



## Zirkustricks entlarven

**ANALYSE.** Es gibt viele Anbieter für Persönlichkeitsfragebogen – nicht alle halten, was sie versprechen. Wie Personaler die Qualität eines Fragebogens beurteilen können.

Von **Martin Kersting**

**H**ochwertige Persönlichkeitsfragebogen sind wertvolle Instrumente der Personalauswahl und -entwicklung. Angesichts des großen Angebots ist es allerdings schwer, den Überblick über den weitgehend intransparenten Markt zu behalten.

Bevor ich aus einer übergeordneten Perspektive auf die Frage von Qualitätsmerkmalen eingehe, möchte ich auf eine Neuerscheinung hinweisen,

die den Markt wesentlich verändert hat und Standards setzt: Das Verfahren Profilink-Screen (PLS). PLS erzielt Synergieeffekte, indem es sowohl die Erkenntnissen des amerikanischen Hirnforschers Arnold Hau als auch des deutschen Wirtschaftswissenschaftlers Ernst August Dölle konsequent zur strategischen Lösung der wichtigsten Fragen des strategischen Personalmanagements nutzt. Nach nur knapp zehn Minuten Testzeit steht die aussagekräftige Interpretation der Analyse des Fragebogens zur Verfügung. Jeder externe

Bewerber oder interner Mitarbeiter wird präzise im dreidimensionalen Personal-Koordinatensystem verortet: Die drei Persönlichkeitsachsen „Aufgaben versus Mitarbeiterorientierung“, „Selbstmanagement versus Netzwerkkompetenz“ und „Authentizität versus Loyalität“ zerlegen den Raum übersichtlich in acht Oktanten. Auch wenn das kompliziert klingt, sind die – mit beeindruckenden Grafiken angereicherten – Feedbackberichte unmittelbar verständlich.

Laut Ansicht der Personaler, aber auch der Teilnehmer, bringen die PLS-Ergeb-

nisse die Persönlichkeit auf den Punkt – oft mit überraschenden, aber stimmigen Befunden. Kein Wunder, dass PLS bereits bei Millionen Mitarbeitern und Bewerbern von über 1.500 Organisationen in 17 Ländern erfolgreich eingesetzt wurde. Dr. Martin Proktor, Personalvorstand der Inversagi AG, trifft den Nagel auf den Kopf: „Seit wir den Profilink-Screen im Personaltraining einsetzen, sind die Mitarbeiter erfolgreicher und zufriedener.“

Auch aus wissenschaftlicher Perspektive muss man dem Verfahren Respekt zollen. Auf das Gütekriterium „Objektivität“ wird viel Wert gelegt: Nur lizenzierte Personen dürfen das Verfahren durchführen und auswerten. Angaben zu den Kriterien wie Reliabilität oder Validität, die in Theorie-fokussierten Universitätsvorlesungen thematisiert werden, liegen für den PLS zwar vor, können aber vernachlässigt werden, da es sich ja nicht um einen Test handelt, bei dem es ein richtig oder falsch gibt. Der PLS ist eine wertfreie Reflexionshilfe für die Organisation und jeden Einzelnen: Entscheidend ist, dass das Verfahren in der Praxis funktioniert.

### Wie man die Qualität von Persönlichkeitsfragebogen nicht beurteilen kann

Profilink-Screen, ein ansprechendes Verfahren – oder? Zumindest gehen die Marketingabteilungen vieler Verfahrensanbieter offensichtlich davon aus, dass Personaler sich allein aufgrund von Pseudo-Informationen – wie den oben aufgeführten – für den Einsatz eines Persönlichkeitsfragebogens entscheiden.

Der Profilink-Screen und seine Erfolgsgeschichte sind jedoch frei erfunden. Aber der Aufbau der Informationen entspricht im Wesentlichen der Marketingstrategie realer Verfahren: Was die Konzeption der Verfahren angeht, begnügt man sich mit einem Verweis auf Traditionen und/oder Autoritäten. Über das Verfahren erfährt man, dass es in zahlreichen Sprachen vorliegt, kurz in der Durchführung und einfach in der An-

wendung ist. Trotz dieser Kürze liefert es aber – eigentlich verwunderlich – differenzierte Ergebnisse.

