

8.2 Qualitätsbeurteilung management-diagnostischer Verfahren und Prozesse

Begriffsbestimmung

„Qualität“, „Qualitätskontrolle“, „Qualitätssicherung“ und „Qualitätsentwicklung“ sind Ausdrücke, die uns innerhalb und außerhalb der Psychologie immer wieder begegnen. Wenn Produkte oder Prozesse zu beurteilen sind, wird in der Regel der Qualitätsbegriff bemüht. Allerdings ist die Frage, was unter „Qualität“ zu verstehen, also wie der Qualitätsbegriff angemessen zu explizieren ist, nicht in allgemeiner Form zu beantworten. Die Beurteilung der Qualität von Produkten oder Prozessen orientiert sich an bestimmten *bereichsspezifischen Kriterien* und erfolgt im Zuge eines Vergleichs der (zum Beurteilungszeitpunkt bekannten) Charakteristika der zu beurteilenden Produkte oder Prozesse mit den in den Beurteilungskriterien enthaltenen Vorgaben bzw. Anforderungen. Die logische Struktur der Qualitätsbeurteilung ist also die folgende: *x* (das Beurteilungssubjekt) *beurteilt die Qualität von y* (des Beurteilungsobjekts) *als z* (das Beurteilungsergebnis) *relativ zu k* (den Beurteilungskriterien) *zur Zeit t*.

Zur Veranschaulichung dieser Explikation der Struktur der Qualitätsbeurteilung als fünfstellige Relation mögen zwei Beispiele aus dem Bereich der psychologischen Diagnostik dienen:

1. Schmidt-Atzert und Rauch (als Beurteilungssubjekt) beurteilen die Qualität der erweiterten und überarbeiteten Fassung des I-S-T 2000 R (als Beurteilungsobjekt) in der in ihrem Beitrag zum Report Psychologie (Schmidt-Atzert & Rauch, 2008) dargestellten Art und Weise (als Beurteilungsergebnis) relativ zum Testbeurteilungssystem des Testkuratoriums (2006; als Beurteilungskriterien) im Jahr 2008 (als Beurteilungszeitpunkt).
2. Ein Sachverständiger (als Beurteilungssubjekt) beurteilt die Qualität des diagnostischen Prozesses (als Beurteilungsobjekt), den ein anderer Sachverständiger bei der Erstellung seines Gutachtens durchlaufen hat, in sehr kritischer Weise (als Beurteilungsergebnis) relativ zu den Richtlinien für den diagnostischen Prozess der European Association of Psychological Assessment (Fernández-Ballesteros et al., 2001; als Beurteilungskriterien) zum Zeitpunkt *t*.

Im ersten Beispiel geht es um die Qualitätsbeurteilung eines diagnostischen Verfahrens, im zweiten um die eines diagnostischen Prozesses. In beiden Fällen wird jeweils auf bestimmte Beurteilungskriterien Bezug genommen, die wiederum als Resultat sozialer Konstruktionen aufzufassen sind. Die logische Struktur dieser Konstruktionen ist die folgende (Westmeyer, 2009, S. 51 f.): *x* (das Konstruktionsobjekt) *konstruiert y* (das Konstruktionsobjekt) *als z* (das Konstruktionsresultat) *zur Zeit t*. Auf die beiden Beispiele bezogen bedeutet das:

- (zu 1) Das Testkuratorium (als Konstruktionsobjekt) konstruiert Beurteilungskriterien für Testverfahren (als Konstruktionsobjekt) als (in Form des) Testbeurteilungssystem(s) des Testkuratoriums (als Konstruktionsresultat) im Jahre 2006 (als Konstruktionszeitpunkt).

- (zu 2) Eine Taskforce der European Association of Psychological Assessment (EAPA, als Konstruktionsobjekt) konstruiert Beurteilungskriterien für diagnostische Prozesse (als Konstruktionsobjekt) als (in Form der) EAPA-Richtlinien für den diagnostischen Prozess (als Konstruktionsresultat) im Jahr 2001.

