Originalarbeit

Das Big Five Inventar 2

Validierung eines Persönlichkeitsinventars zur Erfassung von 5 Persönlichkeitsdomänen und 15 Facetten

Daniel Danner[1], Beatrice Rammstedt[2], Matthias Bluemke[2], Clemens Lechner[2], Sabrina Berres[3], Thomas Knopf[2], Christopher J. Soto[4] und Oliver P. John[5]

[1]Hochschule der Bundesagentur für Arbeit, Mannheim
[2]GESIS – Leibniz-Institut für Sozialwissenschaften, Mannheim
[3]Universität Mannheim
[4]Colby College, Waterville
[5]University of California, Berkeley

Zusammenfassung: Die deutsche Version des Big Five Inventars 2 (BFI-2) erfasst die 5 Persönlichkeitsdomänen Extraversion, Verträglichkeit, Gewissenhaftigkeit, Negative Emotionalität (Neurotizismus) und Offenheit sowie insgesamt 15 Persönlichkeitsfacetten mit 60 Items. Das Inventar wurde im Rahmen eines mehrstufigen Übersetzungsprozesses vom Englischen ins Deutsche übertragen. Anhand einer nach Alter, Geschlecht und Bildungsabschluss quotieren Bevölkerungsstichprobe (N = 1 224) wurden Reliabilität, Validität und Messinvarianz über Geschlechter und Altersgruppen untersucht. Die Ergebnisse der Studien deuten auf gute Reliabilitäten der Domänenwerte und ausreichende Reliabilitäten der Facettenwerte hin. Einzelitems und die Facettenwerte zeigen erwartungskonforme Ladungsmuster. Die Domänen- und Facettenwerte korrelieren erwartungsgemäß mit anderen Persönlichkeitsinventaren und sind mit Kriterien wie Bildungsabschluss, Einkommen, Gesundheit und Lebenszufriedenheit assoziiert. Strukturgleichungsmodelle belegen die residuale Invarianz über Geschlechter. Die durchschnittliche Bearbeitungszeit beträgt 8 Minuten. Normwerte für die deutsche Allgemeinbevölkerung werden berichtet.

Schlüsselwörter: Persönlichkeit, Big Five, Validierung

The German Big Five Inventory 2: Measuring Five Personality Domains and 15 Facets

Abstract: The German version of the 60-item Big Five Inventory 2 (BFI-2) captures the five personality domains Extraversion, Agreeableness, Conscientiousness, Negative Emotionality (Neuroticism), and Open-Mindedness (Openness) and15 facets of personality. The inventory was translated from English into German in multiple steps. On the basis of a population survey (N = 1,224) quoted for age, sex, and education, we examined the reliability, validity, and measurement invariance for gender and age groups. The results suggested good reliability at the domain level and sufficient reliability at the facet level. Overall, items and facets showed the expected loadings on the five personality domains. Also, domain and facet scores correlated with other personality inventories and were associated with external criteria such as educational attainment, income, health, and life satisfaction. Structural equation modeling confirmed residual measurement invariance across gender. Completing the BFI-2 takes 8 min on average. Statistics for the general population in Germany are reported for future reference.

Keywords: personality, Big Five, validation

Wie lässt sich Persönlichkeit sinnvoll messen? Die Persönlichkeitspsychologie der letzten Jahrzehnte liefert eine empirisch untermauerte und weithin akzeptierte Antwort auf diese Frage: Auf globaler Ebene lässt sich Persönlichkeit entlang der fünf Domänen *Extraversion, Verträglichkeit, Gewissenhaftigkeit, Negative Emotionalität (Neurotizismus)* und *Offenheit* beschreiben. Diese fünf Faktoren (Tupes & Christal, 1992) oder *Big Five* (Goldberg, 1992; John, Naumann & Soto, 2008; McCrae & Costa, 1987) erlauben eine ökonomische Beschreibung interindividueller Unterschiede im Denken, Erleben und Verhalten. Darüber hinaus können die Big Five wichtige Lebenserfolgsindikatoren wie Bildungserfolg, Einkommen, Gesundheit und Lebenszufriedenheit vorhersagen (z. B. Paunonen, 2003; Roberts, Kuncel, Shiner, Caspi & Goldberg, 2007). In den vergangenen Jahren verschob sich der Fokus differentialpsychologischer Forschung jedoch von globalen Persönlichkeitsdomänen hin zu hierarchischen Persönlichkeitsmodellen. Auf abstrakterer Ebene beschreibt beispielsweise DeYoung (2006) zwei breite Eigenschaften: Stability, die Tendenz eigene Impulse zu kontrollieren, und Plasticity, die Tendenz sich mit der Umwelt aus-

© 2019 Hogrefe Verlag Veröffentlicht unter der Hogrefe OpenMind-Lizenz (https://doi.org/10.1026/a000002)

${protocol}://econtent.hogrefe.com/doi/pdf/10.1026/0012-1924/a000218 - Wednesday, August 28, 2019 6:50:59 AM - IP Address:137.146.214.250

${protocol}://econtent.hogrefe.com/doi/pdf/10.1026/0012-1924/a000218 - Wednesday, August 28, 2019 6:50:59 AM - IP Address:137.146.214.250

einanderzusetzen. Diese beiden Eigenschaften lassen sich in Persönlichkeitsdomänen wie die Big Five unterteilen, die sich selbst wiederum in spezifische Facetten (Soto & John, 2017) untergliedern. Auf niedrigster Ebene können Persönlichkeitsunterschiede entlang von Nuancen (Mõttus, 2016) beschrieben werden.

Bisher existieren im deutschsprachigen Raum nur wenige Verfahren, die neben den globalen Domänen eine Erfassung der Facettenstruktur der Big Five ermöglichen. Das etablierteste – und im deutschsprachigen Raum lange Zeit einzige – Verfahren ist das *NEO-Persönlichkeitsinventar* (*NEO-PI-R*; Ostendorf & Angleitner, 2004). Mit insgesamt 240 Items bildet es sechs Facetten pro Domäne ab. Das *NEO-PI-R* ermöglicht eine umfassende Beschreibung von Persönlichkeitsfacetten, ist aber zeitaufwändig und wird ausschließlich kommerziell vertrieben, sodass neben Zeit auch Lizenzkosten anfallen. Das alternative *Big Five Inventar* (*BFI*; John, Donahue & Kentle, 1991) wurde ursprünglich zur Erfassung der Big-Five-Domänen konzipiert. Aufgrund des zunehmenden Bedarfs nach einer Abbildung der Big-Five-Facettenstruktur wurde auch für die 45 Items des *BFI* eine Facettenstruktur entwickelt (Rammstedt & Danner, 2016; Soto & John, 2009). Diese bildet aber lediglich zwei Facetten pro Domäne ab. Die meisten bisherigen Verfahren bestehen ferner größtenteils aus positiv gepolten Items (eine Ausnahme bildet das *NEO-Fünf-Faktoren-Inventar* von Borkenau & Ostendorf, 2008), wodurch die Item- und Skalenwerte anfällig für akquieszenten Antwortstil sind. Akquieszenz, die Tendenz Aussagen unabhängig von ihrem Inhalt zuzustimmen, ist vor allem in heterogenen Bevölkerungsstichproben eine Störgröße und kann Mittelwerte, Faktorladungen sowie Korrelationen verzerren (z. B. Danner, Aichholzer & Rammstedt, 2015).

