

Studienerfolgsprognose in der Bundeswehr  
- Evaluation vorhandener und zukünftiger Prädiktoren -

Inaugural-Dissertation  
zur Erlangung der Doktorwürde  
der  
Philosophischen Fakultät  
der  
Rheinischen Friedrich-Wilhelms-Universität  
zu Bonn

vorgelegt von

Larissa Krex  
aus  
Bonn

im Januar 2008

Gedruckt mit der Genehmigung der Philosophischen Fakultät  
der Rheinischen Friedrich-Wilhelms-Universität Bonn

Zusammensetzung der Prüfungskommission:

Privatdozent Dr. Bernd Schlöder  
(Vorsitzender)

Professor Dr. Georg Rudinger  
(Betreuer und Gutachter)

Emeritus Professor Dr. Walter Neubauer  
(Gutachter)

Professor Dr. Rainer H. Kluwe (Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg)  
(weiteres prüfungsberechtigtes Mitglied)

Tag der mündlichen Prüfung: 19. Dezember 2007

Die Dissertation ist auf dem Hochschulschriftenführer der ULB Bonn [http://hss.ulb.uni-bonn.de/diss\\_online](http://hss.ulb.uni-bonn.de/diss_online)  
elektronisch publiziert

## Danksagung

Diese Zeilen möchte ich all denjenigen Personen aus meinem beruflichen und privaten Umfeld widmen, die mich im Rahmen meiner Inauguraldissertation in vielfältiger Weise motivational und emotional unterstützten. Diesen und denen hier nicht namentlich erwähnten Menschen gebührt mein tiefer Dank.

Meinem universitären Betreuer und Erstgutachter Herrn Professor Dr. Georg Rudinger möchte ich für seine stets spontane und bereitwillige Hilfe wissenschaftlicher und praktischer Natur herzlich danken.

Ebenfalls danke ich meinem Zweitgutachter Emeritus Professor Dr. Walter Neubauer, mit dem ich noch lange vor dem Schreiben der ersten Zeilen das Thema konkretisierte.

Sowohl mein Referatsleiter im Bundesministerium der Verteidigung (BMVg) Ministerialrat Dr. Hans-Dieter Hansen, als auch mein Gruppenleiter Leitender Regierungsdirektor außer Dienst Herbert Aschenbrenner trugen seit meinem Praktikum 2001 und mit dem Arbeitsbeginn 2004 entscheidend zur Verwirklichung meines Wunsches einer bundeswehrspezifischen Doktorarbeit bei. Sie standen mir auch in scheinbar ausweglosen Situationen dank ihres langjährigen Erfahrungsschatzes mit Vertrauen und Rat zur Seite. Besonders dankbar bin ich meinem Dezernatsleiter Regierungsdirektor Dr. Albert Melter für zahlreiche wertvolle Gespräche personalpsychologischer Art und die vielen wertvollen und konstruktiven bundeswehrinternen und die Bürokratie überwindenden „Überlebensstipps“. Bei der Umsetzung, Programmierung und technischen Vorbereitung auf das computerassistierte Testsystem standen mir meine lieben Kolleginnen Oberregierungsrätin Bettina Wölm und Ursula Hammes sowie ihr Sachgebietsleiter Oberregierungsrat Bernd Meinardus zur Seite. Bei Computerabstürzen waren Sie telefonisch jederzeit erreichbar. Danken möchte ich auch Sabine Ebeling, die während meiner Abwesenheit in den Pilotphasen die zahlreichen Anfragen der Arbeitsgruppe Personalergänzung und anderer Projekte entgegennahm und dem zweijährigen steinigen Weg der Lebenslaufdatenanforderung wacker trotzte. Förderliche Impulse erhielt ich schließlich nicht minder von Seiten vieler ziviler und militärischer Kollegen und nicht zuletzt durch die Präsidenten und Professoren an den Universitäten der Bundeswehr.

Besonderem Dank gebührt Professor Dr. Rainer H. Kluwe, der als einer der wenigen zivilen Mitarbeiter einer Kooperation zwischen akademischer Lehre mit ihrer zivilen Perspektive und der militärischen Organisation unversperrt gegenüber stand und zu jeder Zeit ein offenes Ohr für meine Anfragen hatte. Beinahe undenkbar gewesen wäre die Durchführung der vielen Testungen und Interviews ohne das Organisationstalent und das „studentische Insiderwissen“ von Oberleutnant zur See Jens Wollert. Ich möchte allen Mitarbeitern im Team aufrichtig danken.

Von besonderem Wert war mir nicht minder der emotionale Beistand meiner Familie und meiner engsten Freunde Nadija Hochstein und René Defossar. Meinem Freund Uwe, Hauptmann der Reserve, bin ich für seinen unschätzbaren liebevollen Rückhalt und seine fachfremden politisch, juristisch sowie insbesondere militärisch geprägten Anregungen während der letzten acht Jahre unserer Partnerschaft sehr verbunden. Ohne seine Liebe und die vielen Diskussionen wäre vieles uninteressant und unentdeckt geblieben.

Mein größtes Dankempfinden gebührt meinen Eltern, denen diese Arbeit gewidmet ist. Sie waren während meiner Dissertation und in allen Lebenslagen immer für mich da, unterstützten mich bei all meinen Plänen und gaben mir von einem immer mehr als andere: Ihre Liebe.

Bonn, im Juni 2007

Das große Ziel der Bildung ist nicht Wissen, sondern Handeln.

Herbert Spencer (1820-1903),  
englischer Philosoph und Sozialwissenschaftler

## Zusammenfassung

Der Offizierberuf stellt an sich und angesichts des veränderten und gestiegenen Aufgabenspektrums einer Einsatzarmee und aufgrund knapper Ressourcen sowie operativer, organisatorischer, verfahrenstechnischer und technologischer Entwicklungen in erhöhtem Maße massive Ansprüche an ein Auswahlssystem. Dazu zählt neben der Bewertung der allgemeinen Offiziereignung die Eignung zum Diplomstudium an einer der beiden Universitäten (UniBw) oder an der Fachhochschule der Bundeswehr. Der Gruppe Wehrpsychologie im Streitkräfteamt wurde im Herbst 2004 das Langzeitprojekt „Entwicklung geeigneter Prädiktoren zur Vorhersage der Studieneignung an den UniBw“ auf Weisung des Bundesministeriums für Verteidigung übertragen. Übergeordnetes Ziel ist die Senkung der Abbruchquoten an den UniBw. Durchschnittlich bricht circa ein Drittel der studierenden Offizieranwärter nach einer bereits absolvierten mindestens 15-monatigen militärischen Offizierausbildung das Studium ab und scheidet kurze Zeit später aus der Bundeswehr aus.

Wesentliches Ziel der vorliegenden Arbeit war die Optimierung der Studienberatung im Eignungsfeststellungsverfahren der Offizierbewerberprüfzentrale mittels einer studienfachspezifischen Reihungstestbatterie für die Studienfelder Geistes-, Ingenieur-, Wirtschaftswissenschaften und Informatik. Die Folgerungen erheben den Anspruch, eine langfristige Orientierungs- und Entscheidungshilfe bei Einführung neuer und Veränderung bestehender Eignungsfeststellungsverfahren für die Zielgruppe zukünftiger Offiziere zu liefern. Zu diesem Zweck wurde die Zusammenstellung und Entwicklung eines computergestützten Testverfahrens mit Modellcharakter, deren Evaluation und nähere Betrachtung kognitiver wie nicht-kognitiver und moderierender Einflussfaktoren geleistet.

Im Mittelpunkt der Evaluation stand die Ermittlung der Testgütekriterien Reliabilität, Konstrukt- und prädiktive Validität sowie Akzeptanz des Verfahrens. Die Voraussetzung einer vorhandenen Reliabilität für die Validität wird mit Ausnahme des spezifischen Testteils des Fachbereichs Informatik und der beiden neu entwickelten Textverständnisaufgaben für den geisteswissenschaftlichen Anteil für alle Testbatterien bestätigt. Die Ergebnisse belegen ferner, dass sich Konstruktvalidität hinreichend nachweisen lässt. Die Analysen zum nomologischen Netzwerk decken das Vorhandensein konvergenter wie diskriminanter Validität für die Persönlichkeitsfragebögen auf. Die Untersuchungen zur prädiktiven Validität unterstützen insbesondere die Bedeutsamkeit weicher Faktoren wie Persönlichkeitseigenschaften und Arbeitshaltungen. Gestützt wird außerdem die hohe Prädiktorkraft der Abiturnote. Berechnungen zum Einfluss moderierender Hintergrundvariablen deckten positive Korrelationen zwischen Kürze der Studienzeit und Studienerfolg auf. Im Unterschied zu Evaluationsuntersuchungen des Auswahlverfahrens in der Offizierbewerberprüfzentrale (OPZ) findet sich kein Einfluss der Variable Geschlecht, TSK oder soziokultureller Hintergrund auf den Studienerfolg. Die Akzeptanz des untersuchten Instruments ist als sehr hoch zu bezeichnen.

Mögliche Konsequenzen für die weitere Forschung und Implikationen für die Praxis in und außerhalb der Bundeswehr werden diskutiert.



A	Präliminarien.....	1
B	Theoretischer Teil .....	6
1.	Konstrukt Studieneignung .....	6
2.	Prädiktoren und Kriterien von Studienerfolg .....	7
2.1	Prädiktoren.....	7
2.1.1	Auswahl über die „Black Box“ Abiturnote.....	8
2.1.2	Auswahl über Einzelfachnoten.....	11
2.1.3	Auswahl über Studieneingangstestleistungen .....	12
2.1.4	Auswahl über biographische Fragebögen oder Referenzschreiben.....	18
2.1.5	Auswahl über Auswahlgespräche.....	20
2.1.6	Auswahlkriterium Berufsausbildung .....	21
2.1.7	Auswahl über Assessment Center .....	21
2.1.8	Trends in der Studierendenauswahl.....	24
2.1.9	Exkurs akademische Intelligenz.....	27
2.1.10	Exkurs Persönlichkeitsmerkmale .....	29
2.1.11	Studiergewohnheiten und Einstellungen zum Studium .....	32
2.2	Zusammenfassung Prädiktoren des Studienerfolgs.....	33
2.2.1	Typologie der Auswahlverfahren .....	35
2.3	Kriterien.....	36
2.3.1	Exkurs Kriterienproblematik.....	36
2.3.2	Kriterium Studienz Zwischen- und -abschlussnoten .....	38
2.3.3	Kriterium Studienabschluss, Studienfachwechsel und Studienabbruch.....	39
2.3.4	Kriterium Studiendauer.....	39
2.3.5	Kriterium Studienzufriedenheit.....	40
2.3.6	Kriterium Beruf .....	40
2.4	Zusammenfassung Kriterien des Studienerfolgs.....	41
3.	Interne Struktur des Konstrukts Studieneignung .....	41
3.1	Theoretische Modelle der Studieneignung .....	41
3.2	Zuordnungsmodelle.....	43
3.3	Prädiktionsmodelle.....	44
3.3.1	Univariate Prädiktion.....	45
3.3.2	Multivariate Prädiktion .....	46
3.3.3	Moderierte Prädiktion.....	46

3.4	Bedingungsmodell des Studienerfolgs .....	47
3.5	Studienabbruchmodelle .....	48
3.6	Modelle der Person-Umwelt-Passung .....	54
3.7	Exkurs Anforderungsanalyse .....	56
4.	Zusammenfassung Konstrukt Studieneignung .....	61
5.	Veränderte Rahmenbedingungen .....	62
6.	Die Offizierauswahl an der Offizierbewerberprüfzentrale (OPZ) .....	64
6.1	Planung .....	65
6.2	Durchführung .....	66
6.2.1	Bewährungswahrscheinlichkeit und Eignungsgrad .....	67
6.2.2	Studienfachrichtungsempfehlung .....	70
6.2.3	Prüfstationen und zu messende Dimensionen .....	72
7.	Das Studium an den Universitäten der Bundeswehr .....	78
7.1	Studienabbruch an den Universitäten der Bundeswehr .....	82
7.2	Vergleich zivile versus militärische Universitäten in Deutschland .....	84
7.3	Zusammenfassung Offizierauswahl und -studium .....	86
C	Empirischer Teil .....	90
8.	Planung .....	91
8.1	Ziele und Nutzen .....	91
8.2	Festlegung und Definition der Anforderungsmerkmale .....	95
8.3	Die Testbatterie Studieneignung .....	96
8.3.1	Allgemeiner Studierfähigkeitstestteil der Pilotphasen .....	97
8.3.2	Spezifischer Studierfähigkeitstestteil der Pilotphasen .....	98
8.4	Werbung freiwilliger Testpersonen .....	100
9.	Durchführung .....	100
9.1	Testdurchführung und -auswertung .....	101
10.	Forschungsfragestellungen .....	101
10.1	Reliabilitätsbestimmung .....	103
10.2	Validitätsbestimmung .....	105
10.2.1	Konstruktvalidität .....	107
10.2.2	Prognostische Validität .....	109
10.3	Annahmen zum Einfluss von Moderatorfaktoren .....	111
10.4	Untersuchte Stichprobe .....	113

10.4.1	Aufbereitung der Stichprobendaten.....	115
10.5	Variablenbeschreibung .....	117
10.5.1	Verteilung der Prädiktoren.....	119
10.5.2	Verteilung des Kriteriums.....	128
10.5.3	Verteilung der Moderatorvariablen .....	129
11.	Ergebnisse.....	130
11.1	Reliabilität.....	131
11.1.1	Itemanalyse Allgemeiner Teil I.....	132
11.1.1.1	Der Untertest Figuren erkennen.....	132
11.1.1.2	Der Untertest Verbale Analogien - adaptive Testform .....	133
11.1.1.3	Testgütemaße.....	134
11.1.2	Normen Allgemeiner Teil I.....	138
11.1.3	Itemanalyse Spezifischer Teil .....	140
11.1.3.1	Studienfeld Geisteswissenschaften - Struktur und Messbereiche ...	140
11.1.3.2	Testgütemaße Geisteswissenschaften .....	141
11.1.3.3	Studienfeld Wirtschaftswissenschaften - Struktur und Messbereiche .....	146
11.1.3.4	Testgütemaße Wirtschaftswissenschaften.....	147
11.1.3.5	Studienfeld Ingenieurwissenschaften - Struktur und Messbereiche .....	151
11.1.3.6	Testgütemaße Ingenieurwissenschaften.....	153
11.1.3.7	Studienfeld Informatik - Struktur und Messbereiche.....	157
11.1.3.8	Testgütemaße Informatik .....	158
11.1.4	Normen Spezifischer Teil .....	160
11.1.5	Itemanalyse Allgemeiner Teil II (Persönlichkeitstests).....	166
11.1.5.1	Persönlichkeitsmerkmale Offizierbewerber .....	167
11.1.5.2	Big Five.....	169
11.1.5.3	Stressbewältigungsressourcen-Fragebogen.....	171
11.1.5.4	Testgütemaße Allgemeiner Teil II (Pilotphasen I und II).....	174
11.1.6	Gesamtreliabilität Allgemeiner Teil I.....	175
11.1.7	Gesamtreliabilität Allgemeiner Teil II.....	176
11.1.8	Gesamtreliabilität Spezifischer Teil .....	177
11.2	Konstruktvalidität.....	177

11.3	Prognostische Validität.....	181
11.4	Einfluss moderierender Hintergrundvariablen.....	196
11.5	Prozessevaluation .....	199
11.5.1	Biographischer Fragebogen.....	199
11.5.2	Akzeptanzfragebogen .....	200
11.5.3	Feedbackgespräch.....	206
D	Diskussion .....	208
12.	Evaluation der Ergebnisse .....	209
12.1	Diskussion der Ergebnisse aus methodischer Sicht .....	209
12.1.1	Reliabilität .....	209
12.1.2	Konstruktvalidität .....	210
12.1.3	Kriteriumsvalidität.....	210
12.1.4	Weitere methodische Einflüsse und Fazit.....	212
12.2	Zusammenfassung und Diskussion der Ergebnisse aus empirischer und theoretischer Sicht .....	213
12.2.1	Reliabilität .....	213
12.2.2	Konstruktvalidität .....	214
12.2.3	Kriteriumsvalidität.....	216
12.2.4	Moderierende Hintergrundvariablen.....	219
12.2.5	Evaluation .....	221
13.	Implikationen für Praxis und Forschung .....	222
13.1.1	Methodische Implikationen.....	222
13.1.2	Fazit für die Praxis.....	224
13.1.3	Ausblick .....	226
E	Abbildungsverzeichnis .....	233
F	Tabellenverzeichnis .....	237
G	Literaturverzeichnis.....	241
H	Anhang .....	261

## A PRÄLIMINARIEN

Mit der Übernahme einer Führungsposition wird die Möglichkeit geschaffen, einen starken Einfluss auf ein Organisationsgeschehen auszuüben. Die Qualität der ausgewählten Personen ist demzufolge eine bedeutende Determinante für den gegenwärtigen und zukünftigen Erfolg einer Organisation. Für die Bundeswehr wie für jede andere Einrichtung sind personelle Entscheidungen umso bedeutsamer, je mehr Einfluss diejenige Person, über die entschieden wird, im Laufe ihrer Zugehörigkeit zu dieser Organisation auf deren Effizienz hat und je schwieriger Korrekturen von Fehlentscheidungen sind. Die Entscheidung, welcher Bewerber<sup>1</sup> angenommen oder abgelehnt wird, ist eine Zuordnungsentscheidung, bei der die deutschen Streitkräfte von der Bewährung der angenommenen Bewerber in ihrer späteren militärischen Verwendung ausgehen. Gerade der Führungsberuf des Offiziers in seiner qualitativen Komplexität und Vielfalt militärischer Tätigkeitsfelder verlangt eine Vielzahl von Anforderungen. Knebel (1995) sieht in ihm ein „Universalgenie“, das situationsabhängig unterschiedliche Führungsstile erfolgreich anwenden kann und dessen Rolle den Führer, Ausbilder und Erzieher von Soldaten in einer Person vereint. Die an den Offizier gerichteten Ansprüche betreffen neben seiner körperlichen Eignung und Leistungsbereitschaft auch seine persönlichen, ethischen und insbesondere intellektuellen Voraussetzungen. Die Bandbreite seiner wahrzunehmenden Aufgaben reicht beim Offizier demnach vom soldatischen Profil des Menschenführers, des Stabsarbeiters, des Spezialisten und Kämpfers über das Spektrum des Retters, Helfers, Beobachters, Beschützers und Bewahrs bis zum vermittelnden Einsatz als Diplomat in einem multinationalen Umfeld. Der Offizierberuf stellt an sich und angesichts des veränderten und gestiegenen Aufgabenspektrums aufgrund knapper Ressourcen sowie operativer, organisatorischer, verfahrenstechnischer und technologischer Entwicklungen in erhöhtem Maße massive Ansprüche an ein Auswahlssystem. Dazu zählt neben der Bewertung der allgemeinen Offiziereignung die Eignung zum Diplomstudium an einer der beiden Universitäten oder an der Fachhochschule der Bundeswehr. Analog zur aktuellen bildungspolitischen Debatte setzen sich auch die deutschen Streitkräfte mit einer optimierten Passung angehender Offiziere auf das ihnen nach Empfehlung zugewiesene Studienfach auseinander.

---

<sup>1</sup> Aus Gründen der Lesbarkeit wird - wenn nicht ausdrücklich darauf hingewiesen - auf weibliche Wortendungen verzichtet.

Hochschuleigene Auswahlverfahren sind aber nicht nur ein politisch äußerst brisantes, sondern auch im Hinblick gewünschter Profilbildung, wachsender Diversifikation bzw. horizontaler Differenzierung (Wolter, 2005) der Studienangebote und steigenden Wettbewerbs der Universitäten und Fachhochschulen im Sinne einer vertikalen Differenzierung ein stark diskutiertes Thema. Ebenso steuern die Hochschulen eine Homogenisierung des Leistungs- bzw. Vorkenntnisniveaus innerhalb der Studiengänge und angesichts zu hoher Abbrecherquoten selbstverantwortliche Zuordnungsverfahren an. Das jüngst novellierte Hochschulrahmengesetz weist in dem Zusammenhang der deutschen Alma Mater eine einflussreichere Rolle zu. Mit dem Wintersemester 2005/06 wird diesen gestattet, „ihre“ Studierenden zu 60 Prozent eigenständig auswählen. Parallel zu den neuen Bedingungen des deutschen Hochschulsektors wachsen auch die Anforderungen an die Qualität der Studien- und Hochschulwahlentscheidung des einzelnen Studieninteressierten. Demgemäß suchen Universitäten unumgänglich nach validen, praktikablen und effizienten Eignungsfeststellungsverfahren, deren Qualität durch hinreichende Evaluationsbefunde im In- und Ausland nachgewiesen wurde. Dabei darf im Kampf um die besten „Köpfe“ Deutschlands nicht übersehen werden, dass zu aufwendig gestaltete Auswahlverfahren leicht das übergeordnete Ziel der Zulassungsentscheidung zum Studium zugunsten einer Zurückweisung zum Studium verfehlen. Jede Hochschule muss sowohl die eigenen Angebote und deren Qualität als auch die Population potentieller Studierender im Blick behalten und danach Auswahlverfahren und -ziele im Zuge einer formativen Evaluation anpassen. Doch nicht nur das Hochschulsystem an sich ist zahlreichen Veränderungen im Wettbewerb um besonders fähige Studierende unterworfen. Ebenso wird die Personenvariable mit dem Hinneinwachsen geburtenschwacher Jahrgänge in das studierfähige Alter auch Art und Ausmaß des Wettstreits zwischen öffentlichen und privaten Hochschulen um die besten Studenten bestimmen. Einerseits bieten die Universitäten maßgeschneiderte Studienangebote an, andererseits suchen sich Bewerber speziell auf ihr Fähigkeitsprofil zugeschnittene Studiengänge aus. Durch ein gezieltes Beratungskonzept kann dahingehend schon im Vorhinein die größtmögliche Passung zwischen Studierenden und Studienangebot erreicht und Korrekturen oder Fehlentscheidungen vermieden werden.

Bis dato waren die eignungsdiagnostischen Methoden der Bundeswehr dem allgemeinen Auswahl- und Verteilungsverfahren der Landesuniversitäten weit voraus. Die Selektion erfolgt hier schon lange nicht allein über die Abiturdurchschnittsnote. Im Falle der medizinischen Studiengänge verfügt die Bundeswehr sogar über einen studienfachspezifischen

Eignungstest. In Anbetracht der zu hohen Anzahl von Studiengangwechslern, der hohen Studienabbecherquoten verbunden mit einem zwangsläufigen und kostspieligen Ausscheiden aus der Bundeswehr ist eine Evaluation und Optimierung der derzeitigen Auswahl unumgänglich.

Der exklusiv für die OPZ entwickelte Studieninteressenfragebogen (STIFBO) kann in Kombination mit der Abiturnote, mit den vorhandenen allgemeinen Intelligenztests sowie dem abschließenden Studienberatungsgespräch lediglich eine ausreichende bis befriedigende Vorhersagevalidität erfüllen. Sowohl die während der Pilotphasen mit allen Testpersonen durchgeführten im Durchschnitt 20-30-minütigen Feedbackgespräche als auch der eingesetzte Akzeptanzfragebogen ließen jedoch einen deutlichen Mangel an standardisierter Studienberatung, fachspezifischer Testnähe und Erhebung von für das Studieren an einer UniBw relevanten Persönlichkeitsmerkmalen sichtbar werden. Schon jetzt kann geschlossen werden, dass auch Persönlichkeitsmerkmale wie Gewissenhaftigkeit, Leistungsmotivation, Offenheit und Selbstdisziplin einen nicht zu vernachlässigenden Einfluss auf den Studienerfolg der Offiziere haben.

Generell kann die *Abiturdurchschnittsnote vor Einzelfachnoten* als ein bedeutsamer Prädiktor für den in Prüfungen festgestellten Studienerfolg gelten. Ihre Heranziehung als ein Kriterium zur Auswahl guter Studienanfänger wird durch empirische Ergebnisse untermauert. Aufgrund vieler Nachteile der Abiturnote als alleiniger Prädiktor des Studienerfolgs sollten *zusätzlich Studieneignungstests* als Alternative zur Erhebung der Fähigkeiten, die durch die Abiturnote nicht oder nicht ausreichend abgedeckt werden, eingesetzt werden. Zudem kann bei der Berücksichtigung von Tests der Kampf um Zehntelnoten und die Oberstufen-Fächerwahl anhand der Notenerwartung vermieden werden.

Die vorliegende Arbeit beschäftigt sich mit Korrelaten vorhandener wie auch zukünftiger Prädiktoren von Studienerfolg im Rahmen der Offizierauswahl in den deutschen Streitkräften. Mit dem zentralen Ziel dieser Arbeit, einen ebenso umfangreichen wie nach Studienfeldern differenzierten Überblick über bestehende Korrelate von Studienerfolg insbesondere im Rahmen der Offizierauswahl und -ausbildung der Bundeswehr liefern zu können, stehen folgende Fragestellungen im Vordergrund:

1. Welche Resultate liefert die empirische Untersuchung an den beiden Universitäten der Bundeswehr in Hamburg und München hinsichtlich geeigneter Prädiktoren zur Vorhersage der Studieneignung?

2. Wie gut werden intendierte Konstrukte mittels untersuchter Reihungstests tatsächlich gemessen?
3. Wie leistungsfähig erweisen sich Qualität, Zuverlässigkeit, Objektivität, Effizienz und Umsetzbarkeit bestehender und untersuchter Prädiktoren?
4. Wo liegen Ansatzpunkte für eine Optimierung des AC-Verfahrens in der Offizierbewerberprüfzentrale?
5. Welche Schlüsse lassen Prozess- und Postevaluation hinsichtlich möglicher interner und externer Einflüsse zu?
6. Wie steht es mit der Akzeptanz der praktizierten und untersuchten Auswahlverfahren bei Offizieranwärtern?

Die Folgerungen erheben den Anspruch, eine langfristige Orientierungs- und Entscheidungshilfe bei Einführung neuer und Veränderung bestehender Eignungsfeststellungsverfahren für Offiziere in der Bundeswehr zu liefern. Dabei soll zugleich das Spezielle der in die Offizierausbildung integrierten Phase des Studiums an den beiden Campus-Universitäten im Vergleich zu Studiengängen an öffentlichen und übrigen privaten Hochschulen Berücksichtigung finden.

Zunächst wird jedoch im theoretischen Teil das Konstrukt Studieneignung hinsichtlich seiner internen Struktur skizziert und am Beispiel der Offizierauswahl und -ausbildung verdeutlicht. Der empirische Teil beschreibt die Projekthintergründe und praktische Umsetzung der Pilotphasen. Die theoretischen Forschungsfragestellungen werden in Form von Annahmen und Hypothesen dargestellt. Daran an schließt sich ihre Operationalisierung, die Deskription der untersuchten Stichprobe und die Variablenbeschreibung. Im Anschluss an eine ausführliche Itemanalyse zur Bestimmung der Verfahrensreliabilität werden eine Konstrukt- und prognostische Validierung sowie eine Untersuchung der Akzeptanz durchgeführt. Als Außenkriterien werden dabei die Vor- und Hauptdiplomnoten sowie zur genaueren Betrachtung der moderierenden Hintergrundvariablen auch die Ergebnisse aus der Theorie- und Praxisausbildung der Offiziere (Offizierlehrgang) herangezogen. Die Zusammenfassung inhaltlicher Kernaussagen, abgeleitet aus zugrunde liegenden Fragestellungen und den Ergebnissen, mündet in eine Erörterung methodischer, empirischer und inhaltlicher Sichtweisen und in eine Diskussion über Perspektiven und Anregungen für die Praxis.

Die vorliegende Inauguraldissertation erhebt den Anspruch, neben einer theoretischen Beschreibung des Themas und Interpretation empirisch gewonnener Erkenntnisse auch die unmittelbare Anwendung in der Praxis zu gewährleisten.

## B THEORETISCHER TEIL

Der theoretische Teil verfolgt das Ziel, anhand detaillierter Definition möglicher Prädiktoren und Kriterien die interne Struktur des Konzepts Studieneignung und somit eine Annäherung an die Diagnose von Studierfähigkeit zu liefern. Dabei sollen in den Kapiteln „Veränderte Rahmenbedingungen“ und „Vergleich zivile versus militärische Universitäten in Deutschland“ die Besonderheiten, denen werdende Offiziere ausgesetzt sind, diskutiert und auf das Ziel der vorliegenden Arbeit hinlenken. In dem Zusammenhang wird das AC Bw und das an die Offizierausbildung anschließende Studium an den UniBw vorgestellt. Zusammenfassungen und Schlussfolgerungen reflektieren die jeweiligen Kapitel und leiten zum empirischen Teil der vorliegenden Arbeit fließend über.

### 1. Konstrukt Studieneignung

Trotz größter Bemühungen in zahlreichen Ländern, die den Studienerfolg bestimmenden Merkmale und Prozesse näher zu ergründen, bleibt „in der Regel mehr als die Hälfte dessen, was Studieneignung ausmacht, unerfasst“ (Trost, 1984, S. 144). Bis heute ist das Konstrukt Studieneignung nicht vollends bestimmbar. Neben den vielen nicht-standardisierbaren moderierenden Hintergrundvariablen während des Studiums (Qualität der Lehre, Finanzierung des Studiums, private Situation, Arbeitsmotivation, Interesse des Studenten usw.) führen eine vergleichsweise geringere Zahl standardisierbarer Bedingungen dazu, dass statt eigentlicher Beziehungen vielmehr Scheinzusammenhänge gemessen werden (Amelang, 1997). In den siebziger Jahren gliedert sowohl der Kreis wissenschaftlicher Berater (1977) als auch Michel (1977) das Gesamtkonstrukt Studieneignung in allgemeine, feldspezifische und fachspezifische Komponenten. Trost (1975) definiert Studierfähigkeit als breites Spektrum individueller Voraussetzungen. Dazu zählt er neben intellektuellen Fähigkeiten und studienrelevantem Vorwissen auch wirkungsvolle Lern- und Arbeitstechniken, Arbeitsdisziplin, Fleiß und Durchhaltevermögen, Selbstkritik sowie emotionale Stabilität. Der Hochschulverband (1984) postuliert neben materialen bzw. fachlichen Bildungsvoraussetzungen auch personale (Persönlichkeitseigenschaften) und formale (allgemeine Fähigkeiten und Fertigkeiten), welche Heldmann (1984) als folgende fünf Dimensionen der Studierfähigkeit vorstellt: Ausbildungsbereitschaft, Vorhandensein elementarer Voraussetzungen für wissenschaftliches Arbeiten, Formen geistigen Tätigseins, Ausprägung der Persönlichkeit, Interesse und Engage-

ment. Die Expertenkommission zur Weiterentwicklung der gymnasialen Oberstufe der Kultusministerkonferenz beschreibt die Obersekunda-, Unter- und Oberprimainhalte mit der Trias vertiefter Allgemeinbildung, Wissenschaftspropädeutik und so genannter Schlüsselqualifikationen (Selbstregulation des Lernens, soziale Kompetenz und Eigenverantwortlichkeit) als Studierfähigkeit im engeren Sinne. Die Mitglieder des Deidesheimer Kreises (1993) kritisieren dagegen den Trugschluss, der Abiturnote die Funktion des zentralen Indikators allgemeiner Studierfähigkeit einzuräumen, ohne die Bedeutsamkeit studienfachspezifischer Dimensionen zu berücksichtigen. Unterschiedlichste Studienfachinhalte unterliegen gemäß ihrer Ansicht heterogenen Anforderungen, für deren spezifische Diagnose globale Auswahlverfahren vorzuziehen seien.

Konegen-Grenier (2001) definiert Studierfähigkeit als in seinem Sinne „komplexes Konstrukt von Fähigkeiten, Einstellungen und Wissensinhalten“, wobei die kognitiven Fähigkeiten eine herausragende Rolle einnehmen. Gleichzeitig existieren enge Verknüpfungen mit persönlichen Merkmalen, insbesondere persönlichen Arbeitshaltungen (z.B. Selbstkontrolle, Gewissenhaftigkeit, Belastbarkeit) und inhaltlichem Interesse.

Im Rahmen der Begriffsbestimmung lässt sich zusammenfassend ein Gefüge bestimmter Leistungsvoraussetzungen annehmen, welches als hypothetisches Konstrukt mit dem Terminus Studieneignung verstanden werden kann (Zimmerhofer, 2003). Die Komponenten Studierfähigkeit, Studierleistung und Studienerfolg sind unmittelbar mit dem Konstrukt Studieneignung verwoben. Die beschriebene Komplexität verlangt weiterhin eine Antwort auf die Frage nach ihrer internen Struktur (vgl. Kapitel 3).

## 2. Prädiktoren und Kriterien von Studienerfolg

Vor einer internen Strukturbestimmung soll sowohl möglichen Wegen zur Prädiktion als auch der Kriteriumsproblematik von Studienerfolg begegnet werden. In den nachstehenden Abschnitten werden die unterschiedlichen Verfahren zur Feststellung der Studieneignung beschrieben und deren Zweckdienlichkeit diskutiert.

### 2.1 Prädiktoren

Wie oben beschrieben erweist sich Studierfähigkeit als vielschichtiges Konstrukt von Fähigkeiten, Einstellungen und Wissensinhalten. Dabei nehmen die kognitiven Fähigkeiten eine herausragende Rolle ein. Gleichzeitig existieren enge Verknüpfungen mit persönlichen

Merkmale, insbesondere mit persönlichen Arbeitshaltungen wie zum Beispiel Selbstkontrolle, Gewissenhaftigkeit, Belastbarkeit oder inhaltliches Interesse (vgl. Kapitel 2.1.10). Zu den üblichen kognitiven Prädiktoren lassen sich Schulnoten, Testleistungen in Eignungs- oder Studierfähigkeitstests und Intelligenz- oder Schulleistungstests subsumieren. Zu den nicht kognitiven sind fachnahe Aufnahmeprüfungen wie z.B. in Sport, Fremdsprachen, Musik oder Kunst sowie außerdem Kreativitätstests, demographische Variablen, Persönlichkeitsmerkmale und durch Exploration erfasste Merkmale zu zählen. Charakteristika der Korrelate werden im Folgenden näher beleuchtet.

### 2.1.1 Auswahl über die „Black Box“ Abiturnote

Während in Ländern wie Großbritannien, in denen bereits die Wahl der Schwerpunktfächer (so genannte A-Level-Kurse) in den Schuljahren die Studienfächerwahl begrenzt, sieht das hiesige Bildungssystem die allgemeine Hochschulreife als hinreichende Voraussetzung für die Aufnahme eines beliebigen Studienfachs.

Der deutsche Hochschulzugang beruht auf zwei Prinzipien, die gleichsam das Spannungsfeld zwischen den beiden Grundelementen des Abiturs - Befähigung und Berechtigung - begründen. Das erste Prinzip besagt, dass eine Prüfung am Ende eines Bildungsgangs oder einer Bildungsstufe eine bedeutsam größere Chance ermöglicht, die Leistungsfähigkeit und den Leistungswillen des zu Prüfenden in einer angemessenen und zuverlässigen Weise zu erfassen, als dies in aller Regel eine Prüfung am Eingang eines neuen Abschnitts im Bildungsweg zu erreichen vermag. In der deutschen Bildungsgeschichte erhält dieser Grundsatz besonderen Nachdruck durch die Humboldtsche Unterscheidung zwischen dem Zweck der gymnasialen Bildung und dem Charakter eines wissenschaftlichen Studiums. Nach Wilhelm von Humboldt muss ein Hochschulstudium Bildung durch Wissenschaft bzw. Teilnahme am wissenschaftlichen Erkenntnisprozess sein. Zum Begriff der Hochschulreife gehört daher notwendigerweise ein solides Maß an Allgemeinwissen und eine entwickelte Methoden- und Sozialkompetenz. Da die Hochschulreife allgemein sein soll und sich prinzipiell in jedem Studiengang zu bewähren hat, muss sie weithin kanonischen Charakter haben, was wiederum einen breiten Konsens in Bezug auf den Inhalt und den Grad der durch das Abitur nachzuweisenden Studierfähigkeit erfordert. Eben diese Vorhersagbarkeit ist für das zuverlässige Funktionieren des Reifezeugnisses an der Schnittstelle zwischen Schule und Hochschule von entscheidender Bedeutung.

Das zweite Prinzip besagt, dass sich aus der allgemeinen Hochschulreife ein allgemeines Einschreiberecht ergibt. Mit Zunahme der Studienbewerberzahl in den letzten Jahrzehnten fand eine Gewichtsverlagerung weg vom Befähigungs- hin zum Berechtigungsprinzip des Abiturs statt, die u.a. zum aktuellen Gegenentwurf des Abiturs führte, nämlich der Einführung von universitätsspezifischen Hochschuleingangsprüfungen. Für zusätzliche Dynamik sorgt die von der Hochschulpolitik forcierte Diversifizierung des Studienangebotes und die Entwicklung standortspezifischer Fächerprofile, die die dominierende Rolle der „Black Box“ Abiturnote als Auswahlkriterium vollends in Frage stellt. Schulische Durchschnittsnoten können zukünftig „nicht die Funktion eines alleinigen Regulativs für den Hochschulzugang haben“ (Heine et al., 2006, S. 9).

Die Debatte um die Abiturnote als Befähigungs- versus Berechtigungsnachweis prägt ein ebenso viel diskutiertes Spannungsfeld zwischen der bekannt eingeschränkten Objektivität und Vergleichbarkeit der Schulnoten einerseits und einer anerkannt hohen Prognosekraft der Abschlussnoten und hoher Ökonomie des Verfahrens andererseits. Länderspezifische Fächerprofile und zahlreiche Wahlmöglichkeiten als Kennzeichen des föderalen Schulsystems in Deutschlands bedingen intransparente Leistungs- bzw. Notenstandards zwischen Lehrern, Schulen und Schulformen. Eine Vielzahl an Studien belegt übereinstimmend die mangelnde Vergleichbarkeit der Noten (Baumert & Watermann, 2000; Köller, Baumert & Schnabel, 1999; Rindermann, 2005). Ihre geringe Objektivität, unzureichende Beurteilungsübereinstimmungen und niedrige Wiederholungszuverlässigkeiten sprechen gegen ihren Einsatz zumindest als alleinigen Prädiktor von Studieneignung.

Ingenkamp (1986) betont überdies neben der diagnostischen die nicht zu unterschätzende pädagogische Funktion, wenn Abschlussnoten zum Beispiel anstelle einer Leistungsrückmeldung als Belohnung des unterrichtsstützenden Verhaltens benutzt werden.

Umso erstaunlicher wiegen die häufig replizierten gegensätzlichen Befunde zur durchschnittlichen Gesamtnote im schulischen Abschlusszeugnis als wichtigstes Einzelmaß zur Vorhersage der Studieneignung (Trost, 1995; Deidesheimer Kreis, 1997). In der oft zitierten Metaanalyse von Baron-Boldt, Schuler und Funke (1988) wird ein mittlerer Zusammenhang zwischen Abiturnote und Examensleistungen von  $r = .46$  festgestellt. Rindermann & Oubaid (1999 und 2005) ermitteln einen Durchschnittskorrelationskoeffizienten von  $r = .39$ . Hinneberg (2003) weist der Abiturnote im Rahmen von Untersuchungen baden-württembergischer Medizinstudenten ebenfalls ein erstaunliches Differenzierungsvermögen zu. Auch in den USA zeigt sich wiederholt, dass der Grade Point Average als mittlere Abschlussnote den

Studienerfolg sogar noch genauer ( $r = .41$  bis  $r = .53$ ) als die Abiturnote in Deutschland vorhersagt. Signifikant höhere Werte können kaum erreicht werden, da der Erfolg im Studium neben den individuellen Eingangsvoraussetzungen von zahlreichen weiteren Faktoren abhängt. Die auf den ersten Blick für alle Kritiker überraschenden Befunde erklärt Schuler (2001) mit der hohen Schnittmenge identischer Anforderungsdimensionen von Schul- und Studienleistungen, die umso größer ausfällt, je verschulter das Studium ist. Hell (2006) führt den Erfolg der Abiturnote als bester Prädiktor für Studienerfolg auf das Zusammenwirken von Persönlichkeitseigenschaften und Arbeitshaltungen mit kognitiven Fähigkeiten sowie auf „Aggregation zahlreicher Einzelmessungen“ zurück. „In statistischen Termini gesprochen führt die Aggregation zahlreicher Einzelmessungen mit normalverteilten Fehlerkomponenten im Mittel zu einem aussagefähigen Leistungsindikator“ (Hell, 2006, S. 3). Aus der Abiturdurchschnittsnote werden vorwiegend Rückschlüsse auf den erreichten schulischen Leistungsstand, aber auch auf allgemeine intellektuelle Fähigkeiten gezogen, die für erfolgreiches Studieren im Allgemeinen bedeutsam sind. Damit sind zugleich die prognostischen Grenzen aufgezeigt. Die Vorhersage von Studienerfolg über die Abiturdurchschnittsnote betrifft eine *generelle* Studieneignung. Sie sagt nichts aus über eine studienfachspezifische Eignung (vgl. Baron-Boldt, 1989; Weingardt, 1989). Die Vorhersagekraft der Abiturnote für den Studienerfolg ist somit für die verschiedenen Studienfächer unterschiedlich groß. Die höchsten Koeffizienten finden sich für Wirtschaftswissenschaftler in Höhe von  $r = .56$  und für Jurastudenten in Höhe von  $r = .38$  (Baron-Boldt, Schuler & Funke, 1988). Schuler (2001) hält diese Forschungslage für evident, da sich Schul- und Studienleistungen generell in einer ganzen Reihe von Anforderungsmerkmalen gleichen. Diese Annahme lässt sich durch empirische Befunde stützen, nach denen die Beziehungen zwischen Schulabschlussnoten und Studienleistungen umso enger ausfallen je „verschulter“ das Studium ist.

Im Rahmen von Abitur-Plus-Konzeptionen wird zudem versucht, die Vorteile der Abiturnote mit den Vorzügen psychometrischer Verfahren zu kombinieren, um die Vorhersagegenauigkeit zu steigern. In einer Studie um Brandstätter, Farthofer & Grillich (2001) erzielten studienfeldbezogene Leistungstests ähnliche Ergebnisse wie die Abiturnote. Erstere sagen etwas besser den Studienabbruch vorher, die Abiturnote sagt zu einem höheren Maße die Hochschul-Prüfungsnoten voraus. Einen deutlich besseren Prognosewert als die Abiturnote allein weist in derselben Untersuchung die Kombination von Abiturnote, studienfeldbezogenen Leistungstests und Persönlichkeitsmerkmalen auf (multiple Korrelation  $R = .60$ ).

Bei der gewichteten Kombination vom Test für medizinische Studiengänge (TMS) und Abitur liegt der durchschnittliche Prognosekennwert bei  $r = .61$ . Studierende, die über dieses Kombinationskriterium zugelassen wurden, traten im Mittel nach kürzerer Studiendauer zur ärztlichen Vorprüfung an (Trost, Klieme & Nauels, 1997). Imposanterweise zeigt sich eine Erfolgsquote für die Testbesten im TMS in Höhe von 97 Prozent. Für per Losverfahren zugelassene Studenten kalkulierte man dagegen eine Erfolgsquote in Höhe von 71 Prozent (Deidesheimer Kreis, 1997). An dieser Stelle muss darauf hingewiesen werden, dass der TMS sich auf das Messen intellektueller Fähigkeiten beschränkt. Unterdessen konnte die Arbeitsgruppe Bildungslebensläufe in einer gewichtigen Längsschnittstudie zeigen, dass die Intelligenz zwar die meisten Anteile an der Abiturnotenvarianz aufklärt, weiteren Merkmalsfaktoren – wie Intelligenz und Wissen, Arbeitshaltung, Interesse und Persönlichkeit, Ausbildungsniveau der Eltern, schulische Umgebung – aber auch Beachtung geschenkt werden müssen (Giese et al., 1981). Im direkten Vergleich zu Intelligenztests spiegeln Schulnoten nicht nur die kognitive Befähigung, sondern auch nicht-kognitive Kompetenzen und motivationale Einstellungen wider, die für alle Studienfächer unentbehrlich sind.

Aus ökonomischer Sicht spricht insbesondere die prinzipielle Verfügbarkeit bei inländischen Schulabgängern für die Verwendung der Abiturdurchschnittsnote als Auswahlkriterium.

### 2.1.2 Auswahl über Einzelfachnoten

Zum Prognosewert gewichteter Einzelfachnoten im Vergleich zur Abiturdurchschnittsnote existieren uneinheitliche Ergebnisse. Das Gros der Studien belegt jedoch, dass der Prognosewert von gewichteten Einzelfachnoten nicht höher bzw. in der Regel unter dem der Abiturnote liegt (Lissmann, 1976; Konegen-Grenier, 2001; Meier, 2003). So ist zum Beispiel die Korrelation von Studienleistungen in Anglistik mit dem Gesamtnotendurchschnitt im Abitur höher als mit der Abschlussnote im Schulfach Englisch (Baron-Boldt, 1989; Köller & Baumert, 2002). Nach Hinzunahme von Einzelfachnoten zur Steigerung der inkrementellen Validität wird keine nennenswert höheren Zusammenhang erzielt als der mittels schulischen Abschlussnoten allein (Steyer, Yousfi & Würfel, 2005). Auch eine gewichtete Einbeziehung der Fachnoten des Abiturzeugnisses weist kaum eine Erhöhung der inkrementellen Validität auf. Vermutet wird eine von Seiten der Abiturdurchschnittsnote bereits optimal erfüllte Gewichtung der Fachnoten im Abiturzeugnis.

Der beste Einzelfachprädiktor ist die Mathematiknote mit einer Validität von  $r = .34$  (Baron-Boldt, Schuler & Funke, 1988) gefolgt von Einzelfachnoten in den Fächern Physik, Franzö-

sich und Deutsch. Nach Giesen et al. (1986) stellt ebenso die Mathematiknote den unter allen Schulfachnoten optimalen Vorhersagewert gefolgt von der Französisch- und Lateinnote dar. Geringere Vorhersagekraft liefern Noten in Biologie, Musik und Kunst. Darüber hinaus lassen sich Studienfeldeffekte finden, wonach generell in Natur-, Ingenieur- und Wirtschaftswissenschaften Studienleistungen besser vorhergesagt werden können als für Geisteswissenschaften, Jura oder Pädagogik. Interpretiert wird diese Befundlage durch die Tatsache, dass Fähigkeiten zu logischem Denken und sprachlichem Ausdrucksvermögen besser Studienerfolg vorhersagen und einzelne Unterrichtsfächer objektiver (schriftliche Prüfung) bewerten als dies zum Beispiel eine mündliche Prüfung vermag.

Abiturienten in Leistungskursen mit engem inhaltlichem Bezug zum Studienfach erzielen auch bessere Examensnoten. Fries (2002) schlussfolgert daraus die Relevanz studienfachspezifischer Fähigkeit. Zudem haben sich solche Schüler vertieft mit dem entsprechenden Unterrichtsfach auseinandergesetzt und somit eine fundierte Basis für ihre Studienwahlentscheidung. Gegen die Berücksichtigung von Leistungskursen bei Auswahlverfahren spricht eine zu erwartende Benachteiligung von studienfachgeeigneten Schülern, die aufgrund ihrer Nichtbelegung bestimmter Leistungskurse keine Zulassung zum Studienfach erhalten. Im Vergleich zur Abiturnote können Fachnoten bei Abwählbarkeit fehlen oder aus wenig oder stark umfangreichen Grund- bzw. Leistungskursen resultieren.

### 2.1.3 Auswahl über Studieneingangstestleistungen

Rindermann und Oubaid (1999) sprechen Studieneingangstests die Funktion einer aktuellen Objektivierung von Schulleistungen oder Ergänzung von Fähigkeiten zu, die durch die Abiturnote nicht oder nicht ausreichend abgedeckt werden. Dieselben Autoren empfehlen eine Einteilung in folgende fünf voneinander zu trennende Ansätze: Schulleistungstests, studienfachspezifische Kenntnistests, Studienfähigkeitstests, Aufsätze, Essays, Beschreibungen, Persönlichkeitstests und Persönlichkeitsfragebögen. Berücksichtigung finden ferner Intelligenz- und Interessentests.

Schulfachbezogene Kenntnis- (engl. Achievement-) oder Schulleistungstests erheben schulbezogene Kenntnisse (Rindermann & Oubaid, 1999) und wirken Nachteilen anderer Prädiktoren wie mangelnde Vergleichbarkeit von Noten entgegen. Da anstelle von fachrelevanten Studieneignungen auf den Schulstoff bezogene und somit kurzfristig trainierbare Kenntnisse erhoben werden, ist eine Bevorteilung von Schülern, die im Rahmen von Testschulungen das notwendige Wissen durch gezielte Vorbereitung bei z.B. kommerziellen Anbietern erwer-

ben, nicht auszuschließen (Becker, 1990; Schneider, 1987). Zahlreiche Colleges in den USA verwenden so genannte Subject Tests wie den Untertest „Scholastic Aptitude Test Subject Test“ für die Zulassungsentscheidung, für die Kurseinstufung, aber auch als Studienberatungsinstrument. Ein in Japan alljährlich durchgeführter standardisierter Schulleistungstest ist der so genannte „National Center for University Entrance Examinations (NCUEE)“-Test, nach dem die Mehrheit der öffentlichen Universitäten Japans ihre Studenten unter häufiger Hinzunahme hochschulspezifischer Leistungstests auswählt. Selbst das NCUEE-Komitee verweist auf private Vorbereitungsschulen, deren kostspielige Trainingsverfahren an den vermeintlich objektiven wie validen Eigenschaften dieses Prädiktors zweifeln lassen.

Studienfachspezifische Kenntnistests wie der Test of English as a Foreign Language (TOEFL) oder andere Aufgabensammlungen mit spezifischen Inhalten beziehen sich im Unterschied dazu nicht rückwärtsgerichtet auf konkrete Schulcurricula, überprüfen aber gleichsam den anzueignenden Wissensstand eines Studienbewerbers in studienrelevanten Bereichen. Der TOEFL ist der vom Educational Testing Service (ETS) in Princeton, New Jersey, entwickelte und international anerkannte englische Sprachtest als notwendige Voraussetzung für den Hochschulzugang ausländischer Studienbewerber an Amerikanischen Universitäten. Validitätsstudien des ETS ermittelten prädiktive Zusammenhänge um  $r = .30$  (Köller & Baumert, 2002). Selbsturteile von Studienanfängern ergeben große Kenntnis- und Kompetenzdefizite für Mathematik, politisches Grundwissen, Naturwissenschaften und wissenschaftliche Arbeitstechniken (Heine et al., 2005). Kenntnistests eignen sich für große Bewerbergruppen mit einer aus Ökonomiegründen notwendigen Multiple-Choice-Verfahrenslösung. Nichtsdestotrotz ist der im Vorhinein getätigte hohe Aufwand für Testentwicklung und -erprobung zu berücksichtigen. Zumal die prädiktive Validität von Kenntnistests im Vergleich zur Abiturnote lediglich als mäßig bis zufriedenstellend bezeichnet werden kann.

Studien- oder Studierfähigkeitstests (engl. Aptitude Tests) messen im Gegensatz zu Kenntnistests statt trainier- oder erlernbarer Wissensbestände geistige Fähig- und Fertigkeiten (Skills), welche „für erfolgreiches Studieren wichtig sind“ (Trost, 2003, S. 11). Ein engerer Zuschnitt der Studierfähigkeitstests auf die Zielgruppe der Studierfähigen einerseits und auf die kognitiven Anforderungen eines Hochschulstudiums andererseits lässt erwarten, dass den Ergebnissen dieser Tests eine höhere Prognosekraft bezüglich des Studienerfolgs zukommt als den allgemeinen Intelligenztests. Die Resultate einschlägiger US-amerikanischer und deutscher Untersuchungen bestätigen diese Erwartungen (Trost, 1975-2003). Es lassen sich

zwei Arten von Studierfähigkeitstests unterscheiden: 1. Allgemeine und 2. spezifische Studierfähigkeitstests. Ihre Unterscheidungsmerkmale werden in Tabelle 1 zusammengefasst.

Allgemeine Studierfähigkeitstests dienen dem Zweck, intellektuelle Fähigkeiten zu erfassen, die mehr oder minder für alle akademischen Ausbildungsgänge wichtig sind. Spezifische Studierfähigkeitstests prüfen Fähigkeiten, die für die Bewältigung bestimmter Studiengänge oder Studienfelder – das sind verschiedene Studiengänge, die ähnliche kognitive Anforderungen an die Studierenden stellen – besonders bedeutsam sind (Trost, 2003, S. 11).

Obschon diese beiden Testtypen überwiegend die gleichen kognitiven Fähigkeiten messen, kann mit einer gezielten Fokussierung auf ein Studienfach erreicht werden, dass die Ergebnisse von Studierfähigkeitstests den Studienerfolg besser vorhersagen als allgemeine Intelligenztestergebnisse (Trost, 1975; 2003). Fachliche Nähe wird auch dadurch geschaffen, indem die kognitiven Funktionen anhand von für den Lehrstoff typischen Inhalten geprüft werden.

Tabelle 1:  
*Arten und Charakteristika von Studierfähigkeitstests*

Studierfähigkeitstests	Allgemeine	Spezifische
Zweck	Erfassung intellektueller Fähigkeiten, die für alle akademischen Ausbildungsgänge wichtig sind.	Erfassung kognitiver Fähigkeiten, die für die Bewältigung bestimmter Studiengänge oder -felder wichtig sind.
Eigenschaften	Meist aus verbalem und quantitativem Teil bestehend, die in Untertests aufgegliedert sind. Aufgabentypen in abwechslungsreicher Reihenfolge sollen die geistige Beweglichkeit prüfen	Keine Bestehensgrenze, sondern Rangreihe - Cut-off-Point ergibt sich aus Anzahl verfügbarer Plätze.
Beispieltests	Scholastic Aptitude Test Reasoning Test (SAT Reasoning Test) Auswahltest der Studienstiftung (ATS) Test der akademischen Befähigung (TAB) Swedish Scholastic Aptitude Test (SweSAT)	Studieneignungstest für Sanitätsanwärter in der Bundeswehr (SETSAN) Test für medizinische Studiengänge (TMS) Medical School Test (MSAT) Graduate Management Admission Test (GMAT) Law School Test (LSAT)
Beispieltests Mischform	Scholastic Aptitude Test (SAT) ⇒ Im 2. Teil (SAT Subject Test) Kenntnistest Graduate Record Examination (GRE) ⇒ "Subject Tests" im 2. Teil spezifisch ACT Assessment" ⇒ Subject Tests" im 2. Teil spezifisch	

Beispielaufgaben	Begriffe präzise definieren und im jeweiligen Kontext richtig anwenden. Komplexe Informationen, die in Texten vorgegeben werden, richtig verarbeiten und zutreffende Schlüsse aus ihnen ableiten. Wort-Analogien	In sprachlichen Gebilden bestimmte klangliche oder rhythmische Strukturen erkennen (wichtig für philologische Studiengänge). Verstehen und richtiges Interpretieren von physiologischen Texten (wichtig für medizinische Studiengänge).
	Kein fachliches Vorwissen erforderlich (alle Informationen, die zur Lösung der Testaufgaben benötigt werden, sind im Aufgabenmaterial enthalten).	

Zu den bekanntesten allgemeinen Studieneignungstests zählt das College-Auswahlinstrument Scholastic Aptitude Test (SAT), welches in den USA als wichtigstes Feststellungsverfahren der Studierfähigkeit gilt. Entwicklung, Durchführung und Organisation unterliegen dem Educational Testing Service (ETS) in Princeton, New Jersey. Mit seinen beiden Teilen, dem allgemeinen „SAT Reasoning Test“ (früher SAT I Reasoning Test) und den schulfachbezogenen „SAT Subject Test“ (früher SAT II Subject Test), verbindet der SAT Studierfähigkeits- und Kenntnistest-Konzepte. Seit 2005 enthält der SAT außerdem eine Aufsatzkomponente (SAT Essay). Der SAT Reasoning Test überprüft allgemeine studienrelevante kognitive Fähigkeiten wie schlussfolgerndes Denken, Problemanalyse- und Problemlösefähigkeiten und besteht aus sprachlichen wie auch mathematischen Aufgabengruppen, für deren Bearbeitung drei Stunden bereitstehen. Whetton, McDonald & Newton (2001) finden für den SAT, der auf Reasoning-Fähigkeiten im sprachlichen und mathematischen Bereich abzielt, Korrelationen mit den Schulnoten um  $r = .50$ . Die Testwiederholungsreliabilität beträgt  $r = .89$  (Donlon, 1984). Das berühmteste deutsche Beispiel für ein studienfachspezifisches Auswahlverfahren stellt der vom Institut für Test- und Begabungsforschung entwickelte Test für medizinische Studiengänge (TMS) dar, der bis in die neunziger Jahre zur Diagnose der Studieneignung in den Studiengängen Medizin, Tiermedizin, Zahnmedizin eingesetzt wurde. Eine exklusive Fassung derselben Testautoren (Studieneignungstest für Sanitätsanwärter (SETSAN)) wendet die Bundeswehr bis heute im Rahmen der Offizierauswahl an. Seit 1998 verwendet die Schweiz das Verfahren als Eignungstest für das Medizinstudium (EMS). Das öffentliche Verfahren enthält neun Untertests (Muster zuordnen, Schlauchfiguren, quantitative und formale Probleme, konzentriertes und sorgfältiges Arbeiten, Figuren lernen, Fakten lernen, Textverständnis und Diagramme und Tabellen). Die TMS-Prädiktionskraft erreicht für medizinische Zwischenprüfungen eine Höhe von  $r = .44$ , die Vorhersagekraft des Abiturs  $r = .38$  und deren Kombination eine prädiktive Validität von

$r = .51$ ; der Einzelprädiktor mathematisches Grundverständnis verfügt über eine prädiktive Validität in Höhe von sogar  $r = .40$  (Troost, 1989). Trotz bewusst angestrebter geringer Korrelation mit der Abiturnote zeigen sich Zusammenhänge von  $r = .40$ . Eine möglichst geringe Beziehung zwischen TMS-Ergebnis und Abiturnote wird deswegen angestrebt, weil nicht die bereits in der Abiturnote vorhandenen, sondern zusätzliche Eignungsmerkmale zu einer Verbesserung der Prognosekraft beitragen können. Der Anstieg mittels Bestimmung der inkrementellen Validität durch die kombinierte Prädiktorbetrachtung von TMS und Abiturnote untermauert das Vorhandensein unterschiedlicher Facetten bzw. der oben beschriebenen Komplexität des Konstrukts Studieneignung. Bei Testwiederholungen zur Messung der Reliabilität berichtet Fay (1982) von Werten um  $r = .80$ .

Studieneignungstests schneiden hinsichtlich der klassischen diagnostischen Gütekriterien sehr gut ab. Bisher untersuchte studienfeldbezogene Leistungstests weisen eine etwa gleich hohe Prognosekraft wie die der Abiturnote auf. Bei zunehmender Differenzierung von Studiengängen wird studienfachspezifischen Fähigkeiten ein stärkeres Gewicht zuteil. Nachteilig wirkt sich die sehr umfangreiche Testentwicklung (Vorreiter in Deutschland: „Institut für Test und Begabungsforschung“ bzw. heutiges „ITB-Consulting“) aus.

Eine in der Praxis international zu findende Mischform nennt sich Reihungstest. Als Kombination eines allgemeinen und spezifischen Studierfähigkeitstests untersuchen Reihungstests die kognitiven sowie intellektuellen Voraussetzungen für erfolgreiches Studieren im Allgemeinen und in Bezug auf den spezifischen Studiengang bzw. Studienfeld. Rindermann empfiehlt auf dem 44. Kongress der Deutschen Gesellschaft für Psychologie in Göttingen (2004) eine „Art SAT<sup>2</sup> für alle Fächer insgesamt, aus dem sich dann die Fächer das, was sie brauchen, auswählen können“. In Kombination mit der Abiturnote können allgemeine und spezifische Studieneignungstest die inkrementelle Validität bedeutsam auf bis zu  $R = .50$  erhöhen (Kuncel, Hezlett & Ones, 2001; Mouw & Khanna, 1993; Robbins et al., 2004). Mit niedriger werdenden Korrelationen zwischen Studierfähigkeitstests und Schulnoten wächst der Zuegewinn durch Testverfahren (Oswald et al., 2004). Reihungstests spricht man gegenüber Schulnoten nicht nur eine höhere Objektivität, sondern auch größere Fairness hinsichtlich sozialer Selektivität zu.

Intelligenztests sind von Studierfähigkeitstests trotz erheblicher Überschneidung ihrer Messbereiche (z.B. Erfassung schlussfolgernden Denkens bzw. Reasoning) deutlich voneinander

---

<sup>2</sup> Scholastic Aptitude Test (SAT) des College Entrance Examination Board besteht aus einem allgemeinen Teil, dem "Reasoning Test", und einer Anzahl schulstoffbezogener "Subject Tests".

zu unterscheiden. So findet Trost (2003) zwar einen Zusammenhang zwischen einem allgemeinen Intelligenztest und der Gesamtleistung im TMS in Höhe von  $r = .70$ , weist aber gleichsam auf die folgenden drei nicht zu vernachlässigenden Gegensätze hin: Erstens fokussieren Studierfähigkeitstest im Vergleich zu Intelligenztests nicht die gesamte Spanne an Intelligenzausprägungen der Bevölkerung, sondern lediglich die der interessierenden Zielgruppe Studienbewerber. Zweitens untersuchen Intelligenztests betreffende Fähigkeiten mittels kleiner Informationseinheiten anstelle des Verarbeitens komplexer Informationen, welche in der Regel in Studieneignungstests zu finden sind. Drittens verfügen beide Testarten über nicht deckungsgleiche Messbereiche. Beide Methodentypen verfügen nicht nur aufgrund standardisierter Durchführungs- und Auswertungsmodalitäten angesichts gleicher Instruktionen und Testaufgaben über sehr gute Objektivitätswerte, sondern erfüllen gleichsam Forderungen nach Messgenauigkeit und Reliabilität. Die Prognosekraft allgemeiner Intelligenztests auf die Leistung im Studium ist jedoch umstritten und mäßiger als die Vorhersagegenauigkeit von Studieneignungstests (Trost & Bickel, 1979; Trost, 2003).

Im universitären Kontext konnte festgestellt werden, dass Ängstlichkeit dazu führt, dass die prognostische Validität von Intelligenz in Bezug auf den Studienerfolg sinkt (Kanekar, 1977). Einen bedeutenden Einfluss auf den Zusammenhang zwischen Intelligenz und Schulerfolg üben Persönlichkeitsfaktoren aus. Ihr Einwirken verändert nicht nur die Korrelationswerte, sondern ihre Kenntnis ermöglicht ihrerseits die Vorhersage des Schul- und Studienerfolgs (Amelang, 1976; Amelang & Hoppensack, 1977b). Persönlichkeitstests und Persönlichkeitsfragebögen ermöglichen standardisierte Aussagen über Vorhersagekraft nichtkognitiver Eigenschaften wie individuelles Erleben und Verhalten. Brandstätter et al. (2001) dokumentieren den Zusammenhang zwischen Höhe der Ausprägung der Persönlichkeitsmerkmale Selbstkontrolle, Gewissenhaftigkeit, Belastbarkeit und Introversion und Höhe des Studienerfolgs hinsichtlich der Uni-Noten und Abbruchquote. Auch Leistungsmotivation weist eine mittlere Validität auf (Schuler, 2003). Die Standardisierung sichert Persönlichkeitsfragebögen eine hohe Durchführungs- und Auswertungsobjektivität (Becker, 2003). Der Gefahr, bei Selbstauskünften nach persönlichem Verhalten sozial erwünschte Antwortmuster zu zeigen, kann sich der Testauswerter hingegen bei der Fremdauswahl nicht verschließen. Hülshager et al. (2004) zeigen dahingehend in einer experimentellen Untersuchung, dass es keine Ideallösung geeigneter Methoden zur Reduzierung sozialer Erwünschtheit in computergestützten eignungsdiagnostischen Testverfahren gibt. Eine generalisierende Aussage zur prognostischen Güte von Persönlichkeitsfragebögen ist daher nicht möglich. Persönlich-

keitsfragebögen gegenüber zu stellen sind die durch die Testperson kaum durchschaubaren objektiven Persönlichkeitstests, welche faktisches Verhalten in definierten Anforderungssituationen messen (Kubinger, 2005).

Essays über Themen oder in Form von Interessen-, Motivations- und Begründungsbeschreibungen vermitteln den Hochschulen einen persönlichen Eindruck des Bewerbers und sollen - unabhängig von der meist nur geringen diagnostischen Güte - Auskunft über Eignungsmerkmale liefern, die mit den oben beschriebenen nicht oder nur unzulänglich gemessen werden. Dazu zu zählen sind Motive für die Wahl des Studienfachs und der jeweiligen Hochschule, sprachliches Ausdrucks- und Argumentationsvermögen, außerschulische Interessen und fachliche Vorerfahrungen des Bewerbers. So nutzt in Deutschland die renommierte Wirtschaftsfakultät der Universität Witten-Herdecke dieses Instrument als Interviewgrundlage. In den USA verlangt der bereits angesprochene SAT Essay von den Teilnehmern, unter höchst kontrollierten Bedingungen schriftlich zu einem Thema Position zu beziehen. Die Beurteiler versprechen sich davon, Aspekte wie Komplexität und Solidität der Argumentation sowie den Umgang mit der Sprache messen zu können. Mittels unabhängiger Mehrfachbewertung eines Essays und webgestützter Zusammenarbeit der Beurteiler erfüllt das SAT-Modul höchste Objektivitätsanforderungen (Kobrin & Schmidt, 2005).

Als üblicher Interessentest ist der Allgemeine Interessen-Struktur-Test (AIST) von Eder & Bergmann (1999) oder das Verfahren „Explorix“ zu nennen; letzteres bedient sich gleicher theoretischer Grundlagen. Darüber hinaus verfügt die Bundeswehr über einen Studieninteressenfragebogen (STIFBO) (Kastner, 2004), der sich derzeit in der Praxiserprobung befindet und absehbar webgestützt durchgeführt werden soll. Die Validität von Aufsätzen liegt wie die Analyse von Studienmotivation usw. deutlich niedriger als die der Abiturnote oder von standardisierten Tests. Zudem ist die Urheberschaft nicht nachweisbar. Alle in diesem Abschnitt beschriebenen Verfahren sind aufgrund ihrer geringen Vorhersagegenauigkeit lediglich als zusätzliche Erhebungsmethode heranzuziehen.

