

Kriteriumsorientierte Diagnostik *via* Prototypen statt normorientierter Diagnostik

Klaus D. Kubinger

Abstract: Ausgehend von einem exemplarischen Bericht über die Schwierigkeiten, heutzutage brauchbare Stichproben für die Kalibrierung und Eichung eines psychologischen Tests erheben zu können, wird grundsätzlich in Frage gestellt, ob für viele Fragestellungen eine normorientierte Diagnostik überhaupt notwendig ist und nicht eine kriteriumsorientierte genügt oder gar vorzuziehen wäre. Als Kriterium wird sowohl ein theoretisch als auch ein empirisch abgeleiteter Prototyp vorgeschlagen. Insbesondere zu ersterem wird am Beispiel der Intelligenzdiagnostik, an Hand des *Adaptiven Intelligenz Diagnostikums* AID 2 (Version 2.2; Kubinger, 2009a), illustriert, wie ein entsprechender Zugang funktionieren könnte. Am Beispiel eines Self-Assessments wird aber auch angedeutet, wie ein empirischer Prototyp gewonnen werden könnte und mit einer darauf bezogenen kriteriumsorientierten Diagnostik die methodischen Probleme von Validierungen in diesem Zusammenhang in den Griff zu bekommen sind.

Einleitung

In der Geschichte der Psychometrie gibt es seit langem die Diskussion: Diagnosen mittels Quantifizierung „entlang“ (latenter, multipel eindimensionaler) Eigenschaften (Traits) *vs.* Diagnosen mittels Qualifizierung in Klassen/Typen von Eigenschaften bzw. Eigenschaftsbündeln. Damit in Zusammenhang steht oft die Diskussion: Diagnosen orientiert an der Norm (also den „Anderen“) *vs.* Diagnosen orientiert an einem (relevanten) Kriterium. Innerhalb der *Psychologischen Diagnostik* ist ein bekannter Vertreter der multiplen (eindimensionalen) Eigenschaftsmessung zum Beispiel das *Adaptive Intelligenz Diagnostikum* AID 2 (Version 2.2; Kubinger, 2009a); es misst auch normorientiert und bietet an, über ein Testprofil, also eine untertestweise Zusammenstellung der (an der Referenzpopulation) geeichten Testwerte, eine individuelle Stärken-Schwächendiagnose zu geben. Der aktuell bekannteste Vertreter des Typenansatzes ist wohl der *Golden Profiler of Personality* GPOP (Deutsche Adaption; Bents & Blank, 2004), der im Wesentlichen auf *C. G. Jungs* Typologie des Menschen aufbaut; sein Vorläufer, der MBTI (*Myers-Briggs Type Indicator*), war dabei noch völlig „normorientierungslos“.

Im vorliegenden Beitrag geht es nun nicht darum, theoretisch inhaltliche oder gar methodische Argumente pro und contra den jeweils diametral zu einander zu stehen scheinenden Ansätzen zu bringen; vielmehr geht es aus der Sicht heute praktisch gegebener Rahmenbedingungen, einen Prototypen-Zugang in der *Psychologischen Diagnostik* auf den Weg zu bringen.

Diese heute praktisch gegebenen Rahmenbedingungen seien mit einem Erfahrungsbericht exemplarisch beschrieben. Er bezieht sich auf die verschiedenen Phasen der Entwicklung des bereits genannten *Adaptiven Intelligenz Diagnostikums*, und zwar von der ursprünglichen Fassung des AID (Kubinger & Wurst, 1985), über dessen zweite Generation in erster Fassung (AID 2 – Version 2.1; Kubinger & Wurst, 2000) und zweiter Fassung (AID 2 – Version 2.2; s. o.), bis zu der Fassung AID 2-Türkisch (Kubinger, 2009a) für in deutschsprachigen Ländern lebende Kinder und Jugendliche mit Türkisch als Muttersprache sowie dem AID 2-English (Kubinger, in Vorb.); schließlich auch auf seine dritte Generation, den AID 3 (Kubinger & Holocher-Ertl, in Vorb.). In dem auf mehr als 25 Jahre sich erstreckenden Zeitraum ist bei der Erhebung von Kalibrierungs- und Eichstichproben ein gravierender „Kultur“-Wandel zu bemerken. Kam es bei der entsprechenden Datenerhebung anfangs, 1982-1983, noch zu einer Teilnahmebereitschaft von etwa 85-90%, so lag diese 2006-2007 nur mehr bei etwa 45-

50%. Genauere „Verlaufsdaten“ finden sich in Tabelle 1. Sie zeigt, dass es dabei nicht nur um eine drastische Abnahme der Teilnahmebereitschaft, sondern vor allem um mangelnde Unterstützungsbereitschaft seitens Dritter geht.

