

Berechtigte Hoffnung auf bessere Studierende durch hochschuleigene Studierendenauswahl?¹

Eine Analyse der Erfahrungen mit Auswahlverfahren in der Hochschulzulassung

Mareike Tarazona

Der Beitrag befasst sich mit empirischen Untersuchungen zu und Erfahrungen mit der Auswahl von Studierenden an deutschen Hochschulen. Entsprechend der Tendenz zur zunehmenden Profilierung von Hochschulen in Deutschland sowie der Sorge um Chancengleichheit und Gerechtigkeit beim Hochschulzugang ist es wichtig, dass Hochschulen Instrumente zur Studierendenauswahl anwenden, die diese Anforderungen berücksichtigen. Eine Vielzahl empirischer Untersuchungen zeigt jedoch, dass die den Hochschulen zur Verfügung stehenden Instrumente weder hinreichend in der Lage sind, später erfolgreiche Studierende zu identifizieren, noch Kriterien einfließen zu lassen, die der Profilbildung der jeweiligen Hochschule Rechnung tragen.

1 Einleitung

Seit wenigen Jahren dürfen Hochschulen einen Teil ihrer Studierenden selbst auswählen. Diese gesetzliche Neuerung ist bisher jedoch nur in wenigen Bundesländern in die regionalen Hochschulgesetze integriert worden. Von Universitäten aus diesen Ländern liegen nun die ersten Erfahrungen vor. Gleichzeitig ist die politische Debatte um den Hochschulzugang im Zuge der Diskussion um die Stärkung der internationalen Wettbewerbsfähigkeit deutscher Hochschulen neu entbrannt. So hat z. B. der Wissenschaftsrat Anfang 2004 eine Stellungnahme zur Neugestaltung des Hochschulzugangs veröffentlicht. Im Juni 2004 wurde im Bundestag beschlossen, dass Hochschulen in Deutschland künftig 60% ihrer Studierenden in Numerus-Clausus-Fächern selbst auswählen dürfen.

Der vorliegende Beitrag thematisiert erstens die aktuelle Debatte um die Hochschulzulassung, zweitens Kriterien für die Angemessenheit der Zulassungsverfahren und gibt schließlich einen Überblick empirischer Ergebnisse zur Validität der Instrumente sowie zum Einsatz der Instrumente in den hochschuleigenen Auswahlverfahren in Baden-

¹ Für kritische Anmerkungen danke ich Prof. Dr. Gerd Grözinger.

Württemberg. Aufgrund der Bedeutung der USA als Referenzsystem in der Hochschulforschung wird hier auch auf die Diskussion des Themas in den USA und in amerikanischen Studien Bezug genommen.

1.1 Zur Debatte in Deutschland

Die verschiedenen Stränge der Debatte in Deutschland, die in engem Zusammenhang mit den Hochschulzulassungsverfahren stehen, betreffen im Wesentlichen drei Bereiche: (mangelnde) Studierfähigkeit, Wettbewerb und Profilierung der Hochschulen sowie soziale Gerechtigkeit in der Hochschulzulassung.

Im Zentrum der Studierfähigkeitsdebatte steht der Wert des Abiturs. Das Recht auf Hochschulbildung wird in Deutschland in der Regel durch das Bestehen des Abiturs erworben. Löwer (1996) macht darauf aufmerksam, dass dies eine Prognoseentscheidung impliziert. Mit dem Abitur wird dem Schüler attestiert, dass er geeignet ist, ein Studium erfolgreich zu absolvieren, wobei der „prognostische Gehalt mit der Abiturnote variieren“ mag (a. a. O., S. 629). Mit anderen Worten: Durch den Erwerb des Abiturs wird dem Abiturienten Studierfähigkeit bescheinigt. Im Sinne einer allgemeinen Hochschulreife gilt dieser Befähigungsnachweis für alle Studiengänge. Eben diese Funktion des Abiturs wird jedoch von einigen Vertretern aus Hochschulen und Wirtschaft mit Skepsis betrachtet. Kritiker des derzeitigen Zulassungssystems plädieren dafür, das Abitur zwar als eine notwendige, aber nicht hinreichende Voraussetzung für das Studium zu sehen.

In einer Studie von Konegen-Grenier (2002) schätzten Hochschullehrer jeden Dritten ihrer Studierenden als nicht studierfähig ein. Die große Streuung der Einschätzungen deutet jedoch darauf hin, dass die Wahrnehmungen der Hochschullehrer kaum Generalisierungen zulassen. Dennoch beurteilen die Studienanfänger selbst ihre Studierfähigkeit kritisch (Lewin/Heublein/Sommer 2000), und ihre Einschätzung ist mit derjenigen der Hochschullehrer nahezu kongruent: Knapp ein Drittel fühlte sich durch die Schule nur schlecht oder unzureichend auf das Studium vorbereitet.

Die Debatte, ob und inwiefern das Abitur als Hochschulzugangsberechtigung im Sinne einer nicht hinreichenden Bedingung für das Studium „an Wert verlieren“ darf, wird stark normativ geführt. Manche Akteure in der hochschulpolitischen Diskussion sind besorgt um die Sozialverträglichkeit von Auswahlverfahren, zumal sich im Hochschulzugang gesellschaftliche Statusdistribution widerspiegelt (Wolter 1999). Andererseits ist bekannt, dass ein größerer sozialer Auswahlprozess schon vor der Bewerbung für einen Studienplatz stattfindet.

In den vergangenen Jahren wurden die Stimmen lauter, die fordern, das deutsche Hochschulsystem wettbewerbler zu gestalten und eine größere Differenzierung der Hochschulen anzustreben. Auf der einen Seite besteht ein reger Wettbewerb um Mittel, um Prestige, um das beste Lehrpersonal wie auch um die besten Studierenden, welcher zu einer vertikalen Differenzierung beiträgt. Die Differenzierung mittels fachlicher Schwerpunktsetzung auf der anderen Seite wird horizontale Differenzierung genannt (*Teichler 1999*). Wenn der Wettbewerb um Studierende – sei es mittels horizontaler oder vertikaler Differenzierung – von Erfolg gekrönt sein soll, stellt sich die Frage, nach welchen Kriterien Studienanfänger ihre Studienorte und Studiengänge tatsächlich auswählen. Wären die wichtigsten Gründe für eine Studienortwahl die Wohnortnähe und der Freizeitwert, bliebe der Wettbewerb um die Studierenden nur bedingt Erfolg versprechend. Antworten auf diese Frage geben entsprechende Untersuchungen. Die Ergebnisse von *Tutt (1997)* und *Lischka (1999)* deuten auf gute Wettbewerbsvoraussetzungen hin, da im Ergebnis die Bedeutung der Heimatnähe hinter die anderer Kriterien, wie fachlicher Neigungen, zurücktritt. In der weit größeren und nicht regional begrenzten Untersuchung von *Lewin/Heublein/Schreiber/Spangenberg/Sommer (2001)* wiederum wird die Heimatnähe am häufigsten als Grund für die Studienortwahl angegeben.

1.2 Zur Debatte in den USA

Während sich in Deutschland die thematischen Schwerpunkte auf die Studierfähigkeit, die Chancengleichheit für alle sozialen Schichten und die Integration der Zulassungsmechanismen in ein wettbewerbleres Hochschulsystem konzentrieren, sind in den USA andere Schwerpunkte erkennbar. Dabei ist aus deutscher Sicht insbesondere die Debatte interessant, welche sich auf die Studierfähigkeitstests bezieht sowie auf die Maßnahmen zur besseren Einbeziehung Studierender aus unterrepräsentierten Minderheiten.

