

Tests statt Uni-Noten: Bessere Erfolgsprognosen für Wirtschaftsprüfer

Ob ein Mitarbeiter das anspruchsvolle Examen zum Wirtschaftsprüfer besteht, lässt sich besser anhand von psychologischen Tests als anhand der Noten vorhersagen. Das zeigt eine Studie aus der Schweiz.

Schul- und Studiennoten sind klassische erste Selektionsfilter für die Auswahl von Berufseinsteigern. Gemeinsam mit anderen Kriterien entscheiden sie häufig über die Zulassung zu weiteren Methoden der Personalauswahl (zum Beispiel Vorstellungsgesprächen). Dies reduziert den Anteil an Bewerbern, der psychologischen Methoden der Eignungsdiagnostik zugeführt wird, obwohl prognostische Mängel wie die mangelnde Vergleichbarkeit von Schulabschlüssen oder die Wohlwollenspflicht bei der Zeugniserstellung bekannt sind (vgl. zusammenfassend Frintrup & Piechowski, 2011). In der hier vorgestellten Studie wurde daher die prognostische Validität von Examennoten und psychometrischen Testverfahren zur Vorhersage des Erfolgs in den Berufsexamina Schweizer Wirtschaftsprüfer untersucht. Zielgruppe waren Hochschulabsolventen, die sich bei einem großen Wirtschaftsprüfungs- und Steuerberatungsunternehmen beworben hatten, eingestellt wurden und berufsbegleitend das schweizerische Berufsexamen zum Wirtschaftsprüfer ablegten. Der Auswahlprozess umfasste mehrere Stufen:

- Personalfragebogen zur Erfassung formaler Voraussetzungen (u. a. Studienrichtung und -noten);
- psychologischer Online-Pre-Test (ungeschützt von zu Hause über das Internet zu bearbeiten);
- Testverfahren in geschützter Umgebung;
- Interview-Verfahren und Fachgespräch sowie das Kennenlernen von Arbeitsplatz und künftigen Kollegen.

Auf jeder Stufe wurden individuelle Merkmalsausprägungen der anforderungsanalytisch ermittelten Erfolgsvoraussetzungen bestimmt, um die Entscheidung (Ablehnung/Einstellung) zu fundieren. Die vorgelagerte Stufe bildete jeweils eine Selektionshürde für die nächstfolgende.

Prozess

Der Validitätsstudie liegen Daten zugrunde, die während des Auswahlprozesses erhoben wurden und zunächst dem Personalauswahlprozess dienten. Sie wurden zusätzlich hinsichtlich ihrer prognostischen Validität für

Leistungen im Wirtschaftsprüferexamen (WP-Examen) untersucht und dabei der Vorhersagekraft von Hochschulabschlussnoten (Uni-Note) gegenübergestellt. Die Analyse konzentrierte sich auf einen Unternehmensteil, in dem vergleichsweise viele Mitarbeiter das Berufsexamen des Wirtschaftsprüfers ablegen, die vor ihrer Einstellung alle identische Auswahltests absolviert haben.

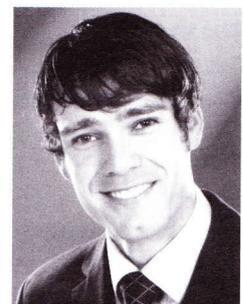
Stichprobe

Alle Personen verfügten über einen abgeschlossenen Hochschulabschluss (Wirtschaftswissenschaften 33,8 Prozent, Master of Business Administration 28,4 Prozent sowie andere Studienfächer 37,8 Prozent), waren mindestens drei Jahre angestellt und legten in dieser Zeit ihr WP-Examen ab. Es handelte sich um 69 Kandidaten, davon waren 27,5 Prozent weiblich. Für 54 Personen lagen Uni-Noten vor. Das Alter kann aufgrund von Datenschutzrichtlinien nicht berichtet werden, allerdings kann die Stichprobe als relativ homogen und für Berufseinsteiger mit Hochschulabschluss als typisch gelten.