Das Big Five Inventar 2 (BFI-2)

Um dem Bedarf nach einem Instrument zu begegnen, das eine hierarchische Persönlichkeitsstruktur abbildet, entwickelten Soto und John (2017) das *BFI-2*. Diese Weiterentwicklung des *BFI* verfolgte vier Ziele: Erstens sollten globale Big-Five-Domänen und spezifische Facetten hierarchisch abgebildet werden. Die Auswahl der insgesamt 15 Facetten erfolgte basierend auf einer Synthese der in der Literatur berichteten Facettenstrukturen (u. a. Goldberg, 1992). Zweitens sollten die Items eindeutig formuliert und leicht verständlich sein, indem Synonyme gekoppelt (z. B. „Ich bin einfühlsam, warmherzig“) oder Begriffe erklärt werden (z. B. „Ich bin systematisch, halte meine Sachen in Ordnung“). Drittens sollte das Instrument sowohl ökonomisch als auch ausreichend reliabel und valide sein. Viertens sollten sämtliche Facetten- und Domänenskalen des Instruments ebenso viele positiv (z. B. „Ich gehe aus mir heraus, bin gesellig“) wie negativ gepolte Items (z. B. „Ich bin eher ruhig.“) enthalten. Dadurch sind die Skalenwerte von Akquieszenz bereinigt, und auch die Itemwerte können durch Ipsativierung von Akquieszenz bereinigt werden (z. B. Rammstedt & Danner, 2016). Soto und John (2017) schlugen des Weiteren vor, die Domäne Neurotizismus in Negative Emotionalität umzubenennen, da der Begriff Neurotizismus im Alltag mit psychischen Störungen konnotiert ist. Zudem bietet Negative Emotionalität eine akkuratere Beschreibung des Konstrukts, das durch das Erleben von Angst und Niedergeschlagenheit gekennzeichnet ist.

Soto und John (2017) untersuchten die psychometrischen Eigenschaften des englischsprachigen *BFI-2* in einer bevölkerungsheterogenen (N = 1 000) und einer studentischen Stichrobe (N = 470). Insgesamt belegten ihre Ergebnisse die faktorielle Validität und die Reliabilität der Domänen- und der Facettenwerte. Um den *BFI-2* mit ähnlich hoher Messqualität im deutschsprachigen Raum zur Verfügung zu stellen, wurden die originalen Formulierungen mittels mehrstufigem *TRAPD-Ansatz* (*Translation, Review, Adjudication, Pretesting, Documentation*; Harkness, 2007) adaptiert. Items wurden zunächst von zwei unabhängigen Übersetzern ins Deutsche übertragen. Die Übersetzungen wurden von Persönlichkeits- und Umfrageforschern geprüft und mit den Übersetzern besprochen.

In einer bevölkerungsheterogenen Stichprobe von N = 770 Personen wurden verschiedene Versionen der Items denselben Probandinnen und Probanden dargeboten. Die Items wurden dann aufgrund ihrer Trennschärfe (auf Domänen- und auf Facettenebene), ihrer inhaltlichen Heterogenität und ihrer Konvergenz mit dem *BFI* ausgewählt (siehe Danner et al., 2016).

Ziel der Studie

Die vorliegende Studie hat zum Ziel, die deutsche Adaptation des *BFI-2* zu validieren. Erstmals werden die psychometrischen Eigenschaften der finalen Version des deutschen *BFI-2* getestet, die konvergente Konstruktvalidität überprüft und der Mehrwert der *BFI-2*-Facetten gegenüber den Domänen demonstriert. Zudem werden Messinvarianztests erstmals für Geschlechter und Altersgruppen berichtet sowie Normwerte für die deutsche Allgemeinbevölkerung zur Verfügung gestellt.

© 2019 Hogrefe Verlag Veröffentlicht unter der Hogrefe OpenMind-Lizenz (https://doi.org/10.1026/a000002)

${protocol}://econtent.hogrefe.com/doi/pdf/10.1026/0012-1924/a000218 - Wednesday, August 28, 2019 6:50:59 AM - IP Address:137.146.214.250

Methode

Stichprobe

Insgesamt wurden Daten von N_1 = 1 224 erwachsenen Personen im Rahmen einer Onlinestudie erhoben. Die Stichprobe war hinsichtlich Alter (dreifach abgestuft), Geschlecht und Bildungsabschluss (dreifach abgestuft) nach dem Zensus 2011 (Statistisches Bundesamt, 2014) quotiert. Die Teilnehmenden wurden aus einem Online-Access-Panel eines kommerziellen Anbieters (Respondi AG) rekrutiert. Zu Beginn der Umfrage wurden alle Teilnehmenden nach ihrem Geschlecht, Alter und Bildungsabschluss befragt. Sobald die Quotierung für eine Zelle des Quotenplans erfüllt war, wurden keine weiteren Teilnehmenden für diese Zelle mehr zur Umfrage zugelassen. Die Verteilung von Geschlecht, Alter und Bildungsabschluss entsprach somit approximativ der des Zensus 2011. Für ihre Teilnahme wurden die Probandinnen und Probanden von Respondi finanziell entlohnt.

Das Alter der Teilnehmenden variierte zwischen 18 und 65 Jahren (M = 43.34, SD = 13.97), 50 % der Teilnehmenden waren weiblich, 36 % hatten einen Hauptschulabschluss, 33 % einen Realschulabschluss, 16 % die (Fach-) Hochschulreife, 15 % einen Hochschulabschluss. Ausgeschlossen wurden Teilnehmende, die eine Kontrollfrage („Bitte klicken Sie hier zur Überprüfung der Funktionsweise unserer Umfrage ‚stimme eher nicht zu' an.") falsch beantworteten sowie jene, die durchschnittlich weniger als 3 s oder mehr als 30 s pro Item benötigten. Ein Teil der Stichprobe (N_2 = 383) wurde nach 6 Wochen erneut befragt.

Instrumente

Alle Teilnehmenden bearbeiteten das *BFI-2*. Unterschiedliche Teilstichproben bearbeiteten zusätzlich jeweils eines der alternativen Big Five Inventare.

BFI-2. Die 60 Items des *BFI-2* erfassen fünf Big-Five-Domänen und 15 Facetten (siehe Tabelle 1). Jede Facette wird mit vier Items erfasst. Die fünfstufige Antwortskala reicht von *„stimme überhaupt nicht zu"* bis *„stimme voll und ganz zu"*. Die Skalenwerte werden durch einfache Mittelwertbildung berechnet. Fragebogen, Instruktion und Auswertungshinweise sind im Elektronischen Supplement 3 dargestellt.

NEO-PI-R. Ein Teil der Stichprobe (n = 192) bearbeitete zusätzlich das *NEO-Persönlichkeitsinventar* (*NEO-PI-R*; Ostendorf & Angleitner, 2004), welches mit insgesamt 240 Items je sechs Persönlichkeitsfacetten pro Big-Five-Domäne erfasst (Abbildung 1). Die fünfstufige Antwortskala reicht von *„starke Ablehnung"* bis *„starke Zustimmung"*. Die internen Konsistenzen der Domänenwerte lagen in einem guten Bereich (Cronbachs α > .86), die internen Konsistenzen der Facettenwerte variierten zwischen α = .26 und α = .89 (siehe Tabelle E1 und E2 im Elektronischen Supplement 1).

NEO-FFI. Als Kurzform des *NEO-PI-R* enthält das *NEO-Fünf-Faktoren-Inventar* (*NEO-FFI*; Borkenau & Ostendorf, 2008) 60 der 240 Items. Das *NEO-FFI* erlaubt keine Erfassung von Facetten. Die Domänenwerte des *NEO-FFI* wurden auf Basis der Antworten auf *NEO-PI-R*-Items berechnet (n = 192). Die internen Konsistenzen der Domänenwerte lagen in einem guten Bereich (α > .71; siehe Tabelle E1 im Elektronischen Supplement 1).

BFI. Ein anderer Teil der Stichprobe (n = 185) bearbeitete anstelle des *NEO-PI-R* zusätzlich das *BFI* (Rammstedt, 1997). Das 45-Item-Inventar bildet die Big-Five-Domänen mithilfe einer fünfstufigen Antwortskala von *„stimme überhaupt nicht zu"* bis *„stimme voll und ganz zu"* ab. Die internen Konsistenzen der Domänen lagen in einem guten Bereich (α > .74; vgl. Tabelle E1 im ESM 1).