Dass in den USA bereits flächendeckend Studierfähigkeitstest eingesetzt werden, ist sowohl historisch zu begründen als auch durch die Heterogenität der Schulen. Darüber hinaus findet die Studierendenauswahl größtenteils dezentral, d. h. hochschulspezifisch, statt. Dadurch bietet die Diskussion in den USA einen Einblick in die mit der Einführung von Tests und hochschuleigenen Auswahlpraxen verbundenen Kontroversen und – damit einhergehend – einen Forschungsüberblick dank einer Vielzahl empirischer Befunde.

Dem Studierfähigkeitstest SAT kommt bei Zulassungsentscheidungen große Bedeutung zu und er betrifft sehr brisante Bereiche. *Adelman (1999)* titelte einen Zeitungsartikel passend zur anhaltenden öffentlichen Diskussion und Kritik des Tests mit „Why can't we

stop talking about the SAT?" und bezeichnet die Verwendung des Tests als „symbolisches Monster“.² Die Tests wurden eingeführt, um den Missstand zu beseitigen, dass alle Universitäten im Land eigene Tests und Anforderungen stellten, was zu einer übertriebenen Anzahl verschiedener Verfahren führte, die ein Bewerber durchlaufen musste, wenn er sich an mehreren Hochschulen bewerben wollte.

Der Hochschulzugang im Allgemeinen und die Verwendung des SAT im Speziellen haben weit reichende Konsequenzen (*Katchadourian/Boli 1994; Bowen/Bok 1998*). Die Entwicklungen in den USA, wo Schüler und Schülerinnen teure SAT-Vorbereitungskurse besuchen, bestärken deutsche Kritiker in der Befürchtung, dass die Einführung nationaler Hochschuleingangstests zur Entstehung einer Bildungsindustrie führen könnte. Wissenschaftler sind sich darüber hinaus weitestgehend einig, dass Vorbereitungskurse für Hochschuleingangstests nicht zur Akkumulation von Humankapital führen (*Iyigun/Levin 1998*).

In den vergangenen Jahren standen darüber hinaus besonders die politischen Maßnahmen zur Herstellung der Chancengleichheit im Mittelpunkt des öffentlichen Interesses. Die Politik, die unter dem Namen „Affirmative Action“ fungierte, sah vor, einen bestimmten Anteil der Studienplätze trotz ungünstigerer Eingangsvoraussetzungen an Bewerber aus Minderheiten zu vergeben, wenn diese unterrepräsentiert waren. Weiße Studienbewerber, die sich aufgrund ihrer Ethnizität benachteiligt fühlten, erreichten schließlich, dass die Maßnahme als nicht verfassungsgemäß eingestuft wurde (*Gose/Seligo 2001*). „Affirmative Action“ wurde daraufhin in manchen Bundesstaaten durch eine andere Maßnahme ersetzt, die „x-Prozent Plan“ genannt wird. Diese Regel garantiert den x Prozent Besten eines jeden Schuljahrgangs aller Schulen einen Studienplatz im eigenen Bundesstaat (Kalifornien 4%, Texas 10%, Florida 20%). Politiker versprechen sich einen indirekten Ethnizitäts-Effekt, welcher sich aus der hohen Korrelation des sozioökonomischen Hintergrunds mit der Ethnizität ergibt. Es gibt jedoch bei beiden Maßnahmen stets einen trade-off zwischen geeigneten Mehrheitskandidaten und Minderheitskandidaten.

Die US-amerikanische Diskussion zeigt somit, dass Studierfähigkeitstests das Dilemma sozialer Ungleichheit nicht lösen und alternative Lösungen ausprobiert werden müssen.

² Der SAT und der ACT sind nicht die einzigen Tests, die flächendeckend als Eingangstests eingesetzt werden. Für weiterführende Hochschulen wird der GRE sehr häufig verlangt. Außerdem gibt es für einzelne Studiengänge wie Management, Medizin, Pharmazie und Jura unterschiedliche Instrumente (GMAT, MCAT, PCAT, LSAT) (vgl. Anhang).

2 Kriterien der Studierendenauswahl

2.1 Bewertungskriterien von Auswahlinstrumenten

Bevor im Weiteren auf empirische Ergebnisse zur Angemessenheit von Auswahlverfahren eingegangen wird, sind zur Systematisierung der Vor- und Nachteile von Auswahlinstrumenten und zu ihrer wissenschaftlichen Bewertung Bewertungskriterien zu bestimmen. So sind grundsätzlich die psychometrischen Gütekriterien der Objektivität, Reliabilität und insbesondere der Validität zu fordern.³ In der Eignungsdiagnostik – und somit auch im Hinblick auf die Diagnose von Studieneignung – spielt die prädiktive (oder auch prognostische Validität) eine wichtige Rolle.

Die Akzeptanz ist ebenfalls ein häufig erwähntes Kriterium für Zulassungsinstrumente, da sie einen Aspekt der Umsetzbarkeit eines Verfahrens darstellt. Die Akzeptanz eines Instruments unterliegt jedoch zeitlichen Variationen, da zunächst heftig umstrittene Instrumente nach ihrer Einführung und einer gewissen Bewährungszeit eine höhere Akzeptanz erfahren können.

Ökonomische Kriterien beziehen sich auf die Kosten eines Verfahrens, die bei seiner Durchführung oder Erstellung anfallen. Zu den direkten Kosten summieren sich die Opportunitätskosten für die am Auswahlverfahren Beteiligten wie Professoren, Dozenten, Vertreter des Studierendensekretariats u. a. Diese Kosten entstehen in hohem Maße bei Verfahren, die sehr zeitaufwändig sind.

Der Begriff der Fairness im Zusammenhang mit der Studierendenauswahl ist mit einer besonderen Brisanz und definitorischen Schwierigkeiten behaftet (*Browne Miller, 1996*). Der Grad der Fairness eines Auswahlinstruments ist vor allem an den Folgen der Beschränkung des Hochschulzugangs festzumachen. Eine Sichtweise auf Fairness ist, dass nur die Studierenden mit dem größten Potential im Hinblick auf den Studienerfolg ausgewählt werden. Dies würde Fairness durch prädiktive Validität bedeuten. Sozial neutral ist dies jedoch nicht, da der sozioökonomische Hintergrund einen deutlichen Einfluss auf die Studienleistung hat.

³ Objektivität bedeutet die Unabhängigkeit der Ergebnisse vom Anwender des Instruments. Reliabilität bezieht sich auf die Genauigkeit der Messungen und Validität darauf, dass ein Instrument tatsächlich das misst, was es messen soll.

Was fair ist, ist entsprechend schwierig abschließend zu beurteilen; noch schwieriger ist Fairness messbar. Als Mindestvoraussetzung von Fairness könnte gelten, dass Studierende nicht nur aufgrund eines einzigen Kriteriums ausgewählt werden, sondern zumindest einem Teil der Studienbewerber die Chance eröffnet wird, seine Qualifikationen für einen Studienplatz aufgrund eines zweiten Kriteriums unter Beweis zu stellen.

Ein nicht zu vernachlässigendes Kriterium ist die Transparenz eines Verfahrens. Die Kommunikation der Instrumente, die Offenlegung der Vorgehensweise und Ergebnisfindung im Vorfeld sowie nach der Entscheidung spielen eine wichtige Rolle. Transparenz kann die Informationskosten minimieren.⁴ Einige Instrumente sind mehr als andere sensibel für mangelnde Transparenz, da der Studienbewerber noch über z.B. Essay- oder Gesprächsinhalte und Bewertungskriterien zu informieren ist.