#### 2.1.4 Auswahl über biographische Fragebögen oder Referenzschreiben

Angaben zur Biographie liefern insbesondere hinsichtlich Alter, Geschlecht und sozioökonomischem Status Zusammenhänge zum Studienerfolg. Laut HIS (2005) lässt sich eine Steigerung des Studienabbruchs unter den deutschen Studienanfängern sowohl unter Frauen als auch Männern feststellen. Der Studienabbrecheranteil bei männlichen Studierenden wächst von 25 Prozent 1999 auf 27 Prozent 2002, der der weiblichen Studierenden von 20 auf 23

Prozent. Die Ursachen dieser Differenzen werden insbesondere in den einzelnen Fächergruppen bzw. Fächern vermutet. Allerdings verzeichnet HIS eine Zunahme der Studienabbruchquoten für Frauen im Bereich der Fachhochschulen, bei Männern erhöhen sich diese besonders an den Universitäten. Zum Geschlecht konstatieren Widmaier et al. (1969) eine Beziehung zum Studienerfolg ähnliche Ergebnisse hinsichtlich der Abbruchquoten. Umgekehrt dargestellt liegen die Erfolgsquoten – ausgedrückt im Anteil der erfolgreichen Examenabsolventen an der Gesamtzahl der Studenten – für männliche Examenkandidaten zwischen vier und fünfzehn Prozent höher als für weibliche. Festerlings (1971) Befunde im Rahmen seiner Untersuchungen bestätigen diese Befunde. Sowohl Kath et al. (1966) als auch Hitpass (1967) finden positive Zusammenhänge zwischen Alter und Studienerfolg. Studenten, die erst sechs oder mehr Jahre nach Ablegen der Hochschulreife mit dem Studium beginnen, schnellt die Abbruchquote auf etwa 75 Prozent hoch. Die exogene Variable sozioökonomischer Status konnte ebenfalls bereits in den sechziger Jahren als auf den Studienerfolg einflussreich bestimmt werden. Eltern von amerikanischen Studenten, die ihr Studium abbrechen, haben im Durchschnitt ein geringeres Einkommen als die erfolgreicher Studierender. In Deutschland ermöglichten unterdes sowohl die Bildungsreform der siebziger Jahre als auch die Einführung der Bundesausbildungsförderung nach und nach breiten Schichten der Bevölkerung eine qualitativ anspruchsvolle Ausbildung. Es ist Aufgabe der heute nach Schwerpunkten eingeteilten gymnasialen Oberstufe, den Abiturienten entsprechende Eingangsstandards zukommen zu lassen. Erst dadurch erhält die soziale Öffnung der Universitäten ihren Sinn. Angesichts der bereits an den meisten öffentlichen Universitäten eingeführten Studiengebühren dürfte diese moderierende Einflussvariable wieder stärker in den Vordergrund der deutschen Bildungslandschaft führen.

Referenzschreiben sind insbesondere in den USA ein bedeutsames Entscheidungskriterium für Universitäten. Analog zu den Aufsätzen können Referenzschreiben als Interviewgrundlage dienen. Für sich alleine genommen kommt ihnen jedoch kein Entscheidungscharakter zu (Deidesheimer Kreis, 1997). Ein Schulgutachten verspricht Bewerberinformationen hinsichtlich Studienmotivation, Lernfähigkeit, Persönlichkeitseigenschaften und Interessen. Lehrerempfehlungen unterscheiden nur schlecht zwischen guten und sehr guten Schülern, können jedoch in der Regel schlechte Schüler erkennen (Schuler, 2003). Biographische Daten und Referenzschreiben erlangen jedoch aufgrund ihrer großen Variationsbreite und Uneinheitlichkeit in der Ausgestaltung niedrigste Objektivitäts- und Reliabilitätskennwerte; ihre Gültigkeit ist umstritten (Deidesheimer Kreis, 1997; Hödl, 2002).

### 2.1.5 Auswahl über Auswahlgespräche

In Form von Auswahlinterviews erhält der Bewerber die Gelegenheit, seine Motivation für das und Identifikation mit dem angestrebten Studium persönlich unter Beweis zu stellen. Gleichzeitig beurteilt die Hochschule auf diesem Wege seine kommunikativen Fähigkeiten und erhält einen Eindruck über den biographischen Hintergrund, seine Selbsteinschätzung, schulische Interessenschwerpunkte, außerschulische Aktivitäten, Art und Intensität der Vorbereitung, Studienmotivation und Vorstellungen über das Studium. Ein nicht außer Acht zu lassender positiver Nebeneffekt dieser sehr persönlichen Form der Studierendenauswahl ist die bessere Identifikation des Hochschullehrers mit seinen zukünftigen Studenten. Gleichzeitig fühlt er sich für die von ihm selbst ausgewählten Personen stärker verantwortlich. Mittels Gesprächen werden solche Eignungsmerkmale mit einbezogen, die über das intellektuelle Potential einer Person hinausgehen und welche im späteren Beruf von großer Bedeutung sein können. Insbesondere die für Campus-Universitäten wie die Bundeswehrhochschulen als wichtig anzunehmende soziale Kompetenz kann laut Schuler (1995) durch Auswahlgespräche verlässlicher erfasst werden als durch andere Verfahren. Um das Risiko sozial erwünschter Antworten zu senken, empfiehlt derselbe Autor das situative Fragen nach Erlebnissen, in denen sich z.B. die Hartnäckigkeit im Verfolgen persönlicher Ziele oder der Grad der Teamfähigkeit zeigen. Im Rahmen eines Interviews besteht die Möglichkeit, ein realistisches Bild des zum Beispiel sehr speziellen bundeswehrtypischen Studiums mit seiner Trimesterregelung, Campusbetrieb, Zusatzstudium EGTWA<sup>3</sup>, Offizierausbildungsanteile et cetera zu vermitteln. Der Person-Job-Fit-Ansatz (Amelang, 1997; Heise et al., 1997), der sich durch eine Optimierung der Passung von Personenmerkmalen (Fähigkeiten, Bedürfnisse, Erwartungen) und Umweltmerkmalen (Anforderungen, Angebote, Fachimages) eine verbesserte Leistung verspricht, ist auf nicht kognitive und kognitive Merkmale anwendbar, da auch ein „Zuviel“ an Erwartungen und Fähigkeiten für das gewählte Studienfach denkbar ist (vgl. Kapitel 3.6). Bei Überforderung und fehlender Person-Umwelt-Kongruenz sind Studienfachwechsel nicht auszuschließen. Neben dem Nutzen, über kognitive Leistungsinformationen hinaus Fähigkeiten beurteilen zu können, finden Interviews höchste Akzeptanz bei Bewerbern wie Hochschullehrern bzw. Hochschulen. So nennen Bewerber das Interview bei Befragung nach der „Wunschauswahlmethode“ an erster Stelle (Fruhner et al., 1991). Anwender schätzen andererseits die individuellen Freiheitsgrade des Interviews. Trotzdem lässt sich die kriteriale

---

<sup>3</sup> Erziehungs-Gesellschafts-Wirtschafts-Technikwissenschaftliche Anteile (EGTWA) zur Vermittlung fachübergreifender Kompetenzen, die in künftigen militärischen wie zivilen Tätigkeiten zunehmend verlangt werden.

Validität von Interviews gerade als Folge der meist nur unzureichenden Objektivität und Reliabilität als gering bis sehr gering einschätzen (Schuler, 2003; Trost, 1996; Trost & von Hayn, 2001; Amelang & Zielinski, 1997). Das Vorstellungsgespräch korreliert mit dem Ausbildungs- bzw. Berufserfolg im Mittel zwischen  $r = .10$  und  $.14$ , während kognitive Fähigkeitstests zu  $r = .54$  mit dem Ausbildungserfolg und zu  $.45$  mit dem Berufserfolg korrelieren. In Abhängigkeit seiner Anforderungsbezogenheit und seines Standardisierungs- und Strukturierungsgrads sowie Einheitlichkeit des Bewertungsmaßstabes kann die prognostische Gültigkeit dessen ungeachtet auch Werte im mittleren Bereich erringen. Die Metaanalyse von Hell et al. (2007) ermittelt für unstrukturierte Interviews eine korrigierte mittlere Validität von  $r = .11$  sowie  $r = .21$  für strukturierte Auswahlgespräche. Herausgestellt wird in dieser Untersuchung der nicht zu unterschätzende Studien- und Platzierungsberatungswert von Interviews. Mit Blick auf den hohen organisatorischen, zeitlichen und somit finanziellen Aufwand zu führender Interviews sollte nicht übersehen werden, dass der Entscheidungsprozess kein eindimensionaler Verlauf ist, sondern trotz bereits positiver Zulassungsbescheide anderer Hochschulen von vielen Bewerbern seinerseits als Beratungs- und Orientierungsangebot genutzt wird (Lewin, 2001). Darüber hinaus gewährleisteten Auswahlinterviews keine Fairness gegenüber den Interviewten (Deidesheimer Kreis, 1993). Der Einsatz von Gesprächen als diagnostisches Instrumentarium in der Studierendenauswahl ist somit nur dann zweckmäßig, wenn ihm beispielsweise strukturierte Interviewmodelle (Pursell, Campion & Gaylord, 1980; Schuler, 1992) zugrunde liegen und über kognitive Merkmale hinausgehende Eignungsaspekte erhoben werden sollen (Campion et. A., 1994; Schuler, 1998).

### 2.1.6 Auswahlkriterium Berufsausbildung

Studierende mit Berufsausbildung erzielen keine besseren Studienabschlüsse und weisen geringere Ausprägungswerte in Intelligenztests auf als direkt nach dem Abitur das Studium aufnehmende Personen. Andererseits ist in dem Zusammenhang zu berücksichtigen, dass vorher Berufstätige einen besseren Praxisbezug und eine fundiertere Informationsbasis für die Studienentscheidung mitbringen, was sich auf das fachliche Interesse, Engagement und eine geringere Abbruchsneigung im Studium positiv auswirken kann (Schuler, 2003).

### 2.1.7 Auswahl über Assessment Center

Das Assessment Center (AC) bedarf im Hinblick auf die im Fokus stehenden studierenden Offizieranwärter und Offiziere und wegen der in ein AC integrierten Studieneignungsemp-

fehlung im Rahmen der Offiziereignungsfeststellung (hier: „AC Bw“) besonderer Aufmerksamkeit. Das AC lässt sich definieren als Kombination mehrerer unterschiedlicher Einzelmethoden zu einem Verfahrenssystem, das von mehreren Beobachtern gleichzeitig für mehrere Bewerber in Bezug auf vorher festgelegte Anforderungen eingesetzt wird (Schuler & Funke, 1981; Jeserich, 1981). Das AC per se gibt es nicht. Je nach Ziel und Präferenzen der Betreiber variieren ACs beträchtlich und machen jeweils eine individuelle Erfolgskontrolle notwendig. „Ein Assessment Center ist ein spezielles führungseignungsdiagnostisches Hypothesensystem, das aus konkreten Konzept-, Verfahrens- und Prädiktor-(kombinations-)Hypothesen besteht“ (Barell, 1992, S. 68). In der Durchführungsphase verfolgen trainierte Beobachter das Verhalten jedes Teilnehmers in jeder Übung und urteilen über die Leistungsfähigkeit des Bewerbers bezüglich der einzelnen Anforderungsarten. Am Schluss steht die gemeinsame Einstufung der Kandidaten durch die Beobachter und eine Prognose über die Bewährungswahrscheinlichkeit. Die Güte der gegebenen Empfehlung hängt dabei von der Berücksichtigung verschiedener Grundprinzipien ab, kann demnach je nach AC variieren. Die in der Literatur und ausdrücklich vom Arbeitskreis Assessment Center (AKAC) proklamierte ausschließliche Zugehörigkeit des AC zu simulationsorientierten Verfahren konnte unter anderem durch metaanalytische Befunde widerlegt werden (Schmidt & Hunter et al., 1993; Hossiep, 1996; Schuler & Funke, 1989). So tragen zum Beispiel auch klassische Testverfahren als AC-Bausteine erheblich im Rahmen des AC zur Gesamtvalidität bei.

Die Messung von Retest- und Paralleltestreliabilität im AC erweist sich jedoch als sehr schwierig, da bereits ein unterschiedlicher Wortlaut im Interview höchst unterschiedliche Antworten auslösen kann. Als ebenso diffizil ist häufig auch die Messung der internen Konsistenz zu bewerten, weil sich bei ganzheitlicher Betrachtung des AC oft keine völlig gleichartigen Teile von Übungen und Beurteilungen abgrenzen lassen. Die Beurteilung der Zuverlässigkeit beschränkt sich maßgeblich auf die Beurteilerübereinstimmung, die auch als Maß der Objektivität gelten kann. Gemäß der deutschen Norm 33430 (DIN 33430, 2002) ergeben sich erfahrungsgemäß Werte zwischen  $r = .70$  bis  $r = .85$ . Scholz (1994), der für verschiedene Studien die Interrater-Reliabilitäten von Gesamtbewertungen auf Übungsebene dargestellt hat, gibt Koeffizienten von üblicherweise  $r = .42$  bis  $r = .95$  an. Moderierende Hintergrundvariablen wie Informationsaustausch vor Urteilsabgabe, Trainiertheit der Beobachter, Übungsart und Dimension erwiesen sich in zahlreichen Studien als einflussreich auf die Übereinstimmung der Beobachter (z.B. Scholz, 1994; Thornton & Byham, 1982; Neidig, Martin & Yates, 1978; Hinrichs & Haanperä, 1976; Obermann, 2002).

Die Validität des AC beschreibt den Grad der Genauigkeit, mit der das AC einen fehlerfreien Rückschluss auf Persönlichkeitseigenschaften des Bewerbers zulässt. Bei vorhandener Konstruktvalidität werden laut Schuler (1989) die Fähigkeiten erfasst, die man zu messen beabsichtigt. Während in älteren Studien den ACs nahezu keine Konstruktvalidität zugesprochen wurde, vermitteln jüngere Untersuchungen von Kleinmann (1993; 1997), Neidig und Neidig (1982) und Sagie und Magnezy (1997) ein gutes Bild über die die Konstruktvalidität moderierenden Faktoren. Werden eine kleine Anzahl konkreter, gut zu beschreibender und voneinander unabhängiger Anforderungsdimensionen eingesetzt und Beurteiler durch erhöhte diagnostische Qualifikation von der Verwendung impliziter Persönlichkeitstheorien abgehalten, lassen sich im AC sehr wohl Dimensionseffekte feststellen. Die relative Ähnlichkeit der Dimensionsbeurteilungen lässt darüber hinaus auf Situationsspezifika und implizite Phänomene wie den Halo-Effekt schließen. In diesem Zusammenhang werden in der Literatur sowohl die Kandidaten als auch die Beurteiler selber, die Übungen sowie die Dimensionen als weitere Einflussfaktoren genannt. Außerdem sind im Rahmen von Studien zum nomologischen Netz Persönlichkeitskonstrukte bekannt geworden, die mit dem AC-Ergebnis kovariieren und den erwarteten Einfluss von Dimensionen über die Übungen hinweg belegen (z.B. Scholz & Schuler, 1993; Klimoski & Brickner, 1987; Crawley, Pinder & Herriot, 1990).

Laut Fecker (1989) besitzt die prognostische Validität bei der Evaluation eines AC die höchste Bedeutung. Sie bezieht sich auf das Ausmaß der Vorhersagegenauigkeit unabhängiger Kriterien. Lattmann (1989) postuliert: „Die in der prognostischen Validität in Erscheinung tretende Verlässlichkeit des Verfahrens ist zwar auf Grund der neuesten Befunde geringer als es den ersten Ergebnisbewertungen entsprach, ist aber immer noch zu den höchsten zu zählen, welche durch Eignungsbeurteilungsverfahren erreicht werden“ (S. 51). Diese Aussage wird auch durch Obermann gestützt (2002), nach dem ACs heute einen relativ gesicherten und günstigeren Platz unter den Prädiktorklassen einnehmen. Sie können subjektive Leistungskriterien und harte Beförderungskriterien im Schnitt und ohne viel Streuung mit einer Validität von knapp  $r = .40$  erfassen. Bahnbrechende Ergebnisse der viel zitierten „Management Progress Study“ bei AT&T (Thornton & Byham, 1982) oder der bisher umfangreichsten Metaanalyse durch Thornton et al. (1992) mit mittleren Korrelationskoeffizienten in Höhe von  $r = .37$  stützen bis heute den postulierten Nutzen von ACs als prognostisch valide Auswahlinstrumente. Als Ausweg aus Reaktivitätseffekten schlägt Bungard (1987) für die Praxis den Einsatz so genannter „Lügenfragen“, wie man sie aus Persönlichkeitstests kennt, sowie nichtreaktive Beobachtungstechniken vor.

Das reziproke Verhältnis zwischen Konstrukt- und Kriteriumsvalidität erklären Klimoski und Strickland (1977) mit dem Wirkmechanismus der „indirekten Kriterienkontamination“, bei der die Beobachter ihre Beurteilung auf die Passung der Bewerber zur „corporate identity“ stützen. Das Problem der eignungsdiagnostischen Forschung besteht darin, befriedigende Maße zu finden, die als Kriterien der Validierung von Auswahlverfahren dienen (Problematik der Indikator-Kriteriumsbeziehungen). Eine exakte Klärung des „ultimativen“ Kriteriums „Erfolg des AC“ ist nicht fehlerlos möglich. Daher orientiert man sich in der Praxis gleichzeitig an weichen Kriterien wie zum Beispiel Vorgesetztenbeurteilung, Peer-Ratings und harten wie Gehalt, Hierarchieebene usw.

Die Forschungsuntersuchungen zur prognostischen Validität lassen das folgende, von Thornton und Byham (1982) gefundene, Fazit gelten. Danach sagt das AC die Führungsleistung genau vorher und differenziert exakt zwischen guten und schlechten Bewerbern.

Our review of the criterion-related validity has shown that assessment centers can accurately predict management performance and progress. A positive validity coefficient indicates that, on average, persons with higher assessment ratings will perform more adequately on the job than those who are assessed poorly (S. 317).

### 2.1.8 Trends in der Studierendenauswahl

Der profilbildende Wettbewerb unter den Hochschulen ist als „Prozess dynamischer Wechselwirkungen“ (Trost & Haase, 2005) zwischen Eigendarstellung der Universität und interessiertem Aspiranten zu verstehen. Letzterer möchte sich hinsichtlich seines Qualifikationsprofils bestmöglich preisgeben. Dabei spielt zukünftig das Internet als Bewerbungsportal eine herausragende Rolle, in dem sowohl umfangreiche in Studierendenmarketing eingebundene Hintergrundinformationen über Universität, Studium und Studiengänge als auch die Möglichkeit von Self-Assessment mit E-Learning-Elementen offeriert wird (Schüpbach & Klauer, 2006). Mit dessen Hilfe lassen sich Auskünfte zur Passung des persönlichen Fähigkeitsprofils mit den Studienanforderungen eruieren.

Nur ein geringer Teil der Studienanfänger gibt an, vor oder während der Studienfach- oder Studienortentscheidung gut informiert worden zu sein (Heine et al., 2005). Im Hinblick auf zunehmende Ausdifferenzierung der Studienfächer und Auswirkungen mangelnder Passgenauigkeit nimmt die Qualität der an das Auswahlverfahren geschalteten Information und Beratung vor und während des Studiums einen größer werdenden Stellenwert ein. Potentielle

Verfahrensentwickler werden bereits heute angehalten, Self-Assessment-Tools für die Beratung von Studierenden bereitzustellen (DGPs, 2005). Der bei den prognostisch viel versprechenden aber meist sehr aufwendig zu konstruierenden Verfahren kritisierte Kosten-Nutzenaspekt wird durch - im besten Falle webbasierte - Selbstselektion bzw. Self-Assessment gemildert. „Im Self-Assessment lässt sich die Angemessenheit der Erwartungen bezüglich des Studiums mittels gezielter Überprüfung der Informiertheit sowie der Erhebung der Gründe für die Studienfachwahl klären“ (Schüpbach, Pixner & Zapf, 2006). Eine einträgliche Selbstselektion vermindert den anschließenden Selektionsaufwand durch die Hochschulen dadurch, dass sich im Idealfall nur geeignete Kandidaten tatsächlich bewerben.

Self-Assessment-Verfahren zeichnen sich durch ihre inhaltliche Vielfalt zur Messung sowohl kognitiver als auch nicht kognitiver Elemente aus. In Form eines Realistic Job Preview seitens der Hochschule, des Fachbereichs bzw. Studiengangs erleichtert den Bewerbern die eigene Stärken-Schwächen- sowie Neigungs- und Leistungseinschätzung im Vergleich zum vorgegebenen Anforderungsprofil die Entscheidung für das persönlich geeignete Fach (vgl. Kapitel 3.6). Die Kommission der Studierendenauswahl in der Deutschen Gesellschaft für Psychologie (DGPs) (2004) regt explizit an, die Studierendenauswahl mittel- und langfristige neben der Entwicklung eines standardisierten Testverfahrens zur Erfassung allgemeiner und spezifischer Studierfähigkeit durch Verfahren für die Beratung von Studierenden in Form eines Self-Assessment zu ergänzen. Dieses soll Studierenden eine fundiertere persönliche Entscheidung für ein Studienfach ermöglichen (DGPs, 2004). Erfahrungen an der Rheinisch-Westfälisch-Technischen Hochschule in Aachen mit dem seit 2002 eingesetzten webbasierten Self-Assessment für die Fächer Informatik, Elektrotechnik und Technische Informatik oder das Beispiel an der Universität Freiburg belegt die nicht zu unterschätzende Steuerfunktion von Selbstselektion vor der eigentlichen Studierendenauswahl (Zimmerhofer, Heukamp & Hornke, 2006). Vor dem Hintergrund „virtuelle Studienberatung und Self-Assessment statt Eignungstest“ geht die Hochschule für Angewandte Wissenschaften Hamburg neue Wege in der Studierendenberatung und -auswahl.

Die Nutzung und rasante Entwicklung des Internet im Sinne von E-Assessment (Personalauswahl und -platzierung unter Nutzung des Internet) oder sogar in Kombination mit E-Recruiting (webbasierte Methoden des Anlockens von potenziellen Bewerbern) (Konradt & Sarges, 2003) verspricht erheblichen Nutzensgewinn auch für die zukünftige Studierendenauswahl. Da sowohl Unternehmensfokus als auch Zweckrichtung und Prozessintegration in den Personalbeschaffungsprozess variabel sind, erweist sich diese kombinierte Form von

Personal- und Studierendenauswahl mit Personalmarketingaspekten gerade für die Offizierauswahl der Bundeswehr als viel versprechend.

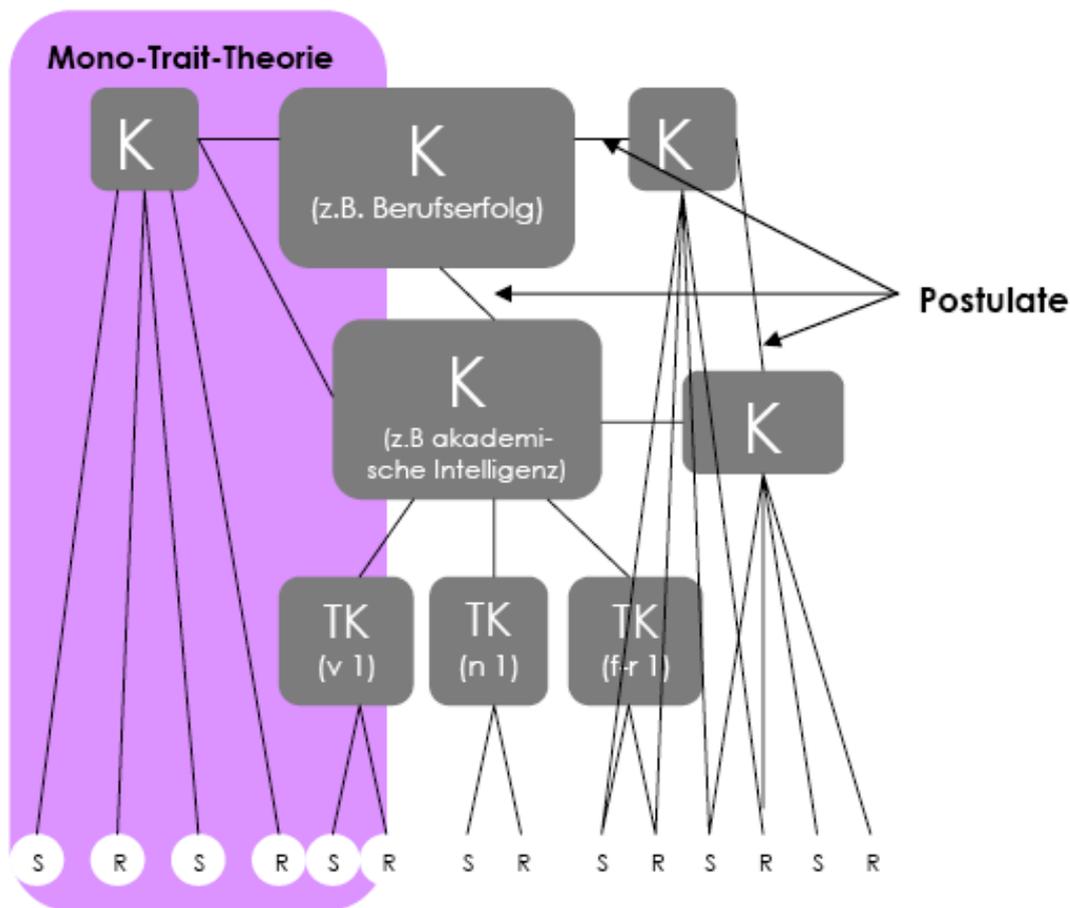
In Online-Assessment integriertes Marketing stellt ein gewichtiges Stellglied der angestrebten konkurrenzorientierten Profilbildung der Fächer und einen entscheidenden Schachzug im Wettbewerb unter den Hochschulen dar. Hornke & Schilling (2004) machen jedoch darauf aufmerksam, dass den meisten Erstsemestern Angaben zu den Forschungsbemühungen der Universitäten unwesentlich erscheinen. Die Perspektive potentieller Bewerber wird vielfach vernachlässigt. Wichtig hingehend sind den Bewerbern transparente und klar verständliche Hintergründe über und die Akzeptanz von Studieninhalten, Zulassungsbedingungen und Auswahlverfahren (Hornke & Zimmerhofer, 2006).

Die Bewerberperspektive wird in den nächsten Jahren wissenschaftlich weiter beleuchtet werden müssen. Zukünftig vollzieht sich an universitätsinternen Beratungsstellen ein Themenwechsel hin zur Hilfestellung bei Prüfungsangst, diffuser Gefühlsarmut, Kontaktproblemen bis hin zu ernsthaften Erkrankungen wie Zwängen, Ess- und Persönlichkeitsstörungen. Aus individueller Perspektive ist ein Studienabbruch dann kritisch, wenn Selbstzweifel und Unsicherheit bei der Neuorientierung überwiegen. Informations- und Kenntnisdefizite, unzureichende Voraussetzungen und sachfremde Hochschulwahlmotive sowie Zweifel an der eigenen Eignung erhöhen das Risiko falscher Erwartungen und Orientierungen, die wiederum einen Studienabbruch begünstigende Leistungs- und Studienmotivationsprobleme nach sich ziehen (Heublein et al., 2003). Dahingehend fanden Schübach, Pixner & Zapf (2006) heraus, dass das Assessment von studien Erfolgskritischen Handlungskompetenzen zu messbaren Verringerungen verlaufsbezogener Studienprobleme führen kann. Studienabbruch, Studienfachwechsel, Unzufriedenheit mit dem Studium und Probleme bei dessen Bewältigung sollten sich – wie in der vorliegenden Arbeit herausgearbeitet – langfristig aufgrund optimierter Studienberatung und Passung des einzelnen auf das richtige Studienfach verringern.

Einen Überblick über aktuelle Trends der Studierendenauswahl und -beratung „for the 21st century“ in den USA geben z.B. Camara & Kimmel (2005). Für den deutschen bzw. europäischen Markt steht – nicht zuletzt durch die Einführung von Bachelor- und Masterstudiengängen – die Internationalisierung der Verfahren in besonderem Fokus.

### 2.1.9 Exkurs akademische Intelligenz

Nach dem gegenwärtigen Forschungsstand unterscheidet man unterschiedliche Domains der Intelligenz: akademische, praktische, soziale und emotionale Intelligenz als Lernfähigkeit und praktische Intelligenz. Hinsichtlich der Thematik der vorliegenden Arbeit und im Hinblick auf zu konstruierende Reihungstestaufgaben soll hier lediglich der Bedeutungsbereich der akademischen Intelligenz Berücksichtigung finden. Intelligenz ist als Konstrukt ein Bestandteil umfassenderer räumlicher Netze mit wieder anderen Konstrukten als Bestandteil. Über die Postulate, d.h. über die Verknüpfungen zwischen Konstrukten (z.B. Intelligenz und Studienerfolg) erhält das Konstrukt Intelligenz weitere Bedeutungskomponenten. Aus diesen Erkenntnissen ergibt sich, dass „die Menge aller Spezifikationen, die für das komplexe Konstrukt Intelligenz zu einem bestimmten Zeitpunkt akzeptiert werden, seinen augenblicklich ausgewiesenen Bedeutungsgehalt bestimmt“ (Brocke, 1990, S. 226). Wie in Abbildung 1 veranschaulicht, lässt sich über die Verknüpfungen zwischen den Konstrukten (Postulate; z.B. zwischen akademischer Intelligenz und Studienerfolg) hinaus innerhalb des nomologischen Netzes der hierarchische Aufbau der Strukturtheorien der Intelligenz abbilden. Dabei werden innerhalb der hierarchischen Intelligenztheorien (u.a. Cattell, 1971; Vernon, 1965) Konstrukte einer höheren Generalitätsebene (z.B. akademische Intelligenz) durch Teilkonstrukte auf der spezifischeren Ebene (z.B. verbale, numerische und figural-räumliche Intelligenz) strukturiert.



*Anmerkung zu den Abkürzungen:* K = Konstrukt; TK = Teilkonstrukt; v 1 = verbale Intelligenz; n 1 = numerische Intelligenz; f-r 1 = figural-räumliche Intelligenz; S = Situationsbedingung; R = Verhaltensvariable.

Abbildung 1: Konstrukte in einem räumlichen Netz (Brocke & Beauducel, 2001)

Die in der Intelligenzforschung sehr bedeutsamen Strukturtheorien liefern den theoretischen Hintergrund des Konstrukts akademische Intelligenz. Der ursprünglichen Zielsetzung psychometrischer Intelligenztests folgend, weisen die akademische Intelligenz und ihre Teilkonstrukte vor allem zu Fähigkeiten einen hohen Zusammenhang auf, wie sie in den gesellschaftlichen Bildungseinrichtungen (Schule, Hochschule) gefordert werden. Hinweise auf die Bedeutung akademischer Intelligenz liefern die Merkmale, die den Aufgaben zur Erfassung dieses Konstrukts gemeinsam sind (Amelang & Bartussek, 1997; Neisser, 1976). Aufgaben zur Erfassung der akademischen Intelligenz zeichnen sich dadurch aus, dass sie von anderen Personen (nicht vom Probanden) formuliert werden, meist nur von geringerem oder gar keinem intrinsischen Interesse sind, alle benötigten Informationen von Beginn an zur Verfügung stellen, von den allgemeinen Erfahrungen mehr oder weniger abgehoben und gut

strukturiert sind, größtenteils nur eine richtige Antwort haben und vorwiegend nur über einen angemessenen Lösungsweg verfügen.

Hinter der Strukturfrage verbirgt sich diejenige nach der Relation der Teilkonstrukte von Intelligenz. Diesbezüglich herrscht Konsens darüber, dass Intelligenz vornehmlich in hierarchischen Modellen erschlossen werden kann (Cattell, 1963, 1971, 1987; Guilford, 1967; Guttman, 1965; Vernon, 1965), die mindestens zwei Hierarchie- bzw. Generalitätsebenen unterscheiden. Dabei bildet das Prinzip der Multi-Trait-Determination von Intelligenzleistungen die Voraussetzung für deren hierarchische Strukturierung. Jede zu erbringende Intelligenzleistung wird gleichzeitig von mehreren Einzelfähigkeiten („Primärfaktoren“, Thurstone, 1938) bestimmt. So können beispielsweise verbale Analogieaufgaben durch die Einzelfähigkeiten „verbale Intelligenz“ und „schlussfolgerndes Denken“ beeinflusst werden. Werden einzelne Intelligenzskalen (z.B. verbale Analogien) als alleiniger Indikator für das Teilkonstrukt „verbale Intelligenz“ herangezogen, dann bleibt oftmals ein Teil der systematischen Varianz unaufgeklärt, da der Einfluss einer weiteren Fähigkeitskomponente nicht berücksichtigt wird (z.B. „schlussfolgerndes Denken“). Tritt diese Intelligenzfähigkeit bei der Manifestation weiterer Teilkonstrukte (z.B. „numerische Intelligenz“) auf, ist sie generellerer Natur und kann auf einer höheren Hierarchieebene angesiedelt werden.

Demgegenüber zu stellen ist, dass einige Faktoren der akademischen Intelligenz den Zusammenhang zwischen Testergebnissen und Schulerfolgskriterien moderieren: So weisen verbale Aufgabengruppen auf Grund der üblichen sprachlichen Ausgestaltung des Unterrichts eine höhere Korrelation auf, als beispielsweise numerische oder figurale (Amelang & Bartussek, 1990). Die Validitäten für die Vorhersage des Studienerfolgs liegen annahm gemäß unterhalb der Koeffizienten für den Schulerfolg (vgl. Giesen, Gold, Hummer & Jansen, 1986; Trost, 1975). Für die Studienfächer Wirtschaftswissenschaften, Jura, Geisteswissenschaften und Pädagogik werden Zusammenhänge zur allgemeinen Intelligenz zwischen  $r = .13$  und  $r = .27$  berichtet. Zwischen Intelligenzwerten und den Fächern Medizin, Mathematik und Naturwissenschaften ließen sich keine systematischen Zusammenhänge ermitteln.

#### 2.1.10 Exkurs Persönlichkeitsmerkmale

Bergmann, Brandstätter und Eder (1994) finden positive Zusammenhänge der Studienleistungen mit Gewissenhaftigkeit, Belastbarkeit und Introversion. Weitere empirische Befunde deuten auf einen positiven Einfluss von Selbstwirksamkeit, Handlungskontrolle und Lernmotivation auf Studienerfolg hin (Schiefele & Urhahne, 2000).

Bei der Messung von Motivation wird Hauptaugenmerk auf die Lernmotivation gelegt, die als „Wunsch bzw. Absicht“ definiert ist, „bestimmte Inhalte oder Fertigkeiten zu lernen“ (Schiefele, 1996, S. 50). Die intrinsische Motivation gilt als wichtige Bedingung für qualitativ anspruchsvolle Formen des Lernens (Ryan & LaGuardia, 1999). Zum Einfluss von Motivation weisen Schiefele et al. (2003) nach, dass Anstrengungsmanagement neben Wettbewerbsmotivation und schulischer Leistungsfähigkeit als wichtigste proximale Einflussgröße fungiert. Die Anstrengung selbst wird primär von der Leistungsmotivation und dem Studieninteresse gesteuert.

Intrinsisch motivierte Studenten erzielen bessere Studienleistungen als extrinsisch motivierte (Skager et al., 1967; Wilcke, 1976). Dabei ist intrinsische Motivation mit dem Begriff des Interesses eng verbunden. Aus einem ausgeprägten Interesse für ein Studienfach oder -feld resultieren ein erhöhter Lernaufwand und tiefergehendes Lernen (Schiefele, Wild & Winterler, 1995). Däumling (1968) berichtet von signifikanten Unterschieden aus einer Untersuchung an der Universität Bonn. Nicht erfolgreiche Studenten haben demnach weniger Interesse an ihrem Studienfach, größere Examensangst und ein gespannteres Verhältnis zu ihrer Familie; ferner bereiten diese sich schlechter auf die Prüfungen vor, sind unkonzentrierter und häufiger an einen Geschlechtspartner gebunden als erfolgreiche.

Leistungsmotivation ist als möglicher Prädiktor für Studienerfolg seit den fünfziger Jahren Gegenstand von Untersuchungen. Hierbei erwiesen sich projektive Verfahren wie der Rorschach-Test und der Thematic Apperception Test als untauglich zur Vorhersage der Leistung im Studium. Durch deutlich höhere Reliabilität waren Fragebogenerhebungen den projektiven Verfahren überlegen. Entwistle & Entwistle (1970) berichten über signifikante Korrelationen zwischen Leistungsmotivations-Scores und Kriterien des Studienerfolgs. Während Personen mit hoher Leistungsmotivation dazu neigen, sich auf die Aufgabe hin zu orientieren und anfängliche Fehlschläge als Steigerung der Attraktivität der Aufgabe zu sehen, geben wenig leistungsorientierte schneller auf. Gewissenhaftigkeit, Zuverlässigkeit und Leistungsmotivation werden von Hitpass (1975) als hoch bedeutsam für Studienerfolg angenommen. Angstneigung, Belastbarkeit, Selbstsicherheit und Empfindsamkeit werden in zahlreichen Studien als für den Studienverlauf relevante Einflussgrößen benannt (Melter, 1977). Der Begriff der intrinsischen Motivation ist eng mit dem des Interesses verbunden. Letzteres definieren Schiefele und Schreyer (1994) als wichtigste Grundlage intrinsischer Lernmotivation. Kühn (1983) wies nach, dass Interesse an einem Schul- oder Studienfach positiv mit formalen Leistungsindikatoren korreliert. Aber Studieninteresse kann nicht nur zur Erläute-

nung von Phänomenen wie Studienfachwahl, Studienwechsel, Studienzufriedenheit und Studienabbruch wertvolle Informationen liefern (Bergmann, 1992; Winteler & Sisterwald, 1987; Winteler et al., 1988). Schiefele et al. (1993) vermuten anhand von Untersuchungen einen positiven Einfluss von Interesse auf die Entstehung intrinsischer Motivation im Studium bei gleichzeitig geringem Einfluss auf die Entstehung bzw. Verhinderung extrinsischer Motivation. Interessenbekundungen in Form von Aufsätzen oder Tests bzw. Fragebögen erweisen sich in Forschungsstudien als valider Prädiktor für Schulerfolg. Schiefele, Krapp und Schreyer (1993) fanden im Rahmen einer Metaanalyse Korrelationen im Mittel von  $r = .30$  und einen korrelativen Anstieg mit zunehmenden Alter. Anhand von Untersuchungen an studierenden Offizieren fanden die Verfasser einen Zusammenhang zwischen Interesse und Zwischenprüfungsnoten in Höhe von  $r = .33$ . In einer Studie zeigten Brandstätter, Farthofer & Grillich (2001), dass für Bewerber die Rückmeldung über ihr Intelligenzniveau wenig bedeutsam ist. Subjektiv wichtiger werten sie Rückmeldungen über ihre Interessen und die spezifische Eignung. Dieselben Autoren entdeckten bei Studieninteressierten, dass deren Interessenkongruenz mit Studienempfehlungen nach einem Test erheblichen Einfluss auf die Studienwahlentscheidung und deren Stabilität hat. Bei Interessenpassung wurde das beabsichtigte Studium in circa 70 Prozent der Fälle tatsächlich begonnen und zu 97 Prozent beibehalten. Bei mangelnder Interessenpassung wurde es in circa 50 Prozent der Fälle begonnen und zu 85 Prozent beibehalten.

Gerpott und Domsch (1988) berücksichtigen in umfangreichen Studien bei Studierenden der Bundeswehrhochschulen als studienerefolgsrelevante Persönlichkeitsvariable „generalisierte Kontrollüberzeugungen (locus of control). Giesen et al. (1981) belegen den signifikanten Einfluss nicht-kognitiver Faktoren auf die Entscheidung, ein Studium einzustellen. Studienabbrecher lassen sich ergo anhand spezifischer Personencharakteristika differenzieren. Sie werden prospektiv von Kommilitonen als weniger fleißig und weniger leistungsmotiviert und leistungsfähig beurteilt (Gold & Kloft, 1991). Ferner finden sich in Studien Hinweise auf das Vorhandensein emotionaler Labilität und geringen Selbstbewusstseins bei Studienabbrechern (Saenger-Ceha, 1970).

Rindermann & Oubaid (1999) unterstreichen ebenfalls die Wichtigkeit von Soft Skills wie Kommunikationsvermögen, Leistungsmotivation, emotionale Stabilität und Offenheit für Studium und Beruf. In neueren Untersuchungen dominieren Schlüsselqualifikationen wie Studienbewältigung und Commitment (vgl. Schüpbach et al., 2006) oder das vorgestellte Modell zum „University Citizenship Behavior“ (Kersting & Gehring, 2006). Der Großteil

der bestehenden Studien zu Organizational Citizenship Behaviors (OCB) basiert auf der Beziehung zwischen OCB und der Arbeitsleistung auf individueller Ebene. Gehring (2006, Diplomarbeit bei Kersting) macht darauf aufmerksam, dass ungeachtet der Tatsache, dass Arbeitgeber von Absolventen überwiegend Soft Skills verlangen, die Diagnostik zur Beratung und Auswahl von Studierenden einseitig an den akademischen Kriterien Studiendauer und -note geschieht. Diesen Kriterien als Indikatoren der aufgabenbezogenen Leistung (Abitur- und Vordiplomnote) steht die umfeldbezogene Leistung (so genannte University Citizenship Behavior (UCB)) entgegen. Die mittels Big Five Inventar (BFI, Rammstedt & John, 2005) sowie einem neuartigen berufsbezogenen Persönlichkeitstest (Start-P, Kersting & Beauducel) gemessenen vier Dimensionen Hilfsbereitschaft, Gewissenhaftigkeit, Unkompliziertheit und Eigeninitiative finden als Prädiktoren von UCB Bestätigung.

Entwistle & Entwistle (1970) belegen ferner, dass emotionale stabile Studenten im Studium erfolgreicher sind als Studenten mit einem hohen Neurotizismuswert (z.B. im Big Five Inventar).

Aus einem anderen Blickwinkel betrachtet spielen Schlüsselqualifikationen im Sinne persönlicher und sozialer Kompetenzen dahingehend eine Rolle, dass diese häufig erst im Studienverlauf ihre Entfaltung finden (Bohleber, 1982).

### 2.1.11 Studiergewohnheiten und Einstellungen zum Studium

Biäsch & Vontobel (1966) fanden in einer Untersuchung an 55 Studenten der Eidgenössischen Technischen Hochschule Zürich, dass erfolgreiche Studenten mehr, systematischer und konzentrierter arbeiten als nicht erfolgreiche. Sie bevorzugen überdies Einzel- vor Gruppenarbeit und trennen schärfer zwischen Arbeit und Freizeit. Das Studieren an sich empfinden erfolgreiche Studenten häufiger als positiv, zentrale Aspekte des Studiums gefallen ihnen eher als der Vergleichsgruppe. Darüber hinaus beurteilen sie das Ausbildungsniveau ihrer Hochschule kritischer. Jahnke (1971) stellte bei erfolgreichen Studenten höhere Wochenarbeitszeiten fest als bei schwächeren Studenten. Sie verwenden beispielsweise mehr Zeit in der Bibliothek, lassen mehr Planung im Studium erkennen und beschreiben differenziertere Lerntechniken. „Studenten mit besseren Leistungen arbeiten systematischer, eher allein, mehr auf das Examen gerichtet, sind selbstkritischer und selbständiger (...)“ (Jahnke, 1971, S. 7). Interessanterweise stellten Brown & Holtzmann bereits 1955 fest, dass Studiergewohnheiten und Einstellungen zum Studium nur niedrig mit herkömmlichen Prädiktoren

korrelieren und aus diesem Grunde einen autonomen Beitrag zur Erhöhung der multiplen Vorhersage leisten können.

## 2.2 Zusammenfassung Prädiktoren des Studienerfolgs

Die bisherigen Ausführungen zu Prädiktoren des Studienerfolgs lassen sich zusammenfassen. Generell kann die Abiturdurchschnittsnote nach gegenwärtigem Forschungsstand noch vor Einzelfachnoten als einer der besten und ökonomischsten Prädiktoren und des in Prüfungen festgestellten Studienerfolgs gelten. Grundsätzlich und trotz ihrer offenkundigen Augenscheininvalidität leisten Einzelfachnoten keinen höheren Beitrag zur Vorhersage des Studienerfolgs als die Gesamtnote im schulischen Abschlusszeugnis. Ihre Heranziehung als Kriterium zur Auswahl guter Studienanfänger dagegen wird durch empirische Ergebnisse untermauert. Aufgrund der genannten Nachteile wie geringe Objektivität, unzureichende Beurteilerübereinstimmungen, niedrige Wiederholungsreliabilität, keine Aussagekraft für ein bestimmtes Fachstudium et cetera sollte die Abiturnote nicht als universeller Prädiktor des Studienerfolgs fungieren. Zusätzlich sind Studieneingangstestleistungen als Alternative zur Erhebung der Fähigkeiten, die durch die Abiturnote nicht oder nicht ausreichend abgedeckt werden, einzusetzen. Nicht zu vergessen ist ferner der Umstand, dass die allgemeine Hochschulreife bestenfalls Aussagen über die allgemeine Studieneignung treffen kann; studienfachspezifische Erfolgsprognose ist durch sie jedoch nicht leistbar. Studienfeldspezifische Eignungstests sind in Kombination mit allgemeinen Studierfähigkeitstests sowohl Kenntnis- als auch der alleinigen Nutzung von Intelligenztests hinsichtlich der klassischen Gütekriterien weit überlegen. Intelligenztestaufgaben finden im Rahmen allgemeiner Studierfähigkeitstests über die Messung allgemeiner kognitiver Fähigkeiten (z.B. schlussfolgerndes Denken) ihre Anwendung. Spezifische Studierfähigkeitstests stellen indessen Anforderungen, die als typisch für ein bestimmtes Studienfach oder Studienfeld gelten. Durch die Kombination von Reihungstests mit der Abiturdurchschnittsnote wird die prognostische Validität ergo weiter erhöht. Zudem kann bei Berücksichtigung von Tests der Kampf um Zehntelnoten und Konsequenzen aufgrund einer Oberstufen-Fächerwahl anhand individueller Notenerwartung vermieden werden. Eine Erhöhung der inkrementellen Validität ist ferner durch Persönlichkeitstests zu erzielen, welche studienerefolgsrelevante Arbeitshaltungen und Merkmale wie Gewissenhaftigkeit, Belastbarkeit, emotionale Stabilität, Offenheit und Leistungsmotivation messen. Wegen ihrer Anfälligkeit für gezielte Beeinflussung und Verfälschung ist die prognostische Gültigkeit von Persönlichkeitstests fragwürdig. Sie eignen sich jedoch ausgezeich-

net als Grundlage für Auswahlgespräche, zumal die Wichtigkeit von Interessen und Schlüsselqualifikationen wie Commitment und Bewältigungsverhalten und Arbeitshaltungen als Prädiktoren positiven University Citizenship Behaviors ausschlaggebend sind. Als in nur sehr geringem Maße prognostisch valide, wenig objektiv und wenig reliabel gelten Aufsätze, Motivations- und Begründungsbeschreibungen, bei denen es primär um motivationale Gesichtspunkte und einen Eindruck vom Argumentations- und Darstellungsvermögen der Studienbewerber geht. Die statistische Gültigkeit von biographischen Fragebögen oder Referenzschreiben ist gar umstritten. Interessensindikatoren, z.B. erfasst durch Interessenfragebögen, eignen sich zur Prognose von Studienzufriedenheit und Studienabbruch. Ein Auswahlgespräch liefert anhand motivationaler Kriterien und nicht kognitiver Persönlichkeitsmerkmale zusätzlich noch ein realistisches Bild über universitäre Gegebenheiten. Darüber hinaus informiert es über individuelle Erwartungen der Bewerber, bewirkt den Aufbau einer Beziehung zwischen Universität und ihren Mitgliedern, erhöht die Identität der Studenten mit der Institution Bundeswehr und fördert die Selbstselektion der Bewerber mit der Folge einer geringeren Quote von Studienfachwechseln. Nicht als alleiniges Vorhersageinstrument, aber kombiniert mit Resultaten anderer Verfahren lässt sich ein Zugewinn an Vorhersagegenauigkeit – insbesondere für Zweifelsfälle und im Sinne einer Platzierungs- und Studienberatungsfunktion – verzeichnen.

Für das AC kann festgehalten werden, dass dieses im Allgemeinen eine gute bis befriedigende Reliabilität besitzt, über eine relativ hohe prognostische Gültigkeit verfügt, Konstruktvalidität jedoch bestenfalls eingeschränkt vorliegt. Aufgrund des hohen Aufwands und der hohen Kosten findet es trotz der hohen Vorhersagegenauigkeit im Kontext der Studierendenauswahl kaum Anwendung.

Selbstselektion vor Beginn eines Bewerbungsprozederes für ein Studium ermöglicht, eine höhere Quote potenziell Erfolgreicher unter den Bewerbern zu erreichen.

Abschließend lässt sich eine Kombination der aussagekräftigsten Prädiktoren befürworten. Nach Rindermann & Oubaid (1999) müsste ein zuverlässiges, faires, ökonomisches und objektives Auswahlverfahren folgende Komponenten enthalten:

1. Abiturdurchschnittsnote
2. Allgemeine & studienfeldspezifische Fähigkeitstests
3. Auswahlgespräche

Im Hinblick auf die besondere Alltagssituation privater Campusuniversitäten wären zusätzlich Persönlichkeitsfragebogenergebnisse bzw. die Integration der drei empfohlenen Komponenten in ein umfangreiches Assessment Center zu empfehlen.

### 2.2.1 Typologie der Auswahlverfahren

Die behandelten Prädiktoren orientieren sich alle an Eignungs- und Leistungskriterien. Sie liefern im Wesentlichen Informationen über kognitive allgemeine und/oder spezifische, Persönlichkeitsvoraussetzungen, Motivationen und Interessen. Sonstige formale Kriterien wie Wartezeitquote oder studienrelevante personale Aspekte (Alter, Nachweis eines Ausbildungsbetriebs usw.) finden an dieser Stelle keine Berücksichtigung. In Tabelle 2 werden beispielhaft Beurteilungskriterien für die verschiedenen Aspekte von Studierfähigkeit aufgelistet.

Tabelle 2:  
*Schematischer Überblick über Prädiktoren zur Vorhersage des Konstrukts Studieneignung*

Aspekte von Studierfähigkeit	Beispiele von Beurteilungskriterien / Prädiktoren
Allgemeine Kognitive Fähigkeiten	Abiturnote Einzelfachnoten Schulfachbezogene Kenntnistests Allgemeine Studierfähigkeitstests Intelligenztests Auswahlgespräche (zur Prüfung allgemeiner kognitiver Voraussetzungen wie z.B. verbale Intelligenz)
Spezifische kognitive Fähigkeiten	Einzelfachnoten Fremdsprachenprüfungen Studienfachbezogene Kenntnistests Studienfachspezifische Studierfähigkeitstests Assessment Center Praktische Eignungsprüfungen Auswahlgespräche (zur Prüfung des Grundverständnisses für einen spezifischen Studiengang) Fachspezifische Klausuren Berufliche Erfahrungen bzw. Praktika Auslandserfahrungen Fachliche Vorbildung, Zusatzausbildung Prüfung im Studium
Interesse und Motivation	Auswahlgespräche Essays, Motivations- und Begründungsschreiben Referenzschreiben Nachweis über außerfachliches Engagement Leistungskurse Außerschulisches Engagement, Auszeichnungen Bewerbungsschreiben, Lebenslauf, Bewerbungsmappe Interessenfragebögen

Persönlichkeitsmerkmale und allgemeine Qualifikationen	Auswahlgespräche Essays und Begründungsschreiben Angaben aus Biographie, Lebenslauf Einschlägige Praxiserfahrungen, Berufstätigkeit Persönlichkeitsfragebögen Persönlichkeitstests (kriterienbezogen) Assessment Center
--	---

## 2.3 Kriterien

Allgemein ist zu bedenken, dass jeder Test so viele Validitätskoeffizienten hat wie Kriterien zur Prüfung seiner diagnostischen Leistungsfähigkeit herangezogen werden. Die Höhe und die Unterschiede dieser Koeffizienten hängen nicht nur von der Stärke der Selektionseffekte, sondern auch von der Güte und Eigenart der herangezogenen Kriterien ab. Die messtheoretische Güte von Kriterien hat einen großen Einfluss auf den Wert des Koeffizienten der prädikativen Validität. Das komplexe Kriterium Studienerfolg umfasst neben objektiven oder objektivierbaren Facetten wie Prüfungsleistungen auch sehr schwer fassbare Faktoren wie anfänglicher Berufserfolg (Abele-Brehm & Stief, 2004), Studienzufriedenheit (Rolfs & Schuler, 2000; Stebler, 2000) oder das Ausmaß erfahrener Persönlichkeitsentwicklung. Letztere wird hier keine tiefer gehende Betrachtung erfahren.

### 2.3.1 Exkurs Kriterienproblematik

Die prognostische Validität bezieht sich auf das Ausmaß an Vorhersagegenauigkeit in Bezug auf unabhängige Kriterien. Die Zeiträume, auf die sich die Vorhersagen erstrecken, können sehr unterschiedlich sein. Sie variieren von wenigen Tagen bis zu Jahren (Lienert, 1969). Eine besondere Schwierigkeit stellt dabei auch die exakte Klärung des Kriteriums „Erfolg eines Auswahlinstrumentes bzw. einer Personalentscheidung“ dar. Dieses eigentliche und ultimative Kriterium, das nicht direkt zu messen ist, umschließt alle Aspekte der Arbeitsleistung. Ein wichtiges Erfordernis an ein messbares Kriterium ist seine Relevanz, die in

Abbildung 2 durch die Schnittstelle (2) von aktuellem (1) und eigentlichem (3) Kriterium veranschaulicht wird (Blum & Naylor, 1968). Sie stellt sozusagen die Güte des aktuell und tatsächlich gemessenen Kriteriums am ultimativen Kriterium sicher.

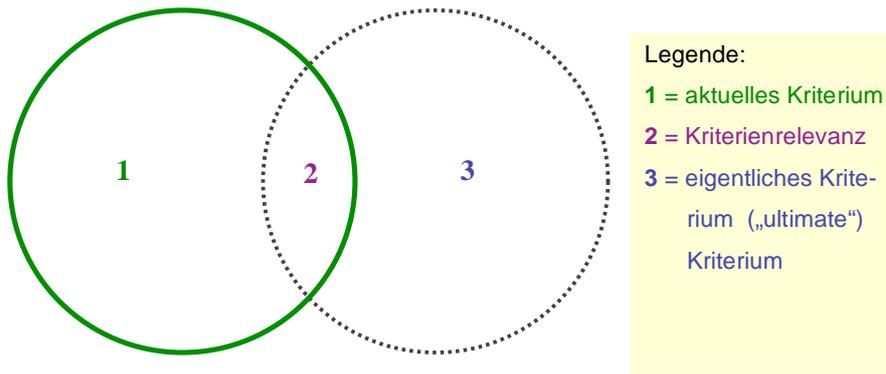


Abbildung 2: Die drei Eigenschaften des Kriteriums

Diese Feststellung erscheint zunächst als Tautologie. Bei näherer Betrachtung aber erweist sich ihre Bedeutung. In den seltensten Fällen wird es nämlich möglich sein, das zu messen, was Thorndike (1947) „ultimate criterion“ genannt hat. Dieser Gesamterfolg einer Organisation oder Person ist eigentlich erst frühestens nach dem Ende der Existenz des Unternehmens feststellbar (Wottawa & Thierau, 1998). Man muss daher aus Gründen der Praktikabilität ein realistischerweise erhebbares Kriterium auswählen. Messverfahren wie das AC sind messfehlerbehaftet. Sie erfassen sowohl für dieses aktuelle Kriterium (siehe Abbildung 3) als auch für die diagnostischen Prädiktoren nur zum Teil die theoretisch ausgewählte Dimension. Aus diesem Grund kann man durch derartige Forschungsarbeiten niemals absolut sichere und den „letzten Erfolg“ einer Entscheidung vollständig prognostizierende Aussagen erwarten.

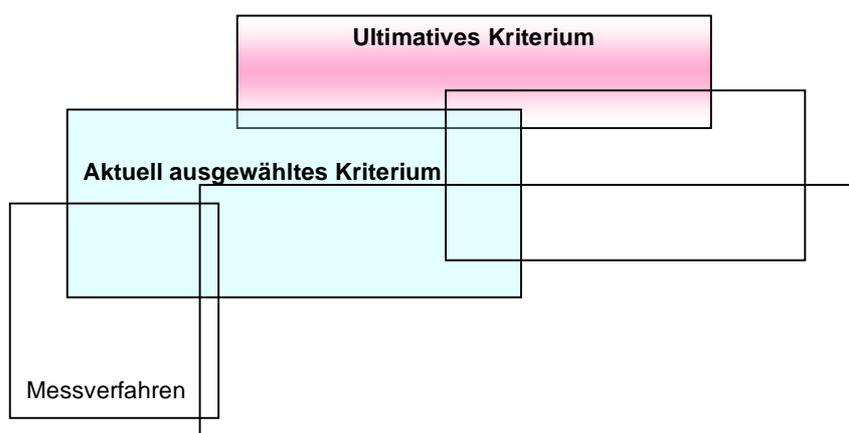


Abbildung 3: Probleme der Indikator-Kriteriumsbeziehungen (nach Wottawa & Hossiep, 1987)

Diese Problematik der Indikator-Kriteriumsbeziehungen ist in Abbildung 3 dargestellt. Die nur partielle und durch die Überschneidungsbereiche symbolisierte Übereinstimmung zwi-

schen dem an sich relevanten ultimativen Kriterium und dem aktuell ausgewählten Kriterium sowie die Unterschiede zwischen den angestrebten Konstrukten (aktuelles Kriterium, diagnostischer Prädiktor) und den tatsächlichen Messwerten (nicht vollständige Reliabilität, zusätzliche systematische Einflüsse durch die Verfahrensauswahl) erschwert die Fundierung der Personalauswahl auf der Basis von Validierungsuntersuchungen.

Auch wenn eine ideale Lösung nicht hundertprozentig durchführbar ist, gibt es eine Vielzahl praktisch relevanter Erfolgskriterien, von denen möglichst viele gleichzeitig untersucht werden sollten. Beispiele sind Vor- und Hauptdiplomnoten, Vorgesetztenbeurteilungen, Kollegenurteile, Lehrgangsergebnisse, Gehalt, Karrierestufen. Neben der Einteilung in „harte/objektive“ vs. „weiche/subjektive“ lassen sich Kriterien im Hinblick auf die Funktionalität, Begründbarkeit und Akzeptabilität in ergebnisorientierte, verhaltensorientierte und eigenschaftsbezogene Kriterien klassifizieren. Verwiesen sei in diesem Zusammenhang auch auf Campbell, Dunette, Lawler & Weick (1970). Thorndike (1947) unterscheidet zwischen unmittelbaren (zum Beispiel Ausbildungserfolg), Zwischen- (zum Beispiel Aufstieg) und endgültigen Kriterien (längerfristiger Berufserfolg). Die Vorhersagegüte nimmt von den unmittelbaren zu den endgültigen Kriterien aus mehreren Gründen ab. Zum einen beeinflussen zu viele unkontrollierte und nicht vorhersehbare Variablen aus dem persönlichen und beruflichen Bereich den Prädiktor-Kriteriums-Zusammenhang; zum anderen werden allgemein multiple Kriterien als aussagekräftiger angesehen als einfache, da jeder Beruf eine Reihe von Tätigkeiten beinhaltet. Bei den Kriterien besteht eine gegenläufige Beziehung zwischen Messbarkeit und Bedeutsamkeit. Besser operationalisierbar und messbar ist sicherlich das Kriterium „Studienerfolg“, bedeutsamer aber der Berufserfolg.

### 2.3.2 Kriterium Studienzwischen- und -abschlussnoten

Travers und Gronlund (1950) sowie Carter (1953) wiesen bereits vor mehr als einem halben Jahrhundert darauf hin, dass Noten stets auch Indices der sozialen Interaktion zwischen Lehrern und Studenten sind. Da nicht nur inter- sondern auch intra-universitäre Bewertungsunterschiede zu finden sind, wird die psychometrische Qualität von Prüfungsnoten insgesamt als gering und sogar niedriger als die der Schulnoten eingeschätzt (Amelang, 1978, 1980; Betz, 1974; Preiser 1975; Rothe 1981; Trost 1975, 1985). Im Falle von mündlichen Zwischen- bzw. Studienabschlussprüfungen sind weitere Fehlerquellen und ihr negativer Einfluss auf Objektivität, Reliabilität und Vergleichbarkeit zu berücksichtigen. Trotz der aufgeführten Artefakte gelten Studienabschluss- oder Zwischenprüfungsnoten aufgrund ihrer In-

haltsvalidität und leichten Erfassbarkeit in der Praxis und insbesondere wegen der fachspezifischen Vorhersagekraft für künftige Arbeitsmarktchancen als wichtigstes und am häufigsten eingesetztes quantitatives Kriterium (Baron-Boldt, Schuler und Funke, 1988). Werden in deutschen Untersuchungen überwiegend Vordiplom- (Hell, Trapmann, Weigand, Hirn & Schuler, 2005) oder Diplomnoten (z.B. Steyer, Yousfi & Würfe, 2005) herangezogen, verwenden amerikanische Studien zumeist Studienleistungen nach dem ersten Studienjahr als Kriterium. Nach Gold & Souvignier (1997) erfüllt dieses Kriterium weniger eine die Leistung rückmeldende als vielmehr eine diagnostisch-filternde und kausale Funktion.

### 2.3.3 Kriterium Studienabschluss, Studienfachwechsel und Studienabbruch

Der Studienabschluss stellt ein essentielles Erfolgskriterium dar, kann doch bei Studienabbruch kaum von einem erfolgreichen Studium gesprochen werden (Gold & Kloft, 1991). Im Streben der Hochschulen, entsprechend ihres Selbstverständnisses und gesellschaftlichen Auftrags die Mehrheit ihrer Studierenden zu einem Examen auf hohem Niveau zu führen, stellt die Quote des Studienabbruchs einen wesentlichen Indikator für das Leistungspotential deutscher Hochschulbereiche oder gar des gesamten Hochschulsystems dar (HIS, 2005). Ein Hochschulabgang ohne Erreichen eines angestrebten Studienabschlusses gilt dagegen als ein besonders hartes Kriterium des Studienmisserfolgs. Erst ein Vergleich der Absolventen mit Studienabbrechern und/oder Universitäts- und/oder Studienfachwechslern lässt in Kombination mit Examens- oder Zwischenprüfungsnoten eine umfassende Bewertung der Prognosekraft von Merkmalen der Studierfähigkeit zu. Das Kriterium Abbrecherquote erlaubt überdies die Aufdeckung von Aspekten finanzieller oder humaner Fehlleistungen. Leider ist es in der Praxis aus organisatorischen oder Datenschutzgründen in der Regel nicht möglich, Zugang zu Studienabbrechern zu erlangen. Besonders schwierig wird es dann, wenn wie im Falle der Bundeswehr die Studienabbrecher nicht nur aus dem Studium, sondern auch aus der Organisation ausscheiden.

### 2.3.4 Kriterium Studiendauer

Die Studiendauer unterliegt zu vielen schwer zu kontrollierenden Einflüssen wie z.B. Art der Studienfinanzierung, familiäre Belastung (Schmidt-Atzert, 2005), nicht optimale Administration und Organisation des Studiums an der Hochschule (Kersting, 2005) usw. als dass sie als ausschließliches Kriterium in Frage käme (Höppel & Moser, 1993). Darüber hinaus wird die Erfassung dieser Variablen durch Defizite der Hochschulstatistik eingeschränkt. Wage-

mann (1987) beschreibt in dem Zusammenhang den verzerrenden Einfluss von Quereinsteigern, Fach- und Ortswechslern oder – wegen finanzieller Vorteile – Immatrikulierten auf die Studienzeitsstatistiken. In Studien fanden Giesen und Gold (1996) demgemäß die niedrigste Korrelation zwischen Studiendauer und Schulleistungs- oder Intelligenzmaßen in Höhe von lediglich  $r = .09$  bzw.  $r = .04$ .

### 2.3.5 Kriterium Studienzufriedenheit

Spies et al. (1998) konzipieren Studienzufriedenheit in Anlehnung an die Arbeitszufriedenheit als Einstellung zum eigenen Studium. Dazu zählen eine Kombination von Merkmalen wie Studienbedingungen, Lehrverhalten der Lehrenden, Benotung, berufliche Relevanz, Anerkennung und unterstützenden Kontakte zu Kommilitonen. Studienzufriedenheit gilt als Korrelat individueller Anpassungs- und Regulationskompetenzen und wird als Passung zwischen objektiven Studienqualitäten und individuellen Bedürfnissen verstanden (Gold & Souvignier, 1997; Heise et al., 1997; Spies et al., 1996; Westermann et al., 1996). Prognostizieren lässt sich Studienzufriedenheit nach Giesen et al. (1986) durch Fleiß, fachliches Selbstbewusstsein, emotionale Stabilität und positive Umweltbeurteilungen. „Studienunzufriedenheit ist als enttäuschte Erfolgserwartung eine wechselseitige Wirkung von mehr oder weniger stabilen Dispositionen der Person (Fähigkeits- und Motivationsmängel, zu wenig wirksame Selbstkontrolle und unzulängliche Lernstrategien) und ungünstigen Umweltbedingungen (Mängel in der Studienorganisation, fehlender sozialer Rückhalt, Notwendigkeit von Erwerbstätigkeit, vom Studium ablenkende Stellenangebote)“ (Brandstätter & Farthofer, 2003). Allerdings konnte in einer Studie von Brandstätter et al. (2001) Studienunzufriedenheit nur als Begleiterscheinung und nicht als unmittelbare Ursache von Studienabbruch nachgewiesen werden.

### 2.3.6 Kriterium Beruf

Allgemeine berufsqualifizierende Kompetenzen wie z.B. Durchsetzungsvermögen, soziale Kompetenz, Führungsfähigkeit und Teamfähigkeit werden aufgrund von Definitions- und Messproblemen in der Regel nicht als Kriterien des Studienerfolgs herangezogen (Rindermann & Oubaid, 1999). Berufserfolg kann im Vergleich dazu müheloser durch Gehaltshöhe, Arbeitszufriedenheit, Arbeitsmarktchancen, erreichte Position usw. definiert und z.B. durch das Vorgesetztenurteil (Hunter, 1986) operationalisiert werden. Obschon Berufserfolg ein

zentrales Kriterium des Studienerfolgs darstellt, lassen sich kaum Studien zum Gegenstand finden.

## 2.4 Zusammenfassung Kriterien des Studienerfolgs

Ungeachtet ihrer geringen und inter- sowie intrauniversitär variierenden psychometrischen Qualität werden Studienabschluss- und Zwischenprüfungsnoten aufgrund ihrer Inhaltsvalidität und leichten Erfassbarkeit in der Praxis und wegen ihrer fachspezifischen Vorhersagekraft in Bezug auf zukünftige Arbeitsmarktchancen am häufigsten als Kriterium von Studienerfolg eingesetzt. Auch ergeben sich beim Blick auf weitere mögliche Kriterien keine Alternativen. Im Rahmen der Diskussion um die Kriterien- oder ultimate-criterion-Problematik kann außerdem aus Gründen der Praktikabilität nur ein realistischerweise erhebbares Kriterium - wie beispielsweise die Examensnote - befürwortet werden.

## 3. Interne Struktur des Konstrukts Studieneignung

Vor einer internen Strukturbestimmung des Konstrukts Studieneignung (und Festlegung des dieser Arbeit zugrunde liegenden theoretischen Modells) wurden mögliche Prädiktoren und Kriterien unter Einbezug der Kriteriumsproblematik von Studienerfolg detailliert diskutiert. Zur genaueren Betrachtung der Komplexität des Konstrukts werden in den nachstehenden Kapiteln sowohl bedeutsame theoretische als auch praxisnahe Prognosemodelle der Studieneignung bzw. Studierendenauswahl bis hin zu in diesem Zusammenhang stehenden Studienabbruchmodellen beschrieben, welche die Schlussfolgerungen oben aufgeführter Abrisse berücksichtigen.

### 3.1 Theoretische Modelle der Studieneignung

Das Gros vieler den Studienerfolgsprädiktoren zugrunde liegenden Modelle setzt entsprechend der dokumentierten empirischen Befunde ihren Schwerpunkt auf kognitive Voraussetzungen. Ferner ist ihnen die Unterscheidung von allgemeinen - für alle Studienfächer geltenden - fach- oder feldspezifischen Merkmalen gemein. Einflussreiche deutsche Exempel werden in den folgenden Abschnitten aufgeführt.

