

# NEO-Persönlichkeitsinventar nach Costa und McCrae, revidierte Fassung (NEO-PI-R)

B. ANDRESEN,  
Universitätsklinikum Hamburg-Eppendorf

A. BEAUDUCEL,  
Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg

Testbeurteilungssystem –  
Testkuratorium  
der Föderation deutscher  
Psychologenvereinigungen

TBS-TK  
Rezension

## Allgemeine Informationen über den Test, Beschreibung des Tests und seiner diagnostischen Zielsetzung

Das NEO-PI-R ist eine autorisierte Fassung des amerikanischen Originals für die Selbst- (Form S) und Fremdbeurteilung (Form F) der Persönlichkeit. Der Fragebogen umfasst die Domänen Neurotizismus (N), Extraversion (E), Verträglichkeit (A), Gewissenhaftigkeit (C) und Offenheit für Erfahrungen (O). Die Domäne O erfasst weniger die Komponente »Intellekt« des lexikalischen Ansatzes als vielmehr einen erlebnisbetonten, kreativen Motivationsfaktor, der auch die Beschäftigung mit Neuem einschließt. Die Durchführung erfordert ca. 35 Minuten Bearbeitungszeit und erfasst mit 240 Items die Domänen und Facetten (sechs pro Domäne) des Fünf-Faktoren-Modells (FFM) der Persönlichkeit. Die Bandbreite seiner Anwendung (ab 16 Jahren) in Forschung und Praxis ist groß. Der zentrale Anspruch des Verfahrens liegt in der erschöpfenden und detaillierten Beschreibung der menschlichen Persönlichkeit auf der Basis einer faktorenanalytisch gesicherten Vollständigkeit der Domänen.

## Theoretische Grundlagen als Ausgangspunkt der Testkonstruktion

Der Konstruktion des NEO-PI-R liegen zahlreiche faktorenanalytische Studien nach dem lexikalischen und dem Fragebogenansatz der Eigenschaftskonzeption zugrunde. Eine wichtige Prämisse ist die hierarchische Untergliederung des Merkmalsraums in Domänen (inhaltlich sehr breit angelegte und unkorrelierte Faktoren) und Facetten (inhaltlich deutlich engere, hoch korrelierte Merkmale) mit der Annahme, dass es definitiv nur fünf Domänenfaktoren gibt, während die Zahl der Facetten pragmatisch limitiert wird. Die vorgestellten faktorenanalytischen Konstruktvalidierungen zielen darauf ab nachzuweisen, dass das NEO-PI-R den Domänenbereich komplett abdeckt. Die Autoren betonen jedoch, dass das NEO-PI-R auch ohne eine Identifikation mit dem FFM diagnostisch wertvoll und ergiebig ist. Die Interkorrelationen der Domänenskalen weichen in den Gesamtstichproben in drei Paarungen deutlich von null ab: Form S: O/E 0,40, N/C -0,37, N/E -0,27; Form F: O/E 0,41, N/E -0,39, N/E -0,21. Dieses Ergebnis kann als moderate Abweichung vom Konstruktionsziel gelten.

## Objektivität

Die fachlichen Voraussetzungen für eine Testanwendung werden spezifiziert. Es besteht weitgehende Durchführungs- und Auswertungsobjektivität durch informative und verständliche Anweisungen und die leicht einzu-

übende Anwendung aller standardisierten Testelemente. Auch der Umgang mit fehlenden Werten wird präzise geregelt. Die Interpretationsobjektivität ist durch eine gute Vorstellung von Anleitungen und Interpretationshilfen gesichert. Gut ist die differenzierte Argumentation und Anleitung zum Umgang mit Differenzwerten. Allerdings könnten die Konsequenzen der Wahl von allgemeinen oder nach Geschlecht, Alter und Schulbildung differenzierten speziellen Normen besser verdeutlicht werden.