Tabelle 1: Übersicht über die Datenerhebung und die Rahmenbedingungen zu den Testentwicklung zum *Adaptiven Intelligenz Diagnostikum* im Laufe der Zeit

Version	Datenerhebung	Erhebung	Rahmenbedingung	Teilnahmebereitschaft (der konkret angesprochenen Schüler)
AID	1982-1983	Klumpenstichprobenerhebung in jeweils 5 Bundesländern in D und A: Schulen (geschichtet nach Schultyp); etwa 2100 Schüler	Kontaktaufnahme mit Schulleitern; Organisation eines Testraums für Einzeltestung (Schüler werden aus dem laufenden Unterricht für 1-2 Schulstunden laut Plan herausgeholt) Freiwilligkeit via Elternbrief und Elternzustimmung;	etwa 85-90%
AID 2 Version 2.1	1995-1997	Klumpenstichprobenerhebung in jeweils 5 Bundesländern in D und A: Schulen (geschichtet nach Schultyp); etwa 1000 Schüler	mit Unterstützung der Landesschulbehörde; einige Absagen von Schulen machte Nachnominierung von Schulen notwendig, Erhebungszeitraum wesentlich vom Plan abweichend	etwa 70-75%
AID 2 Version 2.2 samt AID 2- Türkisch	2006-2007	Klumpenstichprobenerhebung in jeweils 5 Bundesländern in D und A: Schulen (geschichtet nach Schultyp); etwa 800 Schüler plus etwa 350 Schüler mit Türkisch als Muttersprache	mit Unterstützung der Landesschulbehörde; viele Absagen von Schulen ; Zeiträumen zwischen grundsätzlicher Zusage durch Schulleiter und Zustimmung oder Ablehnung (ohne Angabe von Gründen) des Elternforums 5-9 Monate, Erhebungszeitraum wesentlich vom Plan abweichend: 1 Jahr verloren	trotz wiederholten Nachfragens etwa 45-50%
AID 2- English	2008-?	Klumpenstichprobenerhebung in D und A: Schulen mit Unterrichtssprache Englisch; in GB. Geplant 400 Schüler	ohne Unterstützung der Landesschulbehörde; fast nur Absagen von Schulen ; Dauer bis zur allfälligen Zustimmung mindestens 6 Monate, Erhebungszeitraum wesentlich vom Plan abweichend: 1½ Jahr verloren	trotz wiederholter Nachfrage, geschätzt: etwa 30-35%
AID 3	2009-?	Klumpenstichprobenerhebung in jeweils 5 Bundesländern in D und A: Schulen (geschichtet nach Schultyp); etwa 1000 Schüler	mit eingeschränkter Unterstützung der Landesschulbehörde: Großteils untragbare Einschränkung auf 1 Unterrichtseinheit. Viele Absagen von Schulen ; Erhebungszeitraum derzeit nicht abschätzbar	Neu: Rückmeldung über Testergebnisse; Teilnahmebereitschaft dzt. nicht abschätzbar

Die Gründe für die Schwierigkeiten, an entsprechende Stichproben zu kommen, sind vielschichtig und können teilweise nur vermutet werden:

- Es beginnt mit der Schulverwaltung, die trotz Unterstützung von Schulaufsicht und teilweise auch von der Schulpsychologie eine Teilnahme von sich aus verweigert; eine Überlastung der Schule mit Untersuchungen im laufenden Schuljahr, allem voran mit *Large Scale Assessments* (z. B. PISA-Studie), die Unmöglichkeit, Räume zur Verfügung stellen zu können und die befürchtete Störung des Unterrichts oder die zu lange Testdauer werden als Gründe für eine Absage genannt.
- Liegt doch einmal die Zustimmung der Schule vor, sagen häufig die Lehrer mit eben denselben Gründen ab.
- Teilen die Lehrer aber wirklich einmal die Formulare für die Einverständniserklärung der Eltern an die Schüler aus, schwankt die positive Rücklaufquote stark. Ob dabei die Lehrer abschlägige Elternerklärungen gar nicht einsammeln und es somit sein könnte, dass etliche Formulare gar nicht vom Schüler zu den Eltern (und wieder zurück bis zum Lehrer) gelangt sind, oder ob die Lehrer die Schüler zu wenig zum Mitmachen motiviert haben oder die Eltern (oder gar schon die Kinder selbst) sich grundsätzlich allem gegenüber verwehren, was nicht Pflicht ist, bleibt offen.

Abgesehen von der Unbilligkeit dieser heute praktisch gegebenen Rahmenbedingungen für den Forscher bzw. Testentwickler, ist die Problematik dieser Situation wissenschaftlich verheerend. Was letztlich an Daten erhoben wird, kann nicht als repräsentativ für die angepeilte Referenzpopulation gelten. Die sich ergebende „Stichprobe“ – die Mathematische Statistik impliziert mit Stichprobe immer eine Zufallsstichprobe, was offensichtlich hier nicht gegeben ist – besteht eigentlich aus Volunteers im Sinn von „*testophil*“, also Personen mit eher hoher Leistungsbereitschaft, so dass die Generalisierbarkeit der Ergebnisse stark bezweifelt werden muss: Wenn eher nur leistungsmotivierte Schüler der Eichung zugrunde liegen, dann ist es recht wahrscheinlich, dass indizierte Kinder in der psychologischen Fallbehandlung in Bezug auf ihr Leistungsvermögen zu schlecht beurteilt werden.

Das vorläufige Fazit lautet daher: Brauchbare Daten für Kalibrierung und Eichung psychologischer Tests sind heutzutage kaum mehr zu bekommen. Die Zukunft verspricht keine Besserung dieser Situation.

Ein alternativer Zugang

Warum also nicht „aus der Not eine Tugend machen“? Vielfach interessiert eine normorientierte Diagnostik ohnehin nicht. Zum Beispiel geht es bei der Auswahl einer Schreibkraft nicht darum, wie viel Prozent der Referenzpopulation noch besser/nach schlechter in Bezug auf ihre Orthografie sind; so interessiert etwa ein festgestellter Prozentrang von PR = 95 nicht wirklich. Vielmehr interessiert ein absoluter Maßstab, nämlich ob der Stellenbewerber eben keine orthografischen Fehler macht (etwa bei einem Standardtext mit 500 Wörtern). Es geht also um das Kriterium: kein Fehler! Eine kriteriumsorientierte Diagnostik ist also hier gefragt. Da es sich allerdings um eine Bestenauswahl handelt, läuft es ziemlich auf dasselbe hinaus, ob man norm- oder kriteriumsorientiert misst (vgl. dazu das bekannte „Sekretärinnenproblem“ in der Statistik).

Im Zusammenhang mit einem Intelligenztest ist jedoch ein kriteriums- statt eines normorientierten Zugangs „revolutionär“. So braucht man zur Aufnahme in den Club Mensa einen IQ ≥ 130 , das entspricht einem Prozentrang PR $\geq 97,5$. So wurden früher Kinder mit IQ < 85 (PR $< 16,5$) in Sonderschulen (für Lernbehinderte) verweisen. Und so gab es auch lange die „Stimmung“, fürs Gymnasium benötige man einen IQ ≥ 120 (PR ≥ 92). Intelligenz ist also klassisch an der Referenzpopulation orientiert.