BFI-10. Das *BFI-10* ist die Kurzform des *BFI* mit 10 der 45 *BFI*-Items (Rammstedt & John, 2007). Die Domänenwerte des *BFI-10* wurden auf Basis der Antworten auf *BFI*-Items berechnet. Die internen Konsistenzen variierten zwischen α = .37 und α = .64 (vgl. Tabelle E1 im ESM 1).

BFI-S. Ein wiederum anderer Teil der Stichprobe (n = 353) bearbeitete zusätzlich das für das S*ozioökonomische Panel (SOEP)* entwickelte *BFI-S* (Schupp & Gerlitz, 2014). Es umfasst 15 Items, die in der vorliegenden Studie auf einer fünfstufigen Antwortskala von *„stimme überhaupt nicht zu"* bis *„stimme voll und ganz zu"* beantwortet wurden. Die internen Konsistenzen der *BFI-S*-Domänenwerte waren α > .58 (Tabelle E1 im ESM 1).

Lebenserfolgsindikatoren. Um die Kriteriumsvalidität des *BFI-2* zu evaluieren, wurden die Zusammenhänge mit vier zentralen Lebenserfolgsindikatoren erfasst: Höchster erreichter Bildungsabschluss (sechsstufig von *„ohne Abschluss"* bis *„Hochschulabschluss"*), Einkommen (17 ordinale Stufen), selbsteingeschätzter Gesundheitszustand (fünfstufig von *„schlecht"* bis *„sehr gut"*) und Lebenszufriedenheit (11-stufig von *„ganz und gar unzufrieden"* bis *„ganz und gar zufrieden"*). Items und Antwortskalen sind im Elektronischen Supplement 1 in Tabelle E3 dargestellt. Im Rahmen der Befragung bearbeiteten die Teilnehmenden weitere Items zu politischen Einstellungen, existentiellen Fragen und kognitiven Fähigkeiten, die nicht Gegenstand dieser Studie sind.

© 2019 Hogrefe Verlag Veröffentlicht unter der Hogrefe OpenMind-Lizenz (https://doi.org/10.1026/a000002)

${protocol}://econtent.hogrefe.com/doi/pdf/10.1026/0012-1924/a000218 - Wednesday, August 28, 2019 6:50:59 AM - IP Address:137.146.214.250

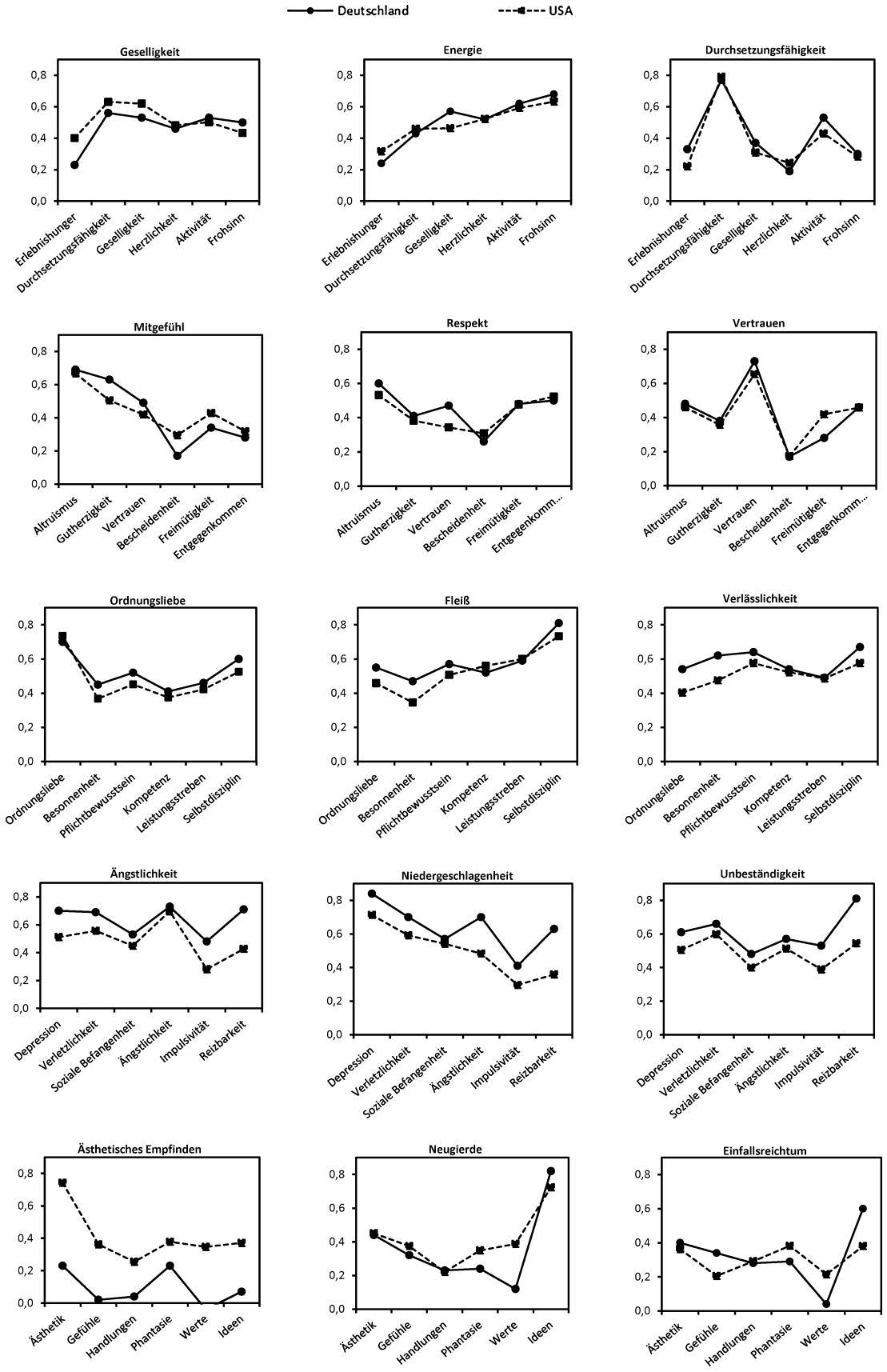

Abbildung 1. Vergleich zwischen deutscher und US-Stichprobe in den Korrelationen der BFI-2-Facetten mit den NEO-PI-R-Facetten.

© 2019 Hogrefe Verlag Veröffentlicht unter der Hogrefe OpenMind-Lizenz (https://doi.org/10.1026/a000002)

${protocol}://econtent.hogrefe.com/doi/pdf/10.1026/0012-1924/a000218 - Wednesday, August 28, 2019 6:50:59 AM - IP Address:137.146.214.250

Ergebnisse

Untersucht wurden die Faktorstruktur der Item- und der Facettenwerte, die Reliabilität der Skalenwerte (manifeste Domänen- und Facettenwerte) sowie die Konstrukt- und Kriteriumsvalidität der Skalenwerte. Abschließend berichten wir Normwerte.

