

1

Lehrereignung – Aufgaben, Probleme, Voraussetzungen und Grundlagen

1.1 Eignungsdiagnostik für angehende Lehrkräfte

Das Für und Wider

Wenn von Lehrereignung gesprochen wird, dann hat man in der Regel die Eignungsabklärung für angehende Lehrkräfte im Auge, die möglichst schon vor dem Studium angestrebt wird. Dies prägt das Verständnis der Eignungsfrage in Bezug auf den Arbeitsplatz Schule. Dabei wird dieses Thema durchaus kontrovers diskutiert. Und so sei

eingangs die Frage erörtert: Braucht man überhaupt eine Eignungsdiagnostik für künftige Lehrerinnen und Lehrer?

Man könnte zunächst ins Feld führen, dass ein Lehramtsstudium auch bislang nur unter klar definierten Bedingungen begonnen werden kann (vgl. Füssel, 2014). Kandidaten müssen die allgemeine Hochschulreife besitzen, die je nach Bewerberzahlen und Ausbildungskapazitäten erst ab einem bestimmten Notenschnitt den Zugang zum Studium öffnet. Unter Umständen ist die Studienzulassung sogar noch an das erfolgreiche Durchlaufen spezieller Vorprüfungen geknüpft (etwa bei Interesse an den Fächern Sport, Musik und Kunst). Diese Filter sind zwar weit weniger restriktiv eingestellt als bei anderen Studiengängen (z. B. bei heiß begehrten Numerus-Clausus-Fächern wie Medizin oder Psychologie, in denen ohne Wartezeit und Härtefallregelungen meist nur Einser-Abiturienten zum Zuge kommen), haben aber dennoch Steuerungsfunktion². In der Entscheidung, eine Bewerbung einzureichen und im Bewilligungsfalle das Studium aufzunehmen, scheint sich zudem eine gewisse Affinität zur Berufsperspektive und zum Fach auszudrücken. Mithin kann ein persönliches Interesse am Ausbildungsgang vermutet werden. Ferner ließe sich – zugegebenermaßen gewagt – argumentieren, dass im Studium selbst zwangsläufig mit einer selektiven Eignungsdiagnostik zu rechnen sei, wenn auch sehr spät: in Form von Abschlussprüfungen. Diese versperren dann Ungeeigneten zu guter Letzt den Weg in den Beruf.

Trotz dieser Mechanismen der Laufbahnsteuerung (denen es mitunter freilich an Auswahlgerechtigkeit, Systematik und methodi-

2 Neuere Untersuchungen belegen allerdings, dass Bewerber für ein Lehramtsstudium (zumal für die Ausrichtung Primarstufe und Sek I) im Schnitt etwas schlechtere Abiturnoten haben als die Kandidaten für andere Studiengänge, also im Gros nicht zur Leistungsspitze der Schulabgängerkohorten zählen (vgl. Wössmann, 2009; Retelsdorf & Möller, 2012; Neugebauer, 2013). Das muss für die spätere Bewährung im Beruf kein Hindernis sein, könnte aber unter Umständen Rückschlüsse auf prekäre Studienmotive (»Wenigstens für ein Lehramtsstudium reicht die Abi-Note!«) oder ungünstigere Lernstile erlauben. (Zur ausführlichen und kritischen Diskussion der vorliegenden Befunde vgl. Rothland, 2014a).