Dazu kommt, dass „Intelligenz“ außerhalb der Psychologie begrifflich schwer fassbar ist und auch innerhalb der Psychologie nicht eindeutig definiert ist. Was also genau sollte das relevante Kriterium sein? Eher einvernehmlich wird dagegen der Zweck von Intelligenztests gesehen: Sie dienen im Aus- und Weiterbildungszusammenhang einer umfassenden kognitiven Stärken- und Schwächendiagnose, und zwar oft genug förderungsorientiert: Wo und wie sind welche Förderprogramme einzusetzen, um die Schwächen möglichst effizient zu egalisieren? Nimmt man dann die spezielle Definition von Kubinger (2009a, S. 23), so kommt man dem modernen „Kompetenz“-Konzept der pädagogischen Psychologie nahe: „*Intelligenz ist die Gesamtheit aller kognitiven Voraussetzungen, die notwendig sind, um Wissen zu erwerben und Handlungskompetenzen zu entwickeln*“ – wobei mit „Kognition“ Wahrnehmung, Erkennen, Vorstellen, Urteilen, Gedächtnis, Lernen, Denken, und Sprache gemeint ist. Und damit nähern wir uns einem kriteriumsorientierten Messen an.

Das gilt insbesondere im Zusammenhang mit schulpsychologischen Begutachtungen (Beratung bei Schulproblemen; Schul- und Laufbahnberatung) bzw. mit bildungspolitischen Fragestellungen. Für bestimmte Inhalte, also für in Form von Testitems abgeprüfte Kompetenzen, mag gut begründbar zu fordern sein: „Dieses Item sollten Schüler (z. B. der 8. Schulstufe) lösen können“. Damit ist jeweils ein itembezogenes Kriterium gegeben. Normorientiertes Diagnostizieren wäre dagegen verfehlt, noch mehr als im Beispiel der Bestenauswahl in Bezug auf orthografische Kenntnisse. Es interessiert nämlich weniger, wie viel Prozent der Referenzpopulation im Test besser abschneiden – selbst ein Prozentrang

von PR = 95 wäre kein Garant dafür, dass das relevante Kriterium erreicht wird; vielmehr interessiert, wie nahe die Testleistung dem angestrebten Kriterium kommt.

Ein sehr anschauliches, in letzter Zeit fast schon wieder in Vergessenheit geratenes Beispiel dessen, worum es eigentlich geht, sind die sog. Lerntests von Guthke (1977). Sie messen, welchen Nutzen ein Kind aus spezifischen Hilfestellungen zieht bzw. welche Hilfestellungen es eben braucht, um doch ein bestimmtes Kriterium an Testleistung zu erzielen. Festgestellte Defizite im Vergleich zu anderen vermögen dies nicht zu leisten.

Im Folgenden sei nun am Beispiel des AID 2 ausgearbeitet, wie genau eine kriteriumsorientierte Intelligenzdiagnostik anstatt einer normorientierten aussehen könnte. Betrachtet wird ein typischer Untertest, der einen besonderen Aspekt „Logisch-/Schlussfolgernden Denkens“ erfassen soll, nämlich die Fähigkeit, durch Abstraktion zu einer Begriffsbildung zu gelangen. Die Aufgabenstellung lautet: „Nun werde ich dir immer zwei Dinge (Begriffe) sagen und du sollst herausfinden, was sie gemeinsam haben oder was man mit beiden tun kann. Beginnen wir: Was ist das Gemeinsame an ...“. Abbildung 1 gibt die ersten fünf für 6-Jährige vorgesehene Items wieder, sowie diejenigen fünf nächsten, die es dann erhält, wenn es bei den ersten fünf so abschneidet, wie es kriteriumsbezogen hiermit gefordert wird. Alle zehn Items sind in der Abbildung auch kommentiert, ob ein 6-jähriges Kind zur Lösung kommen sollte, um ihm die minimalen *kognitiven Voraussetzungen für Wissenserwerb und Handlungskompetenz* zu attestieren. – Das Konzept des AID 2 sieht u. A. für den betreffenden Untertest vor, dass adaptiv getestet wird; d. h., je nach Leistungsverhalten der Testperson bei einer altersmäßig festgesetzten Aufgabengruppe von fünf Items wird eine andere nächste – und in der Folge auch noch eine dritte – Aufgabengruppe mit fünf Items vorgegeben.

