anderen als wertvoll. Für die Durchsetzung von im Team getroffenen Entscheidungen setzen Sie sich aktiv ein und sind gern bereit, Kompetenzen und Entscheidungsbefugnisse auf die Gruppe bzw. einzelne Mitglieder zu übertragen oder mit diesen zu teilen. Sie neigen dazu, Teamleistungen gegenüber addierten Einzelleistungen für überlegen zu halten. Der Austausch mit anderen wirkt anregend auf sie.

Personen mit **niedriger Ausprägung** ist hohe Autonomie bei der Arbeit sehr wichtig. Sie möchten nicht auf die Unterstützung anderer angewiesen sein. Sie sind der Auffassung, dass Teamarbeit der Arbeit Einzelner nicht immer überlegen ist. Niedrig teamorientierte Personen nehmen Arbeitsgruppen bisweilen als ein Auffangbecken wahr, in dem weniger leistungsfähige Mitarbeiter von den Leistungen der Stärkeren profitieren. Sie betonen zwar die Wichtigkeit von Kooperationen, sind aber in der Regel trotzdem der Auffassung, dass sie viele Aufgaben am besten allein lösen könnten. Insgesamt betrachtet sind Sie eher "Einzelkämpfer", die zumindest in ihrer beruflichen Tätigkeit selbstständig und unabhängig agieren wollen.

Durchsetzungsstärke

Sie haben x von 10 Punkten

Personen mit **hohen Werten** auf dieser Skala beschreiben sich häufig als dominant im Umgang mit anderen. In Diskussionen sind sie offensiv und setzen sich nachhaltig für ihren Standpunkt ein. Ihre hartnäckige Argumentation führt häufig dazu, dass sie bei der Auseinandersetzung ihre Ziele durchsetzen können. Vor allem Menschen die sich nicht in gleichem Ausmaß Verhör schaffen können, nehmen diese Menschen nicht selten als sehr überzeugt von eigenen Vorstellungen und mitunter auch als autoritär oder kompromisslos wahr. Vor allem in schwierigen Führungssituationen, kann eine hohe Ausprägung hier durchaus hilfreich sein.

Personen mit **niedriger Ausprägung** versuchen, ihre Ziele durch Kompromisse, nicht durch Dominanz und Autorität zu erreichen. Sie neigen nicht dazu, andere Menschen zu dominieren und in Gruppensituationen für die bedingungslose

Durchsetzung ihrer Auffassungen zu kämpfen. Wenn eine Einigung ausgehandelt werden muss, tendieren sie zur Nachgiebigkeit und beharren nicht auf ihrem Standpunkt. In Situationen, in denen Sie andere von ihren Ideen überzeugen wollen, geben Sie rascher auf, was ihrem Anliegen häufig abträglich ist.

Emotionale Stabilität

Sie haben x von 10 Punkten

Personen mit **hoher Ausprägung** zeichnen sich im Umgang mit Rückschlägen, Misserfolgen und persönlichen Problemen durch ein hohes Maß an Gelassenheit aus. Sie kommen schnell über Rückschläge hinweg und es gelingt Problemen durch ein hohes Maß an Gelassenheit aus. Sie kommen schnell über Rückschläge hinweg und es gelingt ihnen gut, sich nach Niederlagen erneut zu motivieren. Sie verfügen über eine optimistische Lebensauffassung und fühlen sich durch intensive negative Emotionen nicht nachhaltig beeinträchtigt.

Personen mit **niedrigen Werten** benötigen eine gewisse Zeit, um über Niederlagen und Misserfolge hinweg zu kommen. Sie fühlen sich öfter entmutigt, unzulänglich und manchmal überfordert. Wenn sie bedrückt sind, fällt ihnen die Bewältigung ihrer Aufgaben mitunter schwer. Vielen Menschen gelingt es jedoch in angemessener Art und Weise, ihre Schwierigkeiten in diesem Bereich so zu kompensieren, dass keine Minderleistungen oder sonstige Problemsituationen in Erscheinung treten.

Belastbarkeit

Sie haben x von 10 Punkten

Personen mit **hohen Werten** können Belastungen sehr gut verkraften. Es beeinträchtigt sie wenig, sich über längere Zeiträume hinweg auch kontinuierlich hohen Anforderungen stellen zu müssen. Sie verfügen über große Energiereserven. In arbeitsintensiven Phasen schaffen sie es bemerkenswert gut, ihre Kräfte zu bündeln und Arbeitskraft sowie Leistungsfähigkeit zu erhalten. Hohe Werte müssen jedoch nicht unbedingt mit der Abwesenheit von körperlichen Beschwerden einhergehen, denn besonders jene Personen, die sich auf eigenen Wunsch nachhaltig hohen Belastungen aussetzen, müssen sich häufig mit entsprechenden körperlichen Symptomen auseinandersetzen.

Personen mit **niedrigen Werten** geben an, bei starken Belastungen schnell an die Grenzen ihrer Leistungsfähigkeit zu kommen. Wenn sie sich über eine längere Zeit hinweg hohen Anforderungen stellen müssen, fühlen sie sich erschöpft und gereizt oder nervös. In sehr arbeitsintensiven Phasen, lassen Energie und Tatkraft nach. Mit niedrigen Skalenwerten müssen nicht unbedingt gesundheitliche Beschwerden einhergehen. Es geht hierbei eher um das Ausmaß an Belastung, welches man sich zuzumuten bereit ist.

Selbstbewusstsein

Sie haben x von 10 Punkten

Personen mit hohen Werten haben sich in nahezu allen sozialen Situationen gut im Griff. Ihr Auftreten wirkt überzeugend und selbstsicher. Sie haben kaum Schwierigkeiten damit, vor Gruppen zu sprechen oder aus einem anderen Grund im Mittelpunkt der Aufmerksamkeit zu stehen. Sie sind relativ unabhängig davon, was andere über sie denken. Wenn sie das Gefühl haben, abgelehnt zu werden, so belastet sie das nicht nachhaltig. Sie wirken mit sich im Reinen und strahlen dies im Sinne von Souveränität auch aus. Deshalb sind Sie in der Lage, auch deutliche Kritik anzunehmen, ohne dass ihr Selbstwertgefühl darunter leidet.

Personen mit **niedrigen Werten** empfinden in vielen Situationen Besorgnis darüber, wie sie auf andere wirken und welchen Eindruck sie hinterlassen. Die mögliche Bewertung durch andere Personen ist ihnen ständig präsent und schränkt ihren Handlungsspielraum in manchen Fällen ein. Bei nicht wenigen Gelegenheiten würden sie sich wünschen, selbstsicherer und weniger nervös zu sein, etwa wenn Sie vor größeren Gruppen sprechen müssen oder aus andere Gründen im Mittelpunkt der Aufmerksamkeit stehen. Kritik kann sie verunsichern, insbesondere dann wenn sie schroff formuliert ist.

Weiterführende Links zum Thema Bewerbungsprozess und Assessment-Center:

Im folgenden finden Sie einige Links, die hilfreiche Informationen über den Bewerbungsprozess und etwaige Assessment-Center enthalten!

<http://www.squeaker.net/de>

<http://www.e-fellows.net/>

<http://www.karriere.de/startseite/>

Herzlichen Dank für ihre Teilnahme!

3 Anhänge zur Analyse des Verfälschungsverhaltens

3.1 Skaleninterkorrelationsmatrizen (VG)

Versuchsgruppe 1

VG 1	BIP_Soz	BIP_Ko	BIP_TO	BIP_Est	BIP_Bel	BIP_Sb	BIP_LM	BIP_GM	BIP_FM	BIP_Du	BIP_Ge	BIP_FI
BIP_Soz r	1,000											
Sig.	.											
BIP_Ko r	.239	1,000										
Sig.	,007	.										
BIP_TO r	.206	.582	1,000									
Sig.	,020	,000	.									
BIP_Est r	.238	.327	.249	1,000								
Sig.	,007	,000	,005	.								
BIP_Bel r	.251	.301	,127	.633	1,000							
Sig.	,004	,001	,155	,000	.							
BIP_Sb r	-,144	.389	,167	.589	.556	1,000						
Sig.	,106	,000	,060	,000	,000	.						
BIP_LM r	-,105	.276	,085	,140	.525	.367	1,000					
Sig.	,240	,002	,340	,117	,000	,000	.					
BIP_GM r	-,213	.346	,077	.279	.270	.549	.473	1,000				
Sig.	,016	,000	,389	,001	,002	,000	,000	.				
BIP_FM r	-,224	.495	.256	.343	.317	.672	.438	.543	1,000			
Sig.	,011	,000	,004	,000	,000	,000	,000	,000	.			
BIP_Du r	-,167	.447	.224	.269	.332	.598	.442	.629	.766	1,000		
Sig.	,061	,000	,011	,002	,000	,000	,000	,000	,000	.		
BIP_Ge r	.216	-,110	-,085	-,073	,128	-,101	.198	-,030	-,047	,007	1,000	
Sig.	,015	,217	,344	,417	,151	,257	,025	,740	,600	,935	.	
BIP_FI r	,053	.399	.246	.450	.509	.436	.451	.472	.441	.408	-,275	1,000
Sig.	,555	,000	,005	,000	,000	,000	,000	,000	,000	,000	,002	.

VG 1		BIP_Soz	BIP_Ko	BIP_TO	BIP_Est	BIP_Bel	BIP_Sb	BIP_LM	BIP_GM	BIP_FM	BIP_Du	BIP_Ge	BIP_FI
BIP_HO	r	.275	.202	,031	.243	.442	.262	.310	,137	.205	.219	.340	.222
	Sig.	,002	,023	,725	,006	,000	,003	,000	,126	,021	,014	,000	,012
BIP_Sen	r	.566	.556	.289	.180	.245	,067	,100	,118	.190	.256	,076	.285
	Sig.	,000	,000	,001	,043	,006	,456	,262	,185	,032	,004	,398	,001
NEO_N	r	-.090	-.330	-.210	-.535	-.541	-.527	-.261	-,100	-.352	-.326	-,033	-.307
	Sig.	,314	,000	,018	,000	,000	,000	,003	,265	,000	,000	,709	,000
NEO_E	r	.182	.745	.656	.325	.279	.335	.192	.188	.435	.331	-,137	.420
	Sig.	,041	,000	,000	,000	,002	,000	,031	,034	,000	,000	,126	,000
NEO_O	r	.262	.447	.235	.229	.216	.221	.296	.438	.268	.280	-,056	.416
	Sig.	,003	,000	,008	,010	,015	,013	,001	,000	,002	,001	,528	,000
NEO_V	r	.721	.315	.427	,156	,109	-.224	-,148	-.207	-.253	-.237	,111	-,022
	Sig.	,000	,000	,000	,081	,222	,011	,097	,020	,004	,007	,214	,803
NEO_G	r	.215	.189	,165	,089	.334	,126	.348	,028	,108	,104	.597	,007
	Sig.	,015	,033	,063	,322	,000	,157	,000	,757	,229	,247	,000	,935
SozErw	r	.596	.271	.201	.226	.400	,068	,146	-,020	-,021	-,021	.354	,079
	Sig.	,000	,002	,023	,011	,000	,445	,103	,826	,812	,813	,000	,377

VG 1		BIP_HO	BIP_Sen	NEO_N	NEO_E	NEO_O	NEO_V	NEO_G	SozErw
BIP_HO	r	1,000							
	Sig.	.							
BIP_Sen	r	.325	1,000						
	Sig.	,000	.						
NEO_N	r	-.404	-.161	1,000					
	Sig.	,000	,071	.					
NEO_E	r	.209	.378	-.403	1,000				
	Sig.	,018	,000	,000	.				
NEO_O	r	.263	.440	-.082	.342	1,000			
	Sig.	,003	,000	,359	,000	.			
NEO_V	r	.202	.441	-.123	.356	.244	1,000		
	Sig.	,023	,000	,168	,000	,006	.		
NEO_G	r	.626	.183	-.361	.280	.129	.264	1,000	
	Sig.	,000	,040	,000	,001	,149	,003	.	
SozErw	r	.413	.379	-.300	.247	.241	.587	.509	1,000
	Sig.	,000	,000	,001	,005	,006	,000	,000	.

Versuchsgruppe 2

VG 2	BIP_Soz	BIP_Ko	BIP_TO	BIP_Est	BIP_Bel	BIP_Sb	BIP_LM	BIP_GM	BIP_FM	BIP_Du	BIP_Ge	BIP_FI
BIP_Soz r	1,000											
Sig.	.											
BIP_Ko r	,166	1,000										
Sig.	,063	.										
BIP_TO r	,416	,323	1,000									
Sig.	,000	,000	.									
BIP_Est r	,213	,413	,193	1,000								
Sig.	,017	,000	,030	.								
BIP_Bel r	,137	,391	,035	,710	1,000							
Sig.	,126	,000	,700	,000	.							
BIP_Sb r	-,142	,628	,084	,560	,532	1,000						
Sig.	,114	,000	,351	,000	,000	.						
BIP_LM r	-,024	,247	,021	,299	,594	,426	1,000					
Sig.	,792	,005	,813	,001	,000	,000	.					
BIP_GM r	-,260	,279	,107	,233	,387	,504	,493	1,000				
Sig.	,003	,002	,232	,009	,000	,000	,000	.				
BIP_FM r	-,168	,464	,182	,339	,388	,637	,474	,645	1,000			
Sig.	,060	,000	,041	,000	,000	,000	,000	,000	.			
BIP_Du r	-,337	,481	,110	,281	,346	,721	,374	,623	,739	1,000		
Sig.	,000	,000	,218	,001	,000	,000	,000	,000	,000	.		
BIP_Ge r	,109	,189	,043	,214	,413	,189	,385	,123	,238	,193	1,000	
Sig.	,223	,034	,632	,016	,000	,034	,000	,170	,007	,030	.	
BIP_FI r	,129	,367	,139	,549	,562	,476	,528	,469	,379	,281	,000	1,000
Sig.	,150	,000	,121	,000	,000	,000	,000	,000	,000	,001	,997	.

VG 2		BIP_Soz	BIP_Ko	BIP_TO	BIP_Est	BIP_Bel	BIP_Sb	BIP_LM	BIP_GM	BIP_FM	BIP_Du	BIP_Ge	BIP_FI
BIP_HO	r	,173	,243	,037	,553	,615	,346	,393	,214	,315	,273	,569	,368
	Sig.	,052	,006	,679	,000	,000	,000	,000	,016	,000	,002	,000	,000
BIP_Sen	r	,466	,556	,330	,373	,374	,351	,183	,206	,388	,267	,191	,347
	Sig.	,000	,000	,000	,000	,000	,000	,040	,020	,000	,003	,032	,000
NEO_N	r	-,156	-,479	-,113	-,650	-,526	-,560	-,273	-,208	-,414	-,387	-,257	-,413
	Sig.	,081	,000	,208	,000	,000	,000	,002	,019	,000	,000	,004	,000
NEO_E	r	,269	,702	,484	,390	,280	,426	,169	,310	,458	,350	,180	,313
	Sig.	,002	,000	,000	,000	,001	,000	,058	,000	,000	,000	,044	,000
NEO_O	r	,081	,170	,058	,060	,116	,035	,100	,356	,176	,086	-,014	,217
	Sig.	,369	,058	,522	,505	,195	,697	,266	,000	,049	,337	,876	,015
NEO_V	r	,659	,148	,304	,123	-,033	-,241	-,188	-,221	-,198	-,356	,123	,022
	Sig.	,000	,099	,001	,171	,717	,007	,035	,013	,026	,000	,171	,806
NEO_G	r	,157	,189	,011	,381	,477	,195	,349	,089	,255	,164	,722	,099
	Sig.	,079	,034	,903	,000	,000	,029	,000	,324	,004	,066	,000	,270
SozErw	r	,467	,231	,130	,468	,447	,151	,206	-,032	,045	-,013	,375	,229
	Sig.	,000	,009	,146	,000	,000	,092	,021	,722	,616	,886	,000	,010

VG 2		BIP_HO	BIP_Sen	NEO_N	NEO_E	NEO_O	NEO_V	NEO_G	SozErw
BIP_HO	r	1,000							
	Sig.	.							
BIP_Sen	r	.441	1,000						
	Sig.	,000	.						
NEO_N	r	-.464	-.404	1,000					
	Sig.	,000	,000	.					
NEO_E	r	.285	.559	-.485	1,000				
	Sig.	,001	,000	,000	.				
NEO_O	r	,113	.253	-.002	,164	1,000			
	Sig.	,206	,004	,985	,066	.			
NEO_V	r	,104	.317	-.192	.314	.194	1,000		
	Sig.	,246	,000	,031	,000	,030	.		
NEO_G	r	.729	.292	-.436	.245	,041	.231	1,000	
	Sig.	,000	,001	,000	,006	,650	,009	.	
SozErw	r	.495	.331	-.510	.221	,106	.446	.567	1,000
	Sig.	,000	,000	,000	,013	,236	,000	,000	.