scher Kontrolle mangelt) erübrigen sich nach unserer Überzeugung Anstrengungen um eine wissenschaftliche Eignungsabklärung und -beratung keineswegs – und das aus wenigstens zwei Gründen: Zum einen kann mit Hilfe entsprechender Konzepte und Instrumentarien einer Vergeudung von Ressourcen sowohl auf der institutionellen wie auf der persönlichen Ebene effektiver vorgebeugt werden. Falsche Studienentscheidungen und hohe Abbrecher- oder Berufswechslerzahlen bedeuten eine Verschwendung wertvoller Lebenszeit, sind ein enormer Kostenfaktor und erzeugen oft einen großen Leidensdruck bei denjenigen, die mit ihren Berufsplänen scheitern, und ihrem privaten Umfeld (vgl. Henecka & Gesk, 1996; Heublein, Schmelzer & Sommer, 2008). Zweitens bürgt die Bewältigung akademischer Anforderungen, dokumentiert durch den Eignungsindikator *Abchlussnote*, nur bedingt für Erfolg und Zufriedenheit in der späteren Berufsausübung. Nach Befunden der Potsdamer Lehrerstudie leidet z. B. fast jeder dritte in der Schule tätige Pädagoge unter Erschöpfungssymptomen, Unzufriedenheit mit der Berufswahl und Überforderungsgefühlen (► Kap. 5 sowie Schaarschmidt 2005; Schaarschmidt & Kieschke 2013; Kieschke & Schaarschmidt 2008).

Erstaunlich viele Lehramtsstudierende zeigen sich schon *vor* dem Eintritt in die Berufspraxis verunsichert: Zwischen 13 und 26 % schwanken Schätzungen zur Gruppengröße derer, die an der Richtigkeit ihres Studienentschlusses zweifeln (Rothland, 2011, S. 291). Und das ist kein deutsches Phänomen – internationale Vergleichsstudien berichten ähnliche Trends (vgl. Daniels et al., 2006). Bei all dem ist noch etwas anderes im Auge zu behalten: Die Berufswahl einer Person hat Konsequenzen, die über die Belange des Laufbahnanwärters selbst hinausreichen (vgl. Bergmann, 2004). Sie wirkt bei Tätigkeitsaufnahme auch auf das Arbeitsumfeld und dessen Mit-Akteure. Das klingt zunächst trivial, hat große Brisanz aber gerade im Lehrerberuf. Lehrkräfte beeinflussen auf substanzielle Weise die Lern- und Entwicklungschancen ihrer Schüler und Schülerinnen (vgl. z. B. Rowan, Correnti & Miller, 2002; Hattie & Yates, 2014; Nye, Konstantoupolos & Heedges, 2004; Seidel & Shavelson, 2007). Berücksichtigt man, dass jede einzelne Lehrkraft in ihrer Berufsbiogra-

phie Hunderte von Kindern und/oder Jugendlichen unterrichten wird, erschließt sich unmittelbar, warum das Thema »Eignung für den Lehrerberuf« seit einigen Jahren verstärkt in den Fokus der Forschung rückt – national wie international (vgl. Boeger, 2016; Terhart, Bennewitz & Rothland, 2014; Bolle, 2014; Schaarschmidt & Kieschke, 2007; Kieschke, 2015; Watt & Richardson, 2008). Anliegen kann es dabei nicht nur sein, unpassende Kandidaten auszuschließen. Mindestens genauso wichtig wäre es, geeignete Personen für eine pädagogische Berufsperspektive zu gewinnen. Eignungsdiagnostik kann diesen Prozess befördern, wenn sie planvoll in Rekrutierungsstrategien eingebunden und frühzeitig auch zur individualisierten Information genutzt wird (vgl. Kubinger, Frebort & Khorramdel, 2012).