In den siebziger Jahren legte der Kreis wissenschaftlicher Berater (1977) eine Matrix der Lernvoraussetzungen vor, über deren Dimensionen ein Studienanfänger für erfolgreiches

Studieren verfügen sollte. Bereits damals unterschied man zwischen studienfach- und studienfeldspezifischen sowie allgemeinen Lernvoraussetzungen, die als so genannter Fachbezug der Dimension Verantwortlichkeit mit ihrer Unterscheidung in Kenntnisse, Fertig- und Fähigkeiten gegenüberstanden. Weder motivationale Komponenten noch studienrelevante Einstellungen oder Arbeitshaltungen finden im Modell Berücksichtigung.

Im Vergleich zu diesem rein kognitive Aspekte berücksichtigenden Modell beziehen Florin & Rosenstiel (1976) bereits weiche Faktoren mit in das Modell für die Determinanten der Studienleistung ein. Jedoch sollen diese lediglich im Rahmen der Studierendenberatung Berücksichtigung finden (Michel, 1977). Studienleistung wird definiert als Produkt situativer Ermöglichungsbedingungen der Leistung (S), Fähigkeiten als überdauernde Leistungskapazität und erlernbaren Fertigkeiten (FF) sowie den Motivationen zur Leistung und ihren emotionalen Korrelaten (M) (z.B. Arbeitszufriedenheit). Daraus resultiert die Formel  $L = f(S \cdot FF \cdot M)$ .

Die Studienstiftung des deutschen Volkes unterteilt in ihrem Modell studienfeldbezogener Tests das Gesamtkonstrukt Studieneignung in sechs Teilkonstrukte: Informationsaufnahme, -verarbeitung, -speicherung, -weitergabe, motivationale Aspekte und Einstellungen sowie Verhaltensweisen im Arbeitsbereich (Rahn et al. 1976). Das Modell zeichnet sich dadurch aus, dass stets ein Augenmerk auf die Komplexität des Kontextes gelegt wird, die einzelnen Teilkonstrukte also nicht auf isolierte Grundfähigkeiten bezogen werden.

Trost (2003) betont im Modell der Eignung für verschiedene Studienfelder, welches in Abbildung 4 skizziert ist, deren zu berücksichtigende Eigenheiten und unterschiedliche (benachbarte Studienfelder) wie durchgängige (allgemeine kognitive Fähigkeiten studienfeldübergreifend) Überlappungsbereiche. Die optimale Vorhersage der Studieneignung ist mittels einer kombinierten Bewertung aus allgemeinen Fähigkeiten – also für alle akademischen Ausbildungsgänge relevanten – mit studienfach- oder studienfeldspezifischen intellektuellen Fähigkeiten leistbar.

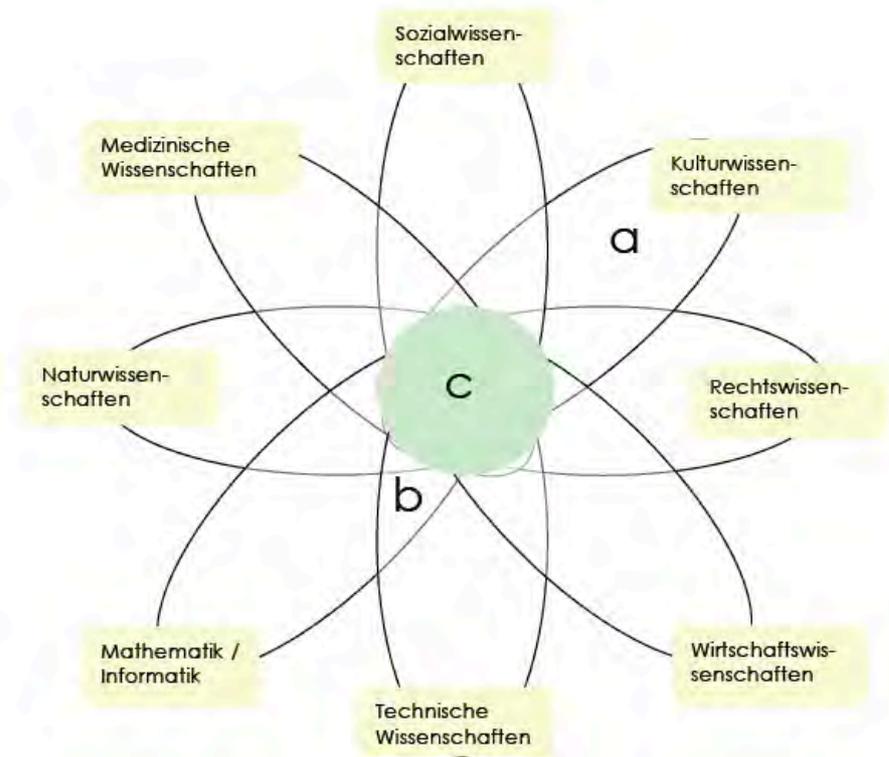


Abbildung 4: Modell der Eignung für verschieden Studienfelder (Trost, 2003, S. 13)

### 3.2 Zuordnungsmodelle

Im Vergleich zu Prädiktionsmodellen zielen die praxisbezogenen Zuordnungs- bzw. Zulassungsmodelle auf eine möglichst zu objektivierende und ökonomisierende Vorhersage ab. Rindermann & Oubaid (1999) benennen bei Verwendung der empfohlenen Kombination eines Auswahlverfahrens aus Abiturnote, Studierfähigkeitstest und Auswahlgespräch folgende drei Modellvarianten: 1. Das Quoten-, 2. Das einstufig-multiple und 3. das sequentielle Modell.

Als besonders ökonomisch gilt das Quotenmodell, nach dem die Studienplätze nur in speziellen Fällen über ein Interview, teils über Abiturnote, ein weiterer Teil über die Kombination von Schulabschlussnote und Studierfähigkeitstestergebnis vergeben werden.

Eher kompensatorischer Natur aber äußerst aufwendig ist hingegen ein einstufig multipel durchgeführtes Auswahlverfahren, in dem alle Bewerber anhand gleicher Kriterien beurteilt werden. Es wird davon ausgegangen, dass ein Defizit durch eine entsprechend höhere Ausprägung anderer Merkmale ausgeglichen werden kann.

Die Vorzüge des Quotenmodells (Ökonomie) und einstufig multiplen Modells (Urteilsgerechtigkeit) vereinen das sequentielle Modell mit vorgegebenen Quoten oder das auf die drei Prädiktoren festgelegte Abitur-Test-Interview-Modell nach Rindermann & Oubaid (1999).

Entsprechend der Verfügbarkeit und Höhe der Prognosekraft für das jeweilige Studienfeld werden Quoten und eine Reihenfolge zu durchlaufender Verfahren festgelegt. So kann beispielsweise im ersten Schritt die Zulassung eines Teils der Studienbewerber aufgrund ihrer herausragenden Abiturnote und die weitere Zulassung eines Teils nach Kombination von Abiturdurchschnittsnote und Fähigkeitstest stehen. In einem letzten Schritt werden die verbliebenen Personen mit unklarer Prognose mittels eines Auswahlgesprächs die übrigen Studienanfänger rekrutiert. Auf diese Weise kann eine Fakultät, die den Schwerpunkt auf motivationale und persönliche Merkmale legt, mit Hilfe beweglicher Quoten beispielsweise die Hälfte aller Studienplätze mit Hilfe des Interviews vergeben, während bei technischen Studiengänge vielleicht 80 Prozent der Studienplätze über eine Auswahl mittels ingenieurwissenschaftlicher Fähigkeitstests vergeben werden.

Tabelle 3:

*Schematischer Aufbau der sequentiellen Prozedur des Abitur-Test-Interview-Modells (ATIM) als Beispiel für ein sequentielles Zulassungsmodell mit vorgegebenen Quoten nach Rindermann & Oubaid (1999)*

Kriterien	Verfahren	Studienplätze
Leistung, Fähigkeiten (indirekt Arbeitshaltung etc.)	Vorauswahl Abiturdurchschnittsnote (eventuell besondere Berücksichtigung studiennaher Fachnoten)	Vergabe von 40 % der Plätze an Abiturbeste
Spezifische und allgemeine Fähigkeiten und Eignungen, eventuell Kenntnisse	2. Stufe Kognitive Fähigkeitstests (studienfachbezogene Tests, allgemeine kognitive Fähigkeiten)	Vergabe von 40 % der Plätze an Abitur- und Testbeste (Kombination), Berücksichtigung von Fachnoten
Studienmotivation, Interessen, soziale Kompetenz, Engagement	3. Stufe Auswahlgespräch (strukturiertes Interview, Leitfaden)	Vergabe von 20 % der Plätze, Verhältnis 2:1 oder 3:1 Interviewte:Plätze

### 3.3 Prädiktionsmodelle

In den vorherigen Kapiteln wurde die Existenz einer Vielzahl von Untersuchungen über die Bedingungen von Studieneignung beschrieben. Eine substantielle Aufklärung des Zustandekommens von Studienerfolg ist jedoch nicht erfolgt. Dabei scheint sich die Vielzahl der Maße von Studieneignung auf wenige Dimensionen zu reduzieren, womit über die tatsächliche Dimensionalität von Studienleistung allerdings noch nichts ausgesagt ist. Damit erweist sich die Prognose von Studienerfolg als Dilemma unbekannter Bedingungswirkungszusammenhänge als optimale Voraussetzung für eine Vorhersage. Eine letztendliche Prognose ist nur mittels Akzeptanz der Prämisse, dass ein kausaler Zusammenhang zunächst unerheblich ist,

möglich. Wie Abbildung 5 veranschaulicht, werden Personenmerkmale entsprechend als Indikatoren bzw. Korrelate des Studienerfolgs aufgefasst, welche unterschiedlich enge Zusammenhänge mit der Studienleistung aufweisen.



Abbildung 5: Zusammenhang zwischen Bewerbermerkmalen und Studienleistung

Prediction represents an effort to describe what will be found concerning an event or outcome not yet observed on the basis of data or information considered to be relevant to this unobserved event (Michael, 1969, S. 101).

Im Folgenden seien die wichtigsten Prädiktionsmodelle kurz vorgestellt.

### 3.3.1 Univariate Prädiktion

Die Vorhersage von Studienerfolg auf Basis einzelner Variablen ist laut Michael (1969) „one of the most widely employed procedure for carrying out prediction studies involving educational and psychological tests“ (S.988). Zwischen der Prädiktorvariablen und dem Kriterium oder den Kriterien für Studienerfolg werden z.B. die Korrelation und die Form der Beziehung zwischen diesen festgestellt (Regressionsrechnung). Im Modell der univariaten Prädiktion ist es also möglich, entweder die Beziehung zwischen Prädiktor und einem einzelnen Kriterium oder mehreren Kriterien zu untersuchen.

Die Vorhersagegleichung hat dann für Rohwerte die Form  $y = a + bx$ , wobei  $y$  den vorhergesagten Kriteriumswert,  $x$  die Prädiktorvariable,  $a$  den Abschnitt auf der  $y$ -Achse und  $b$  die Steigung auf der Geraden repräsentiert.

Kritiker wie Ezekiel & Fox (1959) verweisen bei dieser Form der Vorhersage auf die Notwendigkeit, vor Anwendung eines einfachen linearen Modells die Variablen auf einen linearen Zusammenhang hin zu prüfen. Die vorherzusagende Studieneignung wird durch eine Vielzahl von Bedingungen determiniert, die die alleinige Verwendung eines einzelnen Prädiktors obsolet macht.

### 3.3.2 Multivariate Prädiktion

Im multivariaten Prädiktionsmodell wird aus einer Reihe von Prädiktorvariablen die bestmögliche Variablenkombination zur Vorhersage von Studieneignung gesucht. In fast allen Fällen unter Anwendung des multiplen linearen Regressionsmodells. Die multiple Regressionsgleichung hat (für Rohwerte) dann folgende Form  $y = a + b_1x_1 + b_2x_2, \dots, b_ix_i$ .

Die Verwendung mehrerer unabhängiger Prädiktoren erhöht die Prognosekraft. Einschränkungen gibt es hinsichtlich der maximal erklärbaren Varianz in Höhe von  $R = .70$ , wenn die so genannte „criterion barrier“ erreicht ist. Außerdem werden die Ergebnisse mit größer werdender Variablenzahl im Verhältnis zur Stichprobe unzuverlässiger. Entscheidender Vorteil ist, dass sich die Gewichte verwendeter Variablen für das Zustandekommen der Kriteriumsleistung identifizieren lassen (Gaensslen & Schubö, 1973, Horst, 1971). Resümierend lässt sich sagen, dass die Anwendung des multiplen Regressionsmodells im Vergleich zur univariaten Prädiktion zu einer deutlichen Erhöhung der Vorhersagegenauigkeit führt.

### 3.3.3 Moderierte Prädiktion

Der Moderator-Ansatz trägt dann zu einer Verbesserung der Vorhersage bei, wenn es gelingt, „echte“ Moderator-Variablen ausfindig zu machen. Dies geschieht im Rahmen verschiedenster methodischer Vorgehensweisen. Bei der *Untergruppen-Analyse* wird der Einfluss einer dritten Variablen auf eine Prädiktor-Kriteriums-Beziehung in vorher festgelegten Untergruppen untersucht. Beispielsweise kann die Beziehung zwischen Intelligenz und Examensnoten unterschiedlich für Gewissenhafte und Nicht-Gewissenhafte ausfallen. Letztendlich wird die differentielle Validität beschrieben. Die *Methode differentieller Vorhersagbarkeit* definiert die Untergruppen nach der Vorhersagbarkeit. Zunächst werden Personengruppen identifiziert, für die die Vorhersage relativ genau erstellt werden kann und solche, für die die Prognose weniger genau ist. Im nächsten Schritt werden Variablen (z.B. Vorhersagetests) gesucht, die in der Lage sind, die Zugehörigkeit einer Person zu einer dieser Gruppen zu ermitteln. Bei dem *Ansatz moderierter Regression* wird die Möglichkeit einer Wechselwirkung zwischen Moderator und Prädiktor in Rechnung gestellt und diese Wechselwirkung in der Regressionsgleichung berücksichtigt. Kritiker verweisen auf das fehlende Wissen über die Wirkung von Moderatoren und Verfahren zur Identifizierung von Einflussvariablen.

### 3.4 Bedingungsmodell des Studienerfolgs

Ein Grund für die trotz inkrementeller Validitäten nicht vollständige Vorhersagbarkeit des Studienerfolgs ist neben der Prädiktor-Kriteriumsproblematik insbesondere in den nicht-standardisierbaren Einflussfaktoren während des Studiums zu sehen (Troost, 1984). Dazu zu zählen sind zum Beispiel die Qualität der Lehre, Infrastruktur der Hochschule, persönliche und finanzielle Situation des Bewerbers, rechtliche Rahmenbedingungen usw. In einem *Bedingungsmodell des Studienerfolgs*, welches in Abbildung 6 skizziert ist, wirken universitäre Charakteristika der Ausbildungsqualität ebenso wie individuelle Prädiktoren und gesellschaftliche Rahmenbedingungen aufeinander und/oder auf die Studienbedingungen und letztlich auf den Studienerfolg ein. Die individuellen Eingangsqualifikationen der Studienanfänger stellen dabei die wichtigste Determinante des Studienerfolgs dar, indem ihre positive Ausprägung die Wahrscheinlichkeit eines positiven Outputs erhöht.

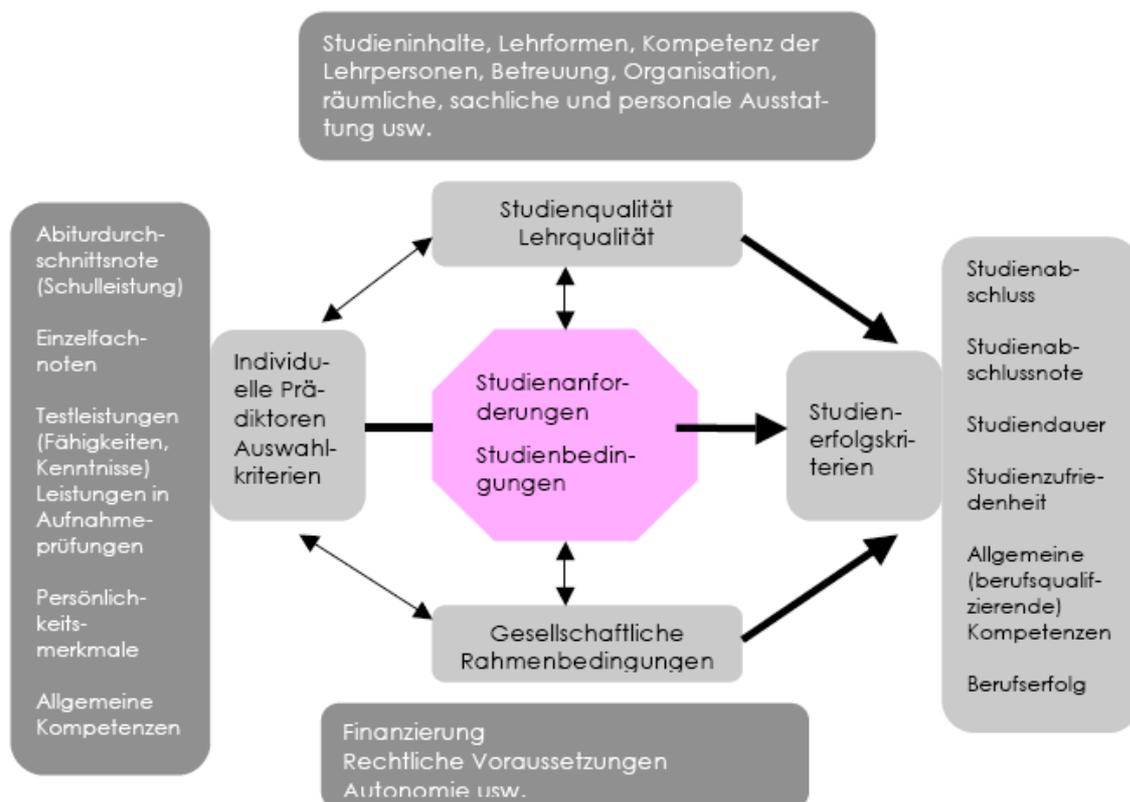


Abbildung 6: Bedingungsmodell des Studienerfolgs (Rindermann & Oubaid, 1999, S. 176)

### 3.5 Studienabbruchmodelle

In den 70er Jahren wurden in den USA zahlreiche Einzeluntersuchungen zur Problematik des Studienabbruches systematisch aufgearbeitet und modellhaft konzipiert. In der amerikanischen Literatur sind es insbesondere die Arbeiten von Spady (1970) und Tinto (1975), die der Debatte um den Studienabbruch entscheidende Impulse verliehen haben. Insbesondere das Drop-Out-Modell von Tinto prägt die Debatte um Studienabbruchgründe bis heute.

Als vermutlich erster Bildungsforscher analysiert Spady (1970) den Studienabbruch an amerikanischen Colleges in Anlehnung an Durkheims Selbstmordtheorie vergleichbar dem persönlichen Bruch mit dem normativen und sozialen System einer Gesellschaft als defizitäre soziale Integration (Durkheim, 1897). Studienabbruch ist demnach für Spady das Ergebnis einer fehlerhaften Integration in das Sozialsystem Hochschule bzw. ein Phänomen, das sich am interaktiven Prozess zwischen dem Studierenden und seiner College-Umgebung entscheidet. Wie in Abbildung 7 zu erkennen, ist als abhängige Variable die Abbruchentscheidung des Studierenden zu sehen. Spady unterscheidet zwei Arten von Studienabbrechern. Während die unfreiwilligen Abbrecher den Leistungsanforderungen nicht mehr gerecht werden, hängt eine freiwillige Studienabbruchentscheidung dagegen von der Verpflichtung des Studierenden gegenüber der Institution (institutional commitment) ab. Jenes *institutionelle Commitment* bildet sich in Abhängigkeit von der Studienzufriedenheit (satisfaction) heraus, welche direkt von der *sozialen Integration* (social integration) als zentrale Variable des Modells beeinflusst wird.

Der Grad der sozialen Integration bestimmt sich dabei aus einem komplexen Prozess, der maßgeblich von der Unterstützung durch Freunde (friendship support), von der intellektuellen Entwicklung (intellectual development), den akademischen Leistungen (grade performance) und der normativen Übereinstimmung (normative congruence) des Studierenden mit der Institution bestimmt wird. Exogene Variable in diesem Modell ist der familiäre Hintergrund (family background), der für das Ausmaß der normativen Übereinstimmung sowie des akademischen Potenzials (academic potential) verantwortlich ist. Dem sozialen Gleichgewicht wird insofern eine besondere Stellung zuteil, als es einerseits – mit Ausnahme von Leistungsversagen – normative Inkongruenzen zwischen Gleichgesinnten und der Institution kompensieren und andererseits akademische Fortschritte begünstigen kann. Besonders für Frauen liegt in der sozialen Anbindung ein wichtiger Integrationsfaktor. Das Modell ist offen für Veränderungen von Interessen und Zielen des Studierenden, da Spady ausdrücklich eine Rückkopplung vom institutionellen Commitment zur normativen Übereinstimmung zulässt.

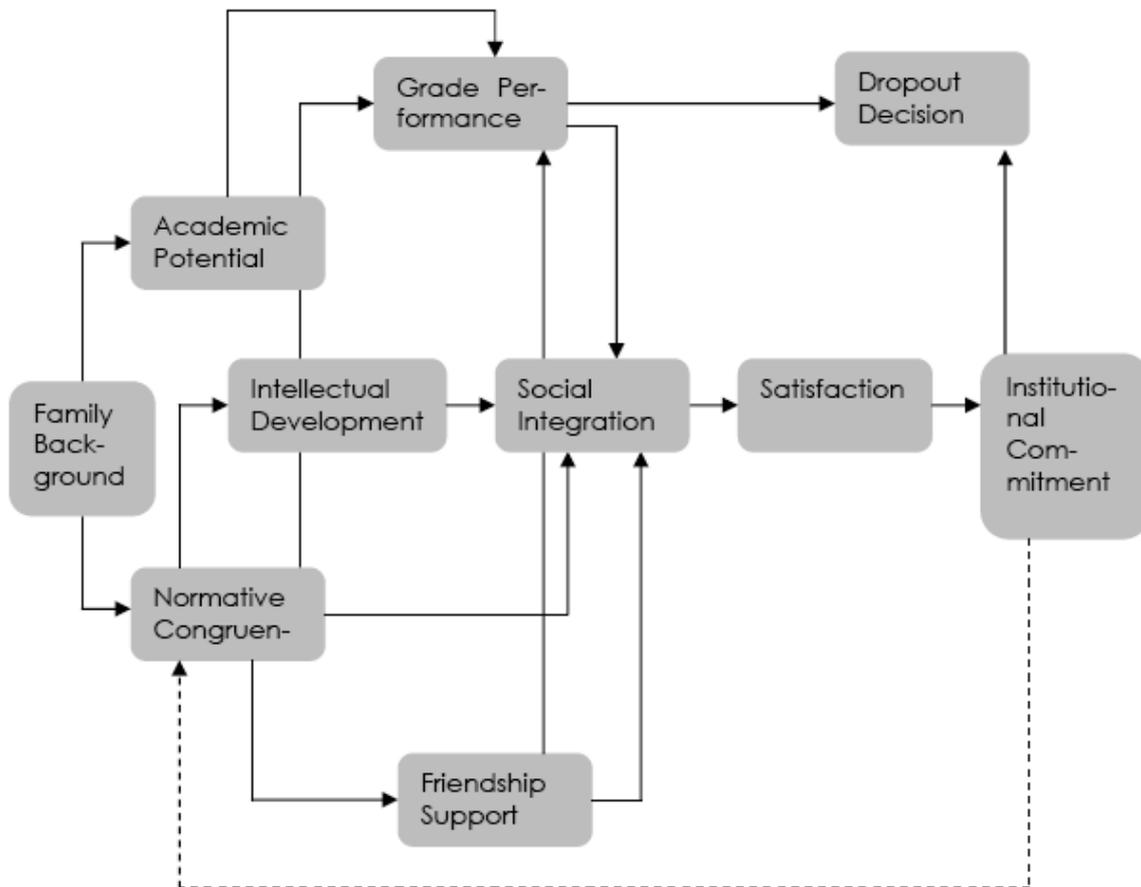


Abbildung 7: Soziologisches Modell des Drop-Out-Prozesses nach Spady (1970, S. 79)

Fünf Jahre später greift *Tintos Modell* (Tinto 1975, 1993) die Überlegungen Spadys auf und entwickelt diese für amerikanische Colleges weiter. Auch er bezieht sich dabei auf Durkheims Selbstmordtheorie, benutzt sie allerdings nur als Deskriptivmodell zur Spezifikation von Bedingungen für verschiedenes Dropout-Verhalten.

Wie in Abbildung 8 veranschaulicht, beschreibt auch in Tintos Modell die Abbruchentscheidung die abhängige Variable, die direkt durch das Ausmaß an institutionellem Commitment (institutional commitment) und durch die Höhe der Verpflichtung gegenüber den eigenen Zielen (goal commitment) beeinflusst wird. In der Integration dieser Selbstverpflichtung liegt der erste Unterschied zu Spadys Modell. Tinto fokussiert den Zusammenhang zwischen der normativen Übereinstimmung und der Verpflichtung gegenüber der Institution nicht auf die soziale Integration (social integration). Er hebt die intellektuelle Studienentwicklung zu einer eigenen, primär dem Studienziel verpflichteten Integrationsleistung (academic integration) heraus, welche eigenständig neben der vielmehr auf die institutionelle Anbindung verwiesene sozialen Integration steht. Soziale Integration bildet sich dabei in Abhängigkeit

von den Interaktionen mit anderen Studierenden (peer-group interactions) sowie mit den Angehörigen der Fakultät (faculty interactions). Die akademische Integration nimmt er als beeinflusst von dem akademischen Erfolg (grade performance) und der intellektuellen Entwicklung (intellectual development) an. Diskrepanzen zwischen intellektuellem Aspirationsniveau und dem vorherrschenden geistigen Klima des Colleges führen zum freiwilligen Studienabbruch. Commitment definiert Tinto nicht nur als sich aus diesen Variablen ergebende Größe, sondern geht davon aus, dass bereits vor dem Eintritt in die Hochschule sowohl *goal* als auch *institutional commitment* in einer bestimmten Form vorhanden sind. Diese Ausgangsgröße wiederum wird in seinem Modell maßgeblich durch die exogenen Variablen familiäre Herkunft (family background), individuelle Merkmale (individual attributes) sowie die Ausbildung vor Eintritt in die Hochschule (pre-college-schooling) beeinflusst.

Noch stärker als bei Spady, der vor allem Kausalbeziehungen abbildet, zeigt das Modell von Tinto den vielschichtigen Interaktionsprozess zwischen dem einzelnen Studierenden und dem akademischen und sozialen System eines College, der schließlich über den Ausgang des Studiums entscheidet. Tinto zeigt durch die Typisierung unterschiedlichen Studienabbruchverhaltens die „Dynamik der Bedingungsvariablen von Studienverlaufsentscheidungen“ (Ströhlein 1983, S. 56) auf. In seinem neueren Werk (Tinto, 1993) führt er zudem noch so genannte externe Commitments, also Verpflichtungen außerhalb des Systems der Hochschule, ein, von denen er annimmt, dass sie ebenfalls das Commitment und damit die Abbruchentscheidung determinieren.

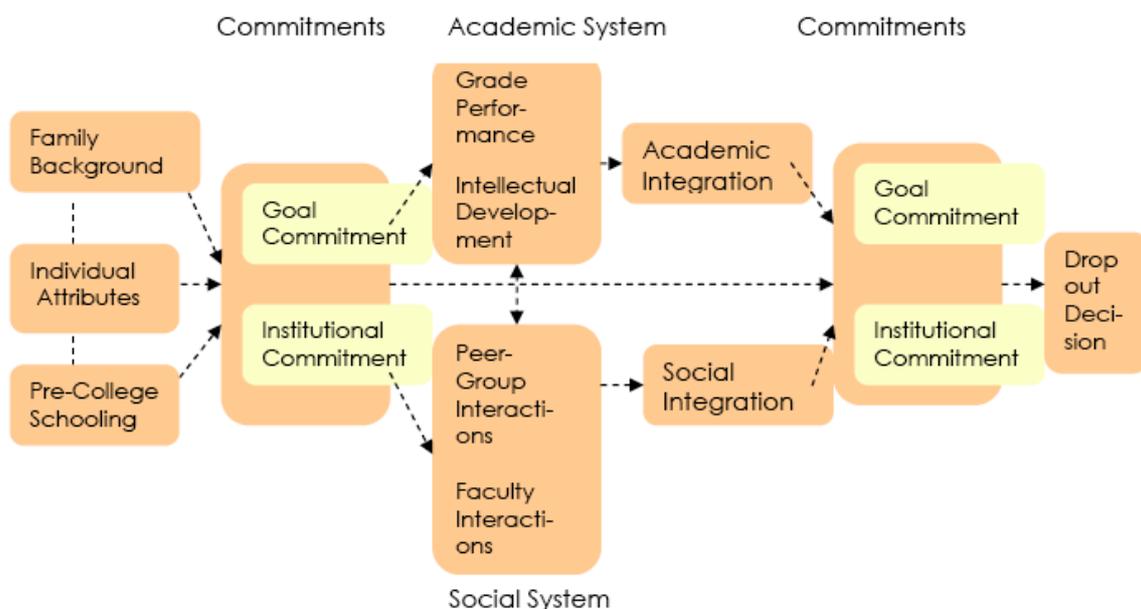


Abbildung 8: Tintos konzeptionelles Studienabbruchmodell (Tinto, 1975, S. 95)

Tintos Drop-Out-Modell bestimmt noch heute die Auseinandersetzung mit dem Studienabbruch in den USA (vgl. Seidman, 1996). In den USA beschloss man infolgedessen eine Vielzahl von Maßnahmen zur Förderung der Integration der Studierenden in das akademische und soziale System der Hochschulen. Diese reichen von Orientierungsprogrammen bis hin zu umfangreichen Beratungsangeboten (vgl. Seidman 1996). Kritische Auseinandersetzungen betreffen das Missachten von Variablen, die im Bedingungsmodell des Studienerfolgs (vgl. oben) als Studien- oder Lehrqualität bezeichnet wird. Weiterhin nimmt die Hochschule in Tintos Modell eine vergleichsweise passive Rolle ein. Einzig und allein der Studierende wird als variabel betrachtet und muss sich den Gegebenheiten anpassen oder die Hochschule verlassen.

Monieren lässt sich an Tintos *Konzept der institutionellen Verpflichtung* die unterstellte Verbundenheit des Studenten mit der speziellen Hochschule. Diese lässt sich außer für amerikanische Colleges in Deutschland nur auf Campus-Universitäten wie die der Bundeswehr anwenden. Trotz der eingeschränkten Übertragbarkeit auf deutsche Hochschulverhältnisse ermöglichte Tintos Denkansatz auch in Deutschland Anregungen zur Entwicklung weiterer Forschungsarbeiten und Prognosemodelle (Gold, 1988). Der Prozesshaftigkeit des Verlaufskriteriums entsprechend wurden Längsschnittuntersuchungen durchgeführt und multivariat analysiert (Gold, 1988, Ströhlein, 1983). Das *konflikttheoretische Modell* von Ströhlein (1983) und das *Bedingungsmodell subjektiven Studienerfolgs* von Gold (1988) sind zwei prominente Beispiele deutscher Forschungsarbeiten. Zwar enthalten beide Modelle auch Bezüge zu den amerikanischen Grundlagenmodellen, jedoch entfalten sie eigene Zugänge zur Klärung der Abbruchproblematik.

Ausgehend von an einem dynamischen Gleichgewicht orientierten konflikttheoretischen Ansatz werden bei Ströhlein (1983) abweichende Studienverlaufsentscheidungen beispielsweise Studienabbrüche als Folgen einer Diskrepanz zwischen Fähigkeiten, Studierenerwartungen und -zielen einerseits und den tatsächlichen Studienanforderungen und -erfahrungen andererseits interpretiert. Ströhlein entwickelt soziodemographische und psychologische Bedingungsfaktoren und prüft, inwieweit diese den Studienabbruch begründen können. Gold (1988) konzeptualisiert im Vergleich zu Ströhlein weniger den Studienabbruch als vielmehr psychologische Bedingungen individuellen Studienerfolgs. Er vergleicht überdauernde Persönlichkeitsmerkmale und Studierenerfahrungen bei ungefährdet Studierenden, Studierenden mit Abbruchneigung und Studienabbrechern und leitet daraus Bedingungsfaktoren des Studienabbruches ab. Zwar resultiert nach ihm der subjektive Studienerfolg generell aus leis-

tungsfördernden Arbeitshaltungen und emotionaler Stabilität, doch messen die Abbrecher und die Abbruchgefährdeten diesen Persönlichkeitsmerkmalen höhere Bedeutung zu als die Gruppe der ungefährdet Studierenden. Überdies hängt der Erfolg für die Studienabbrecher stärker von den beiden anderen Persönlichkeitsfaktoren, den sozialen und intellektuellen Fähigkeiten, ab. Kennzeichnend für „Beinahe-Abbrecher“ sind begrenzte soziale Kompetenzen sowie hervorstechende emotionale Labilität. Auch die Rolle der im universitären Umfeld erfahrenen Anerkennung und Integration sind zur Vorhersage der „Vulnerabilität“ prognostisch relevant. In einer prospektiven Bedingungsanalyse finden Gold und Kloft (1991) Unterschiede zwischen Studienabbrechern und Nichtabbrechern hinsichtlich selbst- und fremdbeurteilter Leistungsvoraussetzungen und leistungsfördernder Arbeitshaltungen sowie der Rolle im universitären Umfeld. Lernprobleme bzw. Leistungsschwierigkeiten, eine unzureichende Motivation und fehlende Anerkennung durch Kommilitonen sind wichtige Vorhersagevariablen für einen Studienabbruch.

In den 90er Jahren steht innerhalb der Abbruchforschung die Exploration studienfachspezifischen Abbruchverhaltens an einzelnen Hochschulen im Vordergrund. Die potentiellen Hintergründe der von Studierenden am häufigsten angegebenen Abbruchgründe wie Distanz zum Studium, Kritik an Didaktik oder Begleitumstände des Studiums oder auch Wunsch nach Praxis (Lewine al. 1995, S. 2) unterstützen die Vermutung suboptimaler Ausbildungsentscheidungen. Zur persönlichen muss jedoch die institutionelle Beratung hinzukommen, damit Anforderungs-, Leistungs- und Fähigkeitsprofile unter überschaubaren Rahmenbedingungen und von beiden Seiten besser abgestimmt werden können.

Das mit bundesweit repräsentativen Untersuchungen im Hochschulbereich (Hochschulplanung und Studentenforschung) befasste Hochschul-Informationssystem in Hannover (HIS) ermittelt den in Tabelle 4 veranschaulichten rasanten Anstieg der Abbruchquoten über einen Zeitraum von beinahe 20 Jahren (HIS, 2002). Die in den 80er Jahren angegebenen Abbruchquoten aus Längsschnittuntersuchungen mit Studienberechtigten aus den 70er Jahren variieren zwischen zehn und 16 Prozent und bestätigen im Großen und Ganzen die Angaben von HIS. Seit den 90er Jahren gelten die durchschnittlichen Abbruchquoten als konstant (ebd.). In anderen Ländern Europas betrug die Abbrecherquote bereits gegen Ende der 70er Jahre mehr als 20 Prozent (Micko & Nippert 1979, S. 19), für die USA wurde sie mit 50 Prozent angegeben (Tinto, 1982, S. 694). Diese sehr hohe Zahl hängt mit der alle vorzeitigen Abgänger von Colleges einschließenden Definition des Kriteriums zusammen.

Tabelle 4:

*Ermittlung der Studienabbruchquoten deutscher Hochschulen in Prozent. Quellen: Griesbach et al., 1977, S. 167; Reissert & Marciszewski, 1987, S. 3; Griesbach et al. 1992, S. 7; Lewin et al. 1995, S. 2*

Studienjahr	1974	1979	1984	1989	1991
Quoten	7	9	16	27	31

Verglichen mit früheren HIS-Untersuchungen zum Studienabbruch zeigen die jüngsten Befragungen eine wachsende Komplexität des Begründungsspektrums. Führte auch vor Jahren selten ein einzelner Grund das vorzeitige Studienende herbei, so scheinen sich inzwischen die Folgen wachsender Studentenzahlen, knapper Studienplätze und unberechenbarer Arbeitsmarktchancen im aufgekünigteten Studierverhalten Vieler abzuzeichnen. Im Zeitvergleich nehmen die falschen Studienvorstellungen, die enttäuschten Studierenerwartungen und das Desinteresse am Studienfach sowie die Kritik an der Hochschuldidaktik und den institutionellen Begleitumständen des Studiums zu. Ferner werden an der Berufspraxis orientierte Studierenerwartungen lauter. Finanzielle Engpässe häufen sich. Auch Überforderung, die Erwartung nachteiliger Arbeitsmarktchancen oder finanzielle Gründe gehörten zu den entscheidenden Abbruchmotiven.

Tabelle 5:

*Studienabbruchquoten für deutsche Studierende an Universitäten nach Fächergruppen und ausgewählten Studienbereichen in Prozent (HIS-Studienabbruchuntersuchung, 2005, S. 19)*

Studienbereiche	insgesamt		männlich		weiblich	
	1999	2002	1999	2002	1999	2002
Bezugsjahrgang: Absolventen						
Sprach-, KulturWi, Sport	33	35	38	39	34	21
	Sprach-, KulturWi	41	45			
	Pädagogik, Sport	28	23			
Rechts-, Wirtschafts-, SoWi	30	28	28	30	31	26
	Sozialwissenschaften, Sozialwesen	42	36			
	Rechtswissenschaften	27	16			
	WiWi	31	32			
Mathematik, NaturWi	23	26	27	28	18	23
	Informatik	37	38			
	Physik, GeoWi	26	30			
	Geographie	36	19			
Ingenieurwissenschaften	26	30	27	30	19	28
	Maschinenbau	25	34			
	Elektrotechnik	23	33			
	Bauwesen	35	30			

*Anmerkung zu den Abkürzungen:* KulturWi: Kulturwissenschaften; SoWi: Sozialwissenschaften; NaturWi: Naturwissenschaften; WiWi: Wirtschaftswissenschaften; GeoWi: Geowissenschaften.

Tabelle 6:

*Studienbereichen in Prozent (HIS-Studienabbruchuntersuchung Studienabbruchquoten für deutsche Studierende an Fachhochschulen nach Fächergruppen und ausgewählten, 2005, S. 23)*

Studienbereiche	Insgesamt		Männliche		Weiblich	
	1999	2002	1999	2002	1999	2002
Bezugsjahrgang: Absolventen	1999	2002	1999	2002	1999	2002
Wirtschafts-, Sozialwesen	16	25	21	29	12	21
Sozialwesen	6	20				
Pädagogik, Sport	25	27				
Mathematik, NaturWi	34	40	34	38	34	49
Informatik	36	39				
Ingenieurwissenschaften	21	20	23	21	14	11
Maschinenbau	25	21				
Elektrotechnik	20	32				
Bauwesen	24	20				

Anmerkung zu den Abkürzungen: NaturWi: Naturwissenschaften.

Trotz von Fachrichtung zu Fachrichtung unterschiedlicher Tendenzen in der Entwicklung der Studienabbruchquoten kam es mit Ausnahme der mathematisch-naturwissenschaftlichen Fächer zu keiner Veränderung im Verhältnis von Universitäten zu Fachhochschulen. Im Laufe der neunziger Jahre beendeten in fast allen Fächergruppen die Studienanfänger an den Fachhochschulen seltener ihr Studium ohne Examen als Kommilitonen anderer Universitäten. Die stärker auf Strukturierung, Berufsbezug und klare Orientierung ausgerichtete Studiengestaltung an den Fachhochschulen könnte sich in dieser Hinsicht als Vorteil erweisen. Jüngste Untersuchungen um Heine (Heine et al., 2005) identifizieren drei hauptsächliche Gründe für unzureichenden Studienerfolg. 1. Informationsdefizite, 2. Sachfremde Kriterien der Hochschulwahl, z.B. in Bezug auf die Gegebenheiten des Hochschulorts und 3. Kenntnis- und Kompetenzdefizite, z.B. Mathematikkenntnisse oder fehlende Techniken wissenschaftlichen Arbeitens. Die ersten beiden Kriterien werden so gut wie gar nicht in heute üblichen statusdiagnostischen Auswahlverfahren berücksichtigt.

### 3.6 Modelle der Person-Umwelt-Passung

Zuordnungsmodelle im Sinne der Person-Umwelt-Passung wie die Trait-And-Factor-Theorien oder die Person-Environment-Fit-Ansätze blicken in der Berufspsychologie auf eine lange Tradition zurück. Zentrale Annahme ist das Vorhandensein positiver Personeffek-

te bei hoher Kongruenz zwischen Person und Umwelt. Eine hohe Übereinstimmung liegt wiederum dann vor, wenn personale Interessen, Fähigkeiten und Eigenschaften den Befriedigungsmöglichkeiten bzw. Anforderungen in der jeweiligen Umwelt entsprechen. Das Passungskonzept, wie es beispielsweise im Person-Job-Fit-Ansatz substituiert (Amelang, 1997) wird, erweitert die Dominanz der kognitiven Erfolgskomponenten bei Studienfachwahl, Studiengangwechsel oder auch Studienabbruch um nicht kognitive Moderatorvariablen (Rindermann & Oubaid, 1999).

Zahlreiche Studien weisen dementsprechend eine Korrelation zwischen subjektiver Passung und Studienerfolg nach: Bereits Parsons (1909) schreibt Personen mit einer klaren Berufsbildvorstellung eine gesteigerte Arbeitszufriedenheit zu. Super et al. (1963) stellt eine Selbsteinschätzung, die den Anforderungen des Berufs annähernd entspricht, ins Zentrum der Theorie der beruflichen Entwicklung. Das Person-Umwelt-Modell nach Holland (1985) stellt berufliche Interessen in den Mittelpunkt. Holland definiert sechs Persönlichkeits- und Umwelttypen mit entsprechenden Interaktionen. Demnach suchen sich Menschen den am besten zu ihren Neigungen und Fähigkeiten passenden Umwelttyp aus. Metaanalysen (Spokane, 1985; Assouline & Meier, 1987) belegen Hollands Hypothese positiver Effekte bei Person-Umwelt-Übereinstimmung. Spokane (1985) klärt darüber hinaus positive Beziehungen von Kongruenz und Studienleistungen sowie persönlicher Ich-Stärke auf. Im Zusammenhang mit Hollands Person-Umwelt-Modell findet Müller (2001) bei Studierenden der Universitäten der Bundeswehr einen signifikanten Zusammenhang zwischen Studieninteresse, Persönlichkeitsstruktur des Studierenden und Studiengang. Ein bedeutsamer Unterschied im Persönlichkeitsprofil zeigt sich zwischen technischen und nicht-technischen Studiengängen. Während sich Techniker eher sachlich orientiert und reservierter präsentieren, weisen Studierende mit nicht-technischem Studienwunsch eine augenfällige zwischenmenschliche Orientierung auf. Wer nicht dem „Idealprofil“ einer Fächergruppe entspricht, verfügt gleichsam über geringeres Studieninteresse. Eine wesentliche Ursache für Studienabbruch sehen auch Winteler, Schiefele & Krapp (1986) in der Divergenz zwischen ursprünglichen Motiven für die Studienfachwahl und Art der Studienangebote, im Studium geforderte Leistungen und andere Umweltbedingungen. Eine Studienwahl ungeachtet fehlender Passung z.B. aufgrund extrinsischer Belohnung (wie Studium bei vollem Gehalt) oder der eigenen Interessen und Studieninhalte führt verstärkt zu Studienabbruch ( $r = .35$ ), ferner zu geringerem subjektiven Wohlbefinden im Studium ( $r = -.67$ ) und zu einer schlechteren Studiumsbewertung ( $r = -.53$ ) (Rolfs, 2001). Nur ungefähr die Hälfte der Studienanfänger zeigen sich nach eige-

nem Urteil gut informiert über die Studienanforderungen im gewählten Studiengang (Heine et al., 2005). Wie sehr der Studienerfolg von der beschriebenen Art von Person-Umwelt-Korrespondenz abhängt, zeigen auch die Ergebnisse der Linzer Arbeiten zum Studienverlauf (vgl. dazu z.B. Bergmann, Brandstätter & Eder, 1994; Brandstätter, Farthofer & Grillich, 2001).

Studierende erhalten offenbar vor allem in den ersten Semestern zu wenig Orientierung, um die Veränderungen des sozialen Umfeldes und die fachlichen Anforderungen bewältigen zu können, die mit der Aufnahme eines Studiums verbunden sind (Wissenschaftsrat, 2004, S. 20f).

Nach der Person-Environment-Fit-Theorie nach Caplan (1987) ist die Kongruenz von Person und Arbeitsumwelt als ein ausschlaggebender Faktor für Stress verantwortlich. Heise et al. (1997) schließen als Folge optimaler Passung des Studierenden mit dem Studienfach eine allgemeine Studienzufriedenheit.

Mit der Etablierung differenzierter valider Eignungsfeststellungsverfahren sind sowohl für die Hochschulen als auch für alle potenziellen Studierenden Vorteile als Auswirkung einer erhöhten Passung zu erwarten.

### 3.7 Exkurs Anforderungsanalyse

Zur Realisierung der beschriebenen Ziele oben aufgeführter Person-Umwelt-Passungsmodelle sollten Studieneignungsfeststellungsinstrumente zum einen in der Lage sein, spezifische Anforderungen des Studiengangs bzw. der für den Studienerfolg entscheidenden Variablen erschöpfend zu berücksichtigen. Zum anderen sollten sie Hinweise über die unterschiedliche Ausprägung von Schlüsselqualifikationen und den Eignungsgrad des Bewerbers geben können. Die Aussagefähigkeit erhobener Daten erweist sich demnach als umso aussagefähiger je mehr die realen Anforderungen mit den im Eignungsprozess berücksichtigten Beurteilungsdimensionen übereinstimmen. Daher stellt eine exakte Anforderungsanalyse die Grundlage der Verfahrenskonstruktion, -durchführung und -auswertung und nicht zuletzt die qualitätssichernde Basis für die Beurteilung einer Passung dar (Rosenstiel, 1980). Doch nicht jedes Auswahlverfahren, das auf der Grundlage von Anforderungsanalysen erstellt wird, verdient blindes Vertrauen seitens der Anwender. Jedes Auswahlinstrument muss nach seiner Konstruktion und anschließend in wiederkehrenden Abständen überprüft werden (Lienert, 1969). Somit sollte stets jedes Verfahren individuell untersucht werden, inwieweit es

den mess- und testtheoretischen Anforderungen gerecht wird und inwiefern der Einsatz eines Verfahrens aufgrund seiner Qualität zu rechtfertigen ist.

Die Studierendenauswahl muss dabei als ständiger in sich systematisch verzahnter Prozess der Informationsgewinnung über die Eignung und berufliche Weiterentwicklung eines Bewerbers betrachtet werden. Ein allgemeines Modell dieses Prozesses der Eignungsfeststellung kann mit Hilfe der drei Phasen Planung, Durchführung und Evaluation in Abbildung 9 veranschaulicht werden.

„Wer das erste Knopfloch verfehlt,  
kommt mit dem Zuknöpfen nicht zu Rande.“  
Johann Wolfgang v. Goethe



Abbildung 9: Phasen der Eignungsfeststellung von Bewerbern - Idealtypische Darstellung

Gemäß DIN-Norm 33430 zu den „Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen“ ist ein grundlegender Standard die Planung des Auswahlverfahrens. „Die berufsbezogene Eignungsbeurteilung ist sorgfältig zu planen, indem alle Aspekte der Durchführung und Auswertung der Verfahren zur Eignungsbeurteilung vorab festgelegt werden“ (DIN 33430, 2002, S. 8). Als Grundlage jeder Personalplanung

und -beurteilung gilt die exakte und umfassende Definition der Anforderungen (Jeserich, 1981). Ein qualitativ hochwertiges Messinstrument zu verwenden ist eine notwendige aber keine hinreichende Voraussetzung für eine effektive Personalauswahl. Ebenso bedeutsam ist die Angemessenheit des Messinstruments für die Anwendungssituation. Gegen diese Regel wurde im Bereich der Eignungsdiagnostik teilweise verstoßen. So wurden klinisch diagnostische Instrumente wie der Minnesota Multiphasic Personality Inventory (MMPI) oder Sechzehn-Persönlichkeitsfaktoren-Test (16 PF) auch für arbeitspsychologische Fragestellungen wie die der Personalauswahl verwendet. Ein Auswahlverfahren wird in der Regel nur dann für eine Berufsanforderung valide sein, wenn die im Auswahlverfahren repräsentierten Anforderungen mit denjenigen des Arbeitsplatzes übereinstimmen (Arbeitskreis Assessment-Center, 1992a).

Vor Gestaltung eines Studieneignungstests müssen demnach die Anforderungen eines spezifischen Studiengangs, Studienerfolgskriterien und prognostisch valide Eingangsmerkmale auf Seiten der Studienanfänger definiert werden, um den Grad der Übereinstimmung zwischen Anforderungen des Studiengangs und der weiteren Hochschulumgebung mit den Leistungsvoraussetzungen der Person zu maximieren. (Amelang, 1997). „Auswahlverfahren können ohne entsprechende Anforderungsprofile der Studiengänge und Berufe nicht gestaltet und legitimiert werden.“ (Rindermann & Oubaid, 1999, S. 185). Zu den möglichen standardisierten Anforderungsanalysemethoden zählen Fragebogen, Checklisten, Beobachtungsinterviews. Diese Daten besitzen ein Maximum an Objektivität und Zuverlässigkeit; ihre Auswertung ist transparent und einfach möglich (DIN 33430). Teilstandardisierte Analysemethoden stellen mittels zugrunde liegender Leitfäden sicher, keine wichtigen Sachverhalte entgehen zu lassen. Als Beispiele sind Expertenbefragungen (Lehrende, Studenten etc.), die „Critical Incident Technique“ nach Flanagan (1954) und die „Repetory Grid-Technik“ nach Kelly (1955) zu nennen. Weitere häufig verwendete Quellen stellen das unstrukturierte Interview, freie schriftliche Berichte, die Dokumentenanalyse z.B. von niedergeschriebenen Curriculum-Anforderungen, Studien- und Prüfungsordnungen, Informations-Broschüren, Material der Studien- und Berufsberatung usw. dar. Diese nicht standardisierten Verfahren der Anforderungsanalyse liefern zusätzlich einen Überblick über den Studienaufbau, involvierte Fächer, die Art der Lehrveranstaltungen und die zu erbringenden Leistungsnachweise zum Beispiel für das Vordiplom. Darüber hinaus können aber auch die Ergebnisse aus Absolventenstudien oder Überlegungen zur Profilbildung und zum geforderten Leistungsniveau Berücksichtigung finden.

Tabelle 7 listet die beschriebenen und weitere Analysemethoden wie den Paarvergleich, Fragebogenverfahren, die empirische Befragung et cetera mit ihren jeweiligen Vor- und Nachteilen auf. An erster Stelle ist die Expertenbefragung aufgeführt, weil sie nicht nur in der Praxis mit 69 prozentiger Anwendung gegenüber den anderen Analyseformen dominiert (Krause, Meyer zu Kniendorf & Gebert, 2001), sondern ebenso Grundlage der Konzeption des Eignungsfeststellungsverfahrens für Offizierbewerber ist. Vorteilen wie Praktikabilität, geringer Aufwand und hohe Akzeptanz müssen wissenschaftlich-methodische Einwände entgegen gesetzt werden. Expertenbefragungen sind abhängig von der subjektiven Meinung gerade anwesender Experten und vorherrschender eignungsdiagnostischer Lehrmeinungen. Ein weiterer Nachteil ist die Bottom-Up-Herangehensweise. Diese versucht zu beantworten, wie sich gegenwärtige oder ehemals erfolgreiche von weniger erfolgreichen Positionsinhabern unterscheiden, statt zukünftige Anforderungen im Profil abzubilden. Außerdem beschränkt sich die Verständlichkeit hauptsächlich auf den Kreis der Formulierer (Jeserich, 1981).

Tabelle 7:

*Auswahl an Methoden der Anforderungsanalyse mit ihren Vor- und Nachteilen (Reimann)*

<b>Methode</b>	<b>Vorteile</b>	<b>Nachteile</b>
<b>Unstandardisiert</b>		
Unstrukturiertes Interview	Vorgabe eines thematischen Rahmens für das Gespräch. Studienbewerber setzt nach offenem Einstieg eigene Schwerpunkte. Flexibel einsetzbar.	Abbild lediglich von Teilaspekten tatsächlicher Anforderungen.
Dokumentenanalyse	Geringer Aufwand	Studienfachbeschreibungs-Material hat unter Umständen andere Zielsetzung (Werbung). Je nach Qualität zu wenig Genaues über situative Herausforderungen.
<b>Teilstandardisiert</b>		
Teilstrukturiertes Interview	Wechsel von offenen und geschlossenen Fragen ermöglicht Filtersetzung bzw. geschicktes offenes Fragen bzgl. Zielen, Motiven und Strategien der Studienbewerber.	
Critical Incident Technik (Flanagan 1954)	Geringer Aufwand (Interviews mit Experten oder Workshop). Das am häufigsten verwendete Verfahren. Auch Berücksichtigung der situativen Rahmenbedingungen – nicht nur Sammeln von Eigenschaftslisten. Berücksichtigung situativer Anforderungen möglich.	Prozess hin zu einem formalen Anforderungsprofil nicht transparent. Zu sehr bottom-up, an bisherigem Stelleninhaber orientiert. Systematischer als Expertenbefragung.
Expertenbefragung (zum Beispiel OSS)	Geringer Aufwand. Geeignet, wenn es nur wenige Exper-	Prozess hin zu einem formalen Anforderungsprofil nicht transparent. Zu sehr

1948, EFV OPZ in der Bundeswehr)	ten gibt (zum Beispiel aktueller Stelleninhaber).	bottom-up, an bisherigem Stelleninhaber orientiert.
Repertory-Grid-Technik (REP) (Kelly 1955)	Einbezug der Subjektivität der Befragten. Hoher heuristischer Wert – am ehesten Vermeidung, dass nur altbekannte Eigenschaftslisten wiederholt werden.	Prozess hin zu einem formalen Anforderungsprofil nicht transparent. Zu sehr bottom-up, an bisherigem Stelleninhaber orientiert. Merkmalslisten vernachlässigen situative Anforderungsmerkmale.
<b>Standardisiert</b>		
Fragebogen-Verfahren zur Anforderungs-Analyse (Frieling & Hoyos 1978)	Identisches Instrument für verschiedene Positionen - damit guter Vergleich von Profilen. Nutzen von Experten-Know-How zum Thema.	Für Führungskräfte kein deutsches Verfahren bewährt. Sprache und Begrifflichkeit standardisiert, nicht für die Studierendenauswahl geeignet.
Top-Down Vorgehen	Einziges Verfahren, das mehr die zukünftigen Anforderungen im Profil abbildet.	<b>Prozess hin zu einem formalen Anforderungsprofil nicht transparent. Einseitig qualitatives Vorgehen.</b>
<i>Anmerkungen.</i> OSS: Office of Strategic Services, EFV OPZ: Eignungsfeststellungsverfahren der Offizierbewerberprüfzentrale.		

Die Durchführungsphase schließt sich an die Prüfung der Eignungsmerkmale während des Eignungsfeststellungsverfahrens bzw. an die Informationsgewinnung an. Die Prüfung der Eignungsmerkmale bildet den Ausgangspunkt, um eine angemessene Einschätzung der Bewährungswahrscheinlichkeit abgeben zu können. Diese Beurteilung mündet in eine Entscheidung über die Ablehnung oder die Einstellung des Bewerbers in der Berufsauswahl bzw. Zulassung zu einem Studiengang in der Studierendenberatung sowie im günstigsten Fall in eine Vereinbarung von Entwicklungsmaßnahmen (Schuler, 1992) bzw. in eine Studienberatung in Bezug auf Alternativen. Auswahlverfahren bieten zunehmend die Möglichkeit zusätzlicher Rückmeldungen für Bewerber und Hochschulen bzw. Institutionen. Die Informationen ermöglichen Entscheidungsprozesse auf Seiten der Bewerber oder strukturelle Optimierungen auf Seiten der auswählenden Einrichtung. „Ausleseverfahren von Organisationen sind Kompromisse, die Validitätskriterien ... genügen müssen.“ (Kompa, 1989, S. 12). Die mittels des Auswahlverfahrens angestrebte Prognose zukünftiger Leistungsunterschiede muss aus ökonomischen Gründen mindestens periodisch überprüft werden (Barell, 1992). Kontinuierliche Bewährungskontrollen bieten anwendungsbegleitende Verbesserung von Anforderungen, Methoden und Regeln. Es handelt sich hier um die Evaluationsphase einer Eignungsfeststellung (vgl. Abbildung 9).

Die drei Phasen stehen keinesfalls für sich alleine. Nur als zusammengefügte Teile eines Modells stellen sie die wissenschaftlich fundierte Grundlage einer jeden Eignungsbewertung sicher. Insbesondere die systematische periodische bzw. kontinuierlich durchzuführende Evaluation der Inhalte, Methodiken und Verfahren im Nachgang liefert im letzten Schritt die

Basis für eine Bewertung und Fortentwicklung. Diese Erfolgskontrolle ist außerdem grundlegend für die Bildung und Überwachung einheitlicher und angemessener Auswahlmaßstäbe. Sie wirkt somit nicht nur auf den Ablauf, sondern auch auf den Vorlauf zurück. Die einzelnen Phasen formatieren zu einem Kreislaufphasenmodell mit der beschriebenen Feedbackschleife (Doppelpfeil in Abbildung 9) zwischen Durchführung und Evaluation.

#### 4. Zusammenfassung Konstrukt Studieneignung

Im Rahmen der Bestimmung des Begriffs Studieneignung wird ein Gefüge bestimmter Leistungsvoraussetzungen angenommen, das als hypothetisches Konstrukt mit dem Terminus Studieneignung verstanden werden kann (vgl. Kapitel 1). Neben individuellen kognitiven Fähigkeiten als einflussreichste Voraussetzungen werden außerdem Schlüsselqualifikationen im Sinne von Persönlichkeitsmerkmalen und Arbeitshaltungen aber auch Interesse subsumiert. Studienerfolg lässt sich ebenso wenig wie Studienabbruch monokausal auf bestimmte Studieneingangsvoraussetzungen oder übrige Faktoren zurückführen. Vielmehr ist er das Ergebnis eines komplexen Wechselwirkungsprozesses, deren Komplexität in den vorangegangenen Kapiteln durch nähere Betrachtung der Korrelate von Studienerfolg, seinen Kriterien und moderierenden Einflussvariablen ausführlich beschrieben wurde. Theoretische und praxisnahe Prognose-, Studienabbruch-, Passungs- und Studienberatungsmodelle ergänzen und unterstützen Forschungsergebnisse zu möglichen Prädiktoren und Kriterien von Studierfähigkeit und beleuchten die interne Struktur des komplexen Konstrukts Studieneignung.

Die Hochschulforschung und -politik prägenden Drop-Out-Modelle stellen Einflüsse interaktiver Prozesse zwischen Student, Hochschule und externen Variablen in den Vordergrund. Umfangreichste Untersuchungen des HIS (2002) bestätigen die Annahme, dass eine genaue Abgrenzung entlang der Abbruchgründe und Risikofaktoren trotz vorheriger Identifizierung abbruchgefährdeter Studenten beinahe unmöglich erscheint. Jedoch konnte in einflussreichen Untersuchungen vor allem fehlendes oder von außen beeinflusstes eigenes und insituionelles oder externes Commitment als wichtige Einflussvariable entsprechenden Drop-Out-Verhaltens bestimmt werden. Im Sinne aller Modelle zur Person-Umwelt-Passung hat das Auswahlverfahren bei Feststellung einer Über- wie auch Unterforderung bzw. Nicht-Kongruenz der Fähigkeiten, Bedürfnisse und Erwartungen des Studierenden mit den Anforderungen im Studium außerdem eine nicht zu unterschätzende Beratungsfunktion. Eine eingehende Auseinandersetzung mit den kognitiven Fähigkeiten, Interessen und Bedürfnissen einerseits und Anforderungen und Rahmenbedingungen des Studienfachs und der Hochschu-

le andererseits kann bereits präventiv Orientierung geben und Rüstzeug für ein aussichtsreiches Studieren sein. Die Aussagefähigkeit erhobener Daten erweist sich als umso aussagekräftiger, je mehr die realen Ansprüche mit den im Eignungsprozess berücksichtigten Beurteilungsdimensionen übereinstimmen. Anforderungsanalysen bilden dahingehend die Basis des dargestellten Dreiphasenmodells der Eignungsfeststellung mit den Dimensionen Verfahrenskonstruktion, -durchführung und -auswertung und somit auch die Grundlage der angestrebten optimalen Passung zwischen Student und Studienfach.

Anhand der in den vorangegangenen Kapiteln dokumentierten Literaturanalyse ist eine optimale und für Organisation und Bewerber effiziente fächerbezogene Studieneignungsempfehlung mittels einer Kombination von Abiturnote und allgemeinen, studienfeldspezifischen Tests sowie Persönlichkeitstests und einem Studienberatungsgespräch abzuleiten und unter Berücksichtigung der jeweiligen institutionellen Begleitumstände bzw. der Hochschul- oder Organisationskultur hin zu überprüfen. Multivariate und moderierte Prädiktion bewirkt eine erhöhte Vorhersagegenauigkeit mit Blick auf Besonderheiten der einzelnen Hochschule bzw. federführenden Institution.

## 5. Veränderte Rahmenbedingungen

Globalisierung und Internationalisierung verändern Politik, Wirtschaft und Gesellschaft und bewirken einen beschleunigten Wandel aller Lebens- und Arbeitsverhältnisse.

Die Veränderungen der sicherheitspolitischen Lage, die gewachsene Verantwortung Deutschlands, seine internationalen Verpflichtungen in den Vereinten Nationen, der Nordatlantischen Allianz und der Europäischen Union haben dazu geführt, dass Deutschland in den vergangenen 15 Jahren zu einem der größten Truppensteller für internationale Friedensmissionen geworden ist. (Weißbuch, 2006, S. 96)

Deutschland benötigt Streitkräfte, die im gesamten Aufgabenspektrum verwendbar sind. Sie müssen für die wahrscheinlicheren Einsätze rasch verfügbar und auf Einsätze höchster Intensität vorbereitet sein. (...). Führungsfähigkeit ist wesentliche Voraussetzung für Informationsüberlegenheit, Führungsüberlegenheit und bestmögliche Wirkung im Einsatz. Effiziente Führungsfähigkeit benötigt ein zuverlässiges und schnelles Informationsmanagement auf und zwischen allen Führungsebe-

nen und allen Truppenteilen der Bundeswehr. Voraussetzung hierfür sind eine straffe Führungsorganisation, klare und einheitliche Führungsverfahren, eine sichere und leistungsfähige Führungsunterstützung sowie die Fähigkeit zur vernetzten Operationsführung. (Weißbuch, 2006, S. 112f).

Gerade der Führungsberuf des Offiziers in seiner qualitativen Komplexität und Vielfalt militärischer Tätigkeitsfelder verlangt eine Vielzahl von Anforderungen. Neben Empathie, Aufgeschlossenheit und Toleranz gegenüber anderen Streitkräften und Kulturen werden Tugenden wie soziale und fremdsprachliche Kompetenz, ökonomisches Denken und Handeln verlangt. Der auch politisch denkende Soldat mit diplomatischen Fähigkeiten definiert das veränderte Leitbild weg vom „soldier-scholar“ der 70er Jahre hin zum „soldier-diplomat“ heute. Die drastisch veränderten Rahmenbedingungen in Politik, Militär, Wissenschaft und Gesellschaft setzen nicht zuletzt auch den UniBw einen zweifachen Reformdruck aus. Zum einen müssen sie als private Hochschulen unter der Kulturhoheit der Bundesländer die Entwicklungen im öffentlichen Hochschulwesen so weit wie möglich mitgehen, damit ihre Anerkennung als qualitativ gleichwertige Bildungseinrichtung neben den Landeshochschulen erhalten bleibt. Zum anderen sind sie als zentrale militärische Dienststellen im Geschäftsbereich des Bundesministers der Verteidigung gezwungen, auf die sicherheits- und haushaltspolitischen Wandlungen adäquat zu reagieren. Nur so werden sie ihrer Aufgabe als akademische Ausbildungseinrichtung für den militärischen Führungsnachwuchs auch weiterhin gerecht werden können. Jedoch haben die UniBw bei der Auswahl ihrer künftigen Studierenden weder Einfluss noch Beteiligungsrechte. Die allgemeine Offiziereignung und somit auch die Studierfähigkeit der angehenden Offiziere werden nicht von den Universitäten selbst sondern im Rahmen der Eignungsfeststellung zum Offizier des Truppendienstes durch die Offizierbewerberprüfzentrale (OPZ) festgestellt. Der Offizierberuf stellt an sich und angesichts des veränderten und gestiegenen Aufgabenspektrums, aufgrund knapper Ressourcen sowie operativer, organisatorischer, verfahrenstechnischer und technologischer Entwicklungen in erhöhtem Maße massive Ansprüche an das Auswahlsystem der OPZ. Der überaus komplexe Prozess der Offizierauswahl wird in den folgenden Abschnitten daher nicht nur auf die Studieneignungsempfehlung bezogen ausführlich beschrieben.

## 6. Die Offizierauswahl an der Offizierbewerberprüfzentrale (OPZ)

Der Auswahlprozess zur Förderung und Beförderung von deutschen Offizieren beginnt mit dem Eignungsfeststellungsverfahren an der Offizierbewerberprüfzentrale (OPZ) in Köln. Dort wird über ein 2 ½ -tägiges AC die grundsätzliche Bewährungswahrscheinlichkeit der Bewerber für die Laufbahn der Offiziere und die Eignung für eine Studienfachrichtung an den UniBw ermittelt (BMVg, 1996). Zur Auswahl kommen zum einen Bewerber aus Schule (meist Abiturienten) oder Beruf, die noch nie Soldat waren. Zum anderen bewerben sich junge Männer und seit Juli 2001 auch Frauen, die bereits Grundwehrdienst leisten oder freiwillig in die Mannschafts- oder Unteroffizier- oder Feldwebellaufbahn eingetreten sind und die Offizieranwärter werden wollen.

Nach Jahren der Stagnation mit meist nicht mehr als 8.000 Bewerbungseingängen pro Jahr kann erstmals ein über 20 Prozent gestiegenes Bewerberaufkommen verzeichnet werden. Knapp 10.000 Abiturienten oder Schüler mit Fachhochschulreife bewarben sich 2005 für eine Einstellung als Offizieranwärter, 12.200 reichten im Jahr 2006 ihre Unterlagen bei der OPZ ein. Aus meist bis zu 4.000 nach dem ACBw als geeignet beurteilten Bewerbern erfolgt eine Bestenauswahl (circa 1.900 Einstellungen).

