

OPQ32- Occupational Personality Questionnaire

C. DORMANN, Institut für Psychologie der Johannes Gutenberg-Universität Mainz
 S. KRUMM, Institut für Psychologie, Westfälische Wilhelms-Universität Münster

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

Der OPQ32 ist ein proprietäres Verfahren der SHL Group. Er will »präferierte oder typische Verhaltensstile am Arbeitsplatz« messen. Er kann für individuelle (z.B. Personalauswahl, Karriereplanung), gruppenbezogene (z.B. Teamentwicklung) und organisationale (z.B. Nachfolgeplanung) Fragestellungen eingesetzt werden. Hauptzielgruppen sind Fach- und Führungskräfte, der Anwendungsbereich erstreckt sich prinzipiell auf alle Berufstätigen mit und ohne Hochschulabschluss. Die Verhaltensstile werden auf 32 Dimensionen erfasst. Diese Dimensionen basieren auf einem Modell mit drei übergeordneten Bereichen: Beziehungen, Denken und Gefühle/Emotionen. Durch spezifische Zuordnungen lassen sich auch die Big Five nach dem Fünf-Faktoren-Modell abbilden. Der OPQ32 stellt eine Weiterentwicklung verschiedener Vorgängerverfahren dar und liegt in einer normativen (Likert-Skalen, OPQ32n, sechs bis acht Items pro Dimension) und einer ipsativen Form (»forced choice ratings«, OPQ32i, 13 Items pro Dimension) vor. Für beide Varianten existieren Online-, (Pocket-)PC-basierte und Papier-und-Bleistift-Versionen. Es liegen umfangreiche Normen für verschiedene Länder, Berufsgruppen und Managementebenen vor.

Theoretische Grundlagen als Ausgangspunkt der Testkonstruktion

Grundlage für die Entwicklung des OPQ war eine Vielzahl existierender Persönlichkeitsmodelle (z.B. Cattell, Eysenck). Die Entwicklung ab den 1980er-Jahren konzentrierte sich auf Facetten, die entweder einen empirischen Bezug zur beruflichen Tätigkeit und zur Leistung aufwiesen oder von den Klienten von SHL als relevant erachtet wurden. Im weiteren Verlauf wurde aufgrund der verfügbaren empirischen Informationen, dem Streben nach internationaler Anwendbarkeit u.a.m. verschiedene Revisionen vorgenommen, die in ein neues theoretisches Modell mündeten. Demnach würden sich die 32 Facetten des Modells drei Domänen (Beziehungen zu Menschen, Denkstile, Gefühle und Emotionen) zuordnen lassen, was jedoch nicht der von den Autoren durchgeführten Faktorenlösung (verschiedene Lösungen mit vier bis 17 Faktoren) entspricht. Die im Manual berichtete empirische Überlappung der OPQ-Skalen

mit den Big Five zeigt aber, dass eine Anbindung an eingeführte Modelle möglich wäre.

Objektivität

Das Benutzerhandbuch des OPQ32 gibt umfangreiche Hinweise zur Durchführung, Auswertung und Interpretation. Folgt der Anwender den auf insgesamt neun Seiten dargestellten Hinweisen (inklusive Checklisten) zur Durchführung der verschiedenen OPQ32-Versionen, so ist von Durchführungsobjektivität auszugehen, zumal die Durchführung eine geringe bis keine Interaktion mit dem Testleiter erfordert. Die Objektivität bei der Durchführung ohne Anwesenheit des Testleiters (»unproctored testing«) ist relativ hoch; der Einfluss auf die Ergebnisse ist gering (in der Regel $< 0,25$ SD). Die Auswertung erfolgt bei der Papier-und-Bleistift-Version mit einem Antwortbogen, anderenfalls ist sie automatisiert. Ein Einscannen des Antwortbogens und eine anschließende automatisierte Auswertung sind möglich. Das Manual des OPQ32 enthält zudem umfangreiche Angaben und Beispiele zur Interpretation der Ergebnisse; es besteht die Möglichkeit, narrative, automatisierte Ergebnisreports in verschiedenen Sprachen zu erstellen. Wird von dieser Möglichkeit Gebrauch gemacht, so kann von hoher Interpretationsobjektivität ausgegangen werden.