Versuchsgruppe 3

VG 3	BIP_Soz	BIP_Ko	BIP_TO	BIP_Est	BIP_Bel	BIP_Sb	BIP_LM	BIP_GM	BIP_FM	BIP_Du	BIP_Ge	BIP_FI
BIP_Soz	r 1,000 Sig.											
BIP_Ko	r .348 Sig.	1,000										
BIP_TO	r .493 Sig.	.328	1,000									
BIP_Est	r .072 Sig.	.294	.062	1,000								
BIP_Bel	r .078 Sig.	.175	.032	.658	1,000							
BIP_Sb	r .035 Sig.	.348	.053	.679	.503	1,000						
BIP_LM	r .029 Sig.	.087	-.171	.191	.469	.336	1,000					
BIP_GM	r .168 Sig.	.274	-.023	.332	.402	.546	.577	1,000				
BIP_FM	r .150 Sig.	.318	.036	.476	.504	.685	.470	.680	1,000			
BIP_Du	r .154 Sig.	.381	.084	.449	.468	.729	.419	.678	.843	1,000		
BIP_Ge	r .045 Sig.	-.111	-.097	-.146	.117	-.055	.361	.082	.031	.059	1,000	
BIP_FI	r .007 Sig.	.250	.182	.503	.477	.478	.249	.530	.455	.426	-.296	1,000
		.939	.004	.036	.000	.000	.004	.000	.000	.000	.001	.

VG 3		BIP_Soz	BIP_Ko	BIP_TO	BIP_Est	BIP_Bel	BIP_Sb	BIP_LM	BIP_GM	BIP_FM	BIP_Du	BIP_Ge	BIP_FI
BIP_HO	r	,122	,136	-,084	,224	,408	,172	,303	,127	,221	,291	,565	,011
	Sig.	,162	,119	,339	,010	,000	,049	,000	,145	,011	,001	,000	,899
BIP_Sen	r	,545	,479	,391	,183	,245	,181	,047	,184	,251	,308	,049	,232
	Sig.	,000	,000	,000	,036	,005	,038	,595	,035	,004	,000	,574	,007
NEO_N	r	,119	-,166	,041	-,525	-,481	-,451	-,194	-,127	-,434	-,388	,023	-,280
	Sig.	,173	,058	,638	,000	,000	,000	,026	,148	,000	,000	,795	,001
NEO_E	r	,244	,613	,386	,272	,186	,272	,092	,358	,323	,304	-,105	,330
	Sig.	,005	,000	,000	,002	,032	,002	,292	,000	,000	,000	,231	,000
NEO_O	r	,158	,174	,139	,158	,083	,127	,107	,391	,110	,166	-,233	,422
	Sig.	,070	,046	,113	,070	,345	,146	,222	,000	,209	,057	,007	,000
NEO_V	r	,770	,328	,287	,045	,014	-,203	-,277	-,210	-,272	-,258	,008	-,019
	Sig.	,000	,000	,001	,606	,869	,019	,001	,016	,002	,003	,923	,831
NEO_G	r	,118	,027	-,105	,012	,236	-,004	,379	,080	,079	,112	,803	-,247
	Sig.	,179	,759	,230	,888	,006	,965	,000	,359	,370	,199	,000	,004
SozErw	r	,326	,142	,126	,030	,069	-,027	-,001	,024	-,192	-,130	,211	-,072
	Sig.	,000	,104	,150	,729	,431	,762	,993	,784	,027	,138	,015	,413

VG 3		BIP_HO	BIP_Sen	NEO_N	NEO_E	NEO_O	NEO_V	NEO_G	SozErw
BIP_HO	r	1,000							
	Sig.	.							
BIP_Sen	r	.272	1,000						
	Sig.	,002	.						
NEO_N	r	-.270	-.122	1,000					
	Sig.	,002	,164	.					
NEO_E	r	,039	.330	-.187	1,000				
	Sig.	,654	,000	,032	.				
NEO_O	r	-.201	.243	,078	,170	1,000			
	Sig.	,021	,005	,375	,051	.			
NEO_V	r	,126	.420	,122	,167	.179	1,000		
	Sig.	,151	,000	,164	,055	,040	.		
NEO_G	r	.742	.191	-.140	-.035	-.232	,122	1,000	
	Sig.	,000	,029	,109	,693	,007	,163	.	
SozErw	r	,164	,150	-.016	,096	,112	.346	.306	1,000
	Sig.	,060	,085	,857	,271	,200	,000	,000	.

3.2 Differenzen der Skaleninterkorrelationen (VG)

Differenz zwischen Versuchsgruppe 1 und 2

VG 1 – VG 2	BIP_Soz	BIP_Ko	BIP_TO	BIP_Est	BIP_Bel	BIP_Sb	BIP_LM	BIP_GM	BIP_FM	BIP_Du	BIP_Ge	BIP_FI
BIP_Soz r_1-r_2												
BIP_Ko r_1-r_2	0,07											
BIP_TO r_1-r_2	-0,21	0,26										
BIP_Est r_1-r_2	0,03	-0,09	0,06									
BIP_Bel r_1-r_2	0,11	-0,09	0,09	-0,08								
BIP_Sb r_1-r_2	0,00	-0,24	0,08	0,03	0,02							
BIP_LM r_1-r_2	-0,08	0,03	0,06	-0,16	-0,07	-0,06						
BIP_GM r_1-r_2	-0,47	0,07	-0,03	0,05	-0,12	0,05	-0,02					
BIP_FM r_1-r_2	-0,06	0,03	0,07	0,00	-0,07	0,04	-0,04	-0,10				
BIP_Du r_1-r_2	0,17	-0,03	0,11	-0,01	-0,01	-0,12	0,07	0,01	0,03			
BIP_Ge r_1-r_2	0,11	-0,30	-0,13	-0,29	-0,28	-0,29	-0,19	-0,15	-0,28	-0,19		
BIP_FI r_1-r_2	-0,08	0,03	0,11	-0,10	-0,05	-0,04	-0,08	0,00	0,06	0,13	-0,27	

Die grauen Markierungen kennzeichnen die Korrelationsdifferenzbeträge $\geq .30$

VG 1 – VG 2		BIP_Soz	BIP_Ko	BIP_TO	BIP_Est	BIP_Bel	BIP_Sb	BIP_LM	BIP_GM	BIP_FM	BIP_Du	BIP_Ge	BIP_FI
BIP_Soz	r_1-r_2												
BIP_Ko	r_1-r_2	0,07											
BIP_TO	r_1-r_2	-0,21	0,26										
BIP_Est	r_1-r_2	0,03	-0,09	0,06									
BIP_Bel	r_1-r_2	0,11	-0,09	0,09	-0,08								
BIP_Sb	r_1-r_2	0,00	-0,24	0,08	0,03	0,02							
BIP_LM	r_1-r_2	-0,08	0,03	0,06	-0,16	-0,07	-0,06						
BIP_GM	r_1-r_2	-0,47	0,07	-0,03	0,05	-0,12	0,05	-0,02					
BIP_FM	r_1-r_2	-0,06	0,03	0,07	0,00	-0,07	0,04	-0,04	-0,10				
BIP_Du	r_1-r_2	0,17	-0,03	0,11	-0,01	-0,01	-0,12	0,07	0,01	0,03			
BIP_Ge	r_1-r_2	0,11	-0,30	-0,13	-0,29	-0,28	-0,29	-0,19	-0,15	-0,28	-0,19		
BIP_FI	r_1-r_2	-0,08	0,03	0,11	-0,10	-0,05	-0,04	-0,08	0,00	0,06	0,13	-0,27	

Die grauen Markierungen kennzeichnen die Korrelationsdifferenzbeträge $\geq .30$

VG 1 – VG 2		BIP_HO	BIP_Sen	NEO_N	NEO_E	NEO_O	NEO_V	NEO_G	SozErw
BIP_HO	r_1-r_2								
BIP_Sen	r_1-r_2	-0,12							
NEO_N	r_1-r_2	0,06	0,24						
NEO_E	r_1-r_2	-0,08	-0,18	0,08					
NEO_O	r_1-r_2	0,15	0,19	-0,08	0,18				
NEO_V	r_1-r_2	0,10	0,12	0,07	0,04	0,05			
NEO_G	r_1-r_2	-0,10	-0,11	0,08	0,04	0,09	0,03		
SozErw	r_1-r_2	-0,08	0,05	0,21	0,03	0,13	0,14	-0,06	

Die grauen Markierungen kennzeichnen die Korrelationsdifferenzbeträge $\geq .30$

Differenz zwischen Versuchsgruppe 2 und 3

VG 2 – VG 3	BIP_Soz	BIP_Ko	BIP_TO	BIP_Est	BIP_Bel	BIP_Sb	BIP_LM	BIP_GM	BIP_FM	BIP_Du	BIP_Ge	BIP_FI
BIP_Soz r_2-r_3												
BIP_Ko r_2-r_3	-0,18											
BIP_TO r_2-r_3	-0,08	-0,01										
BIP_Est r_2-r_3	0,14	0,12	0,13									
BIP_Bel r_2-r_3	0,06	0,22	0,00	0,05								
BIP_Sb r_2-r_3	0,04	0,28	0,03	-0,12	0,03							
BIP_LM r_2-r_3	0,17	0,16	0,19	0,11	0,13	0,09						
BIP_GM r_2-r_3	0,43	0,01	0,13	-0,10	-0,02	-0,04	-0,08					
BIP_FM r_2-r_3	-0,02	0,15	0,15	-0,14	-0,12	-0,05	0,00	-0,04				
BIP_Du r_2-r_3	-0,18	0,10	0,03	-0,17	-0,12	-0,01	-0,05	-0,06	-0,10			
BIP_Ge r_2-r_3	0,06	0,30	0,14	0,36	0,30	0,24	0,02	0,04	0,21	0,13		
BIP_FI r_2-r_3	0,12	0,12	-0,04	0,05	0,09	0,00	0,28	-0,06	-0,08	-0,15	0,30	

Die grauen Markierungen kennzeichnen die Korrelationsdifferenzbeträge $\geq .30$

VG 2 – VG 3	BIP_Soz	BIP_Ko	BIP_TO	BIP_Est	BIP_Bel	BIP_Sb	BIP_LM	BIP_GM	BIP_FM	BIP_Du	BIP_Ge	BIP_FI
BIP_HO r_2-r_3	0,05	0,11	0,12	0,33	0,21	0,17	0,09	0,09	0,09	-0,02	0,00	0,36
BIP_Sen r_2-r_3	-0,08	0,08	-0,06	0,19	0,13	0,17	0,14	0,02	0,14	-0,04	0,14	0,12
NEO_N r_2-r_3	-0,28	-0,31	-0,15	-0,13	-0,05	-0,11	-0,08	-0,08	0,02	0,00	-0,28	-0,13
NEO_E r_2-r_3	0,03	0,09	0,10	0,12	0,09	0,15	0,08	-0,05	0,14	0,05	0,29	-0,02
NEO_O r_2-r_3	-0,08	0,00	-0,08	-0,10	0,03	-0,09	-0,01	-0,04	0,07	-0,08	0,22	-0,21
NEO_V r_2-r_3	-0,11	-0,18	0,02	0,08	-0,05	-0,04	0,09	-0,01	0,07	-0,10	0,11	0,04
NEO_G r_2-r_3	0,04	0,16	0,12	0,37	0,24	0,20	-0,03	0,01	0,18	0,05	-0,08	0,35
SozErw r_2-r_3	0,14	0,09	0,00	0,44	0,38	0,18	0,21	-0,06	0,24	0,12	0,16	0,30

Die grauen Markierungen kennzeichnen die Korrelationsdifferenzbeträge $\geq .30$

VG 2 – VG 3		BIP_HO	BIP_Sen	NEO_N	NEO_E	NEO_O	NEO_V	NEO_G	SozErw
BIP_HO	r_2-r_3								
BIP_Sen	r_2-r_3	0,17							
NEO_N	r_2-r_3	-0,19	-0,28						
NEO_E	r_2-r_3	0,25	0,23	-0,30					
NEO_O	r_2-r_3	0,31	0,01	-0,08	-0,01				
NEO_V	r_2-r_3	-0,02	-0,10	-0,31	0,15	0,02			
NEO_G	r_2-r_3	-0,01	0,10	-0,30	0,28	0,27	0,11		
SozErw	r_2-r_3	0,33	0,18	-0,49	0,12	-0,01	0,10	0,26	

Die grauen Markierungen kennzeichnen die Korrelationsdifferenzbeträge $\geq .30$

Differenz zwischen Versuchsgruppe 3 und 1

VG 3 – VG 1	BIP_Soz	BIP_Ko	BIP_TO	BIP_Est	BIP_Bel	BIP_Sb	BIP_LM	BIP_GM	BIP_FM	BIP_Du	BIP_Ge	BIP_FI
BIP_Soz r_3-r_1												
BIP_Ko r_3-r_1	0,11											
BIP_TO r_3-r_1	0,29	-0,25										
BIP_Est r_3-r_1	-0,17	-0,03	-0,19									
BIP_Bel r_3-r_1	-0,17	-0,13	-0,10	0,03								
BIP_Sb r_3-r_1	-0,04	-0,04	-0,11	0,09	-0,05							
BIP_LM r_3-r_1	-0,09	-0,19	-0,26	0,05	-0,06	-0,03						
BIP_GM r_3-r_1	0,04	-0,07	-0,10	0,05	0,13	0,00	0,10					
BIP_FM r_3-r_1	0,07	-0,18	-0,22	0,13	0,19	0,01	0,03	0,14				
BIP_Du r_3-r_1	0,01	-0,07	-0,14	0,18	0,14	0,13	-0,02	0,05	0,08			
r_3-r_1	-0,17	0,00	-0,01	-0,07	-0,01	0,05	0,16	0,11	0,08	0,05		
BIP_FI r_3-r_1	-0,05	-0,15	-0,06	0,05	-0,03	0,04	-0,20	0,06	0,01	0,02	-0,02	

Die grauen Markierungen kennzeichnen die Korrelationsdifferenzbeträge $\geq .30$

VG 3 – VG 1	BIP_Soz	BIP_Ko	BIP_TO	BIP_Est	BIP_Bel	BIP_Sb	BIP_LM	BIP_GM	BIP_FM	BIP_Du	BIP_Ge	BIP_FI
BIP_HO r_3-r_1	-0,15	-0,07	-0,12	-0,02	-0,03	-0,09	-0,01	-0,01	0,02	0,07	0,23	-0,21
BIP_Sen r_3-r_1	-0,02	-0,08	0,10	0,00	0,00	0,11	-0,05	0,07	0,06	0,05	-0,03	-0,05
NEO_N r_3-r_1	0,21	0,16	0,25	0,01	0,06	0,08	0,07	-0,03	-0,08	-0,06	0,06	0,03
NEO_E r_3-r_1	0,06	-0,13	-0,27	-0,05	-0,09	-0,06	-0,10	0,17	-0,11	-0,03	0,03	-0,09
NEO_O r_3-r_1	-0,10	-0,27	-0,10	-0,07	-0,13	-0,09	-0,19	-0,05	-0,16	-0,11	-0,18	0,01
NEO_V r_3-r_1	0,05	0,01	-0,14	-0,11	-0,09	0,02	-0,13	0,00	-0,02	-0,02	-0,10	0,00
NEO_G r_3-r_1	-0,10	-0,16	-0,27	-0,08	-0,10	-0,13	0,03	0,05	-0,03	0,01	0,21	-0,25
SozErw r_3-r_1	-0,27	-0,13	-0,08	-0,20	-0,33	-0,09	-0,15	0,04	-0,17	-0,11	-0,14	-0,15