Aufgabe der OPZ ist es, gemäß § 37 Abs. 1, Nr. 3. Soldatengesetz den Offizierbewerber zu identifizieren, der die hinreichende *charakterliche, geistige und körperliche Eignung* für die angestrebte Laufbahn aufzuweisen vermag. Die OPZ erfüllt diesen Auftrag auf der Grundlage der „Bestimmungen zur Methodik der Eignungsfeststellung bei Offizierbewerbern (BestMeOB)“ (BMVg, 1996) und der „Annahmebestimmungen für Offizierbewerber (AnBestOB)“ (BMVg, 1998).

Von besonderer Tragweite erweisen sich Personalentscheidungen, da Offizieranwärter bereits während der Ausbildung oder nach kurzer beruflicher Tätigkeit als Vorgesetzte von Soldaten und/oder Verwalter von Material verantwortlich eingesetzt werden. Zwangsläufig stellt dies auch hohe Anforderungen an das Auswahlsystem und an darauf abzustimmende Personalentwicklungsmaßnahmen während einer mehrstufigen militärfachlichen Aus- und Weiterbildung sowie später während des in der Regel zu absolvierenden wissenschaftlichen Studiums. Der Offizier ist zugleich *Führer, Ausbilder* und *Erzieher* der ihm unterstellten Soldaten im Frieden und im Einsatz (Inspekteur des Heeres, 1997). Ziel der Offizierausbildung im Einzelnen ist es, den Offizier zu befähigen, seine Untergebenen einsatzorientiert *auszubilden*, sie unter Berücksichtigung der Grundsätze der Inneren Führung zum leistungs-

willigen Soldaten zu *erziehen* und eine Einheit seiner Truppengattung zu *führen*. Die militärfachliche Ausbildung zum Offizier des Heeres findet im Anschluss an die Eignungsfeststellung zum Offizier und vor Beginn der akademischen Ausbildung statt. Mit Umsetzung vieler Neuerungen der Ordnung in der Ausbildung der Streitkräfte verkürzte man im Jahre 2004 die Ausbildungszeit des Heeres von 39 auf einheitliche 15 Monate und passte sie dadurch im Rahmen der Neuordnung der Ausbildung denen der übrigen Teilstreitkräfte an. Anhand des im Kapitel 3.7 vorgestellten Dreiphasenmodells wird in den folgenden Absätzen das Auswahlverfahren an der OPZ, der Assessment Center der Bundeswehr (AC Bw) anhand der Phasen Planung und Durchführung (vgl. Kapitel 3.7) erschöpfend vorgestellt.

## 6.1 Planung

Die Festlegung der Anforderungsmerkmale, die Bereitstellung der Methoden und Regeln sowie das Personalmarketing laufen unter der Federführung des BMVg ab. Tabelle 8 gibt einen Überblick der durch Expertenratings gewonnenen 12 Eignungsmerkmale und ihrer Definition. Sie unterteilen sich in *charakterliche*, *geistige* und *körperliche* Merkmale.

Tabelle 8:  
Die 12 bei der Offizierbewerberprüfzentrale gemessenen Eignungsmerkmale und ihre Definition

<b>Charakterliche Merkmale</b>	<b>Definition</b>
Gewissenhaftigkeit	→ Fähigkeit, in seinem aufgaben- und sozialbezogenen Verhalten zuverlässig und pflichtbewusst zu sein und Bereitschaft, sich in seinen Überlegungen und Handlungen an Werten und Normen zu orientieren.
Führungsfähigkeit, Durchsetzungsverhalten	→ Fähigkeit und Bereitschaft, zielgerichtet und steuernd auf das Verhalten anderer Einfluss zu nehmen.
Soziale Kompetenz, Ko- operation	→ Fähigkeit und Bereitschaft zur Einordnung in Gemeinschaften.
Psychische Belastbarkeit	→ Fähigkeit zu situationsgerechtem Verhalten und Erhaltung von Leistungsvermögen und Entscheidungs- / Handlungskontrolle unter Belastung.
Verhaltensstabilität	→ Grundlegende Orientierung, Angepasstheit, verhaltensbezogene Zuverlässigkeit und Stabilität über die fünf Merkmalsbereiche: Psychische Funktionsstörungen, soziale Anpassungsstörungen, Gewaltneigung, Dogmatismus und Neigung zu Straffälligkeit.
<b>Geistige Merkmale</b>	<b>Definition</b>
Sprachliche Ausdrucksfä- higkeit, Argumentations- verhalten	→ Beherrschung der deutschen Sprache in Wort und Schrift, verständliches und flüssiges Artikulationsvermögen.
Denkvermögen	→ Intellektuelle Anlagen zum regelmäßigen und ursächlichem Erkennen von Zusammenhängen; kognitive Leistungen bei Aufnahme, Strukturierung und Analyse von Informationen.
Urteilsvermögen	→ Erkennen von komplexen Zusammenhängen über die

Lern- und Leistungsbe- reitschaft	→	unmittelbar erlebte Wirklichkeit hinaus und deren logi- sche Analyse aus unterschiedlichen Perspektiven.
Planungs- / Organisati- onsverhalten	→	Erfolgsorientiertes, ausdauerhaftes Angehen von selbst gesetzten Lern- und Leistungszielen.
Berufsvorstellungen	→	Effizienz und Effektivität im Umgang mit komplexen Planungsaufgaben.
	→	Geistige Auseinandersetzung mit den Pflichten und aufgaben eines Offiziers; Vereinbarkeit von Vorstellun- gen über und Erwartungen an militärische Verwendun- gen mit der Realität.
<b>Körperliches Merkmal</b>		<b>Definition</b>
Sportliche Belastbarkeit und Leistungsfähigkeit	→	Körperliche Spannkraft, Anstrengungsbereitschaft, mo- torische Beweglichkeit, Geschicklichkeit, sportlicher Trainingsstand und Trainierbarkeit.

Bei den Beobachtern der neun Prüfgruppen handelt es sich jeweils um einen Psychologen, einen Prüfstaboffizier sowie einen Prüffizier niedrigeren Ranges. Jeder Prüfer erhält eine Einweisung in die Einzelheiten der Durchführung, ca. sechs Wochen „on the job“, nun sich in einer Prüfkommision mit dem methodischen Vorgehen in den einzelnen Prüfstationen und dem Prüfmaßstab vertraut zu machen. Dies entspricht der in der Wirtschaft üblichen *Beobachterschulung*.

## 6.2 Durchführung

In Anhang A wird ein Überblick über den zeitlichen Ablauf des zweieinhalbtätigen Eignungsfeststellungsverfahrens an der OPZ in Köln vermittelt. Der sequentielle Ablauf (vgl. Kapitel 3.2) gestaltet sich aus organisatorischen Gründen für die täglich neun parallel das AC durchlaufenden Prüfgruppen zeitlich versetzt. Der letztendliche Umfang des Verfahrens richtet sich außerdem nach der angestrebten Verwendung und dem Studienwunsch des Bewerbers und kann bei nicht kompensierbaren Mängeln vorzeitig mit der Feststellung der Nichteignung abgebrochen werden.

98 Prozent der Offizierbewerber müssen im Laufe ihrer Dienstzeit ein Studium absolvieren. Daher wird zusätzlich die *Eignung für die gewünschten Studienfächer* festgestellt. Dabei wird beurteilt, mit welcher Wahrscheinlichkeit die Bewerber das Studium bei Einhaltung der vorgeschriebenen Fristen erfolgreich abschließen können. Die Studieneignung wird in Form von Studienfachrichtungsempfehlungen ausgegeben.

### 6.2.1 Bewährungswahrscheinlichkeit und Eignungsgrad

Die Festlegung von Eignungsmerkmalen bildet die Grundlage für die individuelle Beurteilung eines Offizierbewerbers. Der *Ausprägungsgrad je Dimension* ist in einer Notenskala von eins bis sieben zu erfassen, wobei die Note vier in einem Eignungsmerkmal dem Bewerberdurchschnitt entspricht. Kleinere Noten zeigen überdurchschnittliche, größere Noten unterdurchschnittliche Ausprägungsgrade an.

Unterschieden wird ferner zwischen dem über die Annahme oder Ablehnung entscheidenden *Eignungsgrad* (A - besonders geeignet, M - gut geeignet, X - geeignet, U/Y - nicht geeignet) vom durchschnittlichen *Merkmalsindex*, welcher sich aus dem Mittel aller Dimensionsbewertungen auf einer Skala von eins bis sieben errechnet. Der *Eignungsgrad* ist das abschließende, zusammenfassend bewertende Urteil der Kommission über die Eignungsvoraussetzungen des Offizierbewerbers. Der Bewerber wird einer der Kategorien zugeordnet, je nachdem ob seine Bewährung generell als sehr wahrscheinlich (A), als wahrscheinlich (M), als noch wahrscheinlich (X) oder als unwahrscheinlich (U/Y) gewertet wird.

Im Regelfall - aber nicht stets - korrespondiert der Eignungsgrad eines Bewerbers mit dem durchschnittlichen Merkmalsindex. Bei Indexwerten, die unter 4.0 liegen, ist von Eignung auszugehen. Ausnahmen müssen nachvollziehbar begründet werden. Folgende *Ausnahmen* können und *müssen* berücksichtigt werden:

1. Der Bewerber weist nicht kompensierbare Defizite in einem oder mehreren Eignungsmerkmalen auf (in der Regel Beurteilung als „Y“).
2. Der Bewerber hat eine zwar arithmetisch hinreichende Bewährungswahrscheinlichkeit, weist aber eine inakzeptable Persönlichkeitsstruktur auf (in der Regel Beurteilung als „U“).

Es ist zu erwarten, dass sich auch im Indexbereich 4.0 bis 4.9 eine nennenswerte Zahl hinreichend geeigneter Bewerber identifizieren lässt. Die Abbildung 10 illustriert den Zusammenhang von Eignungsmerkmalen, Bewährungswahrscheinlichkeit und Eignungsgrad in der ab dem gewünschten Einstellungstermin (GEWET) 2001 revidierten Prüfmethodik der OPZ. Es wird der hohe Aufwand durch Kombination verschiedener Prüfstationen („multi-situations“) mit verschiedenen Beobachtern („multi raters“) und bewerteten Eignungsmerkmalen pro Prüfstation („multi traits“) deutlich.

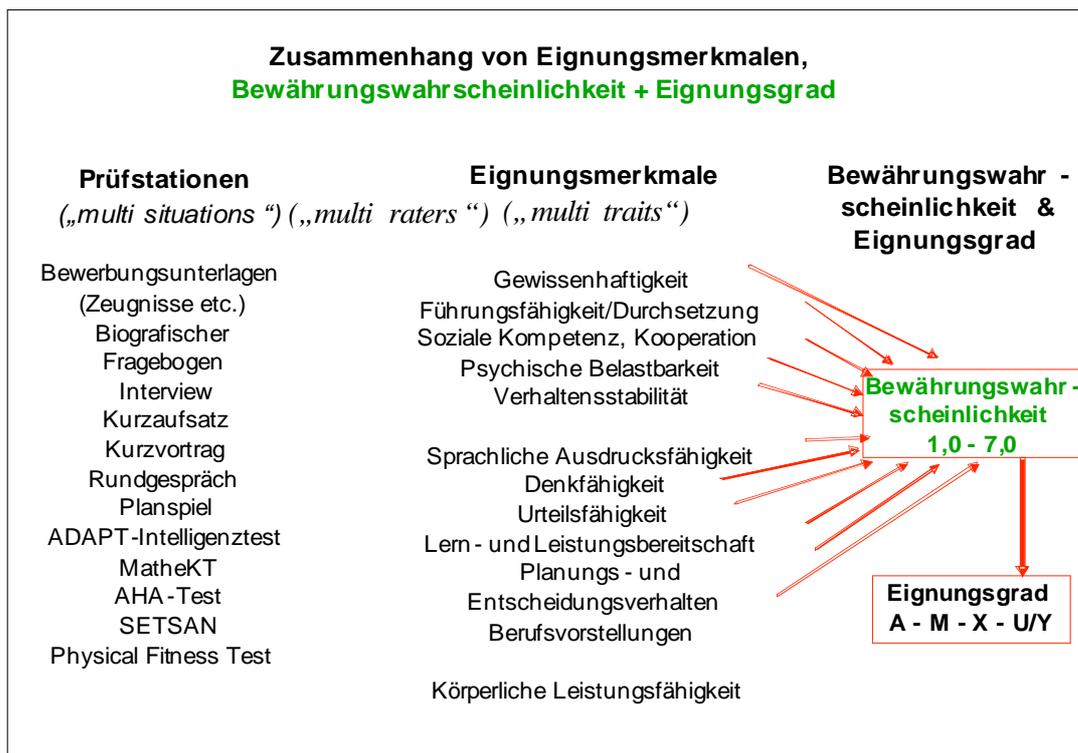


Abbildung 10: Zusammenhang von Eignungsmerkmalen, Bewährungswahrscheinlichkeit und Eignungsgrad im Eignungsfeststellungsverfahren der Offizierberberprüfzentrale

In Abbildung 11 werden außerdem beispielhaft die erwarteten OPZ-Prüfergebnisse bezogen auf die Bundeswehr-Noten-Verteilung veranschaulicht. Wie bereits am Anfang des Kapitels erwähnt, entspricht nicht nur die Note vier dem Bewerberdurchschnitt, also dem Mittelwert der Normalverteilungskurve, sondern bildet in der Regel auch die Obergrenze des Eignungsgrades X. Alle schlechter als der Durchschnitt benoteten Offizierbewerber werden als nicht geeignet bzw. U oder Y angesehen.

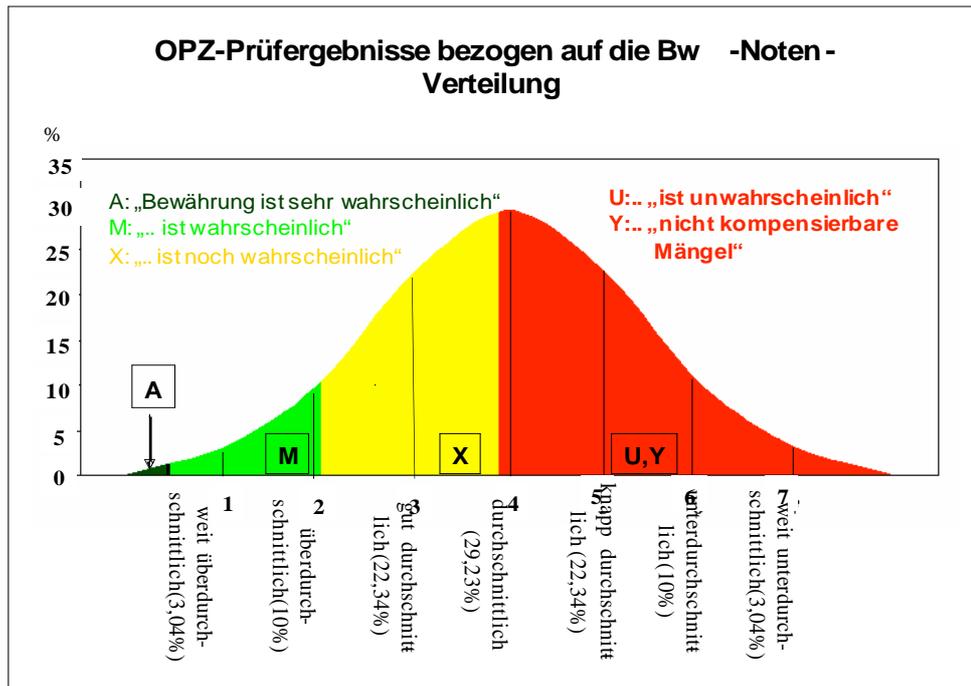


Abbildung 11: Erwartete Prüfergebnisse bezogen auf die Bw-Noten-Verteilung

In der aus Prüfstaboffizier, Diplompsychologe und Prüffoffizier besetzten Prüfgruppe werden sämtliche bewerbenspezifischen Daten gesammelt, geordnet und verdichtet. Offiziere und Psychologen agieren hier im Sinne der ganzheitlichen Erfassung der Bewerberpersönlichkeit. Militärische und persönlichkeitsbestimmende Aspekte werden gleichermaßen berücksichtigt. Jones (1991) betont die Wichtigkeit der Rolle des Psychologen bei der Offizierauswahl insbesondere bezüglich der Einhaltung der Verfahrens-Fairness und politischen Bedeutung:

... the trend has also been to increase the fairness of selection procedures, since selection criteria have been made more explicit and more valid procedures are now in place. The psychologists influence has therefore had a social, and in its broadest sense, political impact (S.110).

Zudem trägt das Kommissionsprinzip insgesamt zur Verbesserung der Testgütekriterien Reliabilität, Validität und Objektivität bei. In der schwedischen Armee wurden aus diesen Gründen zum Beispiel 1994 erstmals nach zwanzig Jahren wieder Psychologen in das Auswahlverfahren für den Offiziernachwuchs integriert (Widen & Carlstedt, 1996).

Die individuellen Merkmalsausprägungen werden in der Prüfgruppe im Sinne eines *dialogischen Prinzips* zusammengeführt. Das bedeutet, jeder Beobachter gibt pro Übung und Di-

mension seine individuelle Bewertung ab, aus der am Ende der Merkmalsindex und die Durchschnittsnote errechnet wird. Im Rahmen einer ausführlichen Kommissionsberatung wird dann im *Konsensverfahren* der Eignungsgrad abgegeben. Im Kommissionsprotokoll findet diese Entscheidung ihren Niederschlag. Individuelle Bewerberbesonderheiten und Einplanungshinweise sind hier außerdem zu erfassen.

### 6.2.2 Studienfachrichtungsempfehlung

Der *Studienberater* gibt im Rahmen eines Einführungsvortrags einen Einblick in die Studiengänge und -fachrichtungen an den Bundeswehruniversitäten und berät den Offizierbewerber im Einzelgespräch über die Studienwünsche. Darüber hinaus soll er prüfend feststellen, ob der Bewerber die nötige Kompetenz, Kenntnis und Motivation für den gewünschten Studiengang vorweist. Die Beurteilung der Studieneignung für das vom Offizierbewerber mit Erst-, Zweit- oder Drittwunsch gewählte Studienfach an einer der beiden Universitäten oder der Fachhochschule der Bundeswehr wird Studienfachrichtungsempfehlung genannt. Festgestellt wird diese in den Entscheidungskommissionen der OPZ unter Federführung des Psychologen und Mitwirkung der Kommissionsmitglieder und des Studienberaters ausschließlich für die zum Offizier als mindestens „geeignet“ bewerteten Offizierbewerber. Das Studienberatungsgespräch ist Teil der eng gestrickten Ablauforganisation der OPZ und dauert zwischen zehn und zwanzig Minuten.

Zur Bestimmung der nötigen Kompetenzen zieht der Studienberater Eingangstestleistungen heran und stellt dem Bewerber studienfachbezogene Wissensfragen. Während bei geisteswissenschaftlichen Fächern vorwiegend die Motivation und die Vorkenntnisse des einzelnen Bewerbers geprüft werden, findet in der Studienberatung für technische Studiengänge eine detaillierte Abfrage fachspezifischen Wissens statt. Konkreter gesprochen werden bei der Bewertung der Erfolgswahrscheinlichkeit in den gewünschten Studienfächern aus dem arithmetischen Mittel der Einzelbewertungen alle für das jeweilige Studienfach bedeutsamen Einzelmerkmale errechnet. Bisher einbezogene Eignungsmerkmale<sup>4</sup> zur Feststellung der fachspezifischen Studieneignung sind:

- Leistungs- und Lernbereitschaft (Kommissionsbewertung aus den Gruppensituationsverfahren im AC Bw auf einer 7er Skala)
- Psychische Belastbarkeit (Kommissionsbewertung AC Bw auf einer 7er Skala)
- Mathematikkenntnisse (Bundeswehrtestnote aus dem mathematischen Kenntnistest)

---

<sup>4</sup> Die für die empirische Teststichprobe nicht relevante und auch nur schwach abgewandelte und im Qualitätssicherungsprozess befindliche Form der Studienberatung seit 2006 findet der Übersicht halber an dieser Stelle keine Berücksichtigung.

- Mathematische Fähigkeiten (Bundeswehrtestnote aus dem adaptiven Rechentest)
- Fachspezifische Kenntnisse und Interessen (Kommissionsbewertung im AC Bw auf einer 7er Skala)
- Einschlägige Fach-/Notenkombinationen der gymnasialen Oberstufe (Kommissionsbewertung im AC Bw auf einer 7er Skala)

Die Bewertung eines Bewerbers in einem dieser Studieneignungsmerkmale resultiert aus dem arithmetischen Mittel der Bewertungen aller Prüfer und des Studienberaters in allen für das Studieneignungsmerkmal einschlägigen Stationen (z.B. Interview, Test der Arbeitshaltungen, Studienberatung, Schulzeugnis).

Abschließend spricht der Studienberater eine Empfehlung aus. Die Erfolgswahrscheinlichkeit wird in drei Stufen bewertet und bei Hochschulstudiengängen durch die arabischen Ziffern 1 (empfohlen), 2 (mit Einschränkung empfohlen), 3 (nicht empfohlen) ausgedrückt, bei Fachhochschulstudiengängen entsprechend durch die Großbuchstaben A, B, C. Es liegt in der Kompetenz des Diplom-Psychologen, die vom Studienberater ausgesprochene Empfehlung zu widerrufen, wenn dieser im Interview einen gegenteiligen Eindruck vom Bewerber erhält. Eine Ausnahme bildet die Studienempfehlung für Sanitätsbewerber. Hier ist das Urteil des medizinischen Studienberaters verbindlich und unanfechtbar.

Tabelle 9 liefert einen genauen inhaltlichen Überblick der Gesamtstudienbewertung.

Tabelle 9:

*Studieneingangsbeurteilungsergebnisse der Offizierbewerberprüfzentrale im Überblick*

1/A	Studienfach empfohlen. Der Bewerber bringt fast in jedem Studieneignungsmerkmal günstige Voraussetzungen mit.
2/B	Studienfach mit Einschränkung empfohlen. Dem Bewerber fehlt es teilweise an den erforderlichen Wissensgrundlagen, oder andere Voraussetzungen geben zu Bedenken Anlass.
3/C	Studienfach nicht empfohlen. Der Studienerfolg des Bewerbers erscheint als zweifelhaft.

Die Abstimmungsnotwendigkeit von Studienwünschen mit dem geforderten Bedarf der Bundeswehr führt in vielen Fällen zu einer Nichtübereinstimmung zwischen Erststudienwunsch und tatsächlichem Studienfach (interne Information Personalamt der Bundeswehr).

Einen Überblick über alle beschriebenen Beurteilungen im AC Bw liefert Tabelle 10 .

Tabelle 10:  
*Überblick über Beurteilungen der Offizierbewerber im Assessment Center*

<b>Bewertung pro Eignungsmerkmal pro Übung</b>		
Ausprägungsgrad je Merkmal	Note	
Weit überdurchschnittlich	= 1	
Überdurchschnittlich	= 2	
Gut durchschnittlich	= 3	
Durchschnittlich	= 4	
Knapp durchschnittlich	= 5	
Unterdurchschnittlich	= 6	
<b>Beurteilung nach abschließender Beobachterkonferenz - „Overall-AC-Rating“</b>		
Bewährungswahrscheinlichkeit	Definition	Eignungsgrad
... sehr wahrscheinlich	Besonders geeignet	= A
... wahrscheinlich	Gut geeignet	= M
... noch wahrscheinlich	Geeignet	= X
... unwahrscheinlich	Nicht geeignet	= U/Y
<b>Beurteilung der Studieneignung</b>		
Studienempfehlung	Eignungsgrad Universität / Fachhochschule	
Empfohlen	1 / A	
Mit Einschränkung empfohlen	2 / B	
Nicht empfohlen	3 / C	

Der Einfluss und die unmittelbare Relevanz von für den Offiziereignungsgrad wichtigen Prädiktoren auf die Studienfachrichtungsempfehlung macht eine ausführliche Betrachtung der Prüfstationen und der zu messenden Dimensionen notwendig.

Trotz der insbesondere bei den männlichen Bewerbern (circa 80 Prozent insgesamt, 85 Prozent Truppendienst, 40 Prozent Sanitätsdienst) zu verzeichnenden Präferenz für ingenieurwissenschaftliche Studiengänge lässt sich der hohe Bedarf der Streitkräfte in diesen Bereichen nur schwer decken. „Nach durchgeführter Prüfung der Studieneignung muss so mancher Wunsch den Realitäten angepasst werden“ (BMVg, 2006, S. 25).

### 6.2.3 Prüfstationen und zu messende Dimensionen

Die Eignung wird auf der Grundlage von Informationen beurteilt, die im Verlauf der persönlichen Vorstellung des Offizierbewerbers bei der OPZ durch schriftliche und mündliche Befragung, Verhaltensbeobachtung und Leistungsproben gewonnen werden. Die aktuellen,

unmittelbar erhobenen Daten werden ergänzt durch biographische und sonstige personenbezogene Informationen aus der Personalakte. Die Prüfungskommission gewinnt aus folgenden Angaben ihren Eindruck über die einzelnen Bewerber<sup>5</sup>. Aus *unmittelbaren Eindrücken* von der Persönlichkeit des Bewerbers durch *Verhaltensbeobachtungen*

- im Interview
- beim Kurzvortrag
- bei der Gruppendiskussion (Rundgespräch) und
- beim Planspiel.

Aus ergänzenden unmittelbaren Eindrücken aus den schriftlichen Äußerungen des Offizierbewerbers, d.h. aus

- Lebenslauf,
- Biographischem Fragebogen (BF),
- Begriffsaufsatz und
- Konzept für den Kurzvortrag.

Aus *vermittelten Leistungsdaten* in Form von Testergebnissen und Messwerten aus dem

- Adaptiven Intelligenztest (ADAPT),
- Test Arbeitshaltungen (AHA),
- Mathematischen Kenntnistest (MathKT) und
- Studieneignungstest für Sanitätsoffizierbewerber (SETSAN),
- Physical-Fitness-Test (PFT).

Außer auf Primärdaten, die von der Kommission erhoben oder zumindest selbst ausgewertet werden, stützt sich die Kommission auf *Sekundärdaten*, d.h. auf solche, die das Ergebnis einer Interpretation und Verdichtung konkreter Informationen und Eindrücke durch andere sind, d.h. auf

- Empfehlungen der Laufbahnberater (Personalberater),
- Stellungnahme der Studienberater zum erwarteten Studienerfolg und
- medizinische Tauglichkeitsgrade und Verwendungsausschlüsse.

Ferner stützt die Prüfungskommission ihren Eindruck mittels *eignungsrelevanter Sekundärdaten auf externe Datenquellen*, z.B. auf

- Schulzeugnisse,
- Ausbildungsnachweise und
- (Laufbahn-)Beurteilungen bereits in der Bundeswehr dienender Bewerber.

<sup>5</sup> Die hier farblich hervorgehobenen Prädiktoren bezieht der Studienberater mit in seine Studieneignungsempfehlung ein.

Diese Daten bedürfen einer Erörterung im Vorstellungsgespräch und einer besonderen Interpretation, da ihre Entstehungsbedingungen in der Regel nicht hinreichend bekannt sind. In Tabelle 11 werden alle in den jeweiligen Übungen gemessenen Dimensionen graphisch dargestellt.

Tabelle 11:

*Erfassung der Anforderungen in den jeweiligen Übungen im AC Bw. Die grauen Kästchen stellen die gemessenen Dimensionen dar. BF: Biographischer Fragebogen, Psych.: Psychische, Sprache: Sprachliche Ausdrucksfähigkeit und Argumentationsvermögen, MathKT: Mathematischer Kenntnistest, Schwarz: In die Studienberatung einfließende Ergebnisse*

	Inter- view	Plan- spiel	Rund- gespräch	Kurz- vortrag	Aufsatz	AHA- Test	ADAPT- Test	BF	MathKT
Gewissenhaftigkeit									
Führungsfähigkeit									
Soziale Kompetenz									
Sprache									
Urteilsvermögen									
Lernbereitschaft									
Psychische Belastbarkeit									
Denkfähigkeit									
Berufsvorstellungen									
Planungsvermögen									
Verhaltensstabilität									
Mathematische Kenntnisse									

Die große Menge an Prüfgruppen und die zeitliche Begrenztheit lassen eine konkrete Reihenfolge der Übungen nicht explizit festlegen. Die einzelnen Prädiktorklassen bzw. AC-Stationen werden, beginnend mit dem Interview, in den nächsten Kapiteln ausführlich skizziert.

Ziel des *Interviews* ist aus Sicht der Prüfungskommission, den Bewerber in einem etwa halbstündigen Gespräch über die Aktenlage hinaus persönlich kennen zu lernen, um so ein möglichst abgerundetes Bild vom Bewerber zu erhalten. Das Interview gliedert sich, entsprechend der Zusammensetzung der Prüfungskommission, in drei Teile. Themen des ersten Teils, den einer der beiden Offiziere einleitet, sind das aktuelle Befinden, die Einschätzung der eigenen Leistungen seitens des Bewerbers, (zivil-) berufliche Alternativen und die Gründe für die Wahl des Offizierberufs. Im zweiten Teil stellt der Diplom-Psychologe, aufbauend auf den Informationen, die er aus dem Lebenslauf und dem biographischen Fragebogen ge-

wonnen hat, Fragen zur Familie, zum schulischen Werdegang und zum Freizeitverhalten. Darüber erhält er Aufschluss über relevante prägnante Persönlichkeitseigenschaften und Motive des Bewerbers. Der abschließende Teil des Interviews dient der Erfragung laufbahnbezogener Kenntnisse und Vorstellungen. Wichtige Informationen liefert das Interview über alle im AC Bw zu beurteilenden und in Tabelle 11 aufgeführten elf Eignungsmerkmale.

Das *Gruppensituationsverfahren* findet in einer Kleingruppe von drei bis vier Bewerbern statt und umfasst die Teilaufgaben Planspiel, Rundgespräch und Kurzvortrag. Die Prüfungskommission beobachtet und bewertet die Interaktion dieser Kleingruppe und erhält so Hinweise auf die Führungsfähigkeit, die soziale Kompetenz, die Ausdrucksfähigkeit im gesprochenen Wort, die psychische Belastbarkeit und das Planungs- und Entscheidungsverhalten jedes einzelnen Bewerbers. Die Gruppe ist selbst gesteuert, d.h. im Vorfeld wird kein Gesprächsleiter eingeteilt, und die Organisation der Aufgabenbewältigung bleibt der Gruppe selbst überlassen. Dadurch wird jeder Bewerber gleichsam motiviert, seine Ideen und Argumente in die Runde einzubringen, erste Führungsqualitäten können sich abzeichnen. Der Einsatz der Gruppensituationsverfahren ist zur Bewertung der Offiziereignung unverzichtbar, da sie wertvolle Informationen über das Verhalten des Bewerbers in der sozialen Interaktion und sein Verhalten bei der konkreten Aufgabenbewältigung liefern. Ziel des *Planspiels* ist die gemeinsame Bearbeitung einer konkreten Aufgabe mit Projektcharakter in der Gruppe. Beim *Rundgespräch* einigen sich die Bewerber auf ein Thema aus einer vorgegebenen Liste und beginnen eine Gruppendiskussion. Die Bewerber werden dazu angehalten, aufgrund der begrenzten Zeit ihre schlagkräftigsten Argumente direkt zu Beginn anzubringen. Die Beurteiler achten besonders auf die Kooperationsfähigkeit dominanter Bewerber und ihre Akzeptanz innerhalb der Gruppe. Aufgabe des Bewerbers beim *Kurzvortrag* ist es, aus einer vorgegebenen Liste von Themen eines auszuwählen und innerhalb von 30 Minuten einen Vortrag vorzubereiten. Die Themeninhalte stammen aus dem gesellschaftspolitischen Bereich und weisen unterschiedliche Schwierigkeitsgrade auf. Für den *Begriffsaufsatz* erhält der Bewerber zwei Themen zur Auswahl, von denen er eines zur Bearbeitung auswählt. Jedes Thema enthält zwei sinnverwandte Begriffe wie zum Beispiel „Soldat und Söldner“, deren Definition, Gemeinsamkeiten und Unterschiede in einem Aufsatz dargestellt werden sollen. Für diese Aufgabe erhält der Bewerber 30 Minuten Bearbeitungszeit. Der Begriffsaufsatz soll Aufschluss über die sprachliche Ausdrucksfähigkeit und die Urteilsfähigkeit des Bewerbers geben.

Die geistige Eignung wird in erster Linie auf der Grundlage *psychologischer Fähigkeits- und Kenntnistests* festgestellt. Diese werden computergestützt unter Einsatz von CAT-Anlagen erhoben (CAT = Computer-Assisted-Testing). Die individuelle Leistung des Offizierbewerbers wird anhand von statistisch errechneten Normen bestimmt, die aus den Offizierbewerbern des vorhergehenden Jahrgangs resultieren. Die Normierung richtet sich nach der Transformationsformel  $Standardwert = 4 - \frac{4}{3}z$  mit Eins als bester, Sieben als schlechtester Testnote und Vier als mittleren Wert der Skala. Das *adaptive Testen* geht nicht wie das konventionelle Testen auf die Theoreme der klassischen Testtheorie zurück, sondern auf ein probabilistisches Testmodell und ist auf das Fähigkeitsniveau jedes einzelnen Probanden zugeschnitten (Fisseni, 1997). Dies macht ein differenziertes und genaues Messen im gesamten Leistungsbereich der Bewerber möglich. Die Auswertung erfolgt demnach nicht gemäß den Testrohwerten. Das adaptive Testen resultiert aus einer Fähigkeitsschätzung (Theta-Wert), deren Wertebereich sich von -3 (geringste Fähigkeit) über 0 (mittlere Fähigkeit) bis zu +3 (höchste Fähigkeit) erstreckt (Erdfelder & Rudinger et al., 1996). Bei diesem adaptiven Intelligenztest soll der Bewerber seine geistigen Fähigkeiten unter Beweis stellen. Gemessen wird sowohl die sprachliche als auch die nicht-sprachliche formal-logische Intelligenz. Die drei computerunterstützten adaptiven Fähigkeitstests (ADAPT) wurden im Rahmen eines Forschungsvertrages des BMVg mit Prof. Dr. Lutz F. Hornke von der Rheinisch-Westfälischen Technischen Hochschule Aachen entwickelt. Sie dienen zur Prüfung der intellektuellen Leistungsfähigkeit auf Basis der drei Teilfähigkeiten. Die ursprünglichen Fähigkeitsschätzungen (Theta) werden zwecks besserer Vergleichbarkeit mit anderen Testresultaten in Standardwerte überführt mit Note eins für „weit überdurchschnittlich“, sieben für „weit unterdurchschnittliche“ und vier für „durchschnittliche“ Leistung.

Das computerunterstützte psychologische *Testverfahren Arbeitshaltungen (AHA)* von Kubinger und Ebenhöf (1996) diente bis 2001<sup>6</sup> der Prüfung der Eignungsmerkmale Lern- und Leistungsbereitschaft und psychische Belastbarkeit (sensu Frustrationstoleranz). Im ersten Untertest „Flächengrößen Vergleichen“ muss eine von zwei gleichzeitig dargebotenen schwer zu unterscheidenden Flächen als größer erkannt werden. Impulsive Offizierbewerber lassen sich zu einer eher schnellen und fehlerhaften Entscheidung verleiten. Im zweiten Untertest „Symbole Kodieren“ muss die Testperson in fünf Etappen nach vorgegebenem Kodierschlüssel zu fortlaufend in unsystematischer Weise am Bildschirm gebotenen - vier ver-

---

<sup>6</sup> Die Testbatterie Arbeitshaltungen wurde aus Kosten-Nutzen-Gesichtspunkten nach 2001 nicht mehr eingesetzt.

schiedenen - abstrakten schwarzweißen Figuren das zugehörige Symbol (eine ebenfalls abstrakte, bunte Figur) zuordnen. Ab der zweiten Etappe wird der Bewerber dazu aufgefordert, eine Vorhersage über seine in der nächsten Etappe bearbeitete Anzahl an Kodierungen abzugeben. Nach der zweiten Etappe wird der Bewerber frustriert durch die Fehlinformation, andere Bewerber seien um 10% besser. Der dritte Untertest „Figuren unterscheiden“ besteht aus theoretisch unendlich vielen Items, d.h. die Testvorgabe erfolgt so lange, bis die Testperson den Test abbricht. Pro Bildschirmseite werden sechs Items, bestehend aus vier einfachen geometrischen Figuren geboten, wovon sich je eine Figur gegenüber den übrigen dreien unterscheidet und identifiziert werden soll. Eine erste Rückmeldung erfolgt nach 170 bearbeiteten Items. Dem Bewerber wird über eine Rückmeldung das immer noch anhaltende Arbeiten anderer Personen suggeriert. Nach 425 Aufgaben wird der Person nahegelegt, bei nachlassender Konzentration aufzuhören, trotz eines angeblich immer noch arbeitenden anderen Bewerbers. „Sie verlieren die Konzentration“ lautet die dritte Rückmeldung nach weiteren 150 Aufgaben oder verlangsamter Arbeitszeit. Beendet der Bewerber den Test nicht eigenständig oder erreicht er nicht die maximal erzielbare Bundeswehrtestnote, ist der Testabbruch durch den Testleiter vorzunehmen.

Der *mathematische Kenntnistest (MathKT)*, basierend auf dem „Mathematiktest für Abiturienten und Studienanfänger“ von Lienert (1969), soll die Offizierbewerber hinsichtlich ihrer Kenntnisse aus den mathematischen Bereichen Algebra, Geometrie sowie Analysis überprüfen und damit *Entscheidungshilfen für die Vergabe von Studienfachrichtungsempfehlungen* liefern. Die Beantwortung der einzelnen Aufgaben erfolgt nach dem Multiple-Choice-Prinzip mit jeweils drei Antwortalternativen. Die Sanitätsbewerber nehmen am Mathematiktest nicht teil und bearbeiten stattdessen einen *Sanitätstest*, der dem Subtest "Mathematisch-naturwissenschaftliches Grundverständnis aus dem ehemaligen "Test für medizinische Studiengänge" (TMS) der ZVS entspricht.

Der *Biographische Fragebogen* fordert den Offizierbewerber zu wertenden Stellungnahmen hinsichtlich seiner bisherigen Entwicklung und zur Selbstdarstellung auf. Konkret wird er zum Beispiel nach der Zufriedenheit mit den eigenen schulischen Leistungen, dem außerschulischen/-beruflichen Freizeitverhalten, den Anregungen aus der Familie, den eigenen charakterlichen Stärken und Schwächen, den Gründen für die Bewerbung und den mit dem Offizierberuf verknüpften Erwartungen gefragt. Somit liefert der biographische Fragebogen wichtige Hinweise über eignungsrelevante Einstellungen, Werthaltungen und Motive und

dient somit auch als Grundlage für die Vorbereitung auf das Interview seitens der Prüfkommision.

Aus dem *Physical-Fitness-Test* erhält die Kommission standardisierte Hinweise auf die allgemeine sportliche Leistungsfähigkeit und körperliche Belastbarkeit der Offizierbewerber. Zu den Übungen gehören der Pendellauf, Situps, der Standweitsprung, Liegestütze sowie der 12-Minuten-Lauf. Der Test und seine Benotung orientieren sich an Übungen und geschlechtsspezifischen Bewertungen des Deutschen Sportabzeichens.

Die Festlegung der Wünsche des Offizierbewerbers hinsichtlich Teilstreitkraft, Verwendung, Status und Studium, seine Zufriedenheit mit der Verwendungszuweisung und damit letztlich auch Ausbildungserfolg und Bewährung hängen stark von der Güte des *Beratungs- und Abschlussgesprächs* ab. Dieses soll dem Offizierbewerber zu eigenständigen Entscheidungen auf der Basis solider Informationen verhelfen. Der *Personalberater* betreut den Offizierbewerber während der gesamten Dauer der Eignungsfeststellung. Er befragt ihn zu seinen formalen Voraussetzungen, Vorstellungen und Wünschen hinsichtlich des militärischen Ausbildungsgangs und zeigt Möglichkeiten und Grenzen ihrer Realisierung auf. Er informiert die Kommissionen über die Ergebnisse seiner Beratung und über dabei gewonnene Erkenntnisse. Nachdem alle Übungen durchlaufen sind, eröffnet der *Prüfgruppenleiter* dem Offizierbewerber in einem Abschlussgespräch das Ergebnis der Eignungsfeststellung, begründet es in allgemeiner Form und erläutert dessen Bedeutung. Beratungs- und Abschlussgespräche erfüllen den sechsten Grundsatz des Arbeitskreises Assessment Center, nachdem individuelle Einsatzmöglichkeiten gesucht werden, die im Bereich der persönlichen Möglichkeiten der Gesamtperson des Bewerbers liegen.

## 7. Das Studium an den Universitäten der Bundeswehr

„Die Bildung des Offiziers verfeinert das Militär nach und nach.“ (Gerhard von Scharnhorst). Jenes vom preußischen Heeresreformer formulierte Ideal ist Richtungspunkt der Offizierausbildung der Bundeswehr. Mit einem progressiven Bildungsansatz - einer universitären Ausbildung innerhalb des allgemeinen Ausbildungsgangs zum Offizier - wird dieser Grundforderung Scharnhorsts am ehesten Rechnung getragen. Der Offizier soll nicht nur fähig und bereit sein, durch sein Führungsverhalten Maßstäbe zu setzen, sondern durch Bildung, Kenntnisse und Wissen dieses gleichsam tragfähig machen. Professionelles, militärisches Können und geistige Weite werden nicht nur aufgrund der immer fortschreitenden Technisierung und Verwissenschaftlichung von Teilfunktionen in den Streitkräften gefordert. Der

Offizier von heute und morgen hat sich darüber hinaus in seiner Führungsfähigkeit neben anderen qualifizierten Berufen zu behaupten, die in der Regel ein akademisches Studium voraussetzen. Außerdem soll der Offizierberuf im Falle der Zeitsoldaten den Einstieg in eine ebenso qualifizierte Anschlussverwendung ermöglichen, welche in der Regel jedenfalls einen akademischen Abschluss voraussetzt.

Die Bundeswehr muss mit der gesellschaftlichen Entwicklung Schritt halten. Sie muss ihre Berufsbilder, ihre Bildungs- und Ausbildungsgänge so gestalten, dass die Soldaten daraus für ihren beruflichen Werdegang innerhalb der Bundeswehr und ebenso später im zivilen Leben den größtmöglichen Nutzen ziehen. Das dient gleichzeitig der Effektivität der Bundeswehr und der Attraktivität der soldatischen Laufbahnen. Ausbildungsgänge und Abschlüsse im gesamten militärischen Bereich sollen so weit wie möglich denen des zivilen Bereichs entsprechen (Bundesministerium der Verteidigung, Weissbuch 1970 zur Sicherheit der Bundesrepublik Deutschland und zur Lage der Bundeswehr, S. 130).

Seit 1973 nehmen jährlich circa 1400 Offizieranwärter ein mindestens dreijähriges Studium an den Universitäten der Bundeswehr als obligatorischen Bestandteil und zeitlich umfangreichsten Abschnitt der Ausbildung zum länger dienenden Offizier mit mindestens 12-jähriger Verpflichtung auf. Eine Ernennung zum Berufssoldaten ist erst nach erfolgreich absolviertem Studium und entsprechenden Lehrgangsbeurteilungen und Vorgesetztenempfehlungen möglich. Beide Universitäten verfügen zusammen über circa 5000 Studienplätze, von denen in etwa 600 dem angegliederten Fachhochschulbereich in München zuzuordnen sind. Wie in Tabelle 12 ersichtlich, werden derzeit 18 universitäre und drei Fachhochschulstudiengänge angeboten.

Tabelle 12:  
*Übersicht der an den beiden Universitäten und einer Fachhochschule der Bundeswehr zu studierende Studienfächer*

	<b>UniBw Hamburg</b>	<b>UniBw München</b>	<b>Fachhochschule der Bw München</b>
<b>GeiWi</b>	Pädagogik		
	Geschichtswissenschaft	Staats- und Sozialwissenschaft	
	Politikwissenschaft	Sportwissenschaft	
<b>WiWi</b>	Betriebswirtschaftslehre	Wirtschafts- und Organisationswissenschaft	Betriebswirtschaftslehre
<b>IngWI</b>	Volkswirtschaftslehre	Wirtschaftsinformatik	Elektrotechnik & technische Informatik

Wirtschaftsingenieurwesen	Geodäsie & Geoinformation	Maschinenbau
Elektrotechnik	Elektrotechnik & Informationstechnik	
Maschinenbau	Bauingenieurwesen & Umwelttechnik	
Rechnergestützte Ingenieurwissenschaften (Computational Engineering)	Luft- & Raumfahrttechnik	
<b>Informatik</b>	Informatik	

*Anmerkungen:* GeiWi: Geisteswissenschaften; WiWi: Wirtschaftswissenschaften; IngWi: Ingenieurwissenschaften.

An den Bundeswehruniversitäten studieren im Trimesterstudienrhythmus bis heute ausschließlich Offizieranwärter und Offiziere der Teilstreitkräfte Heer, Luftwaffe und Marine<sup>7</sup>. Die UniBw in Hamburg und München galten lange Zeit als Reformuniversitäten, nahmen sie doch zu Beginn der 70er Jahre - in einer Zeit des Aufschwungs der Bildungs- und Hochschulpolitik - ihren Lehr- und Studienbetrieb auf. Die Reformuniversität verfolgt im Sinne der Inneren Führung den Wunsch, die Sozialfigur des Staatsbürgers in Uniform an den gesellschaftlichen Wandel - als Reaktion auf die „Bildungskatastrophe“ der 60er Jahre und die daraus folgende größere Studierwilligkeit - anzupassen. Ziel war und ist neben der Attraktivitätssteigerung eine dem Zivilen gleichwertige und zivil anerkannte akademische Ausbildung als Professionalisierung des Offizierberufs. Bis heute zählen die beiden staatlich anerkannten Bundeswehruniversitäten zu den wichtigen deutschen wissenschaftlichen Zentren in Forschung und Lehre und fungieren gleichsam als Brücken zwischen Militär und Gesellschaft (Fleckenstein, 2002). Das kennzeichnende Konzept beinhaltet damals wie heute die Durchsetzung einer sanktionsbewerteten Regelstudienzeit, die Gliederung der Studienzeit in drei Studienjahre mit je drei Trimestern (plus ein Trimester Diplomarbeitszeit), die Ausrichtung der Studien- und Prüfungsordnungen an die allgemein geltenden Rahmenordnungen, die Einführung von fachübergreifenden Studienelementen sowie die Schaffung der erforderlichen Rahmenbedingungen zur Durchsetzung der Regelstudienzeit. Hierzu zählen die Gewährleistung des Lebensunterhalts sowie die Ausstattung in Form von personellen, räumlichen und sonstigen Ressourcen zu zählen. Studierende an den UniBw sind nicht auf einen Nebenverdienst angewiesen. Sie bekommen ihr Gehalt (abhängig von Alter, Dienstgrad und Familienstand) auch während der Studienzeit. Studierende verbringen demzufolge keinerlei Zeit mit der Suche nach einer Studentenwohnung und leben stattdessen in komfortablen Wohnheimen direkt auf dem Campus oder – im Falle Hamburgs - im nahe gelegenen Wohn-

<sup>7</sup> Ausnahme bildet eine hier zahlenmäßig zu vernachlässigende Menge an Gaststudenten privater Wirtschaftskonzerne und hochbegabter Schüler.

bereich der Universität. Auf einer Wohnebene leben normalerweise nur Studenten der gleichen Fachrichtung, aber aus unterschiedlichen Jahrgängen. Das ermöglicht die Unterstützung und Betreuung jüngerer Studenten durch ältere Kommilitonen. Auf dem Campus befinden sich neben den Einrichtungen für Forschung und Lehre auch die Bibliotheken, die Mensa und die Sportanlagen. Ärzte und Zahnärzte, eine evangelische und eine katholische Gemeinde, selbst einen Friseursalon gibt es hier. Das sorgt für kurze Wege spart viel Zeit für wichtigere Dinge. Das Zahlenverhältnis Lehrende zu Studenten ist an ein Kleingruppenkonzept angelehnt, welches sich mit Ausnahme der Vorlesungen an einer Gruppengröße von nicht mehr als 25 Teilnehmern orientiert. Obligatorisch zu belegende erziehungs-, gesellschafts-, technik- und wirtschaftswissenschaftliche Anteile (EGTWA) dienen der inhaltlichen Ergänzung der jeweiligen Fachstudiengänge und zielen auf fächerübergreifende Kompetenzen künftiger militärischer wie ziviler Tätigkeitsfelder. Eine strukturelle Besonderheit der Universitäten der Bundeswehr stellt der Studentenbereich dar, der durch einen Offizier im Rang eines Obersten geführt wird. Die Angehörigen des Studentenbereichs unterstützen die Studierenden in allen militärischen Fragen betreffenden dienstlichen Angelegenheiten wie Urlaub, Versetzung nach dem Studium, Auslandsaufenthalte. Denn auch als Studenten behalten Offizieranwärter und Offiziere der Bundeswehr ihren rechtlichen Status als Soldaten. Das nur zum Herbsttrimester beginnende Studienjahr besteht aus drei Trimestern und insgesamt drei Monaten lehrveranstaltungsfreie Zeit, in der jedoch vorgeschriebene Praktika abgeleistet werden müssen. Die durchschnittliche Studiendauer beträgt Dreieinhalbjahre. Eingebettet in einen Zeit- und Kostenrahmen und als Bestandteil der Offizierausbildung soll das Studium an den UniBw in einer Regelstudienzeit von dreieinviertel Jahren erreicht werden. Diese Zielsetzung ist gleichsam bindend für den Träger, die Universität und die studierenden Offizieranwärter und Offiziere. Eine Überschreitung der Regelstudienzeit um mehr als vier Monate führt bei einem erfolgreichen Studium zu einer Dienstzeitverlängerung um ein Jahr. Eine Verlängerung über vier Jahre hinaus ist nur in begründeten Ausnahmefällen möglich und bedarf der Genehmigung durch den Führungsstab der UniBw.

Noch vor Festlegung von Sanktionsmaßnahmen im Hochschulrahmengesetz von 1976 bei Verletzung der Regelstudienzeit hatten die Gründer der UniBw ihr Studienkonzept bereits auf der Basis einer Studienzeitbegrenzung entwickelt. Ausschlaggebend für eine Studienzeitreduzierung ist das Streben nach einem tragbaren Verhältnis zwischen Ausbildungszeit und praktischer Tätigkeit. Mit Ausnahme der angehenden Berufssoldaten im Offizierstatus verlässt auch heute noch das Gros der Offiziere die Bundeswehr nach 12 Jahren Zeitsoldatentä-

tigkeit. Unter Hinzunahme der dem Studium vorangehenden je nach Teilstreitkraft 15 bis 36 Monate andauernden militärischen Ausbildung ergibt sich eine Ausbildungszeit von bis zu sechs Jahren und drei Monaten. Das Studium ist der zeitlich umfangreichste Abschnitt der Ausbildung zum Offizier des Truppendienstes. Neben seiner gesellschaftspolitischen Bedeutung ist auch seine individuelle Relevanz für eine zukünftige zivilberufliche Karriere oder Berufssoldatenlaufbahn von großer Bedeutung.

## 7.1 Studienabbruch an den Universitäten der Bundeswehr

In Tabelle 13 werden die Abbruchquoten öffentlicher Hochschulen (vgl. Kapitel 3.5) gemäß HIS (2005) mit denen der UniBw gemäß BMVg (2005; 2006b) verglichen. Trotz bester Rahmenbedingungen haben die UniBw mit den Landesuniversitäten vergleichbare Abbruchquoten zu beklagen. Die Fachhochschule der Bundeswehr weist mit einer durchschnittlichen Quote von 34 Prozent sogar hochsignifikant schlechtere Zahlen auf als öffentliche Fachhochschulen mit circa 20 Prozent.

Tabelle 13:

*Studienabbruchquoten für deutsche Studierende an Universitäten und Fachhochschulen nach an den UniBw vorhandenen Fächergruppen und ausgewählten Studienbereichen in Prozent (BMVg, 2006; HIS-Studienabbruchuntersuchung, 2005, S. 19)*

Studienfeld	Studienfach	Studienabbrecher in Prozent			
		UniBw Hamburg 2000	UniBw (FHS) München 2000	Öffentliche Unis (FHS) 1999	Öffentliche Unis (FHS) 2002
Geisteswissenschaften, Pädagogik, Sport	Pädagogik	<b>32</b>	21	28	23
	Sportwissenschaften	-	5	-	-
	Geschichte	23	-	-	-
	Geisteswissenschaften gesamt Uni	28	13	28	23
	Staats- und Sozialwissenschaften	-	14	42 (SoWi) 27 (Jura)	36 (SoWi) 16 (Jura)
Wirtschafts- und Sozialwissenschaften	Politikwissenschaften	17	-	-	-
	BWL	<b>31</b>	-	-	-
	VWL	<b>34</b>	-	-	-
	Wirtschaftsingenieurwesen	25	-	-	-
	WOW	-	17	-	-
	Wirtschaftswissenschaften gesamt Uni (ohne FHS)	26	17	<b>31</b>	<b>32</b>
	BWL (FH)	-	<b>46</b>	25 (WiWi)	27 (WiWi)
eur wi §- Maschinenbau	15	-	25	<b>34</b>	

	Elektrotechnik	24	22	23	<b>33</b>
	Bauingenieurwesen	-	18	<b>35</b>	<b>30</b>
	Luft- und Raumfahrttechnik	-	<b>43</b>	-	-
	Ingenieurwissenschaften gesamt (ohne FHS)	20	28	28	<b>32</b>
	Elektrotechnik (FH)	-	32	20	<b>32</b>
	Maschinenbau (FH)	-	21	25	21
Informa- tik	Wirtschaftsinformatik	-	<b>40</b>	-	-
	Informatik	-	<b>38</b>	<b>37</b>	<b>38</b>
GESAMT Uni (FHS)		25 (-)	26 (34)	24 (20)	26 (22)

*Anmerkungen:* SoWi: Sozialwissenschaften; Jura: Rechtswissenschaften; BWL: Betriebswirtschaftslehre; VWL: Volkswirtschaftslehre; WOW: Wirtschafts- und Organisationswissenschaften; FHS: Fachhochschule.

Jedoch mögen unterschiedlichste Erfolgs- und Abbruchquoten nicht zwingend mit den in der Literaturanalyse beschriebenen möglichen Ursachen im Zusammenhang stehen. Nach Landeck (1988) erweist sich ein direkter Vergleich öffentlicher mit Bundeswehrhochschulen als fraglich. Abbrecher der UniBw sind gleichzeitig auch Ausscheidende aus der Organisation Bundeswehr und aus ihrer Offizierlaufbahn. Mit Ablauf der festgesetzten Zwischendienstzeit (Heer und Marine sechs Jahre, Luftwaffe vier Jahre) ist die Beendigung des Dienstverhältnisses verbunden. „Wer die Ausbildung zum Offizier nicht schafft, soll nicht über Gebühr in der Bundeswehr festgehalten werden“ (Fleckenstein, 2002, S. 144). Auf der 15. Kommandeur-Tagung des zentralen Militärdienstes der Bundeswehr (1998) wurde zur Führung, Ausbildung und Erziehung der studierenden Offiziere an den UniBw das Studium und seine Auslesefunktion herausgestellt, „die in dieser Schärfe an keiner Stelle des Werdegangs vorher zum Tragen kommt“ (S. 5). Überdies ist ein häufig in den ersten Semestern vollzogener Studienfachwechsel bei den Streitkräften nur in Ausnahmefällen und für wenige Offizieranwärter möglich. Das eingeschränkte Fächerangebot erhöht außerdem das Drop-Out-Risiko studierender Offiziere. Bei der häufig stattfindenden Überschreitung der ohnehin straff gegliederten und für technische Studienfachrichtungen sehr knapp bemessenen Regelstudienzeit von 39 Monaten erhöht sich der Leistungs- und Erfolgsdruck und mündet nicht selten in Misserfolg oder Abbruch. Oben beschriebene Abbruchmodelle weisen bereits auf die gerade für Campusuniversitäten besonders relevanten Aspekte erfolgreichen Studierens hin.

Im Vergleich zu den Landesuniversitäten ist das zentral Andere des Studiums der UniBw in der Beibehaltung des Soldatenstatus, in der fortlaufenden Besoldung, im Campusleben und anderer Annehmlichkeiten wie im fortdauernden Erhalt beamtenrechtlicher Privilegien zu sehen. Vor dem Hintergrund hoher Abbruchquoten lässt sich allerdings in Annäherung sagen, dass sich dieser Privilegiertenstatus weder in direkter noch nachvollziehbarer Weise auf den Studienerfolg auszuwirken scheint bzw. von noch einflussreicheren Voraussetzungen oder Moderatorvariablen überlagert wird. Bei einer Studiendauer von im Schnitt Dreieinhalb Jahren erreichen nur knapp zwei Drittel der Studenten das Ziel. Das übrige Drittel scheitert meist schon am Vordiplom. Damit liegen die Abbruchquoten fächerübergreifend ungefähr in Höhe derer öffentlicher Hochschulen (vgl. Untersuchung Krex, 2005 / Kapitel 11.3). Es kann üblicherweise nicht nachvollzogen werden, inwiefern und wie erfolgreich ein Studium im Anschluss außerhalb der Bundeswehr in Angriff genommen wird. Eine erschöpfende Post-Hoc-Untersuchung von Abbruchgründen an den UniBw ist daher und – bzw. dies implizierend – aus Datenschutzgründen nahezu unmöglich.

## 7.2 Vergleich zivile versus militärische Universitäten in Deutschland

Der bestehende Kampf öffentlicher und privater Hochschulen um die besten Studenten erhält nicht nur als Folge des neuen Hochschulrahmengesetzes mit Ziel forcierter Profilbildung und Leistungsdifferenzierung zusätzliche Dynamik, sondern gleichsam durch die in den letzten Jahren zunehmende Zahl privater Hochschulen. Mit ihrem nach außen getragenen individuellen Leitbild im Sinne optimierter akademischer Ausbildung und Förderung eigener Eliten schüren sie den Wettbewerb in einem zunehmend internationalisierten Hochschulraum. Somit nähert man sich einem in den USA vorgelebten Konkurrenzdruck der führenden Universitäten um die besten Studienbewerber und einer generell restriktiveren Eingangsselektion. Auch wenn das neue HRG öffentlichen Hochschulen ermöglicht, verstärkt ihre Studenten selbst auswählen zu dürfen, leisten sich bisher vor allem private Hochschulen wie die Uni Witten-Herdecke (WHU) oder Bucerius Law School in Hamburg aufwändige Auswahlverfahren.

Die Studenten der Bundeswehrhochschulen unterscheiden sich von denen öffentlicher oder übriger privater Hochschulen durch ihre Quasi-Doppelberufswahl bzw. das in den militärischen Kontext der Offizierausbildung integrierte akademische Studium. Primärfunktion des erfolgreichen Studiums ist nicht ein ziviler Berufseinstieg sondern Voraussetzung für das Erreichen des Offizierstatus. Während des Studiums behält der Offizieranwärter bzw. Offi-

zier seinen Soldatenstatus und somit sein Recht auf volles Gehalt, Fürsorge wie Wohnmöglichkeit auf dem Campus und weitere beschriebene Annehmlichkeiten. Der Student erhält seine Studieneignungsempfehlung erst nach bestandenem zweieinhalbtätigem AC und erlangter Offiziereignungsbewertung. Im Vergleich zu Landeshochschulen werden die Weichen für eine Fachrichtung eines Offizierbewerbers bereits bis zu über drei Jahre vor dem eigentlichen Studienbeginn - in der Offizierbewerberprüfzentrale und somit vor Beginn der Offizierausbildung - gestellt. Die größte und schwerwiegendste Besonderheit ist die Verdichtung der sonst üblichen Semester der Landesuniversitäten auf das Studium verdichtende Trimester in Verbindung mit Curricula. Die Trimester werden ohne Pufferzeit bis zum Ende der Regelstudienzeit durchlaufen und sind zusätzlich mit einer dichten Abfolge der zu erbringenden Leistungsnachweise (Zulassungsscheine, Vordiplomprüfung, Diplomarbeit und Diplomhauptprüfung) versehen. Während der vorlesungsfreien Zeit leistet der Student und Offizieranwärter bzw. Offizier Wehrdienst und Praktika bzw. hat seinen Jahresurlaub zu nehmen. Das Zeitbudget der studierenden Offiziere lässt im Trimester sowie in der vorlesungsfreien Zeit wenige Reserven zur Vor- oder Nachbereitung von Stoff oder Prüfungen. Pflichtwochenstunden zwischen 40 und 60, Lernen für Prüfungen parallel zum laufenden Stoff, eventuell auch noch zur Diplomarbeit, Pflichtpraktika, Sprachausbildung, Pflichtsport und militärisches Begleitprogramm müssen bewältigt werden. Im Unterschied zum gewohnten vorgegebenen Alltag in Gruppe oder Truppschule fordert der Beginn des Studiums eine erhebliche Anpassungsleistung. Das charakteristische Andere am Studium gegenüber der Phase der Offizierausbildung vor Studienbeginn besteht in der eigenverantwortlichen Selbstbestimmung des Handelns“ (S. II).

Weitere massive Umstände verschärfen die ohnehin hohen Anforderungen an den militärischen Universitäten. Aufgrund der zu erfüllenden Bedarfsdeckung über alle Studienfächer und mögliche Offizierverwendungen hinweg ist dem bei der Bundeswehr Studierenden weder Studienfach noch eine spätere studienfachnahe Verwendung garantiert. Im Vergleich zu öffentlichen Hochschulen sind die Prüfungswiederholungsmöglichkeiten eingeschränkt. Bonnemann & Posner (2002) stellen im Rahmen einer repräsentativen Längs- und Querschnittsstudie zum Bereich Lehre und Studium ferner auffällige soziodemographische und andere Besonderheiten des militärischen versus zivilen Studenten heraus. Die Untersuchungen bestätigen ein soziales Rekrutierungsmuster des mittelständisch-aufstiegsorientierten familiären Bildungsmilieus. Alleine doppelt so viele Studenten aus niedrigen sozialen Herkunftsgruppen studieren an den UniBw wie an den Landesuniversitäten. Dabei vermuten die

Autoren neben dem Wunsch nach sozialem Aufstieg über das Erreichen des Offizierstatus oder des akademischen Abschlusses gleichsam im Beamtenstatus ein Sicherheitsgefühl und geregeltes berufliches Weiterkommen. Trotz Großstadtangebot am Studienort ist die Zahl der so genannten Wochenendheimfahrer mit fächerübergreifend 40 Prozent auffallend hoch. „Die Einschätzung der Lehrsituation fällt zugunsten der UniBw mit ihrem guten Aufbau der Studiengänge, ihrer inhaltlichen Qualität des Lehrangebots, ihrer intensiveren Betreuung durch die Lehrenden und ihrer räumlichen und materialen Ausstattung.

Möglicherweise ist die ungewöhnliche Zweigleisigkeit der UniBw - ihr Eingebundensein sowohl in zivil-akademischen und militärischen Kontext - ausschlaggebend dafür, dass dieses in der deutschen Hochschullandschaft einmalige und funktionstüchtige, vergleichsweise aber andersartige Studienmodell nach wie vor von der Bildungspolitik im Hochschulbereich oder regelmäßig stattfindenden Hochschulrankings nur selten oder gar nicht zur Kenntnis genommen wird.

### 7.3 Zusammenfassung Offizierauswahl und -studium

Die Bundeswehr hat insbesondere nach Übernahme zusätzlicher Aufgaben in einem multinationalen Umfeld an gesellschaftlicher Bedeutung gewonnen. Dabei ist die Qualität seiner Führungskräfte und somit deren optimale und auf neuartige Anforderungen ausgelegte Ausbildung der Offiziere Bedingung für Informations- und Führungsüberlegenheit und bestmögliche Wirkung im gesamten Aufgabenspektrum der Bundeswehr. Deutschland verfügt dahingehend über Streitkräfte, deren Offiziere zusätzlich zu ihrer militärischen Ausbildung eine akademische Qualifikation erhalten; die damit verbundene berufliche Doppelqualifizierung ist in dieser Form weltweit einzigartig.

Ziel dieses Kapitels war die differenzierte Betrachtung des Hochschulstudiums des Führungsnachwuchses der Streitkräfte an den UniBw und des vorangestellten Offizierauswahlverfahrens der deutschen Streitkräfte an der OPZ hinsichtlich der Offizier- und Studieneignung.

Im Zentrum der eignungsdiagnostischen Arbeit der OPZ steht neben der Prüfung der allgemeinen Offiziereignung die Empfehlung der Studieneignung. Zielgedanke ist es, genau denjenigen Offizierbewerber zu identifizieren, der die hinreichende charakterliche, geistige und körperliche Eignung für die angestrebte Laufbahn aufweist. Die OPZ erfüllt diesen Auftrag auf der Grundlage der BestMeOB und der AnBestOB.

Die von der OPZ verwendeten 12 Eignungsmerkmale zur Bestimmung der Offiziereignung entsprechen bis auf wenige Ausnahmen (wie Verhaltensstabilität) denen eines typischen Führungsberufs. Sie lassen sich in *charakterliche*, *geistige* und *körperliche* Merkmale gliedern. Methodische Gefahrenquelle stellt die nicht präzise definierbare Anforderungsstruktur *des* Offizierberufs dar. Bei allen Offizierbewerbern wird teilstreitkraftübergreifend die allgemeine Eignung zum Offizier geprüft. Sie findet ihr abschließendes Urteil im nach der Beobachterkonferenz ermittelten Eignungsgrad A (besonders geeignet), M (gut geeignet), X (geeignet) und U / Y (nicht geeignet), welcher sich wiederum an den je Übung bewerteten Dimensionsnoten bzw. dem insgesamt berechneten Merkmalsindex orientiert.

Neben dieser Einschätzung über den späteren Ausbildungs- und Berufserfolg wird die Einschätzung über den späteren Studienerfolg für die gewünschten Studienfächer abgegeben. Der Studienberater liefert dem Bewerber einen Einblick in die Studienfachrichtungen an den UniBw und berät jeden Offizieranwärter im zehn- bis zwanzigminütigen Einzelberatungsgespräch hinsichtlich seiner Studienwünsche. Neben der Qualität der Antworten aus von ihm gestellten Wissensfragen subsumiert der Studienberater Einzelbewertungen in allen für das jeweilige Studienfach bedeutsamen Einzelmerkmalen (Leistungs- und Lernbereitschaft, psychische Belastbarkeit, Mathematikkenntnisse, fachspezifische Kenntnisse und Interessen und einschlägige Noten-/Fächerkombinationen in der Oberstufe) seinem Urteil. Abschließend spricht der Studienberater eine Empfehlung aus, die zwischen eins und drei für Hochschul- und A und C für Fachhochschulstudiengänge angesiedelt ist.

In den 1973 in Hamburg und München eingerichteten UniBw absolviert das Gros des Offizier Nachwuchses ein Studium, das auf freier Forschung und Lehre gründet. Die UniBw sind eine vom jeweiligen Bundesland als gleichwertige Hochschule anerkannte nicht staatliche Einrichtung des Bildungswesens im Geschäftsbereich des Bundesministers der Verteidigung und zugleich Dienststelle und Mittelbehörde. Sie sprengen damit die gemeinhin vorhandenen Vorstellungen einer akademischen Bildungs- und Ausbildungsstätte. Was Anfang der siebziger Jahre als durchaus umstrittenes hochschulpolitisches Großexperiment begann, ist heute bewährte Normalität. Die UniBw sind nicht nur anerkannter Teil des Ausbildungssystems in den Streitkräften geworden. Beide Unis sind fest in die nationale und internationale Universitätslandschaft eingebunden. Karriereanalysen von ausscheidenden studierenden Offizieren und Forschungsaufträge an den Instituten belegen die Spitzenstellung, die die UniBw in der deutschen Hochschullandschaft genießen. Das Studium an den UniBw soll den akademisch gebildeten Offizier für das Berufsfeld Streitkräfte wie für einen Zivilberuf qualifizieren.