Normierung (Eichung)

Der OPQ32 unterliegt einem ständigen Normierungsprozess. Die im Manual enthaltenen Normen aus 1998/99 werden durch aktuellere Normen ergänzt. Die UK-Normen aus 2000 bis 2004 beziehen sich auf die ipsative und die normative Form des OPQ32; die internationalen Normen (von 2004 bis 2008 erhoben) allerdings nur auf die ipsative Form. Letztere beinhalten unterschiedliche und regionalisierte Sprachversionen (z.B. Englisch US, Englisch UK, Englisch Südafrika, Französisch). Die Normen des OPQ differenzieren zudem zwischen spezifischen Wirtschaftsbereichen, Jobpositionen u.a.m. Insgesamt gehen in die aktuellste Normierung mehr als 330 000 Datensätze ein. Zur Normierung wird die STEN-Skala verwendet, deren Auflösungsgrad im beruflichen Kontext gut geeignet ist. Standardmessfehler werden berichtet. Weiterhin findet der Anwender umfangreiche Angaben zu den zu erwartenden Geschlechtsunterschieden in den einzelnen Normpopula-

OPQ32	Die TBS-TK-Anforderungen sind erfüllt			
	voll	weitgehend	teilweise	nicht
Allgemeine Informationen, Beschreibung und diagnostische Zielsetzung	●			
Objektivität	●			
Zuverlässigkeit		●		
Validität		●		

tionen. Als proprietäres Verfahren sind die Stichproben qua Definition repräsentativ (Kunden von SHL). Besonders hervorzuheben ist an dieser Stelle, dass die Normen weitgehend im Kontext tatsächlicher Auswahlverfahren erhoben wurden. Inwiefern diese Normen allerdings für die zahlreichen weiteren Anwendungsfelder (z.B. Personalentwicklung, Teambuilding) gültig sind, bleibt offen. Für die normative Variante des OPQ erscheint es erstrebenswert, zukünftig ähnlich umfangreiche internationale Normen bereitzustellen wie für die ipsative Variante.

Zuverlässigkeit (Reliabilität, Messgenauigkeit)

Für den OPQ32n und OPQ32i existieren unterschiedliche Reliabilitätsschätzungen für diverse Stichproben und für unterschiedliche Vorgabemodalitäten. Für den OPQ32n liegen sie im Durchschnitt bei über 0,80, für den OPQ32i nicht ganz bei 0,80. Retest-Reliabilitäten liegen nur für eine kleine studentische Stichprobe im Intervall von vier Wochen vor ($0,63 < r_{tt} < 0,92$; $N = 107$). Werden die OPQ-Werte zu den Big Five aggregiert, ergeben sich Reliabilitäten von $> 0,89$ (Ausnahme: Verträglichkeit mit 0,85 bzw. 0,87 für den OPQ32n bzw. OPQ32i). Insgesamt sind die Reliabilitäten als zufriedenstellend bis gut zu bezeichnen. Die Tatsache, dass bei der ipsativen Version (OPQ32i), die Forced-Choice-Items benutzt, ebenfalls gute Reliabilitäten erzielt werden, lässt den Schluss zu, dass die Reliabilitäten nicht durch bloßen Content-Overlap der Items erzielt werden. Allerdings ist aufgrund der Angaben im technischen Manual nur nach einigem Suchen erkenntlich, wie aktuell die jeweiligen Reliabilitätsprüfungen sind.