Die grauen Markierungen kennzeichnen die Korrelationsdifferenzbeträge $\geq .30$

VG 3 – VG 1		BIP_HO	BIP_Sen	NEO_N	NEO_E	NEO_O	NEO_V	NEO_G	SozErw
BIP_HO	r_3-r_1								
BIP_Sen	r_3-r_1	-0,05							
NEO_N	r_3-r_1	0,13	0,04						
NEO_E	r_3-r_1	-0,17	-0,05	0,22					
NEO_O	r_3-r_1	-0,46	-0,20	0,16	-0,17				
NEO_V	r_3-r_1	-0,08	-0,02	0,24	-0,19	-0,07			
NEO_G	r_3-r_1	0,12	0,01	0,22	-0,31	-0,36	-0,14		
SozErw	r_3-r_1	-0,25	-0,23	0,28	-0,15	-0,13	-0,24	-0,20	

Die grauen Markierungen kennzeichnen die Korrelationsdifferenzbeträge $\geq .30$

4 Anhänge zur Analyse der Antworttendenzen mittels IRT-Modellen

4.1 Beispiel- Syntax der IRT-Analyse mittels mdltm-Software

Syntax zu Bereich Arbeitsverhalten (AV), BIP, eindimensionales Modell: d-Dimension

```
//  
// control file template egpcm 0.07  
//
```

```
title=BIP_123_AV_d_1dim
```

```
//  
// datafile  
//
```

```
data=123_AV_d.txt
```

```
//  
// item description  
//
```

```
maxnrcat=2
```

```
nitems=42
```

```
maxstepwidth=0.2
```

```
//  
// number of classes  
//
```

```
ncl=3
```

```
//  
// group definition  
//
```

```
groups=1.1
```

```
//  
groupid=1,"1"  
groupid=2,"2"  
groupid=3,"3"
```

```
// number of skills
```

```
//  
nskills=1
```

```
//
```

```
// number of skill levels max  
//
```

```
maxlevel=26
```

```
//  
// discrete skill definitions  
//
```

```

skill=1,"dDimension",26,-4.0:4.0

//
// do slope parameters
//
    doslopes=None

    centerslope=None

//
// fix sum of item parameters to be zero
//
    centerscale=Items>0;0
//
    traitdistribution=loglin

    maxmoment=5

    scaletheta=yes

    centertheta=yes

//
// item positions
//
    subjectid = 1.9

//

item=2.2,Rasch,"d_B_12_AV_fl_1",1
item=3.3,Rasch,"d_B_149_AV_fl_10",1
item=4.4,Rasch,"d_B_163_AV_fl_11",1
item=5.5,Rasch,"d_B_177_AV_fl_12",1
item=6.6,Rasch,"d_B_189_AV_fl_13",1
item=7.7,Rasch,"d_B_198_AV_fl_14",1
item=8.8,Rasch,"d_B_27_AV_fl_2",1
item=9.9,Rasch,"d_B_42_AV_fl_3",1
item=10.10,Rasch,"d_B_58_AV_fl_4",1
item=11.11,Rasch,"d_B_74_AV_fl_5",1
item=12.12,Rasch,"d_B_90_AV_fl_6",1
item=13.13,Rasch,"d_B_105_AV_fl_7",1
item=14.14,Rasch,"d_B_120_AV_fl_8",1
item=15.15,Rasch,"d_B_135_AV_fl_9",1
item=16.16,Rasch,"d_B_57_AV_ge_4",1
item=17.17,Rasch,"d_B_11_AV_ge_1",1
item=18.18,Rasch,"d_B_148_AV_ge_10",1
item=19.19,Rasch,"d_B_162_AV_ge_11",1
item=20.20,Rasch,"d_B_176_AV_ge_12",1
item=21.21,Rasch,"d_B_188_AV_ge_13",1
item=22.22,Rasch,"d_B_206_AV_ge_14",1
item=23.23,Rasch,"d_B_26_AV_ge_2",1
item=24.24,Rasch,"d_B_41_AV_ge_3",1
item=25.25,Rasch,"d_B_73_AV_ge_5",1
item=26.26,Rasch,"d_B_89_AV_ge_6",1
item=27.27,Rasch,"d_B_104_AV_ge_7",1
item=28.28,Rasch,"d_B_119_AV_ge_8",1
item=29.29,Rasch,"d_B_134_AV_ge_9",1
item=30.30,Rasch,"d_B_13_AV_ho_1",1
item=31.31,Rasch,"d_B_150_AV_ho_10",1

```

```

item=32.32,Rasch,"d_B_164_AV_ho_11",1
item=33.33,Rasch,"d_B_178_AV_ho_12",1
item=34.34,Rasch,"d_B_190_AV_ho_13",1
item=35.35,Rasch,"d_B_199_AV_ho_14",1
item=36.36,Rasch,"d_B_28_AV_ho_2",1
item=37.37,Rasch,"d_B_43_AV_ho_3",1
item=38.38,Rasch,"d_B_59_AV_ho_4",1
item=39.39,Rasch,"d_B_75_AV_ho_5",1
item=40.40,Rasch,"d_B_91_AV_ho_6",1
item=41.41,Rasch,"d_B_106_AV_ho_7",1
item=42.42,Rasch,"d_B_121_AV_ho_8",1
item=43.43,Rasch,"d_B_136_AV_ho_9",1

```

```

//
// equality constraints
//
//      separator=COMMA

// missing value code for data
//
//      missing = 9

// maxiter and likelihood convergence criterion
//
//      iterations=2500,0.025

//      paramchange=0.00001

//
//
// number of newton raphson iterations per m-step
//

//      msiter=1

//
// read fixed item parameters from file
//
//      outparameters=YES

//
//
//      personparameter=YES

//      printskillprob=no

```

4.2 Beispiel-Output der IRT-Analyse mittels mdltm-Software

Output zu Bereich Arbeitsverhalten (AV), BIP, eindimensionales Modell: d-Dimension

```
=====
= MDLTM 1.077 by MvD - 01-27-2012 =03/06/13 10:55:40
=====

Analysis title= BIP_123_AV_d_1dim
Read data from file= 123_AV_d.txt
Maximum number of response categories = 2
Number of items in definition file = 42
Number of latent classes ncl = 3
read group membership from columns: 0 to: 0
Number of skills: nskills = 1
Maximum number of levels per skill: 26

Skill: 1 label: DDIMENSION levels: 0.. 26
0 -4.00000
1 -3.68000
2 -3.36000
3 -3.04000
4 -2.72000
5 -2.40000
6 -2.08000
7 -1.76000
8 -1.44000
9 -1.12000
10 -0.80000
11 -0.48000
12 -0.16000
13 0.16000
14 0.48000
15 0.80000
16 1.12000
17 1.44000
18 1.76000
19 2.08000
20 2.40000
21 2.72000
22 3.04000
23 3.36000
24 3.68000
25 4.00000
No slopes estimated, CENTERSLOPE does not make sense
centerslope[0][0] NONE
No slopes estimated, CENTERSLOPE does not make sense
centerslope[1][0] NONE
No slopes estimated, CENTERSLOPE does not make sense
centerslope[2][0] NONE
read subject id from columns: 0 to: 8
Global missing value code = 9
deviation 0.025000

Number of subjects      :      386
```

Number of items : 42
 Max. number of categories : 2

assignment of group not found 385

0 0.329958343936165 1 0.327252321274791 2 0.342789334789045 classified:

385.00000 all: 386.00000

item	tried	0	1	r(itm,skill)	logits		
1D_B_12_AV_ncat:	2	323.000	256.000	67.000	0.30861		-1.32957
2D_B_149_AV_ncat:	2	300.000	236.000	64.000	0.32699		-1.29367
3D_B_163_AV_ncat:	2	272.000	150.000	122.000	0.10640		-0.20510
4D_B_177_AV_ncat:	2	311.000	74.000	237.000	-0.12365		1.15478
5D_B_189_AV_ncat:	2	268.000	202.000	66.000	0.29448		-1.10851
6D_B_198_AV_ncat:	2	239.000	143.000	96.000	0.40055		-0.39510
7D_B_27_AV_ncat:	2	295.000	258.000	37.000	0.28353		-1.91924
8D_B_42_AV_ncat:	2	316.000	280.000	36.000	0.32041		-2.02744
9D_B_58_AV_ncat:	2	304.000	147.000	157.000	0.10111		0.06538
10D_B_74_AV_ncat:	2	308.000	174.000	134.000	0.14229		-0.25951
11D_B_90_AV_ncat:	2	288.000	205.000	83.000	0.34526		-0.89706
12D_B_105_AV_ncat:	2	281.000	244.000	37.000	0.37982		-1.86367
13D_B_120_AV_ncat:	2	319.000	265.000	54.000	0.20852		-1.57616
14D_B_135_AV_ncat:	2	300.000	256.000	44.000	0.25956		-1.74241
15D_B_57_AV_ncat:	2	291.000	236.000	55.000	0.42414		-1.44271
16D_B_11_AV_ncat:	2	291.000	195.000	96.000	0.35311		-0.70340
17D_B_148_AV_ncat:	2	313.000	252.000	61.000	0.46578		-1.40626
18D_B_162_AV_ncat:	2	289.000	243.000	46.000	0.46887		-1.64702
19D_B_176_AV_ncat:	2	341.000	297.000	44.000	0.40743		-1.89043
20D_B_188_AV_ncat:	2	303.000	238.000	65.000	0.29984		-1.28681
21D_B_206_AV_ncat:	2	306.000	254.000	52.000	0.35394		-1.57097
22D_B_26_AV_ncat:	2	312.000	257.000	55.000	0.28469		-1.52761
23D_B_41_AV_ncat:	2	310.000	242.000	68.000	0.33086		-1.25895
24D_B_73_AV_ncat:	2	299.000	225.000	74.000	0.40659		-1.10305
25D_B_89_AV_ncat:	2	327.000	311.000	16.000	0.27775		-2.90979
26D_B_104_AV_ncat:	2	307.000	237.000	70.000	0.25649		-1.20959
27D_B_119_AV_ncat:	2	298.000	246.000	52.000	0.46388		-1.53910
28D_B_134_AV_ncat:	2	332.000	306.000	26.000	0.22833		-2.43101
29D_B_13_AV_ncat:	2	354.000	316.000	38.000	0.37025		-2.09534
30D_B_150_AV_ncat:	2	308.000	253.000	55.000	0.51148		-1.51198
31D_B_164_AV_ncat:	2	283.000	214.000	69.000	0.53325		-1.12214
32D_B_178_AV_ncat:	2	358.000	320.000	38.000	0.04290		-2.10788
33D_B_190_AV_ncat:	2	351.000	337.000	14.000	0.36902		-3.11500
34D_B_199_AV_ncat:	2	341.000	321.000	20.000	0.40786		-2.73003
35D_B_28_AV_ncat:	2	312.000	264.000	48.000	0.39303		-1.68791
36D_B_43_AV_ncat:	2	301.000	274.000	27.000	0.40218		-2.28457
37D_B_59_AV_ncat:	2	345.000	306.000	39.000	0.39706		-2.03797
38D_B_75_AV_ncat:	2	310.000	281.000	29.000	0.33307		-2.24071
39D_B_91_AV_ncat:	2	327.000	171.000	156.000	0.37936		-0.09125
40D_B_106_AV_ncat:	2	333.000	323.000	10.000	0.27523		-3.38285
41D_B_121_AV_ncat:	2	317.000	276.000	41.000	0.44270		-1.88635
42D_B_136_AV_ncat:	2	306.000	191.000	115.000	0.43057		-0.50391

Proportion correct scale: DDIMENSION

0 percent 0.09091 *****
3 percent 0.11948 *****
7 percent 0.12468 *****
11 percent 0.14026 *****
15 percent 0.06494 *****
19 percent 0.09610 *****
23 percent 0.06234 *****
26 percent 0.07532 *****
30 percent 0.05974 *****
34 percent 0.03117 *****
38 percent 0.03636 *****
42 percent 0.03117 *****
46 percent 0.02857 *****
50 percent 0.01299 ***
53 percent 0.00779 **
57 percent 0.01039 **
61 percent 0.00519 *
65 percent 0.00000
69 percent 0.00260
73 percent 0.00000
76 percent 0.00000
80 percent 0.00000
84 percent 0.00000
88 percent 0.00000
92 percent 0.00000
96 percent 0.00000

Cases without valid responses: 1.000000

Skill patterns :

0 : 0. : 0.024820916134 : 0.025192596634 : 0.024844309221
1 : 1. : 0.027369345058 : 0.027283629496 : 0.027447708829
2 : 2. : 0.029273049790 : 0.029661710915 : 0.029665402743
3 : 3. : 0.031608239541 : 0.031804419909 : 0.032169773364
4 : 4. : 0.034187841833 : 0.034750717371 : 0.034139713568
5 : 5. : 0.036540048754 : 0.037177409181 : 0.037430989918
6 : 6. : 0.039669818742 : 0.039201921962 : 0.039057428724
7 : 7. : 0.042116307861 : 0.041411305072 : 0.041325398358
8 : 8. : 0.044050346046 : 0.043903572998 : 0.044451137143
9 : 9. : 0.046002017992 : 0.045305770603 : 0.045622199142
10 : 10. : 0.047259976840 : 0.047600696984 : 0.047030927266
11 : 11. : 0.048219365568 : 0.048191378034 : 0.048149970611
12 : 12. : 0.048909305106 : 0.047989106026 : 0.048106341434
13 : 13. : 0.049051216105 : 0.048417976561 : 0.048707801274
14 : 14. : 0.047905077699 : 0.048808738551 : 0.048400157390
15 : 15. : 0.047414021555 : 0.046566732038 : 0.046690678953
16 : 16. : 0.045455039077 : 0.045838225953 : 0.046008078820
17 : 17. : 0.043460876225 : 0.044375347915 : 0.043934194156
18 : 18. : 0.042049954978 : 0.041784548742 : 0.041647669627
19 : 19. : 0.039360112692 : 0.039534035616 : 0.039848202044
20 : 20. : 0.036827976723 : 0.036768888954 : 0.037434696520
21 : 21. : 0.034421674984 : 0.034501533682 : 0.034397298624
22 : 22. : 0.032234386684 : 0.032365402216 : 0.031741832833
23 : 23. : 0.029469999420 : 0.029477680549 : 0.029695106315
24 : 24. : 0.027141070471 : 0.027147101782 : 0.027227550526
25 : 25. : 0.025182014121 : 0.024939552256 : 0.024825432598

QMATRIX:DDIMENSION

D_B_12_AV_FL_1 1.00000
D_B_149_AV_FL_10 1.00000
D_B_163_AV_FL_11 1.00000
D_B_177_AV_FL_12 1.00000
D_B_189_AV_FL_13 1.00000
D_B_198_AV_FL_14 1.00000
D_B_27_AV_FL_2 1.00000
D_B_42_AV_FL_3 1.00000
D_B_58_AV_FL_4 1.00000
D_B_74_AV_FL_5 1.00000
D_B_90_AV_FL_6 1.00000
D_B_105_AV_FL_7 1.00000
D_B_120_AV_FL_8 1.00000
D_B_135_AV_FL_9 1.00000
D_B_57_AV_GE_4 1.00000
D_B_11_AV_GE_1 1.00000
D_B_148_AV_GE_10 1.00000
D_B_162_AV_GE_11 1.00000
D_B_176_AV_GE_12 1.00000
D_B_188_AV_GE_13 1.00000
D_B_206_AV_GE_14 1.00000
D_B_26_AV_GE_2 1.00000
D_B_41_AV_GE_3 1.00000
D_B_73_AV_GE_5 1.00000
D_B_89_AV_GE_6 1.00000
D_B_104_AV_GE_7 1.00000
D_B_119_AV_GE_8 1.00000
D_B_134_AV_GE_9 1.00000
D_B_13_AV_HO_1 1.00000
D_B_150_AV_HO_10 1.00000
D_B_164_AV_HO_11 1.00000
D_B_178_AV_HO_12 1.00000
D_B_190_AV_HO_13 1.00000
D_B_199_AV_HO_14 1.00000
D_B_28_AV_HO_2 1.00000
D_B_43_AV_HO_3 1.00000
D_B_59_AV_HO_4 1.00000
D_B_75_AV_HO_5 1.00000
D_B_91_AV_HO_6 1.00000
D_B_106_AV_HO_7 1.00000
D_B_121_AV_HO_8 1.00000
D_B_136_AV_HO_9 1.00000