Berücksichtigung von Hochschulspezifika: Wenn Hochschulen ihre Studierenden selbst aussuchen sollen und sich mit einem eigenen Profil positionieren wollen, benötigen sie zur Auswahl derjenigen Studierenden, die am besten zu ihnen passen, Instrumente, die eine Berücksichtigung von Spezifika ihrer Hochschule bzw. ihres Instituts ermöglichen. Können die Hochschulen nicht auf entsprechende Instrumente zurückgreifen, kommt es zu erheblicher Redundanz in den Verfahren, d.h. an vielen Hochschulen werden die gleichen Personen angenommen und es ergeben sich wiederholt umfangreiche Nachrückverfahren.

Zusammenfassend ist festzustellen, dass Auswahlverfahren effizient zu gestalten sind, sowohl für die Hochschulen als auch für die Studienbewerber, unter Berücksichtigung der Güte der Instrumente sowie der Fairness des Verfahrens. Durch die Wahl der Auswahlinstrumente ist eine Maximierung der genannten Kriterien (Ökonomie, Reliabilität, Validität, Transparenz und Fairness) anzustreben, in der die prädiktive Validität einen besonderen Stellenwert einnimmt. In einem System hochschuleigener Auswahl muss ausreichend Raum für die Ausarbeitung von Verfahren geboten werden, die sich grundsätzlich unterscheiden, um die erwähnte Redundanz zu vermeiden.

⁴ Zudem reduziert Transparenz weitere Kosten wie die direkten Kosten der Bewerbung an Hochschulen, bei denen man sich bei genauerer Kenntnis des Auswahlverfahrens nicht bewerben würde. Für andere mögen die Opportunitätskosten der Informationsbeschaffung so hoch sein, dass sie sich gegen ein (bestimmtes) Studium entscheiden.

2.2 Kriterien des Studienerfolgs

Der Studienerfolg ist das Kriterium, an dem sich Auswahlverfahren in der Hochschulzulassung zu messen haben, weshalb es einer Determinierung des Konstrukts Studien(miss)erfolg bedarf. Es wird in empirischen Studien eine Reihe verschiedener Maße verwendet, die Studienerfolg operationalisieren. Das Konstrukt des Studien(miss)erfolgs ist nicht als Kontinuum zu verstehen, sondern als zwei verschiedene Ausgänge des Studienverlaufs.

Unter Studienerfolg werden in der deutschen empirischen Literatur vor allem das Erreichen eines Studienabschlusses oder die Examensnote verstanden. Hinzu kommen weitere, seltener evaluierte Kriterien wie die Studiendauer, die Zufriedenheit und der Berufserfolg. In der Befragung von *Konegen-Grenier (2002)* sollten Hochschullehrer mögliche Kriterien des Studienerfolgs bewerten. Die Ergebnisse zeichnen ein humanistisch geprägtes Bild, das an deutsche Bildungstraditionen anknüpft: Die Fähigkeit zum wissenschaftlichen Arbeiten wurde als am bedeutendsten eingestuft, wichtig war auch die persönliche Weiterentwicklung. Nichtsdestotrotz ist die Fähigkeit zum wissenschaftlichen Arbeiten nicht explizit in einer empirischen Studie als Erfolgskriterium verwendet worden (jedenfalls sind der Autorin solche Studien nicht bekannt), vermutlich aufgrund von Schwierigkeiten in der Operationalisierung.

3 Empirische Evidenz zur Angemessenheit von Auswahlverfahren

Die für deutsche hochschuleigene Auswahlverfahren zur Verfügung stehenden Instrumente ergeben sich aus dem Hochschulrahmengesetz und dem Staatsvertrag vom 26.04.1999. Die wichtigsten und am häufigsten verwendeten Instrumente sind Schulnoten, Tests, Auswahlgespräche und Essays. Als im Angelsächsischen häufig verwendetes Instrument sind Empfehlungsschreiben und Gutachten von Schullehrern zu nennen. Fremdsprachentests seien hier als eine spezielle Form von Tests, Essays oder Auswahlgesprächen betrachtet.

Im Weiteren werden empirische Antworten auf die Frage nach der Aussagekraft dieser Auswahlinstrumente gesucht. Aufgrund der Dominanz der Untersuchungen zur prädiktiven Validität und ihrer relativen Bedeutung unter den Kriterien für Auswahlinstrumente wird zunächst nur auf Untersuchungen zur prädiktiven Validität eingegangen. Solche Untersuchungen werden im Bereich der Hochschulzulassung allerdings kaum mit theoretischen Überlegungen gestützt, und eine zufrieden stellende Validität als ausreichend

hingegen. Die dargestellten Forschungsergebnisse beziehen sich auf die am häufigsten eingesetzten Instrumente Tests, Schulnoten, Auswahlgespräche und Essays. Viele andere, oftmals in der US-amerikanischen Praxis angewandte, Instrumente⁵ sind nur ungenügend evaluiert (*Schwirten, 2000*).

Die betrachteten Studien zu den verschiedenen Auswahlinstrumenten zeigen die neueren Entwicklungen in der Forschung zur Hochschulzulassung. Der Fokus liegt daher auf Studien seit dem Jahr 1990.⁶

3.1 Prädiktive Validität von Auswahlinstrumenten und Einflussvariablen

3.1.1 Schulnoten

Baron-Boldt (1989) hat die in Deutschland bisher einzige Metaanalyse dieser Größe mit einer Zahl von 26.867 untersuchten Personen zur prädiktiven Validität von Schulnoten im Bezug auf die Examensnoten vorgenommen. Der Validitätskoeffizient betrug $r = .456$. Ein weiteres bedeutendes Ergebnis war, dass Einzelnoten den Examenserfolg keinesfalls besser vorhersagen als die Abiturnote. *Höppel/Moser (1993)*, die mit der gleichen Korrekturmethode die gefundene Stärke des Zusammenhangs bestätigen, finden gar heraus, dass die Abiturnote selbst nach Einbezug der Vordiplomnote noch einen unabhängigen Erklärungsanteil an der Hauptexamensnote hatte. In neueren Untersuchungen ist die Abiturnote stets ein signifikanter Prädiktor für die Examensnote; dabei wurde mit vielen unterschiedlichen statistischen Methoden gearbeitet (*Giesen/Gold/Hummer/Jansen 1986; Daniel 1996; Friese/Cierpka 1996; Stephan 2002; Streng 1996; Sieglein 1998; Brandstätter/Grillich/Farthofer 2002; Fries 2002; Hinneberg 2004; Meier 2003*). Die Abiturnote ist darüber hinaus in Untersuchungen mit verschiedenen Prädiktoren in der Regel der beste Prädiktor, so z. B. in *Kurz/Fischer/Wagner (1995), Trost/Blum/Fay/Klieme/Maichle/Meyer/Nauels (1998) oder Brandstätter/Grillich/Farthofer (2002)*. Nur für die Ingenieurwissenschaften werden vergleichsweise aussagekräftige Prognosen durch Einzelnoten, nämlich die Mathematik- und Physiknoten, gefunden (*Ulbricht 1986; Kurz/Fischer/Wagner 1995*).

⁵ Zu diesen Instrumenten zählen unter anderem Schulgutachten und extrakurriculare Aktivitäten wie Sport oder Ehrenämter.

⁶ Als Ausnahmen werden auch ältere Studien berücksichtigt. Die Ausnahmen beziehen sich auf Metaanalysen und besonders häufig zitierte Untersuchungen aus den 80er Jahren oder solche Untersuchungen, die durch eine spezielle methodische Herangehensweise charakterisiert sind.