In beiden Fällen werden die Beurteilungskriterien von Gruppen von Personen konstruiert, die von nationalen bzw. internationalen Fachgesellschaften mit Definitionsmacht ausgestattet werden: das Testkuratorium durch die Föderation Deutscher Psychologenvereinigungen und die Taskforce durch die EAPA. Die explizite Nennung des Zeitpunkts der Konstruktion ist deshalb wichtig, weil die Konstruktionsresultate jederzeit revidiert werden können. So gibt es bereits eine geringfügig modifizierte Version des Testbeurteilungssystems des Testkuratoriums aus dem Jahr 2009.

Über die beiden in den Beispielen genannten Kriterien für die Qualitätsbeurteilung diagnostischer Verfahren und Prozesse hinaus gibt es zahlreiche weitere Standards, Richtlinien und Normen. Einen umfassenden Überblick über diese sozialen Konstruktionen geben Kersting, Häcker und Hornke (2011). Eine Auswahl für den Bereich der Management-Diagnostik besonders relevanter Konstruktionen findet sich in Tabelle 8.2.1. Auf ein Konstruktionsresultat soll hier aufgrund seiner zentralen Bedeutung gerade auch für die managementdiagnostische Praxis näher eingegangen werden: die DIN 33430 (vgl. auch 7.1.1 *Empfehlungen zur Eignungsbeurteilung nach der DIN 33430*, in diesem Band).

Tabelle 8.2.1: Einige Standards, Richtlinien und Normen für diagnostische Verfahren und Prozesse

Konstruktions-subjekt	Konstruktionsobjekt	Konstruktions-resultat	Konstruktionszeit-punkt	URL
AERA, APA, NCME	Beurteilungskriterien für die Testkonstruktion und -evaluation sowie eine fachlich kompetente Testanwendung	Standards for educational and psychological testing	1999	
Arbeitskreis Assessment-Center	Beurteilungskriterien für Assessment-Center	Standards der Assessment-Center Technik	2004	http://www.arbeitskreis-ac.de
Deutsches Institut für Normung	Beurteilungskriterien für berufsbezogene Eignungsbeurteilungen	DIN 33430	2002	
EAPA	Beurteilungskriterien für den diagnostischen Prozess	Guidelines for the assessment process	2010	
International Organization for Standardization	Beurteilungskriterien für berufsbezogene Eignungsbeurteilungen	ISO 10667-1 und ISO 10667-2	2011	

Tabelle 8.2.1: Fortsetzung

Konstruktions-subjekt	Konstruktionsobjekt	Konstruktions-resultat	Konstruktionszeit-punkt	URL
International Test Commis-sion	Beurteilungskriterien für eine fachlich kompetente Testanwendung	International guide-lines on test use	2001	http://www.intestcom.org
International Test Commis-sion	Beurteilungskriterien für computerbasiertes Testen und Testen übers Internet	International guide-lines on computer-based and internet-delivered testing	2006	http://www.intestcom.org
SIOP	Beurteilungskriterien für berufsbezogene Eignungsbeurteilungen	Principles for the validation and use of personnel selection procedures	2003	http://www.siop.org/_Principles/principles.pdf
Taskforce on AC Guidelines	Beurteilungskriterien für Assessment-Center	Guidelines and ethical considerations for assessment center operations	2009	http://www.assessmentcenters.org
Testkuratorium	Beurteilungskriterien für Tests	Testbeurteilungs-system des Testkura-toriums (TBS-TK)	2009	http://www.zpid.de/Testkuratorium

Die DIN 33430

Die DIN 33430 (DIN, 2002) formuliert „Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen“. Sie wurde in einem streng formalisierten Prozess entwickelt, der sicherstellt, dass die Forderungen keine Einzelmeinungen wiedergeben, sondern den Konsens über den Stand von Wissenschaft und Technik. Bei der sozialen Konstruktion der Norm waren alle interessierten Kreise teilnahmeberechtigt. Obwohl die Beachtung von DIN-Normen freiwillig erfolgt, kann eine DIN – aufgrund der gesicherten kollektiv-partizipativen Entwicklung – prozessrechtliche Bedeutung entfalten (siehe Kersting, 2008).