6-7 Jahre	1	Korb – Tasche <i>tragen; hinein geben</i>	<i>ja</i>	
	2	Herdplatte – Mikrowelle <i>erhitzen; erwärmen</i>	<i>ja</i>	
	3	3	Gold – Geld <i>Wert; Kaufkraft</i>	<i>alle 6-Jährige sollten für beides wissen, was das ist; aber nur einige werden den gesuchten Oberbegriff dessen, was funktional beiden gemeinsam ist, erklären können</i>
		4	Buch – Computer <i>enthält Text; lesen</i>	<i>ja</i>
		5	Straßenwalze – Bügeleisen <i>glätten</i>	<i>zugestanden werden muss etlichen 6-Jährigen, dass sie nicht wissen, was eine Straßenwalze ist; andere werden (noch) nicht die Funktion von „Walze“ begreifen</i>
10	1	Glocke – Wecker <i>Verständigung (Signal); läuten; klingeln</i>	<i>ja</i>	
	2	Schokolade – Torte <i>Essen; sind süß; Nascherei</i>	<i>ja</i>	
	3	Schreibmaschine (PC) – Klavier <i>Tasten</i>	<i>es mag für 6-Jährige noch zu schwierig sein, auf dem entsprechend geforderten Abstraktionsniveau des Funktionalen und nicht nur produktorientiert zu denken</i>	
	4	Baby – Fohlen	<i>zugestanden werden muss etlichen 6-Jährigen, dass sie nicht wissen, was ein Fohlen ist</i>	

junges Lebewesen; neugeborenes

...

5 **Kanne – Flasche**

Flüssigkeitsbehälter; einfüllen

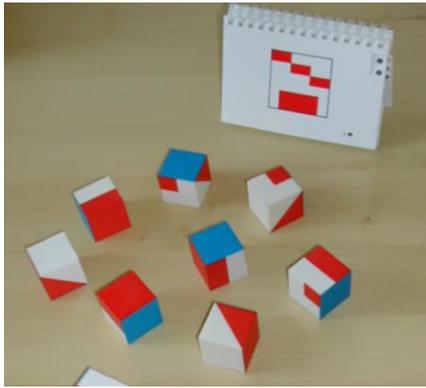
zugestanden werden muss etlichen 6-Jährigen, dass sie (heutzutage) nicht wissen, was eine Kanne ist

Abbildung 1: Zehn exemplarische Items aus dem AID 2 (Kubinger, 2009a) samt Kommentar, ob ein 6-jähriges Kind zur Lösung kommen sollte, um ihm die minimalen *kognitiven Voraussetzungen für Wissenserwerb und Handlungskompetenz* zu attestieren.

Das in Abbildung 1 (minimal) „geforderte“ Lösungsverhalten stellt nun also einen Prototyp dar; nämlich den Prototyp eines „6-Jährigen“ in Bezug auf „Logisch-/Schlussfolgerndes Denken“ (im verbalen Bereich). Dieser Prototyp kann als Kriterium dienen. Alle je getesteten 6-Jährigen können in ihrem abweichenden Verhalten dazu beschrieben und auch quantifiziert werden; zum Beispiel als der Anzahl von Items, die das betreffende Kind vom Prototyp abweichend nicht gelöst hat. Das kann in einem Fall etwa -3 sein. Oder es wird das abweichende Lösungsverhalten gar über eine kompensatorisch verrechenbare Größe bestimmt – was für den AID 2 wegen seiner Kalibrierung nach dem *Rasch-Modell* psychometrisch auch gerechtfertigt wäre. Dann kann ein vom Prototyp negativ abweichendes Verhalten, etwa -3, durch ein vom Prototyp positiv abweichendes Verhalten, etwa +2, (teilweise) kompensiert werden, zu etwa -1 (vielleicht ergänzt um die Abweichung insgesamt, nämlich $3+2=5$). Maßnahmenvorschläge als Konsequenz der Begutachtung sind auf diese Art zielgerichtet, d. h. förderungsorientiert möglich.