Kurze Studienzeiten in Trimestern, Kleingruppen-Konzept, Unterbringung auf dem Campus und Studium Generale kennzeichnen das Studium der Offiziere an den UniBw. EGTW-Anteile sind neben der wissenschaftlichen Fachkompetenz als Basisqualifikation für kompetente Wahrnehmung der Bürger- und Berufsrolle innerhalb und außerhalb der Bundeswehr anzusehen. Durch das auf Trimester verdichtete Studium ist die Belastung des Studenten höher als an öffentlichen Universitäten oder Fachhochschulen. Der Studienbeginn und der damit verbundene eigenverantwortlich zu organisierende Universitätsalltag unterscheiden sich deutlich von dem zuvor erlebten durchstrukturierten Truppenalltag während der Ausbildung zum Truppenoffizier. Als Universität gilt die Bundeswehr-Hochschule als zu militärisch als Militärakademie als zu zivil. Der Student im Status Soldat und in militärische Strukturen eingebunden symbolisiert gleichzeitig den Offizier im Status immatrikulierter Student unter der Freiheit von Forschung, Lehre und Studium dar.

Trotz dieser zahlreichen und die Bundeswehruniversitäten als privilegiert darstellenden Unterschiede reagieren die Landesuniversitäten bereits jetzt auf das neu geordnete Hochschulrahmengesetz mit verstärkter Profilbildung, eingeführter Bachelor-/Master-Studiengänge oder durch den Einsatz studienfeldspezifischer Eignungstests. Die akademische Ausbildung deutscher Offizieranwärter orientiert sich nur zögerlich an den neuen bildungspolitischen Herausforderungen. Ein Schritt in diese Richtung soll die vorliegende Arbeit repräsentieren. In Anbetracht der zu hohen Anzahl von Studiengangwechslern, der hohen Studienabbrecherquoten, verbunden mit einem zwangsläufigen und kostspieligen Ausscheiden aus der Bundeswehr, ist eine Evaluation und Optimierung der derzeitigen Auswahl fernerhin unumgänglich! Ein erfolgreiches Studium an den UniBw ist nur möglich, wenn die ausgezeichneten Rahmenbedingungen wie Kleingruppenkonzept, Campusuniversität, Beratung und Betreuung ergänzt werden durch charakteristische Fähigkeiten und eigene Schlüsselkompetenzen auf Seiten der Studierenden. Die Studienberater der OPZ, die diese Fähigkeiten beurteilen, arbeiten in einem engen Zeitraster und haben oft nur Minuten Zeit für den einzelnen Bewerber. Eine umfassende Beurteilung aller geforderten Leistungskriterien erscheint daher kaum möglich oder sogar fraglich ist. Da es im Verlauf des Prüfverfahrens jedoch kein psychometrisches Verfahren gibt, welches die akademische Befähigung explizit bzw. die von der OPZ selbst an Studenten gestellten Anforderungen konkret misst, muss davon ausgegangen werden, dass Studierfähigkeit im Rahmen des ACBw nicht festgestellt wird. Es werden dort zwar eine Reihe aussagefähiger Tests durchgeführt, die andere bedeutende Fähigkeiten der Bewerber messen. Das derzeit verwendete Verfahren zur Feststellung

der Studierfähigkeit besteht jedoch genau genommen auf der subjektiven Einschätzung der Studierfähigkeit eines Bewerbers durch einen auf diesem Gebiet zwar erfahrenen Studienberater, der hingegen nicht durch ein den Gütekriterien von Tests standhaltendes Instrument unterstützt wird.

Die Messung charakteristischer Fähigkeiten und eigener Schlüsselkompetenzen auf Seiten der Studierenden ist es mehr als wert, hier im Rahmen der vorliegenden Dissertation untersucht zu werden, und wird im folgenden Abschnitt empirisch untermauert.

## C EMPIRISCHER TEIL

Ziel der vorliegenden Inauguraldissertation ist - im Rahmen der personalpsychologischen Qualitätssicherung der Streitkräfte - die Optimierung der Studienberatung im Eignungsfeststellungsverfahren der OPZ mittels einer studienfachspezifischen Reihungstestbatterie. Zu diesem Zweck wird die Zusammenstellung und Entwicklung eines Testverfahrens mit Modellcharakter, dessen Evaluation und eine nähere Betrachtung moderierender Variablen und nicht-kognitiver Einflussfaktoren durchgeführt. Die Folgerungen erheben den Anspruch, eine langfristige Orientierungs- und Entscheidungshilfe bei Einführung neuer und Veränderung bestehender Eignungsfeststellungsverfahren für Offiziere in der Bundeswehr zu liefern. Der empirische Teil beschreibt detailliert die Projekthintergründe und praktischen wie theoretischen Ansatzpunkte der Arbeit. In dem Zusammenhang werden die vorangegangenen Anforderungsanalysen skizziert, untersucht, kommentiert, und die Testbatterien der beiden Pilotphasen an den UniBw sowie vorangehende Entwicklungsarbeiten und Inhalte der Tests auf Subtest- und Itemebene definiert. In den daran angeschlossenen Kapiteln 10.1 bis 10.5 sollen die Forschungsfragestellungen, Grob- und Feinziele und die jeweils verwendeten Hypothesen und statistischen Vorgehensweisen konkretisiert, sowie die untersuchte Stichprobe und die relevanten abhängigen und unabhängigen Variablen der Untersuchung vorgestellt werden. Zu diesem Zweck werden die Verteilungen der Prädiktoren, Kriterien und Moderatorvariablen dargelegt und mittels detaillierter Itemanalyse (Kapitel 11.1) auf die in Kapitel 11.1 bis 11.4 folgende Annahmen- und Hypothesenüberprüfung hingeleitet. Die Ergebnisse unterstützend und den empirischen Prozess evaluierend werden ferner die Instrumente biographischer Fragebogen, der zur Evaluationszwecken eingesetzte Akzeptanzfragebogen sowie das Feedbackgespräch der Pilotphasen näher erläutert und hinsichtlich der Forschungsfragestellungen ausgewertet (Kapitel 11.5). An die Gesamtbewertung und Ergebniszusammenfassung schließt sich der Diskussionsteil in Kapitel 12 an, der alle Resultate integriert und diese in bisherige Forschungsergebnisse einordnet. Abschließend sollen in Kapitel 13 bedeutende Implikationen für Praxis und Forschung abgeleitet und ein Ausblick auf weiterführende Fragestellungen geliefert werden.

In Anlehnung an das in dem Kapitel 3.7 erstellte Modell zur Eignungsfeststellung sollen im Folgenden die ersten zwei der drei Phasen Planung, Durchführung und Evaluation auf die zu optimierende Studierendenauswahl der OPZ mittels des Projekts Studieneignung hinführen und den empirischen Teil einleiten.

## 8. Planung

Die Grundlage der Eignungsfeststellung von Bewerbern bildet neben der Zielfeststellung und exakten wie umfassenden Definition der Anforderungen die Festlegung von Methoden und Regeln sowie das an die beiden Pilotphasen voran geschaltete Personalmarketing zur Werbung ausschließlich freiwilliger Testpersonen.

### 8.1 Ziele und Nutzen

Im Rahmen des Projektes „studienfeldbezogene Eignungstests in der Bundeswehr“<sup>8</sup> soll langfristig eine Weiterentwicklung und Ergänzung der bisherigen Studieneignungsmethoden an der OPZ hin zu einer testtheoretisch abgesicherten und inhaltlich umfassenden Testbatterie mit Modellcharakter für die Studienberatung ermöglicht werden. Übergeordnetes Ziel des Projektes bzw. der vorliegenden Arbeit ist eine optimierte Auswahl der für ein Studium geeigneten Offizieranwärter und die Lieferung eines Beitrags zur Verminderung der Abbruchquoten an den Hochschulen der Bundeswehr. Durch den Einsatz von optimierten Auswahlverfahren wird eine größere Passung zwischen Studierenden und den von ihnen gewählten Studiengängen bzw. Hochschulen im Sinne der Person-Environment-Fit-Ansätze (vgl. Kapitel 3.6) angestrebt. Unzweifelhaft spielen dabei auch ökonomische Gesichtspunkte eine wichtige Rolle. Eine verbesserte Passung zieht zahlreiche Vorteile mit sich, die in Abbildung 12 veranschaulicht werden. Aufgrund hoher Ausbildungskosten sollen mit einer gezielten Auswahl motivierter und fähiger Studierender die Studienzeiten verkürzt und die Abbruchquoten gesenkt werden. Umgekehrt werden transparente und nachvollziehbare Auswahlverfahren auch von den Studierenden akzeptiert und geschätzt; zumal jene weitere Möglichkeiten der Zulassung zu gewünschten Studiengängen neben der Abiturdurchschnittsnote und der Wartezeit eröffnen. Es kann demnach erwartet werden, dass sich als Folge der Einführung neuer studienfeldbezogener Eignungstests sowohl die Studienmotivation erhöhen, der Studienerfolg verbessern als auch die Studiendauer verkürzen wird.

Darüber hinaus ergibt sich ein weit gefächerter Nutzen nicht nur für die Bundeswehr allgemein, sondern auch für die Offizierbewerber und für die Universitäten selbst. Somit erwachsen den Streitkräften verringerte Ausbildungskosten durch kürzere Studiendauer und weniger Studienabbrecher. Der Modellcharakter bzw. Transfer der Ergebnisse auf andere Hoch-

---

<sup>8</sup> Im Folgenden kurz „Projekt Studieneignung“ genannt.

schulen bzw. Fachbereiche im Sinne einer so genannten Open-Source-Lösung, der zu erwartende Imagegewinn der Personalpsychologie der Bundeswehr nach innen und außen, die seit Projektbeginn intensivierte Beteiligung an einschlägigen Fachkongressen und Tagungen zur Vorstellung und Diskussion der Forschungsergebnisse und nicht zuletzt die Senkung der Ausbildungskosten pro Absolvent durch Verkürzung der Studienzeiten nutzen ebenfalls der Bundeswehr insgesamt. Für die Offizierbewerber ist der Mehrwert in Form einer verbesserten Hilfestellung bei der Studienfachwahl, einer gezielteren Vorbereitung auf den Einstieg in das Studium in Form eines verbesserten individuellen Studienerfolgs und einer Steigerung der Zufriedenheit und Motivation zu erwarten. Die Universitäten der Bundeswehr können gleichsam mit einer Erhöhung des Anteils von sehr gut geeigneten Studenten und somit mit insgesamt verbessertem Studienerfolg, mit einer Reduktion der Abbruchquoten, mit verkürzten Studienzeiten, mit einem beträchtlichen Imagegewinn aufgrund zufriedenerer und motivierterer Studierender sowie der Zunahme von Arbeitszufriedenheit und -motivation des Lehrkörpers rechnen.

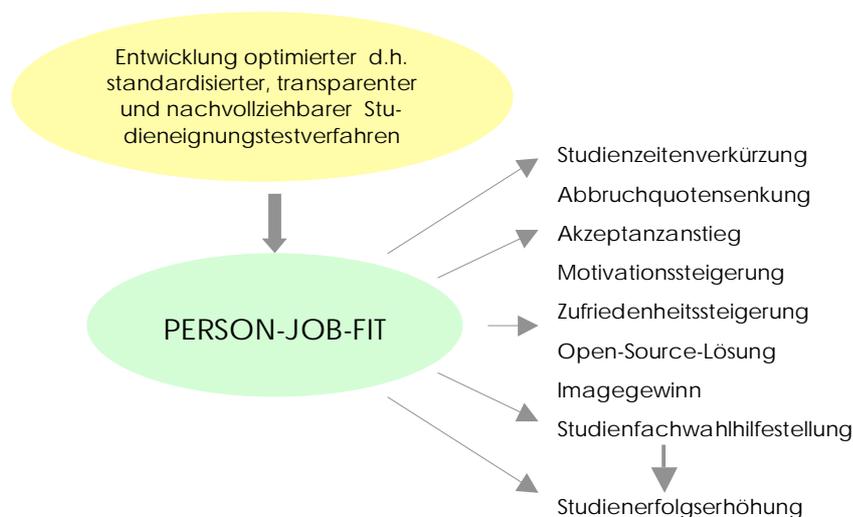


Abbildung 12: Nutzen und positive Effekte optimierter standardisierter Studieneignungstestbatterien

Das zuständige Fachreferat Personal-, Sozial- und Zentralangelegenheiten III 6 im BMVg gibt für die Zusammenstellung der Testbatterie eine Orientierung an den Kompetenzanforderungen zu den Studiengängen laut Abschlussbericht von Kastner (2004) vor. Aufgrund dieser höchst aufwendig und kostenintensiv vorangeschalteten Anforderungsanalyse aller Studienfachrichtungen der UniBw ist im Rahmen der vorliegenden Arbeit keine weitere und den Pilotphasen vorgeschaltete Studienfachanalyse vorgesehen.

Die erfolgreiche und zügige Abwicklung des Projektes erfordert einen detaillierten Zeitplan, der dem besseren Überblick halber in den unten stehenden Grafiken skizziert ist.

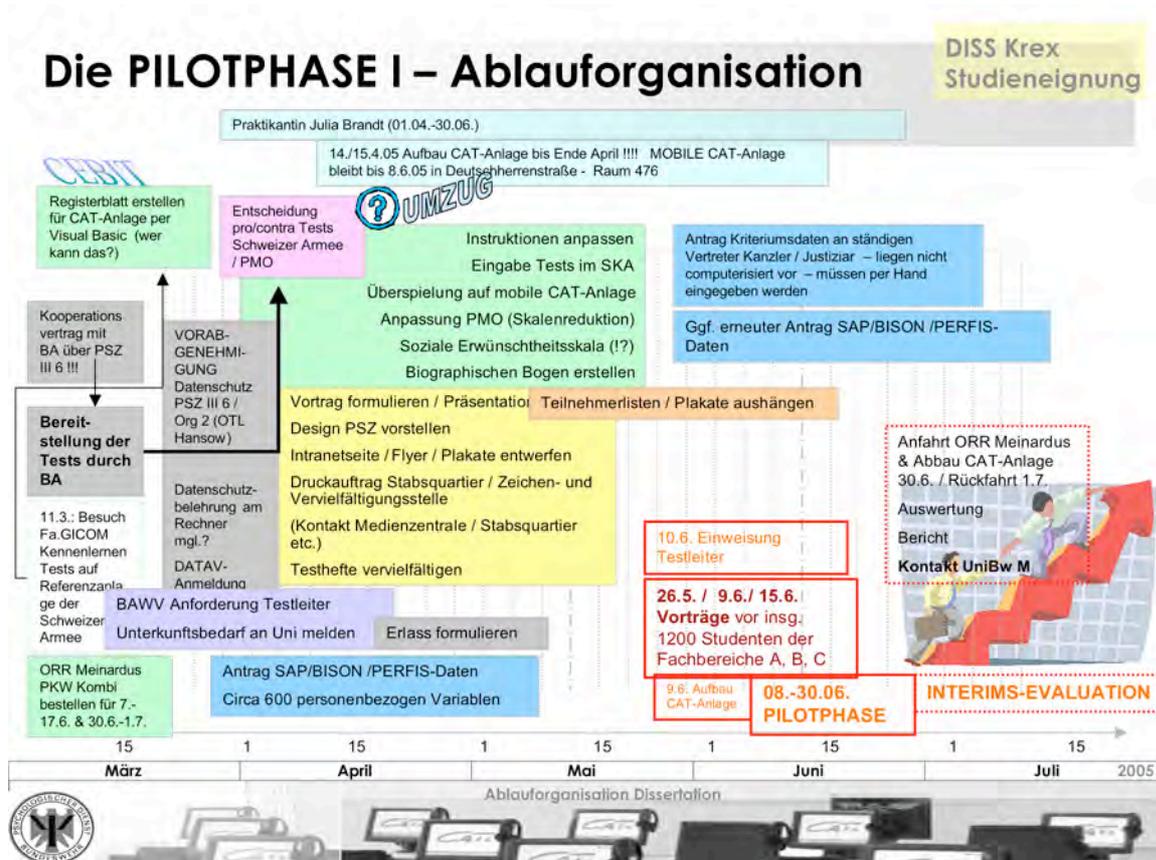


Abbildung 13: Ablauforganisation Pilotphase I (Universität der Bundeswehr Hamburg) Projekt Studieneignung

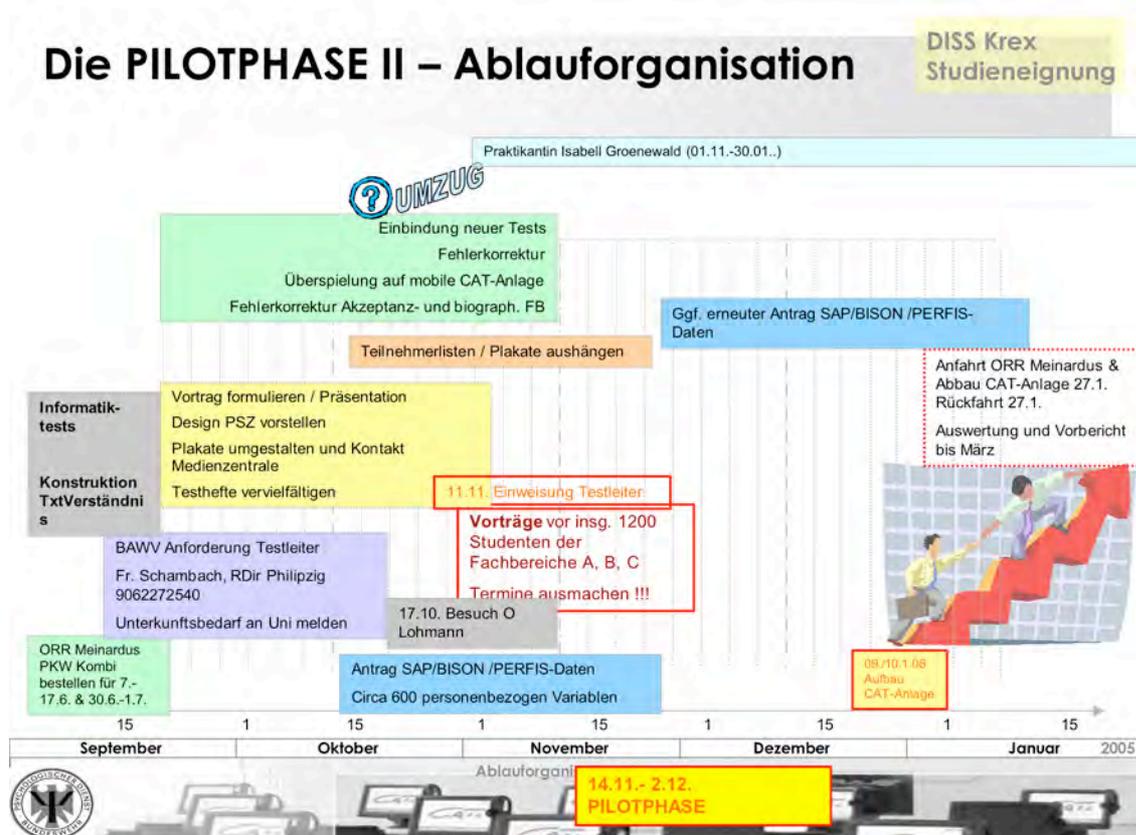


Abbildung 14: Ablauforganisation Pilotphase II (Universität der Bundeswehr München) Projekt Studieneignung

Alle Auswahlmodelle und -kriterien zur Studieneignungsvorhersage beruhen auf der Grundannahme, dass es personenbezogene, stabile, messbare, prognosefähige und für Selektionsentscheidungen heranziehbarere Merkmale gibt, welche in Zusammenhang mit dem späteren Studienerfolg stehen (Deidesheimer Kreis, 1993). Wie in Kapitel 3.4 beschrieben, bilden im Bedingungsmodell des Studienerfolgs studentische Personenvariablen wie private und finanzielle Situation, Motivation und im Verlauf des Studiums erworbene Fertigkeiten einen wichtigen Teil des Bedingungsgefüges. Hinzutreten und zu berücksichtigen sind ferner universitäre Charakteristika der Ausbildungsqualität. Gemeint sind zum Beispiel die Studieninhalte und didaktische Qualität der Veranstaltungen, die Kompetenz der Lehrpersonen, die Betreuung und Organisation sowie räumliche, sachliche und personale Ausstattung. Dennoch ist eine das Studium begleitende und daher prozessorientierte Diagnostik weder für öffentliche noch Bundeswehrhochschulen praktikabel. Die Entscheidung über die Zulassung zum Studium muss schließlich zwingend zu einem Zeitpunkt vor Beginn des ersten Semesters bzw. vor Beginn der Offizierausbildung fallen. *Statische Diagnostik* kann im Vergleich zur Bewerberinnenauswahl z.B. per Los die Qualität der Urteile entscheidend verbessern.

Auswahlverfahren können demnach nur den Input in das Ausbildungssystem Universität optimieren, sie können nicht die universitäre Ausbildung als solche in ihrer Qualität verbessern. Ein adäquates Auswahlverfahren verbessert insofern den Output, da bei alleinigem Einlass von Hochbefähigten die Wahrscheinlichkeit steigt, dass wieder Hochqualifizierte das Ausbildungssystem verlassen. In der Pilotphase des Projekts Studieneignung wird der Schwerpunkt zunächst ausschließlich auf die Komponente computerassistierte studienfeldspezifische Leistungs- und Persönlichkeitstestbatterie sowie Feedbackgespräch gelegt. Im Rahmen der Evaluation wird die Abiturdurchschnittsnote in die Berechnungen gemäß theoretischem Ansatz (vgl. Kapitel 2.2) mit einfließen.

Aufgrund der oben diskutierten empirischen und theoretischen Ergebnisse empfiehlt sich ein *Reihungstest*, der aus einem fächerübergreifenden *allgemeinen Teil* mit drei Untertests und drei Persönlichkeitsfragebögen sowie aus einem *studienfeldspezifischen Teil* mit je nach Studienfeld zwei bis drei Untertests besteht.

## 8.2 Festlegung und Definition der Anforderungsmerkmale

Gemäß Auftrag BMVg war sich bei der Zusammenstellung der Testbatterie an den Kompetenzerfordernissen zu den Studiengängen gemäß „Kastner-Bericht“ zu halten. In Abbildung 15 sind alle mit über zehn Prozent gewichteten Kompetenzen tabellarisch und nach Studienfächern bzw. -feldern geordnet aufgeführt. Die vorliegenden Analysen beruhen auf von den einzelnen Fakultäten, Instituten und Lehrstühlen zugesandten Curricula, Lehrplänen sowie Studienordnungen, in denen alle Studieninhalte festgehalten sind. Darauf basierend wurden prozentuale Anteile der Lehrinhalte ermittelt, nach Grund-, Hauptstudium sowie Studienschwerpunkten gegliedert und ihnen gemäß dem so genannten Kompetenzenkatalog der Bundesagentur für Arbeit ([www.berufenet.de](http://www.berufenet.de)) Kompetenzen zugeordnet. Entsprechend der Anteile der jeweiligen Inhalte pro Studiengang werden diese zugewiesenen Kompetenzen gewichtet dargestellt.<sup>9</sup>

---

<sup>9</sup> Neben einer Spezialisierung auf Schwerpunkte wurden innerhalb der Studiengänge verschiedene *inhaltliche Schwerpunkte* gesetzt, welche sich auf den Kompetenzenkatalog auswirken. Beispielsweise unterscheiden sich die Kompetenzen im Studiengang Elektrotechnik an der Uni Hamburg von denen an der Uni München. Außerdem wirkten sich die verschiedenen Gewichtungen der Studieninhalte verschiedener Studienrichtungen auf den Kompetenzenkatalog aus. Dieser ergibt sich vereinfacht ausgedrückt - aus dem Umfang des Grund- bzw. Hauptstudiums und aus dem Umfang des betreffenden Inhalts. Jede benötigte Kompetenz erhält einen entsprechenden Punktwert und führt somit einen Gesamtpunktwert je Kompetenz. Daraus entstehen Reihenfolgen von Kompetenzen je Studienrichtung bzw. Studienschwerpunkt. Diese werden jeweils wieder zusammengefasst zu einem Kompetenzenkatalog pro Studiengang. Auf diese Weise kann eine allgemeine Tendenz entnommen werden, wie auch spezielle Kompetenzen für die einzelnen zu wählenden Studienrichtungen und Studienschwerpunkte geschlossen werden.

Abbildung 15: Anforderungsdimensionen aufbereitet und gewichtet in Anlehnung an Kastner (2004)

FACH / Uni	Allgemeiner Teil										Feldspezifischer Teil									
	Wings-genaugig./-schwindig.	schriftliches Ausdrucksvermögen	mündliches Ausdrucksvermögen	logische Ausdrucksvermögen	räuml. Denkfähigkeit	Auffassungsvermögen	Planung & Organisation	mathematische Befähigung	Pädagogisches Geschick	Sprachverständnis	technisches Verständnis	schlussfolgerndes Denken	Einfallsreichtum	Textverständnis	Beobachtungsgabe	kaufmännisches Denken	Hand- & Fingergeschicklichkeit	praktische Anstelligkeit		
Pädagogik HH	12	15	14	14	0	4	8	4	12	7	0	5	4	3	1	0	0	0	Geisteswissenschaften	
Geschichtswissenschaft HH	13	16	14	15	0	6	7	3	10	6	0	5	3	4	0	0	0	0		
Pädagogik M	12	15	14	15	0	7	9	3	13	3	0	7	2	2	0	0	0	0		
Staats- u. SoWi M	11	15	14	13	0	8	11	3	2	6	1	3	4	7	3	0	0	0		
SportWi M	14	6	14	6	0	10	4	1	14	11	0	3	3	0	8	0	0	0		
WOW M	13	16	16	7	4	13	5	6	2	11	0	3	1	3	1	0	0	0	Wirtschaftswissenschaften	
Wirtschaftsinformatik M	21	10	10	17	0	15	2	5	2	5	0	12	1	0	0	0	0	0		
WIW HH	13	10	11	10	7	9	8	6	4	4	7	2	5	2	0	2	1	2		
WOW PoWi HH	12	14	13	7	1	6	14	5	1	12	0	4	7	5	3	1	0	0		
WOW BWL HH	13	15	15	9	3	12	5	6	1	9	0	3	1	6	2	1	0	0		
BWL FH M	16	17	17	11	1	10	4	2	1	10	1	2	1	8	0	1	0	0		
WOW VWL HH	11	14	14	9	4	9	10	4	2	8	0	2	4	0	0	2	0	0		
Geodäsie & Geoinformation M	16	11	12	5	12	4	2	14	1	1	3	2	2	1	0	0	8	8	Ingenieurwissenschaften	
Elektrotechnik HH	14	9	10	11	12	5	9	7	1	1	11	1	4	0	0	1	2	4		
Elektrotechnik & Informationstechnik M	15	8	9	12	12	6	7	8	1	1	8	1	4	0	0	1	2	4		
Elektrotechnik & techn. Informatik FH M	12	10	11	13	10	5	8	8	1	2	9	1	6	0	0	2	2	2		
Maschinenbau HH	13	8	8	9	11	7	8	6	0	1	11	0	6	0	0	5	4	6		
Maschinenbau FH M																				
BauIngWesen & Umwelttechn.M	17	12	8	4	14	6	13	7	1	1	11	1	2	0	0	0	2	5		
LRT M	16	4	6	9	14	9	7	11	2	1	12	2	7	0	0	1	1	1		
Rechnerg. Ingenieurwi. HH	12	4	4	9	11	10	4	10	1	1	8	1	8	0	0	1	5	5		
Informatik M	10	13	13	10	6	7	5	14	2	3	6	2	6	2	0	0	0	4		

Es soll darauf hingewiesen werden, dass Zielsetzung dieser theoretischen Anforderungsanalyse die verbesserte Informationsgrundlage für die OPZ-Studienberatung war. Sie erhebt ausdrücklich nicht den Anspruch einer akkuraten Realitätsabbildung und kann vielmehr als Inhaltsanalyse schriftlich festgehaltener Curricula et cetera angesehen werden. Aus diesem Grunde werden im Rahmen dieser Arbeit die in der Forschung und Literatur dokumentierten Studienfeldanalysen zur Festlegung der Prädiktorenzusammenstellung ergänzend und die Ergebnisse der zugrunde zu legenden Anforderungsanalysen absichernd berücksichtigt.

### 8.3 Die Testbatterie Studieneignung

Aufgrund der im Theorieteil konferierten empirischen und theoretisch fundierten Ergebnisse entstand im Rahmen des Projekts Studieneignung ein Reihungstest. Dieser setzt sich aus einem fächerübergreifenden allgemeinen Teil I und II mit drei kognitiven Untertests und drei Persönlichkeitsfragebögen sowie aus einem spezifischen Teil mit je nach Studienfeld zwei Untertests (vgl. Tabelle 14) zusammen.

Tabelle 14:

Übersicht Testbatterie Studieneignung unterteilt nach Studienfelder der Universitäten der Bundeswehr. ' = Minutenangabe

<b>Allgemeiner Teil I</b>			
Figuren erkennen (10')			
Verbale Analogien (adaptive Testform)			
<b>Spezifischer Teil</b>			
GeiWi	WiWi	InWi	Informatik
Textverständnis I (11')	Modellanalyse	Formalisieren technischer Zusammenhänge (60') *	Flussdiagramme (60') *
Textverständnis II (11')	(60') *		
Sprachensysteme (44')*	Diagramme & Tabellen (55') *	Technisches Verständnis (45') *	Formale Systeme (60') *
<b>Allgemeiner Teil II</b>			
Persönlichkeitsmerkmale Offizierbewerber Version Studieneignung (15')			
Big Five Kurzversion (12') **			
Stressbewältigungsressourcen (8') **			
Anmerkungen: GeiWi: Geisteswissenschaften; WiWi: Wirtschaftswissenschaften; IngWi: Ingenieurwissenschaften. * : Test der Bundesagentur für Arbeit für die Pilotphase an den UniBw; ** : Test der Schweizer Armee. , : Minutenangabe.			

### 8.3.1 Allgemeiner Studierfähigkeitstestteil der Pilotphasen

Im allgemeinen Teil I des Reihungstests sollen zentrale intellektuelle Fähigkeiten und Fertigkeiten sowie im allgemeinen Teil II Persönlichkeitsmerkmale wie beispielsweise Selbstkontrolle, Gewissenhaftigkeit, Stressbewältigungsverhalten und Leistungsmotivation, die allgemein für die erfolgreiche Bewältigung eines Studiums von Bedeutung sind, geprüft werden. Der allgemeine Teil ist insgesamt für die vier Bundeswehrstudienfelder Geistes-, Wirtschafts- und Ingenieurwissenschaften sowie Informatik identisch. Laut Kastner (2004) stechen Wahrnehmungsgenauigkeit und Wahrnehmungsgeschwindigkeit, mündliches und schriftliches Ausdrucksvermögen sowie logisches Denkvermögen als fächer- bzw. feldübergreifende Anforderungen hervor. Auch Hell (2005, S. 12) beschreibt das Schlussfolgernde Denken (reasoning) als „die dominierende Intelligenzfacette aller Studierfähigkeitstests“. Tabelle 14 liefert einen Überblick über studienrelevante Eignungsmerkmale und für die Pilotphasen ausgewählte Prädiktoren. Aufgrund immenser zu erwartender Entwicklungskosten von Testverfahren und festgesetztem Einsatztermin in der OPZ zum 01. Juli 2006<sup>10</sup> wurde bei den vorgeschlagenen Testbatterien zur Messung der allgemeinen Intelligenz bzw. allgemeinen Studierfähigkeit zunächst auf in CAT vorhandene Untertests (Figurenerkennen, Verbale Analogien) und für das Kriterium Studienerfolg relevante Skalen des Persönlichkeits-

<sup>10</sup> Aufgrund interner struktureller Veränderungsprozesse und verlagertes Prioritäten wurde der Termin auf Frühjahr 2007 festgelegt.

fragebogens „Persönlichkeitsmerkmale Offizierbewerber“ (PMO) zurückgegriffen. Vorhandene Kooperationsverträge mit der Schweizer Armee ließen außerdem die Verwendung von Paper-Pencil-Versionen sowohl einer Kurzfassung des von der Uni Zürich bereitgestellten Persönlichkeitsfragebogens „Big Five Kurzfassung“ als auch eines eigens für die Schweizer Offizierauswahl entwickelten Stressbewältigungsressourcenfragebogens. Der allgemeine Teil der Pilotphasen besteht somit fächer- bzw. studienfeldübergreifend aus zwei kognitiven und drei Persönlichkeitsuntertests bzw. -fragebögen:

Allgemeiner Teil I:

1. FE (Figurenerkennen)
2. VAADPT (Verbale Analogien als adaptive Testform)

Allgemeiner Teil II:

3. BFI (Big Five Kurzfassung)
4. Stressbewältigungsressourcenfragebogen
5. PMO (Persönlichkeitsmerkmale Offizierbewerber)

Tabelle 15:

*Zu messende fächerübergreifende Basiskompetenzen gemäß Anforderungsprofil des Kastner-Abschlussberichts und in den Pilotphasen an den UniBw angewandte Prädiktoren*

Fächerübergreifende Basiskompetenzen		CAT-Untertest
Wahrnehmungsgenauigkeit und -geschwindigkeit	Fähigkeit zum raschen Auffinden und Unterscheiden klar umrissener Gegenstände innerhalb eines einfach strukturierten visuellen Wahrnehmungsfeldes. Der Abschlussbericht 1997 zum Auswahlverfahren für die Verwendungen im Fliegerischen Dienst der Luftwaffe sieht für den Bereich Wahrnehmung / perceptual speed den Untertest Figurenerkennen (FE) vor.	Figurenerkennen (FE)
Mündliches und schriftliches Ausdrucksvermögen	Fähigkeit, Sachverhalte verständlich darzustellen Bewertungen auszudrücken und Handlungsaufforderungen wirksam an den Empfänger zu richten. Klassische Erhebungsmethode: Interview / Arbeitsprobe (Klausur o. ä.). Als Testverfahren eventuell analog dem Primärfaktor „word fluency“ über Wortproduktionsaufgaben zu ermitteln.	Verbale Analogien als adaptive Testform (VAADPT) (sprachabhängiges logisches Denken)
Logisches Denkvermögen	Konstrukt lässt sich über sprachbezogene, sprachunabhängige oder numerische Analogieaufgaben, Matrizentests u.a. erfassen.	Logisches Denken

### 8.3.2 Spezifischer Studierfähigkeitstestteil der Pilotphasen

Bei den studienfeldbezogenen Beratungstestserien der Bundesagentur für Arbeit (BA) handelt es sich um spezifische Studierfähigkeitstests (vergleiche Kapitel 2.1.3). Sie sind vor allem für diejenigen Bewerber hilfreich, die wie in der Bundeswehr üblich bereits eine bestimmte Studienrichtung in die engere Wahl gezogen haben (Erst- bis Drittwunsch). Die Aufgaben der feldspezifischen Eignungstests werden auf fachtypische Problemstellungen

zugeschnitten und in fachbezogene Inhalte eingekleidet. Sie stellen somit in gewissem Sinne realitätsnahe Arbeitsproben, jedoch keine Wissensproben – dar. Darüber hinaus ermöglichen sie eine Rückmeldung des eigenen feldspezifischen Fähigkeitsniveaus und liefern zugleich Aufschluss über typische Anforderungen des gewünschten Studiums. Voraussetzung für die Gewährleistung der Fachnähe der studienfeldbezogenen Testbatterie ist, dass sie in enger Zusammenarbeit zwischen Fachvertretern bzw. Hochschullehrern und Diplom-Psychologen entwickelt wurden. Der Konzipierung der Aufgabentypen sowie der eigentlichen Itemkonstruktion ging zudem eine sorgfältige Analyse der spezifischen Studienanforderungen voraus (Bundesagentur für Arbeit, 2004). Die heutige Firma ITB Consulting GmbH (ehemals Institut für Test- und Begabungsforschung) erhielt nach jahrelanger Aufgabenüberprüfung an angehenden Offizieren von Seiten der BA den Auftrag, fachbezogenen Eignungstest unter anderem für die philologischen Studiengänge und für die Studienfelder Wirtschafts- und Ingenieurwissenschaften sowie Informatik zu entwickeln. Durch die Erprobung in Quer- und Längsschnittstudien sollten unter anderem Fragen nach den tatsächlichen Messbereichen der Tests, nach deren Messgenauigkeit, Prognosekraft bezüglich des Erfolgs in den betreffenden Studiengängen und nach der Akzeptanz der Verfahren bei Oberstufenschülern beantwortet werden. Anfangs wurden die Tests an circa 7000 Abiturienten in fünf Bundesländern querschnittlich geprüft. Im Anschluss daran wurde ihre Verwendbarkeit zu Beratungszwecken im Rahmen eines Modellversuchs an 4400 Schülern im Bundesland Niedersachsen evaluiert und als äußerst positiv hinsichtlich interner Gütekriterien bewertet. Eine daran anschließende Längsschnittstudie ergab Aufschluss über den Zusammenhang zwischen den Testergebnissen und dem späteren Abschneiden der Teilnehmer in den Vorexamina der betreffenden Studiengänge und damit über die prognostische Validität der Beratungstests bezüglich des Studienerfolgs. Diese Bewährungskontrolle wurde ergänzt durch einen weiteren Feldversuch mit Offizieranwärtern, bei dem die Prognosekraft der Tests im Hinblick auf den Erfolg in einzelnen Studiengängen an den UniBw geprüft werden konnte. Das Gros der Untertests erzielte befriedigende bis günstige Prognosekennwerte. Die BA einschließlich entwickelte die ITB Consulting GmbH fachspezifische Eignungstests für die Zulassung an zahlreichen, meist die Hochschulrankings anführenden privaten und staatlichen Hochschulen im In- und Ausland. Bei der Entwicklung dieser Tests konnten die Vorerfahrungen aus den Modellversuchen genutzt werden, um Testkonzepte weiter oder gänzlich neu zu gestalten. Die im Rahmen von Amtshilfe für das Pilotprojekt an die Bundeswehr übergebenen fach-

spezifischen Studienberatungstests der BA wurden zuvor bis zum Jahre 2001 exklusiv von ITB Consulting weiterentwickelt.

#### 8.4 Werbung freiwilliger Testpersonen

Im Vergleich zu früheren Testevaluationen vom Institut für Test- und Begabungsforschung legte die Autorin in den Pilotphasen des Projekts Studieneignung anstelle von „befohlener Freiwilligkeit“ höchstes Augenmerk auf fakultative Teilnahme aller Testpersonen. Um trotz der auf den ersten Blick eher abschreckenden Gesamtlänge der Testbatterie (je nach Studienfeld bis zu drei Stunden) einen zufriedenstellenden Stichprobenumfang zu erreichen, wurden bereits Wochen vor dem Start des Assessments aufwendige Marketingmaßnahmen veranlasst. Auf Neugierde weckenden DIN-A-0-Werbeplakaten mit dem Titel „Wir testen die Besten“ waren neben überzeugenden Argumenten (Sicherung des eigenen Qualitätsniveaus, Erhalt eines ausführlichen Feedbackgesprächs durch eine Diplom-Psychologin, Erhalt eines Persönlichkeitsprofils und Stärken- Schwächenprofils, Trainingsmöglichkeit für spätere Assessment Center nach Ausscheiden aus der Bundeswehr usw.) auch der Projektablauf und Ansprechpartner sowie die Präsidentin der UniBw München bzw. der Präsident der UniBw Hamburg und höchste militärische Vorgesetzte als Unterstützer und Befürworter aufgeführt. Pro Fachbereich fand eine einstündige Informationsveranstaltung mit Präsentation und Diskussionsteil und Auslegung der Teilnehmerlisten im Anschluss statt. Auf den Wohnebenen des Campus-Betriebs wurde durch ausgewählte Studenten per Mund-zu-Mund-Propaganda für eine Teilnahme geworben und auf die Bedeutsamkeit für die Bundeswehr, die UniBw und jeden einzelnen Bundeswehrangehörigen hingewiesen.

### 9. Durchführung

An die Prüfung der zu messenden relevanten Eignungsmerkmale schließt sich die Durchführungsphase an. Sie beinhaltet unter anderem die Überprüfung der ermittelten Eignungsmerkmale und Prädiktoren. Die Diagnose der Bewährungswahrscheinlichkeit der Testpersonen sowie die Auswahlentscheidung finden sich im Ergebnisteil wider.

## 9.1 Testdurchführung und -auswertung

Die Testzeit der gesamten im Rahmen der vorliegenden Arbeit zusammengestellten und weiterentwickelten Reihungstestbatterie beträgt je nach Studienfach zwischen 120 und 200 Minuten. Die Tests enthalten nur im allgemeinen Teil I Übungsaufgaben, verfügen jedoch alle über eine eigene und auf die Pilotphasen zugeschnittene ausführliche Instruktionsphase sowie über allgemeine Hinweise zur Bearbeitung der Testserie. Die Testauswertung der Bundeswehr-Pilotphasenversion läuft ausnahmslos computergestützt auf mobilen CAT-Anlagen in Form von transportablen Notebooks. Nur so konnte ein reibungsloser Durchlauf einer großen Anzahl von Testpersonen in kurzer Zeit realisiert werden. Gleichzeitig lässt sich ein individueller Test- aber auch freiwilligen Pausenbeginn für jede Person ermöglichen und auf zeitbezogene Instruktionen der Testleiter völlig verzichten. Zusätzlich kann der fortwährend vor Ort befindliche Testleiter individuell auf organisatorische und Verständnisfragen in der Übungsphase oder auf die CAT-Anlage bezogen reagieren.

Die Testinterpretation setzt testtheoretische Kenntnisse auf Seiten des Studienberaters in der Praxis bzw. Diplom-Psychologen in der Pilotphase voraus, um die Aussagekraft der Testergebnisse hinsichtlich der Testgütekriterien einschätzen zu können. Die Testergebnisse auf Ebene der verschiedenen Aufgabengruppen im spezifischen Teil sind vor allem dann von Interesse, wenn sie sich stark voneinander unterscheiden (z.B. erscheint der Untertest Modellanalyse hauptsächlich der Volkswirtschaftslehre nahe). In solchen Fällen können spezifische Stärken oder Schwächen identifiziert werden, die bei alleiniger Betrachtung des Gesamtwertes unerkannt blieben. Zur Ermittlung der Normwerte stand zur Rückmeldung der freiwilligen Testpersonen zunächst die Gesamtstichprobe der Studien durch die BA als Referenzgruppe zur Verfügung. Die Dissertation liefert im Anschluss an die Itemanalysen eine Empfehlung für zukünftig in der Praxisanwendung der OPZ zu nutzenden Normskalenwerte.

Für die Studiengänge wurden diejenigen Tests ausgewählt, welche bei den bisherigen Bewährungskontrollen durch ITB Consulting und die BA die höchste prognostische Gültigkeit bezüglich des Erfolgs in den betreffenden Studiengängen aufweisen.

## 10. Forschungsfragestellungen

Im Mittelpunkt des empirischen Teils dieser Arbeit steht die Entwicklung zukünftiger Prädiktoren von Studienerfolg im Rahmen der Optimierung der Offizierauswahl in den deut-

schen Streitkräften. Mit dem zentralen Ziel dieser Arbeit, Aussagen über die Verbesserung des Offizieranwärterauslese- und Offizierplatzierungsverfahrens in Bezug auf die fächerbezogene Studieneignungsfeststellung treffen zu können, stehen folgende Fragestellungen im Vordergrund dieser Arbeit:

1. Wie genau und messfehlerfrei wird das intendierte Konstrukt Studienerfolg mit Hilfe der Prädiktoren im Vergleich zur Abiturgesamt- bzw. Schulabschlussnote gemessen (Reliabilität)?
2. Wie gut werden die intendierten Konstrukte mittels untersuchter Reihungstests bzw. mittels der einzelnen Subtests tatsächlich gemessen (Konstruktvalidität)?
3. Wie leistungsfähig erweist sich der studienfeldspezifische Eignungstest für die Prognose späterer Studien- und Offiziereignung (prognostische Validität)?
4. Welche Resultate liefert die empirische Untersuchung an den beiden Universitäten der Bundeswehr in Hamburg und München hinsichtlich geeigneter Prädiktoren zur Vorhersage der Studieneignung (prognostische Validität)?
5. Welche Schlüsse lassen Prozess- und Postevaluation hinsichtlich möglicher interner und externer Einflüsse zu?
6. Wie steht es mit der Akzeptanz der praktizierten und untersuchten Auswahlverfahren bei Offizieranwärtern (Akzeptanz)?
7. Wie leistungsfähig erweisen sich Qualität, Objektivität, Effizienz und Umsetzbarkeit bestehender und untersuchter Prädiktoren (Zusatznutzen)?
8. Wo liegen Ansatzpunkte für eine Optimierung des entwickelten studienfeldspezifischen Eignungstest (Ableitungen)?

Der verbesserte Wissensstand über die Zusammenhänge der unterschiedlichen Prädiktoren, Kriterien und der theoretisch relevanten Konstrukte verspricht darüber hinaus ein Beitrag zur Studienberatungsforschung. Die gesonderte Betrachtung von Moderatorvariablen und nicht-kognitiven Einflussvariablen soll zusätzliche Hinweise unter anderem auf beobachterbezogene Ergebnisse, auf militärraffine und geschlechtsspezifische Erwartungs- und Arbeitshaltungen geben und die differentielle Validität der eingesetzten Verfahren beschreiben. Die damit verbundene inhaltliche Fragestellung lautet:

- |   |
|---|
| 9. Bestehen Unterschiede in der Studienerfolgs-Vorhersageleistung für verschiedene Teilgruppen oder aufgrund anderer interner und externer Einflüsse (Moderatorvariablen, Nicht-kognitive Einflüsse)? |
|---|

Entsprechend Wottawa & Thierau (1992) stammen Hypothesen des Wirkmodells einer Evaluation im Idealfall aus folgenden Quellen: Aus experimentellen Studien, die Kausalschlüsse zulassen oder aus gut entwickelten Theorien oder Ergebnissen ähnlicher Interventionen im gleichen Programmsektor. Ein übergeordnetes Ziel dieser Arbeit ist die Prüfung der Verfahrensgüte einer zukünftigen Studierfähigkeitstestbatterie für die Offizierauswahl der deutschen Streitkräfte. Somit werden nur bedingt klassische inhaltliche Hypothesen abgeleitet werden können. Es handelt sich daher in einigen Fällen eher um Annahmen als um aus empirischen Ergebnissen abgeleitete Hypothesen. In den nachstehenden Kapiteln werden nach Gütekriterien und aufgestellten Hypothesen geordnet die Annahmen bzw. Thesen zur Güte sowie inhaltliche Hypothesen aufgeführt (Kapitel 10.1 bis 10.3). Ergänzend wird jeweils eine genaue Darstellung zu ihrer Operationalisierung erarbeitet. Außerdem soll die Gesamtstichprobe beschrieben (Kapitel 10.4) und die Untersuchungsvariablen vorgestellt werden (Kapitel 10.5).

Ausgewählt werden sollen die Subtests, die günstige Kennwerte bezüglich des Schwierigkeitsniveaus und der messtechnischen Zuverlässigkeit erzielen. Außerdem sollen diejenigen Items in die Praxis umgesetzt werden, die in vergleichsweise enger Beziehung zum Gesamtwert im vollständigen Test und zugleich in vergleichsweise schwacher Beziehung untereinander stehen und denen nach Aufgabentyp und inhaltlicher Einkleidung eine hohe Fachspezifität und damit auch eine hohe Augenscheinvalidität für den Bearbeiter zukommen. Mit dem übergeordneten Ziel der vorliegenden Arbeit vor Augen erweisen sich die Aufgaben bzw. Subtests als besonders unerlässlich, welche mindestens zufrieden stellende prädiktive Validität aufweisen.

## 10.1 Reliabilitätsbestimmung

Die Beurteilung der Zuverlässigkeit der Untersuchungsdaten dient als Kriterium der Genauigkeit erhobener Testdaten aus den Pilotphasen an den UniBw und gibt Auskunft über mögliche Fehlereinflüsse. Gemäß der deutschen Industrienorm 33430 zu den Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen (Deutsches Institut für Normierung, 2002) ergeben sich erfahrungsgemäß bei Reliabilitätsuntersuchungen Werte

von mindestens .70 als brauchbar und Werte größer als .85 als erstrebenswert jedoch nur schwer realisierbar. Für die Reihungstestbatterie Studieneignung soll für im Hinblick auf eine spätere Praxisumsetzung Folgendes angenommen werden:

**Annahme:** Alle dasselbe Merkmal messende Items liegen hinsichtlich der Antworten der Probanden nicht weit auseinander.

**Annahme:** Ein zuverlässiges Messen der Reihungstestbatterien ist sowohl auf Item-, Subtestebene, in Bezug auf die einzelnen Teile (Allgemeiner Teil I und II, Spezifischer Teil) und auf den gesamten Reihungstest bezogen gegeben.

Als Kennzahl der Item- und Testreliabilitäten dienen sowohl der Schwierigkeits- als auch der Trennschärfekoeffizient. Beide sind nicht unabhängig voneinander zu sehen. Die Trennschärfe ist für mittlere Schwierigkeitsindizes am größten, während sie für niedrige und hohe sinkt (Rost, 1996). Die Reliabilität wird über den Koeffizienten Cronbachs Alpha ermittelt. Das einfachste Kriterium zur Feststellung der Zweckdienlichkeit eines Items ist der prozentuale Anteil der Richtig-Antworten. Ist  $n$  die Anzahl der Probanden,  $m$  die Anzahl der Items und  $R_j$  die Anzahl der Richtig-Antworten des  $j$ -ten Items, so berechnet sich der *Schwierigkeitsindex*  $P_j$  des  $j$ -ten Items zu  $P_j = \frac{R_j}{n} \cdot 100$  mit  $j = 1, \dots, m$ .

Items mit einem Schwierigkeitsindex kleiner .20 oder größer .80 sollten eliminiert werden.

Da es im Falle der drei Persönlichkeitstests mit ihrer 7er-Skala keine genau definierten Richtig-Antworten gibt, ist für den Schwierigkeitsindex eine modifizierte Formel zu verwenden:

$P_j = \frac{\bar{x}_j - \bar{x}_{\min}}{x_{\max} - x_{\min}} \cdot 100$  mit  $j = 1, \dots, m$ . Dabei sind die  $\bar{x}_j$  Mittelwerte der  $m$  Items über die  $n$

Probanden.  $x_{\min}$  und  $x_{\max}$  bezeichnen die kleinste bzw. größte Item-Codierung (in unserem Falle 1 bzw. 7).

Das wichtigste Kriterium zur Beurteilung der Brauchbarkeit eines Items gibt an, wie gut das betreffende Item zwischen guten und schlechten Testpersonen trennt. Der Trennschärfekoeffizient ist die – im Falle von Richtig-Falsch-Antworten – punktbiserale Korrelation zwischen der Aufgabenantwort und dem Gesamtttestscore. Ist  $n_r$  die Anzahl der Richtig-Antworten des Items und  $n_f$  die Anzahl der Falsch-Antworten, so berechnet sich der *Trennschärfekoeffizient*

T folgendermaßen:  $T = \frac{\bar{x}_r - \bar{x}_f}{n \cdot s} \cdot \sqrt{n_r \cdot n_f}$ .

Seine Signifikanzüberprüfung erfolgt über die t-verteilte Prüfgröße  $t = |T| \cdot \sqrt{\frac{n_r + n_f - 2}{1 - T^2}}$  bei  $df = n_r + n_f - 2$  Freiheitsgraden.

Bei den Stufen-Antwort-Aufgaben wie die der drei Persönlichkeitsuntertests bietet sich die Rangkorrelation nach Spearman zwischen dem betreffenden Item-Wert und dem Gesamtpunktwert an.

Die Streuungen der einzelnen Items können nach der üblichen Formel der Standardabweichung berechnet werden, wobei als Messwerte nur 0 und 1 auftreten. Sie können bei dichotomen Richtig-Falsch-Aufgaben aber auch unmittelbar aus den Schwierigkeitsindizes  $P_j$

bestimmt werden.  $s_j = \sqrt{\frac{P_j}{100} \cdot \left(1 - \frac{P_j}{100}\right)}$  mit  $j = 1, \dots, m$ .

*Cronbachs Alpha* berechnet sich als Maß der internen Konsistenz aus den Itemstreuungen  $s_j$  und der Streuung  $s$  des Gesamtpunktwerts. Interne Konsistenz bedeutet, dass die einzelnen Items mit der Gesamtheit der übrigen Items zusammenhängen. Diese gilt als Maß der Reliabilität der Skala. Alpha kann bei perfekter Konsistenz ein Maximum von + 1.00 erreichen. Mit sinkendem Wert verringert sich auch die Konsistenz.

Die *Reliabilität gesamter Tests* kann unter der Annahme gleicher Untertestgewichte nach der Spearman-Brown Formel folgendermaßen geschätzt werden (Lienert und Raatz, 1994):

$r_{tt} = \frac{k \cdot \bar{r}_{tt}}{1 + (k - 1)\bar{r}_{tt}}$ .  $\bar{r}_{tt}$  stellt dabei das arithmetische Mittel der Untertestreliabilitätskoeffizienten<sup>11</sup> dar, k die Anzahl der Untertests. Die Gesamtreliabilität von Fragebögen bzw. Mehrfachwahlantwortentests gilt gemäß Lienert (1989, S. 380) als sehr gut, wenn  $r > .80$  erreicht wird.

## 10.2 Validitätsbestimmung

Die Validität eines diagnostischen Verfahrens gibt Auskunft über seine Gültigkeit. Lienert (1969) versteht darunter den Grad der Genauigkeit,

..., mit dem dieser Test dasjenige Persönlichkeitsmerkmal oder diejenige Verhaltensweise, das (die) er messen soll oder zu

<sup>11</sup> Über die z-Transformation zu berechnen.

messen vorgibt, tatsächlich misst. Ein Test ist demnach vollkommen valide, wenn seine Ergebnisse einen unmittelbaren und fehlerfreien Rückschluss auf den Ausprägungsgrad des zu erfassenden Persönlichkeits- oder Verhaltensmerkmal zulassen, wenn also der individuelle Testpunktwert eines Probanden diesen auf der Merkmalskala eindeutig lokalisiert (S. 16).

Gemäß der DIN 33430 (2002) „muss die Gültigkeit des Verfahrens aufgrund von empirischen Analysen zur Konstrukt-, Kriteriums- oder Inhaltsvalidität nachgewiesen werden. Die Art der Gültigkeitsbestimmung muss dem Zweck des Verfahrens und der vorliegenden Fragestellung angemessen sein“ (S. 7).

Grundsätzlich können drei Formen der Validität unterschieden werden: *Kriteriums-*, *Konstrukt-* und *Inhaltsvalidität*. Bei der *Kontent-* oder *Inhaltsvalidität* erfolgt der Schluss von Verhalten in der Testsituation auf das Verhalten außerhalb der Testsituation über einen Repräsentationsschluss, nämlich immer, wenn das Testverhalten als direkt repräsentativ für ein bestimmtes Gesamtverhalten angesehen wird. Die Inhaltsvalidität kann meistens als gegeben vorausgesetzt werden, da - wie zum Beispiel bei den hier verwendeten Testverfahren - die Aufgaben von Experten selber erstellt worden sind. Ergänzend führt Barell (1992) die Begriffe *soziale Validität* sowie *innere Validität* auf. Einen Überblick über die fünf Validitätsarten nach Barell bietet Abbildung 16. Die innere Validität ist gleichzusetzen mit der Reliabilität. Obwohl die einzelnen Validitätstypen voneinander unabhängig ermittelt und diskutiert werden können, sind sie miteinander verknüpft. Die folgenden Kapitel beschränken sich auf die für diese Arbeit relevanten Validitätsbegriffe der inneren Validität oder Reliabilität, der Konstrukt- und prognostischen Validität.

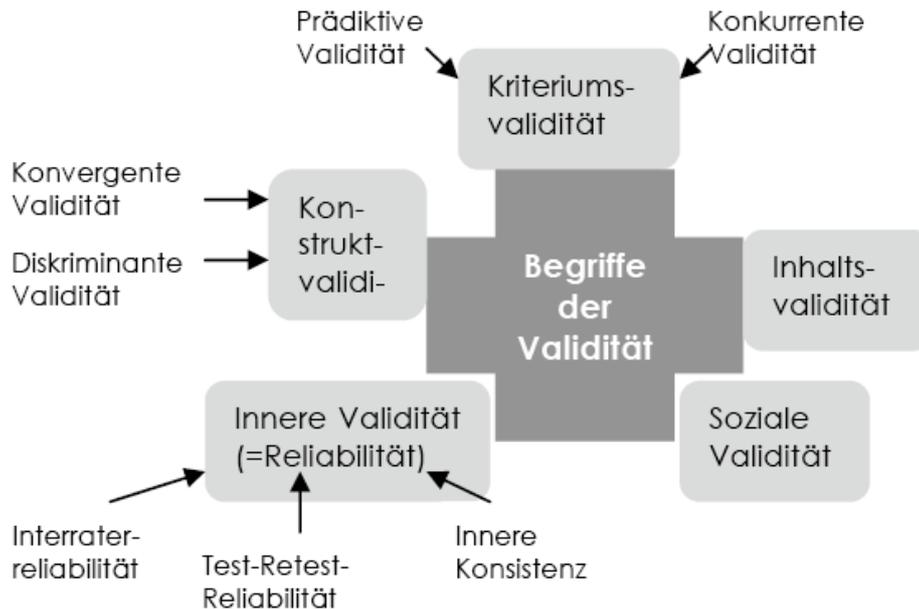


Abbildung 16: Fünfteilung der Validität. Die Ecken des Fünfecks zielen jeweils auf eine der Validitätsformen nach Barell (1992)

### 10.2.1 Konstruktvalidität

Vom Verhalten in der Testsituation wird geschlossen auf Fähigkeiten, Dispositionen oder Persönlichkeitsmerkmalen als Grundlagen oder Bedingungen des Verhaltens.

Erfüllt ein Test (nicht nur einen pragmatischen Anspruch, sondern) gewisse theoretische bzw. theoriegeleitete Vorstellungen hinsichtlich der Messintention, dann ist ihm Konstruktvalidität zuzusprechen. (Kubinger & Proyer, 2004, S. 189)

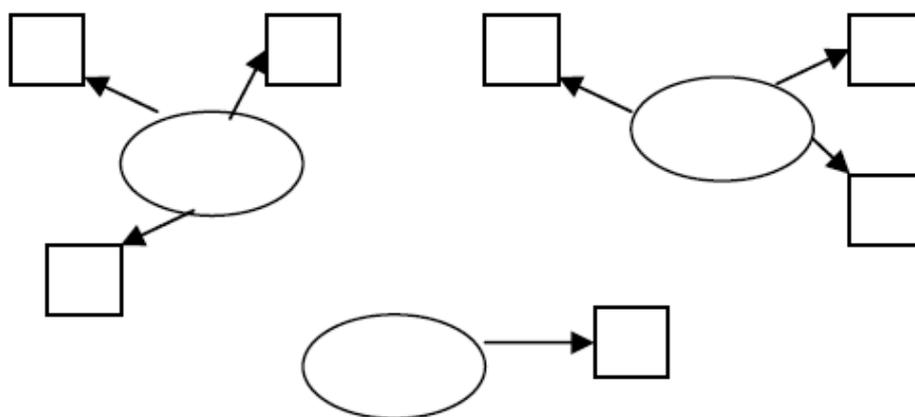
Der Konstruktvalidität zu subsumieren sind die konvergente und diskriminante Validität. *Diskriminante Validität* besagt, dass durch die Messungen sowohl innerhalb einer Methode als auch über verschiedene Methoden hinweg unterschiedliche Traits auseinander gehalten werden können. Die Dimensionsscores für verschiedene Konstrukte korrelieren bei derselben Methode niedrig miteinander, so dass es innerhalb eines Untertests zu einer niedrigen Korrelation zwischen unterschiedlichen Konstrukten kommt. Bei Vorhandensein *konvergenter Validität* korrelieren die Dimensionsscores für dasselbe Konstrukt signifikant über alle sie messenden Testverfahren hinweg hoch miteinander. Konvergente Validität bedeutet, dass die Messungen eines Traits mit verschiedenen Methoden konvergieren.

**Annahme:** Die interessierenden Konstrukte werden nachvollziehbar abgebildet und heben sich von anderen Konstrukten ab.

Die interne Konstruktvalidierung soll ermitteln, ob in dieser Studie die intendierten und insbesondere studienfeldspezifischen Konstrukte mit den eingesetzten Verfahren auch tatsächlich erfasst werden.

**Annahme:** Die intendierten Konstrukte werden mit der eingesetzten Studieneignungstestbatterie studienfeldspezifisch tatsächlich erfasst.

In der Literatur finden sich verschiedene statistische Vorgehensweisen zur Analyse der gewonnenen Daten. Neben der Multi-Trait-Multi-Method-Analyse, die insbesondere für multimodale Verfahren wie dem AC Anwendung finden, lassen sich beispielsweise die zu messenden Konstrukte durch eine Korrelation der Ergebnisse mit verschiedenen psychometrischen bzw. Persönlichkeitstestergebnissen vergleichen. Zunächst werden Hypothesen generiert, um ein so genanntes nomologisches Netzwerk – in der Literatur häufig als externe Validierung bezeichnet – aufstellen zu können. Theoretische Konstrukte werden demnach in ein theoretisches Bezugssystem eingeordnet, aus denen überprüfbare Hypothesen abgeleitet werden können. Je fester ein Test in das theoretische Bezugssystem eingeordnet werden kann, je dichter das nomologische Netz also ist, desto größer ist die Evidenz, dass ein Test das Konstrukt erfasst, das er zu erfassen vorgibt. Auf Grundlage des in *Abbildung 17* skizzierten Systems lassen sich Hypothesen über Beziehungen zwischen den benutzten Indikatoren bzw. Tests machen, die dann empirisch geprüft werden.



*Legende:*  = Konstrukt  = Test/Indikator

*Abbildung 17:* Das nomologische Netzwerk. Die Pfeile stellen die Beziehung zwischen Test bzw. Indikator und Konstrukten sowie jeweils untereinander dar

Neuere Studien konzentrieren sich aus diesem Grund ergänzend auf faktoren-, pfad- und varianzanalytische Ansätze. Thornton und Byham kamen 1982 zu einer äußerst positiven Bilanz von Faktorenanalysen: „Factor analysis show consistency in clusters of dimensions used in several programs“ (1982, S. 216).

### 10.2.2 Prognostische Validität

Studienerfolg ist – wie in Teil I theoretisch erörtert – nur bedingt prognostizierbar. Wirken doch hoch komplexe, kaum im Voraus zu berechnende Wechselwirkungen zwischen Person- und Situationsmerkmalen im zeitlichen Verlauf aufeinander ein.

Findet die Datenerhebung zeitversetzt statt (beispielsweise Prädiktor vor oder zum Zeitpunkt des Studienbeginns und die des Kriteriums als Erfolgsmaß nach frühestens dreieinvierteljähriger Studienzeit), so spricht man von *prädiktiver oder prognostischer Validität*. Erfolgen die Messungen gleichzeitig, spricht man von *konkurrender Validität*. Da sich beide Validitätsformen nicht prinzipiell, sondern nur verfahrensmäßig unterscheiden, werden sie gemeinsam als *kriterienbezogene, Kriteriums- oder empirische Validität* bezeichnet. Beiden Validitätsarten liegt die Fragestellung zugrunde, wie gut mit dem Prädiktor A (zum Beispiel Subtest Formale Systeme) Verhalten (zum Beispiel Erfolg im ingenieurwissenschaftlichen Studiengang) vorhergesagt werden kann.

Wie bereits in Kapitel 2.1.3 ausführlich beschrieben spricht man Reihungstests gegenüber Schulnoten nicht nur eine höhere Objektivität, sondern auch größere Fairness hinsichtlich sozialer Selektivität zu. Ihre Fachspezifität führt zu einer höheren prognostischen Validität als die anderer Testverfahren zur Auswahl von Studierenden (Rindermann, 2004). Wie in Kapitel x zu den Phasen der Eignungsfeststellung und zur Anforderungsanalyse verdeutlicht, wird ein Eignungsbeurteilungsverfahren nur dann für die Anforderung valide sein und den vertretenen Standards gemäß DIN 33430 genügen, wenn die im Auswahlverfahren repräsentierten Anforderungen mit denjenigen des Arbeitsplatzes bzw. Studienfachs übereinstimmen. Somit wird ein über alle Dimensionen hinweg signifikanter Einfluss aller Dimensionen in der Studieneignungstestbatterie auf den Studienerfolg erwartet.

Zur prognostischen Validierung werden als Außenkriterium die Ergebnisse aus Vordiplom und Hauptdiplom herangezogen. Um Prädiktor- und Kriterienwerte miteinander zu vergleichen, empfiehlt es sich, die korrelative sowie regressionsanalytische Höhe der Zusammenhänge zu berechnen. Zusammenhangsberechnungen finden je nach Skalenniveau entweder

mit Hilfe des Spearman-Rang-Korrelationskoeffizienten oder mit der Produkt-Moment-Korrelation nach Bravais-Pearson statt (Bortz, 1993). Zur Signifikanzprüfung der Zusammenhänge wird von dem in der Grundlagenforschung üblichen Signifikanzniveau von fünf Prozent ausgegangen (Bortz & Döring, 1995).

**Hypothese:** Der Gesamtttestscore sagt den späteren Studienerfolg besser vorher als die damalige Studienempfehlung der OPZ (altes Verfahren).

$$H_0: \rho_{\text{Ges,StudE}} \leq \rho_{\text{OPZ,StudE}}$$

$$H_1: \rho_{\text{Ges,StudE}} > \rho_{\text{OPZ,StudE}}$$

Wie im theoretischen Teil ausführlich beschrieben sind studienfeldspezifische Eignungstests in Kombination mit allgemeinen Studierfähigkeitstests sowohl Kenntnis- als auch der alleinigen Nutzung von Intelligenztests hinsichtlich der klassischen Gütekriterien weit überlegen (vgl. Kapitel 2.1.3). Dies führt zur folgenden Hypothesenaufstellung:

**Hypothese:** Der spezifische Teil der Testbatterie Studieneignung ist hinsichtlich seiner prädiktiven Validität dem allgemeinen Teil der Testbatterie Studieneignung überlegen.

$$H_0: \rho_{\text{Spez,StudE}} \leq \rho_{\text{A,StudE}}$$

$$H_1: \rho_{\text{Spez,StudE}} > \rho_{\text{A,StudE}}$$

Bergmann, Brandstätter und Eder (1994) finden positive Zusammenhänge der Studienleistungen mit Gewissenhaftigkeit, Belastbarkeit und Introversion. Weitere empirische Befunde deuten auf einen positiven Einfluss von Selbstwirksamkeit, Handlungskontrolle und Lernmotivation auf Studienerfolg hin (Schiefele & Urhahne, 2000). Gewissenhaftigkeit, Zuverlässigkeit und Leistungsmotivation werden von Hitpass (1975) als hoch bedeutsam für Studienerfolg angenommen. Angstneigung, Belastbarkeit, Selbstsicherheit und Empfindsamkeit werden in zahlreichen Studien als für den Studienverlauf relevante Einflussgrößen benannt (Melter, 1977). In neuartigen Studien zum UCB (DGPs, 2006) fanden die mittels Big Five Inventar (BFI, Rammstedt & John, 2005) sowie einem neuartigen berufsbezogenen Persönlichkeitstest (Start-P, Kersting & Beauducel) gemessenen vier Dimensionen Hilfsbereitschaft, Gewissenhaftigkeit, Unkompliziertheit und Eigeninitiative als Prädiktoren Bestätigung. Entwistle & Entwistle (1970) belegen ferner, dass emotionale stabile Studenten im Studium erfolgreicher sind als Studenten mit einem hohen Neurotizismuswert (z.B. im Big Five Inventar).