Faktorielle Struktur

Die faktorielle Struktur der *BFI-2*-Items wurde mit einer Hauptkomponentenanalyse untersucht. Dies erlaubt die Kontrolle von Akquieszenz und den Vergleich mit den Ergebnissen von Soto und John (2017). Im Sinne des Big-Five-Modells mit fünf orthogonalen Domänen wurden aufgrund theoretischer Überlegungen fünf Komponenten extrahiert und orthogonal rotiert (Varimax). Alle Items wurden vor der Analyse um Akquieszenz bereinigt, indem ein Akquieszenzindikator (Mittelwert über die 30 positiv und die 30 negativ kodierten Items) von jedem Item subtrahiert wurde („Ipsativierung"; Soto & John, 2017; Rammstedt & Danner, 2016). Alle 60 Items luden am höchsten auf der inhaltlich zugeordneten Komponente. Die durchschnittliche Ladung der Items einer Domäne auf der dieser zugerechneten Hauptkomponente betrug .59. Der Durchschnitt (der Absolutbeträge) der standardisierten Ladungen der Items auf andere Komponenten betrug .12. Dieses Ladungsmuster zeigt, dass sich die theoretische Zuordnung der Items zu den fünf Domänen Extraversion, Verträglichkeit, Gewissenhaftigkeit, Negative Emotionalität und Offenheit auch empirisch wiederfindet. Die vollständige Ladungsmatrix ist im Elektronischen Supplement 1 in Tabelle E4 dargestellt. Des Weiteren wurde überprüft, ob das Ladungsmuster der deutschen Adaptation mit der englischsprachigen Version des *BFI-2* übereinstimmt. Dazu wurde die Kongruenz des Ladungsmusters der deutschen Stichprobe mit dem Ladungsmuster einer ähnlichen US-Stichprobe (Soto & John, 2017, Tabelle 6, Internet-Stichprobe, $N = 1\,000$) berechnet. Die Kongruenz der Ladungen für alle fünf Komponenten ergab $\phi > .94$. Die Ladungsmuster stimmen demnach sehr gut überein.

Die faktorielle Struktur der *BFI-2*-Facettenwerte wurde ebenfalls mit einer Hauptkomponentenanalyse (fünf varimax-rotierte Komponenten) untersucht. Alle Facettenwerte luden am höchsten auf die Komponente, der sie inhaltlich zugeordnet werden können. Die durchschnittliche standardisierte Ladung der Facettenwerte einer Domäne auf eine gemeinsame Komponente betrug .79. Die durchschnittliche standardisierte Ladung der Facettenwerte einer Domäne auf andere Komponenten betrug .13. Dieses Ergebnismuster spiegelt die inhaltliche Zuordnung der Facetten zu fünf Domänen auch statistisch wider. Die vollständige Ladungsmatrix ist in Tabelle 1 dargestellt. Der Vergleich mit der US-Stichprobe (Soto & John, 2017, Tabelle 6, Internet-Stichprobe, $N = 1\,000$) belegt die hohe Ähnlichkeit des Ladungsmusters. Die Kongruenz der Ladungen für alle fünf Komponenten betrug $\phi > .88$.

Zusätzlich wurde die faktorielle Struktur der Items und der Facettenwerte mit explorativen Strukturgleichungsmodellen (ESEM mit fünf Faktoren und orthogonaler Target-Rotation; Asparouhov & Muthén, 2009) untersucht, was zu äquivalenten Ladungsmustern führte (siehe Tabelle 1 und Tabelle E4 im Elektronischen Supplement 1).

Reliabilität

Die Reliabilität der Domänen- und Facettenwerte wurde über Cronbachs α, McDonalds Omega und die Retest-Korrelation (6-Wochen-Intervall) geschätzt (Tabelle 2). Cronbachs α wird als geläufigster Reliabilitätsschätzer berichtet; McDonalds Omega, weil es ein präziserer Schätzer für Skalen mit heterogenen Faktorladungen ist und die Retest-Korrelation, weil sie keine Annahmen über die Homogenität der Skalen macht. Die Reliabilitätsschätzungen der Domänenwerte lagen mit .80 bis .90 in einem guten bis sehr guten Bereich: McDonald Omegas lagen zwischen .81 und .87 (durchschnittlich .84), Cronbachs α zwischen .80 und .89 (durchschnittlich .85) und die Retestkorrelationen zwischen .82 und .90 (durchschnittlich .86[1]). Die Reliabilitätsschätzer der Facettenwerte lagen zwischen .56 bis .87. McDonalds Omegas erreichten Werte zwischen .66 und .87 (durchschnittlich .77) und Cronbachs α zwischen .56 und .84 (durchschnittlich .73) sowie die Retestkorrelationen zwischen .58 und .85 (durchschnittlich .79).

Diese Reliabilitätsschätzungen sind ähnlich wie die der US-Stichprobe (Soto & John, 2017, Tabelle 2 und Tabelle 3, Internet-Stichprobe, $N = 1\,000$). Auf Domänenebene berichteten Soto und John (2017) Cronbachs α zwischen .83 und .90 (durchschnittlich .87) und Retestkorrelationen (8 Wochen) zwischen .76 und .84 (durchschnittlich .80). Auf Facettenebene werden Cronbachs α zwischen .66 und .85 (durchschnittlich .76) und Retestkorrelationen zwischen .66 und .83 (durchschnittlich .73) berichtet.

[1] Mittelwert nach Fisher-Z-Transformation der Korrelationen und Rücktransformation der gemittelten Korrelation.

© 2019 Hogrefe Verlag Veröffentlicht unter der
Hogrefe OpenMind-Lizenz (https://doi.org/10.1026/a000002)

${protocol}://econtent.hogrefe.com/doi/pdf/10.1026/0012-1924/a000218 - Wednesday, August 28, 2019 6:50:59 AM - IP Address:137.146.214.250

Tabelle 1. Ladungsmuster der BFI-2-Facettenwerte (PCA und ESEM, orthogonale Rotation)

		PCA					ESEM				
Domäne	Facette	I	II	III	IV	V	I	II	III	IV	V
Extraversion	Geselligkeit	**.87**	.13	.04	-.03	.03	**.75**	.16	-.03	.03	-.08
	Durchsetzungsfähigkeit	**.74**	-.18	.08	-.21	.28	**.73**	-.09	.01	-.08	.17
	Energie	**.74**	.24	.15	-.16	.23	**.72**	.24	.09	-.10	.07
Verträglichkeit	Mitgefühl	.19	**.81**	.19	.09	.18	.16	**.77**	.20	-.04	.09
	Höflichkeit	-.06	**.68**	.40	-.27	.08	.01	**.59**	.39	-.37	.04
	Vertrauen	.05	**.77**	-.10	-.28	-.02	.05	**.53**	-.01	-.33	-.03
Gewissenhaftigkeit	Ordnungsliebe	.02	.07	**.86**	.04	-.01	.34	-.01	**.72**	-.21	.05
	Fleiß	.22	.03	**.80**	-.21	.10	.10	.04	**.74**	-.04	-.05
	Verlässlichkeit	.06	.20	**.76**	-.29	.08	.19	.17	**.67**	-.32	.02
Negative Emotionalität	Ängstlichkeit	-.18	-.05	-.07	**.89**	-.02	-.32	.04	-.01	**.81**	.02
	Niedergeschlagenheit	-.42	-.09	-.20	**.73**	-.07	-.53	-.04	-.13	**.62**	.02
	Unbeständigkeit	.08	-.28	-.22	**.82**	-.12	-.08	-.18	-.19	**.82**	-.10
Offenheit	Neugierde	-.02	.20	.00	.05	**.80**	.16	.21	.04	.02	**.55**
	Ästhetisches Empfinden	.19	.02	.04	-.15	**.83**	.46	.07	.09	-.03	**.49**
	Einfallsreichtum	.36	-.03	.12	-.08	**.69**	.35	.09	.04	-.09	**.76**

Anmerkungen: Der Eigenwerteverlauf der PCA war 4.72, 2.12, 1.56, 1.47, 1.14, 0.63, 0.54, 0.48, 0.42, 0.40, 0.36, 0.34, 0.32, 0.24, 0.23. Das Model erklärte 73 % der Gesamtvarianz. Die Modellpassung des ESEM war RMSEA = .058, CFI = .979. Die höchste Ladung pro Facettenwert ist fett hervorgehoben, *N* = 1 224.