Kersting (2008) hat die normativen (und einige nicht-normative) Forderungen der DIN 33430 in 318 Prüfaussagen zerlegt und thematisch geordnet. Demnach formuliert die DIN 33430 zu den folgenden sechs Bereichen Soll-Werte:

1. *Anforderungen an Verfahrenshinweise (140 Aussagen)*. Gefordert werden u. a. umfassende, zugängliche, wahrheitsgetreue, empirisch abgesicherte und nachvollziehbare Informationen zu den Verfahren. Beispielfhaft wird für Verfahren, die Merkmale erfassen sollen, für die eine zumindest relative Zeit- und Situationsstabilität angenommen wird, verlangt, dass die Zuverlässigkeit über die

Retest-Methode bestimmt wird. Das oben angesprochene Testbeurteilungssystem des Testkuratoriums (2009) greift auf diese 140 Aussagen der DIN 33430 zurück. Die ersten Testrezensionen nach diesem System können unter der URL www.zpid.de/Testkuratorium eingesehen werden. Letztendlich handelt es sich bei den 140 Anforderungen um einen Standard dafür, wie viele (und welche) Informationen gegeben sein müssen, damit ein Test (egal aus welchem Anwendungsbereich) seriös beurteilt werden kann. Die Checkliste (Download: www.Kersting-internet.de/DIN-Buch) mit diesen 140 Anforderungen gilt offiziell als „Standard zur Information und Dokumentation von Instrumenten zur Erfassung menschlichen Erlebens und Verhaltens des Testkuratoriums der Föderation Deutscher Psychologenvereinigungen“.

2. *Planung von berufsbezogenen Eignungsbeurteilungen (33 Aussagen)*. Bei diesen Forderungen geht es u. a. um die Auftragsklärung, um die Klärung der Verantwortlichkeiten und um die – vor dem Verfahren festzulegenden – Regeln zur Durchführung, Auswertung und Interpretation der Verfahrensergebnisse sowie Urteilsbildung und Ergebnisrückmeldung. Ein Beispiel aus diesem Bereich ist die Forderung, dass eine Anforderungsanalyse durchgeführt wird und dass jedem der eingesetzten Verfahren mindestens eine Anforderungsdimension zugeordnet wird, zu dessen Erfassung es eingesetzt wird.
3. *Auswahl und Zusammenstellung von Verfahren (20 Aussagen)*. Die Auswahl und Zusammenstellung von Verfahren der Management-Diagnostik sollte sich der DIN 33430 zufolge u. a. an deren Objektivität, Zuverlässigkeit und Gültigkeit orientieren. Sofern ein Verfahren eingesetzt wird, das einen Vergleich mit Normwerten vorsieht, müssen die herangezogenen Normwerte der Fragestellung und der Referenzgruppe der Zielgruppe entsprechen.
4. *Durchführung, Auswertung und Interpretation der Verfahrensergebnisse sowie Urteilsbildung (55 Aussagen)*. Bezüglich der Durchführung wird z. B. gefordert, dass die Anweisungen an die Kandidaten so weit wie möglich standardisiert sind. Eine überaus praxisrelevante Qualitätsforderung der DIN 33430 betrifft die Interpretation in solchen Fällen, in denen mehrere Personen den Ausprägungsgrad von Personmerkmalen aufgrund von mündlich gewonnenen Informationen bzw. Verhaltensbeobachtungen beurteilen – dies ist die Standardsituation im Assessment-Center, wenn z. B. die in einer Übung gezeigte Teamfähigkeit von mehreren Assessoren beobachtet und beurteilt wird. Die DIN fordert, dass in diesen Fällen der Grad der Übereinstimmung zwischen den beurteilenden Personen berechnet wird.
5. *Dokumentation (22 Aussagen)*. Qualität erfordert Dokumentationsarbeit. Laut DIN 33430 müssen z. B. die Ergebnisse der Anforderungsanalyse ebenso dokumentiert werden wie die Regeln zur Zusammenfügung aller über einen Kandidaten erhobenen Informationen zu einem Eignungsurteil.
6. *Qualifikation der beteiligten Personen (48 Aussagen)*. Die DIN 33430 stellt einerseits Anforderungen an Verfahren, andererseits aber auch Anforderungen an die Qualifikation der an der Eignungsbeurteilung beteiligten Personen. Dabei werden die geforderten Kenntnisse und Erfahrungen für drei unter-

