

Zur prognostischen Güte von Zulassungskriterien
im Psychologiestudium für Studienerfolgsindikatoren

Prognostic Validity of Admission Criteria
for Academic Success in Psychology Programs

Entwurfsfassung: 21. August 2017. Diese Entwurfsfassung entspricht womöglich nicht vollständig der final gesetzten Version, welche im Hogrefe Verlag erschienen ist.

Zu zitieren als: Janke, S. & Dickhäuser, O. (in Druck). Zur prognostischen Güte von Zulassungskriterien im Psychologiestudium für Studienerfolgsindikatoren. *Psychologische Rundschau*.

Stefan Janke & Oliver Dickhäuser

Universität Mannheim

Autorenhinweis

Stefan Janke, Universität Mannheim, Fachbereich Psychologie, 68131 Mannheim;

Oliver Dickhäuser, Universität Mannheim, Fachbereich Psychologie, 68131 Mannheim.

Wir danken Daniel Krauß und Frauke Fried für die Aufbereitung der Daten.

Zusammenfassung

Grundständige Bachelorprogramme im Fach Psychologie unterliegen deutschlandweit Zulassungsbeschränkungen. Um geeignete Bewerber/-innen auszuwählen verwenden Universitäten innerhalb des jeweiligen gesetzlichen Rahmens eine Vielzahl an Zulassungskriterien. Anhand von Zulassungsdaten einer deutschen Universität ($N = 10\,605$ Bachelorbewerber/-innen aus fünf Kohorten) untersuchten wir die prognostische Güte von Zulassungskriterien (Note der Hochschulzugangsberechtigung [HZB], einschlägige Einzelnoten sowie praktische Erfahrungen) für Studienerfolgsindikatoren (Note in der ersten Statistikklausur, Abschlussnote und Studiendauer). Pfadanalysen zeigten, dass die durchschnittliche Note der HZB positiv prädiktiv für die durch Noten operationalisierten Studienerfolgsindikatoren war und an diesen Kriterien auch in bedeutsamen Ausmaß Varianz aufklärte. Die Einzelnoten aus den Fächern Mathematik und Englisch im Abitur zeigten genauso wenig inkrementelle Vorhersagekraft wie einschlägige praktische Erfahrung. Gesetzeskonforme Veränderungen in der Gewichtung der Kriterien zugunsten der HZB verbesserten die Gesamtprognosegüte. Hinsichtlich der Studiendauer erwiesen sich die vier betrachteten Zulassungskriterien als nicht prädiktiv. Der Beitrag illustriert den Nutzen evidenzbasierten Handelns bei der Studienzulassung und diskutiert gleichzeitig Grenzen eines solchen Vorgehens in den verschiedenen Bundesländern.

Schlagwörter: Studierendenauswahl, Zulassungskriterien, Studienerfolg, Hochschulzugangsberechtigung, Bachelorstudium

Abstract

In the presented study, we investigated the predictive power of three kind of admission criteria used to select university attendees, who enlisted to study in German bachelor programs in the field of psychology. The investigated admission criteria were grade point average, single grades in Mathematics and English as well as practical experiences in jobs related to the field of psychology. We used the data of 10 605 university applicants from five bachelor cohorts of a German university in order to investigate whether the different admission criteria would differentially predict academic success operationalized by the first grade in statistics at university, the final grade point average of the bachelor degree and the duration of studies. Path analyses showed that grade point average was the only valid predictor for indicators of academic success indicated by grades, while none of the admission criteria were predictive for the duration of studies. Furthermore, we found that a stronger weighting of grade point average that simultaneously weakened the influence of other admission criteria on the admission process also increased the predictive power of the predictor set for graded academic success. In sum, this study illustrates the benefits of evidence-based university admission procedures and discusses the limits of such procedures set by federal law.

Key words: university admission, grade point average, academic success, admission criteria

Psychologie ist ein ausgesprochen beliebtes Studienfach. In den vergangenen Jahren überstieg die Anzahl an Studieninteressierten jeweils deutlich die Anzahl der an deutschen Präsenzüniversitäten im grundständigen Bachelor Psychologie zur Verfügung stehenden Studienplätze. Margraf (2015) gibt in seinem Bericht zur Lage der Psychologie ein Verhältnis von über acht Bewerbenden je Studienplatz an. Mit dieser Nachfragequote ist die Psychologie beliebter als beispielsweise Fächer wie Biologie, Medizin oder Pharmazie. Die große Nachfrage hat zur Folge, dass das Studium in grundständigen Bachelorstudiengängen der Psychologie an allen deutschen Präsenzüniversitäten einer örtlichen Zulassungsbeschränkung unterliegt. Dies macht eine Personenselektion im Rahmen von Zulassungsverfahren erforderlich. Eine solche Auswahl soll den Nutzen für das ausgewählte Individuum (z.B. mit Blick auf das Vermeiden von Frustrationserlebnissen) und die auswählende Institution (z.B. zur Vermeidung unnötiger Investitionen durch Studienabbruch) maximieren. Die Psychologie selbst kann als Forschungsdisziplin mit dazu beitragen, die Güte des Zulassungsverfahrens durch die Identifikation valider Zulassungskriterien zu verbessern. Hinsichtlich der praktischen Umsetzung entsprechender Forschungsergebnisse sind allerdings gewisse Grenzen gesetzt, da sich nicht nur die empirische Frage stellt in welchem Ausmaß Auswahlkriterien tatsächlich geeignet sind um Studienerfolg vorherzusagen, sondern auch die normative Frage ob bestimmte Auswahlkriterien gesellschaftlich opportun sind. Die Antwort auf die letztgenannte Frage wird im Regelfall im demokratischen Diskurs ausgehandelt und dann ultimativ von gesetzgebenden Instanzen entschieden.