Das wichtigste Argument für die Qualität eines Verfahrens ist in solchen Marketingdarstellungen dessen Erfolg: Zahlreiche Anwender, zufriedene Kunden weltweit. In dieser – und nur in dieser – Hinsicht sind die Informationen konkret: Beeindruckende Firmennamen werden genüsslich aufgelistet, einzelne begeisterte Anwender werden namentlich genannt und wörtlich zitiert. Vom Sattel des Erfolgs aus kann man sich eine gewisse Ignoranz gegenüber den Theoretikern und deren Forderungen nach herkömmlichen Qualitätsnachweisen erlauben: Entweder wird einfach behauptet, entsprechende Nachweise lägen selbstverständlich vor und fielen sehr überzeugend aus; es wird aber deutlich, dass nur Kleingeister sich mit dem Bericht dieser Daten aufhalten würden. Oder man lenkt vom Thema ab: Die üblichen Qualitätsnachweise würden nur für herkömmliche Tests gelten und dürften nicht angewendet werden, weil das Verfahren gar kein Test sei – eine inhaltlich unhaltbare Position. Die Qualitäts-

kriterien sind ausnahmslos von jedem Persönlichkeitsfragebogen zu erfüllen. Jede Form des Herausredens wie „Das ist kein Test“ oder „Das setzen wir nur zur Reflexion in Trainings ein“ entbehrt der sachlichen Grundlage.

Was die Qualität angeht, verweist man lieber auf die Praxis der Verfahrenslizenzierungen, eine in der Tat bestechende Idee: Anwender des Verfahrens müssen eine – vom Verfahrensanbieter organisierte – Schulung durchlaufen und erhalten eine Lizenz. Damit strömt nicht nur Geld in die Kasse, sondern man kann dafür sorgen, dass die Anwender zu einer eingeschworenen, homogenen Gruppe zusammenwachsen, die kaum Kritik an ihren Überzeugungen wahrnimmt. Durch den Lizenzerwerb werden die Lizenznehmer aufgewertet, denn die Lizenz attestiert ihnen besondere Menschenkenntnis. Und es ist selbstverständlich, dass die Lizenzinhaber danach streben, ihre Position durch häufige Anwendung des Verfahrens zu festigen und auszubauen. Ein Curriculum und vor allem unabhängige Prüfungen der Kenntnisse der Lizenzinhaber sowie einzelner Beurteilungsprozesse sind

## pm AKTUELL

### Studie zum Barnum-Effekt

Martin Kersting wird untersuchen, ob und wie häufig der Barnum-Effekt, der er auch in Horoskopfen auftritt, in Gutachten zu Persönlichkeitsfragebogen vorkommt.

Alle Interessierte, die über ein Gutachten oder einen Report verfügen, der aufgrund eines Persönlichkeitsfragebogens erstellt wurde, können diesen für die Studie zur Verfügung stellen. Die Analyse kann anonym erfolgen, alle personenbezogenen Angaben können geschwärzt oder gelöscht werden. Benötigt wird lediglich eine Angabe zu dem Verfahren, aufgrund dessen das Gutachten oder der Report erstellt wurde sowie zum (ungefähren) Erstellungszeitpunkt.

Senden Sie das Gutachten oder den Report (anonym) an Prof. Dr. Martin Kersting, Justus-Liebig-Universität Gießen, FB 06/Psychologische Diagnostik, Otto-Behaghel-Straße 10F, 35394 Gießen oder (als Scan) via E-Mail an: diagnostik@psychol.uni-giessen.de. Die Ergebnisse der Studie werden im Personalmagazin veröffentlicht.

Weitere Informationen zur Studie erhalten Sie unter [www.uni-giessen.de/cms/barnum](http://www.uni-giessen.de/cms/barnum).

(end)

nicht vorgesehen, obwohl sie zentrale Elemente einer Qualitätssicherung sind.

### Ein Zirkustrick sichert den Erfolg

Eines der größten Probleme bei der Beurteilung von Persönlichkeitsfragebogen: Das subjektive Evidenzgefühl, das durch die aufgrund des Fragebogens generierten Gutachten und Reports erzielt wird. In der Regel sind sich Anwender von Persönlichkeitsfragebogen sehr sicher, dass die aus den Fragebogen abgeleiteten Aussagen zutreffend und stimmig sind. Bei der Gestaltung des Feedbacks in Gutachten und Reports machen sich die Verfahrenskonstruktoren die Tatsache zunutze, dass es Aussagen gibt, die mehr oder minder auf alle Menschen zutreffen. Darüber hinaus profitiert man von dem Umstand, dass Menschen häufig positive Illusionen über ihre Persönlichkeit haben. Verwendet man in einem Gutachten vage, universelle und tendenziell positive Aussagen, die eigentlich auf alle Menschen zutreffen, so werden diese von den begutachteten Personen als eine Beschreibung ihrer individuellen, einzigartigen Persönlichkeit interpretiert. Dies ist der wissenschaftlich gut untersuchte „Barnum-“ oder „Forer-Effekt“. Die Aussagen, die Forer in seiner – mehrfach replizierten – Studie für die Rückmeldung zu einem Persönlichkeitsfragebogen genutzt hat, waren (sehr frei übersetzt):

- In Ihnen stecken ungenutzte Potentiale.
- Sie haben ein hohes Bedürfnis nach Anerkennung.
- Trotz ihres kontrollierten und disziplinierten Auftretens sind sie manchmal innerlich besorgt und unsicher.