Kriterium

Beim WP-Examen handelt es sich um eine kognitiv anspruchsvolle, berufsbegleitende Prüfung, die eine qua-

Maik Spengler,
Diplom-Psychologe,
Projektleiter Forschung und Entwicklung bei S&F Personalpsychologie Managementberatung in Stuttgart
spengler@personalpsychologie.de



Andreas Frintrup,
Diplom-Ökonom,
Vorstand HR Diagnostics in Stuttgart
frintrup@HR-Diagnostics.de



	Kriterium 1 WP-Examen bestanden	Kriterium 2 Gesamtnote im WP-Examen
Uni-Note	.09 (.12)	.21 (.29)
Kognition	.41** (.54)	.47** (.61)
Motivation	.34** (.45)	.35** (.47)

** Korrelation signifikant bei $p < .01$ für $N=54$ (Uni-Note) und $N=69$ (Kognition und Motivation). Korrelationen in Klammern sind korrigiert um Reliabilitätseinschränkungen im Kriterium (.80) und Vorselektionseffekte (Selektionsquote = .10)

Tabelle 1: Prädiktive Validitäten zur Vorhersage des Wirtschaftsprüferexamens

lizierende praktische Berufserfahrung voraussetzt. Das Bestehen der Prüfung qualifiziert zum diplomierten Wirtschaftsprüfer und wird von unabhängigen Gutachtern der Schweizerischen Treuhänderkammer abgenommen. Die Bewertung erfolgt anhand mehrerer (mündlicher und schriftlicher) Testteile und gilt als bestanden, wenn ein Kandidat die einzelnen Prüfungsteile mindestens mit der Note „ausreichend“ ablegt. Innerhalb der Branche ist das Bestehen dieser Prüfung eine wichtige Voraussetzung für die individuelle Karriereentwicklung, da diese Qualifikation eine eigenverantwortliche Erstellung von Prüfungsberichten erlaubt.

Prädiktoren

Aufgrund der hohen kognitiven Anforderungen des Berufsexamens wurden anspruchsvolle Verfahren zur Erfassung kognitiver Fähigkeiten eingesetzt: ein Test zum schlussfolgernden Denken auf Basis verbalen Materials (Syllogismen lösen – Syllog) sowie ein als Hybridtest angelegtes Intelligenztestverfahren in Form einer Büroarbeitsprobe (Proffice). Das Bestehen des Berufsexamens setzt aufgrund des berufs begleitenden Charakters, des Anspruchsniveaus und des erforderlichen Verzichts auf außerberufliche Lebensfreuden neben kognitiver Leistungsfähigkeit auch ein hohes Maß an Motivation voraus. Deshalb wurden zusätzlich anforderungsrelevante Skalen der beruflichen Leistungsmotivation (etwa Leistungsstolz und Wettbewerbsorientierung) ausgewählt, um die Ergebnisse der WP-Examen vorherzusagen, da mit den eingesetzten Testverfahren auch andere Anforderungen (zum Beispiel erfolgreiche Kundenberatung) abzubilden waren, die für die Vorhersage

des WP-Examens jedoch als irrelevant einzuschätzen sind.

Ergebnisse

Insgesamt wurden $N=69$ Personen berücksichtigt, wovon 73 Prozent das WP-Examen bestanden. Die weiblichen

Kandidatinnen schnitten etwas besser ab, aber die Mittelwertsunterschiede waren nicht signifikant ($T(67) = -1,68, p > .05$). Auch die unterschiedlichen Studienfächer (Master of Business Administration, Studium der Wirtschaftswissenschaften oder andere Abschlüsse) hatten keinen signifikanten Effekt auf die Noten ($F(69) = 0,12, p > .05$).

Neben dem Bestehen des WP-Examens wurden die vorliegenden Ergebnisse der Einzelprüfungen im WP-Examen zu einem Gesamtwert aggregiert, um auch qualitative Unterschiede in den Prüfungsleistungen zu berichten. Die prädiktive Validität der universitären Abschlussnote (Uni-Note) sowie der psychologischen Testverfahren zur Erfassung kognitiver Fähigkeiten (Kognition) und der beruflichen Leistungsmotivation (Motivation) für diese beiden Kriterien gestaltet sich wie in Tabelle 1 dargestellt. Die Berechnung der Validitätskoeffizienten in Klammern folgt den Empfehlungen von Hunter und Schmidt (2004). Es wurde um die eingeschränkte Reliabilität der Noten im Berufsexamen korrigiert ($r = .80$), darüber hinaus wurde die eingeschränkte Varianz der Leistungen aufgrund der geringen Selektionsquote (10 Prozent) bei der Ermittlung der korrigierten prognostischen Validitäten berücksichtigt.