Grundsätzlich ist die Zulassungsbeschränkung von Studiengängen in Deutschland nur innerhalb klar bestimmter gesetzlich festgelegter Grenzen möglich. Dies ist darin begründet, dass der freie Zugang zu Hochschulen (gegeben die formale Qualifikationen) ein hochgeschätztes und grundgesetzlich gesichertes Gut darstellt (GG Art. 12). Die Gesetzeslage, welche den Rahmen der Hochschulzulassung regelt ist entsprechend komplex. So hat der Bund im Rahmen eines Staatsvertrages über die Vergabe von Studienplätze festgelegt, dass

ein bestimmter Anteil der Studienplätze (maximal 20 Prozent) im Rahmen sogenannter Vorabquoten für bestimmte Bewerber/-innen (z.B. Härtefälle, Zweitstudium) vergeben wird. Der verbleibende Anteil wird über sogenannte Hauptquoten vergeben, welche dem jeweiligen Landesrecht unterliegen. In der Regel wird der kleinere Anteil der Studienplätze nach Wartezeit und/oder rein nach Abschlussnote der Hochschulzugangsberechtigung (HZB) vergeben, während der größere Anteil auf Grundlage eines hochschuleigenen Auswahlverfahren vergeben wird¹. Gegeben diese gesetzlichen Grundlagen der Auswahlverfahren erscheint es insbesondere sinnvoll die hochschuleigenen Auswahlverfahren evidenzbasiert zu optimieren, da der Gesetzgeber für diese eine gewisse Flexibilität bei der Verwendung von Auswahlkriterien erlaubt und die Auswahl über dieses Verfahren die überwiegende Mehrheit der Studienanfänger/-innen betrifft.

Auch der Optimierung hochschuleigener Auswahlverfahren sind allerdings landesrechtliche Grenzen gesetzt. Die Bundesländer setzen dabei unterschiedlich enge bzw. weite Grenzen für die Selektion von Bewerber/-innen (vgl. Tabelle 1). Während das Bundesland Berlin keinerlei gesetzliche Vorgaben macht, definieren die anderen Bundesländer in der Regel sehr klar, welche Kriterien zulässig sind und teilweise auch in welchem Ausmaß diese Kriterien Anwendung finden sollen. Die zulässigen Kriterien sind

¹ Die Bundesländer unterscheiden sich in diesem Punkt teilweise recht stark. Während in Baden-Württemberg, Bayern und Hamburg insgesamt 90 Prozent der nach Abzug der Vorabquoten übrigen Studienplätze durch die Hochschulen vergeben werden sollen, ermöglicht das Berliner Hochschulgesetz als einziges Landesrecht theoretisch sogar den vollständigen Verzicht auf ein hochschuleigenes Auswahlverfahren. Die Hochschulzulassungsgesetze der anderen Bundesländer sortieren sich zwischen diesen Extremen und sehen vor, dass zwischen 50 – 80 Prozent der Studienplätze im Rahmen eines hochschuleigenen Auswahlverfahren vergeben werden.

größteilig Leistungsindikatoren, die entweder Bezug zur Schullaufbahn haben (Note der HZB, fachlich einschlägige Einzelnoten) oder über standardisierte Verfahren (z.B. Studierfähigkeitstests) die Studieneignung feststellen sollen. Diese Indikatoren werden der weit verbreiteten normativen Vorstellung gerecht, dass Studienplätze nach einem Leistungsprinzip vergeben werden sollen (Klomfaß, 2011; Kreckel, 2015). Welche der bezeichneten Kriterien verwendet werden, überlassen die meisten Hochschulzulassungsgesetze größteilig den Hochschulen. Wie Tabelle 1 zeigt, gibt es jedoch mitunter die gesetzliche Einschränkung, dass mehrere Auswahlkriterien zu verwenden sind (drei Bundesländer) und/oder dass die Auswahl überwiegend auf der HZB-Note basieren soll (vier Bundesländer). Besonders restriktiv ist das Hochschulzulassungsgesetz des Landes Baden-Württemberg, welches verlangt, dass mindestens ein Zulassungskriterium außerschulisch sein muss (also z.B. praktische Erfahrungen, Einstellungsinterview).

Die prognostische Güte verschiedener gesetzlich vorgesehenen Kriterien für die Vorhersage von Studienerfolg im Allgemeinen (siehe insbesondere Schuler & Hell, 2008 für einen umfassenden Überblick) und im Psychologiestudium im Besonderen ist in der Vergangenheit verschiedentlich Gegenstand von wissenschaftlichen Debatten (vgl. etwa das Diskussionsforum im Heft 2 der Zeitschrift „Psychologische Rundschau“ des Jahres 2005) sowie empirischer Untersuchung gewesen (für eine beispielhafte Übersicht zu empirischen Studien siehe Troche, Mosimann & Rammsayer, 2014). So untersuchten Steyer, Yousfi und Würfel (2005) beispielhaft den Zusammenhang zwischen Abitur-Gesamtnoten und Abitur-Fachnoten mit Maßen des Studienerfolgs bei 694 Neu-Immatrikulierten der Jahrgänge 1995 – 2001 im Fach Psychologie (Diplom) der Universität Jena. Dabei wies die Abiturnote sowohl zur Gesamtleistung (Durchschnittsnote) im Vordiplom ($r = .39$) wie im Diplom ($r = .28$) die engste Beziehung auf. Fachlich einschlägige Einzelnoten aus dem Abiturzeugnis waren hingegen kaum inkrementell mit den Durchschnittsnoten im (Vor-)Diplom korreliert. Diese Befunde konnten im Wesentlichen auch von Wedler, Troche und Rammsayer (2008) repliziert

werden, welche erneut die überlegene Prädiktionskraft der HZB-Note im Verhältnis zu zahlreichen Einzelnoten für die Leistung in Vor- und Hauptdiplom, sowie die Dauer des Grundstudiums nachweisen konnten. Die vorauslaufenden empirischen Arbeiten sind allerdings vor allem durch zwei Aspekte in ihrer Aussagekraft für die aktuelle Debatte limitiert: Erstens beziehen sich die dargestellten Arbeiten auf die Auswahl von Studierende für Diplomstudiengänge. Die angeführte Empirie kann dementsprechend nur bedingt für die Aussagen zur Auswahl von Bachelorstudierenden herangezogen werden. Zweitens werden die gesetzlichen Rahmenbedingungen häufig mit Blick auf die Analysen nicht berücksichtigt. Dies gilt besonders wenn HZB-Note und Einzelnoten gleichgewichtet in Regressionsgleichungen eingebracht werden, was mit Hochschulzugangsgesetzen, die eine höhere Gewichtung der HZB verlangen, nicht konform ist.