schiedliche Qualifikationsniveaus formuliert. Der (organisationsinterne oder -externe) Auftragnehmer (A) trägt die operative Verantwortung für die gesamte Eignungsbeurteilung, er verantwortet, plant und gestaltet die Eignungsbeurteilung. Mitwirkende sind an der Eignungsbeurteilung beteiligt, z. B. als Mitwirkende an Verhaltensbeobachtungen (MV) und/oder als Mitwirkende an Eignungsinterviews (ME). Einen Überblick über die jeweiligen Anforderungen gibt Kersting (2008). Allerdings sieht die DIN 33430 keine Regelungen vor, wie die notwendigen Qualifikationen erworben und nachgewiesen werden können. Als Ergänzung zur DIN 33430 hat die Föderation Deutscher Psychologenvereinigungen 2004 (revidiert 2011) eine „Fortbildungs- und Prüfungsordnung zur Personenlizenzierung für berufsbezogene Eignungsbeurteilungen nach DIN 33430“ erlassen (<http://www.psychologenakademie.de/download.php?name=ordnungdin>) sowie ein Angebot an Fortbildungen und Lizenzprüfungen erarbeitet.

Zu jedem dieser sechs Themenbereiche gibt es eine Checkliste (die sechs Checklisten sind zur „DIN SCREEN V2“, Kersting, 2008, zusammengefasst), wobei jede Aussage anhand einer zweistufigen Skala mit den beiden Ausprägungen „ja“ und „nein“ zu beantworten ist. Während die Forderungen der DIN 33430 den Soll-Wert darstellen, ermöglicht die Form der Checkliste die Erfassung des Ist-Wertes. Die bearbeiteten Checklisten zeigen auf, ob die eingesetzten Verfahren und Personalbeurteilungen dem aktuellen DIN 33430 Qualitätsstandard genügen. Mögliche Qualitätsmängel können identifiziert werden, sodass in der Folge die Ursachen mangelnder Qualität beseitigt werden können.

Um nachzuweisen, dass die von einer Person oder Organisation durchgeführte Eignungsdiagnostik den Qualitätsforderungen der DIN 33430 (Soll-Wert) entspricht (Konformität), kann man eine Selbsterklärung abgeben: „*Ich/Wir arbeite(n) gemäß DIN 33430*“. Des Weiteren gibt es seit 2004 die Möglichkeit, durch eine gebührenpflichtige schriftliche Prüfung eine Personenlizenz für die berufsbezogene Eignungsbeurteilung nach DIN 33430 zu erwerben und damit einen Qualifikationsnachweis zu erbringen (näheres zu den Personenlizenzen und der Prüfung siehe Kersting, 2010). Organisationen können durch Zertifikate nachweisen, dass sie den Anforderungen der DIN 33430 gerecht werden. Mit dem Zertifikat wird die Fähigkeit einer Organisation zur Realisierung einer definierten Qualität bestätigt. Die Zertifizierung kann im Rahmen des nach ISO 9000 ff. ausgerichteten Qualitätsmanagementsystems erfolgen. Die spezifischen Forderungen der DIN 33430 bauen dann auf dem Fundament der ISO 9000 ff. auf.

Durch verschiedene Initiativen wurde die Idee der DIN 33430 internationalisiert. Besonders hervorzuheben ist die in 2006 begonnene Entwicklung der ISO Norm 10667 (Assessment Service Delivery: Procedures and methods to assess people in work and organizational settings, ISO, 2011), bei der u. a. die DIN 33430 zu Beratungszwecken herangezogen wurde. Die ISO 10667 ist Ende 2010 verabschiedet worden und 2011 in Kraft getreten.