**Hypothese:** Eine Erhöhung der inkrementellen Validität ist durch die Hinzunahme von Persönlichkeitstests zu erzielen.

$$H_0: \rho_{AI/Specz+AI,StudE} \leq \rho_{AI/Specz,StudE}$$

$$H_1: \rho_{AI/Specz+AI,StudE} > \rho_{AI/Specz,StudE}$$

Generell kann die Abiturdurchschnittsnote nach gegenwärtigem Forschungsstand noch vor Einzelfachnoten als einer der besten und ökonomischsten Prädiktoren und des in Prüfungen festgestellten Studienerfolgs gelten (vgl. Kapitel 2.1.1). In Untersuchungen zur inkrementellen Validität sind höhere Kriteriumsvaliditäten des Reihungstests als einzelner Subtests zu vermuten (vgl. Kapitel 2.1). In Kombination mit der Abiturnote können laut Kuncel et al. (2001) allgemeine und spezifische Studieneignungstest die inkrementelle Validität bedeutsam auf bis zu  $R = .50$  erhöhen.

**Hypothese:** Durch die Kombination von Reihungstests mit der Abiturdurchschnittsnote wird die prognostische Validität weiter erhöht.

$$H_0: \rho_{SFT/Abi,StudE} \leq \rho_{SFT,StudE}$$

$$H_1: \rho_{SFT/Abi,StudE} > \rho_{SFT,StudE}$$

### 10.3 Annahmen zum Einfluss von Moderatorfaktoren

Um dem Einfluss möglicher biographischer oder militärraffiner Moderatorvariablen gerecht zu werden, werden in den folgenden Abschnitten Hypothesen zum Einfluss von Hintergrundvariablen auf das Abschneiden in Test und Studium aufgestellt.

Schon aus Gründen der Gleichbehandlung werden in der Bundeswehr alle Testverfahren mit Ausnahme des Physical Fitness Tests nicht geschlechtsspezifisch durchgeführt. Es wird erwartet, dass weder ein Einfluss der moderierenden Hintergrundvariablen Geschlecht auf das Abschneiden im Studieneignungstest noch auf das Kriterium Studienerfolg besteht.

**Hypothese:** Der Studienerfolg von weiblichen und männlichen Offizieranwärtern unterscheidet sich nicht.

$$H_0: U_{\text{männlich}} - U_{\text{weiblich}} = 0$$

Davon ausgehend, dass die Art der Teilstreitkraft keine Effekte auf das Overallergebnis hat, werden folgende Hypothesen formuliert:

Hypothese: Der Studienerfolg der drei Teilstreitkräfte Heer (H), Luftwaffe (Lw) und Marine (M) unterscheidet sich nicht.

$$H_0: H_{He} = H_{Lw} = H_M$$

Die Tatsache, dass bayrischen Abiturienten ein höheres Anspruch- und Leistungsniveau zugesprochen wird und jene Offizierbewerber sich aufgrund des Studienortes für ein Fach an der UniBw München anstelle von Hamburg entschließen, wirft die Frage nach Qualität der Studenten an den beiden UniBw auf.

Hypothese: Der Studienerfolg der UniBw München-Studenten ist höher als der der UniBw Hamburg-Studierenden.

$$H_0: H_{UniBwM} > H_{UniBwH}$$

Da aufgrund vorher gemachter Ausführungen nicht eindeutig anzunehmen ist, ob Studierende mit besonders vielen Trimestern bis zum Vordiplom zu den schlechten oder zu den besseren, die die Zeit für eine optimale Vordiplomnote nutzen wollen, gehören, ist folgende Annahme zu überprüfen.

Hypothese: Studenten mit unterschiedlich langer Trimesterzahl bis zum Vordiplom unterscheiden sich nicht hinsichtlich des Studienerfolgs.

$$H_0: H_{>4Trimester} = H_{\leq 4Trimester}$$

Der soziokulturelle Bildungshintergrund ist zwar kein Kriterium bei der Auswahl von Studierenden, aber eine Einflussvariable, der hinsichtlich der öffentlichen Diskussion um soziale Fairness Aufmerksamkeit zu widmen ist. In Form des elterlichen Berufs kann der sozioökonomische Status eine große Rolle spielen, wenn es um Förderungsmöglichkeiten von Studierenden geht.

Hypothese: Der soziokulturelle Bildungshintergrund des Vaters ist entscheidend für die geistige und bildungstechnische Bildung des studierenden Soldaten.

$$H_0: H_a \neq H_b$$

In Abhängigkeit der jeweiligen Fragestellung werden parametrische bzw. verteilungsgebundene oder non-parametrische bzw. verteilungsfreie statistische Methoden angewendet. Die Zulässigkeit der Applikation parametrischer Verfahren und damit die Genauigkeit der Wahrscheinlichkeitsaussagen hängen von der Erfüllung der Intervallskala und Annahme einer

Normalverteilung ab. Da bei nicht-parametrischen Tests nicht die Messwerte selbst, sondern deren Rangplätze verarbeitet werden, sind sie auch relativ unempfindlich gegenüber Ausreißern (Bortz, 1993). Der Preis für das Ausweichen auf Tests mit schwächeren Anforderungen an die Daten besteht allerdings darin, dass sich mit nicht-parametrischen Tests auch nur weniger scharfe und klare Hypothesen testen lassen. Daher wird man sich bei der Anwendung von verteilungsfreien Tests lediglich auf die Hypothesen bezüglich biographischer und militärraffiner Moderatorvariablen beschränken.

Zunächst werden mit Hilfe des *Kolmogorov-Smirnov-Tests* die Werte der zu untersuchenden Variablen auf ihre Normalverteilung überprüft. Der Test wird anhand der größten (absoluten) Abweichung der empirischen von der theoretischen Verteilung durchgeführt. Die asymptotische Signifikanz entspricht der Irrtumswahrscheinlichkeit und ist mit dem Zurückweisen der Nullhypothese verbunden.

Bei zu geringem Skalenniveau der zu untersuchenden moderierenden Hintergrundvariablen wird für den Vergleich zweier unabhängiger Stichproben der *U-Test nach Mann und Whitney* angewandt. Er prüft die Nullhypothese, ob zwei Stichproben hinsichtlich der Lokation aus derselben Grundgesamtheit stammen. Der *H-Test nach Kruskal und Wallis* ist eine Ausweitung des *U-Tests* beim Vorliegen von mehr als zwei unabhängigen Stichproben. Auch er basiert auf einer gemeinsamen Rangreihe der Werte aller Stichproben. Diese Ein-Weg-Rangvarianzanalyse nach Kruskal und Wallis stellt nach Siegel (1985) das wahrscheinlich effizienteste non-parametrische Verfahren dar, um zu entscheiden, ob k unabhängige Stichproben aus derselben Population stammen.

#### 10.4 Untersuchte Stichprobe

Wegen der zeitlichen Nähe zum Erfolgskriterium Studienerfolg in Form von Vor- und Hauptdiplomnoten entstammen die Testpersonen aus der Gruppe bereits studierender Offizier- und Offizieranwärterjahrgänge. Da es sich bei den Prädiktoren nicht um Wissenstests, sondern allgemeine Intelligenz und studienfeldspezifischen Eignungstests handelt, welche Aufgabe für Aufgabe mögliche unbekannte Größen und Inhalte erklären, sollte der Einfluss möglichen studienbedingten Vorwissens gegenüber Schülern keine große Rolle spielen. Im Vergleich zu untersuchten Schülerstichproben in früheren Jahren durch die Bundesagentur für Arbeit (2004) oder die itb Consulting GmbH konnten hinsichtlich der Anzahl richtiger versus falscher Antworten keine signifikanten Unterschiede festgestellt werden (vgl. Kapitel 11.1). Entscheidend ist, ob die neue Testbatterie zwischen guten und schlechten Studieren-

den bzw. Schülern zu unterscheiden vermag. Eine Messung von originären Offizierbewerbern im Rahmen der täglich laufenden Offizierauswahl in der OPZ wäre im Hinblick auf die dem Studium voran geschaltete und mindestens 15 Monate dauernde militärische Offizierausbildung aus Kosten-Nutzengesichtspunkten ineffizient. Mit Abschluss der Untersuchungen wurden insgesamt 420 Testungen von 320 unterschiedlichen Testpersonen ermittelt. Eine Teilnahme am Testverfahren wurde zwar allen Interessenten vor Ort – demnach auch Angehörigen des Lehrkörpers, Ersttrimesterstudenten, militärischen Vorgesetzten etc. – ermöglicht; in die relevanten Berechnungen flossen jedoch lediglich Ergebnisse der Viert- bis Achtrimesterstudenten ein. An letzteren Probanden orientiert sich die Gesamtstichprobengröße der deskriptiven und Inferenzstatistik in Höhe von  $N = 312$ . Eine Teilnahme an Tests fremder Fachbereiche zwecks Diskriminanzanalyse war gewünscht und wurde betont angeboten, jedoch aufgrund der unverfälschten Freiwilligkeit nicht obligatorisch durchgeführt. Um gehaltvolle Validitätsaussagen zu ermöglichen (Thornton et al. 1992), wird in der vorliegenden Arbeit Wert auf eine große Stichprobe von  $N = 312$  Offizieranwärter der Studienjahrgänge 2001 bis 2004 gelegt. Dies entspricht gewünschten Einstellungsterminen (GEWET) der Jahresspanne 1998 bis 2003.

Die Stichprobe teilt sich entsprechend den Eingangsvoraussetzung in circa 90 Prozent Abiturienten, sechs Prozent Absolventen mit dem Abschluss einer Fachhoch- bzw. höheren Handelsschule gefolgt von einem sehr geringen Anteil an Bewerbern mit mittlerer Reife und erlernter Berufsausbildung (z.B. Einzelhandelskaufmann, Gas-/Wasserinstallateur, Industriekaufmann, Informatiker) bzw. Abschluss in einer Berufsaufbauschule.

Die Aufteilung der Testpersonen auf die einzelnen Studienfelder und Untergliederung in Studienfächer ist in Abbildung 18 skizziert. Demnach sind 33 Prozent den Geisteswissenschaften und 30 Prozent den Wirtschaftswissenschaften zuzuordnen. 30 Prozent der Probanden gehören den Ingenieurwissenschaften und circa sieben Prozent dem Fachbereich Informatik an. Das in der Stichprobe am häufigsten vertretene Studienfach ist Pädagogik (15,7 Prozent), Wirtschafts- und Organisationswissenschaften (10, 8 Prozent) und Betriebswirtschaftslehre (12,5 Prozent inklusive zwei Prozent Fachhochschulstudenten). Aus dem ingenieurwissenschaftlichen Fachbereich nahmen besonders viele Studenten der Studiengänge Elektrotechnik und Luft- und Raumfahrttechnik mit 9,2 und 6,6 Prozent der Gesamtstichprobe an der Testung teil. Die äußerst geringe Studierendenzahl der Fachhochschulstudiengänge wird bei den weiteren Berechnungen den universitären Studiengängen subsumiert

bzw. finden bei Berücksichtigung der Kriterien Vor- und Hauptdiplomnote keinerlei Berücksichtigung.

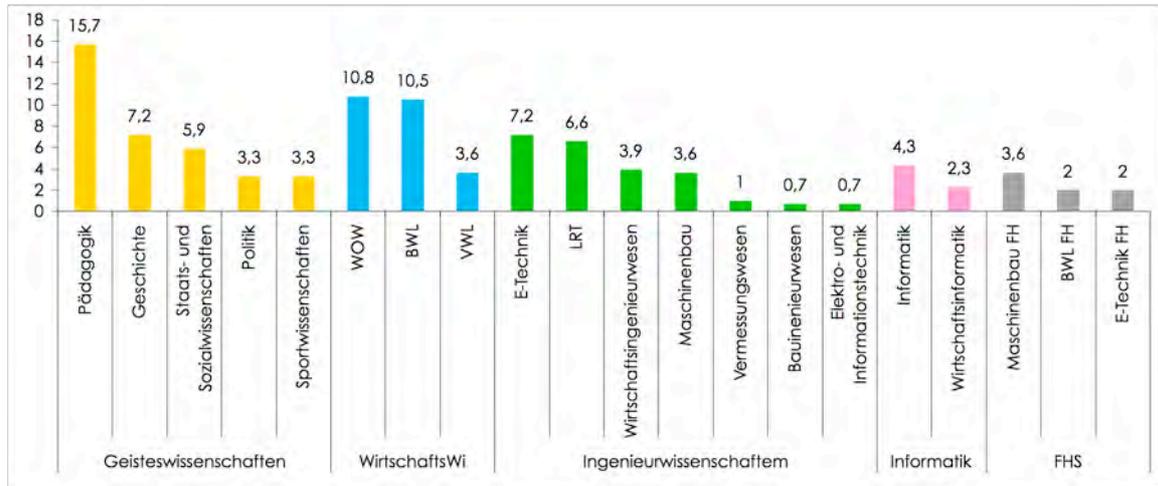


Abbildung 18: Zugehörigkeit der Gesamtstichprobe der Pilotphasen I und II zu Studienfeldern und Studienfächern an den UniBw und der FHS Bw. Angaben in Prozent (WirtschaftsWi: Wirtschaftswissenschaften; FHS: Fachhochschule der Bundeswehr München; E-Technik: Elektrotechnik; LRT: Luft- und Raumfahrttechnik)

#### 10.4.1 Aufbereitung der Stichprobendaten

Bereits mit Beginn der Datenerhebung begannen die Vorbereitungen der Datenbeschaffung so genannter Lebenslaufdaten der Testpersonen. Derartige Daten liefern Informationen zu den Prädiktoren von Berufs- und Studienerfolg, zur Verfahrensgüte der Offizierauswahl, zu Offizierlaufbahnbeurteilungen (Kriterium Berufserfolg), zu soziokulturellem Hintergrund usw. Der frühe Zeitpunkt erwies sich als Vorteil, da es fast zwei Jahre bedurfte, Hürden wie Datenschutzvorschriften zu personenbezogenen Daten, nicht vergleichbare Datenquellen und bereits gelöschte Einzeltestergebnisse der OPZ-Prüfungen sowohl zu durchblicken als auch in Teilen zu überwinden. Beispielsweise sind aufgrund neuester Datenschutzvorschriften wie z.B. der Weisung, Einzeltestergebnisse und vertiefende Prädiktordaten der Offizierauswahlphase bereits nach zwei Jahren zu vernichten, eine Konstruktvalidierung des gesamten ACBw über eine MTMM-Analyse nicht durchführbar.

Für alle Offizieranwärterjahrgänge bis 2001 werden die Daten aus dem Eignungsbeurteilungsverfahren zum einen im Beratungs- und Informationssystem Offizierwachstums (BISON) unter Berücksichtigung des Datenschutzes gespeichert und verwaltet. Nach der Einstellung des Offizieranwärters bis zur Beförderung zum Offizier bzw. Leutnant oder Leutnant zur See ist das Personalstamamt der Bundeswehr für diese Offizieranwärter die personalführende Stelle. Alle während der Verwendung anfallenden persönlichen Daten der Offi-

zieranwärter werden zum anderen für die GEWETs bis 2001 im Dateisystem Personalführungs- und Informationssystem der Bundeswehr (PERFIS) gespeichert. Solche Daten umfassen Grunddaten aus der Offizieranwärterdatei sowie Daten aus Ausbildungsabschnitten, den militärischen Beurteilungen, organisatorischen und strukturellen Daten, sofern zum jeweiligen Zeitpunkt vorhanden. Die in den unterschiedlichen GEWET eingestellten Offizierbewerber wurden in ihrer bisherigen beruflichen Entwicklung betrachtet. Beide Dateien sind dynamisch angelegt, d.h. ihre Datenbestände unterliegen ständiger Änderung. Die in dieser Untersuchung analysierten Daten besitzen den Stand vom 31.02.2007. Vor- und Hauptdiplomnoten wurden unabhängig davon letztmalig im April 2007 direkt über die Prüfungsämter an den Universitäten der Bundeswehr erfragt, an das BMVg übergeben, pseudonymisiert und anhand der vom Datenschützer SKA vorgebenen alphanumerischen Pseudonymisierungsbezeichnung als Schlüsselvariable mit den Daten verknüpft. Im letzten Schritt bestand die Notwendigkeit, alle durch PERFIS und BISON verschlüsselten Variablen und Datenziffern über die jeweiligen Schlüsseldatenbanken der Bundeswehr zu verifizieren.

Erschwerend trat die Tatsache hinzu, dass alle Daten der GEWETe 2002 und jünger nicht mehr in den oben beschriebenen Datenbanken, sondern in der Standardsoftware SASPF<sup>12</sup> und unter neuen Bezeichnungen bzw. auch bei neuen zuständigen Ansprechpartnern im Personalamt der Bundeswehr hinterlegt sind. Diese Verschiedenheit der die Untersuchung voraussetzenden Datensysteme und der Anspruch an einwandfreie verwendbare Daten konnte nur durch zeitaufwendige Recherchen, Datenaufbereitungen und Unterstützung einflussreicher Amtsinhaber auf Ämter- und ministerieller Ebene begegnet werden.

Es wird in dem Zusammenhang darauf hingewiesen, dass im Laufe der Datenanalysen unbegründete Lücken im Datenbestand einzelner Offizieranwärter vorzufinden waren, die in Wirklichkeit bereits hätten erfasst sein müssen. Nachforschungen waren im Rahmen dieser Untersuchung aus rechtlichen Gründen nicht möglich. Da gleichwohl sich ergebende geringere Abweichungen in den einzelnen analysierten Stichproben unter ein bis zwei Prozent lagen, haben sie keine nennenswerten Auswirkungen auf die statistischen Betrachtungen und die daraus zu ziehenden Schlussfolgerungen.

Das wichtigste Kriterium des Studienerfolgs, die Hauptdiplomnote, konnte nicht zu einhundert Prozent ermittelt werden, da teilweise Testpersonen das Studium zum Zeitpunkt der Berechnungen noch nicht abgeschlossen hatten.

---

<sup>12</sup> Standard-Anwendungs-Software-Produkt-Familien (SASPF) zur Einführung betriebswirtschaftlicher Standardsoftware.

Für eine nicht im Vordergrund der Arbeit stehende aber durchaus im weiteren Sinne interessante Prognose des allgemeinen Berufserfolgs des Offiziers im Truppendienst fanden sich im Laufe der Datenanalyse nur unzureichende oder gar keine Hinweise auf kontinuierliche und ausführliche Ausbildungsbeurteilungen. Das einzige zur Verfügung stehende Kriterium stellt die Note aus dem wichtigsten aller Lehrgänge, dem truppengattungsübergreifenden Offizierlehrgang, dar.

## 10.5 Variablenbeschreibung

Zur Bestimmung der prädiktiven bzw. konkurrenten Validität der untersuchten Studieneignungstestbatterie werden in der Ergebnisdarstellung die in Abbildung 19 aufgeführten Prädiktorwerte mit dem Kriteriumswert Diplomnote über Korrelations- und Regressionsanalysen miteinander verglichen. Dabei spielen bei den Prädiktoren nicht nur die erneuerten Verfahren, sondern gleichsam oder insbesondere der Unterschied zwischen neuen und den ursprünglich an der OPZ angewendeten biographischen und psychologischen bzw. Leistungsdaten zur Vorhersage des Studienerfolgs eine große Rolle.

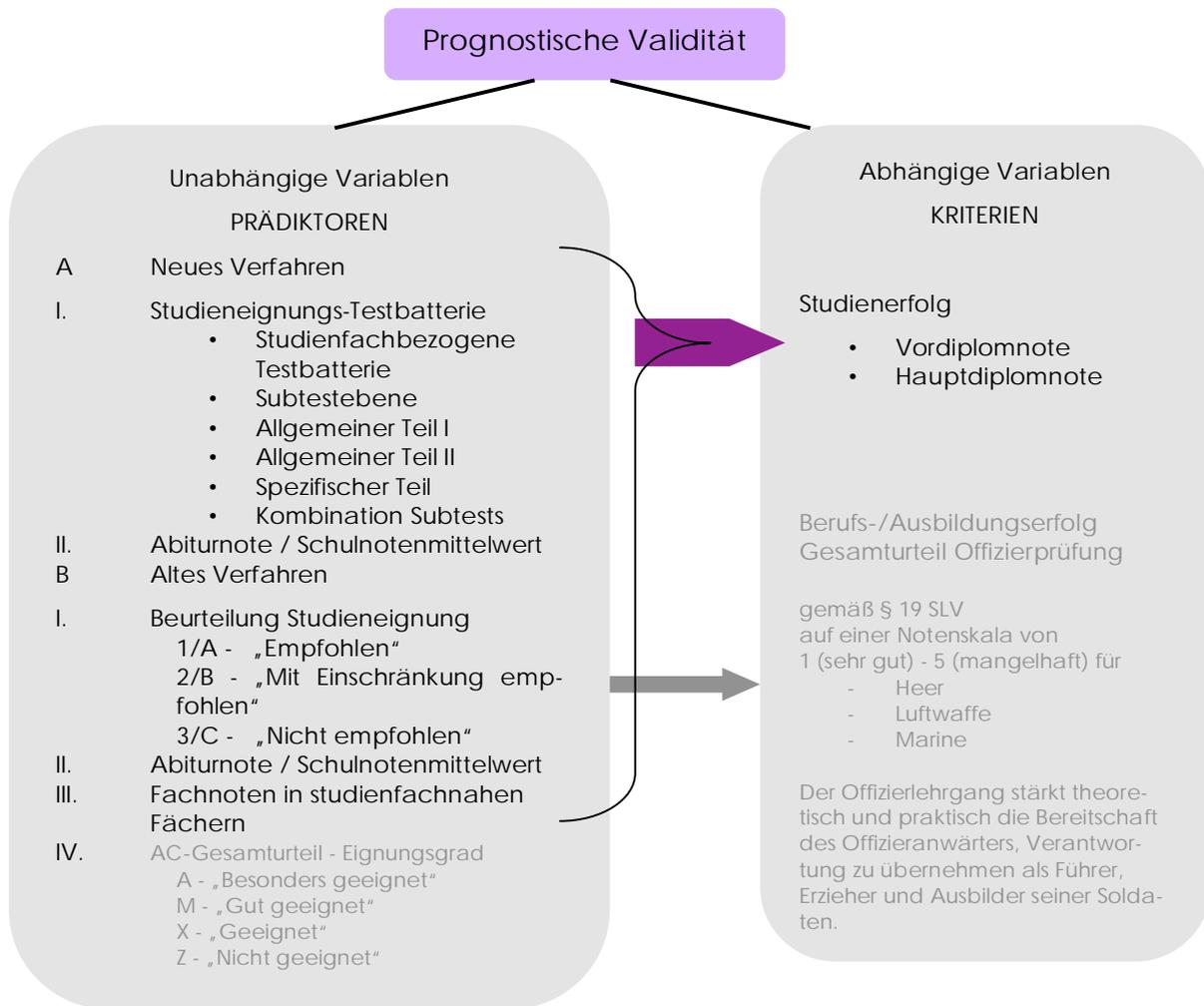


Abbildung 19: Vorhersage des Kriteriumswertes (abhängige Variable) über verschiedene Prädiktorwerte (unabhängige Variable) zur Erfassung der prognostischen Validität im AC Bw und im optimierten Studieneignungstest

Die zu untersuchenden Moderatorvariablen wie auch rein deskriptive Daten lassen sich in die in Tabelle 16 aufgeführten biographische sowie militärraffine Variablen zergliedern.

Tabelle 16:  
Moderatorvariablen in der Studieneignungsvorhersage der Bundeswehr

Biographische Moderatorvariablen	Militärraffine Moderatorvariablen
<ul style="list-style-type: none"> <li>• Geschlecht</li> <li>• Bundesland</li> <li>• Trimesterzahl bis zum Vordiplom</li> <li>• UniBw Hamburg vs. München</li> <li>• Soziokultureller Bildungshintergrund</li> </ul>	<ul style="list-style-type: none"> <li>• Teilstreitkraft</li> </ul>

## 10.5.1 Verteilung der Prädiktoren

Die Prädiktoren sind in ursprüngliche, d.h. bereits während des AC Bw angewendete Leistungsdaten und in optimierte, d.h. im Rahmen der vorliegenden Arbeit entwickelte bzw. neu zusammengestellte Vorhersagevariablen untergliedert. Alte wie neue Auswahlverfahren bzw. Prädiktoren haben gemein, das Kriterium Studienerfolg zu einem möglichst großen Anteil vorherzusagen.

Auf die Inhalte der neuen studienfeldbezogenen Reihungstests wird ausführlich in Kapitel 11.1. eingegangen und anhand der Pilotphasendatenstruktur Normwerte analysiert. Darüber hinaus werden in Tabelle 17 anhand der Aufgabenzahl, Mittelwerte und Varianzen ein Überblick über die Verteilung der Testergebnisse unter den Probanden dargeboten. Die feldspezifischen Subtests weisen bereits jetzt Auffälligkeiten auf, die jedoch einer genaueren Analyse im Ergebnisteil unterzogen werden.

Tabelle 17:  
Überblick Testbatterie Studieneignung über Testergebnisverteilung aller Probanden

<b>Allgemeiner Teil I</b>		<b>Anzahl Aufgaben</b>	<b>Mittelwert Testscore gesamt</b>	<b>Varianz</b>
	<i>Figurenerkennen</i>	30	12,97	<b>28,17</b>
			Note Bw Wehrersatzwesen (1-7): 3,37	
	<i>Verbale Analogien</i>	adaptive Testform	Note Bw Wehrersatzwesen (1-7): 3,07	
<b>Spezifischer Teil</b>				
GeiWi	<i>Sprachensysteme</i>	20	<b>9,67</b>	<b>14,99</b>
	<i>Textverständnis I</i>	5	2,83	1,67
	<i>Textverständnis II</i>	5	1,91	1,16
WiWi	<i>Modellanalyse</i>	20	<b>16,03</b>	<b>28,80</b>
	<i>Diagramme &amp; Tabellen</i>	20	<b>11,92</b>	
IngWi	<i>Formalisieren technischer Zusammenhänge</i>	20	<b>15,18</b>	<b>13,64</b>
	<i>Technisches Verständnis</i>	20	<b>12,73</b>	<b>11,02</b>
Informatik	<i>Flussdiagramme</i>	20	8,26	3,82
	<i>Formale Systeme</i>	20	<b>11,65</b>	<b>10,48</b>
<b>Allgemeiner Teil II</b>				
	<b>Merkmal</b>		<b>Mittelwert Skala</b>	<b>proVarianz</b>
	<i>Persönlichkeitsmerkmale Offizierbewerber</i>			
	Handlungsorientierung		5,01 (5,17)	0,65
	Leistungsmotivation		5,16 (5,22)	0,44
	Risikokompetenz		4,37 (4,36)	0,25
	Bewältigungsverhalten		5,23 (5,24)	0,45
	Teamfähigkeit		4,33 (4,11)	1,06
	Durchsetzungsfähigkeit		4,73 (4,74)	0,29

	Führungspotential	4,79 (4,87)	0,47
	Gewissenhaftigkeit	4,35 (4,48)	0,66
	Selbstdisziplin	4,82 (4,98)	0,61
<i>Big Five</i>			
	Gewissenhaftigkeit	5,20 (5,36)	0,54
	Teamfähigkeit	4,67 (4,49)	0,96
	Neurotizismus	<b>2,89</b> (2,91)	0,59
	Entgegenkommen / Friedfertigkeit	5,03 (4,94)	0,42
	Extraversion	4,61 (4,51)	0,71
	Offenheit für Erfahrungen	4,81 (4,76)	0,35
<i>Stressbewältigungs- ressourcenfragebogen</i>			0,66
	Externale Kontrollüberzeugung	<b>2,78 (2,70)</b>	0,41
	Engagement	<b>5,66 (5,66)</b>	0,39
	Selbstwirksamkeits- überzeugung	<b>5,55 (5,60)</b>	0,45
	Kontrollüberzeugung	5,04 (4,97)	0,67
	Herausforderung	4,94 (4,97)	0,54

*Anmerkungen.* In Klammern zu finden sind Daten der Offiziere mit sehr gutem und guten Berufserfolg bzw. Note 1-2 im Offizierlehrgang.

Ebenso werden die Persönlichkeitsfragebogenergebnisse einer vertiefenden Diagnose in Kapitel 11.1 unterzogen. Die folgende Grafik veranschaulicht das durchschnittliche Persönlichkeitsprofil eines im Studium befindlichen Offizieranwärters.

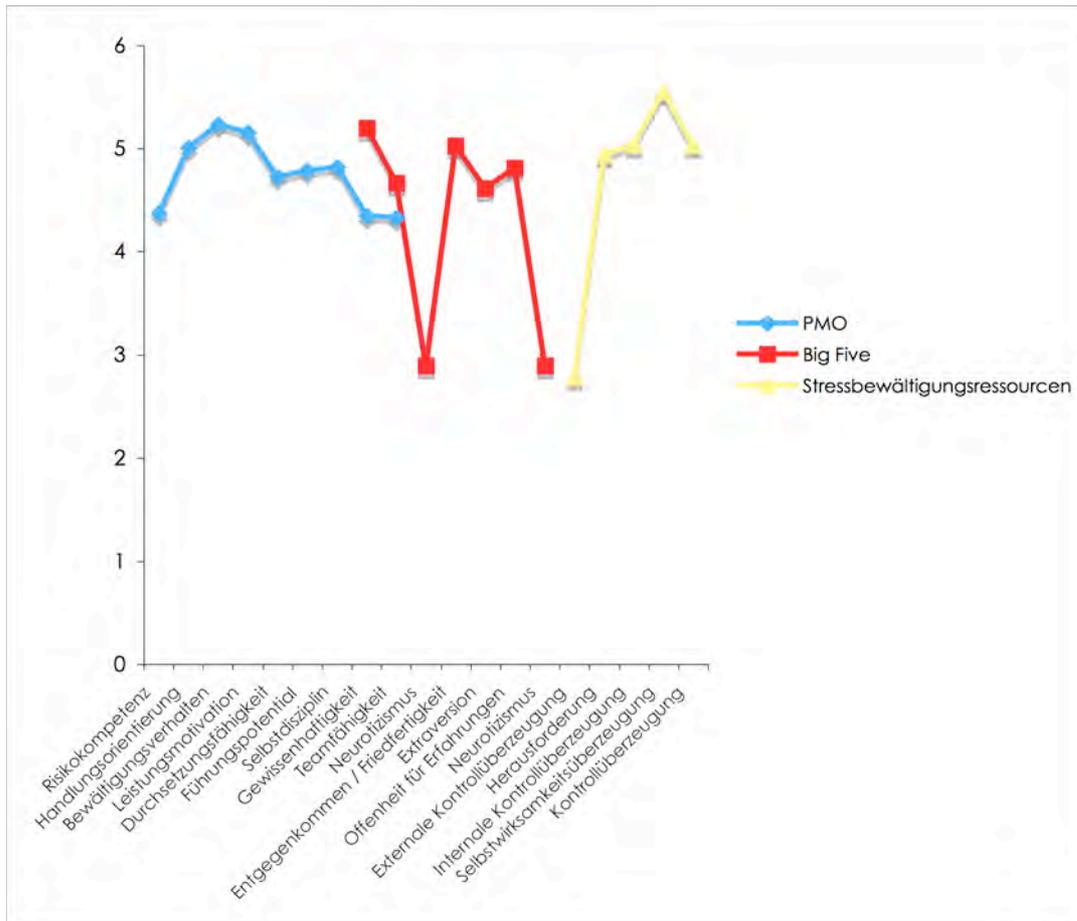
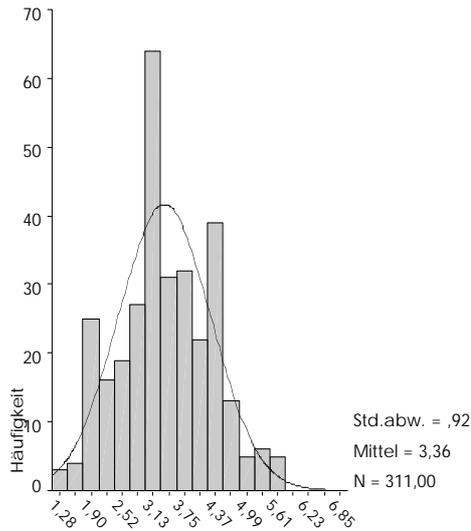


Abbildung 20: Durchschnittliches Persönlichkeitsprofil eines Offizieranwärters

Unter Hinzunahme der in Kapitel 11.1.2 und 11.1.4 aufgezeigten Normwerte für eine zukünftige Praxisanwendung in der OPZ lassen sich die in den nachstehenden Abbildungen veranschaulichten Verteilungen über die fachspezifischen Tests und den Gesamtindex (Allgemeiner Teil I und Spezifischer Teil) darstellen.



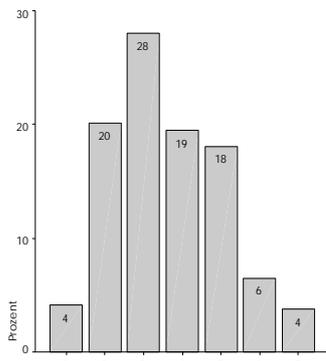
GESAMTINDEX Allgemeiner Teil I und Spezifischer Teil

Abbildung 21: Histogramme der Gesamtindexverteilung (Allgemeiner Teil I und spezifischer Teil) mit Normalverteilungskurve

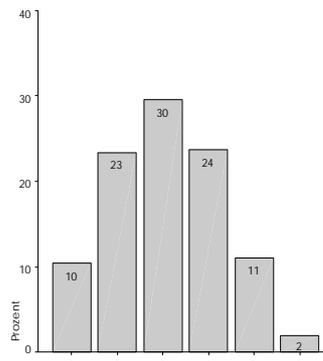
**Allgemeiner Teil I**

Figuren erkennen (n = 293)

Verbale Analogien (adaptive Testform) (n = 308)



MW = 3,62



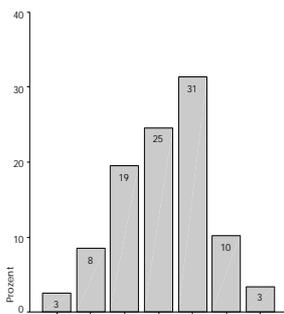
MW = 3,07

Abbildung 22: Verteilung der Prädiktoren (Noten gemäß OPZ-Notenverteilung) im Allgemeinen Teil I der Testbatterie Studieneignung. Anmerkung: MW = Mittelwert

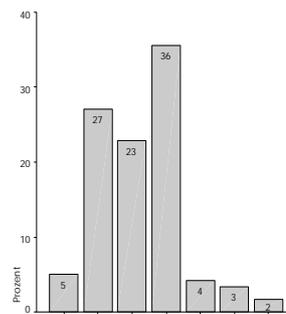
**Spezifischer Teil – Geisteswissenschaften**

Sprachensysteme (n = 118)

Textverständnis I und II (n = 118)



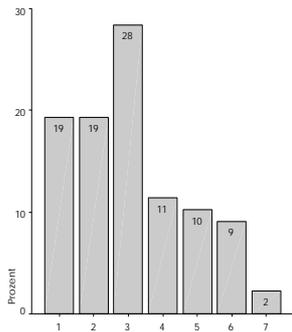
MW = 4,18



MW = 3,24

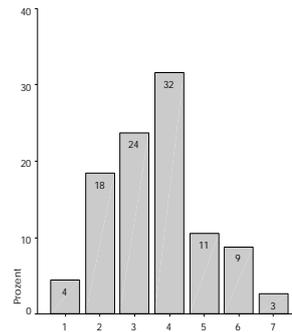
**Spezifischer Teil – Wirtschaftswissenschaften**

Modellanalyse (n = 88)



MW = 3,10

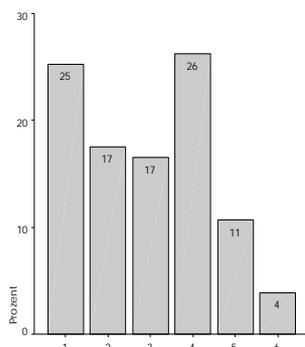
Diagramme und Tabellen (n = 114)



MW = 3,62

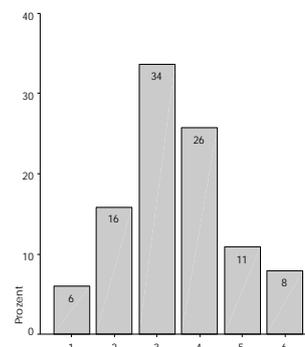
**Spezifischer Teil – Ingenieurwissenschaften**

Formalisieren technischer Zusammenhänge (n = 103)



MW = 2,91

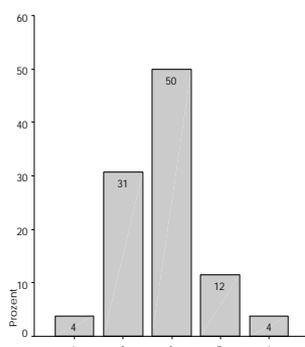
Technisches Verständnis (n = 101)



MW = 3,44

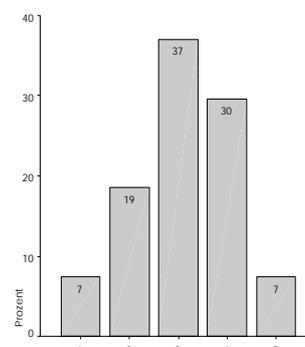
**Spezifischer Teil – Informatik**

Formale Systeme (n = 26)



MW = 2,96

Flussdiagramme (n = 27)



MW = 3,11

Abbildung 23: Verteilung der Prädiktoren (Noten gemäß OPZ-Notenverteilung) im Allgemeinen Teil I der Testbatterie Studieneignung

Vergleicht man die in den folgenden vier Abbildungen dargestellten und nach Studienfeldern gefilterte Gesamtindexnotenverteilungen, bestehen keine signifikanten Mittelwertsunterschiede und beinahe identische Standardabweichungen zwischen den verschiedenen Studienbereichen.

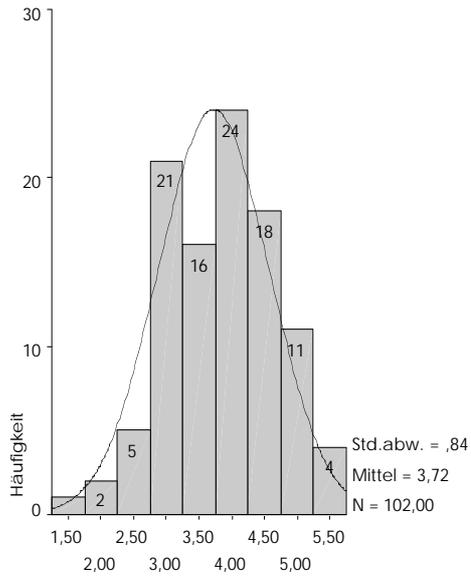


Abbildung 24: Histogramm des Gesamtindex der Testbatterie Studieneignung (Allgemeiner Teil I und spezi- fischer Teil) bezogen auf Studenten der Geisteswissenschaften mit Normalverteilungskurve (n = 102)

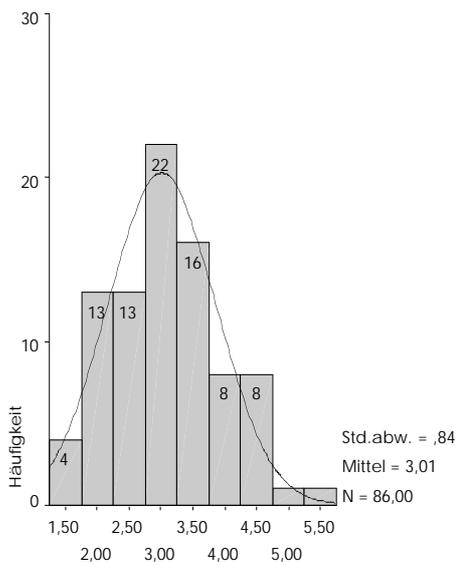


Abbildung 25: Histogramm des Gesamtindex der Testbatterie Studieneignung (Allgemeiner Teil I und spezi- fischer Teil) bezogen auf Studenten der Wirtschaftswissenschaften mit Normalverteilungskurve (n = 86)

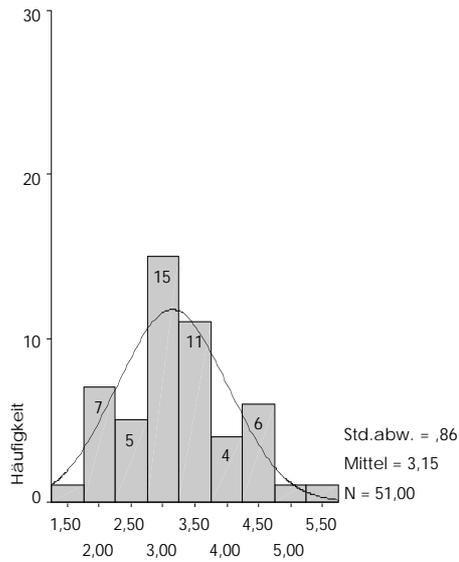


Abbildung 26: Histogramm des Gesamtindex der Testbatterie Studieneignung (Allgemeiner Teil I und spezi-fischer Teil) bezogen auf Studenten der Ingenieurwissenschaften mit Normalverteilungskurve (n = 51)

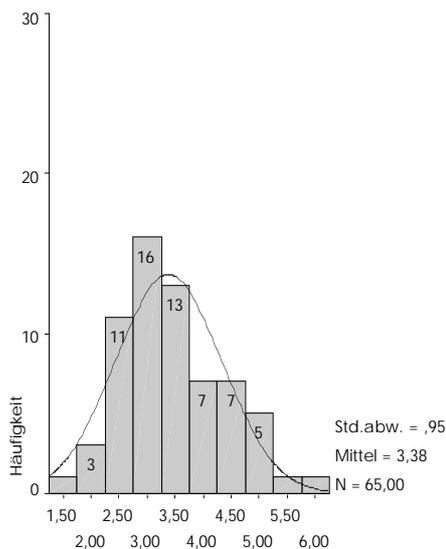


Abbildung 27: Histogramm des Gesamtindex der Testbatterie Studieneignung (Allgemeiner Teil I und spezi-fischer Teil) bezogen auf Studenten der Informatik mit Normalverteilungskurve (n = 65)

Die Schulabschluss- bzw. Abiturnotenverteilung wird im Balkendiagramm der Abbildung 28 veranschaulicht. Die Streuung erweist sich als tendenziell rechtsschief. Mit Hilfe des Kolmogorov-Smirnov-Tests werden die Werte auf Normalverteilung überprüft. Die Voraussetzung einer Normalverteilung ist verletzt bei  $p \leq .05$ . Im vorliegenden Fall der Abiturnotenverteilung liegt mit  $p = .01$  ein deutlich signifikanter p-Wert vor; die Werte sind also nicht hinreichend normalverteilt. Als Folge kommen bei für weitere Berechnungen non-

parametrische Verfahren zum Einsatz. Der durchschnittliche Bewerber für den Offizierberuf verfügt über eine mittlere Abiturnote von 2,3.

Die Schulspornote mit einem Mittel von 1,98 präsentiert sich dem gegenüber gestellt mit eher linksschiefer Verteilungskurve und betont die für den Offizierberuf notwendige Vorbedingung körperlicher Leistungsfähigkeit.

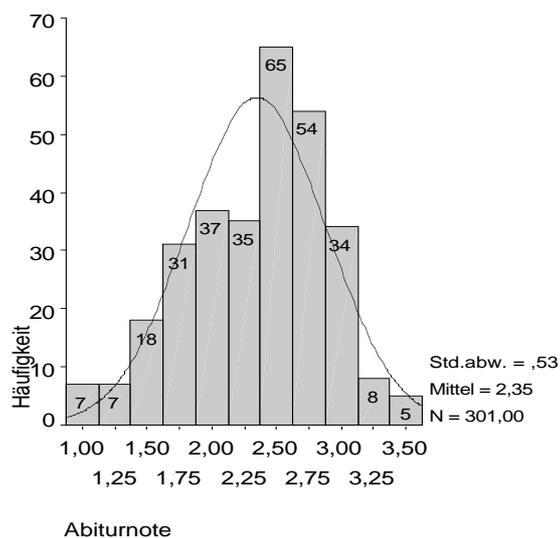


Abbildung 28: Histogramm vorhandener Schulnotenmittelwerte (N = 301) der Testpersonen aus den Pilotphasen I und II mit Normalverteilungskurve

Zu den damals in der OPZ erhobenen Prädiktoren sind neben der Abiturnote insbesondere die Studieneignungsempfehlung des Studienberaters und die Schulfachnoten zu zählen. Ab GEWET 2001 wurde zusätzlich ein mathematischer Kenntnistest eingeführt, dessen Testergebnisdaten jedoch bedauerlicherweise in den Datensätzen nicht mehr vorliegen. Desgleichen wurde Beurteilungen zu Fachnoten in studienfachnahen Fächern gelöscht.

Nach interner Auskunft des Personalamts der Bundeswehr bekommen nur circa 60 Prozent aller Offizierbewerber ihren Erststudienwunsch erfüllt. 40 Prozent wurden „umberaten“ oder auf eine wunschferne Studienfachrichtung eingeplant.

Einer positiven Studieneignungsempfehlung als Voraussetzung für ein erfolgreiches Studieren scheinbar ungeachtet stellt sich die Verteilung der Studieneignungsgrade in Abbildung 29 dar. Die Darstellung verdeutlicht das Missverhältnis zugunsten der nicht für ein Studium geeigneten Offizieranwärter. Nur 27 Prozent aller an den UniBw Studierenden wurde mit 1/A empfohlen, circa die Hälfte aller mit 2/B bzw. mit Einschränkung und circa 11 Prozent nicht für ein Studium empfohlen, jedoch trotzdem – vermutlich der Bedarfsdeckung wegen – eingestellt und einem Studienfach zugewiesen.

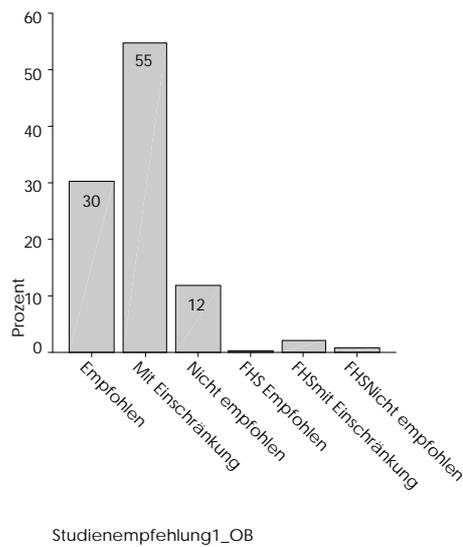


Abbildung 29: Häufigkeitsverteilung Studienempfehlung aller Testpersonen (N = 312) aus den Pilotphasen I und II

Scheinbar nur am Rande betrachtet, aber als nicht nur für den allgemeinen Berufserfolg, sondern auch für die Studieneignung relevant, erweist sich der Gesamteignungsgrad der OPZ. Schließlich ist dies Voraussetzung für die Offizierlaufbahnenempfehlung, ohne die eine Zuweisung zu einem Studienfeld unmöglich ist. Da es sich bei der dieser Arbeit zugrunde liegenden Stichprobe um bereits für die Offizierlaufbahn des Truppendienstes angenommene Offizierbewerber handelt, finden sich nur als geeignet bewertete Bewerber im Datensatz wieder. Abbildung 30 illustriert die Verteilung der ACBw-Eignungsgrade der Gesamtstichprobe. Die an das Urteil angelehnte Merkmalsindexverteilung erweist sich nach Ergebnis des Kolmogorov-Smirnov-Anpassungstest auf Normalverteilung mit  $p = .02$  als nicht normalverteilt. Im vorliegenden Fall der OPZ-Gesamtindexverteilung sind für weitere Berechnungen nonparametrische Verfahren zu verwenden.

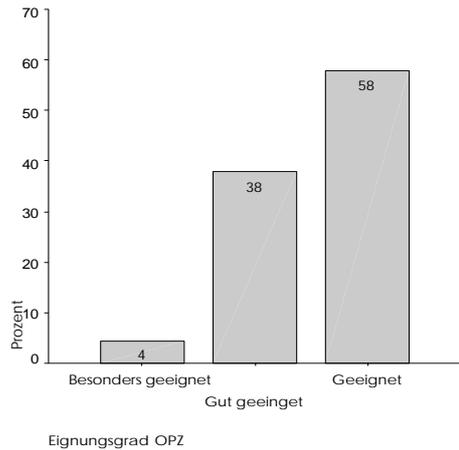


Abbildung 30: Häufigkeitsverteilung des OPZ-Gesamteignungsgrad aller Testpersonen (N = 312) aus den Pilotphasen I und II

### 10.5.2 Verteilung des Kriteriums

Die Kriterien Vor- und Hauptdiplomnote weisen beide eine hohe Varianz auf und unterscheiden sich hinsichtlich ihres Mittelwerts hoch signifikant voneinander. Das Mittel der Vordiplomnoten von 2,66 (n = 309) steht dem der Hauptdiplomnoten in Höhe von 2,06 (n = 148) entgegen. Beide Verteilungen sind gemäß Kolmogorov-Smirnov mit  $p = .14$  (Vordiplom) und  $p = .54$  (Hauptdiplom) normalverteilt. Abbildung 31 und Abbildung 32 stellen die prozentuale Verteilung der Notenschnitte dar.

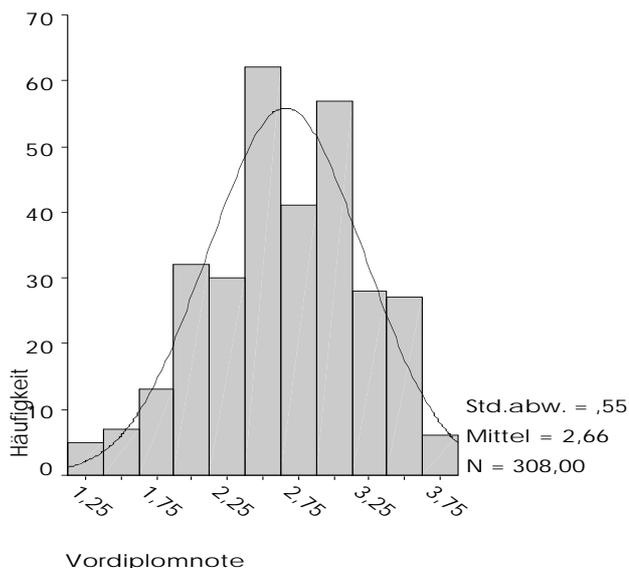


Abbildung 31: Histogramm der Vordiplomnoten der Testpersonen aus den Pilotphasen I und II an den UniBw mit Normalverteilungskurve (n = 308)

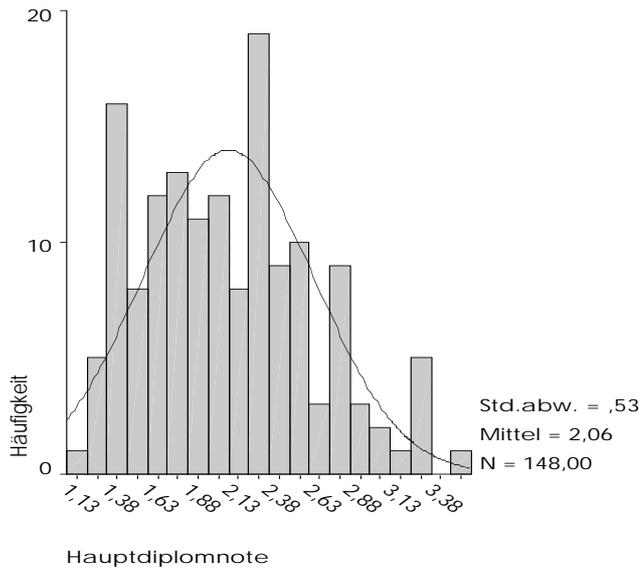


Abbildung 32: Häufigkeits- und Normalverteilung der Vordiplomnoten der Testpersonen (n = 308) aus den Pilotphasen I und II an den UniBw

### 10.5.3 Verteilung der Moderatorvariablen

Gemäß der vom BMVg vorgegebenen Bedarfsdeckungszahlen der einzelnen GEWETe entsprechen die sowohl in Abbildung 33 veranschaulichte Verteilung der TSK Heer, Luftwaffe und Marine als auch Verhältnis von Frauen und Männern im Truppendienst den vorgegebenen Einstellungszahlen und somit einer Zufallsstichprobenverteilung. Die 91 Prozent männlichen und 9 Prozent weiblichen Testpersonen gehören demnach insgesamt zu 57 Prozent dem Heer, zu 30 Prozent der Luftwaffe und zu 13 Prozent der TSK Marine an.

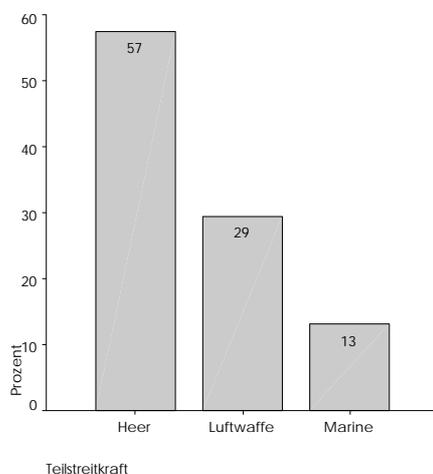


Abbildung 33: Häufigkeitsverteilung Zugehörigkeit Teilstreitkraft aller Testpersonen (N = 312) aus den Pilotphasen I und II

Es wurde Wert gelegt auf eine gleichmäßige Verteilung der Stichprobengröße auf die beiden Universitäten. 48 Prozent der Testpersonen studieren an der Helmut-Schmidt-Universität, UniBw Hamburg, 45 Prozent an der UniBw München, sieben Prozent an der Fachhochschule der Bundeswehr in München.

67 Prozent aller Bewerber stammen aus den alten, die übrigen aus den neuen Bundesländern. 44 der 312 Testteilnehmer waren zum Testzeitpunkt verheiratet, ein Student bereits geschieden.

Das Balkendiagramm in Abbildung 34 skizziert die trotz vorgegebenem Studienablaufplan große Spannweite an Trimesterzahl bis zum Vordiplom. Der Modalwert liegt bei vier Trimestern, der Mittelwert bei 4,4.

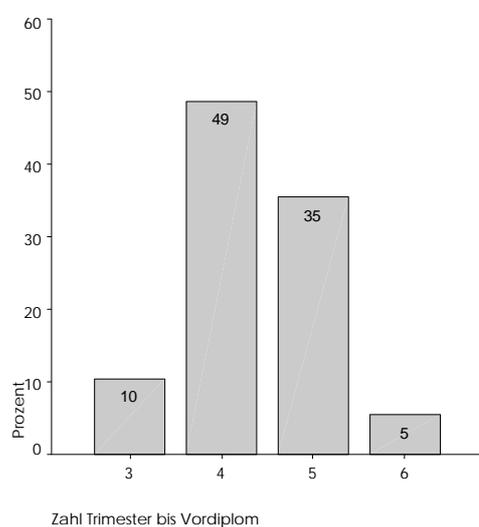


Abbildung 34: Häufigkeitsverteilung der Anzahl Trimester bis Vordiplom aller Testpersonen (N = 312) aus den Pilotphasen I und II

## 11. Ergebnisse

Die Darstellung der Ergebnisse orientiert sich an den in Kapitel 10 beschriebenen Fragestellungen und den daraus resultierenden Annahmen und Hypothesen. Die Testung der Güte und die inhaltlichen Hypothesen zu den Moderatorvariablen geschieht in den Kapiteln 11 entsprechend den beschriebenen methodischen Vorgehensweisen. Die neben der Berechnung der Verfahrensgüte zusätzlich aufgestellten inhaltlichen Hypothesen werden in der unten stehenden Tabelle einschließlich ihrer statistischen Formulierung nochmals übersichtsweise dargestellt.

Tabelle 18:

Die Hypothesen der Evaluation des Eignungsfeststellungsverfahrens an der Offizierbewerberprüfzentrale der Bundeswehr mit statistischer Formulierung

Nr.	Hypothese / Annahme	Statistische Hypothese
1	Der Gesamtttestscore (Ges) sagt den späteren Studienerfolg (StudE) besser vorher als die damalige Studienempfehlung der Offizierbewerberprüfzentrale (OPZ).	$H_0: \rho_{\text{Ges,StudE}} \leq \rho_{\text{OPZ,StudE}}$ $H_1: \rho_{\text{Ges,StudE}} > \rho_{\text{OPZ,StudE}}$
2	Der spezifische Teil der Testbatterie Studieneignung ist hinsichtlich seiner prädiktiven Validität hinsichtlich des Studienerfolgs (StudE) dem allgemeinen Teil der Testbatterie Studieneignung (A) überlegen.	$H_0: \rho_{\text{Spez,StudE}} \leq \rho_{\text{A,StudE}}$ $H_1: \rho_{\text{Spez,StudE}} > \rho_{\text{A,StudE}}$
3	Eine Erhöhung der inkrementellen Validität ist durch die Hinzunahme von Persönlichkeitstests aus dem Allgemeinen Teil II (AII) zu erzielen.	$H_0: \rho_{\text{AI/Spez+AII,StudE}} \leq \rho_{\text{AI/Spez,StudE}}$ $H_1: \rho_{\text{AI/Spez+AII,StudE}} > \rho_{\text{AI/Spez,StudE}}$
4	Durch die Kombination von Reihungstests mit der Abiturdurchschnittsnote (SFT/Abi) wird die prognostische Validität hinsichtlich des Studienerfolgs (StudE) weiter erhöht.	$H_0: \rho_{\text{SFT/Abi,StudE}} \leq \rho_{\text{SFT,StudE}}$ $H_1: \rho_{\text{SFT/Abi,StudE}} > \rho_{\text{SFT,StudE}}$
5	Der Studienerfolg von weiblichen und männlichen Offizieranwärtern unterscheidet sich nicht.	$H_0: U_{\text{männlich}} - U_{\text{weiblich}} = 0$ $H_1: U_{\text{männlich}} - U_{\text{weiblich}} \neq 0$
6	Der Studienerfolg der drei Teilstreitkräfte Heer (H), Luftwaffe (Lw) und Marine (M) unterscheidet sich nicht.	$H_0: H_{\text{He}} = H_{\text{Lw}} = H_{\text{M}}$ $H_1: H_{\text{He}} \neq H_{\text{Lw}} \neq H_{\text{M}}$
7	Der Studienerfolg der UniBw München-Studenten ist höher als der der UniBw Hamburg-Studierenden.	$H_0: H_{\text{UniBwM}} > H_{\text{UniBwH}}$ $H_1: H_{\text{UniBwM}} \leq H_{\text{UniBwH}}$
8	Studenten mit unterschiedlich langer Trimesterzahl bis zum Vordiplom unterscheiden sich nicht hinsichtlich des Studienerfolgs.	$H_0: H_{>4\text{Trimester}} = H_{\leq 4\text{Trimester}}$ $H_1: H_{>4\text{Trimester}} \neq H_{\leq 4\text{Trimester}}$
9	Der soziokulturelle Bildungshintergrund des Vaters ist entscheidend für die geistige und bildungstechnische Bildung des studierenden Soldaten.	$H_0: H_a = H_b$ $H_1: H_a \neq H_b$

## 11.1 Reliabilität

Für jedes Item eines jeden Subtests sind in den unten angehängten Tabellen die Standardabweichung, die Schwierigkeitsindizes und Trennschärfen sowie die Reliabilitäten und deren Mittelwerte der Itemkennwerte über den Test hinweg aufgeführt. Zu Vergleichszwecken werden die mittleren Testgütemaße der BA aus weiter zurückliegenden Untersuchungen mit

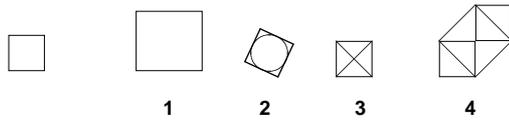
einbezogen. Darüber hinaus veröffentlichte die BA Normen für die Stichprobe der Oberstufenschüler, die unter Berücksichtigung der Normen bereits studierender Offizieranwärter die vorläufige Vergleichsbasis für den OPZ-Einsatz darstellen (vgl. oben). Die Items der Persönlichkeitsfragebögen wurden bereits bei Aufspielung der Aufgaben auf die CAT-Anlage entsprechend ihrer Polung codiert. Vertiefend wird im nächsten Abschnitt kurz der theoretische Hintergrund relevanter Begrifflichkeiten gemäß Lienert (1969) und Lienert & Raatz (1989) geliefert.

### 11.1.1 Itemanalyse Allgemeiner Teil I

In den folgenden Kapiteln werden die einzelnen Subtests des Allgemeinen Teils I näher erläutert und auf Test- wie auch Itemebene hinsichtlich Schwierigkeit, Trennschärfe, Reliabilität und Normierung analysiert.

#### 11.1.1.1 Der Untertest Figuren erkennen

Der Untertest „Figuren erkennen“ erfasst Fähigkeiten wie die flexible Gestaltwahrnehmung, die Wahrnehmungsgeschwindigkeit sowie Konzentrationsvermögen und eignet sich zur Messung allgemeiner Studieneignung. Wie in Abbildung 35 ersichtlich, besteht die Aufgabe des Probanden darin, bestimmte Figuren aus einer sozusagen mehr oder weniger guten Tarnung „heraus zu sehen“. Für 30 Testaufgaben stehen 10 Minuten Zeit zur Verfügung. Der Test wurde Anfang der 90er-Jahre im damaligen Personalstammamt der Bundeswehr (heute: Personalamt der Bundeswehr) von der damaligen CAT-Zentrale für die Anwendung bei Offizierbewerbern zusammengestellt und erprobt. Vorläufer des Tests war die Papier-Bleistift-Testform der Gottschaldt-Figuren, die seit den 70er-Jahren in der fliegerpsychologischen Eignungsfeststellung eingesetzt worden war. Die Daten zur psychometrischen Qualität (Itemkennwerte, Reliabilität) und die Vergleichsmessdaten (Testnormen, deskriptive Statistiken) für die Anwendung des Tests stammen aus Erprobungen und Anwendungen bei speziellen Verwendungen und an den ZNWG. Erkenntnisse aus anderen Stichproben sind im Text kenntlich gemacht.



Ihre Aufgabe ist es zu prüfen, ob die Figur in einer der vier rechten Zeichnungen 1 bis 4 enthalten ist. Haben Sie festgestellt, in welchen Zeichnungen die Figur genau enthalten ist, geben Sie die darunter stehende Zahl bzw. die darunter stehenden Zahlen in aufsteigender Reihenfolge ein!

Abbildung 35: Beispielaufgabe Untertests Figurenerkennen

### 11.1.1.2 Der Untertest Verbale Analogien - adaptive Testform

Der von Professor Hornke, RWTH Aachen, konzipierte adaptive Untertest „Verbale Analogien“ erfasst die Fähigkeit zum verballogischen Denken bzw. Fähigkeit zum Erkennen des sprachlogischen Zusammenhangs zwischen Begriffen (vgl. Abbildung 36).

Verbale Analogie:

Freund : Feind = dick : ?

Lösung:

- 1) breit
- 2) schmal
- 3) dünn
- 4) klein
- 5) tief

Abbildung 36: Beispielaufgabe des ADAPT-Untertests Verbale Analogien

Die Testung beginnt mit einem mittelschwierigen Item, wobei die vorläufige Fähigkeit zu Beginn auf Null gesetzt wird. Löst der Proband dieses richtig oder falsch, sucht der Algorithmus das Item, bei dem der Informationsgehalt maximal ist. Es ist voreingestellt, dass mindestens fünf Items bearbeitet werden müssen. Die maximale Bearbeitungszeit pro Item ist mit drei Minuten voreingestellt.

Das adaptive Testen erzielt höchste Testgenauigkeit auf allen Fähigkeitsniveaus bei größter Testökonomie. Die Aufgaben werden nach ihrer Schwierigkeit und Trennschärfe exakt an die Leistungsfähigkeit des Probanden angepasst, daher genügen im Mittel etwa 8,4 Items

(KWEÄ, 1998) bzw. 10,7 Items (USLw, 1999), um die Fähigkeitsausprägung mit einer Konfidenz von 95 Prozent zu messen. Diese Vorteile einer kürzeren Testzeit bei verbesserter Messgenauigkeit und hoher Motivation der Probanden aufgrund einer leistungsadäquaten Aufgabenauswahl wurden durch einen ersten Vergleich von adaptivem mit konventionellem Testen seitens des Streitkräfteamtes überprüft (Storm, 1998). Das adaptive Testen kann seine Vorteile nur ausspielen, wenn die Software regelmäßig gepflegt und infolge dessen die Qualität sichernder Arbeiten gegebenenfalls modifiziert wird. So müssen zum Beispiel die Start-Items periodisch geändert werden.

Der Vorteil des adaptiven Testmodus ist, dass die gewünschte Messgenauigkeit, der Standardmessfehler, der für den Theta-Wert des Probanden zugelassen wird, genau eingestellt werden kann. Für die Anwendung ist je adaptive Testung eines Probanden eingestellt, dass der resultierende Messwert einen Standardmessfehler von höchstens 0,387 hat, was einer Reliabilität von 0,840 entspricht. Die Testung bei einem Probanden endet, wenn diese Messgenauigkeit erreicht ist, wobei mindestens fünf Items bearbeitet sein müssen. Wenn bereits bei vier bearbeiteten Items der Theta-Wert mit der voreingestellten Messgenauigkeit von 0,387 erreicht ist, so wird dennoch das fünfte Item zur Bearbeitung vorgegeben, was u. U. wieder zur Unterschreitung des Theta-Wertes nach dem fünften Item führen kann. Bei einer Erprobungsstichprobe von 9841 Probanden der KWEÄ zeigte sich, dass die adaptiven Untertests bei einem Viertel der Probanden bereits nach 5 Items mit einem Theta-Wert der voreingestellten Messgenauigkeit beendet waren (Storm, 1998).

### 11.1.1.3 Testgütemaße

Wie in den folgenden Tabellen zu entnehmen, beträgt die interne Konsistenz für den Allgemeinen Teil I über beide Pilotphasen hinweg 0,946. Die internen Konsistenzwerte für die einzelnen Tests variieren je nach Pilotphase und Test zwischen 0,403 und 0,877.

Tabelle 19:

*Übersicht über die internen Konsistenzkoeffizienten für den allgemeinen Teil der Testbatterie Geisteswissenschaften bezogen auf beide Pilotphasen ohne Untertest Figurenreihen*

Reliabilität	Cronbachs Alpha
Figuren erkennen	0,869
Verbale Analogien adaptive Testform	0,840
<b>Allgemeiner Teil I</b>	<b>0,946</b>

Die Testgütemaße pro Pilotphase bzw. über beide Pilotphasen hinweg pro Test sind in den folgenden Abschnitten und Tabellen aufgeführt.