## Normierung (Eichung)

Für Form S werden T-Werte (auch Umrechnung auf Stanine-Werte und Prozentränge) mit optional geschlechts- und altersgruppenspezifischen Normierungen angeboten. Die Bestimmung der Werte wird direkt auf den Profildbögen vorgenommen. Es werden Normen und »Repräsentativ«-Normen unterschieden, Letztere mit kleinerem N und ohne Altersdifferenzierung. Die Repräsentativität bezieht sich nur auf Geschlecht, Alter und Bildungsstand (sekundäre Quotenstichprobe). Für Skala O werden im Anhang Normen für Schulbildungsgruppen (bei Reduktion der Altersgruppen) angeboten, da O signifikant mit der Schulbildung korreliert. Für Form F sind die Normen altersmäßig weniger differenziert, hier liegen keine Repräsentativnormen vor. Für Patienten liegen keine gesonderten Normen vor, obwohl klinische Gruppen psychometrisch separat analysiert werden. Die Normen wurden nicht in einer einzelnen Repräsentativerhebung gewonnen; sie stammen vielmehr aus den Stichproben einer größeren Zahl von Einzelstudien. Im Manual wird als »Schätzung« für den Zeitpunkt der mehrjährigen Normierungserhebung das Jahr 1999 angegeben (S. 82). Offenbar hatten die verschiedenen Erhebungszeitpunkte keine relevanten Einflüsse auf die Normwerte, was ein Hinweis auf die zeitliche Stabilität der vorliegenden Normen ist.

## Zuverlässigkeit (Reliabilität, Messgenauigkeit)

Interne Konsistenzen wurden für Form S in verschiedenen Alters- und Geschlechtsgruppen (N zwischen 139 und 11724) bestimmt. Die Domänen zeigen sehr gute interne Konsistenzen (Cronbachs Alpha) um 0,90, wobei Skala A mit 0,85 leicht abfällt. Die Facetten haben annehmbare Konsistenzen um 0,70. Facette O6 (»Werte«) ist mit einem Alpha von 0,50 deutlich weniger konsistent als im amerikanischen Original. Hierbei spielen Kulturunterschiede eine besondere Rolle. Die Facettenskalen für A fallen ebenfalls etwas ab. Für Form F liegen bei Gruppengrößen von 221 – 1547 die Konsistenzen durchweg etwas höher. Die Konsistenzwerte der repräsentativen Stichproben (423 – 871 Pbn) bestätigen obige Angaben. Retest-Reliabilitä-

	NEO-Persönlichkeitsinventar nach Costa und McCrae, Revidierte Fassung	Die TBS-TK-Anforderungen sind erfüllt			
		voll	weitgehend	teilweise	nicht
Allgemeine Informationen, Beschreibung und diagnostische Zielsetzung		●			
Objektivität			●		
Zuverlässigkeit			●		
Validität			●		

ten (nur Form S) wurden bei Abständen von einem Monat bis fünf Jahren für teilweise sehr kleine Stichproben von 10 – 391 Pbn bestimmt. Für Stichprobengrößen von 70 aufwärts wurden bei Abständen von ein bis zwei Monaten Retest-Koeffizienten für Domänen von 0,82 – 0,91 (Facetten: 0,48 – 0,91), bei Abständen von fünf Jahren Werte zwischen 0,74 und 0,78 (Facetten: 0,53 – 0,78) ermittelt.

**Gültigkeit (Validität)**

Die Konstruktvalidierung konzentriert sich auf die Domänenfaktoren, für die in verschiedenen Unterstichproben und im Vergleich zum amerikanischen Original eine hohe Kongruenz gezeigt wird. Die rationale Konstruktion der Facetten wird nicht näher erläutert. Die Kriteriumsvalidität wird auf der Basis internationaler Studien im Abschnitt »Einsatzbereiche« selektiv kommentiert. Zur Kriteriumsvalidität des originalen NEO-PI-R und seiner vielen Übersetzungen liegen diverse Studien vor, die in der Handanweisung nicht gebündelt dargestellt werden. Im Besonderen werden biopsychologische und verhaltensgenetische Validierungsstudien nur pauschal angesprochen. Demgegenüber werden eingehend solche Studien vorgestellt, die bei gemeinsamer Faktorenanalyse von NEO-PI-R und anderen Verfahren einheitlich das FFM bestätigen.