Soziale versus persönliche Konstruktionen von Qualitätskriterien

Alle Standards, Richtlinien und Normen, wie z. B. die in Tabelle 8.2.1 aufgeführten, sind das Resultat sozialer Konstruktionen. Es ist immer eine mit Definitionsmacht ausgestattete Gruppe von Personen, die im Auftrag einer Fachgesellschaft oder einer anderen gesellschaftlichen Institution oder Organisation die Qualitätskriterien formuliert, festgelegt, also konstruiert hat. Aber auch wenn diese sozialen Konstruktionen in der Regel den Stand der Wissenschaft und Technik weitestgehend berücksichtigen, stoßen sie in der Praxis nicht selten auf erhebliche Vorbehalte. Viele berufsbezogene Eignungsuntersuchungen in der diagnostischen Praxis folgen nicht den strengen Anweisungen der DIN 33430. Die Parallele zur ungebrochenen Beliebtheit klinischer Urteilsbildung im Unterschied zur weitverbreiteten Zurückhaltung gegenüber der statistischen Urteilsbildung trotz allem, was über die Mängel klinischer Urteilsbildung bekannt ist (siehe dazu den Beitrag 7.1.3 *Klinische Urteilsbildung*, in diesem Band), liegt auf der Hand. Wie ist die Zurückhaltung bei der Anwendung der genannten sozialen Konstruktionen zu erklären? Sind die Ursachen in der Unkenntnis neuerer Entwicklungen, in der Schwierigkeit, liebgewordene Diagnostizierungsgewohnheiten zu verändern oder gar in mangelndem Interesse an qualitätsvollem praktischen Handeln zu sehen? Das ist keineswegs ausgemacht. Qualitätskriterien können und werden auch persönlich konstruiert: Ein Praktiker (als Konstruktionssubjekt) konstruiert Kriterien für die Beurteilung diagnostischer Verfahren bzw. Prozesse (als Konstruktionsobjekt) in einer bestimmten persönlichen Art und Weise (als Konstruktionsresultat) zu einem bestimmten Zeitpunkt und beurteilt dann (als Beurteilungssubjekt) die Qualität eines diagnostischen Verfahrens bzw. Prozesses (als Beurteilungsobjekts) in einer bestimmten Art und Weise (als Beurteilungsergebnis) relativ zu seinen eigenen Beurteilungsmaßstäben (als Beurteilungskriterien) wiederum zu einem bestimmten Zeitpunkt. Ebenso können einzelne Gruppen (z. B. eine Personalabteilung, eine Personalberatung) eigene Maßstäbe etablieren.

Diese persönlichen Konstruktionen müssen nicht willkürlich oder beliebig sein und sind es in der Regel auch nicht. Sie sind vielmehr orientiert am Kriterium der *Viabilität*, der Lebbarkeit, im Sinne von Glasersfeld (1996). Eine Diagnose und/oder Prognose, die der Praktiker stellt, ist/sind z. B. dann viabel, wenn der Praktiker auf der Grundlage dieser Diagnose und/oder Prognose im gegebenen Fall effektiv intervenieren kann. Die Effektivität seiner Interventionen hat er dann allerdings im Rahmen einer *kontrollierten Praxis* (siehe dazu 8.7 *Grundlagen kontrollierter Praxis*, in diesem Band) zu belegen. Wenn es nicht um die Vorbereitung von Interventionen geht, kann Viabilität auch in der Akzeptanz der gestellten Diagnose und/oder Prognose durch den Auftraggeber zum Ausdruck kommen. Kundenzufriedenheit ist sicher ein wichtiger Faktor, der mit darüber entscheidet, welche Praktiken sich in der Praxis durchsetzen und erhalten. Dieser Umstand mag vor dem Hintergrund einer realistischen Wissenschaftsauffassung schwer zu akzeptieren sein, vor dem Hintergrund einer konstruktivistischen Wissenschaftsauffassung (z. B. Gergen,

2002; siehe auch Westmeyer, 2011) stellt er keine Überraschung dar. Aus dieser Sicht ist die Viabilität wenigstens ebenso wichtig wie die Validität. Nur wenn die möglichst genaue Vorhersage eines Außenkriteriums (z. B. Ausbildungserfolg oder Berufserfolg) das mit dem Auftraggeber vereinbarte Ziel der diagnostischen Bemühungen ist, fallen Validität und Viabilität zusammen.