Start Time: 10:55:41

2	-5150.863603258463600	386.0000000000000000	0.012925829666869
0.063327353875811	386.0000000000000000	0.012925829666869	
-- Adjust step parameter: 1.0000000000000000			
3	-5441.334362027708300	-290.470758769244640	0.012399171121904
0.066574216739616	322.313839179496430	0.000526658544965	
4	-5414.097698623773800	27.236663403934472	0.011955179680313
0.069987549822231	125.595721995788450	0.000443991441591	
5	-5399.571737869436200	14.525960754337575	0.011590979252339
0.073575888234288	51.549214501487867	0.000364200427974	
6	-5391.606078678045400	7.965659191390841	0.011309993173024
0.077348204690900	22.493510961423183	0.000280986079315	
7	-5386.242013865242400	5.364064812803008	0.011079092953348
0.081313931948120	11.073880195676400	0.000230900219676	
8	-5381.882467130490100	4.359546734752257	0.010856177520364
0.085482986389745	6.597657888393639	0.000222915432984	

9	-5378.040751299313300	3.841715831176771	0.010608931661661
0.089865792823444	4.760363183582394	0.000247245858702	
10	-5374.580792314009200	3.459958985304183	0.010326759717233
0.094473310548203	3.893427051396920	0.000282171944428	
11	-5371.474168043504500	3.106624270504653	0.010009115364883
0.099317060758282	3.368891864135409	0.000317644352350	
12	-5368.701831942748900	2.772336100755638	0.009656998395007
0.104409155352203	2.971188021882228	0.000352116969876	
13	-5366.245394896006500	2.456437046742394	0.009274126209630
0.109762327218805	2.628020705122339	0.000382872185377	
14	-5364.092443367115300	2.152951528891208	0.008861434289807
0.115389962076097	2.311307920968252	0.000412691919822	
15	-5362.217691959390300	1.874751407724943	0.008420221478800
0.121306131942527	2.020270245472713	0.000441212811007	
16	-5360.604206404360400	1.613485555029911	0.007952288284255
0.127525630324355	1.749080451844178	0.000467933194545	
17	-5359.224270313571700	1.379936090788760	0.007461317478267
0.134064009207128	1.502984211140566	0.000490970805988	
18	-5358.056413407564800	1.167856906006819	0.006951645455734
0.140937617943743	1.279566007718068	0.000509672022533	
19	-5357.080514378459200	0.975899029105676	0.006431796357049
0.148163644136344	1.077121355309807	0.000519849098686	
20	-5356.267837702971500	0.812676675487637	0.005907344520754
0.153482353768513	0.900824902095027	0.000524451836295	
21	-5355.601800487277600	0.666037215693905	0.005303093012933
0.158923230353266	0.744299777827612	0.000604251507822	
22	-5355.060917967751300	0.540882519526349	0.004729950279052
0.164702834797608	0.608688272293437	0.000573142733880	
23	-5354.624132654478700	0.436785313272594	0.004195954745260
0.170828588753683	0.494086299612875	0.000533995533792	
24	-5354.276154889912500	0.347977764566167	0.003706867266028
0.177309621666999	0.396680609581736	0.000489087479232	
25	-5353.991765819289900	0.284389070622638	0.003260935768062
0.184156605858081	0.321819583609004	0.000445931497966	
26	-5353.766416807294300	0.225349011995604	0.002858536168543
0.182047845325368	0.257505869200071	0.000402399599520	
27	-5353.399354279902600	0.367062527391681	0.002388019761510
0.180060977100248	0.330543641327811	0.000470516407032	
-- Adjust step parameter:	0.100000000000000		
28	-5353.412589261914000	-0.013234982011454	0.002008639995056
0.178184522624022	0.119004535116906	0.000379379766454	
29	-5353.302869023174000	0.109720238740010	0.001884253678559
0.176408462277124	0.112815004198976	0.000124386316496	
30	-5353.182203566937500	0.120665456236566	0.001734443948326
0.174724005731977	0.118048638890702	0.000149809730234	
31	-5352.998316237883000	0.183887329054414	0.001489188896087
0.173123404906676	0.161941098999844	0.000245255052238	
32	-5352.943377036775000	0.054939201108027	0.001270671076840
0.171599800501605	0.090606500405299	0.000218517819248	
33	-5352.903201984327100	0.040175052447921	0.001080053064739
0.170147095230800	0.056985535100381	0.000190618012100	
34	-5352.870723763404400	0.032478220922712	0.000916689321330
0.168759848437637	0.040647325648602	0.000163363743409	
35	-5352.843328412994200	0.027395350410188	0.000778334570609
0.167433187965374	0.031812675489659	0.000138354750721	
36	-5352.819324640080600	0.024003772913602	0.000662201635352
0.166162736045314	0.026606740438954	0.000116132935257	
37	-5352.797566223781400	0.021758416299235	0.000565240486679
0.164944546645496	0.023374524345808	0.000096961148673	

38	-5352.777277732877300	0.020288490904022	0.000483877796900
0.163775052245638	0.021317168717951	0.000081362689779	
39	-5352.757937534001000	0.019340198876307	0.000416573433413
0.162651018409090	0.019999188823522	0.000067304363487	
40	-5352.739197411438000	0.018740122563031	0.000360463400996
0.161569504838635	0.019159811316528	0.000056110032417	
41	-5352.720827766490400	0.018369644947597	0.000313916900962
0.160527831851432	0.018633033737241	0.000046546500034	
-- Adjust step parameter:	0.018856491423232		
42	-5352.731056494286300	-0.010228727795948	0.000280156032270
0.159523551404911	0.013030163109712	0.000033760868692	
43	-5352.715010441326700	0.016046052959609	0.000381475995500
0.158554421961867	0.015040756342977	0.000101319963229	
44	-5352.711007958557300	0.004002482769465	0.000389082033065
0.157618386608318	0.007681907293969	0.000007606037565	

EAP based estimates for class "1": Respondents: 127.07

Proportion completely without responses: 0.00787

EAP Reliability: 0.73420

Adjusted EAP Reliability: 0.73850

Marginal EAP mean:-1.97544

Variance of posterior means: 0.53793

Average posterior variance: 0.19475

Sum of the two variance terms: 0.73267

Model based latent variable variance: 0.73183

EAP based estimates for class "2": Respondents: 126.01

Proportion completely without responses: 0.00000

EAP Reliability: 0.80964

Adjusted EAP Reliability: 0.80964

Marginal EAP mean:-0.36527

Variance of posterior means: 0.93377

Average posterior variance: 0.21954

Sum of the two variance terms: 1.15331

Model based latent variable variance: 1.15040

EAP based estimates for class "3": Respondents: 132.92

Proportion completely without responses: 0.00000

EAP Reliability: 0.81013

Adjusted EAP Reliability: 0.81013

Marginal EAP mean: 0.04020

Variance of posterior means: 0.83495

Average posterior variance: 0.19568

Sum of the two variance terms: 1.03063

Model based latent variable variance: 1.02987

Crosstable of class membership:

"1"	127.00000	0.329015544	1.00000	0.00000	0.00000
"2"	126.00000	0.326424870	0.00000	1.00000	0.00000
"3"	133.00000	0.344559585	0.00000	0.00000	0.99942

Start Time: 10:55:41 ... End Time: 10:55:45

Number of estimated item thresholds	126
Number of estimated item slopes	0
Minus determinacy constraints on items (negative)	-1
Number of estimated skill distribution parameters	15 (15)

Minus skills determinacy constraints (negative) 0
 Number of class-size and cluster parameters 2 (percent classified: 100)
 Total number of parameters: 142

Likelihood: -5352.7040 Deviance: 10705.40800

AIC penalty term : 284.00 AIC Index: 10989.40800
 BIC penalty term : 845.73 BIC Index: 11551.13691

CAIC penalty term : 987.73 CAIC Index: 11693.13691

Model based log penalty per item: 0.4120952 .. Akaike: 0.4230275
 Independence log penalty per item: 0.4520498
 Average Model based odds $P(X|M)/P(X|I)$ 1.0407635

Model based log-lik per respondent : -13.8671088
 Independence log-lik per respondent : -15.2115920

Unweighted N: 386 Average # of items per respondent: 33.65

Iterations 44
 Estimated skill distribution(s):
 Class: "1" size: 0.3292842
 Scale: 1
 Targeted Centered Mean: -1.9751062 .. Stdev: 0.8554727

-3.4319 : 0.0331061 *****
 -3.2233 : 0.0535164 *****
 -3.0146 : 0.0697033 *****
 -2.8060 : 0.0790168 *****
 -2.5973 : 0.0825478 *****
 -2.3887 : 0.0826725 *****
 -2.1800 : 0.0813175 *****
 -1.9714 : 0.0794405 *****
 -1.7627 : 0.0771240 *****
 -1.5541 : 0.0738399 *****
 -1.3454 : 0.0687762 *****
 -1.1367 : 0.0612591 *****
 -0.9281 : 0.0512267 *****
 -0.7194 : 0.0395285 *****
 -0.5108 : 0.0277597 *****
 -0.3021 : 0.0176012 ****
 -0.0935 : 0.0100778 **
 0.1152 : 0.0052666 *
 0.3238 : 0.0025720
 0.5325 : 0.0012202
 0.7412 : 0.0005949
 0.9498 : 0.0003217
 1.1585 : 0.0002129
 1.3671 : 0.0001951
 1.5758 : 0.0002875
 1.7844 : 0.0008152

After EM-cycle Mean(theta)= -1.96776 stdev(theta)= 0.86814

Log linear model parameters for skill profile distribution in class 1:

Intercept term: 1.75853
Skill 1 (1.) moment: -0.80829
Skill 1 (2.) moment: -0.43095
Skill 1 (3.) moment: -0.07173
Skill 1 (4.) moment: 0.01846
Skill 1 (5.) moment: 0.00583

RMSEA of loglinear skill profile model: 0.0010677

Class: "2" size: 0.3264684

Scale: 1

Targeted Centered Mean: -0.3652630 .. Stdev: 1.0725690

-2.9506 :	0.0088997 **
-2.7273 :	0.0053057 *
-2.5039 :	0.0057254 *
-2.2806 :	0.0087504 **
-2.0572 :	0.0156761 ****
-1.8339 :	0.0286450 *****
-1.6105 :	0.0485538 *****
-1.3872 :	0.0721229 *****
-1.1639 :	0.0915884 *****
-0.9405 :	0.0995598 *****
-0.7172 :	0.0946379 *****
-0.4938 :	0.0814956 *****
-0.2705 :	0.0663920 *****
-0.0471 :	0.0535421 *****
0.1762 :	0.0445458 *****
0.3996 :	0.0394471 *****
0.6229 :	0.0377490 *****
0.8463 :	0.0387675 *****
1.0696 :	0.0412555 *****
1.2929 :	0.0424529 *****
1.5163 :	0.0378672 *****
1.7396 :	0.0250624 *****
1.9630 :	0.0099997 **
2.1863 :	0.0018437
2.4097 :	0.0001129
2.6330 :	0.0000015

After EM-cycle Mean(theta)= -0.36298 stdev(theta)= 1.07650

Log linear model parameters for skill profile distribution in class 2:

Intercept term: 2.01882
Skill 1 (1.) moment: -0.67768
Skill 1 (2.) moment: 0.05849
Skill 1 (3.) moment: 0.21385
Skill 1 (4.) moment: -0.02800
Skill 1 (5.) moment: -0.01495

RMSEA of loglinear skill profile model: 0.0022581

Class: "3" size: 0.3442474

Scale: 1

Targeted Centered Mean: 0.0403144 .. Stdev: 1.0148234

-2.7415 :	0.0010510
-2.5206 :	0.0038915 *
-2.2996 :	0.0094406 **
-2.0786 :	0.0168499 ***
-1.8576 :	0.0243252 *****
-1.6367 :	0.0306270 *****
-1.4157 :	0.0356304 *****
-1.1947 :	0.0399406 *****
-0.9738 :	0.0443607 *****
-0.7528 :	0.0495863 *****
-0.5318 :	0.0560728 *****
-0.3109 :	0.0639214 *****
-0.0899 :	0.0726986 *****
0.1311 :	0.0812189 *****
0.3520 :	0.0874704 *****
0.5730 :	0.0889800 *****
0.7940 :	0.0837946 *****
1.0149 :	0.0717429 *****
1.2359 :	0.0550541 *****
1.4569 :	0.0375397 *****
1.6778 :	0.0227163 *****
1.8988 :	0.0122962 ***
2.1198 :	0.0060673 *
2.3407 :	0.0028167
2.5617 :	0.0012885
2.7827 :	0.0006182

After EM-cycle Mean(theta)= 0.04078 stdev(theta)= 1.01836

Log linear model parameters for skill profile distribution in class 3:

Intercept term: 2.33039
Skill 1 (1.) moment: 0.34881
Skill 1 (2.) moment: -0.13063
Skill 1 (3.) moment: -0.09691
Skill 1 (4.) moment: -0.00965
Skill 1 (5.) moment: 0.00444

RMSEA of loglinear skill profile model: 0.0014854

Skill patterns :

0 : 0. :	0.010901321944 :	0.002905486066 :	0.000361791857
1 : 1. :	0.017622102193 :	0.001732127405 :	0.001339629566
2 : 2. :	0.022952204485 :	0.001869161281 :	0.003249911167
3 : 3. :	0.026018995703 :	0.002856725041 :	0.005800543880
4 : 4. :	0.027181696406 :	0.005117763083 :	0.008373890036
5 : 5. :	0.027222747467 :	0.009351685677 :	0.010543256918
6 : 6. :	0.026776574162 :	0.015851272855 :	0.012265675447
7 : 7. :	0.026158490290 :	0.023545833877 :	0.013749455844
8 : 8. :	0.025395710789 :	0.029900714534 :	0.015271047424
9 : 9. :	0.024314303861 :	0.032503131503 :	0.017069965232
10 : 10. :	0.022646909847 :	0.030896290854 :	0.019302923683
11 : 11. :	0.020171655649 :	0.026605727856 :	0.022004773115
12 : 12. :	0.016868142253 :	0.021674875123 :	0.025026314846
13 : 13. :	0.013016113622 :	0.017479805855 :	0.027959394827
14 : 14. :	0.009140847736 :	0.014542797182 :	0.030111440516

15 : 15. : 0.005795790285 : 0.012878236350 : 0.030631147622
 16 : 16. : 0.003318462964 : 0.012323858210 : 0.028846071837
 17 : 17. : 0.001734216940 : 0.012656354909 : 0.024697314484
 18 : 18. : 0.000846920766 : 0.013468620428 : 0.018952227388
 19 : 19. : 0.000401783785 : 0.013859516522 : 0.012922947781
 20 : 20. : 0.000195891716 : 0.012362453684 : 0.007820025422
 21 : 21. : 0.000105931176 : 0.008182070709 : 0.004232930653
 22 : 22. : 0.000070111888 : 0.003264594812 : 0.002088642311
 23 : 23. : 0.000064236468 : 0.000601907386 : 0.000969657587
 24 : 24. : 0.000094658908 : 0.000036863665 : 0.000443566974
 25 : 25. : 0.000268425909 : 0.000000502532 : 0.000212828973

Final Item Parameter Estimates

Class "1" :

item i.label slopes ... | intercept ; difficulty...