Selbstverständlich ist zu hinterfragen, wie es grundsätzlich zur Festlegung der zum Beispiel in Abbildung 1 gegebenen itembezogenen Kriterien kommen kann. Im vorliegenden Beitrag wurde diese Entscheidung vom Verfasser wohl apodiktisch getroffen, d. h. seiner persönlichen, wenn auch fachkundigen Einschätzung gemäß, was denn alles belegt, dass ein 6-jähriges Kind unserer Zeit die minimalen *kognitiven Voraussetzungen für Wissenserwerb und Handlungskompetenz* erbringt. Sicher sind solche Festlegungen der kritische Punkt des hier vorgestellten neuen Zugangs. Trotzdem sollte es möglich sein, dass „jemand“ die bildungspolitische Verantwortung dafür übernimmt.

Fraglich ist jedenfalls die Generalisierbarkeit des gezeigten Zugangs. Noch auf die Intelligenzdiagnostik beschränkt, bedeutet es bereits eine Herausforderung, die minimalen *kognitiven Voraussetzungen für Wissenserwerb und Handlungskompetenz* für solche Tests nachvollziehbar festzulegen, die manuell-visuelle Fähigkeiten prüfen. Abbildung 2 zeigt entsprechendes Material samt den ersten vier Items eines anderen Untertests des AID 2. Mit den gegebenen Würfeln und ihren verschiedenen Seiten sind Muster nach Vorlage zu legen. Gemessen werden soll die Fähigkeit, komplexe (abstrakte) Gestalten durch eine geeignete Strukturierung reproduzieren zu können. Die Festlegung zum Beispiel, „Muster 1 muss von einem 6-Jährigen gelöst werden, die übrigen Muster nicht“, mag darauf fußen, dass Kinder dieses Alters erst die kognitive Entwicklungsstufe des vertikalen (90°) und horizontalen (180°) Strukturierens erreicht haben sollten, nicht aber notwendigerweise schon die Fähigkeit zum Analysieren und Synthetisieren von 45° -Strukturelementen.



Muster 1	1 min	
Muster 2	45 sek	
① Muster 3	1 min	
Muster 4	(nicht: wenn RW = 0) 1 min 30 sek	

Abbildung 2: Die ersten vier Items aus einem manuell-visuelle Fähigkeiten prüfenden Untertest des AID 2 (Kubinger, 2009a). Mit den gegebenen Würfeln ist ein Muster nach Vorlage zu legen.

Erst recht steht die Generalisierbarkeit des gezeigten Zugangs außerhalb der Intelligenzdiagnostik zur Diskussion. Geht also eine kriteriumsorientierte Diagnostik *via* Prototypen auch in der Persönlichkeitsdiagnostik? Unter Umständen sogar leichter. In der Personalauswahl stellt sich der Personalchef zumeist einen ganz bestimmten Typ von Mitarbeiter vor. Z. B. weiß dieser ganz genau, wie Bewerber auf die Frage zur Skala Extraversion im NEO-PI-R (*NEO-Persönlichkeitsinventar nach Costa und McCrae*; Ostendorf & Angleitner, 2004) antworten sollen: „*Ich ärgere mich oft darüber, wie andere Leute mich behandeln.*“ Konkret umgesetzt wurde das sogar schon im Zusammenhang mit der Auswahl von Tierpflegeschülern (Frebort, Litzenberger & Kubinger, in Vorb.); die Ausbildungsleiter wurden gebeten, einen Persönlichkeits-Interessenfragebogen so als Kriterium auszufüllen, wie sie es von erfolgreichen Bewerbern für die vorgesehene Ausbildung erwarten.

Und was die Generalisierbarkeit des gezeigten Zugangs auf Messungen außerhalb der Psychologischen Diagnostik betrifft, z.B. in der Evaluationsforschung, gibt es sicher auch dort Möglichkeiten. So ist ohnehin in der Therapieeffizienzforschung ein kriteriumsorientierter Ansatz eigentlich alltäglich, wenn auch vielleicht kein prototypischer. Geht es zum Beispiel um die Evaluation einer politischen Kampagne, dann mag mit ihr ein bestimmtes prototypisches Verhalten der Bürger angestrebt werden; es interessiert, wie sehr sich ihr tatsächliche Verhalten dem prototypischen durch die Kampagne annähernd, es interessiert aber wenig(er), inwieweit sich das Verhalten eines Landes, normorientiert, hin oder weg zum europäischen Durchschnitt verändert.

