

## »Golden Profiler of Personality (GPOP). Deutsche Adaptation des Golden Personality Type Profiler von John P. Golden«

S. HÖFT, Hochschule der Bundesagentur für Arbeit, Mannheim  
P. M. MUCK, Universität Bielefeld

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

Der GPOP erfasst die Persönlichkeitsstruktur von Individuen mit Hilfe von vier bipolaren Globalskalen (Energie, Wahrnehmen, Entscheiden, Lebensstil). Durch die Kombination der unterschiedlichen Pole werden insgesamt 16 Persönlichkeitstypen unterschieden. Zusätzlich wird jede Skala durch fünf zugeordnete Facetten genauer spezifiziert. Eine weitere globale Stressskala erfasst die Art, wie der jeweilige Typus geäußert wird.

Der Test kann bei Personen ab 16 Jahren eingesetzt werden und dauert ca. 35 Minuten. Die Durchführung kann computergestützt (Hogrefe-Testsystem) oder als Papier-Bleistift-Verfahren erfolgen, allerdings fehlen zu beiden Varianten genauere Durchführungshinweise. Die Auswertung muss zwingend computergestützt erfolgen. Ein vollständiger Ergebnisbericht umfasst 97 Seiten mit allgemeinen Beschreibungen, typenspezifischen Darstellungen und grafischen (Sub-)Skalenauswertungen. Die für den Probanden bestimmte Kurzversion ist 22 Seiten lang.

### Theoretische Grundlagen als Ausgangspunkt der Testkonstruktion

Als theoretische Grundlage werden unterschiedliche Persönlichkeitstheorien (vorrangig Jung, Allport, Eysenck, Adler) angegeben, neuere Ansätze (z.B. das Fünf-Faktoren-Modell) werden nur gestreift. Die Referenzliteratur ist stark veraltet. Generell ist die theoretische Begründung zu kurz und erfolgt zu global. Dadurch wirkt das verwendete Typenmodell mit seiner speziellen Logik (z.B. zur dynamischen Interaktion der Präferenzen und der dominanten Funktion) sehr apodiktisch. Theoretisch nicht eingebunden sind die zusätzlich verwendete Facettenebene sowie die fünfte Globalskala. Auch das Begleitbuch »Typisch Mensch« bleibt bei einer einfachen Deskription des Modells und bietet keine wissenschaftlich fundierte Begründung. Für die Herleitung der Skalen und Items wird ohne ausführliche Erläuterung auf das amerikanische Original verwiesen. Die deutschsprachige Adaptation erfolgt sehr frei und richtet sich nicht nach geltenden Adaptationsstandards (z.B. kein Einsatz von bilingualen Muttersprachlern, keine Rückübersetzung). Anders als von den Autoren behauptet, können Befunde zum Originalverfahren somit nicht übertragen werden.

### Objektivität

Es fehlen standardisierte Vorgaben für die Testdurchführung. Die computerisierte Vorgabe erfolgt aber über das Hogrefe-Testsystem und ist in diesem Rahmen kontrolliert durchführbar. Mangels Dokumentation kann nur indirekt geschlossen werden, dass die ausgefüllten Papier-Bleistift-Fragebogen durch Assistenzkräfte händisch in das System

eingetragen werden müssen, wodurch Eingabefehler entstehen können. Die Auswertung erfolgt komplett computergestützt und verdeckt. Verrechnungsalgorithmen werden nicht transparent gemacht. Dadurch ist nur schwer rückverfolgbar, auf welche Itembeantwortung sich die unterschiedlichen Passagen des standardisierten Gutachtentextes stützen. Das Manual beschränkt sich darauf, Gutachtentexte im Wortlaut wiederzugeben. Darüber hinausgehende Interpretationshinweise fehlen. Es ist zu hoffen, dass in dem begleitend angebotenen Lizenzierungseminar tiefer gehende Strategien zur Ergebniskommunikation und weitergehenden Beratung vermittelt werden.

