

Die Qual der Vorauswahl: Online-Tests identifizieren geeignete Kandidaten besser als Schulnoten

Eine Studie zeigt, dass Online-Tests besser geeignet sind als Schulnoten, um eine Vorselektion von Bewerbern im Hinblick auf einen beaufsichtigten und somit rechtssicheren Leistungstest zu treffen. Auch hinsichtlich der Objektivität und Fairness sind Online-Tests gegenüber Schulnoten im Vorteil.

Standardisierte Intelligenz- und Leistungstests zählen neben Assessment-Centern heute zum Handwerkszeug professioneller Personalauswahl. Hülshager et al. (2006) berichten in ihrer Metaanalyse etwa, dass Intelligenztests hervorragende Prognosen für den Ausbildungserfolg liefern. Bei großen Bewerberzahlen stoßen diese Verfahren jedoch an praktische und ökonomische Grenzen, sodass eine Vorselektion der Bewerber unabdingbar ist. Das Ziel besteht dabei weniger in der Auswahl geeigneter, sondern in der Identifikation nicht geeigneter Kandidaten. Sie sollen bereits früh im Auswahlprozess erkannt werden, damit die Zahl an Bewerbern, die zu zeit- und kostenintensiven Leistungstests eingeladen werden, bewältigt werden kann. Immer mehr Unternehmen und Verwaltungen greifen zu diesem Zweck auf Online-Tests zurück. Die Vorteile liegen auf der Hand: Bewerber können den Test zeitlich und örtlich flexibel durchführen und ersparen sich im Fall einer ne-

gativen Eignungsprognose die Anreise und Teilnahme an weiteren Auswahlritten. Bei jüngeren Bewerbern stoßen Online-Tests auf hohe Akzeptanz, was im Wettbewerb um junge Talente eine nicht unerhebliche Rolle spielt. Weitaus häufiger als Online-Tests werden allerdings Schulnoten zur Vorauswahl verwendet. In einer Befragung unter 109 deutschen Unternehmen fanden Benit et al. (2014) heraus, dass knapp zwei Drittel der befragten Unternehmen zur Vorselektion Schulzeugnisse heranziehen. Deren Beliebtheit lässt sich dadurch erklären, dass sie gängiger Bestandteil von Bewerbungsunterlagen sind und eine hohe Augenscheinvalidität haben. Zudem liegen neben Anschreiben und Lebenslauf meist keine weiteren „harten“ Auswahlkriterien vor. Die große Bedeutung schulischer Leistungen in Auswahlprozessen mag auch einem Mangel an kostengünstigen Alternativen geschuldet sein, die in Form von Online-Tests noch nicht allzu lange zur Verfügung stehen. Doch welches Auswahlkriterium führt zu valideren Vorhersagen und sollte daher zur Vorselektion eingesetzt werden? Um einen empirischen Beitrag zu dieser Frage zu liefern, wurde in der vorliegenden Studie anhand zweier unterschiedlicher Personalauswahlverfahren untersucht, wie sich Schulnoten und ein Online-Test zur Vorauswahl von Bewerbern im Hinblick auf den Erfolg in einem darauf folgenden beaufsichtigten Leistungstest eignen.

Ablauf der Untersuchung

Die Daten wurden im Rahmen des Personalauswahlprozesses einer deutschen Landesverwaltung erhoben. Ausgeschrieben waren Ausbildungsstellen für ein duales Studium zum Diplom-Finanzwirt sowie für eine Ausbildung zum Finanzwirt. Nach Bewerbungsschluss wurden alle Bewerber per E-Mail zu einem Online-Vortest eingeladen. Kandidaten, die darin eine zuvor festgelegte Punktzahl erzielt hatten, wurden zu einem beaufsichtigten Haupttest eingeladen, der vier Wochen später stattfand.