In Ansätzen geschieht das bereits. Viele eignungsdiagnostische Angebote brechen mittlerweile mit der klassischen Engführung auf reine Selektionsziele. Großen Raum nimmt mehr und mehr die Beschreibung des künftigen Tätigkeitsfeldes ein. Indem man bei der Verfahrensgestaltung betont auf die Transparenz von Anforderungsstrukturen und die adressatengerechte Rückmeldung von Ergebnissen setzt, wird das Leistungsspektrum von Eignungsdiagnostik bewusst auf Aspekte wie Aufklärung und Beratung ausgeweitet. Verschiedene Selbsterkundungsbögen informieren den Probanden beispielsweise explizit über die Relevanz eines bestimmten Verhaltensmerkmals, das im Verfahren getestet wird (► Kap. 2 sowie Nieskens, Mayr & Meyerdieks, 2011). Gelangen einschlägige Instrumente schon im Rahmen der allgemeinen Berufsorientierung zum Einsatz, wird das helfen, die Wissensbasis für die Prüfung eigener Wünsche und Vorhaben beizeiten zu verbreitern. Es dürfte noch aus einem anderen Grunde von Vorteil sein, tätigkeitsspezifische Eignungsdiagnostik in einer sehr frühen Phase der Entscheidungsfindung anzusiedeln: Je eher Betroffene über Defizite oder Passungsprobleme im Bilde sind, desto schneller und komplikationsärmer kann über Gegenstrategien nachgedacht werden. Vorstellbar wäre bei Bedarf eine große Bandbreite von Optionen: vom individuell zugeschnittenen Coaching über studienbegleitende Praxistrainings bis hin zum Entschluss, andere Berufsziele anzupeilen (vgl. a. Lehberger & Schaarschmidt, 2009).

Eignungsdiagnostik wird einen solchen Beratungs- und Aufklärungsanspruch allerdings nur dann einlösen können, wenn sie auf gründlicher Kenntnis der Spezifika des Zielberufes fußt: Welche Aufgaben sind zentral und unter welchen persönlichen Voraussetzungen wird man ihnen am ehesten gerecht? Bleiben Antworten hierauf allzu vage und unkonkret, steigt die Gefahr, am beruflichen Anforderungsprofil, an den Realitäten des Arbeitsplatzes »vorbeizudiagnostizieren«.

Voraussetzungen und Besonderheiten eignungsdiagnostischer Aussagen

Eine seriöse Eignungsdiagnostik muss auch die Voraussetzungen und Besonderheiten deutlich machen, auf die sich ihre Aussagen gründen. Skepsis und Ablehnung bezüglich der Eignungsdiagnostik – und das gilt generell, für lehrerbezogene Aussagen aber vielleicht besonders – haben oftmals auch damit zu tun, dass einige ihrer Prämissen nicht ausreichend beachtet und damit u. U. falsche Erwartungen an die Aussagekraft eignungsdiagnostischer Urteile geknüpft werden. Es seien dazu einige grundsätzliche Anmerkungen festgehalten (vgl. Kieschke, 2015; ausführlicher zu den einzelnen Aspekten: Schmidt-Atzert & Amelang, 2012; Krohne & Hock, 2015).

Eignungsaussagen sind Wahrscheinlichkeitsaussagen

Eignungsdiagnostiker hantieren ausdrücklich mit Wahrscheinlichkeitsaussagen. Sie fällen keine apodiktischen Urteile (»Hopp oder Top!«); sie äußern sich zur *relativen* Passung zwischen Person und Tätigkeit und suchen gemäß wissenschaftlicher Argumentationsstandards plausible Handlungsempfehlungen abzuleiten (besagte Argumentationsstandards und Qualitätsdesiderate sind inzwischen in einer DIN-Norm verankert: der DIN 33430, vgl. Kersting, 2008). Dass solche Abwägungsprozesse am Ende sehr wohl in kategoriale Entscheidungen einmünden (z. B.: Arbeitsvertrag ja oder nein), ändert nichts an der Wahrscheinlichkeitsnatur eignungsdiagnosti-

scher Auskünfte, veranschaulicht aber sinnfällig, dass pragmatische Festlegungszwänge und Ansprüche an die klare Verwertbarkeit begründeter Urteile durch wissenschaftstheoretische Prinzipienverweise kaum auszuhebeln sind. Das ist auch gut so, sonst käme man an kein Ende: Eignungsdiagnostik ist eine Anwendungsdisziplin und deshalb durch die Gravitationskräfte konkreter Problemstellungen (»Welcher Bewerber wird der zum Termin X verfügbaren Stelle am ehesten genügen?«) mitbestimmt.