In den USA werden die Validitätswerte der Schulnoten meist in Studien mit der Validität von Studierfähigkeitstests (in der Regel SAT) verglichen. Die US-amerikanischen Studien bestärken die deutschen Ergebnisse und weisen zudem ähnliche Korrelationskoeffizienten aus (z.B. *Elliot/Strenta 1988; Bowen/Bok 1998; Geiser/Studley 2001; Kobrin/Camara/Milewski 2002; Harackiewicz/Barron/Tauer/Elliot 2002*). Die Stärke des Zusammenhangs von Schulabschlussnoten (in den USA: high school grade point average) und Abschlussnoten im Studium wird in der Literatur damit begründet, dass beide ein Aggregat von Noten umfassen, in denen abweichende Einzelnoten ausgeglichen werden. Dies gilt für das deutsche und das US-amerikanische System gleichermaßen.

Die Stärke des Abiturs liegt vermutlich darin begründet, dass die Notengebung von Schullehrern ähnlich wie die von Hochschullehrern ein Aggregat von Kriterien umfasst. Die vergleichsweise hohe prognostische Validität besteht trotz einer mangelhaften Reliabilität und Objektivität und obwohl in Deutschland bundeslandspezifische Unterschiede in der Vergabe von Abiturnoten bestehen (*Baumert/Artelt/Klieme/Neubrand/Prenzel/Schiefele/Schneider/Tillmann/Weiß 2003*). Selbst einige Kriterien des Berufserfolgs weisen noch einen Zusammenhang zum Abitur auf. Andere Kriterien wie Studiendauer und Studienabbruch können weniger gut prognostiziert werden, es lassen sich jedoch auch hier zum Teil signifikante Korrelationen finden (z.B. *Höppel/Moser 1993; Daniel 1996; Streng 1996*).

3.1.2 Tests

Als Hochschuleingangstests sind Studierfähigkeitstests von Kenntnistests zu unterscheiden. Kenntnistests überprüfen schulfachbezogenes oder studienfachbezogenes Wissen, während Studierfähigkeitstests kognitive Fähigkeiten wie Gedächtnisstärke oder mathematisch-technische Problemlösekompetenz erheben (*Deidesheimer Kreis 1997*).

Da in Deutschland nur der Studierfähigkeitstest TMS weitflächig zur Auswahl von Studierenden eingesetzt wurde, konzentrieren sich die Überprüfungen von Tests in Deutschland auf den Mediziner-test. Die Ergebnisse jahrelanger Evaluation wurden in *Trost/Blum/Fay/Klieme/Maichle/Meyer/Nauels (1998)*⁷ veröffentlicht. Der Korrelationskoeffizient der Gesamtnote aus der Ärztlichen Vorprüfung mit dem TMS war beinahe identisch mit der Validität der Abiturnote (.45 zu .47). Durch die Kombination der beiden Kriterien wird

⁷ Die Prognosefähigkeit des Tests wurde mit einer Stichprobe von 34.301 Studierenden für den Studiengang Medizin an drei Testzeitpunkten überprüft.

eine inkrementelle Verbesserung durch den TMS von ca. 7% erreicht. Vergleichbare Ergebnisse zeigt eine Reihe von Untersuchungen aus den USA zum MCAT und PCAT für Medizin und Pharmazie, so z.B. die Metaanalyse von *Kuncel/Hezlett/Ones (2001)*.

Ergebnisse aus Kenntnistests sind in Untersuchungen, die zum Teil mit sehr großen Studierendenzahlen durchgeführt wurden, prognostisch valider als Studierfähigkeitstests (z.B. *Elliott/Strenta 1988; Geiser/Studley 2001; Bridgeman/Burton/Cline 2003; Kobrin/Camara/Milewski 2002*) und darüber hinaus weniger durch den sozioökonomischen Hintergrund beeinträchtigt (*Bridgeman/Burton/Cline 2003*). Daran schließt sich die Vermutung an, dass die intellektuellen Faktoren stärker durch den sozioökonomischen Hintergrund geprägt sind als das Wissen, welches ein Studienbewerber bis zur Bewerbung akkumuliert.

Beratungseffekte mit dem Einsatz studienfeldbezogener Tests sind nicht konsistent positiv: Während *Brandstätter/Grillich/Farthofer (2002)* von weniger Studienabbrüchen berichten konnten, fand *Ulbricht (1986)* keine statistisch signifikante Korrelationen zwischen Beratung und Diplomnote für die Ingenieurwissenschaften.

Die prognostische Validität für Tests, die nicht speziell für die Auswahl bzw. Identifikation von potentiell guten Studierenden entwickelt wurden, ist enttäuschend oder allenfalls inkonsistent; dies gilt auch für Intelligenztests (z.B. *Giesen/Gold/Hummer/Jansen 1986; Meulemann 1995; Brandstätter/Grillich/Farthofer 2002*).

Die Trainierbarkeit des TMS stellte sich als vernachlässigbar heraus (*Trost/Blum/Fay/Klieme/Maichle/Meyer/Nauels 1998*). Für den SAT gibt es wiederum keine eindeutigen Ergebnisse, da die Trainingseffekte je nach Studie sehr unterschiedlich groß ausfallen (*Powers 1993*). *Allalouf/Shakhar (1998)* zeigen allerdings, dass Coaching-Effekte die prädiktive Validität von Tests nicht mindern. Dieses ist aus eignungsdiagnostischer Sicht ein beruhigendes Resultat, entkräftet aber nicht das soziale Argument der Entstehung einer Bildungsindustrie.

3.1.3 Auswahlgespräche

Auswahlgespräche in der Hochschulzulassung wurden bisher deutlich seltener evaluiert als Tests oder Schulnoten. Einige Studien zu Interviews liegen dennoch vor, in Deutschland größtenteils für die über eine Gesprächsquote vergebenen Studienplätze der Medizin. *Basler/Bolm/Bärwinkel (1992)* zeigen mittels Varianzanalyse, dass die über die Gesprächs-

quote zugelassenen Studierenden eine signifikant höhere Motivation und Identifikation mit der Universität aufweisen als jene aus der Test-Abiturnoten-Quote. Im Hinblick auf die Noten ist letztere Gruppe signifikant besser; keine Unterschiede gibt es bei den Bestehensquoten. *Hinneberg (2004)* findet heraus, dass die über Auswahlgespräche zugelassenen Studierenden keine besseren Examensleistungen erbringen als nachgerückte Studierende mit der gleichen Abiturnote. In vielen Untersuchungen sind die Korrelationen zumindest statistisch signifikant, wobei die Stärke der Korrelationen variiert.

Ittner (1993) stellt fest, dass der Korrelationskoeffizient zwischen Examensnoten und Test-Abiturnoten-Wert bei über die Gesprächsquote Zugelassenen erheblich geringer ist als bei der Test-Abiturnoten-Quote. Es scheint im Auswahlgespräch gelungen zu sein, Bewerber zu identifizieren, deren Leistungen nicht gut mit den in Tests überprüften Konstrukten erklärbar sind. *Lewin (2004)* zeigt darüber hinaus, dass Auswahlgespräche zu einer größeren Informiertheit der Studienanfänger führen.

Am Einsatz von Auswahlgesprächen wird kritisiert, dass diese hinsichtlich der zu beurteilenden Kriterien überschätzt werden. Insbesondere von der Kommunikationsfähigkeit wird im Allgemeinen aber angenommen, dass sie in einer Interviewsituation beurteilt werden kann. Eine Untersuchung von *Kelman/Canger (1994)* zeigt jedoch, wie schwierig auch eine solche Beurteilung ist. Sie setzten Skalen ein, mit deren Hilfe Kommunikationsfähigkeit, Problemlösungsfähigkeit und soziale Verantwortlichkeit erfasst werden sollten. Die Korrelation zwischen den Ratings der Interviewer und den Ratings vorgesetzter Ärzte ein paar Jahre später anhand derselben Skalen war nicht nur insignifikant, sondern praktisch Null (-.068).