1	D_B_12_AV	1.00000		0.41050		-0.41050	
2	D_B_149_AV	1.00000		0.74694		-0.74694	
3	D_B_163_AV	1.00000		1.30339		-1.30339	
4	D_B_177_AV	1.00000		3.38566		-3.38566	
5	D_B_189_AV	1.00000		-0.01966		0.01966	
6	D_B_198_AV	1.00000		1.08201		-1.08201	
7	D_B_27_AV	1.00000		-0.54745		0.54745	
8	D_B_42_AV	1.00000		-0.77820		0.77820	
9	D_B_58_AV	1.00000		1.71676		-1.71676	
10	D_B_74_AV	1.00000		1.26525		-1.26525	
11	D_B_90_AV	1.00000		0.60219		-0.60219	
12	D_B_105_AV	1.00000		-0.29605		0.29605	
13	D_B_120_AV	1.00000		0.13429		-0.13429	
14	D_B_135_AV	1.00000		-0.19515		0.19515	
15	D_B_57_AV	1.00000		0.17484		-0.17484	
16	D_B_11_AV	1.00000		0.72956		-0.72956	
17	D_B_148_AV	1.00000		-0.00427		0.00427	
18	D_B_162_AV	1.00000		-0.17611		0.17611	
19	D_B_176_AV	1.00000		-0.90354		0.90354	
20	D_B_188_AV	1.00000		0.51846		-0.51846	
21	D_B_206_AV	1.00000		-0.43117		0.43117	
22	D_B_26_AV	1.00000		0.09960		-0.09960	
23	D_B_41_AV	1.00000		0.24291		-0.24291	
24	D_B_73_AV	1.00000		0.41364		-0.41364	
25	D_B_89_AV	1.00000		-1.91790		1.91790	
26	D_B_104_AV	1.00000		0.39704		-0.39704	
27	D_B_119_AV	1.00000		0.02397		-0.02397	
28	D_B_134_AV	1.00000		-1.04568		1.04568	
29	D_B_13_AV	1.00000		-0.64542		0.64542	
30	D_B_150_AV	1.00000		0.22417		-0.22417	
31	D_B_164_AV	1.00000		0.38522		-0.38522	
32	D_B_178_AV	1.00000		-0.22505		0.22505	
33	D_B_190_AV	1.00000		-3.09991		3.09991	
34	D_B_199_AV	1.00000		-1.60955		1.60955	
35	D_B_28_AV	1.00000		-0.05163		0.05163	
36	D_B_43_AV	1.00000		-0.97708		0.97708	
37	D_B_59_AV	1.00000		-0.84309		0.84309	
38	D_B_75_AV	1.00000		-0.73008		0.73008	
39	D_B_91_AV	1.00000		1.79575		-1.79575	
40	D_B_106_AV	1.00000		-1.46199		1.46199	
41	D_B_121_AV	1.00000		-0.90490		0.90490	
42	D_B_136_AV	1.00000		1.20310		-1.20310	

By skill means for slope 0: 1.0000000

By skill means for betas 0: -0.0004859

Average difficulty across: -0.00020 sum of categories 42

Values marked with * are fixed values

Values marked with # represent equality constrained parameters

Class "2" :

item i.label slopes ... | intercept ; difficulty...

1	D_B_12_AV_	1.00000	-1.21822	1.21822
2	D_B_149_AV_	1.00000	-1.69757	1.69757
3	D_B_163_AV_	1.00000	0.30097	-0.30097
4	D_B_177_AV_	1.00000	1.82342	-1.82342
5	D_B_189_AV_	1.00000	-1.06177	1.06177
6	D_B_198_AV_	1.00000	-0.04616	0.04616
7	D_B_27_AV_	1.00000	-1.94914	1.94914
8	D_B_42_AV_	1.00000	-1.97733	1.97733
9	D_B_58_AV_	1.00000	0.56301	-0.56301
10	D_B_74_AV_	1.00000	0.29475	-0.29475
11	D_B_90_AV_	1.00000	-0.77059	0.77059
12	D_B_105_AV_	1.00000	-1.78701	1.78701
13	D_B_120_AV_	1.00000	-1.93673	1.93673
14	D_B_135_AV_	1.00000	-1.70903	1.70903
15	D_B_57_AV_	1.00000	-1.34641	1.34641
16	D_B_11_AV_	1.00000	-0.58780	0.58780
17	D_B_148_AV_	1.00000	-1.46516	1.46516
18	D_B_162_AV_	1.00000	-1.69980	1.69980
19	D_B_176_AV_	1.00000	-2.23363	2.23363
20	D_B_188_AV_	1.00000	-1.86241	1.86241
21	D_B_206_AV_	1.00000	-1.38417	1.38417
22	D_B_26_AV_	1.00000	-1.82262	1.82262
23	D_B_41_AV_	1.00000	-1.48834	1.48834
24	D_B_73_AV_	1.00000	-0.95005	0.95005
25	D_B_89_AV_	1.00000	-3.05913	3.05913
26	D_B_104_AV_	1.00000	-1.41557	1.41557
27	D_B_119_AV_	1.00000	-1.69276	1.69276
28	D_B_134_AV_	1.00000	-2.60224	2.60224
29	D_B_13_AV_	1.00000	-2.91333	2.91333
30	D_B_150_AV_	1.00000	-1.81261	1.81261
31	D_B_164_AV_	1.00000	-1.49865	1.49865
32	D_B_178_AV_	1.00000	-2.80907	2.80907
33	D_B_190_AV_	1.00000	-3.23270	3.23270
34	D_B_199_AV_	1.00000	-3.21031	3.21031
35	D_B_28_AV_	1.00000	-1.77126	1.77126
36	D_B_43_AV_	1.00000	-2.14958	2.14958
37	D_B_59_AV_	1.00000	-2.13137	2.13137
38	D_B_75_AV_	1.00000	-2.58115	2.58115
39	D_B_91_AV_	1.00000	-0.24889	0.24889
40	D_B_106_AV_	1.00000	-3.59304	3.59304
41	D_B_121_AV_	1.00000	-1.61027	1.61027
42	D_B_136_AV_	1.00000	-0.47140	0.47140

By skill means for slope 0: 1.0000000

By skill means for betas 0: -1.5432173

Average difficulty across: -1.54322 sum of categories 42

Values marked with * are fixed values

Values marked with # represent equality constrained parameters

Class "3" :

item i.label slopes ... | intercept ; difficulty...

1	D_B_12_AV_	1.00000	-1.70420	1.70420
---	------------	---------	----------	---------

2 D_B_149_AV	1.00000	-1.55160	1.55160
3 D_B_163_AV	1.00000	-0.17090	0.17090
4 D_B_177_AV	1.00000	1.14443	-1.14443
5 D_B_189_AV	1.00000	-0.93866	0.93866
6 D_B_198_AV	1.00000	-0.30803	0.30803
7 D_B_27_AV	1.00000	-2.00563	2.00563
8 D_B_42_AV	1.00000	-2.07006	2.07006
9 D_B_58_AV	1.00000	0.14540	-0.14540
10 D_B_74_AV	1.00000	-0.27918	0.27918
11 D_B_90_AV	1.00000	-0.81522	0.81522
12 D_B_105_AV	1.00000	-2.15234	2.15234
13 D_B_120_AV	1.00000	-1.69474	1.69474
14 D_B_135_AV	1.00000	-1.96965	1.96965
15 D_B_57_AV	1.00000	-1.49724	1.49724
16 D_B_11_AV	1.00000	-0.35600	0.35600
17 D_B_148_AV	1.00000	-1.35356	1.35356
18 D_B_162_AV	1.00000	-1.52957	1.52957
19 D_B_176_AV	1.00000	-1.59918	1.59918
20 D_B_188_AV	1.00000	-1.21261	1.21261
21 D_B_206_AV	1.00000	-1.59463	1.59463
22 D_B_26_AV	1.00000	-1.54768	1.54768
23 D_B_41_AV	1.00000	-1.13965	1.13965
24 D_B_73_AV	1.00000	-1.21316	1.21316
25 D_B_89_AV	1.00000	-2.99718	2.99718
26 D_B_104_AV	1.00000	-1.18239	1.18239
27 D_B_119_AV	1.00000	-1.57400	1.57400
28 D_B_134_AV	1.00000	-2.58424	2.58424
29 D_B_13_AV	1.00000	-1.85650	1.85650
30 D_B_150_AV	1.00000	-1.54149	1.54149
31 D_B_164_AV	1.00000	-0.84194	0.84194
32 D_B_178_AV	1.00000	-2.19775	2.19775
33 D_B_190_AV	1.00000	-3.01628	3.01628
34 D_B_199_AV	1.00000	-2.52281	2.52281
35 D_B_28_AV	1.00000	-1.90662	1.90662
36 D_B_43_AV	1.00000	-2.50777	2.50777
37 D_B_59_AV	1.00000	-2.06833	2.06833
38 D_B_75_AV	1.00000	-2.15005	2.15005
39 D_B_91_AV	1.00000	0.26068	-0.26068
40 D_B_106_AV	1.00000	-4.30816	4.30816
41 D_B_121_AV	1.00000	-1.89996	1.89996
42 D_B_136_AV	1.00000	-0.45509	0.45509

By skill means for slope 0: 1.0000000

By skill means for betas 0: -1.4943703

Average difficulty across: -1.49437 sum of categories 42

Values marked with * are fixed values

Values marked with # represent equality constrained parameters

Final Item Parameter Standard Errors

Class "1" :

item i.label slopes ... | intercept ; difficulty...

1	D_B_12_AV_	1.#INFO		0.25647
2	D_B_149_AV_	1.#INFO		0.23771
3	D_B_163_AV_	1.#INFO		0.24448
4	D_B_177_AV_	1.#INFO		0.27708
5	D_B_189_AV_	1.#INFO		0.33310
6	D_B_198_AV_	1.#INFO		0.26145
7	D_B_27_AV_	1.#INFO		0.36394
8	D_B_42_AV_	1.#INFO		0.38666
9	D_B_58_AV_	1.#INFO		0.21922
10	D_B_74_AV_	1.#INFO		0.21967
11	D_B_90_AV_	1.#INFO		0.25864
12	D_B_105_AV_	1.#INFO		0.35333
13	D_B_120_AV_	1.#INFO		0.28537
14	D_B_135_AV_	1.#INFO		0.31476
15	D_B_57_AV_	1.#INFO		0.29587
16	D_B_11_AV_	1.#INFO		0.26021
17	D_B_148_AV_	1.#INFO		0.29462
18	D_B_162_AV_	1.#INFO		0.32475
19	D_B_176_AV_	1.#INFO		0.38197
20	D_B_188_AV_	1.#INFO		0.26283
21	D_B_206_AV_	1.#INFO		0.34840
22	D_B_26_AV_	1.#INFO		0.28837
23	D_B_41_AV_	1.#INFO		0.28390
24	D_B_73_AV_	1.#INFO		0.27520
25	D_B_89_AV_	1.#INFO		0.61477
26	D_B_104_AV_	1.#INFO		0.28220
27	D_B_119_AV_	1.#INFO		0.29781
28	D_B_134_AV_	1.#INFO		0.40563
29	D_B_13_AV_	1.#INFO		0.34563
30	D_B_150_AV_	1.#INFO		0.27908
31	D_B_164_AV_	1.#INFO		0.28564
32	D_B_178_AV_	1.#INFO		0.29756
33	D_B_190_AV_	1.#INFO		1.22885
34	D_B_199_AV_	1.#INFO		0.55025
35	D_B_28_AV_	1.#INFO		0.31609
36	D_B_43_AV_	1.#INFO		0.48248
37	D_B_59_AV_	1.#INFO		0.38108
38	D_B_75_AV_	1.#INFO		0.38352
39	D_B_91_AV_	1.#INFO		0.21346
40	D_B_106_AV_	1.#INFO		0.48404
41	D_B_121_AV_	1.#INFO		0.40149
42	D_B_136_AV_	1.#INFO		0.23249

Class "2" :

item i.label slopes ... | intercept ; difficulty...

1	D_B_12_AV_	1.#INFO		0.25880
2	D_B_149_AV_	1.#INFO		0.29904
3	D_B_163_AV_	1.#INFO		0.25293
4	D_B_177_AV_	1.#INFO		0.26202
5	D_B_189_AV_	1.#INFO		0.28300
6	D_B_198_AV_	1.#INFO		0.27439
7	D_B_27_AV_	1.#INFO		0.34482
8	D_B_42_AV_	1.#INFO		0.32976
9	D_B_58_AV_	1.#INFO		0.23839
10	D_B_74_AV_	1.#INFO		0.23634
11	D_B_90_AV_	1.#INFO		0.25413

12 D_B_105_AV	1.#INFO		0.34724
13 D_B_120_AV	1.#INFO		0.32429
14 D_B_135_AV	1.#INFO		0.31766
15 D_B_57_AV	1.#INFO		0.29683
16 D_B_11_AV	1.#INFO		0.24081
17 D_B_148_AV	1.#INFO		0.27817
18 D_B_162_AV	1.#INFO		0.32611
19 D_B_176_AV	1.#INFO		0.34516
20 D_B_188_AV	1.#INFO		0.32031
21 D_B_206_AV	1.#INFO		0.28265
22 D_B_26_AV	1.#INFO		0.30456
23 D_B_41_AV	1.#INFO		0.27951
24 D_B_73_AV	1.#INFO		0.26068
25 D_B_89_AV	1.#INFO		0.47952
26 D_B_104_AV	1.#INFO		0.28547
27 D_B_119_AV	1.#INFO		0.32227
28 D_B_134_AV	1.#INFO		0.39886
29 D_B_13_AV	1.#INFO		0.40637
30 D_B_150_AV	1.#INFO		0.33527
31 D_B_164_AV	1.#INFO		0.29676
32 D_B_178_AV	1.#INFO		0.42670
33 D_B_190_AV	1.#INFO		0.48773
34 D_B_199_AV	1.#INFO		0.48401
35 D_B_28_AV	1.#INFO		0.30787
36 D_B_43_AV	1.#INFO		0.34683
37 D_B_59_AV	1.#INFO		0.31820
38 D_B_75_AV	1.#INFO		0.39200
39 D_B_91_AV	1.#INFO		0.22194
40 D_B_106_AV	1.#INFO		0.63412
41 D_B_121_AV	1.#INFO		0.29165
42 D_B_136_AV	1.#INFO		0.23353

Class "3" :

item i.label slopes ... | intercept ; difficulty...