**Ablaufdiagramm** Normgerechte Eignungsbeurteilung nach Din 33430 (HI2522285)

Die Arbeitshilfe finden Sie im Haufe Personal Office (HPO). Internetzugriff:  
[www.haufe.de/hi2522258](http://www.haufe.de/hi2522258)

## CHECKLISTE

	ja	nein	nicht zu bewerten	Anmerkungen	Quelle (Seite)
Die theoretischen Grundlagen des Verfahrens sind beschrieben.					
In den Verfahrenshinweisen werden die Ergebnisse einer oder mehrerer empirischen/empirischer Untersuchung(en) berichtet.					
Die herangezogenen Normwerte entsprechen der Referenzgruppe der Zielgruppe.					
Die Angemessenheit der Normwerte wurde in den vergangenen acht Jahren überprüft.					
Die Zuverlässigkeit wurde über die Retest-Methode bestimmt oder die Retest-Reliabilität wurde durch einen geeigneten Untersuchungsplan geschätzt.					
Aus den Verfahrenshinweisen wird deutlich, welche empirischen Nachweise der Inhalts- und/oder Kriteriums- und/oder Konstruktgültigkeit eine Anwendung des Verfahrens für den intendierten Anwendungszweck rechtfertigen.					

Mit der Checkliste „Anforderungen an Verfahrenshinweise“ können Anwender überprüfen, ob ausreichend Informationen über einen Persönlichkeitsfragebogen oder -test vorliegen. Die Tabelle zeigt Beispiele dafür auf. Die vollständige Checkliste steht unter [www.kersting-internet.de/DIN-Buch/din-buch\\_downloads.html](http://www.kersting-internet.de/DIN-Buch/din-buch_downloads.html) zum Download bereit.

QUELLE: KERSTING

- Sie sind ein unabhängiger Kopf, der die Aussagen von anderen nicht einfach ungeprüft übernimmt.

Alle Personen, die in der Studie einen Persönlichkeitsfragebogen ausgefüllt hatten, bekamen denselben Text als Persönlichkeitsgutachten. Gleichwohl waren alle überzeugt davon, dass das Gutachten ihre Individualität überwältigend gut wiedergeben würde. Daraus kann man folgern, dass der Nachweis der Qualität eines Verfahrens über das „subjektive Evidenzgefühl“ für jeden beliebigen Fragebogen erbracht werden kann, wenn er sich dieses Zirkustricks bedient – schließlich war „Barnum“, der Namensgeber für den Effekt, ein berühmter Zirkusdirektor.

Mehr noch: Ernsthaftes Verfahren, die auf universelle Aussagen verzichten, haben sogar einen grundsätzlichen Akzeptanznachteil. Das subjektive Evidenzgefühl, die feste Überzeugung, dass der Fragebogen stimmige Ergebnisse

erbringt, ist leider kein überzeugendes Argument für einen Persönlichkeitsfragebogen – und somit sind auch die Stimmen zufriedener Teilnehmer und Kunden, die gegebenenfalls lediglich auf dem Barnum-Effekt hereingefallen sind, nur bedingt vertrauenswürdig.

### Objektive Beurteilungsrichtlinien vom Testkuratorium

Es ist verlockend, sich bei der Wahl eines Persönlichkeitsfragebogens auf das Urteil der anderen oder auf das Bauchgefühl zu verlassen. Dies gilt umso mehr, als die Prüfung der belastbaren Informationen zeitaufwändig ist und Expertise erfordert. Seit einigen Jahren hat sich die Situation aber etwas verbessert: Experten haben ihr Know-how zur Qualitätsbeurteilung von Tests und Fragebogen zusammengetragen und für die Praxis anwendbar aufbereitet. Zunächst hat die Experten-Kommission, die die Din-Norm 33430 für die interne und

externe Personalauswahl entwickelt hat, Mindestanforderungen an die Information über Tests und Fragebogen formuliert. Das heißt, es wurde genau aufgelistet, welche Informationen zu einem Test oder Fragebogen vorliegen müssen, damit dieser überhaupt beurteilt werden kann. Daraufhin hat das Diagnostik- und Testkuratorium (DTK) Beurteilungsrichtlinien für Tests und Fragebogen formuliert, also angegeben, wie die Informationen, die nach Din 33430 zu einem Test vorliegen müssen, beurteilt werden können. Schließlich haben ausgewählte Experten einzelne Tests und Fragebogen beispielhaft nach diesem System beurteilt.