Gültigkeit (Validität)

Ein herausragendes Merkmal dieses Verfahrens ist seine umfangreiche Validierung. Zur Konstruktvalidierung wurden neben der intendierten Anwendungspopulation (Manager, Vorgesetzte) auch studentische Stichproben eingesetzt. Die Kriteriumsvalidierung erfolgte ausschließlich an der intendierten Anwendungspopulation. Es werden auch Angaben zur inkrementellen Validität (z.B. über Leistungstests hinaus) gemacht. Für den OPQ32 gibt es neben konkurrenten Validierungen nur eine retrospektive (mit früherer Leistung), aber keine prädiktive Validierung. Insgesamt zeigen die Angaben zur Kriteriumsvalidität, dass eine Kombination verschiedener OPQ32-Skalen in bekannter Größenordnung mit beruflicher Leistung zusammenhängt. Insgesamt ist

die Konstruktvalidität gut und die leistungsbezogene Kriteriumsvalidität gering; sie liegt damit aber im Rahmen der Kriteriumsvaliditäten, wie sie in Metaanalysen zum Zusammenhang zwischen Persönlichkeit und beruflicher Leistung aufgezeigt worden sind. Bezüglich der Konstruktvalidität der einzelnen Subskalen des OPQ32 präsentieren die Autoren eine große Vielzahl nicht immer einheitlich ausfallender Korrelationen mit anderen Persönlichkeitsinventaren. Zumindest für einige Subskalen stehen klare Validitätsbelege daher noch aus (z.B. keiner der Validitätskoeffizienten für die Skalen »Competitive« und »Decisive« war $> .30$).

Weitere Gütekriterien (Störanfälligkeit, Unverfälschbarkeit und Skalierung)

Aufgrund seiner langen anforderungsbezogenen Konstruktionsgeschichte und seiner weiten Verbreitung erfüllt der OPQ das Nebengütekriterium der Nützlichkeit. Kritisch zu hinterfragen ist der Einsatz eines prinzipiell verfälschbaren Verfahrens zu Selektionszwecken. Bei der ipsativen Version ist aufgrund des Antwortformats mit geringerer Verfälschung zu rechnen, und die hohen Korrelationen der ipsativen mit der normativen Version sprechen indirekt auch für eine geringe Verfälschung der normativen Version. Ebenso belegen die geringen Unterschiede zwischen »proctered« und »unproctered testing« indirekt den geringen situativen Einfluss. Jedoch würde man sich im Manual direkte Evidenz zum Einfluss der Verfälschung (z.B. Selektion vs. Konsultation) wünschen.

Abschlussbewertung/Empfehlung

Als besonderes Qualitätsmerkmal des OPQ ist die umfangreiche empirische Fundierung zu sehen. Detaillierte psychometrische Analysen (Faktorenladungen, Schwierigkeiten, Trennschärfen) sind zwar auch im technischen Manual nicht enthalten, dieses stellt aber auf 200 Seiten eine Fülle von Untersuchungen zusammen, sodass an psychometrischer Qualität und praktischem Nutzen wenig Zweifel bestehen. Im Bereich der berufsbezogenen Diagnostik stellt der OPQ damit eine erfreuliche Ausnahme dar. Lediglich bei zwei Aspekten bleiben Fragezeichen. Zum einen ist die Konstruktvalidität mancher der 32 Skalen nicht bzw. nicht ausreichend geklärt. Die vergleichsweise mittleren bis niedrigen Interkorrelationen der 32 Skalen deuten aber indirekt auf deren diskriminante Validität hin. Zum zweiten sind die psychometrischen Eigenschaften des ipsativen OPQ32 nicht immer hundertprozentig überzeugend, was teilweise in dem ipsativen Format begründet ist. Insgesamt kann die Konstruktion des OPQ als sorgfältig bezeichnet werden.

Obwohl die Autoren zunächst ein von etablierten Modellen abweichendes Modell zugrunde legen, unternehmen sie nachträglich viele Anstrengungen, um die Konvergenz der Skalen des OPQ32 mit dem Fünf-Faktoren-Modell der Persönlichkeit aufzuzeigen. Hier würde eine größere konzeptuelle Klarheit den OPQ noch weiter von anderen Verfahren der beruflichen Diagnostik abheben. In diesem Zusammenhang muss auch die Darstellung der für den Anwender wichtigen Infor-