### Normierung (Eichung)

Die im Gutachten sehr intensiv genutzte Normierung stützt sich auf eine Stichprobe von insgesamt 2.568 Personen, die über eine internetgestützte Befragung erhoben wurde. Weitergehende Informationen zu den situativen Rahmenbedingungen der Stichprobenziehung (Waren die Teilnehmer z.B. prinzipiell identifizierbar? Welche Konsequenzen ergaben sich aus der Teilnahme?) sowie zur Repräsentativität (für die Gesamtbevölkerung oder für Berufsgruppen?) fehlen. Relevante demografische Randvariablen werden zwar erfasst (Alter, Geschlecht, Branchenzugehörigkeit, Hierarchiestufe), Analysen zu Subgruppenunterschieden und damit notwendigen spezifischen Normen werden aber nicht durchgeführt.

Im Gutachten werden nicht dokumentierte Normwerte für die Facettenscores eingesetzt. Auf Ebene der Globalskalen wird wegen deutlicher Abweichungen von einer Normalverteilung auf eine Prozentrangnormierung zurückgegriffen.

### Zuverlässigkeit (Reliabilität, Messgenauigkeit)

Reliabilitätsschätzungen erfolgen ausschließlich über interne Konsistenzanalysen und Split-Half-Berechnungen. Die ermittelten Werte sind auf der Ebene der Globalskalen gut bis sehr gut (zwischen 0,82 und 0,90), auf der untergeordneten Facettenebene liegen sie mehrheitlich unter 0,60. Auf die eingeschränkte bis nicht vorhandene Aussagekraft der Facettenskalen wird zwar hingewiesen (S. 38) und eine entsprechende Vorsicht bei späteren »Validierungsgesprächen« angeraten, erforderliche Warnhinweise fehlen allerdings in dem sonst sehr ausführlichen Gutachtentext. An dieser Stelle wäre eine zusätzliche Retest-Reliabilitätsschätzung zwingend notwendig, um die zeitliche Stabilität der ermittelten Werte zu erfassen. Schließlich fehlen ebenfalls Angaben zur Profilreliabilität der Typenstruktur sowie zur Stabilität der errechneten Typen. Letzteres wäre besonders wichtig, da die Zuordnung schon ab einem Prozent Abweichung zwischen zwei bipolaren Skalen einer Globalskala erfolgt.

### Das Testkuratorium

(TK) der Föderation Deutscher Psychologinnen und Psychologen dankt den Rezensentinnen und Rezensenten der Testverfahren, die im Jahr 2008 nach dem Testbeurteilungssystem (TBS) evaluiert worden sind, für ihre ausgezeichnete Arbeit im Dienste der Qualitätssicherung von Testverfahren.

Gerhard Stemmler  
(Beauftragter des TK für TBS-TK Rezensionen)

Rezensionen wurden 2008 erstellt von  
Burghard Andresen  
André Beauducel  
Susanne Goldberg  
Nina Heinrichs  
Philipp Y. Herzberg  
Tuulia M. Ortner  
Wolfgang Rauch  
Lothar Schmidt-Atzert  
Kurt Sokolowski

**Gültigkeit (Validität)**

Die größte Schwachstelle des GOP sind die fehlenden Validitätsbefunde. Eine explorative Faktorenanalyse auf Itemebene bestätigt ansatzweise die angenommene Fünf-Faktoren-Struktur. Informativer und besser darstellbar wäre an dieser Stelle eine Faktorenanalyse auf Basis der Facettenskalen gewesen. Im Manual wird auf die teilweise sehr hohen Interkorrelationen der Globalskalen (bis zu  $r = -0,68$ ) zwar hingewiesen, bei der späteren Interpretation im Gutachten werden aber keine Konsequenzen gezogen. Eine konvergente und/oder diskriminante Validierung mit externen Konstrukten (z.B. anderen Persönlichkeitsinventaren) fehlt ebenso wie eine Validierung an relevanten Außenkriterien. Dies verwundert umso mehr, als der Anwendungszweck Aussagen zu Führungsstil, Lernstil sowie Beiträgen im Team und Beiträgen zur Organisation verspricht. Letztlich sind wesentliche Aussagen im Gutachten empirisch nicht belegt.

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

Hinweise zur Störanfälligkeit und Verfälschbarkeit fehlen. Sie dürften aber vergleichbar mit anderen Persönlichkeitsverfahren sein. Der GOP wird explizit zur Personalentwicklung empfohlen, der Einsatz für die Personalauswahl wird zwar nicht ausgeschlossen, aber nur im Rahmen eines umfangreichen Batterieinsatzes empfohlen (S. 9). Insoweit ist bei den meisten Anwendungen die Gefahr intentional beschönigter Selbstdarstellungen als relativ gering einzuschätzen. Die Testkonstruktion folgt klassisch testtheoretischen Grundprinzipien, setzt etablierte Standards aber nur ansatzweise um. Weitergehende Überlegungen zu Skalierungsproblemen werden nicht angestellt.