Bisher wurden ausschließlich theoretische Prototypen angesprochen; solche also, wie sie sich aus Theorien, vielleicht auch nur aus kompetenten Meinungen ableiten. Es lässt sich aber durchaus auch ein Ansatz mit empirischen Prototypen umsetzen. Als Beispiel diene die virulente Diskussion im Zusammenhang mit Self-Assessments vor allem zur Studienwahlberatung. Da es bei den stets geforderten Validierungsstudien zum methodischen Phänomen der Varianzeinschränkung kommt (vgl. z.B. Kubinger, 2009b), sind entsprechende Korrelationsstudien schlicht obsolet. Zum **Wiener Self-Assessment Psychologie** (s. Frebort & Kubinger, 2008; Sonnleitner, Kubinger & Frebort, 2009) wird deshalb derzeit versucht, das prototypische Testergebnis von außerordentlich erfolgreichen Absolventen des Psychologiestudiums zu erfassen; die Eignung von künftigen Bewerbern kann dann an der Abweichung zu diesem Prototyp valide beurteilt werden.

Diskussion

Es stellt sich die Frage, ob nicht in Zukunft psychologisches Diagnostizieren dahingehend zu reflektieren ist, ob im Sinne der Fragestellung eine an Prototypen orientierte kriteriumsbezogene Begutachtung statt

einer normorientierten erfolgen kann (oder vielleicht sogar soll). Wie das gangbar ist, wurde im vorliegenden Beitrag illustriert. Aussteht, wer es sich dabei „getraut“, die (theoretischen) Prototypen festzulegen; und welche Akzeptanz diese dann in der Gesellschaft finden. Nimmt man sich für die Psychologie ein Beispiel an der Medizin, die schlechthin Gesundheits-Prototypen, etwa in Form des bekannten Beispiels des Body-Maß-Index’ propagiert, dann sollte die kriteriumsorientierte Diagnostik via Prototypen durchaus populär werden können.

Literatur:

- Bents, R. & Blank, R. (2004). *Golden Profiler of Personality – Deutsche Adaption (GPOP)*. Bern: Huber.
- Frebort, M. & Kubinger, K.D. (2008). Qualitätsansprüche an ein Self-Assessment zur Studienwahlberatung: Der Wiener Ansatz. In H. Schuler & B. Hell (Hrsg.), *Studierendenauswahl und Studienentscheidung* (S. 95-101). Göttingen: Hogrefe.
- Guthke, J. (1977). *Zur Diagnostik der intellektuellen Lernfähigkeit*. Berlin: Deutscher Verlag der Wissenschaften.
- Kubinger, K. D. (2009a). *Adaptives Intelligenz Diagnostikum – Version 2.2 (AID 2) samt AID 2-Türkisch*. Göttingen: Beltz.
- Kubinger, K.D. (2009b). *Psychologische Diagnostik – Theorie und Praxis psychologischen Diagnostizierens* (2., überarb. u. erweiterte Aufl.). Göttingen: Hogrefe.
- Kubinger, K. D. & Wurst, E. (1985). *Adaptives Intelligenz Diagnostikum (AID)*. Weinheim: Beltz.
- Kubinger, K. D. & Wurst, E. (2000). *Adaptives Intelligenz Diagnostikum – Version 2.1 (AID 2)*. Göttingen: Beltz.
- Ostendorf, F. & Angleitner, A. (2004). *NEO-Persönlichkeitsinventar nach Costa und McCrae, Revidierte Fassung (NEO-PI-R)*. Göttingen: Hogrefe.
- Sonnleitner, P., Kubinger, K.D. & Frebort, M. (2009). Das Wiener Self-Assessment Psychologie mit seinen Verfahren der experimentalpsychologischen Verhaltensdiagnostik zur Messung studienfachübergreifender *Soft Skills*. In G. Rudinger & K. Hörsch. (Hrsg.), *Self-Assessment an Hochschulen: Von der Studienfachwahl zur Profilbildung* (S. 63-72). Bonn: V&Runipress.