Tabelle 2. Reliabilitätsschätzer für BFI-2-Domänen- und Facettenwerte

	Interne Konsistenz		
	ω (*N* = 1 224)	α (*N* = 1 224)	Retest Korrelation (*N* = 383)
Extraversion	.83	.85	.90
Geselligkeit	.79	.76	.83
Durchsetzungsfähigkeit	.78	.75	.83
Energie	.71	.67	.80
Verträglichkeit	.81	.80	.82
Mitgefühl	.81	.71	.75
Höflichkeit	.69	.65	.76
Vertrauen	.66	.56	.74
Gewissenhaftigkeit	.84	.86	.87
Ordnungsliebe	.87	.84	.85
Fleiß	.76	.72	.83
Verlässlichkeit	.66	.58	.74
Negative Emotionalität	.85	.89	.87
Ängstlichkeit	.73	.69	.78
Niedergeschlagenheit	.87	.82	.84
Unbeständigkeit	.84	.80	.81
Offenheit	.87	.85	.82
Neugierde	.71	.65	.74
Ästhetisches Empfinden	.83	.82	.58
Einfallsreichtum	.82	.81	.78
Mittelwert Domänen	.84	.85	.86
Mittelwert Facetten	.78	.73	.78

Anmerkungen: Zur Schätzung von ω wurden für jede Domäne bifaktorielle Messmodelle spezifiziert, alle RMSEA $\leq$ .081, alle CFI $\geq$ .933, alle SRMR $\leq$ .048. Das Retestinvervall betrug 6 Wochen.

© 2019 Hogrefe Verlag Veröffentlicht unter der Hogrefe OpenMind-Lizenz (https://doi.org/10.1026/a000002)

Tabelle 3. Korrelationen der BFI-2-Domänen mit anderen Big-Five-Inventaren

	NEO-PI-R (*N* = 192)	NEO-FFI (*N* = 192)	BFI (*N* = 185)	BFI-S (*N* = 353)	BFI-10 (*N* = 185)
Extraversion	.81	.77	.86	.78	.71
Verträglichkeit	.78	.72	.78	.70	.67
Gewissenhaftigkeit	.81	.82	.85	.73	.77
Negative Emotionalität	.87	.84	.88	.81	.78
Offenheit	.65	.64	.87	.73	.74
Mittelwert	.79	.77	.85	.75	.74

Anmerkung: Alle Korrelationen $p < .001$.

Tabelle 4. Korrelationen zwischen den manifesten BFI-2-Domänenwerten

	Verträglichkeit	Gewissenhaftigkeit	Negative Emotionalität	Offenheit
Extraversion	.21	.25	-.38	.41
Verträglichkeit		.33	-.38	.22
Gewissenhaftigkeit			-.37	.17
Negative Emotionalität				-.21

Anmerkungen: N = 1 224. Alle Korrelationen $p < .001$.

Konstruktvalidität

Die Korrelationen der *BFI-2*-Domänenwerte mit den entsprechenden Domänenwerten des *NEO-PI-R*, *NEO-FFI*, *BFI*, *BFI-S* und *BFI-10* (Tabelle 3) lagen zwischen .64 und .88 (durchschnittlich .78) und wiesen damit insgesamt auf eine gute konvergente Konstruktvalidität hin. Die höchsten Korrelationen zeigten sich mit dem *BFI* (durchschnittlich .85).

Soto und John (2017, Tabelle 9) berichteten ebenfalls Korrelationen zwischen den Domänenwerten des *BFI-2* mit den Domänenwerten des *NEO-PI-R*, des *NEO-FFI* und des *BFI*, die von uns in Tabelle E5 im Elektronischen Supplement 1 zu Vergleichszwecken wiedergegeben werden. Die deutsche Stichprobe erzielte im Vergleich zur US-Stichprobe etwas höhere Korrelationen mit dem *NEO-PI-R* sowie *NEO-FFI* und geringere Korrelationen mit dem *BFI* (v. a. der Domäne Verträglichkeit).

Die divergente Konstruktvalidität der Domänenwerte wurde anhand der Interkorrelationen der *BFI-2*-Domänenwerte bewertet. Die mittlere Korrelation der *BFI-2* Domänenwerte untereinander erreichte .30 (absolute Werte, Fisher-Z-transformiert, gemittelt, rücktransformiert). Die höchste Korrelation ergab sich zwischen Extraversion und Offenheit (.41). Richtung und Höhe der Interkorrelationen sind ähnlich wie die Interkorrelationen des *NEO-PI-R*, *NEO-FFI*, des *BFI*, *BFI-S* und *BFI-10* (siehe Tabelle 4 und Tabelle E6 im Elektronischen Supplement 1) und ähneln der in der Literatur berichteten Interkorrelationen zwischen manifesten Domänenwerten (Goldberg, 1992; Soto & John, 2017).

Die Konstruktvalidität der Facettenwerte wurde anhand der Korrelationen mit den Facettenwerten des *NEO-PI-R* bewertet, welches mit sechs Facetten pro Domäne (trotz teilweise identischer Bezeichnung) keine konzeptionell identischen Facettenwerte liefert. Daher ist keine perfekte Übereinstimmung zu erwarten. Stattdessen wurden die Korrelationen der deutschen Stichprobe mit den Korrelationen der US-Stichprobe von Soto und John (2017) verglichen. Abbildung 1 zeigt den grafischen Vergleich der Korrelationen der deutschen Stichprobe mit den von Soto und John (2017) berichteten Korrelationen zwischen den *BFI-2* und den *NEO-PI-R*-Facetten. Insgesamt zeigt sich ein ähnliches Korrelationsmuster. Zusätzlich wurde für jede *BFI-2*-Facette die Vektorkorrelation (Korrelation zwischen deutschen und US-Korrelationen über die sechs *NEO-PI-R*-Facetten) bestimmt. Die Vektorkorrelationen variierten zwischen .62 (Verlässlichkeit) und .97 (Ordnungsliebe). Durchschnittlich betrug die Vektorkorrelation .85, was eine hohe Ähnlichkeit der Facettenkorrelationen in der deutschen und in der US-Stichprobe belegt. Die Korrelationen zwischen den *BFI-2* und den *NEO-PI-R*-Facetten sind im Elektronischen Supplement 1 in Tabelle E7 dargestellt.

Kriteriumsvalidität und Nützlichkeit

Die Kriteriumsvalidität und die Nützlichkeit der Domänen- und Facettenwerte wurde anhand der Zusammenhänge mit dem höchsten erreichten Bildungsabschluss, dem Einkommen, der selbsteingeschätzten Gesundheit

${protocol}://econtent.hogrefe.com/doi/pdf/10.1026/0012-1924/a000218 - Wednesday, August 28, 2019 6:50:59 AM - IP Address:137.146.214.250

© 2019 Hogrefe Verlag Veröffentlicht unter der Hogrefe OpenMind-Lizenz (https://doi.org/10.1026/a000002)

Tabelle 5. Standardisierte Regressionskoeffizienten für Vorhersage von Bildung, Einkommen, Gesundheit und Lebenszufriedenheit durch die BFI-2-Domänen- bzw. Facettenwerte