Es liegen zahlreiche Untersuchungen zu Auswahlgesprächen in der Personalauswahl vor, welche zu ähnlichen Ergebnissen führen wie Untersuchungen im Zusammenhang mit der Hochschulzulassung. In ihrer Metaanalyse ermittelten *Hunter/Hunter (1984)* Validitätskoeffizienten von .03 bis .14, abhängig vom Kriterium des Berufserfolgs. Verbesserungen sind durch Schulung der Interviewer, Strukturierung und Anforderungsbezug zu erreichen (*Wiesner/Cronshaw 1988*).

3.1.4 Essays

In Deutschland gibt es kaum Erfahrungen zur Validität von Essays in der Hochschulzulassung. Im Gegensatz dazu, widmen sich US-amerikanische Wissenschaftler diesem Thema in einer Reihe von Studien, insbesondere den Revisionstests im GRE oder SAT II.

Breland (1999) fasst zusammen, dass die Reliabilität von Essays bei Ratings durch erfahrene Beurteiler bis zu .90 betragen kann. Auch die Aufsatzform hat hierauf Einfluss: Während die Reliabilität bei Aufsätzen im engeren Sinne eher niedrig ist, wird bei Aufsatz-Revision oder Satz-Revision⁸ eine wesentlich höhere Reliabilität erreicht. Dies trifft auch auf die prädiktive Validität zu. Die prädiktive Validität ist in der Regel nicht sehr hoch. Bei *Kirchner/Holm (1997)* und *Gutierrez/Cuervo/Rosselli (1999)* liegt sie bei um die $r=.20$ für die Vorhersage von Noten; *Sadler (2003)* berichtet einen insignifikanten Zusammenhang. Eine etwas stärkere Korrelation können *Kuncel/Hezlett/Ones (2001)* für den Revisionstest des GRE in ihrer Metaanalyse bestätigen.

3.1.5 Sozioökonomischer Hintergrund

Der sozioökonomische Hintergrund ist zwar kein Kriterium bei der Auswahl von Studierenden, aber eine Einflussvariable, der hinsichtlich der öffentlichen Diskussion um soziale Fairness Aufmerksamkeit zu widmen ist. Eine Studie aus dem EU Labour Force Survey 2000 (*Ianelli 2002*) findet sehr unterschiedliche, länderspezifische Ergebnisse zur sozialen Mobilität. Eine Regressionsanalyse zeigt allerdings, dass der Effekt der sozialen Herkunft über alle Länder statistisch hoch signifikant ist. In den USA hat der sozioökonomische Hintergrund einen sehr starken Einfluss auf den Studienabbruch (*Bowen/Bok 1998*). Selbst innerhalb der Gruppe der Hochbegabten spielt der familiäre Hintergrund noch eine große Rolle (*Benbow/Arjmand 1990*). *Studley (2003)* korrigiert statistisch die Prädiktoren für das Haushaltseinkommen, die Ausbildung beider Eltern, die Nachbarschaft und die besuchte Schule. Allein diese Faktoren erklären 32% der SAT-Werte in der Stichprobe.

3.1.6 Zusammenfassung der empirischen Ergebnisse

Die empirischen Ergebnisse zeigen, dass die Abiturnote am ehesten geeignet ist zur Vorhersage von Examensleistungen. Eine ähnliche Validität erreicht lediglich ein Studierfähigkeits- oder Kenntnistest. Auswahlgespräche haben zwar positive Externalitäten wie eine bessere Informiertheit von Studienbewerbern, sind aber, ähnlich wie Essays, wenig valide im Hinblick auf Examensleistungen. Der sozioökonomische Hintergrund hat wiederum einen deutlichen Effekt auf die mithilfe der Auswahlinstrumente ermittelten Ergebnisse.

⁸ Es handelt sich hierbei um Korrektur und Verbesserung bereits verfasster Aufsätze (Aufsatz-Revision) und Sätze (Satz-Revision).

3.2 Beurteilung der Instrumente anhand weiterer Bewertungskriterien

Die Akzeptanz verschiedener Hochschulzulassungskriterien wurde in einer aktuellen Studie überprüft (*TNS Emnid 2004*). Grundsätzlich scheinen Studierende eine Reihe von Kriterien als sinnvoll zu erachten – am ehesten Auswahlgespräche und studienfachspezifische Wissensfragen, am wenigsten Essays. Abiturnoten befinden sich im mittleren Akzeptanzbereich. Die gute Akzeptanz der Auswahlgespräche wird von Interview-Teilnehmern bzw. Studierenden aus Studiengängen mit Zulassung über Gesprächsquoten bekräftigt, die sich im Allgemeinen positiv zu Interviews äußern (*Basler/Bolm/Bärwinkel 1992; Ittner 1993; Lewin 2004*).

Die Kriterien Ökonomie und Berücksichtigung von Hochschulspezifika sind empirisch nur schwer bewertbar. Die ökonomische Bewertung der Instrumente ist zum Teil abhängig von der konkreten Ausgestaltung der Verfahren. Auswahlgespräche sind in jedem Fall aufwändig, ebenso wie die Bewertung von Schulgutachten und Essays (mit Ausnahme von Revisionstests). Die Kosten von Tests sind vor allem davon abhängig, ob sie maschinell ausgewertet sowie von zentralen Instituten ausgearbeitet und evaluiert werden. Hochschulspezifika bzw. Spezifika eines Studiengangs an einer bestimmten Institution können nur mit solchen Instrumenten überprüft werden, die mindestens fachspezifische Unterschiede erkennen lassen sowie mit solchen, die erlauben, Einstellungen, Motivation, oder Interessen im Hinblick auf das hochschulspezifische Anforderungsprofil abzuwägen. Hierzu zählen beispielsweise die (wenig validen) Auswahlgespräche und Einzelnoten. Die Handhabung der Instrumente in den hochschuleigenen Verfahren ist schließlich für die Transparenz entscheidend.

4 Hochschulautonome Studierendenauswahl in Baden-Württemberg

Wie die Auswahl der Instrumente an deutschen Hochschulen gehandhabt wird und ob die Auswahlverfahren den dargestellten Bewertungskriterien gerecht werden, soll im Folgenden aufgezeigt werden. Dabei wird besonders Baden-Württemberg genauer betrachtet, weil dort die Hochschulen zur Durchführung einer hochschuleigenen Auswahl bei örtlichen Zulassungsbeschränkungen verpflichtet sind (90% der zu vergebenden Studienplätze). § 10 Hochschulvergabeverordnung sieht vor, dass die in der Oberstufe belegten Kernfächer Deutsch, eine fortgeführte Fremdsprache und Mathematik berücksichtigt werden müssen. Schulnoten sind mit einem anderen Kriterium zu kombinieren, können aber nicht allein miteinander kombiniert werden.

4.1 Praxis und Erfahrungen in der hochschulautonomen Studierendenauswahl

Eine Untersuchung zur Auswahlpraxis von Universitäten in Baden-Württemberg und Bayern haben *Lewin/Lischka (2004)* für das Wintersemester 2003/2004 unternommen; für Baden-Württemberg wurde von *Caspar/Kimmler-Schad (2004)* eine detaillierte Analyse durchgeführt. Eigene Analysen wurden zu den Auswahlverfahren der Universitäten in Baden-Württemberg realisiert, indem die Auswahlsetzungen der zulassungsbeschränkten Studiengänge gesammelt wurden.⁹ Für die drei Untersuchungen wurde aufgrund unterschiedlicher Einordnung eine differierende Anzahl an Studiengängen ausgewertet.¹⁰

Caspar/Kimmler-Schad (2004) nennen eine bestimmte Kriterien-Kombination in Baden-Württemberg aufgrund der Häufigkeit ihrer Anwendung das „Standard-Auswahlmodell“. Dies sieht die Verwendung der Kriterien (a) Note(n) der Hochschulzugangsberechtigung und (b) Berufsausbildung, praktische Tätigkeit und/oder außerschulische Leistungen vor. In 89% der Fälle wird dieses Modell angewandt. Bei einem Großteil dieser Auswahlverfahren handelt es sich wiederum um ein Standard-Auswahlmodell mit erweiterten Anforderungen an die Hochschulzugangsberechtigung, da neben den gesetzlich geforderten Einzelnoten noch weitere Einzelnoten einbezogen werden (67,5%). Tests wurden in 12,4% der Fälle vorgesehen, Auswahlgespräche in 6,7%. Eine Berufsausbildung oder praktische Tätigkeiten werden in 86,6% honoriert.