Martin Kersting und Hans Westmeyer

Literatur

- American Educational Research Association (AERA), American Psychological Association (APA) & National Council on Measurement in Education (NCME) (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Arbeitskreis Assessment Center (2004). *Standards der Assessment Center Technik*. Hamburg: Arbeitskreis Assessment Center e. V.
- DIN (2002). *DIN 33430: Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen*. Berlin: Beuth.
- Fernández-Ballesteros, R., De Bruyn, E. E. J., Godoy, A., Hornke, L. F., Ter Laak, J., Vizcarro, C., Westhoff, K., Westmeyer, H. & Zaccagnini, J. L. (2001). Guidelines for the assessment process (GAP): A proposal for discussion. *European Journal of Psychological Assessment*, 17, 187–200. In deutscher Übersetzung: Westhoff, K., Hornke, L. F. & Westmeyer, H. (2003). Richtlinien für den diagnostischen Prozess – Zur Diskussion gestellt. *Report Psychologie*, 28, 504–517.
- Gergen, K. J. (2002). *Konstruierte Wirklichkeiten: Eine Hinführung zum Sozialen Konstruktivismus*. Stuttgart: Kohlhammer.
- Glasersfeld, E. von (1996). *Radikaler Konstruktivismus*. Frankfurt a. M.: Suhrkamp.
- International Test Commission (ITC) (2001). International guidelines on test use. *International Journal of Testing*, 1, 95–114.
- International Test Commission (ITC) (2006). International guidelines on computer-based and internet-delivered testing. *International Journal of Testing*, 6, 143–171.
- ISO (2011). *ISO 10667. Assessment Service Delivery: Procedures and methods to assess people in work and organizational settings*. Geneva, Switzerland: International Organization for Standardization.
- Kersting, M. (2008). *Qualität in der Diagnostik und Personalauswahl: Der DIN Ansatz*. Göttingen: Hogrefe.
- Kersting, M. (2010). Diagnostische Fortbildung am Beispiel des Trainings zur Eignungsbeurteilung nach DIN 33430. In U. P. Kanning, L. von Rosenstiel & H. Schuler (Hrsg.), *Jenseits des Elfenbeinturms. Psychologie als nützliche Wissenschaft* (S. 223–240). Göttingen: Vandenhoeck & Ruprecht.
- Kersting, M., Häcker, H. & Hornke, L. F. (2011). Qualitätsstandards in der Diagnostik. In M. Amelang, L. F. Hornke & M. Kersting (Hrsg.), *Enzyklopädie der Psychologie. Themenbereich B, Methodologie und Methoden. Serie II, Psychologische Diagnostik. Band 1, Grundlagen psychologischer Diagnostik* (S. 1–86). Göttingen: Hogrefe.
- Schmidt-Atzert, L. & Rauch, W. (2008). TBS-TK Rezension: Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R). 2., erweiterte und überarbeitete Auflage. *Report Psychologie*, 33, 303–304.
- Taskforce on AC Guidelines (2009). Guidelines and ethical considerations for assessment center operations. *International Journal of Selection and Assessment*, 17, 243–254.
- Testkuratorium der Föderation Deutscher Psychologenvereinigungen (2006). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologenvereinigungen. *Report Psychologie*, 31, 492–499.
- Testkuratorium der Föderation Deutscher Psychologenvereinigungen (2009). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologenvereinigungen. Revidierte Fassung vom 09. September 2009. *Report Psychologie*, 34, 470–478.
- Westmeyer, H. (2009). Wissenschaftstheoretische Grundlagen der Evaluationsforschung. In H. Holling (Hrsg.), *Grundlagen und statistische Methoden der Evaluationsforschung. Band 1 der Serie Evaluation der Enzyklopädie der Psychologie* (S. 35–58). Göttingen: Hogrefe.
- Westmeyer, H. (2011). Communicamus ergo sum oder Am Anfang stehen die Beziehungen: Hans Westmeyer über Kenneth Gergens Konstruierte Wirklichkeiten. In B. Pörksen (Hrsg.), *Schlüsselformen des Konstruktivismus* (S. 411–424). Wiesbaden: VS Verlag für Sozialwissenschaften.

14:

Management- Diagnostik

herausgegeben von
Werner Sarges

2013

4., vollständig überarbeitete und erweiterte Auflage

HOGREFE



GÖTTINGEN · BERN · WIEN · PARIS · OXFORD
PRAG · TORONTO · BOSTON · AMSTERDAM
KOPENHAGEN · STOCKHOLM · FLORENZ