	Bildung (N = 1 224)			Einkommen (N = 1 082)			Gesundheit (N = 1 220)			Zufriedenheit (N = 1 223)		
Modell	I	II	III	I	II	III	I	II	III	I	II	III
Extraversion	.03			.06			.15***			.14***		
Geselligkeit		-.14***			-.03			-.09**			-.07*	
Durchsetzungsfähigkeit		.22***	.20***		.10*	.08*		-.05			-.01	
Energie		.02			-.02			.25***	.20***		.13***	.11***
Verträglichkeit	-.02			-.13***			-.13***			.00		
Mitgefühl		-.02			-.06			-.08**	-.13***		.00	
Respekt		-.09*			-.16***	-.14***		-.05			.01	
Vertrauen		.12***	.06*		.09*			-.03			.02	.01
Gewissenhaftigkeit	-.10***			.08*			.02			.04		
Ordnungsliebe		.00			.07			.02	.01		.05	.03
Fleiß		-.14***	-.12***		.09*	.13***		.00			-.02	
Verlässlichkeit		.09*			.00			-.01			.01	
Negative Emotionalität	-.12***			-.19***			-.42***			-.41***		
Ängstlichkeit		-.07	-.09**		-.13**	-.18***		-.05			-.10***	
Niedergeschlagenheit		-.05			-.10*			-.48***	-.45***		-.50***	-.50***
Unbeständigkeit		.03			.02			.03			.10*	
Offenheit	.27***			.04			-.10***			-.05		
Neugierde		.18***			.00			.00			-.02	
Ästhetische Empfinden		.18***	.23***		.05	.05		.00			.04	
Einfallsreichtum		-.07			.00			-.14***	-.15***		-.08*	-.09**
Adjustiertes R^2	.09	.15	.12	.06	.08	.07	.20	.29	.29	.23	.30	.30

Anmerkungen: * $p < .05$, ** $p < .01$, *** $p < .001$.

und der Lebenszufriedenheit mit linearen Regressionsmodellen analysiert. In einem ersten Schritt wurde jedes Kriterium durch die fünf Domänenwerte vorhergesagt (Modell I). In einem zweiten Schritt wurde jedes Kriterium durch die 15 Facettenwerte vorhergesagt (Modell II). Der Vergleich der (adjustierten) aufgeklärten Varianzen in Modell I und Modell II offenbarte den Mehrwert der Facettenstruktur gegenüber einer reinen Domänenstruktur. Um zu überprüfen, ob der Mehrwert der Facettenstruktur durch die inhaltlich spezifischeren Facetten oder nur durch die höhere Anzahl der Prädiktoren zustande kommt, wurden in einem dritten Schritt die Kriterien durch nur jeweils die Facette vorhergesagt, die in Modell II den höchsten Zusammenhang mit dem Kriterium aufwies (Modell III). Die standardisierten Regressionskoeffizienten und die aufgeklärten Varianzanteile sind in Tabelle 5 dargestellt.

Facetten (Modell II) klärten allgemein mehr Kriteriumsvarianz auf als Domänen (Model I). Für Bildung stieg der aufgeklärte Varianzanteil um 6 Prozentpunkte, für Einkommen um 2 Prozentpunkte, für Gesundheit um 9 Prozentpunkte und für Lebenszufriedenheit um 7 Prozentpunkte. Der Anstieg aufgeklärter Varianz war laut Steigers Z-Werten in allen Fällen signifikant, $Z_H > 2.78$, $p < .005$. Dieser Anstieg aufgeklärter Varianz kann auch nicht alleine durch die Anzahl der Prädiktoren erklärt werden, da der Zuwachs an aufgeklärter Varianz im Wesentlichen bestehen blieb, wenn pro Domäne jeweils nur eine einzige Facette als Prädiktor herangezogen wurde (Modell III). Auch dieser Anstieg aufgeklärter Varianz war im Vergleich zu Modell I in allen Fällen signifikant, $Z_H > 1.96$, $p < .05$.

Einen indirekten Hinweis auf den Mehrwert der Facettenstruktur lieferten die häufig unterschiedlichen Kriterienzusammenhänge von Facetten derselben Domäne. Beispielsweise resultierte ein nicht signifikanter Zusammenhang zwischen Bildung und Extraversion auf Domänenebene, jedoch ein positiver Zusammenhang zwischen Bildung und der Extraversionsfacette Durchsetzungsfähigkeit und ein negativer Zusammenhang zwischen Bildung und der Extraversionsfacette Geselligkeit. Solche unterschiedlichen Zusammenhänge zwischen Facetten einer Domäne und Kriterien belegen, dass sich die Effekte der Facetten bei Analysen auf Domänenebene wechselseitig maskieren können.

${protocol}://econtent.hogrefe.com/doi/pdf/10.1026/0012-1924/a000218 - Wednesday, August 28, 2019 6:50:59 AM - IP Address:137.146.214.250

© 2019 Hogrefe Verlag Veröffentlicht unter der Hogrefe OpenMind-Lizenz (https://doi.org/10.1026/a000002)

Tabelle 6. Messinvarianz der BFI-2-Domänenwerte über Geschlechter und Altersgruppen

	Geschlecht			Alter		
Invarianzlevel	CFI	RMSEA	SRMR	CFI	RMSEA	SRMR
konfigural	.977	.060	.016	.972	.066	.020
metrisch	.972	.052	.032	.963	.054	.048
skalar	.963	.057	.034	.955	.058	.052
residual	.962	.056	.042	.947	.058	.065

Anmerkungen: Altersgruppen: Bis 30 Jahre, 31 – 40 Jahre, 41 – 50 Jahre, über 50 Jahre. *N* = 1 224.

${protocol}://econtent.hogrefe.com/doi/pdf/10.1026/0012-1924/a000218 - Wednesday, August 28, 2019 6:50:59 AM - IP Address:137.146.214.250

Messinvarianz über Geschlechter und Altersgruppen

Ob Skalenwerte des *BFI-2* zwischen Frauen und Männern und zwischen verschiedenen Altersgruppen verglichen werden können, prüften wir durch Testen der Messinvarianz des *BFI-2* über Geschlechter und Altersgruppen. *Konfigurale Invarianz* beschreibt eine äquivalente faktorielle Struktur der Konstrukte. *Metrische Invarianz* beschreibt eine äquivalente faktorielle Struktur der Konstrukte und eine vergleichbare Gewichtung der Indikatoren. *Skalare Invarianz* beschreibt eine äquivalente faktorielle Struktur der Konstrukte, eine äquivalente Gewichtung und Schwierigkeit der Indikatoren. *Residuale Invarianz* beschreibt zusätzlich gleiche Fehlervarianzen der Indikatoren. Nur im Falle residualer Invarianz sollten manifeste Skalenwerte zwischen Gruppen verglichen werden.

Die Messinvarianz der *BFI-2*-Domänenwerte wurde mit explorativen Strukturgleichungsmodellen (ESEM; Asparouhov & Muthén, 2009) für mehrere Gruppen untersucht, wobei die 15 Facettenwerte des *BFI-2* als manifeste Indikatoren für fünf latente Variablen spezifiziert wurden. Die Facettenwerte wurden als manifeste Indikatoren gewählt, da diese die kleinsten interpretierbaren Einheiten des *BFI-2* darstellen, nicht durch Akquieszenz verzerrt sind und mit drei Indikatoren pro Domäne eine ausreichende Identifizierung des Modells ermöglichen. Die Ladungen aller Facetten auf alle Domänen wurden frei geschätzt, die Korrelationen zwischen den Domänen auf null fixiert. Für das konfigural invariante Modell wurden alle zu schätzenden Parameter des Modells frei geschätzt. Für das metrisch invariante Modell wurden die Faktorladungen über alle Gruppen gleichgesetzt. Für das skalar invariante Modell wurden auch die Intercepts (Schwierigkeiten) über die Gruppen gleichgesetzt. Für das residual invariante Modell wurden zudem die Fehlervarianzen über die Gruppen gleichgesetzt. Die Unterschiede zwischen den Modellen wurde anhand der Kriterien von </litr=3>Chen (2007) bewertet, wonach bei einem Abfall im CFI < .010, ein Anstieg im RMSEA < .015 und SRMR < .030 vom konfiguralen zum metrischen Modell das metrische Modell angenommen werden soll. Bei einem Abfall im CFI < .010, einem Anstieg im RMSEA < .015 und SRMR < .010 vom metrischen zum skalaren Modell soll das skalare Modell angenommen werden und bei einem Abfall im CFI < .010, einem Anstieg im RMSEA < .015 und SRMR < .010 vom skalaren zum residualen Modell wurde das residuale Modell angenommen.