Dem Nützlichkeitsprinzip folgend müssen die Tests gegenüber den Uni-Noten einen Mehrwert für das Unternehmen bieten. Anhand einer Hierarchischen Multiplen Regression wurde der inkrementelle Nutzen der eingesetzten Testverfahren zur Vorhersage der Leistungen im Wirtschaftsprüferexamen – über die vorliegenden Uni-Noten hinaus – untersucht (siehe Tabelle 2). Wie die Daten zeigen, erlaubt die Uni-Note keine verlässliche Vorhersage der Leistung im WP-Examen, wohingegen sowohl Testverfahren zur Erfassung kognitiver Fähigkeiten als auch die Ergebnisse im beruflichen Leistungsmotivationstest zu einer bedeutsamen Steigerung der prognostischen Validität führen und insgesamt 27 Prozent der Leistungsunterschiede der Kandidaten vorhersagen. Ferner wurde die Treffsicherheit ermittelt, mit der die zum Zeitpunkt der Einstellung vorliegenden Testergebnisse die Leistung im mindestens drei Jahre später abgelegten WP-Examen vorhersagen können. Die in Tabelle 3 dargestellten prozentualen Häufigkeiten geben den Anteil richtiger beziehungsweise falscher Klassifikationen der Kandidaten für ihre Prüfungsleistung im WP-Examen an. Die Ergebnisse sind sowohl für die Vorhersage der Uni-Note als auch für die Prognose der psychologischen Testverfahren bei einem unteren Cut-Off von jeweils $Z=95$ im Prädiktor angegeben, welcher sich am Anteil der Personen (ca. 70 Prozent) orientiert, die das WP-Examen üblicherweise bestehen.

Wie Tabelle 3 auf Seite 57 zeigt, würden 37 Prozent aller Kandidaten wegen ihrer Universitätsabschlussnote eine falsche Prognose für ihr Abschneiden im WP-Examen erhalten. Basiert die Prognose hingegen auf den psychologischen Testverfahren,

Tabelle 2: Inkrementelle Validität der Testverfahren über Universitätsabschlussnoten hinaus

Modell	1			2			3		
	R	R ²	ΔR^2	R	R ²	ΔR^2	R	R ²	ΔR^2
1. Uni-Note	.09	.01	-.01						
2. Uni-Note Kognition				.09 .40	.01 .16	-.01 .5*			
3. Uni-Note Kognition & Motivation							.09 .53	.01 .28	-.01 .27***

N=54 * $p < .05$ ** $p < .01$ *** $p < .001$

reduziert sich der Anteil falscher Klassifikationen merklich. Der Anteil richtiger Klassifikationen steigt im Vergleich zur Prognose auf Grundlage der Uni-Noten von 63 auf 83 Prozent. Der Anteil ungeeigneter Mitarbeiter, die fälschlicherweise als geeignet klassifiziert werden, lässt sich durch den Einsatz von Testverfahren von 15 auf 6 Prozent reduzieren. Gleichzeitig sinkt auch der Anteil geeigneter Kandidaten um die Hälfte (11 statt 22 Prozent), die auf Basis der Uni-Noten fälschlicherweise keine Empfehlung für die Aufnahme einer Weiterbildung zum Wirtschaftsprüfer erhalten hätten.

Diskussion

Die vorgestellten Ergebnisse zeigen, dass psychologische Testverfahren eine signifikante Vorhersagekraft für das WP-Examen aufweisen – im Gegensatz zu den Uni-Noten. Wegen der hohen Attraktivität großer Wirtschaftsprüfungsgesellschaften als Arbeitgeber kann auch das betreffende Unternehmen unter vielen geeigneten Universitätsabsolventen auswählen. Durch die kontinuierliche interne Selektion („Up or Out“-Prinzip) während der ersten Berufsjahre kommt es zu einer hohen Fluktuation weniger geeigneter Kandidaten. Die aus diesen Rahmenbedingungen resultierende Ähnlichkeit der Leistungsfähigkeit der Bewerber reduziert aber auch die empirisch bestimmten Validitätskennzahlen. Vor diesem Hintergrund können die ermittelten Kennwerte für die psychologischen Testverfahren als konservative Schätzung und daher im Ergebnis als umso bemerkenswerter gelten. Die zusätzliche Darstellung korrigierter Kennzahlen berücksichtigt die vorliegenden Rahmenbedingungen eines attraktiven Arbeitgebers sowie anzunehmende Beurteilerfehler im WP-Examen. Da es auch mündliche Prüfungsteile enthält, die in der Regel keine zuverlässigen Messungswiederholungen ermöglichen, wäre eine stärkere Korrektur um Reliabilitäts-Einschränkungen ebenfalls vertretbar gewesen. Die prognostischen Validitäten der kognitiven Testverfahren erreichen damit Kennwerte wie sie metaanalytisch an anderer Stelle für allgemeinen Berufserfolg berichtet werden