## Pilotphase I UniBw Hamburg

Tabelle 20:  
Kennwerte des Tests Figuren erkennen bezogen auf Pilotphase I

„Figuren erkennen“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw HH N = 156	0,408	0,392	0,877

Item	Standardabweichung	Schwierigkeitsindex	Trennschärfe
1	0,5052	0,5405	0,3838
2	0,3466	0,1351	<b>0,0189<sup>13</sup></b>
3	0,5022	0,5676	0,3753
4	0,4634	0,7027	0,7130
5	0,1644	0,0270	<b>0,1578</b>
6	0,4502	0,7297	0,7247
7	0,4977	0,5946	0,4549
8	0,5052	0,5405	0,3933
9	0,4917	0,6216	0,7002
10	0,5052	0,5405	0,3274
11	0,3466	<b>0,1351</b>	0,3080
12	0,2767	<b>0,0811</b>	<b>-0,2158</b>
13	0,4502	0,2703	0,4551
14	0,4746	0,6757	0,4171
15	0,5067	0,5135	0,4794
16	0,5067	0,5135	0,5759
17	0,4634	0,7027	0,5850
18	0,4634	0,2973	0,2621
19	0,5022	0,4324	0,4059
20	0,4634	0,7027	0,7237
21	0,5022	0,4324	0,5799
22	0,5052	0,5405	0,5466
23	0,4634	0,2973	0,4766
24	0,5067	0,5135	0,5759
25	0,2292	<b>0,0541</b>	<b>0,1100</b>
26	0,3466	<b>0,1351</b>	0,3348
27	0,4917	0,3784	0,1465
28	0,2292	<b>0,0541</b>	<b>0,1297</b>
29	0,5022	0,4324	0,4154
20	0,2767	<b>0,0811</b>	0,1876

Tabelle 21:  
Cronbachs Alpha des Tests Verbale Analogien bezogen auf Pilotphase I

„Verbale Analogien“	Cronbachs Alpha
UniBw HH N = 156	Voreingestellt 0,84

<sup>13</sup> Fett gedruckt sind kritische und zu löschende Items.

**Pilotphase II UniBw München**

Tabelle 22:

*Kennwerte des Tests Figuren erkennen bezogen auf Pilotphase II*

„Figuren erkennen“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw M N = 163	0,473	0,379	0,858

Item	Standardabweichung	Schwierigkeitsindex	Trennschärfe
1	,4767	,6667	,1806
2	,4045	,2000	,2916
3	,5055	,5111	,2349
4	,4584	,7111	,3223
5	,3438	<b>,1333</b>	<b>,0856</b>
6	,3438	<b>,8667</b>	,4603
7	,5025	,5556	,4708
8	,5055	,5111	,2190
9	,4045	,8000	,6017
10	,4682	,6889	,2879
11	,3866	<b>,1778</b>	,2412
12	,3665	<b>,1556</b>	,2747
13	,4346	,2444	,3879
14	,4472	,7333	,5352
15	,5045	,5333	,4533
16	,4903	,6222	,5419
17	,4472	,7333	,6200
18	,5055	,5111	,4619
19	,4903	,6222	,4055
20	,4346	,7556	,5054
21	,4682	,3111	,4576
22	,4954	,6000	,5191
23	,4204	,2222	,4470
24	,5045	,5333	,5282
25	,3866	<b>,1778</b>	,2515
26	,2878	<b>,0889</b>	,2925
27	,5025	,4444	,1706
28	,4204	,2222	,4470
29	,4682	,6889	,4629
30	,3866	<b>,1778</b>	,2002

Tabelle 23:

*Cronbachs Alpha des Tests Verbale Analogien bezogen auf Pilotphase II*

„Verbale Analogien“	A-	Cronbachs Alpha
UniBw M N = 163	Voreingestellt	0,84

## Pilotphase I und II

Tabelle 24:

*Kennwerte des Tests Figuren erkennen bezogen auf Pilotphase I und II*

„Figuren erkennen“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw M N = 319	0,444	0,390	0,869

Item	Standardabweichung	Schwierigkeitsindex	Trennschärfe
1	0,4908	0,6098	0,2910
2	0,3786	0,1707	0,1841
3	0,5017	0,5366	0,2848
4	0,4578	0,7073	0,4959
5	0,2811	<b>0,0854</b>	<b>0,1261</b>
6	0,3987	<b>0,8049</b>	0,6050
7	0,4977	0,5732	0,4494
8	0,5025	0,5244	0,2897
9	0,4520	0,7195	0,6597
10	0,4879	0,6220	0,3230
11	0,3675	<b>0,1585</b>	0,2748
12	0,3293	<b>0,1220</b>	<b>0,0932</b>
13	0,4392	0,2561	0,4084
14	0,4578	0,7073	0,4815
15	0,5025	0,5244	0,4620
16	0,4977	0,5732	0,5649
17	0,4520	0,7195	0,6004
18	0,4957	0,4146	0,3908
19	0,5017	0,5366	0,4223
20	0,4458	0,7317	0,6093
21	0,4846	0,3659	0,4814
22	0,4977	0,5732	0,5335
23	0,4392	0,2561	0,4379
24	0,5025	0,5244	0,5456
25	0,3292	<b>0,1220</b>	0,2186
26	0,3145	<b>0,1098</b>	0,2956
27	0,4957	0,4146	<b>0,1676</b>
28	0,3556	<b>0,1463</b>	0,3492
29	0,4977	0,5732	0,4582
30	0,3429	<b>0,1341</b>	0,2095

Tabelle 25:

*Kennwerte des Tests Figurenreihen bezogen auf Pilotphase I und II*

„Figurenreihen“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw M N = 319	0,846	0,113	0,403

Item	Standardabweichung	Schwierigkeitsindex	Trennschärfe
------	--------------------	---------------------	--------------

1	0,1358	<b>0,9813</b>	<b>-0,0132</b>
2	0,1808	<b>0,9663</b>	<b>0,0296</b>
3	0,1485	<b>0,9775</b>	<b>-0,0080</b>
4	0,0864	<b>0,9925</b>	<b>0,0653</b>
5	0,4894	0,6067	<b>0,2871</b>
6	0,1991	<b>0,9588</b>	<b>0,0572</b>
7	0,0000	<b>1,0000</b>	
8	0,0000	<b>1,0000</b>	
9	0,2233	<b>0,9476</b>	<b>0,0194</b>
10	0,3117	<b>0,8914</b>	<b>-0,0099</b>
11	0,3070	<b>0,8951</b>	0,2471
12	0,2811	<b>0,9139</b>	0,3047
13	0,2697	<b>0,9213</b>	<b>0,0813</b>
14	0,1358	<b>0,9813</b>	<b>0,1571</b>
15	0,5009	0,5056	<b>0,0285</b>
16	0,0864	<b>0,9925</b>	<b>0,0653</b>
17	0,3254	<b>0,8801</b>	<b>0,0800</b>
18	0,3648	<b>0,8427</b>	0,2155
19	0,4637	0,6891	0,1920
20	0,4503	0,2809	0,2314

Tabelle 26:

*Kennwerte des Tests Verbale Analogien bezogen auf Pilotphase I und II*

„Verbale A- nalogien“		Cronbachs Alpha
UniBw M N = 319	Voreingestellt	0,84

### 11.1.2 Normen Allgemeiner Teil I

In den folgenden Tabellen werden ermittelte Normen für die jeweilige Stichprobe (Studenten UniBw, BA und (abgeleitet) OPZ-Bewerber) dargestellt.

Tabelle 27:

*Normen der Tests Figuren erkennen und Verbale Analogien*

**Notentabelle für Figuren erkennen HH**  
N = 156

Note	7	6	5	4	3	2	1
Anzahl	0	2	6	11	16	19	21
Richtige	bis 1	bis 5	bis 10	bis 15	bis 18	bis 20	bis 30

**Notentabelle für Figuren erkennen HH und M**  
N = 319

Note	7	6	5	4	3	2	1
Anzahl	0	3	7	12	16	19	22
Richtige	bis 2	bis 6	bis 11	bis 15	bis 18	bis 21	bis 30

**Notentabelle für Figuren erkennen M**  
N = 163

Note	7	6	5	4	3	2	1
Anzahl	0	4	8	13	16	20	24
Richtige	bis 3	bis 7	bis 12	bis 15	bis 19	bis 23	bis 30

**Notentabelle für Figuren erkennen ZNWG 2002**  
N = 456

Note	7	6	5	4	3	2	1
Anzahl	0	2	4	7	11	15	18
Richtige	bis 1	bis 3	bis 6	bis 10	bis 14	bis 17	bis 30

Notentabelle für Figuren erkennen **OPZ**

Note	7	6	5	4	3	2	1
Anzahl	0	3	6	10	13	17	20
Richtige	bis 2	bis 5	bis 9	bis 12	bis 16	bis 19	bis 30

Cut-off-Theta-Werte für Untertest Verbale Analogaen **OPZ**

Notentabelle für Verbale Analogien HH

N = 156

SW = 4 - 4/3 z							
Note	7	6	5	4	3	2	1
Anzahl	∞	,055	,721	1,147	1,639	2,160	2,478
Richtige	bis ,054	bis ,720	bis 1,146	bis 1,638	bis 2,159	bis 2,478	bis ∞

Notentabelle für Verbale Analogien M

N = 163

SW = 4 - 4/3 z							
Note	7	6	5	4	3	2	1
Anzahl	∞	-,032	,377	1,095	1,654	2,009	2,500
Richtige	bis -,033	bis ,376	bis 1,094	bis 1,653	bis 2,008	bis 2,500	bis ∞

Notentabelle für Verbale Analogien HH und M

N = 319

SW = 4 - 4/3 z							
Note	7	6	5	4	3	2	1
Anzahl	∞	,049	,555	1,135	1,638	2,085	2,501
Richtige	bis ,048	bis ,554	bis 1,134	bis 1,637	bis 2,084	bis 2,500	bis ∞

Notentabelle für Verbale Analogien OB 2002

N = 6333

SW = 4 - 4/3 z							
Note	7	6	5	4	3	2	1
Anzahl	∞	-,391	-,058	,559	1,148	1,635	2,142
Richtige	bis -,391	bis -,059	bis ,558	bis 1,147	bis 1,635	bis 2,142	bis ∞

Notentabelle für Verbale Analogien **OPZ**

SW = 4 - 4/3 z							
Note	7	6	5	4	3	2	1
Anzahl	∞	0,1	0,31	0,71	1,31	1,71	2,3
Richtige	bis 0	bis 0,3	bis 0,7	bis 1,3	bis 1,7	bis 2,2	bis ∞

### 11.1.3 Itemanalyse Spezifischer Teil

In den folgenden Kapiteln werden die einzelnen Subtests der studienfeldspezifischen Testbatterieanteile näher erläutert und auf Test- wie auch Itemebene hinsichtlich Schwierigkeit, Trennschärfe, Reliabilität und Normierung analysiert.

#### 11.1.3.1 Studienfeld Geisteswissenschaften - Struktur und Messbereiche

Die Geisteswissenschaften gelten aufgrund ihrer heterogenen Zusammensetzung am schwersten mit Testaufgaben zu fassendste Studienfeld. Allein die UniBw verfügen über einen Fächerkanon, welcher angefangen von Pädagogik über Staatswissenschaften hin zu Sport- und Geschichtswissenschaften reicht.

Für die Pilotphasen wurde sich in Anlehnung an die Anforderungsanalyseergebnisse nach Kastner (2004) und nach Rücksprache mit Experten folgender Untertest der Studienberatungs-Testserie Neuphilologien ausgewählt:

„Sprachensysteme“, bestehend aus 20 Aufgaben, Bearbeitungszeit 44 Minuten

Wie in Abbildung 37 ersichtlich, sind im Test „Sprachensysteme“ aus vorgegebenen Ausdrücken einer fiktiven Sprache Wortbedeutungen und grammatikalische Regeln abzuleiten. Dies erfordert die Fähigkeit zum Erkennen von Gesetzmäßigkeiten und Strukturen in Sprachmustern.

**Ausdrücke:**

Lolem neung	=	Lebende Sprachen
Saam kop hasip	=	Menschen und Tiere
Lolem kop saam	=	Sprachen und Menschen

**„lebende Tiere“  
heißt in der fremden Sprache:**

**Lösung:**

- 1 hasip saam
- 2 hasip neung
- 3 neun hasip
- 4 saam neung
- 5 neung kop

Abbildung 37: Beispielaufgabe des Untertests Sprachensysteme

Aufgrund von Empfehlungen der Professoren entstand außerdem ein vom SKA konstruierter

„Textverständnis“, bestehend aus 2 Texten (Textverständnis I und II) mit jeweils fünf Aufgaben.

Dieser Subtest soll die Fähigkeit messen, umfangreiches Textmaterial mit seinen komplexen Informationen sachgerecht zu erfassen. Aufgrund der Textlänge und der nur in geringem Maße zur Verfügung stehenden Lesezeit kommt es entscheidend darauf an, in kurzer Zeit die wichtigen Informationen von den unwichtigen Informationen zu trennen. Bei der Beantwortung der auf den Text bezogenen Fragen bzw. Ableitungen soll dabei nicht auf eventuell vorhandenes Wissen zurückgegriffen werden können (vgl. Abbildung 38).

**Aussagen zu Text II**

I. Hypothesen lassen sich überprüfen in dem man ihren Wahrheitsgehalt der Falsifikation unterzieht.

II. „Tautologie“ ist ein außerempirischer logischer Begriff, weil er einen Satz ohne Rücksicht auf Änderungen der empirischen Welt kennzeichnet.

**Lösung:**

1 Nur Aussage I lässt sich ableiten.

2 Nur Aussage II lässt sich ableiten.

3 Beide Aussagen lassen sich ableiten.

4 Keine der beiden Aussagen lässt sich ableiten.

Abbildung 38: Beispielaufgabe des Untertests Textverständnis I (bezogen auf den zugrunde liegenden Text)

### 11.1.3.2 Testgütemaße Geisteswissenschaften

Die interne Konsistenz beträgt für den gesamten Spezifischen Teil Geisteswissenschaften bezogen auf beide Pilotphasen 0,742 (vgl. Formel S. 11). Die internen Konsistenzwerte für die einzelnen Tests variieren je nach Pilotphase und Test zwischen 0,142 und 0,791.

Tabelle 28:

*Übersicht über die internen Konsistenzkoeffizienten für den Spezifischen Teil der Testbatterie Geisteswissenschaften bezogen auf beide Pilotphasen*

Reliabilität	Cronbachs Alpha
Sprachensysteme	0,772
Textverständnis I	0,437
Textverständnis II	0,142
<b>Spezifischer Teil Geisteswissenschaften</b>	<b>0,742</b>

Der Subtest Textverständnis II liefert weder dem Personalpsychologischen Qualitätssicherungskonzept der Bundeswehr noch den Bestimmungen der DIN 33430 entsprechende Itemkennwerte. Bei Nicht-Berücksichtigung dieses Untertests steigt die Gesamttestreliabilität von 0,742 auf ein Cronbachs Alpha von 0,773.

Die Testgütemaße pro Pilotphase und Test sind in den folgenden Abschnitten aufgeführt.

### Pilotphase I UniBw HH

Tabelle 29:

*Kennwerte des Tests Sprachensysteme bezogen auf Pilotphase I*

„Sprachensysteme“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw HH N = 75	0,452	0,346	0,790
BA= 295	0,585	0,341	0,791

Item	Standard- abweichung	Schwierigkeits-index	Trennschärfe
1	0,3551	<b>0,8548</b>	<b>0,2696</b>
2	0,3551	<b>0,8548</b>	<b>0,2233</b>
3	0,4576	0,2903	0,4002
4	0,4824	0,6452	0,5493
5	0,4648	0,3065	0,4649
6	0,4911	0,3871	0,3363
7	0,4999	0,5645	0,5141
8	0,4771	0,3387	<b>-0,1457</b>
9	0,4824	0,3548	0,4467
10	0,4648	0,3065	0,3072
11	0,4975	0,5806	0,5265
12	0,5038	0,5161	0,5541
13	0,4911	0,3871	0,5513
14	0,4911	0,3871	0,4514
15	0,5041	0,5000	<b>0,0000</b>
16	0,4999	0,5645	0,3730
17	0,4945	0,4032	0,3058
18	0,4104	0,2097	0,5064
19	0,3380	<b>0,1290</b>	<b>0,1016</b>
20	0,5030	0,4677	<b>0,1863</b>

Tabelle 30:

*Kennwerte des Tests Textverständnis I bezogen auf Pilotphase I*

„Textverständnis I“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw HH N = 75	0,528	0,160	0,324

Item	Standard- abweichung	Schwierigkeitsindex	Trennschärfe
1	0,4932	0,6000	<b>0,0819</b>
2	0,4997	0,4400	<b>0,1975</b>
3	0,4832	0,3600	0,2038
4	0,3562	0,8533	<b>0,1960</b>
5	0,4903	0,3867	<b>0,1217</b>

Tabelle 31:

*Kennwerte des Tests Textverständnis II bezogen auf Pilotphase I*

„Textverständnis II“	Mittlere Schwierig- keit	Mittlere korrigierte Trennschärfe	Cronbachs Alpha
UniBw HH N = 75	0,3787	0,074	0,142

Item	Standard- abweichung	Schwierigkeits- index	Trennschärfe
1	0,3691	<b>0,1600</b>	0,2467
2	0,4791	0,6533	<b>0,1001</b>
3	0,4979	0,4267	<b>-0,1063</b>
4	0,5022	0,4667	<b>0,0914</b>
5	0,3923	<b>0,1867</b>	<b>0,0388</b>

### Pilotphase II UniBw M

Tabelle 32:

*Kennwerte des Tests Sprachensysteme bezogen auf Pilotphase II*

„Sprachensysteme“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw M N = 47	0,502	0,278	0,720
BA= 295	0,585	0,341	0,791

Item	Standard- abweichung	Schwierigkeits- index	Trennschärfe
1	0,3051	<b>0,9000</b>	<b>-0,2537</b>
2	0,3051	<b>0,9000</b>	<b>0,0727</b>
3	0,4901	0,3667	0,4535
4	0,4983	0,6000	0,2532
5	0,4795	0,3333	0,3422
6	0,5074	0,4667	0,1500
7	0,4983	0,6000	0,5423
8	0,4795	0,3333	<b>0,1773</b>
9	0,4795	0,3333	0,4273
10	0,5074	0,5333	0,2840
11	0,4795	0,6667	0,3562
12	0,5040	0,5667	0,4082
13	0,4983	0,4000	<b>0,0074</b>
14	0,5074	0,4667	0,3450
15	0,5040	0,4333	0,3905

16	0,4983	0,6000	0,3949
17	0,4795	0,3333	0,3633
18	0,4661	0,3000	<b>0,1288</b>
19	0,4302	0,2333	0,2289
20	0,4795	0,6667	0,4417

Tabelle 33:

*Kennwerte des Tests Textverständnis I bezogen auf Pilotphase II*

„Textverständnis I“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw M N = 47	0,630	0,335	0,552

Item	Standardabweichung	Schwierigkeitsindex	Trennschärfe
1	0,3977	<b>0,8085</b>	0,2412
2	0,5025	0,5532	0,3076
3	0,5025	0,4468	0,3431
4	0,2040	<b>0,9574</b>	0,3862
5	0,4914	0,3830	0,3973

Tabelle 34:

*Kennwerte des Tests Textverständnis II bezogen auf Pilotphase II*

„Textverständnis II“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw M N = 47	0,379	0,075	0,164

Item	Standardabweichung	Schwierigkeitsindex	Trennschärfe
1	0,4408	0,2553	<b>-0,0329</b>
2	0,4623	0,7021	<b>0,1400</b>
3	0,4623	0,2979	<b>0,1389</b>
4	0,5053	0,5106	<b>0,0377</b>
5	0,3373	<b>0,1277</b>	<b>0,0921</b>

## Pilotphasen I und II

Tabelle 35:

*Kennwerte des Tests Sprachensysteme bezogen auf Pilotphase I und II*

„Sprachensysteme“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw N = 122	0,469	0,330	0,772
BA= 295	0,585	0,341	0,791

Item	Standard- abweichung	Schwierigkeits- index	Trennschärfe
1	0,3386	<b>0,8696</b>	0,1300
2	0,3386	<b>0,8696</b>	0,1878
3	0,4671	0,3152	0,4200
4	0,4853	0,6304	0,4464
5	0,4671	0,3152	0,4264
6	0,4951	0,4130	0,2939
7	0,4969	0,5761	0,5218
8	0,4753	0,3370	0,0404
9	0,4789	0,3478	0,4343
10	0,4882	0,3804	0,3103
11	0,4907	0,6087	0,4806
12	0,5017	0,5326	0,5110
13	0,4907	0,3913	0,3778
14	0,4951	0,4130	0,4225
15	0,5023	0,4783	0,1036
16	0,4969	0,5761	0,3802
17	0,4882	0,3804	0,3103
18	0,4289	0,2391	0,3852
19	0,3714	<b>0,1630</b>	0,1555
20	0,5017	0,5326	0,2705

Tabelle 36:

*Kennwerte des Tests Textverständnis I bezogen auf Pilotphase I und II*

„Textverständnis I“	Mittlere Schwierig- keit	Mittlere Trennschär- fe	Cronbachs Alpha
UniBw N = 122	0,567	0,233	0,437

Item	Standard- abweichung	Schwierigkeits- index	Trennschärfe
1	0,4683	0,6803	0,1607
2	0,5018	0,4836	0,2557
3	0,4905	0,3734	0,2702
4	0,3098	<b>0,8934</b>	0,2606
5	0,4887	0,3852	0,2186

Tabelle 37:

*Kennwerte des Tests Textverständnis II bezogen auf Pilotphase I und II*

„Textverständnis II“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw N = 122	0,379	0,064	0,142

Item	Standard- abweichung	Schwierigkeits- index	Trennschärfe
1	0,3992	<b>0,1967</b>	<b>0,1158</b>
2	0,4714	0,6721	<b>0,1134</b>

3	0,4866	0,3770	<b>-0,0309</b>
4	0,5018	0,4836	<b>0,0694</b>
5	0,3717	<b>0,1639</b>	<b>0,0547</b>

### 11.1.3.3 Studienfeld Wirtschaftswissenschaften - Struktur und Messbereiche

Die Testserie Wirtschaftswissenschaften besteht aus den beiden Tests, die in nachstehender Reihenfolge vorgegeben werden:

„Modellanalyse“, bestehend aus 20 Aufgaben, Bearbeitungszeit 60 Minuten

„Diagramme und Tabellen“, bestehend aus 20 Aufgaben, Bearbeitungszeit 55 Minuten.

Wie in Abbildung 39 veranschaulicht, werden im Test „Modellanalyse“ durch abstrakte Definitionen und Gleichungen Bestimmungsgrößen und Funktionszusammenhänge eines wirtschaftlichen Geschehens dargestellt. Die Aufgaben zielen primär auf die Fähigkeit ab, komplexe, formalisierte Funktionssysteme zu analysieren und die gewonnenen Erkenntnisse auf konkrete Fragestellungen anzuwenden.

D1	$Y_t = C_t + I_t$	$Y_t$ : Sozialprodukt der Volkswirtschaft im Jahr t (in Milliarden €)
D2	$C_t = 100 + 0,6 Y_t$	$C_t$ : Gesamtwirtschaftlicher Konsum im Jahr t (in Milliarden €) - Für das aktuelle Jahr t bezugs der gesamtwirtschaftliche Konsum 700 Milliarden Euro
D3	$I_t = 0,6 (Y_t - Y_{t-1}) + 0,1 K_{t-1}$	$I_t$ : Gesamtwirtschaftliche Investitionen im Jahr t (in Milliarden €)
D4	$K_{t-1} = 1400$	$K_t$ : Kapitalbestand im Jahr t (in Milliarden €). Für das vergangene Jahr t-1, bezugs des Kapitalbestand 1400 Milliarden €
D5	$M_t = 30 + 0,3 Y_t$	$M_t$ : Importe im Jahr t (in Milliarden €)

Angenommen (nur für diese Aufgabe), die gesamtwirtschaftlichen Investitionen im aktuellen Jahr ( $I_{10}$ ) betragen 300 Milliarden €. Wie groß sind dann die Importe im aktuellen Jahr ( $M_{10}$ )?

#### Lösung:

- 1 60 Milliarden €
- 2 270 Milliarden €
- 3 330 Milliarden €
- 4 970 Milliarden €
- 5 Die Frage ist anhand der vorgegebenen Informationen nicht beantwortbar.

Abbildung 39: Beispielaufgabe des Untertests Modellanalyse

Im Test „Diagramme und Tabellen“ sind Graphiken, Schaubilder und tabellarisch dargestellte Informationen aus dem wirtschaftlichen Bereich zu analysieren und zu interpretieren (vgl. Abbildung 40). Es gilt, eine Reihe von Aussagen daraufhin zu überprüfen, ob sie mit den präsentierten Informationen übereinstimmen oder nicht. Dieser Untertest erfordert die Fähigkeit, abstrakte Informationen zu konkretisieren und konkrete Informationen in eine abs-

trakte Darstellung zu übertragen. Ferner müssen wichtige von unwichtigen Informationen unterschieden werden.

Unter dem Bruttoinlandsprodukt (BIP) eines Staates versteht man den Wert aller Güter, die auf dem Gebiet dieses Staates in einem bestimmten Jahr produziert worden sind (einheimische Güter). Beim nominalen BIP werden die Preise des betreffenden Jahres zugrunde gelegt. Beim realen BIP gibt man den Wert der Güter dagegen zu Preisen eines bestimmten Bezugsjahres an. Will man z.B. die Leistungsfähigkeit einer Wirtschaft in zwei verschiedenen Jahren miteinander vergleichen, so bietet es sich an, die BIP beider Jahre auf der Preisbasis desselben Bezugsjahres zu berechnen; der Vergleich wird dann nicht durch die unterschiedlichen Preisniveaus in den beiden Jahren verfälscht. Die folgende Tabelle zeigt die nominalen und die realen BIP von Österreich in den Jahren 1974 bis 1984. Die realen BIP sind alle auf der Preisbasis von 1976 berechnet.

Welche der folgenden Aussagen lässt bzw. lassen sich aus den gegebenen Informationen ableiten?

- I. Der Anstieg des nominalen BIP von 1980 auf 1981 ist ausschließlich durch einen Anstieg der Preise bedingt und nicht auf einem Anstieg der Güterproduktion zurückzuführen.
- II. Von 1976 auf 1978 sind die Preise in Österreich prozentual weniger stark gestiegen als die Güterproduktion.
- III. 1974 waren in Österreich einheimische Güter billiger als 1975.

JAHR	Bruttoinlandsprodukt (+ Milliarden Schilling)	
	Nominal	Real (Preisjahr: 1976)
1974	518	580
1975	656	603
1976	723	725
1977	796	756
1978	842	760
1979	918	766
1980	968	820
1981	936	819
1982	1136	827
1983	1205	845
1984	1380	860

**Lösung:**

- 1. Nur Aussage I lässt sich ableiten.
- 2. Nur Aussagen I und II lassen sich ableiten.
- 3. Nur Aussagen I und III lassen sich ableiten.
- 4. Nur Aussagen II und III lassen sich ableiten.
- 5. Alle drei Aussagen lassen sich ableiten.

Abbildung 40: Beispielaufgabe des Untertests Diagramme und Tabellen

### 11.1.3.4 Testgütemaße Wirtschaftswissenschaften

Die interne Konsistenz beträgt für den gesamten spezifischen Testteil Wirtschaftswissenschaften bezogen auf beide Pilotphasen 0,883. Die internen Konsistenzwerte für die einzelnen Tests variieren je nach Pilotphase und Test zwischen 0,699 bis 0,837.

Tabelle 38:  
Übersicht über die internen Konsistenzkoeffizienten für den Spezifischen Teil der Testatterie Wirtschaftswissenschaften bezogen auf beide Pilotphasen

Reliabilität	Cronbachs Alpha
Modellanalysen	0,837
Diagramme und Tabellen	0,711
<b>Spezifischer Teil Wirtschaftswissenschaften</b>	<b>0,883</b>

### Pilotphase I UniBw HH

Tabelle 39:  
Kennwerte des Tests Modellanalyse bezogen auf die Pilotphase I

„Modellanalyse“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
<b>UniBw HH</b>	0,623	0,417	0,836
<b>N = 65</b>			
<b>BA= 536</b>	0,609	0,326	0,767

Item	Standard- abweichung	Schwierigkeits- index	Trennschärfe
1	0,4369	0,7500	0,3902
2	0,4009	<b>0,8036</b>	0,4820
3	0,4778	0,6607	0,5350
4	0,5045	0,5000	0,2620
5	0,4994	0,4714	0,3805
6	0,4778	0,6607	0,4975
7	0,5042	0,4821	0,5003
8	0,5042	0,4821	0,5092
9	0,4835	0,3571	0,2369
10	0,4009	0,8036	0,2372
11	0,4558	0,7143	0,6073
12	0,4469	0,7321	0,5587
13	0,4713	0,3214	0,0809
14	0,4260	0,7679	0,7249
15	0,3865	<b>0,8214</b>	0,4455
16	0,4713	0,6786	0,3177
17	0,4885	0,6250	0,3285
18	0,4928	0,6071	0,4160
19	0,4140	0,7857	0,5397
20	0,4778	0,3393	0,2820

Tabelle 40:

*Kennwerte des Tests Diagramme und Tabellen bezogen auf die Pilotphase I*

„Diagramme und Tabellen“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw HH N = 65	0,583	0,268	0,699
BA= 536	0,522	0,265	0,694

Item	Standard- abweichung	Schwierigkeits- index	Trennschärfe
1	0,3955	<b>0,8103</b>	<b>0,1790</b>
2	0,4207	0,7759	0,2063
3	0,3072	<b>0,8966</b>	<b>0,0722</b>
4	0,4895	0,6207	0,5208
5	0,4848	0,6379	0,4522
6	0,4895	0,6207	<b>0,1292</b>
7	0,4848	0,6379	0,1892
8	0,4935	0,6034	0,3110
9	0,4935	0,6034	0,5675
10	0,3810	<b>0,8276</b>	0,2926
11	0,4795	0,6552	0,3735
12	0,5041	0,5172	0,3435
13	0,5017	0,5517	0,3648
14	0,5017	0,4483	0,3212
15	0,5041	0,5172	<b>0,0768</b>
16	0,4795	0,3448	<b>-0,2572</b>
17	0,4968	0,4138	0,4509
18	0,4995	0,4310	0,3739
19	0,4995	0,4310	0,1841
20	0,4667	0,3103	0,2122

## Pilotphase II UniBw München

Tabelle 41:  
Kennwerte des Tests Modellanalyse bezogen auf die Pilotphase II

„Modellanalyse“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw M	0,698	0,413	0,833
<b>N = 53</b>			
<b>BA= 536</b>	0,609	0,326	0,767

Item	Standardabweichung	Schwierigkeitsindex	Trennschärfe
1	0,3471	<b>0,8636</b>	0,3632
2	0,3902	<b>0,8182</b>	0,3467
3	0,4080	0,7955	0,4366
4	0,5011	0,4318	0,3403
5	0,4712	0,6818	0,5635
6	0,4615	0,7045	0,5276
7	0,4712	0,6818	0,4984
8	0,4866	0,6364	0,4488
9	0,5058	0,5000	0,3952
10	0,3210	<b>0,8864</b>	0,2645
11	0,4505	0,7273	0,3474
12	0,4239	0,7727	0,5389
13	0,4866	0,3636	0,2127
14	0,4380	0,7500	0,5416
15	0,2908	<b>0,9091</b>	0,5255
16	0,4615	0,7045	0,2453
17	0,4380	0,7500	0,3776
18	0,4795	0,6591	0,4723
19	0,3471	<b>0,8636</b>	0,3965
20	0,5037	0,4545	0,4207

Tabelle 42:  
Kennwerte des Tests Diagramme und Tabellen bezogen auf die Pilotphase II

„Diagramme und Tabellen“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw M	0,614	0,302	0,728
<b>N = 53</b>			
<b>BA= 536</b>	0,522	0,265	0,694

Item	Standardabweichung	Schwierigkeitsindex	Trennschärfe
1	0,3117	<b>0,8936</b>	0,3888
2	0,3799	<b>0,8298</b>	0,2318
3	0,3799	<b>0,8298</b>	0,5693
4	0,4522	0,7234	0,6391
5	0,4790	0,6596	0,2670

6	0,5025	0,5532	0,3355
7	0,4857	0,6383	0,3290
8	0,4790	0,6596	0,2274
9	0,4408	0,7447	0,4288
10	0,3977	<b>0,8085</b>	0,2785
11	0,5025	0,5532	0,1960
12	0,4522	0,7234	<b>0,0951</b>
13	0,4857	0,6383	0,2241
14	0,4961	0,5957	0,3892
15	0,5025	0,5532	0,2335
16	0,4961	0,4043	<b>-0,2154</b>
17	0,5044	0,4681	0,3217
18	0,4961	0,4043	0,5451
19	0,4961	0,4043	0,2908
20	0,3977	<b>0,1915</b>	0,2631

## Pilotphasen I und II

Tabelle 43:

*Kennwerte des Tests Modellanalyse bezogen auf die Pilotphase I und II*

„Modellanalyse“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw N = 118	0,657	0,416	0,837
BA= 536	0,609	0,326	0,767

Item	Standardabweichung	Schwierigkeitsindex	Trennschärfe
1	0,4020	<b>0,8000</b>	0,3977
2	0,3943	<b>0,8100</b>	0,4146
3	0,4513	0,7200	0,5089
4	0,5016	0,4700	0,2743
5	0,4878	0,6200	0,4608
6	0,4688	0,6800	0,5061
7	0,4976	0,5700	0,5172
8	0,5000	0,5500	0,4944
9	0,4960	0,4200	0,3229
10	0,3487	<b>0,8600</b>	<b>0,1980</b>
11	0,4513	0,7200	0,4865
12	0,4352	0,7500	0,5507
13	0,4761	0,3400	<b>0,1399</b>
14	0,4292	0,7600	0,6260
15	0,3487	<b>0,8600</b>	0,4799
16	0,4648	0,6900	0,2928
17	0,4688	0,6800	0,3621
18	0,4852	0,6300	0,4431
19	0,3861	<b>0,8200</b>	0,4838
20	0,4902	0,3900	0,3517

Tabelle 44:

*Kennwerte des Tests Diagramme und Tabellen bezogen auf die Pilotphase I und II*

„Diagramme und Tabellen“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw N = 118	0,597	0,212	0,711
BA= 536	0,522	0,265	0,694

Item	Standardabweichung	Schwierigkeitsindex	Trennschärfe
1	0,3611	<b>0,8476</b>	0,2653
2	0,4019	<b>0,8000</b>	0,2209
3	0,3416	<b>0,8667</b>	0,3013
4	0,4737	0,6667	0,5744
5	0,4800	0,6476	0,3683
6	0,4941	0,5905	0,2139
7	0,4829	0,6381	0,2508
8	0,4855	0,6286	0,2768
9	0,4737	0,6667	0,5099
10	0,3868	<b>0,8190</b>	0,2826
11	0,4902	0,6095	0,2765
12	0,4902	0,6095	0,2417
13	0,4941	0,5905	0,3061
14	0,5022	0,5143	0,3570
15	0,5013	0,5333	0,1490
16	0,4855	0,3714	-0,2323
17	0,4985	0,4381	0,3941
18	0,4958	0,4190	0,4452
19	0,4958	0,4190	0,2282
20	0,4392	0,2571	0,2147

### 11.1.3.5 Studienfeld Ingenieurwissenschaften - Struktur und Messbereiche

Die Testbatterie Ingenieurwissenschaften setzt sich aus zwei Tests zusammen, die in nachstehender Reihenfolge vorgegeben werden:

„Formalisieren technischer Zusammenhänge“, bestehend aus 20 Aufgaben, Bearbeitungszeit: 60 Minuten

„Technisches Verständnis“, bestehend aus 20 Aufgaben, Bearbeitungszeit: 45 Minuten.

Der Test „Formalisieren technischer Zusammenhänge“ prüft die Fähigkeit, verbal oder graphisch vorgegebene technisch-naturwissenschaftliche Sachverhalte in eine formelmäßige Darstellung zu übertragen und die auftretenden Größen zueinander in Beziehung zu setzen (vgl. Abbildung 41). Dies erfordert vor allem die Fähigkeit, verbal vorgegebene Informationen zu formalisieren bzw. zu abstrahieren, sowie das Verstehen von formalisiert dargebotenen Informationen.

Die Rakete „Speed of Light“ startet mit dem Startgewicht  $G_s$ . Nach der Zeit  $T$  hat sich das Gewicht um  $G_T$  verringert, wobei der Gewichtsverlust proportional zur Zeit verlaufen ist.



Welche der folgenden Formeln gibt das Raketengewicht  $G$  nach der Zeit  $t$  wieder, wenn  $0 \leq t \leq T$  gilt (Start:  $t = 0$ )?

**Lösung:**

- 1  $G = G_s - G_T \cdot (T - t)$
- 2  $G = G_s - G_T \cdot t$
- 3  $G = (G_s - G_T) \cdot \frac{t}{T}$
- 4  $G = G_s - G_T \cdot \frac{t}{T}$

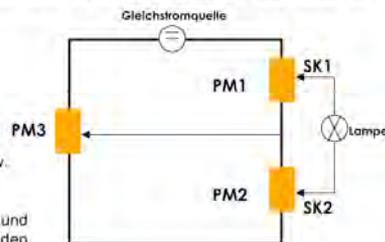
Abbildung 41: Beispielaufgabe des Untertests Formalisieren technischer Zusammenhänge

Im Test „Technisches Verständnis“ sind einfache, unmittelbar praxisbezogene Problemstellungen zu bearbeiten (vgl. *Abbildung 42*). Sie zielt primär auf die Fähigkeit ab, sich einfache technische Funktionsabläufe vorzustellen und elementare technisch-naturwissenschaftliche Ursache-Wirkungs-Zusammenhänge zu analysieren.

Die abgebildete Schaltung enthält drei gleiche Potenziometer (regelbare Widerstände) mit den Widerstandsdrähten PM1, PM2 und PM3 sowie den Schleifkontakten SK1, SK2 und SK3. Die Helligkeit, mit der die Lampe leuchtet, ist abhängig davon, in welchen Positionen sich die Schleifkontakte auf den dazugehörigen Widerstandsdrähten befinden. Sie ist umso geringer, je größer die Gesamtlänge der vom Strom durchflossenen Abschnitte der Widerstandsdrähte ist.

Welche der folgenden Aussagen trifft bzw. treffen in diesem Zusammenhang zu?

- I. Greifen die Schleifkontakte SK1 und SK2 an den betreffenden Widerstandsdrähten in der Mitte an, so leuchtet die Lampe genau so hell, wie wenn sie jeweils ganz unten angreifen.
- II. Die Lampe leuchtet dann am hellsten, wenn alle drei Schleifkontakte an den betreffenden Widerstandsdrähten ganz oben angreifen.



**Lösung:**

- 1 Nur Aussage I trifft zu.
- 2 Nur Aussage II trifft zu.
- 3 Beide Aussagen treffen zu.
- 4 Keine der beiden Aussagen trifft zu.

Abbildung 42: Beispielaufgabe des Untertests Technisches Verständnis

## 11.1.3.6 Testgütemaße Ingenieurwissenschaften

Die interne Konsistenz beträgt für den gesamten Spezifischen Teil Ingenieurwissenschaften bezogen auf beide Pilotphasen 0,870. Die internen Konsistenzwerte für die einzelnen Tests variieren je nach Pilotphase und Test zwischen 0,641 und 0,878.

Tabelle 45:

*Übersicht über die internen Konsistenzkoeffizienten für den Spezifischen Teil der Testbatterie Wirtschaftswissenschaften bezogen auf beide Pilotphasen*

Reliabilität	Cronbachs Alpha
Formalisieren technischer Zusammenhänge	0,828
Technisches Verständnis	0,684
<b>Spezifischer Teil Ingenieurwissenschaften</b>	<b>0,870</b>

**Pilotphase I UniBw Hamburg**

Tabelle 46:

*Kennwerte des Tests Formalisieren technischer Zusammenhänge bezogen auf die Pilotphase I*

„Formalisieren technischer Zusammenhänge“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
<b>UniBw HH</b> <b>N = 50</b>	0,766	0,489	0,878
<b>BA= 509</b>	0,43	0,37	0,81

Item	Standardabweichung	Schwierigkeitsindex	Trennschärfe
1	0,3734	<b>0,8367</b>	0,4852
2	0,3312	<b>0,8776</b>	0,7568
3	0,3058	<b>0,8980</b>	0,3398
4	0,4738	0,6735	0,1625
5	0,3912	<b>0,8163</b>	0,8679
6	0,3912	<b>0,8163</b>	0,6929
7	0,4809	0,6531	0,5741
8	0,3536	0,8571	0,7407
9	0,4216	0,7755	0,5880
10	0,3536	<b>0,8571</b>	0,5316
11	0,3536	<b>0,8571</b>	0,5457
12	0,4345	0,7551	0,4284
13	0,4461	0,7347	0,5059
14	0,4809	0,6531	0,1999
15	0,4564	0,7143	0,2223
16	0,4923	0,6122	0,3252
17	0,3912	0,8163	0,5619
18	0,4345	0,7551	0,4630
19	0,3312	<b>0,8776</b>	0,3951
20	0,5051	0,4898	0,3797

Tabelle 47:

Kennwerte des Tests technisches Verständnis bezogen auf die Pilotphase I

„Technisches Verständnis“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw HH N = 49	0,640	0,319	0,758
BA= 509	0,49	0,24	0,66

Item	Standard- abweichung	Schwierigkeits- index	Trennschärfe
1	0,2040	<b>0,9574</b>	<b>0,1968</b>
2	0,3799	<b>0,8298</b>	<b>0,0287</b>
3	0,2471	<b>0,9362</b>	0,3342
4	0,3977	<b>0,8085</b>	0,2459
5	0,4408	0,7447	0,3640
6	0,1459	0,0213	<b>-0,0723</b>
7	0,4712	0,6809	0,4128
8	0,4790	0,6596	0,1316
9	0,4623	0,7021	0,4869
10	0,4522	0,7234	0,3455
11	0,4280	0,7660	0,4006
12	0,4998	0,5745	<b>0,1052</b>
13	0,4712	0,6809	0,3429
14	0,4522	0,7234	0,5977
15	0,4790	0,6596	0,3712
16	0,5053	0,5106	0,3715
17	0,5025	0,5532	0,5842
18	0,4961	0,5957	0,4619
19	0,4998	0,4255	<b>0,1010</b>
20	0,4408	0,2553	0,4475

### Pilotphase II UniBw München

Tabelle 48:

Kennwerte des Tests Formalisieren technischer Zusammenhänge bezogen auf die Pilotphase II

„Formalisieren technischer Zusammen- hänge“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
UniBw M N = 59	0,756	0,329	0,756
BA= 509	0,43	0,37	0,81

Item	Standard- abweichung	Schwierigkeits- index	Trennschärfe
1	0,3837	<b>0,8246</b>	0,3909
2	0,2577	<b>0,9298</b>	0,4636
3	0,3311	<b>0,8772</b>	0,2584
4	0,4113	0,7895	<b>0,1583</b>
5	0,2854	<b>0,9123</b>	0,3256
6	0,3837	<b>0,8246</b>	0,3607

7	0,4690	0,6842	0,4726
8	0,2577	<b>0,9298</b>	0,4858
9	0,3679	<b>0,8421</b>	0,3376
10	0,3981	<b>0,8070</b>	0,2260
11	0,3311	<b>0,8772</b>	0,3607
12	0,4616	0,7018	0,3156
13	0,4867	0,6316	0,3620
14	0,4867	0,6316	0,2658
15	0,5006	0,4386	0,0635
16	0,5025	0,5439	0,2090
17	0,3504	<b>0,8596</b>	0,5471
18	0,4616	0,7018	0,4571
19	0,2253	<b>0,9474</b>	0,2360
20	0,4867	0,3684	0,2758

Tabelle 49:

*Kennwerte des Tests technisches Verständnis bezogen auf die Pilotphase II*

„Technisches Verständnis“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
<b>UniBw M</b>	0,634	0,227	0,641
<b>N = 59</b>			
<b>BA= 509</b>	0,49	0,24	0,66

Item	Standard- abweichung	Schwierigkeits- index	Trennschärfe
1	0,3072	<b>0,8966</b>	0,3709
2	0,3478	<b>0,8621</b>	<b>-0,0083</b>
3	0,2831	<b>0,9138</b>	0,2172
4	0,3810	<b>0,8276</b>	<b>0,1110</b>
5	0,4509	0,7241	0,3429
6	0,3810	<b>0,8276</b>	0,0809
7	0,4317	0,7586	<b>-0,0027</b>
8	0,4592	0,7069	0,1428
9	0,4968	0,5862	0,2363
10	0,4592	0,7069	0,4364
11	0,4509	0,7241	0,2747
12	0,4895	0,3793	0,1358
13	0,4667	0,6897	0,3118
14	0,4968	0,5862	0,3997
15	0,4995	0,5690	0,2782
16	0,4667	0,3103	<b>0,1288</b>
17	0,4968	0,4138	0,2208
18	0,4895	0,6552	0,3776
19	0,4667	0,3103	0,2436
20	0,4207	0,2241	0,2508

## Pilotphase I und II

Tabelle 50:

*Kennwerte des Tests Formalisieren technischer Zusammenhänge bezogen auf die Pilotphase I und II*

„Formalisieren technischer Zusammenhänge“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
<b>UniBw</b>	0,761	0,424	0,828
<b>N = 115</b>			
<b>BA = 509</b>	0,43	0,37	0,81

Item	Standardabweichung	Schwierigkeitsindex	Trennschärfe
1	0,3773	<b>0,8302</b>	0,4352
2	0,2937	<b>0,9057</b>	0,6314
3	0,3184	<b>0,8868</b>	0,2956
4	0,4430	0,7658	<b>0,1532</b>
5	0,3402	<b>0,8679</b>	0,6386
6	0,3854	<b>0,8208</b>	0,5338
7	0,4725	0,6698	0,5211
8	0,3064	<b>0,8962</b>	0,6350
9	0,3931	<b>0,8113</b>	0,4728
10	0,3773	<b>0,8302</b>	0,3727
11	0,3402	<b>0,8679</b>	0,4584
12	0,4479	0,7264	0,3688
13	0,4690	0,6792	0,4269
14	0,4818	0,6415	0,2276
15	0,4980	0,5660	<b>0,1339</b>
16	0,4966	0,5755	0,2679
17	0,3687	<b>0,8396</b>	0,5496
18	0,4479	0,7264	0,4523
19	0,2801	<b>0,9151</b>	0,5826
20	0,4966	0,4245	0,3285

Tabelle 51:

*Kennwerte des Tests technisches Verständnis bezogen auf die Pilotphase I und II*

„Technisches Verständnis“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
<b>UniBw</b>	0,634	0,261	0,684
<b>N = 115</b>			
<b>BA = 509</b>	0,49	0,24	0,66

Item	Standardabweichung	Schwierigkeitsindex	Trennschärfe
1	0,2666	<b>0,9238</b>	0,2900
2	0,3611	<b>0,8476</b>	0,0093
3	0,2666	<b>0,9238</b>	0,2676
4	0,3868	<b>0,8190</b>	0,1765
5	0,4443	0,7333	0,3518
6	0,5013	0,4667	-0,0909
7	0,4493	0,7238	0,1986

8	0,4665	0,6985	0,1353
9	0,4829	0,6381	0,3481
10	0,4539	0,7143	0,3905
11	0,4392	0,7429	0,3326
12	0,5013	0,4667	0,1164
13	0,4665	0,6857	0,3258
14	0,4800	0,6476	0,4822
15	0,4902	0,6095	0,3201
16	0,4922	0,4000	0,2421
17	0,5018	0,4762	0,3898
18	0,4855	0,6286	0,4155
19	0,4829	0,3619	0,1701
20	0,4280	0,2381	0,3475

### 11.1.3.7 Studienfeld Informatik - Struktur und Messbereiche

Die Testbatterie Informatik setzt sich aus zwei Tests zusammen, die in nachstehender Reihenfolge vorgegeben werden:

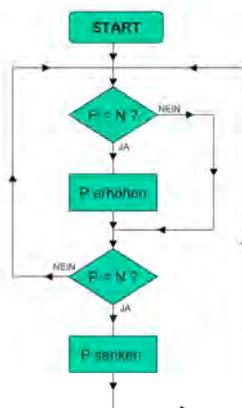
„Flussdiagramme“, bestehend aus 20 Aufgaben, Bearbeitungszeit 60 Minuten

„Formale Systeme“, bestehend aus 20 Aufgaben, Bearbeitungszeit 60 Minuten

Die Messbereiche der einzelnen Aufgabengruppen lassen sich wie folgt umschreiben:

Im Test „Flussdiagramme“ sind verbal beschriebene technische oder wirtschaftliche Prozesse in ein formalisiertes Ablaufschema zu übertragen und vorgegebene Ablaufschemata zu analysieren (vgl. *Abbildung 43*). Sie spricht damit vor allem die Fähigkeit an, gegebene technische oder wirtschaftliche Sachverhalte und Prozesse modellhaft darzustellen, sowie die Fähigkeit, mit formalisierten Systemen umzugehen.

Beim Computerhersteller BIM ist man bestrebt, den Produktionsumfang  $P$  für ein bestimmtes Produkt möglichst genau an die jeweilige Nachfrage  $N$  anzupassen. Das abgebildete Flussdiagramm zeigt die Strategie, mit der man dieses Ziel zu erreichen versucht.



- I. Das Flussdiagramm zeigt eine Strategie, bei der ein zu großer Produktionsumfang unverändert bleibt, ein zu geringer Produktionsumfang dagegen korrekterweise erhöht wird.
- II. Das Flussdiagramm zeigt eine Strategie, bei der es vorkommen kann, dass der Produktionsumfang fälschlicherweise gesenkt wird.

**Lösung:**

- |                          |   |                                      |
|--------------------------|---|--------------------------------------|
| <input type="checkbox"/> | 1 | Nur Aussage I trifft zu.             |
| <input type="checkbox"/> | 2 | Nur Aussage II trifft zu.            |
| <input type="checkbox"/> | 3 | Beide Aussagen treffen zu.           |
| <input type="checkbox"/> | 4 | Keine der beiden Aussagen trifft zu. |

Abbildung 43: Beispielaufgabe des Untertests Flussdiagramme

Wie Abbildung 44 veranschaulicht, befasst sich der Test „Formale Systeme“ mit der Umformung von Zeichenreihen nach bestimmten Regeln. Das Lösen dieser Aufgaben setzt die Fähigkeit voraus, zu erkennen, welche Möglichkeiten der Bildung von Zeichenreihen die Regeln und ihre Verknüpfungen zulassen. Dazu gilt es, neben der folgerichtigen Anwendung abstrakter Regeln die charakteristischen Eigenschaften des jeweiligen Regelsystems zu erfassen (beispielsweise bleiben bestimmte Eigenschaften der Zeichenreihen, z.B. Symmetrien, bei Anwendung gewisser Regeln erhalten). Zum Lösen derartiger Aufgaben ist auch ein hohes Maß an Konzentration erforderlich sowie die Bereitschaft, sich in abstrakt dargebotene Problemstellungen hinein zu denken.

Welche(s) der unter I und II aufgeführten Wörter lässt bzw. lassen sich mit Hilfe der Regeln

S → FSF  
 S → GSG  
 S → HSH

aus S erzeugen?

I.  $F^3S^3F^3$   
 II.  $(GH)^{10}S(GH)^{10}$

**Lösung:**

- 1 Nur I lässt sich erzeugen.  
 2 Nur II lässt sich erzeugen.  
 3 I und II lassen sich erzeugen.  
 4 Weder I noch II lassen sich erzeugen.

Abbildung 44: Beispielaufgabe des Untertests Formale Systeme

### 11.1.3.8 Testgütemaße Informatik

Die interne Konsistenz beträgt für den gesamten Spezifischen Teil Informatik 0,639. Die internen Konsistenzwerte für die einzelnen Tests variieren zwischen 0,131 und 0,622.

Tabelle 52:

Übersicht über die internen Konsistenzkoeffizienten für den Spezifischen Teil der Testbatterie Informatik

Reliabilität	Cronbachs Alpha
Flussdiagramme	0,131
Formale Systeme	0,622
<b>Spezifischer Teil Informatik</b>	<b>0,639</b>

**Pilotphase I UniBw München (keine Testung in Hamburg)**

Tabelle 53:

*Kennwerte des Tests Flussdiagramme*

„Flussdiagramme“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
<b>UniBw M</b>	0,424	0,037	0,131
<b>N = 30</b>			
<b>BA= 499</b>	0,33	0,19	0,59

Item	Standard-abweichung	Schwierigkeits-index	Trennschärfe
1	0,5061	0,4483	0,3622
2	0,4123	0,7931	<b>0,0917</b>
3	0,4837	0,6552	<b>0,1219</b>
4	0,4355	0,2414	<b>-0,0274</b>
5	0,5061	0,5517	<b>0,0386</b>
6	0,5085	0,4828	<b>0,1430</b>
7	0,3509	<b>0,1379</b>	<b>0,1838</b>
8	0,2579	0,0690	<b>-0,1838</b>
9	0,4837	0,6552	<b>-0,1643</b>
10	0,4355	0,7586	<b>0,0842</b>
11	0,5012	0,4138	<b>-0,3477</b>
12	0,4123	0,2069	0,2828
13	0,3099	<b>0,1034</b>	<b>0,1050</b>
14	0,3509	0,8621	<b>-0,0737</b>
15	0,4123	0,2069	<b>0,1450</b>
16	0,5061	0,4483	<b>-0,3334</b>
17	0,4549	0,2759	0,2576
18	0,4938	0,3793	<b>0,1418</b>
19	0,4837	0,6552	<b>0,0840</b>
20	0,3509	<b>0,1379</b>	<b>-0,1618</b>

Tabelle 54:

*Kennwerte des Tests Formale Systeme*

„Formale Systeme“	Mittlere Schwierigkeit	Mittlere Trennschärfe	Cronbachs Alpha
<b>UniBw M</b>	0,595	0,217	0,622
<b>N = 30</b>			
<b>BA= 499</b>	0,47	0,31	0,75

Item	Standard-abweichung	Schwierigkeits-index	Trennschärfe
1	0,3509	<b>0,8621</b>	<b>-0,2069</b>
2	0,4549	0,7241	0,6069
3	0,4837	0,6552	0,1946
4	0,3844	<b>0,8246</b>	0,2121
5	0,4837	0,6552	0,2668
6	0,3844	<b>0,8276</b>	0,3936
7	0,4938	0,6207	<b>0,0901</b>

8	0,4837	0,6552	0,4161
9	0,5061	0,4483	0,4661
10	0,4938	0,6207	<b>0,1128</b>
11	0,4938	0,6207	0,3246
12	0,4837	0,6552	<b>0,1007</b>
13	0,4938	0,6207	<b>0,1819</b>
14	0,5085	0,5172	0,3938
15	0,5085	0,4829	<b>0,1357</b>
16	0,5012	0,5862	0,3607
17	0,5061	0,4483	<b>-0,0177</b>
18	0,5061	0,4483	0,4413
19	0,4837	0,3448	<b>-0,0781</b>
20	0,4549	0,2759	<b>-0,0466</b>

11.1.4 Normen Spezifischer Teil

**Geisteswissenschaften**

Tabelle 55:  
*Normen der Tests Sprachensysteme und Textverständnis I und II*

**Notentabelle für Sprachensysteme HH**

N = 75

Note	7	6	5	4	3	2	1
Anzahl	0	3	5	8	11	14	18
Richtig	bis 2	bis 4	bis 7	bis 10	bis 13	bis 17	bis 20

**Notentabelle für Sprachensysteme M**

N = 47

Note	7	6	5	4	3	2	1
Anzahl	0	3	7	10	12	15	18
Richtige	bis 2	bis 6	bis 9	bis 11	bis 14	bis 17	bis 20

**Notentabelle für Sprachensysteme HH & M**

N = 122

Note	7	6	5	4	3	2	1
Anzahl	0	3	6	9	11	15	18
Richtige	bis 2	bis 5	bis 8	bis 10	bis 14	bis 17	bis 20

**Notentabelle für Sprachensysteme BAA**

N = 295

Note	7	6	5	4	3	2	1
Anzahl	0	5	7	10	13	17	19
Richtige	bis 4	bis 6	bis 9	bis 12	bis 16	bis 18	bis 20

**Notentabelle für Sprachensysteme OPZ**

Note	7	6	5	4	3	2	1
Anzahl	0	3	6	10	12	15	18
Richtige	bis 2	bis 5	bis 9	bis 11	bis 14	bis 17	bis 20

**Notentabelle für Textverständnistest 1 HH**

N = 75

Note	7	6	5	4	3	2	1
Anzahl							
Richtige		0	1	2	3	4	5

**Notentabelle für Textverständnistest 1 M**

N = 47

Note	7	6	5	4	3	2	1
Anzahl							
Richtige		0	1	2	3	4	5

**Notentabelle für Textverständnistest 2 HH**

N = 75

Note	7	6	5	4	3	2	1
Anzahl							
Richtige		0	1	2	3	4	5

**Notentabelle für Textverständnistest 2 M**

N = 47

Note	7	6	5	4	3	2	1
Anzahl							
Richtige		0	1	2	3	4	5

**Notentabelle für Textverständnis 1 HH & M**

N = 122

Note	7	6	5	4	3	2	1
Anzahl							
Richtige		0	1	2	3	4	5

**Notentabelle für Textverständnis 2 HH & M**

N = 122

Note	7	6	5	4	3	2	1
Anzahl							
Richtige		0	1	2	3	4	5

**Notentabelle für Textverständnis 1 + 2 OPZ**

Note	7	6	5	4	3	2	1
Anzahl				3		6	8
Richtige	0	1	2	bis 4	5	bis 7	bis 10

**Wirtschaftswissenschaften**

Tabelle 56:

*Normen der Tests Modellanalyse und Diagramme und Tabellen***Notentabelle für Modellanalyse HH**

N = 65

Note	7	6	5	4	3	2	1
Anzahl		1	5	11	14	17	19
Richtige	0	bis 4	bis 10	bis 13	bis 16	bis 18	bis 20

**Notentabelle für Modellanalyse M**

N = 53

Note	7	6	5	4	3	2	1
Anzahl	0	5	8	13	15		19
Richtige	bis 4	bis 7	bis 12	bis 14	bis 17	18	bis 20

**Notentabelle für Modellanalyse HH & M**

N = 118

Note	7	6	5	4	3	2	1
Anzahl	0	4	7	12	15		19
Richtige	bis 2	bis 6	bis 11	bis 14	bis 17	18	bis 20

**Notentabelle für Modellanalyse BAA**

N = 284

Note	7	6	5	4	3	2	1
Anzahl	0	5	8	11	14	17	19
Richtige	bis 4	bis 7	bis 10	bis 13	bis 16	bis 18	bis 20

**Notentabelle für Modellanalyse OPZ**

Note	7	6	5	4	3	2	1
Anzahl	0	4	7	11	14	17	19
Richtige	bis 3	bis 6	bis 10	bis 13	bis 16	bis 18	bis 20

**Notentabelle für Diagramme & Tabellen HH**

N = 65

Note	7	6	5	4	3	2	1
Anzahl	0	4	7	11	13	16	18
Richtige	bis 3	bis 6	bis 10	bis 12	bis 15	bis 17	Bis 20

**Notentabelle für Diagramme & Tabellen M**

N = 53

Note	7	6	5	4	3	2	1
Anzahl	0	4	9	12	14	16	18
Richtige	bis 3	bis 8	bis 11	bis 13	bis 15	bis 17	bis 20

Notentabelle für Diagramme & Tabellen HH & M  
N = 118

Note	7	6	5	4	3	2	1
Anzahl	0	4	8	11	14	16	18
Richtige	bis 3	bis 7	bis 10	bis 13	bis 15	bis 17	bis 20

Notentabelle für Diagramme & Tabellen BAA  
N = 284

Note	7	6	5	4	3	2	1
Anzahl	0	4	8	10	13	15	18
Richtige	bis 3	bis 7	bis 9	bis 12	bis 14	bis 17	bis 20

Notentabelle für Diagramme & Tabellen OPZ

Note	7	6	5	4	3	2	1
Anzahl	0	4	8	10	13	15	18
Richtige	bis 3	bis 7	bis 9	bis 12	bis 14	bis 17	bis 20

## Ingenieurwissenschaften

Tabelle 57:

*Normen der Tests Formalisieren technischer Zusammenhänge und Technisches Verständnis*

Notentabelle für Formalisieren technischer Zusammenhänge HH  
N = 50

Note	7	6	5	4	3	2	1
Anzahl	0	5	9	16			
Richtige	bis 4	bis 8	bis 15	bis 17	18	19	20

Notentabelle für Formalisieren technischer Zusammenhänge M  
N = 59

Note	7	6	5	4	3	2	1
Anzahl	0	6	12	15	17		
Richtige	bis 5	bis 11	bis 14	bis 16	bis 18	19	20

Notentabelle für Formalisieren technischer Zusammenhänge HH & M  
N = 109

Note	7	6	5	4	3	2	1
Anzahl	0	5	10	15	17		
Richtige	bis 4	bis 9	bis 14	bis 16	bis 18	19	20

Notentabelle für Formalisieren technischer Zusammenhänge BAA  
N = 509

Note	7	6	5	4	3	2	1
Anzahl	0	2	4	6	10	14	17
Richtige	bis 1	bis 3	bis 5	bis 9	bis 13	bis 16	bis 20

Notentabelle für Formalisieren technischer Zusammenhänge **OPZ**

Note	7	6	5	4	3	2	1
Anzahl	0	3	7	12	13	15	18
Richtige	bis 2	bis 6	bis 11	bis 15	bis 16	bis 17	bis 20

Notentabelle für Technisches Verständnis HH  
N = 49

Note	7	6	5	4	3	2	1
Anzahl	0	4	8	12	15	17	19
Richtige	bis 3	bis 7	bis 11	bis 14	bis 16	bis 18	bis 20

Notentabelle für Technisches Verständnis M  
N = 58

Note	7	6	5	4	3	2	1
Anzahl	0		9	11	11	17	19
Richtige	bis 7	8	bis 10	bis 14	bis 14	bis 18	bis 20

Notentabelle für Technisches Verständnis HH & M  
N = 109

Note	7	6	5	4	3	2	1
Anzahl	0	7	9	12	15	17	19
Richtige	bis 6	bis 8	bis 11	bis 14	bis 16	bis 18	bis 20

Notentabelle Technisches Verständnis **BAA**  
N = 509

Note	7	6	5	4	3	2	1
Anzahl	0	4	6	8	11	14	16
Richtige	bis 3	bis 5	bis 7	bis 10	bis 13	bis 15	bis 20

Notentabelle für Technisches Verständnis **OPZ**

Note	7	6	5	4	3	2	1
Anzahl	0	5	8	10	13	16	18
Richtige	bis 4	bis 7	bis 9	bis 12	bis 15	bis 17	bis 20

**Informatik**

Tabelle 58:  
Normen der Tests Flussdiagramme und Formale Systeme

**Keine Testung in Hamburg erfolgt (kein Studiengang Informatik).**

**Notentabelle für Flussdiagramme M**

**N = 30**

Note	7	6	5	4	3	2	1
Anzahl	0			8		11	>
Richtige	bis 5	6	7	bis 9	10	bis 12	

**Notentabelle für Flussdiagramme BAA**

**N = 499**

Note	7	6	5	4	3	2	1
Anzahl	0		3	5	7	10	13
Richtige	bis 1	2	bis 4	bis 6	bis 9	bis 12	bis 20

**Notentabelle für Flussdiagramme OPZ**

Note	7	6	5	4	3	2	1
Anzahl	0		4	6	8	10	12
Richtige	bis 2	3	bis 5	bis 7	bis 9	bis 11	bis 20

**Keine Testung in Hamburg erfolgt (kein Studiengang Informatik).**

**Notentabelle für Formale Systeme M**

**N = 30**

Note	7	6	5	4	3	2	1
Anzahl	0	6	9	11	14	16	19
Richtige	bis 5	bis 8	bis 10	bis 13	bis 15	bis 18	bis 20

Notentabelle für Formale Systeme **BAA**

N = 499

Note	7	6	5	4	3	2	1
Anzahl	0	3	5	8	11	14	17
Richtige	bis 2	bis 4	bis 7	bis 10	bis 13	bis 16	bis 20

Notentabelle für Formale Systeme **OPZ**

Note	7	6	5	4	3	2	1
Anzahl	0	4	6	9	12	14	17
Richtige	bis 3	bis 5	bis 8	bis 11	bis 13	bis 16	bis 20

## 11.1.5 Itemanalyse Allgemeiner Teil II (Persönlichkeitstests)

Der Allgemeine Teil II nennt sich Persönlichkeitsfragebogen Studieneignung. Dieser besteht aus Skalen folgender drei Persönlichkeitsfragebögen, die die in Tabelle 59 aufgeführten Persönlichkeitseigenschaften und Arbeitshaltungen messen.

1. Persönlichkeitsmerkmale Offizierbewerber (PMO)
2. Kurzfassung des Fünf-Faktorenmodells der Persönlichkeit (Big Five)
3. Stressbewältigungsressourcen-Fragebogen

Tabelle 59:

Übersicht aller Skalen des Allgemeinen Teils II (Persönlichkeitsfragebögen)

PMO	Big Five	Stressbewältigungsressourcen
	Neurotizismus	
	Extraversion	
	Offenheit für Erfahrungen	
	Verträglichkeit:	
	Entgegenkommen / Friedfertigkeit	
	Verträglichkeit:	
Teamfähigkeit	Teamfähigkeit	
Gewissenhaftigkeit	Gewissenhaftigkeit	
		Selbstwirksamkeitserwartung
Bewältigungsverhalten		Externale Kontrollüberzeugung
		Internale Kontrollüberzeugung
		Engagement
		Herausforderung
Risikokompetenz		
Handlungsorientierung		
Leistungsmotivation		
Durchsetzungsfähigkeit		
Führungspotential		
Selbstdisziplin		

Alle Items müssen zwecks späterer Praxistauglichkeit auf der Basis der siebenstufigen Bundeswehr-Antwortskala (1 = „ich stimme gar nicht zu“ bis 7 = „ich stimme vollkommen zu“) hinsichtlich Zustimmung bzw. Ablehnung bearbeitet werden. Eine andere Form der Skalierung (z.B. Vierer-Skalierung) ist aufgrund des Anwendungsbezugs nicht umsetzbar.

#### 11.1.5.1 Persönlichkeitsmerkmale Offizierbewerber

Der Fragebogen PMO wurde im Rahmen der Studierendenauswahl auf folgende neun zur Messung der Studieneignung relevante Persönlichkeitsmerkmale reduziert:

1. Teamfähigkeit
2. Gewissenhaftigkeit
3. Bewältigungsverhalten
4. Risikokompetenz
5. Handlungsorientierung
6. Leistungsmotivation
7. Durchsetzungsfähigkeit
8. Führungspotential
9. Selbstdisziplin

Der PMO dient der Absicherung allgemeiner Offiziereignungsbeurteilung in der OPZ. Zu diesem Zweck wurden durch die Universität Dortmund im Auftrag des BMVg Persönlichkeitsmerkmale oder individuelle Einstellungen in Fragen umgesetzt. Die Merkmale entstammen dem Bestand der bisher in der OPZ diagnostizierten Merkmale (z.T. liegen leicht veränderte Begrifflichkeiten vor) sowie aus Anforderungsermittlungen, die besonders neue und zukünftige Tätigkeitsmerkmale von Offizieren berücksichtigen. Im Rahmen der Studierendenauswahl sind nur die aufgeführten Skalen relevant, die im Folgenden im Sinne einer exakten Definition näher erläutert werden.