Eine Ausnahme von den vorher genannten Limitationen stellt die Studie von Formazin, Schroeders, Köller, Wilhelm und Westmeyer (2011) dar, in welcher die Güte zusätzlicher Studierfähigkeitstests in einem Bachelorzulassungsverfahren im Land Berlin geprüft wurde. Hier zeigte sich, dass die Testleistung neben der HZB-Note inkrementell Varianz an der Studienleistung nach zwei Studienjahren im Bachelorstudium aufklären kann. Empirische Forschung, welche die Diversität der gesetzlichen Vorgaben berücksichtigt und dabei auch endgültigen Studienerfolg (Bachelornote und Studiendauer) bei Bachelorstudierenden in den Blick nimmt, steht allerdings auch in Folge dieser Studie noch aus.

Gegenstand der im Folgenden beschriebenen Untersuchung ist es entsprechend herauszufinden, wie sich die Berücksichtigung verschiedener gesetzlich möglicher Zulassungskriterien (untersucht wurden HZB, Einzelnoten, praktische Erfahrungen) auf die Zulassungspraxis und -güte bei der Auswahl von Bachelorstudierenden im Fachbereich Psychologie auswirken kann. Dabei soll betrachtet werden, ob die verschiedenen gesetzlich vorgesehenen Zulassungskriterien auch tatsächlich inkrementell bedeutsam für die Vorhersage

von Studienerfolg sind. Anschließend wird geprüft, wie verschiedene gesetzlich mögliche Gewichtungen der Zulassungskriterien sich in ihrer Vorhersagekraft für den Studienerfolg unterscheiden.

Methode

Stichprobe

Zur Untersuchung der Forschungsfragen wurden die Daten aus dem hochschuleigenen Auswahlverfahren einer baden-württembergischen Universität für die Bachelorjahrgänge 2007 bis 2011 im Fach Psychologie verwendet. Insgesamt lagen Angaben bezüglich der örtlich eingesetzten Zulassungskriterien für 10 605 Bewerber/-innen vor. Zusätzlich zu den Zulassungsdaten lagen uns die Angaben zu Studienerfolgsindikatoren für 411 eingeschriebenen Studierende der entsprechenden Kohorten vor. Eine Zuordnung von Zulassungsdaten zu diesen Daten war für 298 Studierende möglich (72.5% der Stichprobe). Dabei ist zu bedenken, dass eine vollständige Zuordnung auch deshalb nicht möglich war, weil ein Teil der Studierenden unabhängig vom hochschuleigenen Zulassungsverfahren auf Basis von Vorabquoten und landesspezifischen Hauptquoten vergeben werden (siehe Einleitung). Für die ersten drei erfassten Kohorten lagen uns zusätzlich Angaben zum Geschlecht der Bewerber/-innen vor. In dieser Teilstichprobe zeigte sich, dass 80.8 Prozent der Studienbewerber/-innen, sowie 75.4 Prozent der zugelassenen Studierenden weiblichen Geschlechts waren. Weitergehende demographische Daten zur genaueren Beschreibung der Stichprobe lagen uns auch auf Grund von Datenschutzbestimmungen nicht vor.

Modellparameter

Uns lagen Informationen zu allen an der entsprechenden Universität verwendeten Zulassungskriterien vor. Dabei handelt es sich im Einzelnen um die *HZB-Note*, die letzten Einzelnoten in *Englisch* und *Mathematik*, sowie *Zusatzpunkte für vor Studienbeginn absolvierte Praktika und einschlägige Berufserfahren (Praxispunkte)*. Die Praxispunkte wurden von der örtlichen Zulassungsstelle anhand eines Auswertungsschemas ermittelt,

welches Informationen zu Berufsbereich und Zeitdauer der Tätigkeit zu einem eindeutigen Punktwert zwischen 0 und 5 Punkte aggregiert. Um diese Praxispunkte vergleichbar zu den anderen Kriterien zu halten, wurden sie von uns zu einer Schulnotenskala von 1 bis 6 transformiert, wobei bei 5 Praxispunkten die Note 1 und bei 0 Punkten die Note 6 kodiert wurde.²

Als Studienerfolgsindikatoren wurden die *Note in der ersten Statistiklausur*, die *Studiendauer in Fachsemestern* und die *Abschlussnote* herangezogen. Einzelne zugelassene Studierende wiesen auf verschiedenen Kriterien keine Werte auf, weil sie über ihren Studienverlauf beispielsweise die Universität oder das Studienfach gewechselt oder gar das Studium ganz abgebrochen hatten. Der Grund für das Ausscheiden aus dem Bachelorprogramm ließ sich auf Basis der vorliegenden Daten allerdings nicht feststellen. Hinsichtlich der Statistiklausur fehlten nur für eine Person Angaben, während sich für 16 Studierende keine Angaben zur Endnote oder Gesamtstudiendauer machen ließen.