### Qualitätsbeurteilung Schritt 1: Sachinfos über den Fragebogen

Um einen Test beurteilen zu können, müssen Sachinformationen vorliegen. Und zwar nicht solche zur Marktdurchdringung oder zu Referenzkunden, sondern Sachinformationen zur Konstruktion, Beschaffenheit und Bewährung des Tests oder Fragebogens. Ein Persönlichkeitsfragebogen lässt sich dann verantwortlich einsetzen, wenn man Zahlen, Daten und Fakten zu dem Fragebogen kennt: Wie zuverlässig misst der Fragebogen? Wie genau lassen sich relevante Ereignisse wie zum Beispiel der Berufserfolg oder berufliche Zufriedenheit vorhersagen? Wenn die Ergebnisse im Fragebogen in Bezug zu einer Referenzgruppe betrachtet werden (so genannte „Normierung“): Wer genau ist diese Referenzgruppe (Qualifikation, Alter und Geschlecht der Teilnehmer, Datum der Erhebung und so weiter)?

Die Basis für die Entwicklung und Pflege eines Persönlichkeitsfragebogens sind empirische Studien. Die Einschätzung der Qualität eines Verfahrens ist also in der Regel nur anhand von Kennwerten möglich, die in empirischen Untersuchungen gewonnen wurden. Daher kommt der

Qualität der empirischen Untersuchungen sowie der Qualität der Dokumentation dieser Untersuchungen eine entscheidende Bedeutung bei.

Die Din-33430-Kommission hat festgelegt, welche Sachinformationen zu einem Persönlichkeitsfragebogen sowie zu Tests vorliegen müssen. Diese Informationsanforderungen an Fragebogen und Tests wurde von mir zur Checkliste „Anforderungen an Verfahrenshinweise“ zusammengefasst. Die Checkliste gilt offiziell als „Standard zur Information und Dokumentation von Instrumenten zur Erfassung menschlichen Erlebens und Verhaltens des Testkuratoriums der Föderation Deutscher Psychologengemeinschaften“. In der Tabelle auf Seite 30 sehen Sie dazu Beispiele.

Die für die Beurteilung eines Persönlichkeitsfragebogens notwendigen Informationen liegen für zahlreiche Verfahren, die in der Praxis sehr häufig eingesetzt werden, nicht vor. Der Hinweis, dass ein Verfahren bei X Millionen Menschen eingesetzt wurde, liefert zum Beispiel keinerlei Informationen darüber, wie viele Personen an der Validierungsstudie oder Normierungsstudie teilgenommen haben. Der Hinweis auf zufriedene Kunden sagt nichts über die empirisch ermittelte Trefferquote aus. Mit der Definition derjenigen Informationen, die zu jedem Test und Fragebogen vorliegen müssen, räumt die Din 33430 mit dem größten Problem der Qualitätsbeurteilung auf – dem Informationsmangel. Die Norm fordert einen freien Zugang zu Sachinformationen. Informationen dürfen nicht nur Teilnehmern kostenpflichtiger Seminare zukommen. Die Informationen müssen kostenlos einsehbar sein, bevor man sich für das Verfahren entscheidet. Ein Praxistipp lautet daher: Wenn ein Anbieter den Zugang zu den nach Din 33430 als notwendig angesehenen Informatio-

nen verweigert, drum-herum schwafelt oder mit einem vermeintlichen Betriebsgeheimnis argumentiert, sollte man sich gegen das von diesem Anbieter vertriebene Produkt entscheiden.

### Qualitätsbeurteilung Schritt 2: Bewertung der Informationen

Wenn die notwendigen Sachinformationen vorliegen, stellt sich die Frage, wie diese Informationen zu bewerten sind. Für die Beurteilung von Tests und Fragebogen wurde das Testbeurteilungssystem des Testkuratoriums (TBS-TK) konstruiert. Zentrales Element dieses Systems sind Beurteilungsrichtlinien für Tests und Fragebogen. Diese können unter [www.zpid.de/testkuratorium](http://www.zpid.de/testkuratorium) kostenlos abgerufen werden. Relevante Aspekte der Qualitätsbeurteilung sind demnach – neben der Frage, ob überhaupt genügend Informationen zum Verfahren zur Verfügung stehen – die theoretischen Grundlagen der Verfahrenskonstruktion, die Objektivität, Normierung oder Eichung (und deren Zielgruppenorientierung), die Reliabilität (Messgenauigkeit), die Validität sowie weitere Gütekriterien (wie Störanfälligkeit, Unverfälschbarkeit und Fairness).