Stichprobe

Stichprobe 1 (siehe Tabelle Seite 18) umfasst die Daten von 1227 Personen (58 Prozent weiblich), die sich für das duale Studium zum Diplom-Finanzwirt bewarben; Stichprobe 2 (siehe Tabelle Seite 18) besteht aus 469 Personen (60 Prozent weiblich), die am Auswahlprozess für die Aus-



*Benjamin Haarhaus,
Diplom-Psychologe,
B.Sc. in Wirtschaftswissenschaften,
Mitarbeiter im Bereich Test-
entwicklung und Eignungsdiagnostik
bei der Deutschen Gesellschaft für
Personalwesen e. V. in Düsseldorf,
haarhaus@dgp.de*



*Birk Diedenhofen,
Diplom-Psychologe,
wissenschaftlicher Mitarbeiter an der
Universität Düsseldorf, Institut für
Experimentelle Psychologie, Abtei-
lung für Diagnostik und Differentielle
Psychologie,
birk.diedenhofen@uni-duesseldorf.de*

bildung zum Finanzwirt teilnahmen. Das Durchschnittsalter in Stichprobe 1 betrug 24,71 Jahre (*SD*: 5,91) und in Stichprobe 2 waren es 26,94 Jahre (*SD*: 6,90). Das schulische Bildungsniveau war in Stichprobe 1 erwartungsgemäß höher ausgeprägt als in Stichprobe 2. Während in Stichprobe 1 über 90 Prozent der Bewerber mindestens die (Fach-)Hochschulreife erreichten, gab der Großteil (82 Prozent) der Bewerber in Stichprobe 2 mittlere Reife als höchsten Bildungsabschluss an. Am beaufsichtigten Haupttest nahmen 561 Bewerber aus Stichprobe 1 (46 Prozent) und 272 Bewerber aus Stichprobe 2 (58 Prozent) teil.

Testverfahren und Schulnoten

Beide Bewerbergruppen bearbeiteten einen Online-Vortest, der auf dem Berliner Intelligenzstrukturmodell basiert, etwa eineinhalb Stunden dauert und die drei Dimensionen numerische und verbale Verarbeitungskapazität sowie Verarbeitungsgeschwindigkeit erfasst. Die internen Konsistenzen der gemessenen Fähigkeitsdimensionen liegen zwischen $\alpha = .77$ und $\alpha = .87$. Im schriftlichen Haupttest bearbeiteten beide Gruppen unterschiedliche Testbatterien, die sich hauptsächlich im Niveau und in der Auswahl und Gewichtung der Subtests unterscheiden. Personen, die sich für das duale Studium bewarben (Stichprobe 1), bearbeiteten eine auf den gehobenen Dienst zugeschnittene Testbatterie, die numerische und verbale Verarbeitungskapazität, das Arbeitsverhalten sowie verschiedene anforderungsrelevante Kenntnisbereiche erfasst. Bewerber für die Ausbildung (Stichprobe 2) absolvierten eine Testbatterie, die auf den mittleren Dienst ausgerichtet ist (zum Beispiel Verwaltungsfachangestellte und Bürokaufleute). Neben den Konstrukten, die auch von der in Stichprobe 1 verwendeten Testbatterie abgedeckt werden, misst die Testbatterie in Stichprobe 2 darüber hinaus die Merkfähigkeit und Rechtschreibkompetenz der Bewerber. Die Bearbeitung der beiden Testbatterien dauert jeweils etwa 3,5 bis vier

Stunden. Die internen Konsistenzen der Fähigkeitsdimensionen liegen zwischen $\alpha = .80$ und $\alpha = .90$. Alle Bewerber wurden vor dem Online-Test gebeten, ihre letzten Zeugnisnoten in Deutsch, Englisch und Mathematik anzugeben. Da sie in verschiedenen Skalierungen vorlagen, wurden sie in eine einheitliche Skala von 0 bis 15 transformiert.

Ergebnisse

Um zu bestimmen, inwiefern sich der Erfolg im Haupttest durch das Abschneiden im Online-Vortest und die Schulnoten vorhersagen lässt, wurden für beide Stichproben Regressionsanalysen durchgeführt. Im ersten Schritt wurden nur die Ergebnisse des Online-Tests berücksichtigt, im zweiten Schritt zählten auch die Noten mit. In beiden Stichproben vermochte der Online-Vortest die Leistungen im Haupttest vorherzusagen (Tabelle, Seite 18). Als die Noten im zweiten Schritt in die Regressionsgleichungen integriert wurden, verbesserte sich die Vorhersage leicht, jedoch trug lediglich die Englischnote signifikant zur Vorhersage der Testergebnisse bei. Da nur Personen mit guten Leistungen im Online-Vortest zum Haupttest zugelassen wurden, wird die Stärke des Zusammenhangs zwischen beiden Tests unterschätzt. Korrigiert man die Korrelation zwischen Online-Test und Haupttest um die Einschränkung der Prädiktorvarianz, ergeben sich Korrelationen von $r = .54$ in Stichprobe 1 und $r = .44$ in Stichprobe 2.