Die Veränderungen der Model-Fitindices (Tabelle 6) deuten auf residuale Invarianz für Geschlechter hin, sodass die manifesten Domänenwerte zwischen Frauen und Männern verglichen werden können. Für die Altersgruppen sind die Befunde heterogen: Während der RMSEA und CFI auf residuale Invarianz hindeuteten, legt der SRMR nur skalare Invarianz nahe. Eine Vergleichbarkeit der manifesten Skalenwerte über Altersgruppen ist daher nicht gewährleistet, sondern nur bei korrekter Spezifikation im Strukturgleichungsmodell mit unterschiedlichen Fehlervarianzen für verschiedene Altersgruppen sinnvoll.

Die Ergebnisse für die Facettenwerte sind im Elektronischen Supplement 1 in Tabelle E8 dargestellt. Die Ergebnisse deuten auf residuale Invarianz für Frauen und Männer hin, lassen aber keine eindeutige Schlussfolgerung für die Invarianz über Altersgruppen zu.

Normwerte

Im Elektronischen Supplement 1 in Tabelle E9 sind Normwerte (manifeste Mittelwerte) für die fünf Domänen und in Tabelle E10 sind Normwerte für die 15 Facetten dargestellt. Damit sowohl Vergleiche zwischen als auch innerhalb verschiedener Bevölkerungsgruppen durchgeführt werden können, werden Normwerte für die erwachsene Allgemeinbevölkerung und getrennt für Frauen und Männer und vier Altersgruppen berichtet. Mit den Tabellen E11 bis E30 im Elektronischen Supplement 2 können außerdem Rohwerte in Standard-T-Werte und Prozentränge übersetzt werden.

© 2019 Hogrefe Verlag Veröffentlicht unter der
Hogrefe OpenMind-Lizenz (https://doi.org/10.1026/a000002)

${protocol}://econtent.hogrefe.com/doi/pdf/10.1026/0012-1924/a000218 - Wednesday, August 28, 2019 6:50:59 AM - IP Address:137.146.214.250

Diskussion

Ziel des vorliegenden Beitrags war die Evaluierung der deutschen Version des *BFI-2*. Insgesamt weist das Inventar gute psychometrische Eigenschaften auf. Alle Items laden am höchsten auf die Domäne, der sie inhaltlich zugeordnet werden. Einige Items weisen substantielle Nebenladungen auf andere Domänen auf. Dies bedeutet, dass diese Items keine „reinen" Indikatoren einer einzigen Domäne sind, sondern dass sie teilweise auch andere Domänen abbilden. Beispielsweise lädt das Item „Ich bin voller Energie und Tatendrang" (Nr. 41), welches inhaltlich der Domäne Extraversion und der Facette Energie zugeordnet wird, am höchsten auf Extraversion, weist aber auch eine substantielle Nebenladung auf Negative Emotionalität auf. Auch die faktorielle Struktur der Facettenwerte spiegelt die erwartete Struktur wider. Alle Facetten laden am höchsten auf die Komponente, der sie inhaltlich zugeordnet werden, und Nebenladungen auf andere Komponenten sind inhaltlich plausibel. Beispielsweise bedeuten Nebenladungen der Facetten Ängstlichkeit und Niedergeschlagenheit auf Extraversion, dass Personen, die sich als ängstlich oder niedergeschlagen beschreiben, sich auch als weniger gesellig, durchsetzungsfähig oder energisch beschreiben.

Die Reliabilitätsschätzer der Domänenwerte liegen in einem sehr guten bis guten Bereich. Die Schätzungen sind mindestens so hoch wie jene des *NEO-PI-R*, des *NEO-FFI* oder des *BFI* (Borkenau & Ostendorf, 2008; Ostendorf & Angleitner, 2004; Rammstedt, 1997). Für empirische Untersuchungen auf Gruppenebene, für korrelative Untersuchungen als auch für Individualdiagnostik auf Personenebene können die manifesten Domänenwerte damit genutzt werden.

Die Reliabilitätsschätzer der Facettenwerte können als ausreichend für empirische Untersuchungen auf Gruppenebene und korrelative Untersuchungen angesehen werden. Die Höhe der Reliabilitätsschätzungen – sowohl auf Domänen- als auch auf Facettenebene – ist ähnlich wie die von Soto und John (2017), was darauf hindeutet, dass die deutsche und die US-Version des *BFI-2* ähnlich genaue Messungen ermöglichen.

Die konvergente Konstruktvalidität der *BFI-2*-Domänenwerte wird durch die hohen Korrelationen mit den konvergenten Domänenwerten des *NEO-PI-R*, des *NEO-FFI*, des *BFI* sowie der Kurzskalen *BFI-S* und *BFI-10* belegt. Im Vergleich mit den US-Daten zeigen sich tendenziell höhere Korrelationen mit dem *NEO-PI-R* und *NEO-FFI*, aber niedrigere Korrelationen mit dem *BFI*. Die konvergente Konstruktvalidität der *BFI-2*-Facettenwerte wird durch die konvergenten Korrelationen der *BFI-2*-Facetten mit den *NEO-PI-R*-Facetten belegt. Der Vergleich mit den Korrelationsmustern in der US-Stichprobe zeigt, dass die Zusammenhänge zwischen *BFI-2* und *NEO-PI-R*-Facetten in beiden Ländern ähnlich sind. Damit weisen die Ergebnisse insgesamt auf eine gute konvergente Konstruktvalidität der *BFI-2*-Skalen hin.

Die divergente Konstruktvalidität kann anhand der Korrelationen der Big-Five-Domänen untereinander bewertet werden. Erwartungsgemäß sind die Korrelationen zwischen divergenten Domänenwerten innerhalb des *BFI-2* geringer als die Korrelationen zwischen konvergenten Domänenwerten mit anderen Big-Five-Inventaren. Des Weiteren zeigen sich die in der Literatur berichteten positiven Zusammenhänge zwischen den manifesten Extraversions- und Offenheitswerten, zwischen Verträglichkeits- und Gewissenhaftigkeitswerten und negative Zusammenhänge zwischen den manifesten Negative Emotionalität- und den anderen Domänenwerten (Goldberg, 1992; Soto & John, 2017).

Die Zusammenhänge mit Bildung, Einkommen, Gesundheit und Lebenszufriedenheit belegen die Kriteriumsvalidität und die Nützlichkeit der Domänen- und der Facettenwerte. Die Domänenwerte können zwischen 6 % (Einkommen) und 23 % (Lebenszufriedenheit) der Varianz erklären. Die Facettenwerte können zwischen 8 % und 30 % an Kriteriumsvarianz erklären. Dies deutet darauf hin, dass spezifische Facetten die Zusammenhänge mit Lebenserfolgsindikatoren besser erklären können als globale Domänen. Die Ergebnisse der Regressionsanalysen deuten außerdem darauf hin, dass der Zugewinn an erklärter Varianz nicht alleine durch die höhere Anzahl der Prädiktoren, sondern durch höhere Spezifität der Facetten erklärt werden kann. Selbst wenn nur je eine Facette pro Domäne ausgewählt wird und selbst wenn diese manifesten Facettenwerte eine geringere Reliabilität als die Domänenwerte besitzen, können sie dennoch mehr Kriteriumsvarianz erklären als die globalen Domänen. Des Weiteren zeigen sich differenzielle Zusammenhänge zwischen den Kriterien und den Facetten einer Domäne. Diese Ergebnisse belegen eine gute Kriteriumsvalidität der Domänenwerte und darüber hinaus auch die inkrementelle Kriteriumsvarianz der Facetten gegenüber den Domänen.