	Laut Cut-off (Z=95) Universitätsabschlussnote:		Laut Cut-off (Z=95) psychologischer Testwerte:	
	nicht empfohlen	empfohlen	nicht empfohlen	empfohlen
WP-Prüfung bestanden	22 %	50 %	11 %	61 %
nicht bestanden	13 %	15 %	22 %	6 %
N=54; Uni-Note: Eta=.15; Testverfahren Eta=.61				

Tabelle 3: Ergebnisse der Sensitivitätsanalyse

(Schmidt & Hunter, 2004). Wohingegen die Prognosekraft der Motivations-Skalen andernorts berichtete Validitäten aus dem Temperamentsbereich noch übertrifft (Barrick, Mount & Judge, 2001).

Für das Unternehmen kann ein hoher monetärer Nutzen angenommen werden, der durch den Einsatz von Testverfahren entsteht, da für Mitarbeiter mit Wirtschaftsprüfer-Diplom bei den Kunden höhere Tagessätze fakturiert werden können. Die geringe Übereinstimmung der Uni-Note mit dem Bestehen eines WP-Examen zeigt auf, dass der mit 50 Prozent berichtete Prozentsatz richtiger Klassifikationen der Uni-Note vor allem eine Folge der hohen Basisrate geeigneter Kandidaten ist und noch geringer ausfallen würde, wenn bei der Einstellungsentscheidung keine validen Testverfahren verwendet worden wären.

Einschränkungen

Da die Analyse in einem spezifischen Unternehmensbereich durchgeführt wurde, konnten wichtige Einflussfaktoren (gestellte Aufgaben und verwendete Normen der Auswahltests und WP-Examen) parallelisiert werden. Darüber hinaus kommt aus diesem Unternehmensbereich der insgesamt größte Anteil potenzieller Wirtschaftsprüfer, sodass die Ergebnisse eine hohe Relevanz für das Unternehmen besitzen. Allerdings reduziert dieser Umstand die Generalisierbarkeit der Ergebnisse auf andere Unternehmensbereiche und über die Unternehmensgrenzen hinaus. Die Datenanalyse wurde in weiten Teilen auf das dichotome Merkmal (bestandene oder nicht bestandene WP-Prüfung) beschränkt, um dem praktischen Nutzen für das Unternehmen Rechnung zu tragen,

der aus einem bestandenen WP-Diplom resultiert und Unterschiede in der Abschlussnote weit übertrifft. Eine gesonderte Analyse der Daten (siehe Tabelle 1 auf Seite 56) des intervallskalierten Notenkriteriums bestätigte die hier berichteten Ergebnisse und ergab erwartungsgemäß eine leicht erhöhte prognostische Validität der verwendeten Prädiktoren gegenüber dem dichotomen Kriterium. Die berichtete prognostische Validität der eingesetzten Testverfahren für die WPPrüfung wurde infolge einer Sekundäranalyse, also auf Grundlage der allgemeinen Einstellungstests, berechnet. Eine tätigkeitsspezifische Konstruktion eines Hybridtests (Kombination aus Intelligenztest und Arbeitsprobe) in Verbindung mit ausgewählten Persönlichkeitsverfahren oder einem spezifischen Auswahlinterview könnte die Prognosekraft der psychologischen Testverfahren weiter steigern.

Weiterführende Literatur

- Barrick, M. R., Mount, M. K. & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, 9, 9–30.
- Frintrup, A. & Piechowski, S. (2011). Effiziente Prozesse der Mitarbeiterrekrutierung: Integriertes E-Recruiting. *Industrie Management*, 27, 19–22.
- Hunter, J. E. & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA: Sage.
- Schmidt, F. L. & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, 86, 162–173.