Eignungsdiagnostische Verfahren müssen Gütekriterien genügen

Die Güte diagnostischer Verfahren bemisst sich nach drei Hauptkriterien, die hierarchisch aufeinander aufbauen (zur empirischen Schätzung der betreffenden Indikatoren und zur Diskussion verschiedener Nebengütekriterien vgl. Schmidt-Atzert & Amelang, 2012; Moosbrugger & Kelava, 2012).

Das Merkmal *Objektivität* bildet das Fundament dieser hierarchischen Struktur. Objektivität meint dabei schlicht die Testleiterunabhängigkeit der Messresultate. Jeder Proband absolviert demnach den psychologischen Test unter vergleichbaren Bedingungen und Auswertungskonditionen (Aufgabenmaterial, Instruktionen, eventuelle Zeitbeschränkungen, Interpretationsschlüssel etc. sind für alle Untersuchungsteilnehmer identisch)³. Ob Testleiter X, Y oder gar ein Com-

3 Der Vollständigkeit halber sei angemerkt, dass es im Rahmen bestimmter Skalierungsmodelle (der sogenannten Item-Response-Modelle) durchaus möglich und statthaft ist, von jener Uniformitätsforderung abzuweichen (vgl. Rost, 2004). Die Auswahl und Reihenfolge der im weiteren Testverlauf zu bearbeitenden Aufgaben kann zum Beispiel an das Fähigkeitsniveau einzelner Probanden angepasst werden, ohne dass dies die Vergleichbarkeit der Testergebnisse mindern würde. (Wohlverstanden: Personen hatten dann möglicherweise unterschiedliche Anforderungstypen und Aufgabenmengen zu bewältigen.) Das ist die Grundidee des adaptiven Testens, das auf modelltheoretisch recht restriktiven Annahmen basiert und sehr hohe Ansprüche an das Aufgabenmaterial stellt (vgl. Frey, 2012).

puterprogramm die Durchführung, numerische Aufbereitung und die Übersetzung der erreichten Ergebnisse in Sachbescheide (»Die Testergebnisse liegen im Normalbereich bzw. sind über- oder unterdurchschnittlich ...«) verwaltet, sollte unerheblich sein.

Ist Objektivität im gerade umrissenen Sinne nachgewiesen, bleibt die *Reliabilität* des Verfahrens zu prüfen. Dieses zweite Gütekriterium beschreibt die formale Exaktheit und technische Verlässlichkeit der Merkmalerfassung, ohne die inhaltliche Frage zu berühren, ob eine Messung das intendierte Phänomen auch wirklich abbildet. Reliabilität zeigt lediglich an, wie gut Messergebnisse unter sonst gleichen Bedingungen reproduzierbar sind, ob also mit einem Test einmal festgestellte Personenunterschiede bei neuerlicher Untersuchung der gleichen Personengruppe mit demselben Verfahren ähnlich wieder sichtbar werden. Verfielen man z. B. auf die immerhin preiswerte und zeitsparende Idee, Intelligenzdiagnostik mit einem Zentimeter-Maß zu betreiben, dann wären die so erhobenen Daten zumindest reliabel (die heute beobachteten Größenunterschiede in Kopfumfang oder Längenwachstum sind in derselben Stichprobe mit hoher Wahrscheinlichkeit morgen oder übernächste Woche wiederzufinden; die Rangfolge der Untersuchten in der Messtabelle dürfte sich kaum wesentlich umsortieren). Herbe Kritik schlägt allerdings demjenigen entgegen, der solche Körpergrößen-Messungen als Aussagen über die Intelligenz der Untersuchten lesen wollte. (Wissenschaftsgeschichtlich gab es durchaus derartige Versuche: z. B. in der von Gall und Spurzheimer Anfang des 19. Jahrhunderts entwickelten Phrenologie, einem Ansatz, der physische Merkmale des Schädels in psychologische Diagnosen ummünzte; vgl. Hagner, 2004).