Die Messinvarianztests deuten darüber hinaus auf residuale Messinvarianz des *BFI-2* für Frauen und Männer hin. Dies bedeutet, dass Unterschiede zwischen Frauen und Männern echte Unterschiede auf Konstruktebene widerspiegeln und auf manifester Ebene miteinander verglichen werden können.

Darüber hinaus regen die Ergebnisse der Studie weitere Forschung an, beispielsweise in Bezug auf hierarchische Strukturmodelle der Persönlichkeit, in Bezug auf den Mehrwert von spezifischen Persönlichkeitsfacetten für die Vorhersage von spezifischen Kriterien, in Bezug auf bis-

© 2019 Hogrefe Verlag Veröffentlicht unter der Hogrefe OpenMind-Lizenz (https://doi.org/10.1026/a000002)

${protocol}://econtent.hogrefe.com/doi/pdf/10.1026/0012-1924/a000218 - Wednesday, August 28, 2019 6:50:59 AM - IP Address:137.146.214.250

herige Befunde zum Zusammenhang von globalen Persönlichkeitsdomänen und Lebenserfolg und in Bezug auf die Messung von Persönlichkeit in verschiedenen Altersgruppen und die Implikationen der Messinvarianztests für entwicklungspsychologische Längsschnittstudien. Die vorliegende Arbeit und die hier veröffentlichte deutsche Version des *BFI-2* bieten die Grundlagen für diese Forschung.

Insgesamt belegen die Ergebnisse die gute psychometrische Qualität der *BFI-2*-Domänen- und Facettenwerte. In ca. 8 Minuten Bearbeitungszeit können die Big-Five-Domänen sowie 15 Facetten mit 60 Items reliabel, valide und messinvariant über Geschlechter gemessen werden. Das Instrument ist für nichtkommerzielle Forschung frei verfügbar und im Elektronischen Supplement 3 dargestellt.

Elektronische Supplemente (ESM)

Die elektronischen Supplemente sind mit der Online-Version dieses Artikels verfügbar unter https://doi.org/10.1026/0012-1924/a000218

ESM 1. Tabelle E1 - Tabelle E10
ESM 2. Tabelle E11 – Tabelle E30
ESM 3. Big Five Inventar 2 (BFI-2)

Literatur

Asparouhov, T. & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16*, 397 – 438. https://doi.org/10.1080/10705510903008204

Borkenau, P. & Ostendorf, F. (2008*). NEO-Fünf-Faktoren-Inventar nach Costa und McCrae (NEO-FFI).* Göttingen: Hogrefe.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*, 464 – 504. https://doi.org/10.1080/10705510701301834

Danner, D., Aichholzer, J. & Rammstedt, B. (2015). Acquiescence in personality questionnaires: Relevance, domain specificity, and stability. *Journal of Research in Personality, 57*, 119 – 130. https://doi.org/10.1016/j.jrp.2015.05.004

Danner, D., Rammstedt, B., Bluemke, M., Treiber, L., Berres, S., Soto, C. et al. (2016). Die deutsche Version des Big Five Inventory 2 (BFI-2). *Zusammenstellung sozialwissenschaftlicher Items und Skalen.* https://doi.org/10.6102/zis247

DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology, 91*, 1138 – 1151. https://doi.org/10.1037/0022-3514.91.6.1138

Goldberg, L. R. (1992). The development of markers for the Big Five factor structure. *Psychological Assessment, 4* (1), 26 – 42. https://doi.org/10.1037/1040-3590.4.1.26

Harkness, J. A. (2007). Improving the comparability of translations. In R. Jowell, C. Roberts, R. Fitzgerald & G. Eva (Eds.), *Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey* (pp. 79 – 94). New York, NY: Guilford Press, SAGE Publications Ltd. https://doi.org/10.4135/9781849209458

John, O. P., Donahue, E. M. & Kentle, R. L. (1991). *The Big Five Inventory – Versions 4a and 54.* Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.

John, O. P., Naumann, L. P. & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and Research* (pp. 114 – 158). New York, NY, US: Guilford Press.

McCrae, R. R. & Costa, P. T. Jr. (1987). Validation of the five-factor model of personality and across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81 – 90.

Mõttus, R. (2016). Towards More Rigorous Personality Trait-Outcome Research. *European Journal of Personality, 30*, 292 – 303. https://doi.org/10.1002/per.2041

Ostendorf, F. & Angleitner, A. (2004). *NEO-Persönlichkeitsinventar nach Costa und McCrae*, *Revidierte Fassung* (NEO-PI-R). Göttingen: Hogrefe.

Paunonen, S. V. (2003). Big Five factors of personality and replicated prediction of behavior. *Journal of Personality and Social Psychology*, 84, 411 – 424. http://doi.org/10.1037/0022-3514.84.2.411

Rammstedt, B. (1997). *Die deutsche Version des Big Five Inventory (BFI): Übersetzung und Validierung eines Fragebogens zur Erfassung des Fünf-Faktoren-Modells der Persönlichkeit.* Unveröffentlichte Diplomarbeit, Universität Bielefeld.

Rammstedt, B. & Danner, D. (2016). Die Facettenstruktur des Big Five Inventory (BFI): Validierung für die deutsche Adaptation des BFI. *Diagnostica, 63*, 70 – 84. https://doi.org/10.1026/0012-1924/a000161

Rammstedt, B. & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41*, 203 – 212. https://doi.org/10.1016/j.jrp.2006.02.001

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A. & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*, 313 – 345. https://doi.org/10.1111/j.1745-6916.2007.00047.x

Schupp, J. & Gerlitz, J.-Y. (2014). Big Five Inventory-SOEP (BFI-S). *Zusammenstellung sozialwissenschaftlicher Items und Skalen.* https://doi.org/10.6102/zis54

Soto, C. J. & John, O. P. (2009). Ten facet scales for the Big Five Inventory: Convergence with NEO-PI-R facets, self-peer agreement, and discriminant validity. *Journal of Research in Personality, 43*, 84 – 90. https://doi.org/10.1016/j.jrp.2008.10.002

Soto, C. J. & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology, 113*, 117 – 143. https://doi.org/10.1037/pspp0000096

Statistisches Bundesamt (2014). *Zensus 2011.* Zugriff am 02.12.2016 unter http://www.zensus2011.de

Tupes, E. C. & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality, 60*, 225 – 251.

Historie
Onlineveröffentlichung: 31.01.2019

Danksagung
Wir danken dem Hogrefe Verlag für seine Unterstützung sowie Olli Schulz und den anonymen Gutachterinnen bzw. Gutachtern für Ihre hilfreichen Kommentare.

© 2019 Hogrefe Verlag Veröffentlicht unter der Hogrefe OpenMind-Lizenz (https://doi.org/10.1026/a000002)

Prof. Dr. Daniel Danner
Hochschule der Bundesagentur für Arbeit (HdBA)
Seckenheimer Landstraße 16
68163 Mannheim
daniel.danner@hdba.de

Prof. Dr. Beatrice Rammstedt
Dr. Matthias Bluemke
Dr. Clemens Lechner
Thomas Knopf
GESIS – Leibniz-Institut für Sozialwissenschaften
Postfach 122155
68072 Mannheim

Sabrina Berres
Universität Mannheim
B6, 30 – 32
68161 Mannheim

Christopher J. Soto, Ph.D.
Colby College
5550 Mayflower Hill, Waterville
Maine 04901 – 8855
USA

Prof. Dr. Oliver P. John
University of California
Berkeley
California 94720 – 1650
USA

${protocol}://econtent.hogrefe.com/doi/pdf/10.1026/0012-1924/a000218 - Wednesday, August 28, 2019 6:50:59 AM - IP Address:137.146.214.250

© 2019 Hogrefe Verlag Veröffentlicht unter der
Hogrefe OpenMind-Lizenz (https://doi.org/10.1026/a000002)