Diskussion

Die vorliegenden Ergebnisse zeigen, dass Online-Tests den Erfolg in darauf folgenden schriftlichen Leistungstests zuverlässig vorhersagen können. Damit empfiehlt es sich, zur Bewältigung großer Bewerberzahlen die ökonomischen Vorteile online-basierter Vorauswahltests mit der hohen Validität von Intelligenz- und Leistungstests zu kombinieren (Hülshager et al., 2006). Die hohe prognostische Validität

des Online-Vortests in der vorliegenden Studie lässt sich auf die gleiche Methodik und die ähnlichen Prüfbereiche in Vor- und Haupttest zurückführen.

Bei der Auswahl der Testverfahren sollte darauf geachtet werden, dass sich Vor- und Haupttest methodisch und inhaltlich nicht zu stark unterscheiden. Darüber hinaus ist beim Einsatz von Online-Tests zu beachten, dass Täuschungsversuche nicht zuverlässig ausgeschlossen werden können. Zwar lassen Informationen, wie die Bearbeitungsdauer einzelner Aufgaben, gewisse Rückschlüsse auf Täuschungshandlungen zu, praktisch sind sie jedoch nicht nachweisbar. Um die Motivation zur Täuschung zu senken, sollten Bewerber vor dem Online-Test darauf hingewiesen werden, dass ihre Leistung im weiteren Verlauf auch in einem beaufsichtigten Test vor Ort überprüft wird. Die Ergebnisse der vorliegenden Studie zeigen außerdem, dass die Schulnoten der Bewerber nur wenig zur Treffsicherheit der Vorauswahl beitragen. Lediglich die Englischnote verbesserte die prognostische Validität über den Online-Test hinaus. Das lässt sich eventuell dadurch erklären, dass fremdsprachliche Kompetenzen einen Teilbereich allgemeiner Intelligenz widerspiegeln, den der durchgeführte Online-Test nicht erfasst. Dass sich dieses Ergebnis konsistent über beide Stichproben hinweg zeigt, deutet zudem darauf hin, dass sich das schulische Niveau nicht auf die Validität der Noten auswirkt. Die geringe Vorhersagekraft der Schulnoten bedeutet praktisch, dass Selektionsentscheidungen allein auf Basis von Schulnoten Kosten wegen unzureichender Testungen verursachen und dazu führen, dass potenziell geeigneten Kandidaten die Teilnahme am weiteren Auswahlverfahren zu Unrecht verwehrt wird. Da die Betrachtung von Schulnoten die Prognose zumindest leicht verbesserte, scheint es jedoch gerechtfertigt, Schulnoten zwar nicht ausschließlich, aber zusätzlich zu Online-Tests in die Selektionsentscheidung einfließen zu lassen. Hierbei sind allerdings folgende Punkte zu beachten:

	Stichprobe 1 (Duales Studium)		Stichprobe 2 (Ausbildung)	
	Modell 1	Modell 2	Modell 1	Modell 2
Konstante (B)	60,13	58,24	73,34	70,06
Online-Vortest	,40**	,37**	,35**	,33**
Schulnoten				
Deutsch		–,01		–,03
Englisch		,14*		,15*
Mathematik		,08		,08
R ²	,16	,18	,12	,14
ΔR ²		,025**		,022**

* p < .05; ** p < .01

Vorhersage der Leistung im Haupttest durch die Ergebnisse im Online-Vortest und Schulnoten

- Steht keine Bewerbermanagement-Software zur Verfügung, in die Bewerber ihre Schulnoten selbst eintragen, müssen die Schulnoten einzeln aus den Bewerbungsunterlagen herausgesucht, gegebenenfalls vereinheitlicht und händisch in eine Datenbank oder Tabelle übertragen werden. Hierdurch entstehen möglicherweise Kosten, die zu dem geringen Zusatznutzen von Schulnoten in einem Missverhältnis stehen.
- Schulnoten und Online-Tests lassen sich auf verschiedene Arten kombinieren: Bei einem *sequenziellen Vorgehen* stellen Online-Test und Schulnoten getrennte Auswahlsschritte dar. Entweder lädt man nur solche Personen zum Online-Test ein, die eine vorher festgelegte Notengrenze überschreiten, oder man betrachtet die Schulnoten nur bei denjenigen Bewerbern, die bereits eine Mindestpunktzahl im Online-Test erreicht haben. Dieses Vorgehen führt jedoch aufgrund der geringen Validität der Noten – unabhängig von der Reihenfolge – zu Fehlentscheidungen und wird nicht empfohlen. Alternativ lassen sich Schulnoten auch mit dem Ergebnis des Online-Tests *verrechnen*. Hier stellt sich jedoch die Frage, nach welcher Vorschrift die Verrechnung erfolgen soll. Um Online-Test und Schulnoten optimal zu gewichten, benötigt man nicht nur statistisches Know-how, sondern auch zuverlässige Validierungsdaten.
- Selbst wenn sich Schulnoten ohne zu-