Die Frage, ob solche inhaltlichen Bezugnahmen sinnvoll sind, ob tatsächlich gemessen wurde, was gemessen werden sollte, beschäftigt den Diagnostiker unter dem Stichwort *Validität*. Mit Studien zur Validität wird der Gültigkeitsrahmen für diagnostische Befunde genauer abgesteckt. Man will hier wissen, welcher Informationsgehalt einem bestimmten diagnostischen Resultat überhaupt zugetraut werden kann (vgl. Borsboom, Mellenbergh & van Heerden, 2004). Erlaubt es Rückschlüsse auf den ursprünglich fokussierten Verhaltensbe-

reich? Es leuchtet sofort ein, dass Validität das entscheidende Gütekriterium eines Verfahrens ist (zuma! es Objektivität und Reliabilität ohnehin voraussetzt). Der schönste Eignungstest ist nichts wert, wenn er mit dem eigentlich angezielten Inhaltsfeld nur wenig zu tun hat.

Eignungsdiagnostik baut auf Vergleichen auf, deren normative Grundlage jeweils klar sein sollte

Eignungsdiagnostik beruht wesentlich auf Vergleichen. Das impliziert für jede Verfahrensentwicklung eine doppelte Frage: Hinsichtlich welcher Merkmale soll jemand mit welcher Referenzgruppe verglichen werden? Die zwei Teile der Frage sind auf unterschiedliche Arten von Stichproben ausgerichtet: zum einen auf *Merkmalsstichproben*, also eine Auswahlmenge eignungsrelevanter Indikatoren aus der Vielzahl psychologischer Beschreibungsmöglichkeiten, und zum anderen auf *Personenstichproben*, also eine Auswahlmenge von Untersuchungsteilnehmern, deren Datenprofil dann als Maßstab für die Einordnung des individuellen Testergebnisses dient. Als Bezugspunkte kämen prinzipiell auch Sachkriterien in Betracht, die vorab festgelegt wurden (vgl. Klauer, 1987).⁴ Statt *soziale* Vergleiche anzustellen (etwa gemäß der Leitfrage: Gehört jemand zu den besten 10 % der Referenzstichprobe?), rasterte man die Testergebnisse jetzt nach einem einheitlichen Bewertungsschema, wie es z. B. für die Zensu-

4 Eignungsdiagnostiker verfügen sogar noch über eine dritte Vergleichsoption: die Gegenüberstellung *intraindividuell*er Daten. Man schaut dann entweder auf das Zueinander unterschiedlicher Merkmale bei derselben Person (das wären so genannte Profilanalysen, Bsp: »X ist bei großer Einsatzbereitschaft und überdurchschnittlicher Intelligenz in sozialen Kontakten eher schüchtern und gehemmt.«) oder auf die Entwicklungsdynamik, die bei ein und demselben Probanden innerhalb einer oder mehrerer Merkmalsdomänen auftritt (Bsp: »X verbessert seine soziale Kompetenz im viertägigen Training bedeutsam.«; vgl. zu Varianten solchen dynamischen Testens Guthke, Beckmann & Wiedl, 2003; Stangel-Meseke, 2005).

renvergabe bei Klausuren üblich ist. Sind mindestens 95 % der Aufgaben richtig gelöst, wird mit »1« benotet; 94 % bis 82 % richtige Antworten entsprechen einer »2« etc. In eignungsdiagnostischen Kontexten ist es aber zuweilen schwierig, verbindliche Grenzwerte nach Vorbild der Notenvergabe zu fixieren. Ab wann hätte man z. B. den Persönlichkeitstest gerade noch bestanden, gewissermaßen mit »Vier minus«? Eignungsurteile sind zwingend jeweils inhaltlich zu begründen.

Verantwortungsbewusste Diagnostiker dürfen diesen diskursiven Aufwand nicht scheuen; psychologische Sachkunde ist nun einmal eine komplexe Angelegenheit. Zwei kleine Beispiele mögen das illustrieren.