⁹ Zuerst wurden die Hochschulen um die Auswahlsetzungen derjenigen Studiengänge gebeten, die im hochschuleigenen Verfahren vergeben werden. Dazu wurden die Studiensekretariate aller Hochschulen zweimal angeschrieben. Fehlende Satzungen wurden schließlich aus dem Internet gewonnen, und einzelne Universitäten ein weiteres Mal angeschrieben.

¹⁰ Aufgrund ihrer Sonderstellung werden Sport- Kunst- oder Musikstudiengänge in der eigenen Untersuchung nicht betrachtet. Zwischen den Untersuchungen durch *Lewin/Lischka (2004)* und *Caspar/Kimmler-Schad (2004)* liegen deutliche Unterschiede: Erstere zählten 61 Studiengänge an Universitäten, letztere 206. Dies mag verschiedene Ursachen haben, denn *Lewin/Lischka* schlossen Lehramtsstudiengänge aus und bezogen nur diejenigen ein, an denen tatsächlich ein Verfahren durchgeführt wurde. *Caspar/Kimmler-Schad* analysierten alle grundständigen Studiengänge mit Auswahlsetzungen. Auch zur eigenen Untersuchung gibt es eine Diskrepanz, da hier 165 (von 173 identifizierten) Studiengänge einfließen. Diese Diskrepanz ergibt sich vermutlich größtenteils aus der Tatsache, dass in der eigenen Untersuchung Auswahlsetzungen mit dem gleichen Verfahren für gleiche Studiengänge mit unterschiedlichen Abschlussarten als ein Studiengang gewertet wurden sowie Musik-, Kunst- und Sportwissenschaftliche Studiengänge bei *Caspar/Kimmler-Schad* einbezogen wurden.

Die eigenen Analysen der Auswahlsetzungen ermöglichen einen noch differenzierteren Blick auf einige Details der Verfahren. Die Analysen replizieren zunächst eine Reihe bereits beschriebener Auswahlpraxen: Die Anwendungshäufigkeit von Tests und Interviews unter den Verfahren ist sehr dicht an den Ergebnissen von *Caspar/Kimmler-Schad (2004)* und deutet daher auch darauf hin, dass die Ergebnisse der eigenen Analyse trotz einiger fehlender Auswahlsetzungen repräsentativ für die Baden-Württembergischen Universitäten sind; dies gilt auch für die Honorierung von Berufstätigkeit. In den eigenen Analysen wurden die Kategorien „nur Berufstätigkeit“ und „sonstige Leistungen“ (welche auch Berufstätigkeit enthalten kann, wenn sie mit Anderem kombiniert wird) unterschieden. Gemeinsam finden die beiden Kategorien in ca. 80% der Fälle Anwendung, in 30% wird die Berufstätigkeit als einzige zusätzliche Leistung berücksichtigt. In den verbleibenden 50% findet eine Kombination aus Berufstätigkeit und außerschulischen Leistungen oder Auslandsaufenthalten Anwendung. Der Prozentsatz an sich sagt nur wenig über den Stellenwert der Berufstätigkeit und sonstiger Tätigkeiten aus, da dessen Gewichtung stark variiert. Dadurch entsteht der Eindruck, dass manche Hochschulen viel Wert auf vorherige Berufserfahrungen legen in der Hoffnung auf motivierte und engagierte Studierende, während andere die Tätigkeiten nur aus dem Grund in die Auswahlentscheidung aufnehmen, dass sie gesetzlich zur Kombination von Noten mit einem weiteren Kriterium verpflichtet sind.

Das oben nach *Caspar/Kimmler-Schad (2004)* „Standard-Modell“ genannte Auswahlmodell kann bei näherer Betrachtung entsprechend der Gewichtung der Kriterien Noten und sonstige Leistungen bzw. Berufstätigkeit noch differenziert werden. Das am häufigsten verwendete Modell ist demnach die Drei-Drittel-Lösung: Ein Drittel Abiturnote, ein Drittel Einzelnoten und ein Drittel sonstige Leistungen bzw. Berufstätigkeit. In fast einem Fünftel aller Studiengänge findet dieses Modell Anwendung.

Insgesamt kann festgehalten werden, dass der Kreativität bei der Gewichtung von Kriterien keine Grenzen gesetzt sind. Eine Profilierung in der Auswahl von Studierenden wird dadurch zumindest zum Teil erreicht. Ein positiver Effekt ist, dass die Studierenden dazu gezwungen sind, sich näher mit den Auswahlkriterien und somit auch mit dem Studiengang zu beschäftigen. Es ist jedoch kritisch zu sehen, dass in vielen Studiengängen die Abiturnote trotz ihrer vergleichsweise hohen prognostischen Validität für die Examensnote nicht als Auswahlkriterium herangezogen wird. Einzelnoten werden bei einer Kombination mit der Abiturnote sehr oft mindestens genau so hoch gewichtet wie die Abiturnote.

4.2 Erfahrungen mit der Studierendenauswahl in Baden-Württemberg

Die Hochschulen wurden neben ihren Satzungen auch um Erfahrungsberichte gebeten. Die meisten Hochschulen haben jedoch aufgrund der kurzen Zeit, in der die Auswahlverfahren durchgeführt werden, noch keine Erfahrungsberichte zusammengestellt (eine Ausnahme ist z. B. die Universität Mannheim).

Viele Verantwortliche sehen das Auswahlgespräch offensichtlich als besonders geeignet zur Auswahl von Studierenden an, was den dargestellten empirischen Erkenntnissen nicht entspricht. Die Beobachtungen verschiedener Verantwortlicher, die in die Auswahlverfahren einbezogen wurden, sind größtenteils pessimistischer als schriftliche Erfahrungsberichte, die Durchführbarkeit und erste Ergebnisse als positiv beschreiben. So bescheiden sie den Verfahren mangelnde Transparenz und Übersichtlichkeit und die Entstehung von Verunsicherung bei den Studienbewerbern sowie von umfangreichen Nachrückverfahren. Letztere kommen dadurch zustande, dass die Hochschule trotz der Möglichkeit unterschiedlicher Gewichtungen nur eine begrenzte Zahl an Instrumenten zur Verfügung hat, die sie einbeziehen kann, und ausgewählte Studierende an mehreren Hochschulen angenommen werden. Die Restriktion auf wenige Auswahlinstrumente ist daher zwar einerseits wünschenswert, weil viele Instrumente wenig valide und kaum evaluiert sind, jedoch kann dadurch eine effiziente hochschulspezifische Auswahl nicht stattfinden.

5 Schlussfolgerungen

Welches sind nun die Schlussfolgerungen aus der vorangegangenen Analyse? Sind die Hoffnungen begründet, dass durch hochschuleigene Auswahlverfahren eine angemessenere Studierendenauswahl erfolgen kann als bislang? Dies ist meines Erachtens zu verneinen. Zur Begründung sei mit der Validität der Instrumente begonnen.