**Weitere Gütekriterien (Störanfälligkeit, Unverfälschbarkeit und Skalierung)**

Möglichkeiten der Vorgabe mit oder ohne Anwesenheit eines Testleiters werden nicht klar beschrieben. Antworttendenzen (z. B. sozial erwünschtes Antworten) bleiben teilweise mit dem Hinweis auf nicht schlüssige Konzepte und Methoden zur Aufdeckung von Antwortstilen und fehlerhaftem Antwortverhalten ohne verbindliche bzw. schlüssige Regeln. Andere Kontrollversuche werden an letztlich willkürlichen Grenzwerten für die Zahl gewählter Antwortkategorien festgemacht. Die 21 verschiedenen Profilbögen sind in der Farbgebung insgesamt eher verwirrend gestaltet, was fehlerhaften Zugriff zur Folge haben könnte. Geschlecht und Schulbildung sollten ergänzend zum Alter vom Testleiter direkt eingetragen werden können, auch um Auswahlfehler der Bögen zu verhindern. Die Testkonstruktion folgt ausschließlich der klassischen Testtheorie.

**Abschlussbewertung/Empfehlung**

Es handelt sich insgesamt um die gelungene Adaptierung eines solide konstruierten amerikanischen Persönlich-

keitstests mit grundwissenschaftlichem Schwerpunkt, sehr großer Verbreitung und weltweiter Marktführerschaft. Die fünf Domänenskalen sind hoch konsistent und hinreichend zeitstabil gemessen, aber teilweise in den Faktorenpaaren E/O, N/C und N/E stärker korreliert als in konkurrierenden Verfahren, die bei den entsprechenden Paarungen angenähert Null-Korrelationen erzielen. Hier könnten Rekonstruktionen sinnvoll sein. Auf der Facetenebene besteht bei einzelnen Facetten, vor allem O6, Optimierungsbedarf hinsichtlich der Reliabilität. Generell wäre eine faktorenanalytische Überprüfung der Facetenebene anzuraten. Die Zeitstabilität ist für die Formen S und F asymmetrisch berücksichtigt, hier sollte die Form F ebenfalls einbezogen werden. Eine gleichgewichtig für die Formen S und F angelegte Neunormierung auf der Grundlage repräsentativer Stichproben wäre anzustreben. Diese könnte auch das unübersichtliche Auswertungsmaterial vereinfachen. In der zentralen Frage der Domänenvollständigkeit wird empfohlen, auch das FFM offen und fair in Frage zu stellen und Falsifikationsversuche zu fördern. Das durchgängige Prinzip, das NEO-PI-R-Skalenensemble als vollständig und faktorenanalytisch inventarübergreifend konstruktvalide darzustellen, hat möglicherweise dazu geführt, neuere Modelle mit mehr als fünf Domänenfaktoren, wie die von Comrey, Tellegen und Waller, Cloninger et al., Hogan und Hogan, Jackson et al., Ashton et al., Andresen oder Saucier, empirisch und konzeptuell nicht vergleichend einzubeziehen. Trotz der wünschenswerten Weiterentwicklungen handelt es sich beim NEO-PI-R aber um ein – gemessen am Konstruktionsziel – sehr gut konstruiertes Verfahren, dessen Einsatz eine fachgerechte und differenzierte Persönlichkeitsdiagnostik nach dem FFM gestattet.

*Diese Testrezension wurde im Auftrag des Testkuratoriums der Föderation deutscher Psychologengvereinigungen (DGPs und BDP) gemäß den TBS-TK-Richtlinien (Testkuratorium, 2006, 2007) erstellt.*

*Testkuratorium (2006). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologengvereinigungen. Report Psychologie, 31, 492-499.*

*Testkuratorium (2007). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologengvereinigungen. Psychologische Rundschau, 58, 25-30.*

**TESTINFORMATIONEN**

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

Bezugsquelle: Testzentrale Göttingen, Robert-Bosch-Breite 25, 37079 Göttingen.

Test komplett, 210 €. Computerversion inkl. 50 Durchführungen und Manual, 980 €. Computerversion 50 weitere Durchführungen, 80 €. Eine Durchführung (Testing on Demand) 8 €.

Bitte zitieren Sie diesen Artikel wie folgt: Andresen, B. & Beauducel, A. (2008). TBS-TK Rezension: »NEO-Persönlichkeitsinventar nach Costa und McCrae, Revidierte Fassung (NEO-PI-R)«. Report Psychologie, 11/12;2008.