1 D_B_12_AV	1.#INFO		0.25611
2 D_B_149_AV	1.#INFO		0.27199
3 D_B_163_AV	1.#INFO		0.25466
4 D_B_177_AV	1.#INFO		0.26686
5 D_B_189_AV	1.#INFO		0.23735
6 D_B_198_AV	1.#INFO		0.24579
7 D_B_27_AV	1.#INFO		0.28664
8 D_B_42_AV	1.#INFO		0.28870
9 D_B_58_AV	1.#INFO		0.23480
10 D_B_74_AV	1.#INFO		0.23194
11 D_B_90_AV	1.#INFO		0.24167
12 D_B_105_AV	1.#INFO		0.29520
13 D_B_120_AV	1.#INFO		0.25114
14 D_B_135_AV	1.#INFO		0.28313
15 D_B_57_AV	1.#INFO		0.26904
16 D_B_11_AV	1.#INFO		0.23561
17 D_B_148_AV	1.#INFO		0.25056
18 D_B_162_AV	1.#INFO		0.27530
19 D_B_176_AV	1.#INFO		0.24191
20 D_B_188_AV	1.#INFO		0.23504
21 D_B_206_AV	1.#INFO		0.25199
22 D_B_26_AV	1.#INFO		0.24942
23 D_B_41_AV	1.#INFO		0.22739
24 D_B_73_AV	1.#INFO		0.23330
25 D_B_89_AV	1.#INFO		0.38098
26 D_B_104_AV	1.#INFO		0.22650

27 D_B_119_AV	1.#INFO		0.25735
28 D_B_134_AV	1.#INFO		0.34090
29 D_B_13_AV	1.#INFO		0.25899
30 D_B_150_AV	1.#INFO		0.25747
31 D_B_164_AV	1.#INFO		0.23683
32 D_B_178_AV	1.#INFO		0.28487
33 D_B_190_AV	1.#INFO		0.38882
34 D_B_199_AV	1.#INFO		0.33984
35 D_B_28_AV	1.#INFO		0.26400
36 D_B_43_AV	1.#INFO		0.34165
37 D_B_59_AV	1.#INFO		0.27682
38 D_B_75_AV	1.#INFO		0.32322
39 D_B_91_AV	1.#INFO		0.21347
40 D_B_106_AV	1.#INFO		0.72851
41 D_B_121_AV	1.#INFO		0.28449
42 D_B_136_AV	1.#INFO		0.21626

Expected category frequencies by class and scale:

Class: "1"

class.size: 0.329284

item	cats:	tried	0	1	Expected	
1 D_B_12_AV	1	107.07	0.7945	0.2055	0.2055	
2 D_B_149_AV	1	108.07	0.7409	0.2591	0.2591	
3 D_B_163_AV	1	91.07	0.6376	0.3624	0.3624	
4 D_B_177_AV	1	97.07	0.2171	0.7829	0.7829	
5 D_B_189_AV	1	82.07	0.8538	0.1462	0.1462	
6 D_B_198_AV	1	80.07	0.6878	0.3122	0.3122	
7 D_B_27_AV	1	98.07	0.9082	0.0918	0.0918	
8 D_B_42_AV	1	106.07	0.9246	0.0754	0.0754	
9 D_B_58_AV	1	103.07	0.5537	0.4463	0.4463	
10 D_B_74_AV	1	107.07	0.6451	0.3549	0.3549	
11 D_B_90_AV	1	97.07	0.7734	0.2266	0.2266	
12 D_B_105_AV	1	90.07	0.8890	0.1110	0.1110	
13 D_B_120_AV	1	102.07	0.8334	0.1666	0.1666	
14 D_B_135_AV	1	100.07	0.8701	0.1299	0.1299	
15 D_B_57_AV	1	98.07	0.8361	0.1639	0.1639	
16 D_B_11_AV	1	96.00	0.7604	0.2396	0.2396	
17 D_B_148_AV	1	99.07	0.8479	0.1521	0.1521	
18 D_B_162_AV	1	95.07	0.8730	0.1270	0.1270	
19 D_B_176_AV	1	112.07	0.9286	0.0714	0.0714	
20 D_B_188_AV	1	98.07	0.7749	0.2251	0.2251	
21 D_B_206_AV	1	95.07	0.8941	0.1059	0.1059	
22 D_B_26_AV	1	100.07	0.8394	0.1606	0.1606	
23 D_B_41_AV	1	97.07	0.8241	0.1759	0.1759	
24 D_B_73_AV	1	93.07	0.7951	0.2049	0.2049	
25 D_B_89_AV	1	107.00	0.9720	0.0280	0.0280	
26 D_B_104_AV	1	92.07	0.8037	0.1963	0.1963	
27 D_B_119_AV	1	96.07	0.8431	0.1569	0.1569	
28 D_B_134_AV	1	110.07	0.9364	0.0636	0.0636	
29 D_B_13_AV	1	116.07	0.9132	0.0868	0.0868	
30 D_B_150_AV	1	102.07	0.8237	0.1763	0.1763	
31 D_B_164_AV	1	93.07	0.8058	0.1942	0.1942	
32 D_B_178_AV	1	119.07	0.8740	0.1260	0.1260	
33 D_B_190_AV	1	114.07	0.9912	0.0088	0.0088	
34 D_B_199_AV	1	113.07	0.9640	0.0360	0.0360	
35 D_B_28_AV	1	97.07	0.8661	0.1339	0.1339	
36 D_B_43_AV	1	93.07	0.9463	0.0537	0.0537	
37 D_B_59_AV	1	109.07	0.9267	0.0733	0.0733	
38 D_B_75_AV	1	101.00	0.9208	0.0792	0.0792	

39	D_B_91_AV_	1	104.07	0.5388	0.4612	0.4612
40	D_B_106_AV	1	114.07	0.9562	0.0438	0.0438
41	D_B_121_AV	1	106.07	0.9340	0.0660	0.0660
42	D_B_136_AV	1	95.07	0.6416	0.3584	0.3584

Class: "2"

class.size: 0.326468

item	cats:	tried	0	1	Expected	
1	D_B_12_AV_	1	108.01	0.7870	0.2130	0.2130
2	D_B_149_AV	1	101.01	0.8416	0.1584	0.1584
3	D_B_163_AV	1	86.01	0.5117	0.4883	0.4883
4	D_B_177_AV	1	106.01	0.2359	0.7641	0.7641
5	D_B_189_AV	1	89.01	0.7641	0.2359	0.2359
6	D_B_198_AV	1	73.01	0.5754	0.4246	0.4246
7	D_B_27_AV_	1	94.01	0.8830	0.1170	0.1170
8	D_B_42_AV_	1	105.01	0.8857	0.1143	0.1143
9	D_B_58_AV_	1	95.01	0.4527	0.5473	0.5473
10	D_B_74_AV_	1	96.01	0.5209	0.4791	0.4791
11	D_B_90_AV_	1	97.01	0.7217	0.2783	0.2783
12	D_B_105_AV	1	88.01	0.8636	0.1364	0.1364
13	D_B_120_AV	1	103.01	0.8738	0.1262	0.1262
14	D_B_135_AV	1	95.01	0.8526	0.1474	0.1474
15	D_B_57_AV_	1	94.01	0.8191	0.1809	0.1809
16	D_B_11_AV_	1	100.00	0.6900	0.3100	0.3100
17	D_B_148_AV	1	107.01	0.8224	0.1776	0.1776
18	D_B_162_AV	1	94.01	0.8616	0.1384	0.1384
19	D_B_176_AV	1	111.01	0.9009	0.0991	0.0991
20	D_B_188_AV	1	98.01	0.8673	0.1327	0.1327
21	D_B_206_AV	1	103.01	0.8252	0.1748	0.1748
22	D_B_26_AV_	1	105.01	0.8571	0.1429	0.1429
23	D_B_41_AV_	1	103.01	0.8155	0.1845	0.1845
24	D_B_73_AV_	1	97.01	0.7422	0.2578	0.2578
25	D_B_89_AV_	1	105.00	0.9524	0.0476	0.0476
26	D_B_104_AV	1	102.01	0.8039	0.1961	0.1961
27	D_B_119_AV	1	93.01	0.8494	0.1506	0.1506
28	D_B_134_AV	1	110.01	0.9273	0.0727	0.0727
29	D_B_13_AV_	1	119.01	0.9411	0.0589	0.0589
30	D_B_150_AV	1	102.01	0.8628	0.1372	0.1372
31	D_B_164_AV	1	94.01	0.8191	0.1809	0.1809
32	D_B_178_AV	1	117.01	0.9402	0.0598	0.0598
33	D_B_190_AV	1	118.01	0.9576	0.0424	0.0424
34	D_B_199_AV	1	116.01	0.9568	0.0432	0.0432
35	D_B_28_AV_	1	101.01	0.8515	0.1485	0.1485
36	D_B_43_AV_	1	104.01	0.8942	0.1058	0.1058
37	D_B_59_AV_	1	116.01	0.8879	0.1121	0.1121
38	D_B_75_AV_	1	111.00	0.9279	0.0721	0.0721
39	D_B_91_AV_	1	110.01	0.6091	0.3909	0.3909
40	D_B_106_AV	1	110.01	0.9727	0.0273	0.0273
41	D_B_121_AV	1	109.01	0.8440	0.1560	0.1560
42	D_B_136_AV	1	105.01	0.6666	0.3334	0.3334

Class: "3"

class.size: 0.344247

item	cats	tried	0	1	Expected	
1 D_B_12_AV_	1	107.92	0.7961	0.2039	0.2039	
2 D_B_149_AV_	1	90.92	0.7800	0.2200	0.2200	
3 D_B_163_AV_	1	94.92	0.5049	0.4951	0.4951	
4 D_B_177_AV_	1	107.92	0.2587	0.7413	0.7413	
5 D_B_189_AV_	1	96.92	0.6595	0.3405	0.3405	
6 D_B_198_AV_	1	85.92	0.5345	0.4655	0.4655	
7 D_B_27_AV_	1	102.92	0.8348	0.1652	0.1652	
8 D_B_42_AV_	1	104.92	0.8475	0.1525	0.1525	
9 D_B_58_AV_	1	105.92	0.4430	0.5570	0.5570	
10 D_B_74_AV_	1	104.92	0.5235	0.4765	0.4765	
11 D_B_90_AV_	1	93.92	0.6380	0.3620	0.3620	
12 D_B_105_AV_	1	102.92	0.8543	0.1457	0.1457	
13 D_B_120_AV_	1	113.92	0.7893	0.2107	0.2107	
14 D_B_135_AV_	1	104.92	0.8380	0.1620	0.1620	
15 D_B_57_AV_	1	98.92	0.7784	0.2216	0.2216	
16 D_B_11_AV_	1	95.00	0.5579	0.4421	0.4421	
17 D_B_148_AV_	1	106.92	0.7482	0.2518	0.2518	
18 D_B_162_AV_	1	99.92	0.7906	0.2094	0.2094	
19 D_B_176_AV_	1	117.92	0.7880	0.2120	0.2120	
20 D_B_188_AV_	1	106.92	0.7201	0.2799	0.2799	
21 D_B_206_AV_	1	107.92	0.7783	0.2217	0.2217	
22 D_B_26_AV_	1	106.92	0.7763	0.2237	0.2237	
23 D_B_41_AV_	1	109.92	0.7096	0.2904	0.2904	
24 D_B_73_AV_	1	108.92	0.7253	0.2747	0.2747	
25 D_B_89_AV_	1	115.00	0.9304	0.0696	0.0696	
26 D_B_104_AV_	1	112.92	0.7173	0.2827	0.2827	
27 D_B_119_AV_	1	108.92	0.7896	0.2104	0.2104	
28 D_B_134_AV_	1	111.92	0.9017	0.0983	0.0983	
29 D_B_13_AV_	1	118.92	0.8241	0.1759	0.1759	
30 D_B_150_AV_	1	103.92	0.7787	0.2213	0.2213	
31 D_B_164_AV_	1	95.92	0.6464	0.3536	0.3536	
32 D_B_178_AV_	1	121.92	0.8688	0.1312	0.1312	
33 D_B_190_AV_	1	118.92	0.9327	0.0673	0.0673	
34 D_B_199_AV_	1	111.92	0.9024	0.0976	0.0976	
35 D_B_28_AV_	1	113.92	0.8244	0.1756	0.1756	
36 D_B_43_AV_	1	103.92	0.8942	0.1058	0.1058	
37 D_B_59_AV_	1	119.92	0.8499	0.1501	0.1501	
38 D_B_75_AV_	1	98.00	0.8673	0.1327	0.1327	
39 D_B_91_AV_	1	112.92	0.4244	0.5756	0.5756	
40 D_B_106_AV_	1	108.92	0.9816	0.0184	0.0184	
41 D_B_121_AV_	1	101.92	0.8332	0.1668	0.1668	
42 D_B_136_AV_	1	105.92	0.5665	0.4335	0.4335	

Estimates of Person parameters (EAPs) by Raw Score by Skill in Class: "1"

Skill variable 1

	Average	Pooled	By Raw by Skill
Raw- Estimate Standard Deviation Sum of Weights	M(Theta Raw) S(Theta Raw) N(Raw Skill)		
Score			
0	-2.967706	0.379115	6.00
1	-2.900476	0.390560	5.00
2	-2.740353	0.436498	15.00
3	-2.587043	0.463769	15.00
4	-2.337668	0.485338	15.00
5	-2.125853	0.484845	13.00
6	-1.811302	0.469154	11.00
7	-1.719495	0.455672	5.00
8	-1.528203	0.430325	10.00
9	-1.356720	0.405252	7.00
10	-1.067003	0.390447	5.00
11	-1.197566	0.371845	5.00
12	-0.885283	0.369698	6.00
13	-0.940432	0.348564	2.00
14	-0.550600	0.354546	2.07
15	-0.578664	0.347163	1.00
16	*****	*****	***** **
17	-0.473418	0.331962	1.00
18	*****	*****	***** **
19	*****	*****	***** **
20	*****	*****	***** **
21	-0.062285	0.336295	2.00
22	*****	*****	***** **
23	*****	*****	***** **
24	*****	*****	***** **
25	*****	*****	***** **
26	*****	*****	***** **
27	*****	*****	***** **
28	*****	*****	***** **
29	*****	*****	***** **
30	*****	*****	***** **
31	*****	*****	***** **
32	*****	*****	***** **
33	*****	*****	***** **
34	*****	*****	***** **
35	*****	*****	***** **
36	*****	*****	***** **
37	*****	*****	***** **
38	*****	*****	***** **
39	*****	*****	***** **
40	*****	*****	***** **
41	*****	*****	***** **
42	*****	*****	***** **

e: -1.97550 vb: 0.51000 vw: 0.19031 vt: 0.70031 w: 126.07111

Derived Reliability: 0.7282443

Estimates of Person parameters (EAPs) by Raw Score by Skill in Class: "2"

Skill variable 1

	Average	Pooled	By Raw by Skill
Raw- Estimate Standard Deviation Sum of Weights			
Score M(Theta Raw) S(Theta Raw) N(Raw Skill)			
0	-1.895825	0.639067	5.00
1	-1.525346	0.576167	10.00
2	-1.252892	0.511670	8.00
3	-1.026509	0.474634	13.00
4	-0.916882	0.447768	20.00
5	-0.594429	0.481791	15.00
6	-0.352992	0.492912	8.00
7	-0.411101	0.449480	5.00
8	0.303666	0.499263	6.00
9	0.089398	0.468079	4.00
10	0.537134	0.447890	5.00
11	0.706236	0.418946	6.00
12	0.640519	0.412244	3.00
13	0.980248	0.381796	3.00
14	1.040590	0.356572	3.01
15	1.296642	0.328732	3.00
16	1.310142	0.319024	3.00
17	*****	*****	*****
18	1.530458	0.278112	2.00
19	1.397879	0.290610	2.00
20	1.596662	0.265680	1.00
21	*****	*****	*****
22	1.840840	0.227570	1.00
23	*****	*****	*****
24	*****	*****	*****
25	*****	*****	*****
26	*****	*****	*****
27	*****	*****	*****
28	*****	*****	*****
29	*****	*****	*****
30	*****	*****	*****
31	*****	*****	*****
32	*****	*****	*****
33	*****	*****	*****
34	*****	*****	*****
35	*****	*****	*****
36	*****	*****	*****
37	*****	*****	*****
38	*****	*****	*****
39	*****	*****	*****
40	*****	*****	*****
41	*****	*****	*****
42	*****	*****	*****

e: -0.36527 vb: 0.88687 vw: 0.21954 vt: 1.10641 w: 126.00637

Derived Reliability: 0.8015742

Estimates of Person parameters (EAPs) by Raw Score by Skill in Class: "3"