Die beste Vorhersagbarkeit von Examensnoten bietet die Abiturnote; sie ist außerdem als Instrument am ökonomischsten. Daneben sind Studierfähigkeits- und Kenntnistests prognostisch relativ valide, jedoch weniger ökonomisch. Es gibt aber auch andere Instrumente, mit denen teilweise Studierende identifiziert werden können, die erfolgreich und engagiert studieren werden; die Validitätswerte reichen jedoch nicht an die der Abiturnote heran und Validitätsverbesserungen sind eher gering. Was *Markl (1997)* für die Forschungsförderung bereits formuliert hat, gilt auch für die Studierendenauswahl: Letztlich weiß man immer erst im Nachhinein, wer die besten waren. Es gibt zu wenige Instrumente, die wissenschaftlichen Kriterien gerecht werden, als dass durch deren

Kombination eine hohe Redundanz in den Verfahren und darauf folgend umfangreiche Nachrückverfahren vermieden werden könnten.

Daher sollte zumindest die Hälfte der Studienplätze in Studiengängen mit Bewerberüberhang an die Schüler mit den besten Noten vergeben werden. Die restlichen Plätze könnten nach nicht-eignungsdiagnostischen Gesichtspunkten vergeben werden. Hierzu sollte die Mission einer Hochschule oder eines Fachbereichs – sofern vorhanden – Berücksichtigung finden, z. B. Interdisziplinarität, Internationalität oder verschiedene Sozialkriterien. Dabei ist zu bedenken, dass ein anregendes, (leistungs)motiviertes Studiumfeld schlechtere Eingangsvoraussetzungen einer Studierendengruppe zu einem großen Teil ausgleichen kann.

Abschließend sei festgehalten, dass mit der Hochschulquote durchaus sehr unterschiedliche, nicht-eignungsdiagnostische Modelle ausprobiert und wissenschaftlich angemessen begleitet werden sollten. Nur auf diese Weise wird ein besserer Erkenntnisstand über eine Vielzahl möglicher Effekte verschiedener Auswahlstrategien erreicht werden.

Literatur

Adelman, C. (1999): Why can't we stop talking about the SAT? In: The chronicle of higher education, B4

Allalouf, A.; Shakhar, G. B. (1998): The effect of coaching on the predictive validity of scholastic aptitude tests. In: Journal of educational management 35 (1998), S. 31–47

Baron-Boldt, J. (1989): Die Validität von Schulabschlussnoten für die Prognose von Ausbildungs- und Studienerfolg. Eine Metaanalyse nach dem Prinzip der Validitätsgeneralisierung. Frankfurt

Basler, H.-D.; Bolm, G.; Bärwinkel, P. (1992): Auswahlgespräche für den humanmedizinischen Studiengang an der Philipps-Universität Marburg: Beurteilung durch die Teilnehmer und prognostische Qualität des Verfahrens. Frankfurt a. M.

Baumert, J.; Artelt, C.; Klieme, E.; Neubrand, M.; Prenzel, M.; Schiefele, U.; Schneider, W.; Tillmann, K.-J.; Weiß, M. (Hrsg.). (2003): Pisa 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland. Opladen

Benbow, C. P.; Arjmand, O. (1990): Predictors of high academic achievement in mathematics and science by mathematically talented students: a longitudinal study. In: Journal of educational psychology 82 (1990), S. 430–441

Bowen, W. G.; Bok, D. (1998): The shape of the river: long-term consequences of considering race in college and university admissions. Princeton

Brandstätter, H.; Grillich, L.; Farthofer, A. (2002): Studienverlauf nach Studienberatung. In: Zeitschrift für Pädagogische Psychologie 16 (2002), S. 15–28

Breland, H. M. (1999): From 2 to 3 Rs: The expanding use of writing in admissions. In: Messick, S. J. (Hrsg.): Assessment in higher education. Issues of quality, student development, and public policy. Mahwah, S. 91–111

Bridgeman, B.; Burton, N.; Cline, F. (2003): Substituting SAT II: Subject Tests for SAT I: Reasoning Tests: Impact on admitted class composition and quality. In: Research in higher education 44 (2003), S. 83–98

Browne Miller, A. (1996): Shameful admissions. San Francisco

Caspar, R.; Kimmler-Schad, S. (2004): Die Neuregelung des Hochschulzugangs in Baden-Württemberg ab Wintersemester 2003/04. Stuttgart

Daniel, H.-D. (1996): Korrelate der Fachstudierendauer von Betriebswirten. Ergebnisse einer Absolventenbefragung an der Universität Mannheim. In: Ergänzungsheft der Zeitschrift für Betriebswirtschaft: Betriebswirtschaftslehre und der Standort Deutschland 65 (1996), S. 95–115

Deidesheimer Kreis (1997): Hochschulzulassung und Studieneignungstests. Studienfeld-bezogene Verfahren zur Feststellung der Eignung für Numerus-clausus- und andere Studiengänge. Göttingen

Elliot, R.; Strenta, A. C. (1988): Effects of improving the reliability of the GPA on prediction generally and on comparative predictions for gender and race particularly. In: Journal of educational measurement 25 (1988), S. 333–347

Fries, M. (2002): Abitur und Studienerfolg. In: Beiträge zur Hochschulforschung 24 (2002), 1, S. 30–51

Friese, M.; Cierpka, R. (1996): Erfolg in Studium und Beruf. Befunde einer empirischen Untersuchung und deren Konsequenzen für die Universitätsausbildung. In: Marketing als Schnittstellenwissenschaft und Transfertechnologie (S. 353–373)

Geiser, S.; Studley, R. (2001): UC and the SAT: Predictive validity and differential impact of the SAT I and SAT II at the University of California www.ucop.edu/sas/research/researchandplanning/pdf/sat_study.pdf, 01–2004

Giesen, H.; Gold, A.; Hummer, A.; Jansen, R. (1986): Prognose des Studienerfolgs. Ergebnisse aus Längsschnittuntersuchungen (Schlußbericht). Frankfurt

- Gose, B.; Selingo, J. (2001):* The SAT's greatest test: Social, legal and demographic forces threatening to dethrone the most widely used college entrance exam. In: *The chronicle of higher education* A 10–15
- Gutierrez, J. M.; Cuervo, C.; Rosselli, D. (1999):* Selecting medical students in Colombia. In: *Medical education* 33 (1999), S. 863–864
- Harackiewicz, J. M.; Barron, K. E.; Tauer, J. M.; Elliot, A. J. (2002):* Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. In: *Journal of educational psychology* 94 (2002), S. 562–575
- Hinneberg, H. (2004):* Zulassung zum Studium – bessere Studenten durch Auswahlgespräche? In: *Das Hochschulwesen* 52 (2004), 1, S. 12–14
- Höppel, D.; Moser, K. (1993):* Die Prognostizierbarkeit von Studiennoten und Studierendauer durch Schulabschlußnoten. In: *Zeitschrift für Pädagogische Psychologie* 7 (1993), S. 25–32
- Hunter, J. E.; Hunter, R. F. (1984):* Validity and utility of alternative predictors of job performance. In: *Psychological Bulletin* 96 (1984), S. 72–98
- Ianelli, C. (2002):* Parental income and young people's market outcomes: A comparison across Europe. In: *Arbeitspapiere – Mannheimer Zentrum für Europäische Sozialforschung* 45, www.mzes.uni-mannheim.de, 08-2004
- Ittner, E. (1993):* Das Auswahlgespräch bei der Hochschulzulassung zu den medizinischen Studiengängen (Dissertation). Regensburg
- Iyigun, M. F.; Levin, A. T. (1998):* Macroeconomic implications of competitive college admissions (International Finance Discussion Papers No. 613): Board of Governors of the Federal Reserve System
- Katchadourian, H.; Boli, J. (1994):* Cream of the crop: The impact of elite education in the decade after college. New York
- Kelman, E. G.; Canger, S. (1994):* Validity of interviews for admission evaluation. In: *Journal of veterinary medical education* 21, www.utpjournals.com/JVME/JVME.html, 07-2004
- Kirchner, G. L.; Holm, M. B. (1997):* Prediction of academic and clinical performance of occupational therapy students in an entry-level master's program. In: *American journal of occupational therapy* 1997, S. 775–779
- Kobrin, J. L.; Camara, W. J.; Milewski, G. B. (2002):* The utility of the SAT I and SAT II for admission decisions in California and the nation (College Board Research Report No. 2002-6). New York