In Tabelle 2 finden sich die deskriptiven Daten bezüglich Zulassungskriterien und Studienerfolgsindikatoren für die zugelassenen Studierenden, sowie eine Korrelationsmatrix, welche beide Variablengruppen in Bezug zueinander setzt. Der Mittelwert der HZB illustriert,

² In den ersten beiden erfassten Kohorten wurde seitens der Zulassungsstelle davon abgesehen für verhältnismäßig rangniedrige Bewerber/-innen Praxispunkte zu ermitteln. Dies bedeutet, dass die in den Rohdaten enthaltenen Praxispunkte auf den unteren Rangplätzen nicht valide waren, da ohne Prüfung der angegebenen Praktika jeweils null Praxispunkte vergeben wurden. Der genaue Rangplatz ab dem die Praxispunkte nicht mehr ermittelt wurden, ließ sich aus den Daten nicht direkt entnehmen. In Konsequenz wurde der Rangplatz der letzten Bewerber/-innen, welchen Praxispunkte zugesprochen wurden als Approximation für den Abbruch der Beurteilung praktischer Tätigkeit verwendet. Die Praxispunkte auf den folgenden Rangplätzen wurden als fehlende Daten kodiert.

dass vorwiegend Studierende zugelassen wurden, welche auch schon an der Schule besonders leistungsstark waren.

Vorgehen

Vorhersagekraft gesetzlich vorgesehener Zulassungskriterien

Zunächst wurde überprüft, wie prädiktiv die einzelnen Zulassungskriterien für die verschiedenen Studienerfolgsindikatoren waren. Dabei wurde mittels Mplus Version 7.2 (Muthén & Muthén, 1998-2012) ein umfassendes Pfadmodell unter Einbezug aller Zulassungskriterien und Studienerfolgsindikatoren gerechnet. In dem zu Grunde liegenden Strukturmodell wurden außerdem alle Korrelationen zwischen Prädiktoren und Kriterien freigesetzt, da gerade auf Seite der Prädiktoren von Interkorrelationen auszugehen war (bspw. gehen die verwendeten Einzelnoten auch in die Abiturnote ein). Die Schätzung der Modellparameter in dem resultierenden saturierten Pfadmodell erfolgte durch Einsatz des MLR-Schätzers.

Bedingt durch die Zulassungspraxis lagen nur für einen Bruchteil der Studieninteressenten auch tatsächlich Daten zu Studienerfolg vor (2.8 %). Da nur sehr leistungsstarke Bewerber/-innen zum Studium zugelassen wurden, musste von einer deutlichen Varianzeinschränkung auf Seite der Studienerfolgsindikatoren in unserer Stichprobe ausgegangen werden. Um die Unterschätzung der prädiktiven Kraft von Studienauswahlkriterien zu vermeiden, verwendeten wir die von Formazin und Kollegen (2011) vorgeschlagene Methode zur Schätzung der populationsbezogenen Vorhersagekraft von Studienzulassungskriterien. Dabei werden die Populationsparameter unter Einbezug der fehlenden Werte mit Hilfe des Full-Information-Maximum-Likelihood- Algorithmus (FIML) geschätzt (Graham, 2009). Dieses Verfahren führt nach Formazin und Kollegen (2011) insbesondere dann zu einer akkuraten Schätzung der Populationsparameter, wenn das Fehlen der Werte nicht in der Variablen selbst, sondern höchstens in anderen Variablen begründet ist (Missing at random; entsprechend Little & Rubin, 2002). Davon ist bei den vorliegenden

Daten auszugehen, da das Fehlen von Studienerfolgsindikatoren zwar in der fehlenden Studienzulassung, jedoch eher nicht in der potentiellen Ausprägung dieser Indikatoren begründet liegt.

Um die Modellschätzung unter Berücksichtigung fehlender Daten zu verbessern, ist es nach Graham, Olchowski und Gilreath (2007) ratsam möglichst viele Hilfsvariablen einzuschließen. Diese Variablen werden nicht in das Strukturmodell eingeschlossen, sondern nur zur Bestimmung der fehlenden Werte eingesetzt (Asparouhov & Muthén, 2008). Auf Grund des umfassenden Datenschutzes lagen uns als Hilfsvariablen nur das Geschlecht sowie die Kohortenzugehörigkeit der Bewerber/-innen vor. Beide Variablen wurden als Hilfsvariablen zur akkurateren Parameterschätzung verwendet, wobei die Kohortenzugehörigkeit in vier Dummyvariablen kodiert wurde. Zusätzlich zur Schätzung der populationsbezogenen Pfadkoeffizienten wurde außerdem ein Pfadmodell in der Teilstichprobe der 298 zugelassenen Bewerber/-innen gerechnet. Diese Zusatzanalyse sollte dabei helfen, die grundsätzliche Robustheit des Zusammenhangsmusters auf der Basis größtenteils geschätzten Daten zu beurteilen.

Vorhersagegüte verschiedener Gewichtungsmodelle im gesetzlichen Rahmen

Im Rahmen der folgenden Analysen wurden drei verschiedene Gewichtungsmodelle hinsichtlich ihrer prädiktiven Vorhersagekraft überprüft. Gewichtungsmodell 1 sieht eine maximale Gewichtung der HZB (99%), bei minimaler Gewichtung der Einzelnoten (jeweils 0.5 %) vor (HZB-Modell). Gewichtungsmodell 2 sieht hingegen vor, dass die HZB nur mit 51% gewichtet, während die Einzelnoten zu jeweils 24.5% in das Modell eingebracht wurden (HZB/Einzelnoten-Modell). Das HZB- und das HZB/Einzelnoten-Modell sind mit fast allen Landesbestimmungen zur Zulassungspraxis an Universitäten konform. Dies ist der Fall, da beide Gewichtungsmodelle – wie in vielen Bundesländern gefordert – auf mehreren Kriterien beruhen und – wiederum wie in einigen Bundesländern legislativ vorgegeben – die HZB als maßgebliches Kriterium einbeziehen. Gewichtungsmodell 3 folgte der lokal üblichen

Zulassungspraxis, welche der Hochschulzugangsberechtigung das größte Gewicht zuweist (80% des Gesamtscores), während jede Einzelnote und die Zusatzpunkte für absolvierte Praktika gleich gewichtet (jeweils 6.67% des Gesamtscores) einfließen (Plus-Praxis-Modell). Dieses Gewichtungsmodell unterscheidet sich somit von den anderen Ansätzen dadurch, dass noch ein zusätzliches außerschulisches Kriterium verwendet wurde (Erfahrung durch Praktika). Damit wird sie auch den restriktivsten Zulassungsbestimmungen (Hochschulzulassungsgesetz Baden-Württemberg) gerecht.