Skill variable 1

	Average		Pooled		By Raw by Skill
--	---------	--	--------	--	-----------------

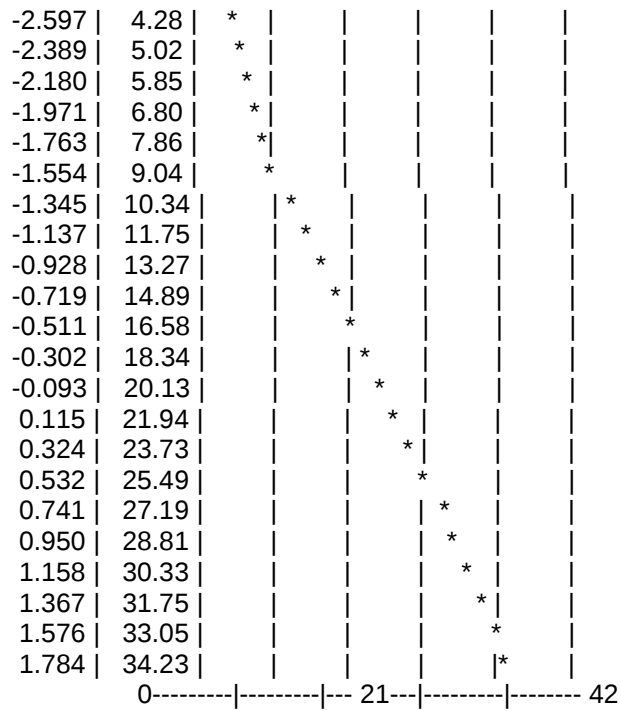
Raw- Score	Estimate M(Theta Raw)	Standard Deviation S(Theta Raw)	Sum of Weights N(Raw Skill)
0	-1.762653	0.487931	3.00
1	-1.499534	0.533847	7.00
2	-1.261737	0.548628	11.00
3	-1.007359	0.546389	6.00
4	-0.669830	0.535437	10.00
5	-0.321827	0.505913	10.00
6	-0.076389	0.472114	8.00
7	0.013373	0.440891	14.00
8	0.161155	0.414835	7.00
9	0.463314	0.393240	9.00
10	0.409295	0.378955	5.00
11	0.581496	0.365385	7.00
12	0.906473	0.364543	4.00
13	0.982756	0.356780	4.00
14	0.707644	0.331972	3.92
15	1.024968	0.337059	6.00
16	1.030063	0.325202	4.00
17	1.219540	0.328390	7.00
18	*****	*****	*****
19	1.382229	0.326338	1.00
20	1.273523	0.307697	2.00
21	*****	*****	*****
22	1.481757	0.309006	2.00
23	*****	*****	*****
24	1.714645	0.309760	1.00
25	*****	*****	*****
26	*****	*****	*****
27	2.230116	0.322500	1.00
28	*****	*****	*****
29	*****	*****	*****
30	*****	*****	*****
31	*****	*****	*****
32	*****	*****	*****
33	*****	*****	*****
34	*****	*****	*****
35	*****	*****	*****
36	*****	*****	*****
37	*****	*****	*****
38	*****	*****	*****
39	*****	*****	*****
40	*****	*****	*****
41	*****	*****	*****
42	*****	*****	*****

e: 0.04020 vb: 0.80399 vw: 0.19568 vt: 0.99967 w: 132.92252

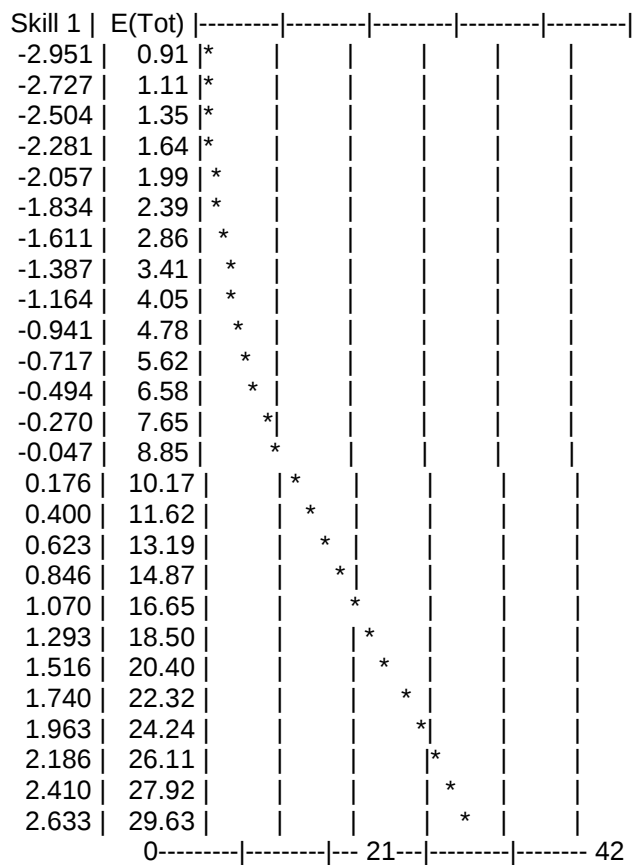
Derived Reliability: 0.8042534

Model based expected total (TCC) and skill score (STCCs) given skill pattern in class: "1"

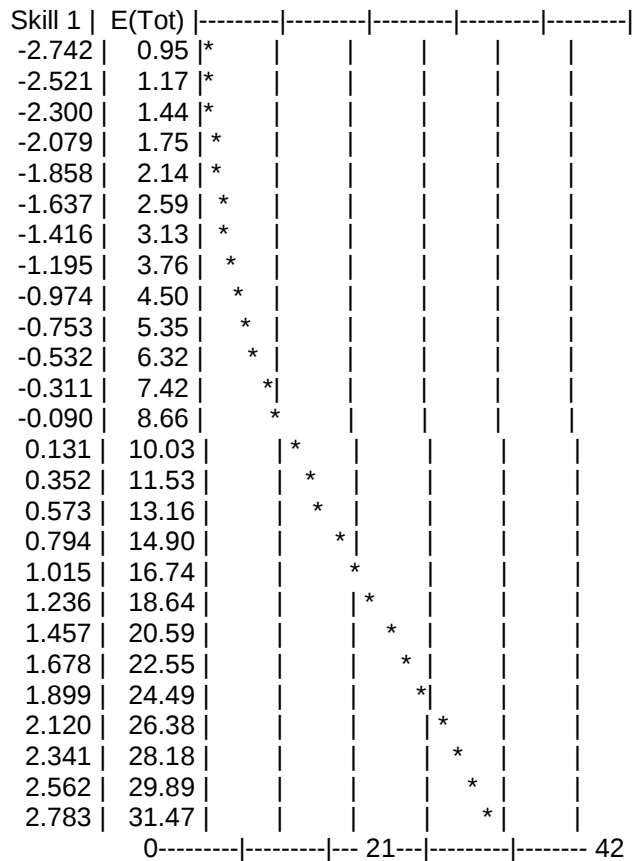
Skill 1	E(Tot)						
-3.432	2.19	*					
-3.223	2.61	*					
-3.015	3.09	*					
-2.806	3.64	*					



Model based expected total (TCC) and skill score (STCCs) given skill pattern in class: "2"



Model based expected total (TCC) and skill score (STCCs) given skill pattern in class: "3"



		chisquare	rmsea	meanD			
item	label Class:	"1" Class:	"2" Class:	"3"			
1	D_B_12_AV_	0.59 0.0483 -0.0002	0.77 0.0405 -0.0005	0.47 0.0314 -0.0002			
2	D_B_149_AV_	0.93 0.0641 -0.0001	0.75 0.0525 -0.0005	0.88 0.0363 -0.0002			
3	D_B_163_AV_	5.32#0.1287#-0.0002	2.27#0.1249# 0.0002	7.76#0.1971# 0.0001			
4	D_B_177_AV_	2.53#0.1411# 0.0005	2.93#0.1420# 0.0005	5.65#0.1777# 0.0003			
5	D_B_189_AV_	1.24 0.0591 -0.0003	0.70 0.0733 -0.0004	1.65 0.0886 -0.0000			
6	D_B_198_AV_	0.51 0.0548 -0.0001	1.56 0.0949 -0.0001	2.32#0.1138# 0.0001			
7	D_B_27_AV_	0.19 0.0299 -0.0002	0.90 0.0394 -0.0003	1.08 0.0670 -0.0001			
8	D_B_42_AV_	1.30 0.0497 -0.0002	0.68 0.0383 -0.0003	0.86 0.0426 -0.0001			
9	D_B_58_AV_	2.21#0.1020# 0.0001	2.39#0.1207# 0.0001	8.67#0.1753# 0.0002			
10	D_B_74_AV_	0.52 0.0526 -0.0001	2.90#0.1288# 0.0001	5.72#0.1701# 0.0000			
11	D_B_90_AV_	1.06 0.0485 -0.0002	0.46 0.0422 -0.0004	2.51#0.1130#-0.0000			
12	D_B_105_AV_	1.08 0.0554 -0.0002	1.97 0.0840 -0.0003	0.35 0.0385 -0.0001			
13	D_B_120_AV_	1.65 0.0635 -0.0002	1.71 0.0599 -0.0004	2.99 0.0911 -0.0001			
14	D_B_135_AV_	0.35 0.0530 -0.0002	1.32 0.0548 -0.0003	0.93 0.0494 -0.0001			
15	D_B_57_AV_	1.26 0.0809 -0.0002	2.67 0.0722 -0.0004	4.11 0.0966 -0.0000			
16	D_B_11_AV_	1.03 0.0697 -0.0002	0.10 0.0257 -0.0003	1.91#0.1137# 0.0002			
17	D_B_148_AV_	0.49 0.0293 -0.0003	0.38 0.0424 -0.0004	5.04#0.1131#-0.0000			
18	D_B_162_AV_	0.82 0.0361 -0.0002	2.48 0.0589 -0.0004	5.01#0.1045# 0.0000			
19	D_B_176_AV_	1.26 0.0381 -0.0002	3.62 0.0651 -0.0003	0.41 0.0398 -0.0000			
20	D_B_188_AV_	1.33 0.0804 -0.0002	0.44 0.0308 -0.0005	2.45 0.0571 -0.0000			
21	D_B_206_AV_	0.39 0.0339 -0.0003	0.81 0.0368 -0.0003	2.49 0.0671 -0.0001			
22	D_B_26_AV_	0.56 0.0402 -0.0002	1.13 0.0546 -0.0004	0.76 0.0376 -0.0001			
23	D_B_41_AV_	0.82 0.0448 -0.0002	0.37 0.0306 -0.0005	0.45 0.0398 0.0000			
24	D_B_73_AV_	1.52 0.0478 -0.0003	1.00 0.0535 -0.0004	1.15 0.0495 -0.0000			

25 D_B_89_AV_	1.51 0.0433 -0.0001	0.79 0.0345 -0.0002	1.13 0.0346 -0.0001
26 D_B_104_AV	2.01 0.0620 -0.0002	6.11#0.1085#-0.0005	0.23 0.0238 -0.0000
27 D_B_119_AV	1.60 0.0463 -0.0003	2.74 0.0732 -0.0004	2.95 0.0788 -0.0000
28 D_B_134_AV	0.62 0.0360 -0.0002	8.02 0.0679 -0.0003	2.48 0.0748 -0.0001
29 D_B_13_AV_	1.16 0.0433 -0.0002	0.58 0.0307 -0.0003	2.63 0.0606 -0.0000
30 D_B_150_AV	0.77 0.0620 -0.0002	6.67#0.1197#-0.0004	1.98 0.0666 -0.0001
31 D_B_164_AV	2.90 0.0790 -0.0002	1.68 0.0555 -0.0005	1.95 0.0644 0.0001
32 D_B_178_AV	4.04 0.0798 -0.0002	7.46 0.0740 -0.0003	2.33 0.0606 -0.0001
33 D_B_190_AV	4.14 0.0537 -0.0001	3.07 0.0541 -0.0002	2.82 0.0612 -0.0000
34 D_B_199_AV	2.95 0.0670 -0.0001	2.30 0.0488 -0.0002	2.60 0.0628 -0.0000
35 D_B_28_AV_	1.71 0.0558 -0.0001	1.83 0.0661 -0.0004	1.24 0.0580 -0.0001
36 D_B_43_AV_	3.03 0.0534 -0.0001	1.38 0.0624 -0.0003	1.75 0.0683 -0.0001
37 D_B_59_AV_	0.80 0.0286 -0.0002	0.88 0.0518 -0.0004	3.17 0.0787 -0.0001
38 D_B_75_AV_	0.23 0.0319 -0.0001	1.33 0.0510 -0.0003	2.66 0.0788 -0.0000
39 D_B_91_AV_	0.13 0.0279 0.0001	0.95 0.0620 -0.0004	0.27 0.0663 0.0003
40 D_B_106_AV	2.01 0.0543 -0.0001	3.14 0.0549 -0.0001	1.33 0.0225 -0.0001
41 D_B_121_AV	0.63 0.0199 -0.0002	1.85 0.0622 -0.0003	1.04 0.0509 -0.0001
42 D_B_136_AV	0.91 0.0478 -0.0002	0.95 0.0603 -0.0004	0.59 0.0379 0.0001

4.3 Modellpassung der Skala Soziale Erwünschtheit (GSP)

GESAMTSTICHPROBE					
d-, e-, m-Items	<i>AIC-Index</i>	16325,87	15750,02	15765,17	16003,74
	<i>BIC-Index</i>	16986,06	16564,39	16496,52	16735,09

Die grauen Markierungen kennzeichnen die vergleichsweise geringsten AIC/BIC -Werte der berechneten Modelle innerhalb eines Bereichs, die folglich die höchste Modellanpassungsgüte aufweisen.

4.4 Beobachtete Interkorrelations-Matrizen zu Modellvergleich I (d-Items, GSP)

d-Items				
Arbeitsverhalten (AV)	Flexibilität	Gewissenhaftigkeit	Handlungsoptimierung	
Flexibilität	1,000			
Gewissenhaftigkeit	-0,175	1,000		
Handlungsoptimierung	0,179	0,501	1,000	
Beruf. Orientierung (BO)	Führungsmotivation	Gestaltungsmotivation	Leistungsmotivation	
Führungsmotivation	1,000			
Gestaltungsmotivation	0,644	1,000		
Leistungsmotivation	0,466	0,481	1,000	
Psych. Konstitution (PK)	Belastbarkeit	Emotionale Stabilität	Selbstbewusstsein	
Belastbarkeit	1,000			
Emotionale Stabilität	0,615	1,000		
Selbstbewusstsein	0,527	0,632	1,000	
Original-Items (BIP)				
Arbeitsverhalten (AV)	Flexibilität	Gewissenhaftigkeit	Handlungsoptimierung	
Führungsmotivation	1,000			
Gestaltungsmotivation	0,159	1,000		
Leistungsmotivation	0,423	0,610	1,000	
Beruf. Orientierung (BO)	Führungsmotivation	Gestaltungsmotivation	Leistungsmotivation	
Führungsmotivation	1,000			
Gestaltungsmotivation	0,651	1,000		
Leistungsmotivation	0,604	0,631	1,000	
Psych. Konstitution (PK)	Belastbarkeit	Emotionale Stabilität	Selbstbewusstsein	
Belastbarkeit	1,000			
Emotionale Stabilität	0,711	1,000		
Selbstbewusstsein	0,569	0,619	1,000	

d-Items					
Soziale Kompetenz (SK)	Durchsetzungsstärke	Kontaktfähigkeit	Sensitivität	Soziabilität	Teamorientierung
Durchsetzungsstärke	1,00				
Kontaktfähigkeit	0,380	1,00			
Sensitivität	0,179	0,347	1,00		
Soziabilität	-0,130	0,241	0,341	1,00	
Teamorientierung	0,144	0,355	0,267	0,376	1,00
NEO-FFI (NEO)	Extraversion	Gewissenhaftigkeit	Neurotizismus	Offenheit	Verträglichkeit
Extraversion	1,00				
Gewissenhaftigkeit	0,062	1,00			
Neurotizismus	-0,224	-0,193	1,00		
Offenheit	0,116	-0,068	0,039	1,00	
Verträglichkeit	0,156	0,094	-0,103	0,107	1,00
Original-Items (BIP)					
Soziale Kompetenz (SK)	Durchsetzungsstärke	Kontaktfähigkeit	Sensitivität	Soziabilität	Teamorientierung
Durchsetzungsstärke	1,00				
Kontaktfähigkeit	0,425	1,00			
Sensitivität	0,389	0,541	1,00		
Soziabilität	0,165	0,405	0,539	1,00	
Teamorientierung	0,239	0,430	0,404	0,409	1,00
NEO-FFI (NEO)	Extraversion	Gewissenhaftigkeit	Neurotizismus	Offenheit	Verträglichkeit
Extraversion	1,00				
Gewissenhaftigkeit	0,282	1,00			
Neurotizismus	0,073	0,053	1,00		
Offenheit	0,369	0,224	0,166	1,00	
Verträglichkeit	0,371	0,36	0,134	0,359	1,00