- ♦ *Beispiel eins:* Die Behauptung »Je ausgeprägter ein Persönlichkeitszug, desto besser (bzw. schlechter)« greift in manchen Eignungsfeldern zu kurz. Die Wirkzusammenhänge sind mitunter vielschichtiger und vertrackter, auch im schulischen Arbeitsalltag. Einfühlungsvermögen, Zuwendungsbereitschaft und Sensibilität für die Problemlagen Heranwachsender gelten etwa als erwünschte Eigenschaften einer Lehrkraft (► Abschnitt 3 dieses Kapitels). Ein Zuviel an Anteilnahme (oder anders ausgedrückt: ein Zuwenig an emotionaler Distanzierung) dürfte jedoch pädagogisches Handeln und subjektives Stressmanagement oftmals erschweren (vgl. Schaarschmidt & Kieschke, 2007).
- ♦ *Beispiel zwei:* Eignungsdiagnostik stützt sich in der Regel auf die simultane Betrachtung mehrerer Personeneigenschaften. Es ist nun vorstellbar, dass günstige Ausprägungen in Merkmal X ungünstige Gegebenheiten in Merkmal Y im Gesamtbild ausbalancieren oder gar überstrahlen. Ob solche Kompensationschancen tatsächlich bestehen und wie sie im Detail beschaffen sind, ist für eignungsdiagnostische Einschätzungen allemal eine interessante Frage (die Anwender übrigens viel zu selten aufwerfen). Beantworten können wird man sie indessen erst, wenn man die Anfangsaufgabe einer jeden ernst zu nehmenden Eignungsdiagnostik erledigt hat: eine umfängliche Anforderungsanalyse.

1.2 Zur Anforderungsstruktur des Lehrerberufs

Rahmenbedingungen der Lehrerarbeit und daraus resultierende Anforderungen

Der Versuch, Anforderungsstrukturen des Lehrerberufs zu thematisieren, scheint zunächst dem Vorhaben zu ähneln, mit angeknipster Taschenlampe einen bereits voll ausgeleuchteten Raum erkunden zu wollen. Eigentlich liegt doch vor aller Augen ausgebreitet, was Lehrkräfte tun (und was sie lassen, wie manch einer kritisch hinzufügen würde). Schulpädagogische Tätigkeit ist eine hochgradig öffentliche Angelegenheit (vgl. Gold, 2015, S.7). Kaum ein anderer Beschäftigter versieht sein Hauptarbeitspensum vor solch einem großen Publikum. Nahezu jeder (auch jede spätere Lehrkraft) zählte einmal zu einem speziellen Segment jenes Publikums: zur Schülerschaft. Der typische Schüler verlebt – grob kalkuliert – etwa 15.000 Stunden in der Schule und sammelt dabei eine Vielzahl von Erfahrungen mit den unterschiedlichsten Vertretern der Lehrprofession (vgl. Rutter, Maughan, Mortimore & Ouston, 1980). Die Verbindungen zum Berufsfeld sind meist selbst nach Ende der Schulzeit keineswegs gekappt. Ob als Elternteil, als Verwandter oder Bekannter von Schulpflichtigen, als aufmerksamer Beobachter der Bildungsdebatte, vielleicht gar von Berufs wegen: Man bleibt in Fühlung mit dem Thema Lehrerarbeit. Von daher mag es aufs Erste überraschen, wie fest bestimmte Vorurteile und Fehleinschätzungen bezüglich der Besonderheiten des Schuldienstes eingewurzelt sind. Das fängt bei der gern kolportierten Mär vom »Halbtagsjob mit Dauerferien« an und hört bei Witzeleien über den als anstrengend empfundenen Berufshabitus längst nicht auf (»Lehrer als ewige Besserwisser«).

Hier bahnt sich jedoch eine Änderung der öffentlichen Wahrnehmung an. Lehrkräfte rangieren mittlerweile in der Achtung weiter Bevölkerungskreise vor Rechtsanwälten, Politikern und Unternehmern (vgl. Institut für Demoskopie Allensbach, 2011). Das