Auch in diesem Schritt wurden die fehlenden Werte mit Hilfe des FIML-Algorithmus in drei unabhängigen saturierten Pfadanalysen geschätzt. Die Kohortenzugehörigkeit und das Geschlecht wurden erneut als Hilfsvariable einbezogen. Die Varianz des Gewichtungsfaktors wurde auf einen in vorauslaufenden Analysen ermittelten Wert fixiert, da eine Schätzung von Pfadkoeffizienten unter Berücksichtigung fehlender Werte mit Hilfe des FIML-Algorithmus nicht für manifeste exogene Variablen ohne zusätzliche Modellannahmen (bspw. Varianzfixierung oder Schätzung von Kovarianzen) möglich ist (Arbuckle, 1996, Geiser, 2012). Um die Robustheit der Analysen zu prüfen, wurden drei weitere Pfadmodelle in der Teilstichprobe der zugelassenen Bewerber/-innen gerechnet.

Ergebnisse

Inkrementelle Prädiktivität der unterschiedlichen Kriterien für Studienerfolg

Das Pfadmodell zur Ermittlung der Vorhersagekraft der Zulassungskriterien zeigte sowohl unter Einbezug aller Studienbewerber/-innen, als auch bei konservativerer Testung in der Teilstichprobe der schlussendlich zugelassenen Studierenden das gleiche Muster: In beiden Fällen erwies sich einzig die Abiturnote als positiv prädiktiv für den Studienerfolg gemessen durch Noten (Note in der ersten Statistiklausur, Bachelorendnote). Die anderen Zulassungskriterien (Einzelnoten und Praxispunkte) wiesen hingegen keine zusätzliche prädiktive Güte auf. Hinsichtlich der Studiendauer zeigte sich zwar ein signifikantes beta-Gewicht für die Englischnote in beiden Pfadmodellen, jedoch keine signifikante

Varianzaufklärung. Die Pfadkoeffizienten der verschiedenen Zulassungskriterien und die Varianzaufklärung an den verschiedenen Studienerfolgsindikatoren sind in Tabelle 3 abgetragen³.

Auswirkung der Kriteriengewichtung auf die Vorhersagbarkeit von Studienerfolg

Unterschiedliche Gewichtungen der Zulassungskriterien zeigten einen klaren Effekt auf die Vorhersagegüte des Gesamtmodells. Die Vorhersagegüte stand dabei in klarer Abhängigkeit von der Stärke der Gewichtung der HZB. Bei stärkerem Einfluss der HZB auf die Auswahl der Studierenden, stieg die Varianzaufklärung für die durch Noten erfassten Studienerfolgsindikatoren. Dies zeigte sich insbesondere, wenn die Vorhersagekraft des HZB/Einzelnoten-Modells (51% Gewichtung HZB) mit der Vorhersagekraft des HZB-Modells (99% Gewichtung HZB) verglichen wurde ($\Delta R^2 = .01 - .03$). Die Tatsache, dass das Plus-Praxis-Modell in der Gesamtstichprobe, trotz verhältnismäßig hoher Gewichtung der HZB, eine vergleichsweise schlechte Vorhersagekraft zeigte, liegt höchstwahrscheinlich in der unterschiedlichen Stichprobengröße begründet. Da für 2085 Bewerber/-innen Angaben zu den Praxispunkten fehlen, basieren die Schätzung lediglich auf einer Teilstichprobe von 8520 Bewerber/-innen. Der robustere Vergleich in der Teilstichprobe der zugelassenen Bewerber/-innen verdeutlicht, dass das Plus-Praxis-Modell in seiner Vorhersagekraft für durch Noten

³ Im Rahmen der vorgestellten Analysen wurde die letzte im Zeugnis vermerkte Englisch- bzw. Mathematiknote als Zulassungskriterium verwendet, wie auch im Rahmen des lokalen Zulassungsverfahrens. Für vier von fünf Kohorten war es allerdings auch möglich gemittelte Noten für die Oberstufe zu bilden, deren Eignung als Zulassungskriterium ebenfalls geprüft wurde. Die Prüfung auf der Basis dieser Daten trägt der Überlegung Rechnung, dass gemittelte Noten reliabler (vgl. Trapmann, Hell, Weigand & Schuler, 2007)) und deshalb möglicherweise stärker prognostisch valide sind als Einzelnoten. Die Verwendung gemittelter Noten veränderte das Gesamtmuster der Befunde allerdings nicht bedeutsam.

gemessene Studienleistung zwischen dem HZB- und dem HZB/Einzelnoten-Modell liegt. Keines der Gewichtungsmodelle konnte statistisch bedeutsam Varianz an der Studiendauer aufklären. Die genaue Vorhersagegüte der verschiedenen Gewichtungsmodelle für die Gesamtstichprobe, sowie die Teilstichprobe der zugelassenen Studierenden sind in *Tabelle 4* abgetragen. Die Bedeutsamkeit der Befunde wird auch dadurch gestärkt, dass verschiedene Gewichtungsmodelle deutliche Konsequenzen für die Zusammensetzung der Studierendenschaft haben. So wären im Mittel über die Jahre hinweg bei einer starken Gewichtung der Einzelnoten (HZB/Einzelnoten-Modell) lediglich 71.54 Prozent der Studierenden zugelassen worden, die auch bei starker Gewichtung der HZB zugelassen worden wären (HZB-Modell).