sätzlichen Aufwand erheben und in optimaler Weise mit dem Ergebnis des Online-Tests verrechnen ließen, bleiben ihre Objektivität und Fairness fraglich. Vielfach wurde in empirischen Studien nachgewiesen, dass sich die Benotungsmaßstäbe zwischen Bundesländern und Schultypen unterscheiden (etwa Klieme, 2003; Neumann & Nagy, 2009). Auch die Tatsache, dass Mädchen in der Regel bessere Schulnoten erhalten als Jungen (Buchmann, DiPrete & McDaniel, 2008) und in Abhängigkeit des Fachs unterschiedliche Benotungsmaßstäbe für Mädchen und Jungen gelten (Blossfeld et al., 2009), lässt Schulnoten für eine objektive und faire Personalauswahl in einem schlechten Licht erscheinen.

Trotz der angesprochenen Probleme haben Schulnoten durchaus Vorzüge. Sie erfassen Leistungen mit verschiedenen Methoden (etwa schriftliche Prüfungen, mündliche Mitarbeit, Praxisarbeiten), und das über einen langen Zeitraum. Es wäre daher voreilig, in der Personalauswahl komplett auf sie zu verzichten. In strukturierten Interviews besteht etwa die Möglichkeit, sich nach den Hintergründen besonders guter oder besonders schlechter Leistungen zu erkundigen. Ferner lassen Schulnoten möglicherweise Rückschlüsse auf Entwicklungsprozesse zu, die einmalige Testungen nicht bieten können. Fragen zu schulischen Leistungen könnten so helfen, anforderungsrelevante Merkma-

le wie Leistungsmotivation oder Gewissenhaftigkeit im Interview zu erfassen.

Fazit

Bei großen Bewerberzahlen benötigen Unternehmen und Verwaltungen eine zuverlässige Vorselektion. Die Ergebnisse der Studie deuten darauf hin, dass sich Online-Tests besser als Schulnoten eignen, um eine Vorselektion für einen beaufsichtigten und so rechtssicheren Leistungstest zu treffen. Auch bei Objektivität und Fairness sind Schulnoten im Nachteil. Schulnoten sollten daher nicht als „hartes“ Selektionskriterium, sondern nur zur Formulierung anforderungsbezogener Fragen im Interview verwendet werden.

Weiterführende Literatur

- Benit, N., Mojzisch, A. & Soellner, R.** (2014). Preselection methods prior to the internal assessment center for personnel selection in German companies. *International Journal of Selection and Assessment*, 22, 253–260.
- Blossfeld, H.-P., Bos, W., Hannover, B., Lenzen, D., Müller-Böling, D. & Prenzel, M.** (2009). *Geschlechterdifferenzen im Bildungssystem: Jahresgutachten 2009*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Buchmann, C., DiPrete, T. A. & McDaniel, A.** (2008). Gender inequalities in education. *Annual Review of Sociology*, 34, 319–337.
- Hülshager, U. R., Maier, G. W., Stumpp, T. & Muck, P. M.** (2006). Vergleich kriteriumsbezogener Validitäten verschiedener Intelligenztests zur Vorhersage von Ausbildungserfolg in Deutschland. *Zeitschrift für Personalpsychologie*, 5, 145–162.
- Klieme, E.** (2003). Benotungsmaßstäbe an Schulen: Pädagogische Praxis und institutionelle Bedingungen. Eine empirische Analyse auf der Basis der PISA-Studie. In H. Döbert, B. von Kopp, R. Martini & M. Weiß (Hrsg.), *Bildung vor neuen Herausforderungen* (S. 195–211). Neuwied: Luchterhand.
- Neumann, M. & Nagy, G.** (2009). Vergleichbarkeit von Abiturleistungen. *Zeitschrift für Erziehungswissenschaft*, 12, 691–714.