Konegen-Grenier, C. (2002): Studierfähigkeit und Hochschulzugang (Projekt im Rahmen des von der Informedia-Stiftung – Gemeinnützige Stiftung für Gesellschaftswissenschaften und Publizistik, Köln, geförderten Rahmenprojektes „Egalität und Effizienz – das deutsche Modell auf dem Prüfstand“). Köln

Kuncel, N. R.; Hezlett, S. A.; Ones, D. S. (2001): A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate students. In: *Psychological bulletin* 127 (2001), S. 163–181

Kurz, G.; Fischer, W.; Wagner, H. (1995): Prognose des Studienerfolgs in Studiengängen des Maschinenbaus – Untersuchungen an der FHT Esslingen und an der FH Karlsruhe. In: *Empirische Pädagogik* 9 (1995), S. 331–360

Lewin, D. (2004): Qualität aufgrund hochschuleigener Auswahlverfahren? In: *Das Hochschulwesen* 52 (2004), S. 15–18

Lewin, D.; Lischka, I. (2004): Passfähigkeit beim Hochschulzugang als Voraussetzung für Qualität und Effizienz von Hochschulbildung. Wittenberg

Lewin, K.; Heublein, U.; Schreiber, J.; Spangenberg, H.; Sommer, D. (2001): Studienanfänger im Wintersemester 2000/2001: Trotz Anfangsschwierigkeiten optimistisch in die Zukunft. Hannover

Lewin, K.; Heublein, U.; Sommer, D. (2000): Differenzierung und Homogenität beim Hochschulzugang. In: *HIS-Kurzinformation A* 2000, 7

Lischka, I. (1999): Studentische Nachfrage und Erwartungen der Wirtschaft. In: Olbertz; J.-H.; Pasternack, P. (Hrsg.): *Profilbildung – Standards – Selbststeuerung. Ein Dialog zwischen Hochschulforschung und Reformpraxis*. Weinheim, S. 125–135

Löwer, W. (1996): Der richtige Weg zur Reform? Hochschuleingangsprüfungen aus juristischer Perspektive. In: *Forschung & Lehre* 3 (1996), S. 629–631

Markl, H. (1997): Die Auswahl des Besten – Methoden und Wirkungen. In: Max-Planck-Gesellschaft (Hrsg.): *Internationalität der Forschung*. München, S. 278–294

Meier, B.-D. (2003): Ist der Erfolg im Jurastudium vorhersagbar? Empirische Befunde zum Zusammenhang zwischen Schulnoten und Abschneiden im Ersten Juristischen Staatsexamen. In: *Beiträge zur Hochschulforschung* 25 (2003), 4, S. 18–35

Meulemann, H. (1995): Die Geschichte einer Jugend. Lebenserfolg und Erfolgsdeutung ehemaliger Gymnasiasten zwischen dem 15. und 30. Lebensjahr. Opladen

Meyer, H. J. (1996): Studierfähigkeit und Hochschulzugang. In: Meyer H. J.; Müller-Böling, D. (Hrsg.): *Hochschulzugang in Deutschland. Status quo und Perspektiven*. Gütersloh, S. 15–40

- Powers, D. E. (1993):* Coaching for the SAT: A summary of the summaries and an update. In: Educational measurement: Issues and practice 12 (1993), S. 24–30
- Sadler, J. (2003):* Effectiveness of student admission essays in identifying attrition. In: Nurse education today, 23 (2003), S. 620–627
- Schwirten, C. (2000):* Stellenwert und Kriterien der Studentenauswahl an US-Hochschulen. In: Ergänzungsheft der Zeitschrift für Betriebswirtschaft 69 (2000), S. 37–55.
- Sieglein, J. (1998):* Erklärung und Vorhersage außergewöhnlicher beruflicher Leistungen. Frankfurt a. M.
- Stephan, H. (2002):* Lässt sich aus der Abiturgesamtnote die Studieneignung erkennen? In: Zeitschrift für Rechtspolitik 35 (2002)
- Streng, F. (1996):* Determinanten und Indikatoren von Examenserfolg und Studiendauer im Jurastudium. Ergebnisse einer Absolventenbefragung der Juristischen Fakultät Erlangen. In Herrmann, D.; Tag, B. (Hrsg.): Die universitäre Juristenausbildung. Bonn, S. 32–55
- Studley, R. (2003):* Inequality, student achievement, and college admissions: a remedy for underrepresentation <http://ishi.lib.berkeley.edu/cshe/>, 05-2004. Berkeley
- Teichler, U. (1999):* Profilierungspfade der Hochschulen im internationalen Vergleich. In Olbertz J.-H.; Pasternack, P. (Hrsg.): Profilbildung – Standards – Selbststeuerung. Ein Dialog zwischen Hochschulforschung und Reformpraxis. Weinheim, S. 27–38
- TNS Emnid. (2004):* Die Hochschulreform aus Sicht der Studierenden. Eine Studie im Auftrag der Initiative Neue Soziale Marktwirtschaft www.chancenfueralle.de/Downloads/PDF_-_Dateien/040421_Hochschulreform.pdf, 07-2004
- Trost, G.; Blum, F.; Fay, E.; Klieme, E.; Maichle, U.; Meyer, M.; Nauels, H.-U. (1998):* Evaluation des Tests für medizinische Studiengänge (TMS): Synopse der Ergebnisse. Bonn
- Tutt, L. (1997):* Der Studienentscheidungsprozeß. Informationsquellen, Informationswünsche und Auswahlkriterien bei der Hochschulwahl: Universität Gesamthochschule Duisburg, Fachbereich Wirtschaftswissenschaft. Duisburg
- Ulbricht, K. (1986):* Studienerfolgsprognose. Chancen und Grenzen für die Beratung vor Studienbeginn. In: Technik-Didakt 11 (1986), S. 133–142
- Wiesner, W. H.; Cronshaw, S. F. (1988):* A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview. In: Journal of occupational psychology 72 (1988), S. 484–487
- Wolter, A. (1999):* Neuordnung des Hochschulzugangs durch hochschuleigene Auswahlverfahren. Motive, Modelle, Erfahrungen, Perspektiven. In: Olbertz, J.-H.; Pasternack, P.

(Hrsg.): Profilbildung – Standards – Selbststeuerung. Ein Dialog zwischen Hochschulforschung und Reformpraxis. Weinheim, S. 269–299

Anhang: Abkürzungsverzeichnis:

ACT	American College Testing
GMAT	Graduate Management Admission Test
GRE	Graduate Record Examination
LSAT	Law School Admission Test
MCAT	Medical College Admission Test
PCAT	Pharmacy College Admission Test
SAT	ursprünglich: Scholastic Aptitude Test
TMS	Test für medizinische Studiengänge

Anschrift der Verfasserin:

Mareike Tarazona
Zentrum für Bildungsforschung
Universität Flensburg
Auf dem Campus 1
24943 Flensburg