Diskussion

Die Ergebnisse unserer Analysen bestätigen und erweitern die Befunde vorangegangener Studien zur Eignung von Einzelnoten als zusätzliches Kriterium für die Studierendenauswahl neben der HZB-Note. Einzelnoten erwiesen sich dabei als nicht inkrementell valide für den Studienerfolg von Bachelorstudierenden. Im Gegenteil: Das Ausmaß der an Studienerfolgsmaßen aufgeklärten Varianz scheint sogar zu sinken, je stärker Einzelnoten Berücksichtigung finden. Ein Grund hierfür könnte sein, dass die HZB-Note als über zwei Jahre und zahlreiche Fächer aggregierter Prädiktor reliabler ist als Einzelnoten, welche stärker systematischen Verzerrungen auf Ebene der beurteilenden Lehrkraft unterliegen (Trapmann, et al., 2007). Es lässt sich entsprechend schlussfolgern, dass eine verstärkte Berücksichtigung der Einzelnoten klar gegen das Leistungsprinzip (kompetentere Studierende erhalten wahrscheinlicher den Studienplatz) als Grundlage der Studierendenauswahl verstößt. Berufspraktische Erfahrungen tragen ebenfalls in keinerlei Weise inkrementell zur Vorhersage von Studienerfolgsmaßen bei.

Weitergehend können wir die Befunde aus vorherigen Studien mit Diplomstudierenden nicht replizieren, welche zeigen konnten, dass die HZB-Note auch für

die Studiendauer prädiktiv ist (Wedler, 2008). Hierbei stellt sich allerdings die Frage ob dieses Studienerfolgskriterium tatsächlich noch zeitgemäß ist, da die Studiendauer in Bachelorstudiengängen in der Regel begrenzt ist (sog. Maximalstudiendauer). Während im Diplomstudium recht lange Studiendauern erreicht werden konnten, ist dies in Bachelorprogrammen nicht mehr möglich. In letzteren kann eine längere Studiendauer im Rahmen der Maximalstudienzeit außerdem auch ein (erwünschtes) Auslandssemester reflektieren, weshalb fraglich ist, ob in Bachelorprogrammen kürzere Studiendauer tatsächlich größeren Studienerfolg anzeigt.

Auch die HZB-Note als Prädiktor ist wahrscheinlich in ihrer Reliabilität und Validität begrenzt, u.a. wegen der fehlenden Vergleichbarkeit über Bundesländer hinweg (siehe auch Rindermann, 2005). Entsprechend wäre eine zentrale Studienplatzvergabe auf Basis von Länderquoten möglicherweise zielführender als hochschuleigene Verfahren, welche die mangelnde Vergleichbarkeit der Abiturnote zwischen den Bundesländern meist völlig ausblenden. Schlussendlich könnte auch die Suche nach prädiktiven leistungsbezogenen Verfahren abseits der HZB-Note zielführend sein, wie beispielsweise von Formazin und Kollegen (2011) vorgeschlagen, welche den Einsatz von Studierfähigkeitstests empfehlen. In diesem Punkt sollten aber die Praktiker an den Hochschulen bedenken, dass Chancengleichheit ebenfalls eine normative Stellgröße für Auswahlverfahren sein kann. Sollten flächendeckend Studierfähigkeitstests eingeführt werden, müsste überlegt werden, wie sich Anreise- und Durchführungskosten so gestalten lassen, dass kein weiterer struktureller Nachteil für Studienanfänger aus weniger finanzkräftigen Elternhäusern entsteht.

Die vorliegende Arbeit fokussiert vor allem auf Zulassungskriterien, welche der Überzeugung gerecht werden, dass besonders leistungsstarke Bewerber/-innen bei der Studierendenauswahl bevorzugt werden sollten (Leistungsprinzip; Klomfaß, 2011; Kreckel, 2015). Es wäre allerdings auch denkbar Studierende nach einem Interessensprinzip auszuwählen, welches vorsieht, dass besonders für das Studienfach motivierte Bewerber/-

innen zugelassen werden. Dies wäre besonders im Hinblick darauf sinnvoll, dass persönliche Interessen in direktem Bezug zum Aufbau domänenspezifischer Kompetenzen stehen (Ackerman, 1996). Des Weiteren ist davon auszugehen, dass das studentische Wohlbefinden und damit womöglich auch die Studienpersistenz in starkem Ausmaß von dem Fit zwischen Eigeninteressen und Studieninhalten abhängen (Le, Robbins & Westrick, 2014).

Die meisten Hochschulzulassungsgesetze ermöglichen die Verwendung von Zulassungskriterien, von denen sich vermuten lässt, dass sie einem solchen Interessensprinzip Rechnung tragen. So könnte man die Berücksichtigung von berufspraktischen Vorerfahrungen als eine erste Annäherung an das Interessensprinzip betrachten, wenn man annimmt, dass gewählte Berufe und Praktika mit Bezug zu dem Studienfach fachliches Interessen indizieren. Die aktuellen Analysen zeigen jedoch, dass die Berücksichtigung dieses Kriteriums bei der Studienausswahl zu keinem substantiellen Ertrag in Bezug auf Leistungsindikatoren führt. Hieraus abzuleiten, dass eine Auswahl auf Basis des Interessensprinzips nicht zielführend sei, würde die Eignung von berufspraktischen Vorerfahrungen als Interessensindikator allerdings überschätzen. Besser wäre wahrscheinlich die Berücksichtigung differenzierter Interessensprofile beispielsweise mit Hilfe des RIASEC-Modells (Holland, 1985; siehe im Speziellen auch Le, Robbins & Westrick, 2014). Eine abschließende Evaluation von interessengeleiteten Zulassungsverfahren oder kombinierten Zulassungsverfahren, die sowohl auf Interessens- als auch auf Leistungsindikatoren setzen, steht dementsprechend noch aus.