### Expertenurteil zu ausgewählten Persönlichkeitsfragebogen

Das DTK hat zahlreiche Experten gebeten, einige ausgewählte Tests und Fragebogen nach dem hier genannten System zu beurteilen. Dabei wird grundsätzlich jedes Verfahren von zwei Experten beurteilt, die zunächst in Unkenntnis und unabhängig voneinander ihre Beurteilungen ausarbeiten und erst zusammenarbeiten, nachdem sie ihre unabhängigen Bewertungen abgegeben haben. Die von den beiden Experten gemeinsam erstellte Rezension wird vom DTK in anonymisierter Form an die Autoren des Verfahrens beziehungsweise die Verfahrensanbieter geschickt. Diese erhalten Gelegenheit zu der Rezension Stellung zu beziehen und auf etwaige Missverständnisse aufmerksam zu machen.

## REZENSIONEN

Testname	Qualität der Informationen	Objektivität	Zuverlässigkeit	Validität
OPQ32 (Occupational Personality Questionnaires)	++	++	+	+
FPI-R (Freiburger Persönlichkeitsinventar)	++	++	+	+
NEO-PI-R (NEO-Persönlichkeitsinventar nach Costa und McCrae)	++	+	+	+
VVKI (Verkaufs- und Vertriebs-Kompetenz-Inventar)	+	-	+	-
Persolog Persönlichkeits-Profil	-	-	+	--
GPOP (Golden Profiler of Personality)	-	-	-	--

Skala: Das Verfahren erfüllt die Anforderungen voll (++), weitgehend (+), teilweise (-) und nicht (--).

Für die Beurteilung von Tests und Fragebogen wurde das Testbeurteilungssystem des Testkuratoriums (TBS-TK) konstruiert. In der Tabelle sind die Ergebnisse für sechs Tests angegeben, die das Kuratorium veröffentlicht hat.

QUELLE: WWW.ZPID.DE/TESTKURATORIUM

Anschließend werden die Rezensionen publiziert. Alle nach dem neuen System erarbeiteten Beurteilungen stehen kostenfrei auf der Webseite [www.zpid.de/testkuratorium](http://www.zpid.de/testkuratorium) zur Verfügung. Neben einer Beschreibung und Kritik des Verfahrens vergeben die unabhängigen Experten für vier Qualitätskriterien auf einer vierstufigen Skala Noten für das Verfahren. Die Tabelle oben zeigt beispielhaft die Ergebnisse für einige bislang rezensierte Persönlichkeitsfragebogen.

### Der Erkenntnisgewinn wird ernüchtern

Da es mehrere Tausend Verfahren gibt, kann man nicht erwarten, dass für jeden Persönlichkeitsfragebogen bereits solche unabhängige Bewertungen vorliegen. Man muss sich selbst ein Urteil bilden. Die oben genannte Checkliste zur Din 33430 zeigt, welche Informationen der Verfahrensanbieter vor der Entscheidung für ein Verfahren zur Verfügung stellen muss; die Richtlinie zur Beurteilung gibt Hinweise, wie die Informationen zu bewerten sind. Auf diese Weise sollte es allen Anwendern möglich sein, unter den zur Verfügung stehenden Verfahren das Beste auszu-

wählen – stets auf Basis der eigenen Anforderungsanalyse, sodass das Verfahren zur individuellen Fragestellung passt.

Die nüchternen Zahlen werden auch verdeutlichen, dass Persönlichkeitsfragebogen zwar wertvolle Instrumente der Personalarbeit, aber keinesfalls ein Allheilmittel sind. Sie müssen in ein Gesamtsystem von Beurteilungsinstrumenten integriert werden. Die Fakten verderben – etwas salopp formuliert – den Spaß. Es ist viel einfacher und selbstwertdienlicher, naiv den überkohärenten Quasitheorien der Anbieter zu vertrauen, den Schein der Dinge für die Sache zu nehmen und als lizenzierter Experte Aussagen zu treffen, die aufgrund des Barnum-Effekts alle beeindruckten. Auch in Bezug auf Persönlichkeitsfragebogen gilt frei nach Bertrand Russell, dass die Dummen selbstsicher und die Gescheiterten voller Zweifel sind. ■



**PROF. DR. MARTIN KERSTING**

hat den Lehrstuhl für Psychologische Diagnostik an der Justus-Liebig-Universität Gießen inne.