4.5 Marginale IRT-basierte Reliabilitäten der Faktoren zur Skala Soziale Erwünschtheit (VG und GSP)

Versuchsgruppen 1, 2 und 3

Soziale Erwünschtheit SES-Skala	Kriterium	1- dimensional	3- dimensional	2- dimensional	2- dimensional
		SES-Skala	Pseudo-Item Skalen	d-&e-Skala, m-Skala	d-&m-Skala, e-Skala
VERSUCHSGRUPPE 1					
d-, e-, m-Items	SES	,692			
	d-Dimension		,690		
	e-Dimension		,780		,751
	m-Dimension		,651	,630	
	d&e-Dimension			,789	
	d&m-Dimension				,667
VERSUCHSGRUPPE 2					
d-, e-, m-Items	SES	,637			
	d-Dimension		,664		
	e-Dimension		,776		,743
	m-Dimension		,619	,615	
	d&e-Dimension			,840	
	d&m-Dimension				,468
VERSUCHSGRUPPE 3					
d-, e-, m-Items	SES	,583			
	d-Dimension		,591		
	e-Dimension		,740		,711
	m-Dimension		,537	,511	
	d&e-Dimension			,746	
	d&m-Dimension				,512

Gesamtstichprobe													
Soziale Erwünschtheit SES-Skala	Kriterium	1-dimensional			3-dimensional			2-dimensional			2-dimensional		
		SES-Skala			Pseudo-Item Skalen			d-&e-Skala, m-Skala			d-&m-Skala, e-Skala		
		VG 1	VG 2	VG 3	VG 1	VG 2	VG 3	VG 1	VG 2	VG 3	VG 1	VG 2	VG 3
d-, e-, m-Items	<i>SES</i>	,692	,650	,575									
	<i>d-Dimension</i>				,622	,654	,589						
	<i>e-Dimension</i>				,763	,775	,740				,752	,740	,718
	<i>m-Dimension</i>				,629	,617	,538	,638	,617	,515			
	<i>d&e-Dimension</i>							,783	,825	,739			
	<i>d&m-Dimension</i>										,664	,502	,527

4.1 Geschätzte Mittelwerte und Standardabweichungen der Faktoren der Skala Soziale Erwünschtheit (GSP)

Gesamtstichprobe													
SES- Skala	Kriterium	1-dimensional			3-dimensional			2-dimensional			2-dimensional		
		SES-Skala			Pseudo-Item Skalen			d-&e-Skala, m-Skala			d-&m-Skala, e-Skala		
		1	2	3	1	2	3	1	2	3	1	2	3
d-, e-, m-Items	<i>SES</i>	0,20	0,02	-0,12									
		0,69	0,57	0,48									
	<i>d-Dimension</i>				3,52	0,26	-0,01						
					1,29	1,64	1,06						
	<i>e-Dimension</i>				0,25	-0,11	-0,03				0,11	-0,12	-0,15
					1,11	1,17	-0,02				1,15	1,10	1,04
MW Std. Abw.	<i>m-Dimension</i>				-1,15	0,21	-0,02	-1,23	-0,05	-0,08			
					1,06	1,04	0,88	1,08	1,08	0,87			
	<i>d&e-Dimension</i>							1,39	0,07	-0,08			
								0,96	1,08	0,79			
	<i>d&m-Dimension</i>										0,34	0,08	-0,17
											0,78	0,57	0,58

4.2 Marginale IRT-basierte Reliabilitäten der Faktoren der übrigen 19 Skalen auf Basis der einzelnen VG (I-III)

Einzelne Versuchsgruppen (VG1, 2 und 3), Modellvergleiche I-III

Arbeitsverhalten (AV)	Kriterium	A	B
		1-dimensional d-/e-/m-Dimension	3-dimensional Fragebogen- Skalen
VERSUCHSGRUPPE 1			
d-Items	d-Dimension	,726	
	Flexibilität		,743
	Gewissenhaftigkeit		,780
	Handlungsoptimierung		,687
e-Items	e-Dimension	,000	
	Flexibilität		,000
	Gewissenhaftigkeit		,000
	Handlungsoptimierung		,000
m-Items	m-Dimension	,780	
	Flexibilität		,695
	Gewissenhaftigkeit		,650
	Handlungsoptimierung		,615
VERSUCHSGRUPPE 2			
d-Items	d-Dimension	,826	
	Flexibilität		,774
	Gewissenhaftigkeit		,757
	Handlungsoptimierung		,734
e-Items*	e-Dimension	,882	
	Flexibilität		,802
	Gewissenhaftigkeit		,717
	Handlungsoptimierung		,826
m-Items	m-Dimension	,770	
	Flexibilität		,678
	Gewissenhaftigkeit		,677
	Handlungsoptimierung		,667
VERSUCHSGRUPPE 3			
d-Items	d-Dimension	,809	
	Flexibilität		,824
	Gewissenhaftigkeit		,788
	Handlungsoptimierung		,760
e-Items*	e-Dimension	,865	
	Flexibilität		,811
	Gewissenhaftigkeit		,748
	Handlungsoptimierung		,726
m-Items	m-Dimension	,728	
	Flexibilität		,687
	Gewissenhaftigkeit		,591
	Handlungsoptimierung		,627

Die Markierung * bezeichnet jene(s) Modell(e), für welche(s) Items ausgeschlossen wurden um eine eindeutige oder höhere Passung des ein oder anderen Modells zu erzielen. Die Reliabilitäten entstammen den finalen Modellen (mit Item-Ausschluss).

Berufl. Orientierung (BO)	Kriterium	A	B
		1-dimensional d-/e-/m-Dimension	3-dimensional Fragebogen- Skalen
VERSUCHSGRUPPE 1			
d-Items *	d-Dimension	,798	
	Führungsmotivation		,760
	Gestaltungsmotivation		,638
	Leistungsmotivation		,674
e-Items	e-Dimension	,871	
	Führungsmotivation		,737
	Gestaltungsmotivation		,623
	Leistungsmotivation		,746
m-Items	m-Dimension	,766	
	Führungsmotivation		,664
	Gestaltungsmotivation		,562
	Leistungsmotivation		,661
VERSUCHSGRUPPE 2			
d-Items *	d-Dimension	,853	
	Führungsmotivation		,815
	Gestaltungsmotivation		,743
	Leistungsmotivation		,710
e-Items	e-Dimension	,847	
	Führungsmotivation		,774
	Gestaltungsmotivation		,675
	Leistungsmotivation		,760
m-Items	m-Dimension	,696	
	Führungsmotivation		,560
	Gestaltungsmotivation		,466
	Leistungsmotivation		,626
VERSUCHSGRUPPE 3			
d-Items *	d-Dimension	,857	
	Führungsmotivation		,825
	Gestaltungsmotivation		,695
	Leistungsmotivation		,738
e-Items	e-Dimension	,844	
	Führungsmotivation		,752
	Gestaltungsmotivation		,695
	Leistungsmotivation		,723
m-Items	m-Dimension	,811	
	Führungsmotivation		,703
	Gestaltungsmotivation		,603
	Leistungsmotivation		,672

Die Markierung * bezeichnet jene(s) Modell(e), für welche(s) Items ausgeschlossen wurden um eine eindeutige oder höhere Passung des ein oder anderen Modells zu erzielen. Die Reliabilitäten entstammen den finalen Modellen (mit Item-Ausschluss).

Psych. Konstitution (PK)	Kriterium	A 1-dimensional d-/e-/m-Dimension	B 3-dimensional Fragebogen- Skalen
VERSUCHSGRUPPE 1			
d-Items *	d-Dimension	,855	
	Belastbarkeit		,670
	Emot. Stabilität		,772
	Selbstbewusstsein		,771
e-Items	e-Dimension	,850	
	Belastbarkeit		,762
	Emot. Stabilität		,729
	Selbstbewusstsein		,693
m-Items	m-Dimension	,839	
	Belastbarkeit		,751
	Emot. Stabilität		,756
	Selbstbewusstsein		,670
VERSUCHSGRUPPE 2			
d-Items	d-Dimension	,904	
	Belastbarkeit		,747
	Emot. Stabilität		,834
	Selbstbewusstsein		,822
e-Items*	e-Dimension	,876	
	Belastbarkeit		,817
	Emot. Stabilität		,755
	Selbstbewusstsein		,731
m-Items*	m-Dimension	,722	
	Belastbarkeit		,688
	Emot. Stabilität		,635
	Selbstbewusstsein		,609
VERSUCHSGRUPPE 3			
d-Items	d-Dimension	,897	
	Belastbarkeit		,793
	Emot. Stabilität		,796
	Selbstbewusstsein		,846
e-Items	e-Dimension	,864	
	Belastbarkeit		,775
	Emot. Stabilität		,741
	Selbstbewusstsein		,728
m-Items	m-Dimension	,802	
	Belastbarkeit		,748
	Emot. Stabilität		,696
	Selbstbewusstsein		,663

Die Markierung * bezeichnet jene(s) Modell(e), für welche(s) Items ausgeschlossen wurden um eine eindeutige oder höhere Passung des ein oder anderen Modells zu erzielen. Die Reliabilitäten entstammen den finalen Modellen (mit Item-Ausschluss).

Soz. Kompetenz(SK)	Kriterium	A 1-dimensional d-/e-/m-Dimension	B 5-dimensional Fragebogen- Skalen
VERSUCHSGRUPPE 1			
d-Items	d-Dimension	,861	
	Durchsetzungsstärke		,714
	Kontaktfähigkeit		,794
	Sensitivität		,570
	Soziabilität		,677
	Teamorientierung		,788
e-Items	e-Dimension	,896	
	Durchsetzungsstärke		,657
	Kontaktfähigkeit		,783
	Sensitivität		,784
	Soziabilität		,767
	Teamorientierung		,715
m-Items	m-Dimension	,856	
	Durchsetzungsstärke		,752
	Kontaktfähigkeit		,754
	Sensitivität		,700
	Soziabilität		,705
	Teamorientierung		,785
VERSUCHSGRUPPE 2			
d-Items	d-Dimension	,853	
	Durchsetzungsstärke		,779
	Kontaktfähigkeit		,810
	Sensitivität		,540
	Soziabilität		,675
	Teamorientierung		,770
e-Items	e-Dimension	,890	
	Durchsetzungsstärke		,658
	Kontaktfähigkeit		,795
	Sensitivität		,811
	Soziabilität		,724
	Teamorientierung		,723
m-Items	m-Dimension	,854	
	Durchsetzungsstärke		,745
	Kontaktfähigkeit		,787
	Sensitivität		,689
	Soziabilität		,682
	Teamorientierung		,793
VERSUCHSGRUPPE 3			
d-Items	d-Dimension	,884	

	Durchsetzungsstärke		,772
	Kontaktfähigkeit		,807
	Sensitivität		,612
	Soziabilität		,722
	Teamorientierung		,844
e-Items	e-Dimension	,907	
	Durchsetzungsstärke		,727
	Kontaktfähigkeit		,812
	Sensitivität		,825
	Soziabilität		,771
	Teamorientierung		,705
m-Items	m-Dimension	,828	
	Durchsetzungsstärke		,772
	Kontaktfähigkeit		,692
	Sensitivität		,723
	Soziabilität		,628
	Teamorientierung		,728
NEO-FFI (NEO)	Kriterium	A 1-dimensional d-/e-/m-Dimension	B 5-dimensional Fragebogen- Skalen
VERSUCHSGRUPPE 1			
d-Items	d-Dimension	,668	
	Extraversion		,686
	Gewissenhaftigkeit		,709
	Neurotizismus		,521
	Offenheit		,663
	Verträglichkeit		,746
e-Items	e-Dimension	,901	
	Extraversion		,759
	Gewissenhaftigkeit		,807
	Neurotizismus		,802
	Offenheit		,741
	Verträglichkeit		,766
m-Items	m-Dimension	,827	
	Extraversion		,632
	Gewissenhaftigkeit		,716
	Neurotizismus		,692
	Offenheit		,642
	Verträglichkeit		,630
VERSUCHSGRUPPE 2			
d-Items	d-Dimension	,605	
	Extraversion		,665
	Gewissenhaftigkeit		,642
	Neurotizismus		,525
	Offenheit		,714
	Verträglichkeit		,733

e-Items	<i>e-Dimension</i>	,909	
	<i>Extraversion</i>		,759
	<i>Gewissenhaftigkeit</i>		,855
	<i>Neurotizismus</i>		,840
	<i>Offenheit</i>		,737
	<i>Verträglichkeit</i>		,756
m-Items	<i>m-Dimension</i>	,801	
	<i>Extraversion</i>		,670
	<i>Gewissenhaftigkeit</i>		,676
	<i>Neurotizismus</i>		,765
	<i>Offenheit</i>		,667
	<i>Verträglichkeit</i>		,653
VERSUCHSGRUPPE 3			
d-Items *	<i>d-Dimension</i>	,891	
	<i>Extraversion</i>		,766
	<i>Gewissenhaftigkeit</i>		,834
	<i>Neurotizismus</i>		,729
	<i>Offenheit</i>		,785
	<i>Verträglichkeit</i>		,749
e-Items	<i>e-Dimension</i>	,890	
	<i>Extraversion</i>		,741
	<i>Gewissenhaftigkeit</i>		,832
	<i>Neurotizismus</i>		,761
	<i>Offenheit</i>		,783
	<i>Verträglichkeit</i>		,748
m-Items	<i>m-Dimension</i>	,748	
	<i>Extraversion</i>		,542
	<i>Gewissenhaftigkeit</i>		,684
	<i>Neurotizismus</i>		,673
	<i>Offenheit</i>		,629
	<i>Verträglichkeit</i>		,664

Die Markierung * bezeichnet jene(s) Modell(e), für welche(s) Items ausgeschlossen wurden um eine eindeutige oder höhere Passung des ein oder anderen Modells zu erzielen. Die Reliabilitäten entstammen den finalen Modellen (mit Item-Ausschluss).

VICTORIA GROTHE

Porzellangasse 33/13 | A-1090 Wien | +43 650 2000795 | victoria.albertz@gmail.com

PERSÖNLICHE ANGABEN

Name	Victoria Grothe, geb. Albertz
Geburtsdatum	24. Oktober 1985
Geburtsort	Essen, Deutschland
Familienstand	verheiratet, zwei Töchter geboren 07/2010 und 08/2013
Staatsangehörigkeit	Deutsch

SCHUL- UND STUDIENVERLAUF

seit 03/2007	Diplom-Studiengang Psychologie , Universität Wien
03/2007 – 06/2011	Bachelor in Betriebswirtschaftslehre , Universität Wien; Abschlussnote: 2,4
10/2005 – 10/2006	Studium der Humanmedizin, Medizinische Universität Wien
10/2004 – 10/2005	Studium der Humanmedizin, Karlsuniversität Prag
03/2001 – 07/2003	Malvern College, Great Malvern, UK; International Baccalaureate
10/1995 – 03/2001	Erzbischöfliches Clara-Fey Gymnasium, Bonn
03/1993 – 10/1995	Grundschule Klein Flottbecker Weg, Hamburg
10/1991 – 03/1993	Grundschule Burgschule, Bonn

PRAKTISCHE ERFAHRUNG

02/2010 - 04/2010	Praktikum, Klinik Königshof, Fachklinik für Psychiatrie, Neurologie und Psychotherapie, Krefeld
12/2009-01/2010	Praktikum, Lackner & Kabas, Organisationspsychologie & Coaching, Wien
07/2008 – 09/2008	Praktikum, ISG Personalmanagement, Wien
08/2007 – 10/2007	Praktikum, Investa Immobilienentwicklungs-GmbH, München
01/2007 – 03/2007	Praktikum, Investa Immobilienentwicklungs-GmbH, München
03/2004 – 07/2004	Pflegepraktikum, Alfried-Krupp-Krankenhaus, Essen
02/2003	Schul-Praktikum, Augenklinik Dardenne, Bonn Bad Godesberg

SPRACHKENNTNISSE

Deutsch	Muttersprache
Englisch	Fließend
Spanisch	Grundkenntnisse
Französisch	Grundkenntnisse
Latein	

EDV – KENNNTNISSE

Microsoft Office Word	sehr gute Kenntnisse
Microsoft Office Excel	sehr gute Kenntnisse
Microsoft Office PowerPoint	sehr gute Kenntnisse
SPSS	gute Kenntnisse