Abschließend lässt sich sagen, dass eine Studierendenauswahl die vornehmlich auf Basis der HZB-Note geschieht, immer auch als Kosten-Nutzen-Abwägung betrachtet werden muss. Eine solche Auswahl ist wie dargestellt gesetzeskonform (wenn ein Minimum an Alternativkriterien einbezogen wird) und verhindert zusätzliche Kosten für Institutionen und Individuum. Nichts desto trotz muss auch festgestellt werden, dass diese Form von Zulassungsverfahren keinesfalls optimal ist. So wird sowohl das Wissen über die

unterschiedlichen Länderstandards hinsichtlich der HZB, als auch die vielversprechenden Erkenntnisse zu Studierfähigkeitstests ignoriert. Zusätzlich bleibt auch unklar inwiefern eine solche Auswahl sich auch auf andere Indikatoren von Studienerfolg jenseits der Hochschulnote auswirkt (z.B. wissenschaftliche Neugier, Befähigung zur Ausübung psychologischer Berufe). Aus unserer Sicht ergeben sich daraus Herausforderung für den Gesetzgeber und die psychologische Forschung. Die Forschung ist gefordert herauszufinden, welche Auswahlkriterien geeignet sind um über Leistung hinausgehenden Studienerfolg sicherzustellen (beispielsweise durch Entwicklung interessensgebundener Zulassungskriterien). Der Gesetzgeber ist hingegen in der Zwischenzeit gefragt die Vergleichbarkeit der HZB-Note über die Länder hinweg zu erhöhen oder zumindest die zentrale Zulassung von Studierenden auf Basis von Länderquoten wieder stärker zu befördern. Sollte der Gesetzgeber außerdem die Wichtigkeit der Hochschulautonomie in den Vordergrund stellen wollen, müssen die Universitäten finanziell auch so ausgestattet werden, dass sie die Studierendenauswahl mit Hilfe valider Zusatzkriterien (z.B. fachspezifische Studierfähigkeitstests) eigenständig durchführen können, ohne dass unzumutbare Kosten an die Bewerber/-innen weitergegeben werden. Solange dies nicht erfolgt, werden Universitäten wahrscheinlich auch in Zukunft auf Kriterien zurückgreifen, die zwar gesetzeskonform und kostengünstig sind, die Validität des Auswahlprozesses jedoch eher gefährden (z.B. Einzelnoten).

Literatur

- Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, 22, 227-257.
- Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G.A. Marcoulides & R. E. Schumacker (Hrsg.), *Advanced structural equation modeling: Issues and techniques* (S. 243-277), NJ: Lawrence Erlbaum Associates, Inc.
- Asparouhov, T. & Muthén, B. O. (2008). *Auxiliary variables predicting missing data*.
Abgerufen am 02. August, 2017 unter <http://www.statmodel.com/download/AuxM2.pdf>.
- Formazin, M., Schroeders, U., Köller, O., Wilhelm, O., & Westmeyer, H. (2011).
Studierendenauswahl im Fach Psychologie. *Psychologische Rundschau*, 62, 221–236.
- Geiser, C. (2012). *Data analysis with Mplus*. New York: The Guilford Press.
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60, 549-576.
- Holland, J. L. (1985). *Making vocational choices: A theory of vocational personalities and work environments*. Englewood Cliffs, NJ: Prentice-Hall.
- Klomfaß, S. (2011). Die Institution des Hochschulzugangs. In S. Klomfaß (Hrsg.), *Hochschulzugang und Bologna-Prozess* (S. 131-190). VS Verlag für Sozialwissenschaften: Wiesbaden.
- Kreckel, R. (2015). Struktur der Studierendenauswahl im expandierenden Hochschulsystem der Bundesrepublik Deutschland. In W. Helsper & H.-H. Krüger (Hrsg.), *Auswahl der Bildungsklientel* (S. 405-419). Springer Fachmedien: Wiesbaden.
- Le, H., Robbins, S. B., & Westrick, P. (2014). Predicting student enrollment and persistence in college STEM fields using an expanded PE fit framework: A large-scale multilevel study. *Journal of Applied Psychology*, 99, 915-947.
- Little, R. J. A. & Rubin, D. B. (2002). *Statistical analysis with missing data (2nd ed.)*. Hoboken, NJ: John Wiley & Sons.

- Margraf, J. (2015). Zur Lage der Psychologie. *Psychologische Rundschau*, 66, 1-30.
- Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus User's Guide. Seventh Edition*. Los Angeles, California: Muthén & Muthén.
- Rindermann, H. (2005). Für ein bundesweites Auswahlverfahren von Studienanfängern über Fähigkeitsmessung. *Psychologische Rundschau*, 56, 127-129.
- Schuler, H., & Hell, B. (2008). *Studierendenauswahl und Studienentscheidung*. Göttingen: Hogrefe.
- Steyer, R., Yousfi, S. & Würfel, K. (2005). Prädiktoren von Studienerfolg: Der Zusammenhang zwischen Schul- und Studiennoten im Diplomstudiengang Psychologie. *Psychologische Rundschau*, 56, 129-131.
- Trapmann, S., Hell, B., Weigand, S., & Schuler, H. (2007). Die Validität von Schulnoten zur Vorhersage des Studienerfolgs-eine Metaanalyse. *Zeitschrift für pädagogische Psychologie*, 21, 11-27.
- Troche, S., Mosimann, M., & Rammsayer, T. (2014). Die Vorhersage des Studienerfolgs im Masterstudiengang Psychologie durch Schul- und Bachelorstudienleistungen. *Beiträge zur Hochschulforschung*, 36, 30-45.
- Wedler, B., Troche, S., Rammsayer, T. (2008). Studierendenauswahl – Eignungsdiagnostischer Nutzen von Noten aus Schule und Studium. *Psychologische Rundschau*, 59, 123-125.