**Abschlussbewertung/Empfehlung**

Der GOP ist ein Repräsentant der in der wirtschaftlichen Anwendung weitverbreiteten typologischen Persönlichkeitsverfahren. Die Grundstruktur des Verfahrens ist dimensional ausgerichtet. Erst durch eine vergrößerte Testauswertung (es wird mittels Differenzbildung untersucht, ob der positive oder negative Pol der Globalskala bzw. Facette überwiegt) werden 16 Typen erzeugt. Über eine separate Stressskala sollen spezielle Ausdrucksformen betrachtet werden.

Der GOP weist in der jetzigen Form mehrere Problem-bereiche auf: Das Verfahren basiert auf einer veralteten, nicht empirisch geprüften Theorie. Nicht weiter belegte Behauptungen zur Ermittlung sogenannter dominanter, sekundärer und tertiärer Funktionen im Rahmen des sich ergebenden Typus werden als allgemeingültig dargestellt. Die Integration des Stressgrads als fünfte Globalskala wirkt aufgesetzt. Die Verwendung von Facettenskalen hat keinen theoretischen Hintergrund.

Die Verfahrensentwicklung folgt klassisch testtheoretischen Grundsätzen, weist dabei aber deutliche Lücken auf. So folgte die Verfahrensübersetzung nicht den geltenden Standards. Schwer wiegen die Reliabilitätsmängel speziell auf der Facettenebene sowie die zu grobe Normierung. Bei der Verfahrenvalidierung zeigen sich die größten Lücken: Konstruktvalidierungsstudien unter Verwendung anderer Konstrukte fehlen ebenso wie Studien

## Testbeurteilungssystem – Testkuratorium der Föderation deutscher Psychologinnenvereinigungen

# TBS-TK

Rezension

zur Kriteriumsvalidierung. Befunde zur Übereinstimmung von Selbst- und Fremdbild liegen ebenfalls nicht vor. Für Anwender ist das Verfahren aufgrund fehlender Information im Manual nur schwer zu durchschauen. Zusammengefasst handelt es sich beim GOP um ein in theoretischer, empirischer und pragmatischer Sicht unzureichendes Verfahren, das nicht den TBS-TK-Ansprüchen an psychometrische Testverfahren genügt. Es bleibt zu hoffen, dass in der von den Autoren angekündigten zweiten Auflage die angesprochenen Mängel behoben werden.

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

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

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

**T E S T I N F O R M A T I O N E N**

TBS-TK Rezension	Golden Profiler of Personality (GOP). Deutsche Adaptation des Golden Personality Type Profiler von John P. Golden	Die TBS-TK-Anforderungen sind erfüllt			
		voll	weitgehend	teilweise	nicht
Testbeurteilungssystem – Testkuratorium der Föderation deutscher Psychologinnenvereinigungen	Allgemeine Informationen, Beschreibung und diagnostische Zielsetzung			●	
	Objektivität			●	
	Zuverlässigkeit			●	
	Validität				●

Golden, J. P., Bents, R., & Blank, R. (2004). Golden Profiler of Personality (GOP). Deutsche Adaptation des Golden Personality Type Profiler von John P. Golden. Bern: Huber.

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

Manual, 135 €. Fragenheft, 9,80 €. Buch zum Test [Blank, R. & Bents, R. (2005). Typisch Mensch – Eine Einführung in die Typentheorie (3., überarbeitete Aufl.). Göttingen: Hogrefe], 15,80 €.

Starter-Kit, bestehend aus Manual, 4 Fragenheften, 20 Punkten für computergestützte Testdurchführungen/Auswertungen inkl. Ergebnisbericht und Koffer, 360 €, 20 weitere Punkte, 260 €, 100 weitere Punkte, 1160 €, in Vorbereitung: Testauswerteservice (Stand: 5.8.2008).

Bitte zitieren Sie diesen Artikel wie folgt:  
Höft, S. & Muck, P. M. (2009). TBS-TK Rezension: »Golden Profiler of Personality (GOP). Deutsche Adaptation des Golden Personality Type Profiler von John P. Golden«. Report Psychologie, 7/8-2009.