Tabellenverzeichnis

Tabelle 1

Vorgaben verschiedener Bundesländer zur Verwendung von Zulassungskriterien bei hochschuleigenen Auswahlverfahren.

	Note HZB	Einzelnoten	Test	Praxis*	Interview	Anmerkung
Baden-Württemberg	x	x	x	x	x	Mindestens zwei Kriterien; ein Kriterium muss außerschulisch sein
Bayern	x	a	a	a	a	Mindestens ein a-Kriterium. Note HZB mindestens gleichgewichtig
Berlin						Keine Vorgaben
Brandenburg	x	a	a	a	a	Mindestens ein a-Kriterium. Note HZB mindestens gleichgewichtig
Bremen	x	x	x	x	x	Eines oder mehrere Kriterien; Note HZB gibt maßgeblichen Ausschlag
Hamburg	x	x	x	x	x	Eines oder mehrere Kriterien
Hessen	x	x	x	x	x	Keine Vorgaben
Mecklenburg-Vorpommern	x	x	x	x	x	Eines oder mehrere Kriterien
Niedersachsen	x	b	b	b	b	Note HZB oder Note HZB plus b-Kriterien
Nordrhein-Westfalen	x	x	x	x	x	Keine weiteren Vorgaben
Rheinland-Pfalz	x	x	x	x	x	Keine weiteren Vorgaben
Saarland	x	x	x	x	x	Keine weiteren Vorgaben
Sachsen	x	x	x	x	x	Mindestens ein Kriterium
Sachsen-Anhalt	x	x	x	x	x	Einzelnen oder additiv
Schleswig-Holstein	x	x	x	x	x	Keine weiteren Vorgaben
Thüringen	x	a	a	a	a	Mindestens ein a-Kriterium. Note mindestens gleichgewichtig

Anmerkung: x, a, b: Kriterium vorgesehen. * Einschlägige praktische Erfahrungen / Berufsausbildung

Tabelle 2

Deskriptive Statistiken und Zusammenhänge von Zulassungskriterien und Studienerfolgsindikatoren für zugelassene Bachelorstudierende.

	<i>M</i>	<i>SD</i>	(1)	(2)	(3)	(4)	(5)	(6)
(1) HZB	1.34	0.47						
(2) Letzte Englischnote	1.63	0.88	.58**					
(3) Letzte Mathematiknote	1.78	0.97	.63**	.37**				
(4) Praxiserfahrungen	3.92	1.91	-.05	-.06	-.11			
(5) Erste Statistiklausur	2.70	0.73	.33**	.16**	.25**	-.05		
(6) Bachelorendnote	1.77	0.32	.41**	.25**	.27**	.00	.59**	
(7) Studiendauer	6.38	0.94	.03	.11	.05	.01	.06	.19**

Anmerkung: Bis auf die Studiendauer wird in Bezug auf alle angeführten Variablen eine Schulnotenmetrik eingesetzt (1 = „sehr gut“; 6 = „ungenügend“). Die Studiendauer wird in Fachsemestern bei Studienabschluss angegeben. ** $p < .001$

Tabelle 3

Regressionsgewichte verschiedener Zulassungskriterien für die Vorhersage von Studienerfolgsindikatoren.

Studienerfolgsindikator	HZB	Englischnote	Mathematiknote	Praxiserfahrungen	Aufgeklärte Varianz
Erste Statistiklausur	$\beta = .38/.27, p < .001/<.001$	$\beta = -.06/-.04, p = .417/.493$	$\beta = .09/.05, p = .358/.533$	$\beta = .00/.00, p = .982/.986$	$R^2 = .18/.08, p = .006/.050$
Bachelorendnote	$\beta = .48/.34, p < .001/<.001$	$\beta = .09/.08, p = .181/.161$	$\beta = .06/.03, p = .501/.680$	$\beta = .03/.04, p = .441/.523$	$R^2 = .32/.15, p < .001/.001$
Studiendauer	$\beta = -.10/-.08, p = .320/.246$	$\beta = .23/.16, p = .016/.013$	$\beta = .08/.05, p = .353/.447$	$\beta = .07/.07, p = .189/.239$	$R^2 = .05/.02, p = .153/.163$

Anmerkung: Die erste Zahl indiziert jeweils die Koeffizienten im Gesamtmodell, während die zweite Zahl (hinter dem Schrägstrich) die Koeffizienten in der Teilstichprobe der zugelassenen Studierenden indiziert.

Tabelle 4

Vorhersagekraft verschiedener Gewichtungsmodelle für Studienerfolgsindikatoren.

Studienerfolgsindikator	GM 1 (HZB = 99%)	GM 2 (HZB = 51%)	GM 3 (HZB = 80%)
Erste Statistiklausur	$R^2 = .17/.11, p = .006/.008$	$R^2 = .14/.09, p = .020/.020$	$R^2 = .11/.11, p = .015/.013$
Bachelorendnote	$R^2 = .30/.21, p < .001/<.001$	$R^2 = .28/.20, p < .001/<.001$	$R^2 = .21/.20, p < .001/<.001$
Studiendauer	$R^2 = .00/.00, p = .791/.793$	$R^2 = .01/.01, p = .506/.522$	$R^2 = .01/.00, p = .574/.685$

GM = Gewichtungsmodell

Anmerkung: GM 1 = 99% HZB, 0.5% Mathematiknote, 0.5% Englischnote (HZB-Modell); GM 2 = 51% HZB, 24.5% Mathematiknote, 24.5% Englischnote (HZB/Einzelnoten-Modell); GM 3 = 80% HZB, 6.67% Mathematiknote, 6.67% Englischnote, 6.67% Praxiserfahrung (Plus-Praxis-Modell). Die erste Zahl indiziert jeweils die Koeffizienten im Gesamtmodell, während die zweite Zahl (hinter dem Schrägstrich) die Koeffizienten in der Teilstichprobe der zugelassenen Studierenden indiziert.