#### **Definition Teamfähigkeit**

Die Skala Teamfähigkeit entstand aus dem Alltagsverständnis heraus, dass Teamkompetenz als Fähigkeit und Bereitschaft zur Zusammenarbeit und Kooperation in einem Team versteht. Als Facetten sind sowohl grundsätzliche Einstellungen zur Teamarbeit enthalten, als auch Selbsteinschätzungen zum eigenen Arbeitsverhalten.

#### **Definition Bewältigungsverhalten**

Die Dimension des Bewältigungsverhaltens bezieht sich auf das Anforderungs-Ressourcen-Modell von Becker (1998). Es wird postuliert, dass die habituelle Gesundheit und Stressbewältigungsfähigkeit eines Individuums um so besser ist, je mehr es dem Individuum mit Hil-

fe der ihm zur Verfügung stehenden internen und externen Ressourcen gelingt, interne und externe Anforderungen und Belastungen zu bewältigen. Externe Ressourcen sind in der Umwelt liegende Ressourcen, die einem Individuum zur Verfügung stehen und die es nutzen kann, um Anforderungen besser zu bewältigen (z.B. soziale Unterstützung). Interne Ressourcen beschreiben die Kompetenzen und Problemlösefähigkeiten, die sich eine Person im Laufe ihrer Entwicklung und Sozialisation durch Lernen erworben hat (z.B. individuelle und soziale Kompetenzen).

### **Definition Gewissenhaftigkeit**

In Anlehnung an die Skala Gewissenhaftigkeit des NEO-FFI (Borkenau & Ostendorf, 1993) beschreibt diese Dimension nicht die Impulskontrolle, sondern die Art der Selbstkontrolle, die sich auf den aktiven Prozess der Planung, Organisation und Durchführung von Aufgaben bezieht. In der Absicht, Verhaltensweisen über Einstellungen zu erfassen, wurde ein niedriges Abstraktionsniveau gewählt. Die Items beschreiben die Facetten Zuverlässigkeit, Pünktlichkeit, Sorgfalt und Organisation.

### **Definition Risikokompetenz**

Die Skala Risikokompetenz orientiert sich an der Disposition Sensation Seeking (Erlebnissuche), die 1979 von Zuckerman beschrieben wurde. Gemeint ist das individuell unterschiedliche Bedürfnis nach abwechslungsreichen, neuen und komplexen Eindrücken und der zugehörigen Bereitschaft, um solcher Eindrücke willen physische und soziale Risiken in Kauf zu nehmen. Nach Rheinberg (2000) fördert neben Thrill und Adventure Seeking eine stark ausgeprägte Leistungsthematik die Motivation, sich neuen Herausforderungen zu stellen. Die ist die Facette „Risikobereitschaft“. Als bedeutsam wird weiter die adäquate Einschätzung von Risiken und die Art und Weise, Handlungen abzuleiten angesehen.

### **Definition Handlungsorientierung**

Die Skala Handlungsorientierung bezieht sich auf das Konstrukt der Handlungs- und Lageorientierung von Kuhl. Diese Theorie ist den Selbstregulations- und Motivationstheorien zuzuordnen. Nach Schwarzer (2000) befasst sie sich vor allem mit individuellen Unterschieden in dem Unvermögen, selbst gewählte Intentionen in Handeln umzusetzen und diese Handlung gegen widerstrebende Tendenzen aufrechtzuerhalten. Entsprechend Kuhl (Fragebogen HAKEMP) kann das Konstrukt am besten über die drei bipolaren Dimensionen Präokkupation vs. Disengagement, Zögern vs. Initiative und Unbeständigkeit vs. Ausdauer erfasst werden.

**Definition Leistungsmotivation**

Die Dimension Leistungsmotivation stützt sich auf das Selbstbewertungsmodell der Leistungsmotivation von Heckhausen (1989). Nach Rheinberg (2000) wird in diesem Modell das Motiv nicht mehr eigenschaftsähnlich als ein stabiles und in sich einheitliches Personmerkmal aufgefasst, sondern als ein sich selbst stabilisierendes System der Selbstbewertung beschrieben. Es wird über die drei Teilprozesse Zielsetzung/Anspruchsniveau, Ursachenzuschreibung und Selbstbewertung erfasst.

**Definition Durchsetzungsfähigkeit:**

In dieser Skala wird ein alltagspraktisches Begriffsverständnis zu Grunde gelegt: die Fähigkeit, sich Schwierigkeiten und Widerständen zu stellen, verbunden mit einem starken Interesse, andere von eigenen Vorstellungen zu überzeugen.

**Definition Führungspotenzial**

Die Skala zu Führungspotenzial basiert auf einem Führungsverständnis der intentionalen sozialen Einflussnahme und berücksichtigt in der Itemformulierung die häufig geringen expliziten Führungserfahrungen der Bewerbergruppe.

**Definition Selbstdisziplin**

Die Skala Selbstdisziplin erfasst weitere Facetten des Arbeitsverhaltens wie Umgang mit Zeit und Einstellung zu Fehlern.

## 11.1.5.2 Big Five

Im Laufe der letzten Jahrzehnte hat das Fünf-Faktoren-Modell der Persönlichkeit eine breite Akzeptanz erreicht (John & Srivastava, 1999; McCrae & Costa, 1999). Werden wie im vorliegenden Fall freiwillig teilnehmende Stichproben verwendet, stellt der Zeitbedarf der Befragung einen zentralen Faktor für die Teilnahmebereitschaft dar. Im hier befindlichen Untersuchungskontext bietet sich anstelle der zeitaufwendigen Langversion eine evaluierte Kurzversion der Schweizer Armee an. Der Fragebogen Big Five besteht aus 48 Items und erfasst die folgenden näher definierten Skalen mit jeweils 12 Items:

1. Extraversion
2. Verträglichkeit: Entgegenkommen / Friedfertigkeit
3. Verträglichkeit: Teamfähigkeit
4. Gewissenhaftigkeit
5. Neurotizismus
6. Offenheit für Erfahrungen

Der Fragebogen Big Five basiert auf dem Fünf-Faktoren-Modell der Persönlichkeit (FFM, „The Big Five“) (Goldberg, 1981), welches die menschliche Persönlichkeit mit folgenden fünf Merkmalen beschreibt (Norman, 1963; de Raad & Perugini, 2002):

**Definition Extraversion**

(Borkenau & Ostendorf, 1993, S. 5):

„Personen mit hohen Werten in Extraversion sind gesellig, aktiv, gesprächig, personenorientiert, herzlich, optimistisch und heiter. Sie mögen Anregungen und Aufregungen.“

**Definition Verträglichkeit**

(Borkenau & Ostendorf, 1993, S. 5):

„Personen mit hohen Werten in der Skala Verträglichkeit sind altruistisch, mitfühlend, verständnisvoll und wohlwollend. Sie neigen zu zwischenmenschlichem Vertrauen, zur Kooperation, zur Nachgiebigkeit, und sie haben ein starkes Harmoniebedürfnis.“

**Definition Gewissenhaftigkeit**

(Borkenau & Ostendorf, 1993, S. 5):

„Die Skala Gewissenhaftigkeit unterscheidet ordentliche, zuverlässige, hart arbeitende, disziplinierte, pünktliche, penible, ehrgeizige und systematische von nachlässigen und gleichgültigen Personen.“

**Definition Neurotizismus**

Die Skala Neurotizismus umfasst Ängstlichkeit, Reizbarkeit, Depressivität, Befangenheit, Impulsivität und Verletzlichkeit.

**Definition Offenheit für Erfahrungen**

Die Dimension Offenheit für Erfahrungen misst Phantasie, Sinn für Ästhetik, Offenheit für Gefühle, Offenheit für Handlungen, für Ideen und für das Reflektieren über Werte und Normen.

Die fünf Merkmalsbereiche gehen aus dem lexikalischen Ansatz hervor, welcher sich aus der Sedimentationshypothese von Ludwig Klages (1926) und Raymond B. Cattell (1943) herleitet. „Die *Sedimentationshypothese* besagt, dass alle Aspekte individueller Differenzen, welche bedeutsam, interessant oder nützlich sind oder waren, in die Sprache Eingang gefunden haben; je bedeutender eine solche individuelle Differenz, desto größer die Wahrscheinlichkeit, dass sie ein gesondertes Wort hervorbrachte“ (Borkenau & Ostendorf, 1993, S. 5). Der *lexikalische Ansatz* orientiert sich an besagter Hypothese. Es besteht die Annahme, dass in Lexika vermerkte Begriffe, welche individuelle Unterschiede beschreiben, die Gesamtheit aller bedeutenden individuellen Unterschiede abdecken. Je wichtiger ein Persönlichkeits-

merkmal ist, umso größer ist die Wahrscheinlichkeit, ein oder mehrere eigene Wörter dafür in der Sprache zu finden (Hardegger, 2002). Daher können aus der Analyse der Sprache die wichtigsten Persönlichkeitsdimensionen der Gesamtpersönlichkeit gefunden werden (Amelang & Bartussek, 1997). Von zahlreichen Untersuchungen wurden die oben genannten fünf Faktoren als repräsentativ für die Gesamtpersönlichkeit angesehen (Norman, 1963; Guthrie & Bennet, 1971; Amelang & Bartussek, 1997). Ostendorf (1990, S. 197) konnte die Gültigkeit des FFM in umfangreichen Studien speziell für den deutschen Sprachraum bestätigen: „Die Untersuchungen bestätigten die strukturelle Validität des Fünf-Faktoren-Modells im deutschen Sprachraum sowie die Robustheit der Faktoren über verschiedene Variablenstichproben, Ratingverfahren, Beurteilergruppen und verschiedene Methoden der Faktorenanalyse hinweg.“

Zur Generierung der Fragebogenitems wurde im Auftrag der Schweizer Armee durch die ETH Zürich in einem ersten Schritt bestehende Verfahren zusammengetragen, welche die Big Five oder Teile davon erfassen: NEO-FFI, NEO-PI-R deutsch, NEO-PI-R englisch, MMQ, EPI-A, EPI-B, FPI, FPI-R, PRF Form KA, PRF Form KB, MRS 45, BIP, CPI, Persönlichkeitsskalensystem 25, 16 PF und 16 PF-R. In zwei Pretests (N = 110 resp. 150) konnten mittels Items- und Skalenanalyse unbrauchbare Items bestimmt und zum Teil durch neue ersetzt werden. Der dritte Pretest wurde an der regulären Aushebung mit 350 Stellungspflichtigen der Schweizer Armee durchgeführt. Aufgrund der Daten ist es gelungen, vier homogene Skalen mit je zwölf Items zu bilden.

Im Rahmen des Projekts Studieneignung wurde der Fragebogen um die im Rahmen der Rekrutierung der Schweizer Armee nicht verwendeten Skalen Neurotizismus sowie Offenheit erweitert.

### 11.1.5.3 Stressbewältigungsressourcen-Fragebogen

Der Fragebogen zur Erfassung von Stressbewältigungsressourcen erfasst fünf Skalen mit jeweils 8 Items, welche eine erfolgreiche Stressbewältigung prognostizieren sollen.

1. Selbstwirksamkeitserwartung
2. Externale Kontrollüberzeugungen
3. Internale Kontrollüberzeugungen
4. Engagement
5. Herausforderung

Bei der Konstruktion des Fragebogens zur Erfassung von Stressbewältigungsressourcen diente das *transaktionale Stressmodell* von Lazarus et al. (1978, 1984) als Basis. Stress be-

zieht sich dabei auf einen transaktionalen Prozess kognitiver Bewertungen, die eine erlebte Beanspruchung bzw. Überforderung eigener Handlungskompetenzen durch eine situativ vorliegende Problemstruktur zum Inhalt haben. Kognitive Bewertungen werden verstanden als kontinuierliche Überprüfung der Umweltbedingungen darauf hin, ob sie bedeutsam für das Wohlbefinden sind und ob entsprechende Bewältigungsressourcen vorhanden sind. Handlungen und Bewältigungsprozesse dienen der Aufrechterhaltung bzw. der Wiederherstellung des Wohlbefindens. Ob eine Anforderung als stressreich erlebt wird, hängt unter anderem von der Problemstruktur (Umwelt) und der individuellen Belastbarkeit (Person) ab. Wird Stress erlebt, versucht die betroffene Person, diesen zu bewältigen. Diese Stressbewältigung (Coping) ist ein „prozesshaftes Geschehen, das alle sich laufend verändernden verhaltensorientierten und intrapsychischen Anstrengungen umfasst, mit umweltbedingten und internen Anforderungen, die als belastend erlebt werden oder die Ressourcen des Individuums überschreiten, fertig zu werden, das heißt, sie zu meistern, sie zu tolerieren, zu reduzieren oder zu minimieren“ (Lazarus & Folkmann, 1984, S. 141). Dieses Bewältigungsverhalten hat zum Ziel, das bestehende Ungleichgewicht zwischen den Anforderungen der Umwelt und den individuellen Ressourcen aufzuheben. Die Persönlichkeitskonstrukte Selbstwirksamkeitserwartung, Kontrollüberzeugungen und Hardiness stellen kognitive Einstellungs-Konzepte dar, welche den Fokus auf stressrelevante Einschätzungsprozesse lenken (Jerusalem, 1990; Krampen, 1991; Maddi, 1999). Sie wurden bei der Konstruktion des Fragebogens zur Erfassung von Stressbewältigungsressourcen operationalisiert. Dies ermöglicht die Prognostizierung einer erfolgreichen Stressbewältigung von Personen. Empirische Untersuchungen bestätigen deren prognostischen Wert im Stressbewältigungsprozess, deren positive Stress-Leistungsbeziehung und deren Relevanz im militärischen Kontext.

### **Definition Selbstwirksamkeitserwartung**

„Selbstwirksamkeitserwartung ist eine stabile Persönlichkeitsdimension, welche die subjektive Überzeugung zum Ausdruck bringt, aufgrund eigenen Handelns schwierige Anforderungen bewältigen zu können“ (Schwarzer, 1994, S. 105). Die *Selbstwirksamkeitserwartung* stellt ein wichtiger Prädiktor des Stresserlebens dar. Hoch selbstwirksame Personen erweisen sich in einer Studie von Jerusalem (1991) als wenig vulnerabel für belastende Situationen. Ihre Stresseinschätzungen bleiben überwiegend positiv und sie befinden sich trotz der belastenden Situation im motivationalen Zustand der Herausforderung. Personen mit einer niedrigen Selbstwirksamkeit hingegen fühlen sich durch stressinduzierte Anforderungen stark belastet und erbringen deutlich schlechtere Leistungen.

**Definition externale Kontrollüberzeugungen**

„Externale Kontrollüberzeugungen liegen vor, wenn eine Person Verstärkungen und Ereignisse, die eigenen Handlungen folgen, als nicht kontingent zum eigenen Verhalten wahrnimmt, sondern sie als Ergebnis von Glück, Pech, Zufall, Schicksal, als unvorhersehbar wegen der Komplexität der Umwelt oder als von anderen Personen abhängig wahrnimmt und interpretiert“ (Rotter, 1966, 1975; zit. nach Krampen, 1981, S. 5). Das Konstrukt *Kontrollüberzeugungen* spielt im Stressbewältigungsprozess eine tragende Rolle. So erwähnt Kohlmann (1990) beispielsweise, dass der Grad der Kontrollierbarkeit ein zentrales Charakteristikum eines bedrohlichen Ereignisses darstellt. Dabei zeigt die Wahrnehmung von Kontrollierbarkeit im Stress-Leistungszusammenhang eine günstige moderierende Wirkung, indem der schädigende Effekt von Stress auf kognitive, motivationale und affektive Reaktionen minimiert wird (Weaver, Morgan, Adkinson-Holmes & Hall, 1992). Das Konstrukt Kontrollüberzeugungen wird unterteilt in externale und internale Kontrollüberzeugungen.

**Definition internale Kontrollüberzeugungen**

„Internale Kontrollüberzeugungen liegen vor, wenn eine Person Verstärkungen und Ereignisse, die eigenen Handlungen folgen, als kontingent zum eigenen Verhalten oder zu eigenen Persönlichkeitscharakteristika wahrnimmt“ (Rotter, 1966, 1975; zit. nach Krampen, 1981, S. 6).

**Definition Engagement**

„Personen mit einem hohen Engagement sind stark in ihre Aktivitäten und Beziehungen involviert, sodass sie auch in belastenden Situationen nicht leicht aufgeben. Sie wählen aktive Bewältigungsstrategien anstelle von Passivität und Vermeidung“ (Kobasa, Maddi & Kahn, 1982, S. 169).

**Definition Herausforderung**

„Personen mit einer hohen Ausprägung in dem Konstrukt Herausforderung sehen in schwierigen Ereignissen und Veränderungen eine Herausforderung, die ihnen eine Lernchance ermöglicht und die Gelegenheit bietet, sich persönlich weiterzuentwickeln. Sie zeichnen sich aus durch Offenheit, kognitive Flexibilität und Ambiguitätstoleranz“ (Kobasa, Maddi & Kahn, 1982, S. 170).

Items aus sechs bestehenden Fragebogen, welche die Gütekriterien erfüllen und für die Zielgruppe der Stellungspflichtigen der Schweizer Armee angemessen sind, dienen als Referenz für die Konstruktion neuer Items und deren anschließende statistische Überprüfung. Aus den

folgenden Fragebögen wurden Referenzitems gewählt: Fragebogen für Kompetenz und Kontrollüberzeugungen (FKK) (Krampen, 1991), Fragebogen zur Erfassung von allgemeiner Selbstwirksamkeit (Jerusalem & Schwarzer, 1999), Hardi Survey III-R (HS-III-R) (Maddi, 2001), Cognitive Hardiness Scale (Nowack, 1989), Personal Views Survey III-R (PVS-III) (Maddi, 1999), Alienation Test (Maddi, Kobasa & Hoover, 1979). Die Items des Fragebogens zur Erfassung von Stressbewältigungsressourcen beziehen sich alle auf generalisierte stressinduzierende Situationen. Es wurden folgende formalen Charakteristika von generalisierten Situationen berücksichtigt: Neuigkeit, Vorhersehbarkeit, zeitliche Bedingungen und Ambiguität. Um der Gefahr von Verzerrungen des Antwortverhaltens vorzubeugen, sind die Items in positiver und negativer Richtung des zu erfragenden Konstrukts formuliert, zudem werden keine Items zu den gleichen Konstrukten hintereinander erfragt. 92 neu konstruierte Items und 41 Referenzitems wurden darauf in Form eines Pretests einer Stichprobe ( $N = 133$ ) vorgelegt. Bei der Konstruktion der ersten Version des Fragebogens zur Erfassung von Stressbewältigungsressourcen wurden die Referenzitems wieder ausgeschlossen. Mittels einer Item-Analyse wurde danach entschieden, welche Items beibehalten und in die endgültige Form des Fragebogens aufgenommen werden sollen. Dieser Fragebogen wurde dann in der Aushebung über 400 Stellungspflichtigen vorgelegt und nochmals überprüft.

Im Rahmen des Projekts Studieneignung wurde der Fragebogen den Gegebenheiten der neuen Zielgruppe Offizieranwärter sowie dem deutschen Sprachgebrauch angepasst.

#### 11.1.5.4 Testgütemaße Allgemeiner Teil II (Pilotphasen I und II)

##### **PMO**

Die internen Konsistenzen (Cronbach's Alpha;  $N = 319$ ) der acht Skalen reichen über beide Pilotphasen hinweg von 0,580 bis 0,923. Die Gesamtreliabilität des Tests beträgt 0,971 (vgl. Formel Seite 9) und ist somit überdurchschnittlich gut.

Tabelle 60:

*Übersicht über die internen Konsistenzkoeffizienten für den Persönlichkeitstest PMO über beide Pilotphasen hinweg*

<b>Reliabilität</b>	<b>Cronbachs Alpha</b>
Teamfähigkeit	0,923
Gewissenhaftigkeit	0,889
Bewältigungsverhalten	0,580
Risikokompetenz	0,482
Handlungsorientierung	0,812
Leistungsmotivation	0,740
Durchsetzungsfähigkeit	0,771
Führungspotential	0,816
Selbstdisziplin	0,786

**Gesamtreliabilität PMO** **0,971**

### Big Five

Die internen Konsistenzen (Cronbach's Alpha; N = 319) der sechs Skalen variieren über beide Pilotphasen hinweg von 0,637 bis 0,936. Die Gesamtreliabilität des Tests beträgt 0,965 und ist somit hervorragend.

Tabelle 61:

*Übersicht über die internen Konsistenzkoeffizienten für Persönlichkeitstest BIG V über beide Pilotphasen hinweg*

Reliabilität	Cronbachs Alpha
Extraversion	0,865
Verträglichkeit: Entgegenkommen / Friedfertigkeit	0,711
Verträglichkeit: Teamfähigkeit	0,936
Gewissenhaftigkeit	0,824
Neurotizismus	0,823
Offenheit für Erfahrungen	0,637
<b>Gesamtreliabilität BIG V</b>	<b>0,965</b>

### Stressbewältigungsressourcen

Die internen Konsistenzen (Cronbach's Alpha; N = 319) der fünf Skalen reichen über die Pilotphasen hinweg von 0,757 bis 0,852. Die Gesamtreliabilität beträgt 0,962.

Tabelle 62:

*Übersicht über die internen Konsistenzkoeffizienten für Persönlichkeitstest Stressbewältigungsressourcen über beide Pilotphasen hinweg*

Reliabilität	Cronbachs Alpha
Selbstwirksamkeitserwartung	0,845
Externale Kontrollüberzeugungen	0,852
Internale Kontrollüberzeugungen	0,757
Engagement	0,807
Herausforderung	0,805
<b>Gesamtreliabilität BIG V</b>	<b>0,962</b>

Die Gesamtreliabilität des Fragebogens zur Erfassung von Stressbewältigungsressourcen nimmt unter Berücksichtigung aller fünf Skalen einen Wert von  $r = .94$  ein und ist somit ausgezeichnet. Die internen Konsistenzwerte variieren zwischen und 0,757 und 0,852.

#### 11.1.6 Gesamtreliabilität Allgemeiner Teil I

Die interne Konsistenz beträgt für den allgemeinen Teil I 0,946 und für den allgemeinen Teil II 0,988 (vgl. Tabelle 63).

Die interne Konsistenz beträgt zusammenfassend somit für den gesamten allgemeinen Teil (Teil I und II) ohne den Untertest Figurenreihen **0,989**.

Tabelle 63:

*Übersicht über die internen Konsistenzkoeffizienten für den allgemeinen Teil insgesamt über beide Pilotphasen hinweg ohne Subtest Figurenreihen*

<b>Reliabilität</b>	<b>Cronbachs Alpha</b>
Figuren erkennen	0,869
Verbale Analogien adaptive Testform	0,840
Teamfähigkeit	0,923
Gewissenhaftigkeit	0,889
Bewältigungsverhalten	0,580
Risikokompetenz	0,482
Handlungsorientierung	0,812
Leistungsmotivation	0,740
Durchsetzungsfähigkeit	0,771
Führungspotential	0,816
Selbstdisziplin	0,786
Extraversion	0,865
Verträglichkeit: Entgegenkommen/Friedfertigkeit	0,711
Verträglichkeit: Teamfähigkeit	0,936
Gewissenhaftigkeit	0,824
Neurotizismus	0,823
Offenheit für Erfahrungen	0,637
Selbstwirksamkeitserwartung	0,845
Externale Kontrollüberzeugungen	0,852
Internale Kontrollüberzeugungen	0,757
Engagement	0,807
Herausforderung	0,805
<b>Gesamtreliabilität Allgemeiner Teil</b>	<b>0,989</b>

### 11.1.7 Gesamtreliabilität Allgemeiner Teil II

Wie in Tabelle 64 dokumentiert beträgt die Gesamtreliabilität des allgemeinen Teils II 0,988 (vgl. Formel S. 11).

Tabelle 64:

*Übersicht über die internen Konsistenzkoeffizienten für den allgemeinen Teil II über beide Pilotphasen hinweg*

	<b>Cronbachs Alpha</b>
Teamfähigkeit	0,923
Gewissenhaftigkeit	0,889
Bewältigungsverhalten	0,580
Risikokompetenz	0,482
Handlungsorientierung	0,812
Leistungsmotivation	0,740
Durchsetzungsfähigkeit	0,771
Führungspotential	0,816
Selbstdisziplin	0,786
Extraversion	0,865
Verträglichkeit: Entgegenkommen/Friedfertigkeit	0,711
Verträglichkeit: Teamfähigkeit	0,936

Gewissenhaftigkeit	0,824
Neurotizismus	0,823
Offenheit für Erfahrungen	0,637
Selbstwirksamkeitserwartung	0,845
Externale Kontrollüberzeugungen	0,852
Internale Kontrollüberzeugungen	0,757
Engagement	0,807
Herausforderung	0,805
<b>Gesamtreliabilität Allgemeiner Teil II</b>	<b>0,962</b>

### 11.1.8 Gesamtreliabilität Spezifischer Teil

Der spezifische Teil der Testbatterie Studieneignung erlangt einen Gesamtreliabilitätswert in Höhe von 0,950.

Tabelle 65:

*Übersicht über die internen Konsistenzkoeffizienten für den gesamten spezifischen Teil über beide Pilotphasen hinweg ohne Subtests Textverständnis I und II und Flussdiagramme*

<b>Reliabilität</b>	<b>Cronbachs Alpha</b>
Sprachensysteme	0,772
Modellanalysen	0,837
Diagramme und Tabellen	0,711
Formalisieren technischer Zusammenhänge	0,828
Technisches Verständnis	0,684
Formale Systeme	0,622
<b>Gesamtreliabilität Spezifischer Teil</b>	<b>0,950</b>

## 11.2 Konstruktvalidität

Die Untersuchung auf Vorhandensein von Konstruktvalidität soll die Annahme bestätigen, dass die Testbatterie Studieneignung sowohl die interessierenden Konstrukte nachvollziehbar abbildet (konvergenter Validität) als auch diese von anderen Konstrukten abgrenzt (diskriminante Validität). Im Folgenden wird die Zusammenhangsstruktur der Subtests aus dem allgemeinen Teil I und spezifischem Teil anhand einer explorativen Faktorenanalyse geprüft. Um zu eruieren, welche Dimensionen tatsächlich gemessen werden, werden im Anschluss an die Faktorenanalyse die Resultate zum nomologischen Netzwerk (vgl. Kapitel 11.1.8) vorgestellt.

### **Analyse der Konstruktvalidität mittels der Faktorenanalyse**

Es wird untersucht, ob sich die gesuchten Konstrukte mittels Hauptkomponentenanalyse bestätigen lassen. „Mit der Faktorenanalyse können Variablen gemäß ihrer korrelativen Beziehungen in voneinander unabhängige Gruppen klassifiziert werden“ (Bortz, 1993, S. 473).

Es wird sowohl für den Testteil als auch den Persönlichkeitsfragebogenteil jeweils eine Faktorenanalyse mit quadrierten multiplen Korrelationen als Kommunalitätsschätzung und mit anschließender Varimax-Rotation durchgeführt. Als Entscheidungshilfe für die Zahl substantieller Hauptkomponenten wurde das Eigenwertkriterium ( $\lambda > 1$ ) gewählt. Wie in Tabelle 66 abgebildet, weisen vier Faktoren einen Eigenwert größer als eins auf und erklären insgesamt circa 58 Prozent der Varianz der Variablen.

Tabelle 66:

*Allgemeiner Teil I und spezifischer Teil der Testbatterie Studieneignung: Erklärter Prozentanteil der vier Faktoren an der Gesamtvarianz.*

	Faktor / Hauptkomponente			
	1	2	3	4
Eigenwert $\lambda$	2.19	1.37	1.17	1.12
Erklärter Anteil an der Gesamtvarianz	18.03	16.11	12.81	11.48

Es wird deutlich dass die erste Hauptkomponente bei einem Eigenwert von  $\lambda = 2.19$  circa 18 Prozent der Gesamtvarianz aufklärt. Die zweite Hauptkomponente trägt bei einem Eigenwert von  $\lambda = 1.37$  16 Prozent zur Varianzaufklärung bei. Die Faktoren drei und vier erklären bei einem Eigenwert von  $\lambda = 1.17$  und  $\lambda = 1.12$  gemeinsam noch circa ein viertel der Gesamtstreuung. Für die Lösung mit vier Hauptkomponenten sind in der folgenden Tabelle 67 die varimax-rotierten Ladungen der Variablen auf den Hauptkomponenten dargestellt. Um die Betrachtung rein substantieller Faktorladungen zu erleichtern, werden nur solche aber alle größer .30 angeführt und schattiert abgebildet.

Tabelle 67:

*Ergebnisse der Hauptkomponentenanalyse (varimax-rotierte Komponentenmatrix). Die Variablen sind den Testteilen zugeordnet.*

	Faktor / Hauptkomponente			
	1	2	3	4
<b>Allgemeiner Teil I</b>				
Figuren erkennen		.445	.303	
Verbale Analogien	.419		.316	
<b>Spezifischer Teil</b>				
<b>Geisteswissenschaften</b>				
1 Sprachsysteme			.738	
2 Textverständnis 1 & 2			.713	
<b>Wirtschaftswissenschaften</b>				
3 Modellanalyse	.881			
4 Diagramme & Tabellen	.871			
<b>Ingenieurwissenschaften</b>				

5 Formalisieren technischer Zusammenhänge	.783
6 Technisches Verständnis	.824
<b>Informatik</b>	
8 Flussdiagramme	.763
9 Formale Systeme	.737

Die varimax-rotierte Faktorladungsmatrix bildet deutlich die vier Studienfelder ab sowie die Aufteilung der allgemeinen Intelligenztestanteile auf den wirtschaftswissenschaftlichen (verbale Analogien) und ingenieurwissenschaftlichen (Figuren erkennen) Testteil. Beide laden außerdem mit einem Faktor größer als .30 auch auf die geisteswissenschaftlichen Subtests Sprachensysteme und Textverständnis. Da die Faktorladungen der Korrelation zwischen einer Variablen und einem Faktor entsprechen, lässt sich ablesen, dass der größte Teil der Varianz der Einzeldimensionen im allgemeinen Teil I und im spezifischen Teil durch den jeweiligen Studienfeldfaktor erklärt wird.

Zur Ermittlung der Konstruktvalidität der Persönlichkeitsfragebögen geschieht gleichsam mittels Hauptkomponentenanalyse. Wie in Tabelle 68 demonstriert, weisen fünf Faktoren einen Eigenwert größer als eins auf und erklären insgesamt circa 70 Prozent der Varianz der Variablen. Es wird deutlich dass die erste Hauptkomponente bei einem beachtlichen Eigenwert von  $\lambda = 6.7$  circa 18 Prozent der Gesamtvarianz aufklärt. Die zweite Hauptkomponente trägt bei einem Eigenwert von  $\lambda = 2.72$  ungefähr 17 Prozent zur Varianzaufklärung bei. Die Faktoren drei und vier erklären bei einem Eigenwert von  $\lambda = 1.84$  und  $\lambda = 1.55$  gemeinsam noch circa ein viertel der Gesamtstreuung. Der schwächste Faktor mit einem knapp über eins eigenen Wert verfügt über einen erklärten Anteil an der Gesamtvarianz in Höhe von etwa 10 Prozent.

Tabelle 68:

*Allgemeiner Teil II der Testbatterie Studieneignung: Erklärter Prozentanteil der vier Faktoren an der Gesamtvarianz.*

	Faktor / Hauptkomponente				
	1	2	3	4	5
Eigenwert $\lambda$	6.70	2.72	1.84	1.55	1.10
Erklärter Anteil an der Gesamtvarianz	18.23	17.21	12.54	11.03	10.48

Für eine Lösung mit fünf Hauptkomponenten sind in der folgenden Tabelle 69 die Ladungen der Variablen auf den Hauptkomponenten dargestellt. Um die Betrachtung rein substantieller Faktorladungen zu erleichtern, werden nur solche aber alle größer .30 angeführt und farblich schattiert abgebildet.

Die varimax-rotierte Faktorladungsmatrix bildet deutlich die fünf Hauptkomponenten ab. Der erste Faktor setzt sich aus den Persönlichkeitsdimensionen Handlungsorientierung, Leistungsmotivation, Gewissenhaftigkeit (PMO und Big Five), Selbstdisziplin und internale Kontrollüberzeugung. Faktor zwei lädt auf die Eigenschaften Neurotizismus (negative Korrelation), Engagement, Selbstwirksamkeitsüberzeugung, Extraversion, externale Kontrollüberzeugung (negative Korrelation) und Risikokompetenz. Bewältigungsverhalten und Teamfähigkeit (PMO und Big Five) bilden den dritten der fünf Faktoren, Offenheit für Erfahrungen und Herausforderung den vierten. Den fünften Faktor bildet die Persönlichkeitseigenschaft Durchsetzungsfähigkeit, Führungspotential und Entgegenkommen bzw. Friedfertigkeit.

Tabelle 69:  
Ergebnisse der Hauptkomponentenanalyse (varimax-rotierte Komponentenmatrix). Die Variablen sind den Testteilen zugeordnet.

	Faktor / Hauptkomponente				
	1	2	3	4	5
<b>PMO</b>					
Handlungsorientierung	.593	.466		.560	
Leistungsmotivation	.508			.314	.473
Risikokompetenz	-.315	.449			
Bewältigungsverhalten	.331	.381	.472		
Teamfähigkeit			.936		
Durchsetzungsfähigkeit					.759
Führungspotential					.669
Gewissenhaftigkeit	.860				
Selbstdisziplin	.877				
<b>Big Five</b>					
Gewissenhaftigkeit	.868				
Teamfähigkeit			.934		
Neurotizismus		-.841			
Entgegenkommen / Friedfertigkeit		.421			.659
Extraversion		.426	.342		
Offenheit für Erfahrungen				.838	
<b>Stressbewältigungsressourcen</b>					
Externale Kontrollüberzeugung		-.778			
Engagement	.414	.592		.323	
Selbstwirksamkeitsüberzeugung		.636		.439	.327
Internale Kontrollüberzeugung	.408				
Herausforderung		.397		.645	.309

### Analyse der Konstruktvalidität mittels des nomologischen Netzwerkes

Der Test Figuren erkennen misst wie in Kapitel 11.1.1.1 beschrieben die flexible Gestaltwahrnehmung, die Wahrnehmungsgeschwindigkeit sowie Konzentrationsvermögen. Da sich ein hochsignifikanter Zusammenhang zwischen dem Test Figuren erkennen und den Ingeni-

eurstudienerfolg messenden Subtests, die auf dieselben oder ähnliche Intelligenzaspekte abzielen, ermitteln lässt, wird die Annahme vorhandener Konstruktvalidität für diese Dimension gestützt.

Die Bewertungen der die geistes- und wirtschaftswissenschaftlichen Fähigkeiten messenden Eigenschaftsdimensionen korrelieren darüber alle hoch signifikant mit dem Untertest Verbal Analogien, der das sprachabhängige logische Denkvermögen und die Fähigkeit zum Erkennen des sprachlogischen Zusammenhangs zwischen Begriffen misst. Mehr als in den Studienfeldern Ingenieurwesen oder Informatik ist die Notwendigkeit sprachabhängiger Intelligenz in Fächern wie Pädagogik oder BWL anzunehmen. Somit kann für das Konstrukt fachspezifischen Studienerfolg bzw. alle zugehörigen Dimensionen mindestens befriedigende bis sehr gute Konstruktvalidität angenommen werden.

Die Annahme, bzw. dass gleiche oder ähnliche Persönlichkeitsdimensionen wie z.B. Führungspotential und Durchsetzungsfähigkeit testübergreifend auf demselben Faktor hoch laden, erweist sich als richtig. Ebenso gehen Bewertungen der Dimension Teamfähigkeit im Test PMO einher mit nicht signifikant unterschiedlichen Teamfähigkeitswerten im Test Big Five. Gleiches gilt für das Persönlichkeitsmerkmal Gewissenhaftigkeit im PMO und Big Five sowie Selbstdisziplin.

### 11.3 Prognostische Validität

In den voran stehenden Kapiteln wurde die Verteilung der Prädiktoren, Kriterien und moderierenden Einflüsse von Studienerfolg beschrieben. Darüber hinaus wurde die Testbatterie Studieneignung hinsichtlich ihrer Reliabilität und Konstruktvalidität umfassend untersucht. Aber erst die Zusammenführung der Erkenntnisse aus der alten und neuen Studieneignungsempfehlung mit den im Studium erbrachten Leistungen bildet den Kern von Erfolgs- und Bewährungskontrollen. In den nächsten Unterkapiteln soll daher aufgezeigt werden, in welchem Zusammenhang und Ausmaß sich die Vorhersagen des neuen im Vergleich zum alten Verfahren bestätigen lassen.

Zur Prüfung der Hypothese 1 wird die Vorhersagekraft des OPZ-Studieneignungsempfehlungsgrads auf den Studienerfolg ermittelt. Gemessen wird ein mit  $r = .21$  als niedrig zu bewertender Spearman-Korrelationskoeffizient zwischen Prädiktor und dem Kriterium Vordiplomnote. Im Vergleich dazu sind bei Vorhersage der Gesamtestbatterie Studieneignung auf den Studienerfolg sehr signifikante Regressionskoeffizienten in Höhe von  $r = .34$  (Vordiplom) und  $r = .42$  zu verzeichnen. Die Höhe des Zusammenhangs zwischen OPZ-

Studieneignungsbeurteilung und Hauptdiplom ist mit  $r = .20$  als ebenfalls gering valide zu beurteilen (vgl. Tabelle 70). Insgesamt betrachtet ist die konkurrente Validität der damaligen Studieneignungsempfehlung als schwach zu bewerten, zumal in das Urteil des Studienberaters bereits weitere Prädiktoren wie Fachnoten, Ergebnis des mathematischen Kenntnistests, Ergebnis des Beratungsgesprächs und Abiturnote eingeflossen waren. Alternativhypothese 1 kann demnach bestätigt werden.

Tabelle 70:

*Ergebnisse der Korrelationsanalyse zur Überprüfung der prognostischen Validität und Gegenüberstellung früherer OPZ-Studieneignungsbeurteilung versus neuer Gesamttestbatterie Studieneignung (studienfachübergreifend).*

Prädiktor	Kriterium	r/R	p
OPZ-Studieneignungsbeurteilung	Vordiplom	<b>.212**</b>	.000
OPZ-Studieneignungsbeurteilung	Hauptdiplom	<b>.204*</b>	.016
Testbatterie Studieneignung (Gesamttestscore)	Vordiplom	<b>.415**</b>	.000
Testbatterie Studieneignung (Gesamttestscore)	Hauptdiplom	<b>.343**</b>	.019

*Anmerkungen:* r: Korrelationskoeffizient, p: ermitteltes Signifikanzniveau zweiseitig, \*\* - Korrelation ist auf dem Niveau von  $p = .01$  signifikant, \* - Korrelation ist auf dem Niveau von  $p = .05$  signifikant. N = 317.

Zur Untersuchung der Hypothese 2 werden einleitend und zu Vergleichszwecken Ergebnisse aus der Entwicklungszeit der fachspezifischen Testverfahren der BA aufgenommen.

### Geisteswissenschaften

Der im Auftrag der BA weiterentwickelte Test für Philologien lässt sich mit dem im Rahmen der vorliegenden Arbeit angewendeten studienfeldspezifischen Test für Geisteswissenschaften an den UniBw nicht vergleichen. Der Ursprungsversion wurde lediglich der Untertest Sprachensysteme entnommen. Zusätzlich wurden die beiden Textverständnisaufgaben I und II entwickelt und bilden einen Teil der Testbatterie. Berechnungen zur prädiktiven Validität ergeben, dass weder in Bezug auf das Vor- noch auf das Hauptdiplom signifikante Validitäten erzielt werden können. Der spezifische Teil Geisteswissenschaften vermag es demnach nicht, zwischen späteren guten und schlechten Absolventen zu unterscheiden bzw. liefert keine Vorhersagegüte, die eine spätere Praxisanwendung der Testbatterie rechtfertigen würde. Hypothese 2 findet in Bezug auf das Studienfeld Geisteswissenschaften keine Bestätigung.

### **Wirtschaftswissenschaften**

Vorläufige Ergebnisse der BA zur prognostischen Validität der Testbatterie Wirtschaftswissenschaften zeigten, dass bei Studenten eines wirtschaftswissenschaftlichen Faches einer regulären Hochschule ( $N = 69$ ) eine Korrelation von  $r = .37$  ( $p < .01$ , einseitiger Test) zwischen dem SFBT-Gesamtwert und der Vorexamensnote bestand. Allerdings beruhte dieses Ergebnis auf einem SFBT-Gesamtwert einer Vorläuferversion der SFBT Wirtschaftswissenschaften. Diese Vorläuferversion bestand aus neun Tests, von denen lediglich zwei in modifizierter Form die aktuelle Testversion bilden. In weiteren Untersuchungen ergaben sich für die aktuelle Testversion folgende Ergebnisse (Didi, 2004, pers. Mitteilung): An einer privaten Hochschule für Wirtschaftswissenschaften wurde ein Aufnahmetest durchgeführt, der u.a. die Studienberatungs-Tests Modellanalyse und Diagramme und Tabellen enthielt. Von 294 Personen, die an diesem Aufnahmetest teilnahmen, wurden ebenfalls die Vordiplomnoten erhoben. Das Gesamtergebnis des Aufnahmetests hatte eine gute Prognosekraft bezüglich der Vordiplomnote ( $r = .27$ ,  $p < .01$ , einseitiger Test). Bei einer Untersuchung an einer Bundeswehrhochschule wurden von 69 Studierenden aus dem Bereich Wirtschaftswissenschaften der SFBT-Gesamtwert sowie die Vordiplomnote erhoben. Es ergab sich hier ein Validitätskoeffizient von  $r = .37$  ( $p < .01$ , einseitiger Test). Der Test Diagramme und Tabellen ging darüber hinaus auch in einen Auswahltest für wirtschaftswissenschaftliche Studiengänge einer österreichischen Fachhochschule ein. Seit 2000 wurde der Auswahltest mit ca. 500 Personen durchgeführt. Für den Test Diagramme und Tabellen offenbarte sich eine überlegene Prognosekraft für die durchschnittliche Fachhochschulnote nach dem 4. Semester ( $r = .42$ ,  $p < .01$ , einseitiger Test).

Die beiden die neue Testbatterie Studieneignung testenden Pilotphasen an den UniBw können trotz ähnlicher Testbedingungen – z.B. Testung bereits studierender Offizieranwärter – auch im Falle dieses studienfeldspezifischen Teils keine signifikanten prognostischen Validitäten hervorbringen. Hypothese 2 findet in Bezug auf das Studienfeld Wirtschaftswissenschaften keine Bestätigung.

### **Ingenieurwissenschaften**

Im Rahmen einer früheren Untersuchung bei der Bundeswehr ergab sich die Möglichkeit einer Nacherhebung bei ehemaligen Teilnehmern an einer groß angelegten Studie mit ca. 4500 Schülern in Niedersachsen (Blum, Hensgen & Trost, 1985). Bei dieser Nacherhebung konnten von 512 Studierenden des Fachbereichs Ingenieurwissenschaften Studienleistungen

als Kriterienvariablen aufgenommen werden (Trost & Freitag, 1990). Der SFBT-Gesamtwert Ingenieurwissenschaften korrelierte signifikant mit der Vordiplomnote ( $r = .34$ ;  $p < .01$ ). An der Universität Linz sind Untersuchungen mit Studienfeldbezogenen Beratungstestserien der ITB Consulting durchgeführt worden (Brandstätter, 2001). Im Fachbereich Ingenieurwissenschaften wurden 163 Personen (vor Studienbeginn) mit der SFBT Ingenieurwissenschaften untersucht. Der SFBT-Gesamtwert Ingenieurwissenschaften korrelierte ebenfalls signifikant mit der Vordiplomnote ( $r = .31$ ;  $p < .01$ ).

Alleinig der spezifische Teil der Testbatterie Studieneignung für Ingenieurwissenschaften verfügt im Vergleich zu den übrigen studienfeldspezifischen Tests über eine auf dem 5%-Niveau signifikante und mehr ausreichende als befriedigende Validität von  $r = .20$  (Kriterium Vordiplom). Hypothese 2 kann in Bezug auf das Studienfeld Ingenieurwissenschaften angenommen werden.

### **Informatik**

Im Rahmen einer früheren Untersuchung bei der Bundeswehr ergab sich die Möglichkeit einer Nacherhebung bei ehemaligen Teilnehmern an einer groß angelegten Studie mit circa 4500 Schülerinnen und Schülern in Niedersachsen (Blum, Hensgen & Trost, 1985). Bei dieser Nacherhebung konnten von 51 Studierenden des Fachbereichs Informatik Studienleistungen als Kriterienvariablen aufgenommen werden (Trost & Freitag, 1990). Der SFBT-Gesamtwert korrelierte signifikant mit der Vordiplomnote ( $r = .46$ ;  $p < .01$ ).

An der Universität Linz wurden Untersuchungen mit den studienfeldbezogenen Beratungstestserien der ITB Consulting GmbH durchgeführt (Brandstätter, 2001). In den Fachbereichen Informatik und Mathematik konnten 120 Personen (vor Studienbeginn) mit der Testbatterie untersucht werden. Der Gesamtwert korrelierte ebenfalls signifikant mit der Vordiplomnote ( $r = .41$ ;  $p < .01$ ).

Auch im Falle des Fachbereichs Informatik korreliert der studienfeldspezifische Gesamtwert mit der Studienleistung nicht signifikant. Hypothese 2 findet in Bezug auf das Studienfeld Informatik keine Bestätigung.

### **Allgemeiner Teil I und II**

Entgegen aller Annahmen ergeben sich auch für den die akademische Intelligenz messenden Allgemeinen Teil I mit den Untertests Figuren erkennen und verbale Analogien keine nennenswerten Gültigkeitswerte.

Völlig anders aber erwartungskonform verhalten sich im Vergleich dazu die Persönlichkeits-tests aus dem Allgemeinen Teil II. Insbesondere der Untertest PMO und Big Five Kurversion liefern in entscheidenden Dimensionen höchst aussagekräftige Koeffizienten und bestätigen das Vorhandensein prognostischer Validität. Nach Bortz (1993) müssen die an einer Regressionsrechnung beteiligten Variablen mindestens Intervallskalenniveau besitzen. Als ordinalskalierte Variable kann die von den Testpersonen beurteilte Skala von „ich stimme gar nicht zu“ (1) bis „ich stimme vollkommen zu“ (7) diese Voraussetzung nicht erfüllen, so dass keine regressionsanalytische Vorhersage des Studienerfolgs auf Basis der Skalenwerte ermittelt werden kann. Somit wird die Stärke des Zusammenhangs zwischen den Variablen durch den Spearman-Rangkorrelationskoeffizienten gemessen. Die Korrelation für die einzelnen Persönlichkeitsskalen sind in Tabelle 71 aufgeführt. Besonders aussagekräftig für den Studienerfolg in Form der Hauptdiplomnote erweisen sich die Persönlichkeitsmerkmale Selbstdisziplin, Handlungsorientierung, Führungspotential, Gewissenhaftigkeit, Teamfähigkeit und Leistungsmotivation. Zur Vorhersage der Vordiplomnote kristallisieren sich die drei Prädiktordimensionen Selbstdisziplin, Gewissenhaftigkeit und Risikokompetenz als besonders valide heraus.

Untersucht man die Vorhersagekraft gemessener Persönlichkeitsfaktoren auf den Studienerfolg in Abhängigkeit des Studienfelds, sind folgende Besonderheiten zu nennen (Studienfachbereich Informatik wird dabei mangels ausreichender Kriteriumszahl nicht einbezogen): Für die Geisteswissenschaften finden sich als hoch signifikante Prädiktoren des Hauptdiploms neben Selbstdisziplin und Gewissenhaftigkeit eine beträchtliche internale Kontrollüberzeugung ( $r = -.31$ ) sowie Führungspotential ( $r = -.35$ ). Dagegen spielen Gewissenhaftigkeit für Ingenieurstudenten eine geringere Rolle als Leistungsmotivation ( $r = -.40$ ) und Selbstdisziplin ( $r = -.30$ ).

Darüber hinaus lässt sich für besonders „schnelle Studenten“ – operationalisiert durch die besonders kurze Zeit von drei Trimestern bis zum Vordiplom – als alleinig hoch signifikanter Prädiktor das Merkmal Gewissenhaftigkeit mit  $r = -.93$  mittels Big Five und  $r = -.83$  mittels PMO ermitteln.

Tabelle 71:

Vergleich der Rangkorrelationskoeffizienten der Prädiktoren aus dem Allgemeinen Teil II nach Spearman mit Vor- und Hauptdiplomnote als Kriterien für Studienerfolg.

Spearman-Rho		Vordiplomnote	Hauptdiplomnote
PMO	<b>Selbstdisziplin</b>	<b>Korrelationskoeffizient</b> <b>-.301**</b>	<b>-.396**</b>
		Signifikanz (2-seitig) .000	.000
PMO	<b>Handlungsorientierung</b>	<b>Korrelationskoeffizient</b> <b>-.093</b>	<b>-.288**</b>
		Signifikanz (2-seitig) .108	.000
PMO	<b>Führungspotential</b>	<b>Korrelationskoeffizient</b> <b>-.029</b>	<b>-.250**</b>
		Signifikanz (2-seitig) .620	.003
PMO	<b>Gewissenhaftigkeit</b>	<b>Korrelationskoeffizient</b> <b>-.261**</b>	<b>-.246**</b>
		Signifikanz (2-seitig) .000	.003
PMO	<b>Teamfähigkeit</b>	<b>Korrelationskoeffizient</b> <b>.109</b>	<b>.179*</b>
		Signifikanz (2-seitig) .059	.032
PMO	<b>Leistungsmotivation</b>	<b>Korrelationskoeffizient</b> <b>-.060</b>	<b>-.169*</b>
		Signifikanz (2-seitig) .299	.043
PMO	<b>Risikokompetenz</b>	<b>Korrelationskoeffizient</b> <b>.137*</b>	<b>.155</b>
		Signifikanz (2-seitig) .017	.169
PMO	Durchsetzungsfähigkeit	Korrelationskoeffizient <b>-.029</b>	<b>-.053</b>
		Signifikanz (2-seitig) .620	.531
PMO	Bewältigungsverhalten	Korrelationskoeffizient <b>-.003</b>	<b>-.098</b>
		Signifikanz (2-seitig) .959	.244
Big Five	<b>Gewissenhaftigkeit</b>	<b>Korrelationskoeffizient</b> <b>-.195**</b>	<b>-.287**</b>
		Signifikanz (2-seitig) .001	.000
Big Five	Teamfähigkeit	Korrelationskoeffizient <b>.086</b>	<b>.154</b>
		Signifikanz (2-seitig) .134	.065
Big Five	Entgegenkommen/Friedfertigkeit	Korrelationskoeffizient <b>-.001</b>	<b>-.029</b>
		Signifikanz (2-seitig) .993	.729
Big Five	Offenheit für Erfahrungen	Korrelationskoeffizient <b>-.079</b>	<b>-.085</b>
		Signifikanz (2-seitig) .172	.313
Big Five	Extraversion	Korrelationskoeffizient <b>.049</b>	<b>-.103</b>
		Signifikanz (2-seitig) .398	.218
Big Five	Neurotizismus	Korrelationskoeffizient <b>-.082</b>	<b>.092</b>
		Signifikanz (2-seitig) .155	.272
Stressbewältigungsressourcen	Selbstwirksamkeitsüberzeugung	Korrelationskoeffizient <b>.005</b>	<b>-.162</b>
		Signifikanz (2-seitig) .924	.052
Stressbewältigungsressourcen	Externale Kontrollüberzeugung	Korrelationskoeffizient <b>.076</b>	<b>.135</b>
		Signifikanz (2-seitig) .186	.107
Stressbewältigungsressourcen	Internale Kontrollüberzeugung	Korrelationskoeffizient <b>.052</b>	<b>-.045</b>
		Signifikanz (2-seitig) .371	.590
Stressbewältigungsressourcen	Engagement	Korrelationskoeffizient <b>-.064</b>	<b>-.128</b>
		Signifikanz (2-seitig) .265	.126
Stressbewältigungsressourcen	Herausforderung	Korrelationskoeffizient <b>.004</b>	<b>-.093</b>
		Signifikanz (2-seitig) .945	.268

Anmerkung:  $N_{\text{Vordiplom}}=301$ ;  $N_{\text{Hauptdiplom}}=144$ ; \*\* - Korrelation ist auf dem Niveau von  $p = .01$  signifikant (2-seitig); \* - Korrelation ist auf dem Niveau von  $p = .05$  signifikant; PMO: Fragebogen Persönlichkeitsmerkmale Offizierbewerber; Stressbewältigungsressourcen: Stressbewältigungsressourcenfragebogen.

Zwecks Ermittlung inkrementeller Validitäten wird in einem multiplen Modell mittels sequentieller multipler linearer Regression überprüft, welche der Prädiktoren in welcher Kombination einen signifikanten Einfluss auf den Studienerfolg haben (vgl. Tabelle 72 und Tabelle 73). Als Persönlichkeitsfaktoren werden die in Tabelle 71 aufgeführten relevanten Eigenschaften Selbstdisziplin, Handlungsorientierung, Führungspotential, Gewissenhaftigkeit, Teamfähigkeit, Leistungsmotivation und Risikokompetenz aus dem PMO und Big Five Kurzversion einbezogen. Die Ergebnisse der multiplen Korrelations- und Regressionsanalyse zeigen bei Anwendung relevanter Prädiktorkombinationen und insbesondere unter Hinzunahme relevanter Persönlichkeitsfaktoren und Abiturnote eine gute Prognosekraft hinsichtlich des Studienerfolgs. Dabei wächst die aufgeklärte Varianz prozentual besonders bei Hinzunahme „weicher Faktoren“ zur Vorhersage der Hauptdiplomnote und weniger bei zusätzlicher Berücksichtigung des Schulnotenmittelwerts. Letzterer erweist sich einflussreicher zur Vorhersage des Kriteriums Vordiplomnote. Darüber hinaus veranschaulichen die Grafiken der Abbildung 45 und Abbildung 46, dass gerade unter Hinzunahme der relevanten Persönlichkeitsmerkmale die Korrelation zum Kriterium Studienerfolg verbessert wird. Am meisten und hoch signifikant profitiert das Studienfeld Wirtschaftswissenschaften von einer zusätzlichen Varianzaufklärung zur Vorhersage der Hauptdiplomnote.

Tabelle 72:

*Inkrementelle Validitäten - Ergebnisse der Regressionsanalyse zur Überprüfung des Einflusses verschiedener Prädiktorkombinationen auf den Studienerfolg in Form der Vordiplomnote*

Studienfeld	n	Allgemeiner Teil I	Allgemeiner Teil I + spezifischer Teil	Allgemeiner Teil + Spezifischer Teil + Persönlichkeitsfaktoren	Allgemeiner Teil + Spezifischer Teil + relevante Persönlichkeitsfaktoren + Abiturnote
GeiWi	116	.038	.039	<b>.476**</b>	<b>.516**</b>
WiWi	109	.112	.118	<b>.398*</b>	<b>.429*</b>
IngWi	94	.122	.197	<b>.478**</b>	<b>.529**</b>
Informatik	25	.057	.101	.482	.616

*Anmerkungen.* R: Multipler Regressionskoeffizient nach sukzessivem Einbezug der Prädiktorvariablen allgemeiner Teil I, spezifischer Teil, PMO-Persönlichkeitsfaktoren, Abiturnote; \*\* - Korrelation ist auf dem Niveau von  $p = .01$  signifikant (2-seitig); \* - Korrelation ist auf dem Niveau von  $p = .05$  signifikant; GeiWi: Geisteswissenschaften; WiWi: Wirtschaftswissenschaften; IngWi: Ingenieurwissenschaften.

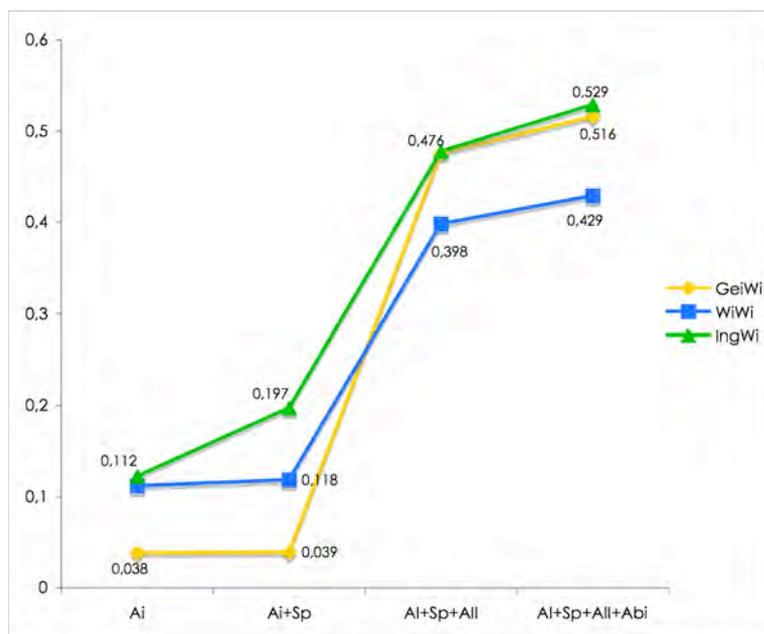


Abbildung 45: Grafische Gegenüberstellung des inkrementellen Validitätswachstums der Regressionsanalyse zur Überprüfung verschiedener Prädiktorkombinationen auf den Studienerfolg in Form der Vordiplomnote. GeiWi: Geisteswissenschaften; WiWi: Wirtschaftswissenschaften; IngWi: Ingenieurwissenschaften; AI: Allgemeiner Teil I; Sp: Spezifischer Teil; All: relevante Persönlichkeitsmerkmale aus Allgemeinem Teil II; Abi: Abiturnote

Wie in Kapitel 2.1.1 eruiert, weisen unterschiedlichste Untersuchungen der Abiturnote ein erstaunliches Differenzierungsvermögen zwischen  $r = .30$  und  $r = .50$  zu. Erwartungsgemäß repliziert die Prädiktorkraft der Abiturnote mit  $r = .32$  bezogen auf die Vordiplomnote und mit  $r = .24$  (signifikant auf dem .01-Niveau zweiseitig) bezogen auf die Hauptdiplomnote frühere Forschungsergebnisse.

Die Kombination mit der Abiturnote mit allgemeinen und spezifischen Studieneignungstests soll die inkrementelle Validität bedeutsam erhöhen können. Im vorliegenden Falle der Testbatterie Studieneignung und unter Hinzunahme von Persönlichkeitsfaktoren mittels PMO und Big Five Kurzversion erhöht sich – wie in Tabelle 73 dokumentiert – die inkrementelle Validität auf  $R = .68$  (Kriterium Hauptdiplomnote). Die Alternativhypothese 3 wird demnach bestätigt. Die Hypothese, nach der eine Erhöhung der Validität durch Hinzunahme von Persönlichkeitstests aus dem allgemeinen Teil II zu erwarten ist, findet ihre Bestätigung für die Studienfelder Geistes- und Wirtschaftswissenschaften. Die Ergebnisse für den Ingenieurwissenschaftstest sind nicht signifikant, die Statistik für den Informatiktestteil kann wegen fehlender Korrelationen nicht berechnet werden. Bei Hinzunahme der Abiturnote steigt die Validität entgegen der Hypothese vier nicht signifikant verschieden vom oben zitierten Ergebnis an.

Tabelle 73:

*Inkrementelle Validitäten - Ergebnisse der Regressionsanalyse zur Überprüfung des Einflusses verschiedener Prädiktorkombinationen auf den Studienerfolg in Form der Hauptdiplomnote.*

Studienfeld	n	Allgemeiner Teil I	Allgemeiner Teil I + spezifischer Teil	Allgemeiner Teil + Spezifischer Teil + Persönlichkeitsfaktoren	Allgemeiner Teil + Spezifischer Teil + relevante Persönlichkeitsfaktoren + Abiturnote
<b>R</b>					
GeiWi	68	.049	.089	<b>.598**</b>	<b>.610**</b>
WiWi	57	.031	.087	<b>.680**</b>	<b>.682**</b>
IngWi	39	.007	.131	.617	.639

Informatik 25 Wegen fehlender Korrelationen kann Statistik nicht berechnet werden.

Anmerkungen. R: Multipler Regressionskoeffizient nach sukzessivem Einbezug der Prädiktorvariablen allgemeiner Teil I, spezifischer Teil, PMO-Persönlichkeitsfaktoren, Abiturnote; \*\* - Korrelation ist auf dem Niveau von  $p = .01$  signifikant (2-seitig); \* - Korrelation ist auf dem Niveau von  $p = .05$  signifikant; GeiWi: Geisteswissenschaften; WiWi: Wirtschaftswissenschaften; IngWi: Ingenieurwissenschaften.

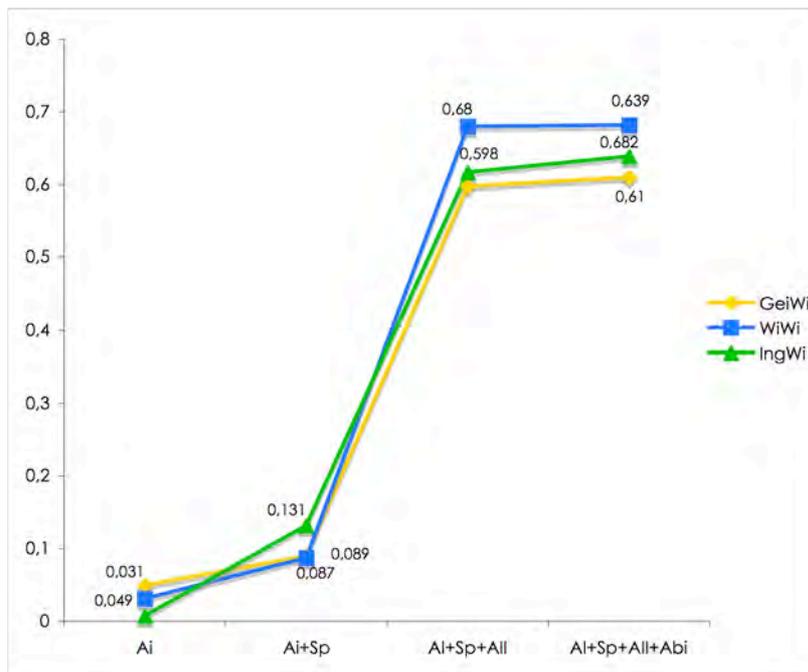


Abbildung 46: Grafische Gegenüberstellung des inkrementellen Validitätswachstums der Regressionsanalyse zur Überprüfung verschiedener Prädiktorkombinationen auf den Studienerfolg in Form der Hauptdiplomnote. GeiWi: Geisteswissenschaften; WiWi: Wirtschaftswissenschaften; IngWi: Ingenieurwissenschaften; AI: Allgemeiner Teil I; Sp: Spezifischer Teil; All: relevante Persönlichkeitsmerkmale aus Allgemeinem Teil II; Abi: Abiturnote

Im Rahmen einer parallel laufenden und für die Diskussion der Ergebnisse der vorliegenden Arbeit durchaus wertvollen Untersuchung von Storm & Krex (2005) werden Antworten auf folgende Fragen gegeben.

1. Wie hat sich die Studienerfolgsquote im Ablauf von zehn Jahren im Vergleich zur schulischen Leistungsfähigkeit entwickelt?
2. Wie hat sich die Studienerfolgsquote im Ablauf von zehn Jahren im Vergleich zur schulischen Leistungsfähigkeit differenziert nach Studiengängen entwickelt?
3. Wie eng ist der Zusammenhang zwischen durchschnittlicher Abiturnote und Studien-erfolg innerhalb der einzelnen Studiengänge?

Bei dieser Untersuchung wird zum einen die durchschnittliche Examensnote herangezogen, zum anderen - zur Ermittlung der „Studienerfolgsquote“ - der prozentuale Anteil bestandener Hauptdiplomnoten. Die Datengrundlage liefert BISON. Unter der Prämisse, methodisch absolut exakte Messungen hervorzubringen, wurden lediglich die Studentengänge 1989 bis 1998 berücksichtigt. Bei vorhergehenden und späteren erweist sich zum Zeitpunkt der Untersuchung die Datenpflegesituation und bei den nachfolgenden die Studienabschluss-situation als uneindeutig.

„Die genaueste und damit methodisch beste Möglichkeit zur Berechnung von Studienerfolgs- bzw. Studienabbruchquoten bietet eine Studienverlaufsstatistik, in der das Studienverhalten jedes einzelnen Studierenden – von der Aufnahme seines Studiums bis zu seinem Ausscheiden – statistisch erfasst wird“ (HIS, 2006, S. 4). Diese aufwendige Berechnungsmöglichkeit scheidet nicht nur für das öffentliche Hochschulsystem in Deutschland sondern aus in den letzten Jahren verschärften Datenschutzgründen auch für die Bundeswehr bzw. vielmehr für den Psychologischen Dienst der Bundeswehr aus.

Als Studienabbrecher werden in Anlehnung an das HIS (2006) ehemalige Studierende charakterisiert, die zwar durch Immatrikulation ein Studium aufgenommen haben, dann aber ohne Abschlussexamen die Universität sowie – im Falle der vorliegenden Stichprobe – kurze Zeit später die gesamte Organisation Bundeswehr verlassen. Die Studienabbruchquote wird definiert als Anteil der Studienanfänger eines Studentengangs an den UniBw, die das Studium beenden, ohne es mit einem Examen abzuschließen. Es soll an dieser Stelle unterstrichen werden, dass sich in den Studienerfolgskriterien keine Noten von Studienabbrechern mehr finden und diese somit das Ergebnis nicht konfundieren. Studienabbrecher wurden im Vorhinein sorgfältig gefiltert und in dem Zusammenhang Unterschiede zwischen Studienwechslern vs. -abbrechern, Differenzen verschiedener Abbrucharten und Abbruchgründe wie Tod oder persönliche Härtefälle versus mangelnde akademische Leistung konsequent geprüft. Aufgrund der besseren Vergleichbarkeit werden ausschließlich Hochschulabsolventen der UniBw Hamburg und München einbezogen. Gemessen an der Laufbahn handelt es sich

auch bei dieser untersuchten Population demnach um Offizieranwärter des Truppendienstes, nicht jedoch um Soldaten des Sanitäts-, militärischen Musik-, Geoinformations- oder militärischen Fachdienstes.

In den nachstehenden Abschnitten werden die Ergebnisse der Gegenüberstellung von Studienerfolg und schulischer Leistungsfähigkeit pro Studentenjahrgang über einen Zeitraum von zehn Jahren beschrieben. In Abbildung 47 sind die über alle Studienfachrichtungen gemittelten Studienerfolgsquoten und Abiturdurchschnittsnote der Studentenjahrgänge 1989 bis 1998 aufgeführt. Zur verständlicheren graphischen Darstellung des Verlaufs sind Letzgenannte hier und in folgenden Abbildungen mit 10 multipliziert. Demnach liegen z.B. für den Studienjahrgang 1989 der gemittelte Abiturdurchschnitt bei 2,6 und die Erfolgsquote bei 67 Prozent. Die niedrigsten Erfolgsquoten finden sich für Studienbeginner der Jahre 1992 und 1995 mit jeweils 61 Prozent. Signifikant höhere Quoten erzielen die Studentenjahrgänge 1989 und 1990 mit 67 und 68 Prozent, was einer Abbruchquote von 33 bzw. 32 Prozent entspricht. Auffällig ist der statistisch bedeutsame Unterschied von Vor- und Hauptdiplomnote. Erstere weicht mit einem Mittel von 3,0 deutlich von der Hauptdiplomnote mit durchschnittlich 2,4 ab.

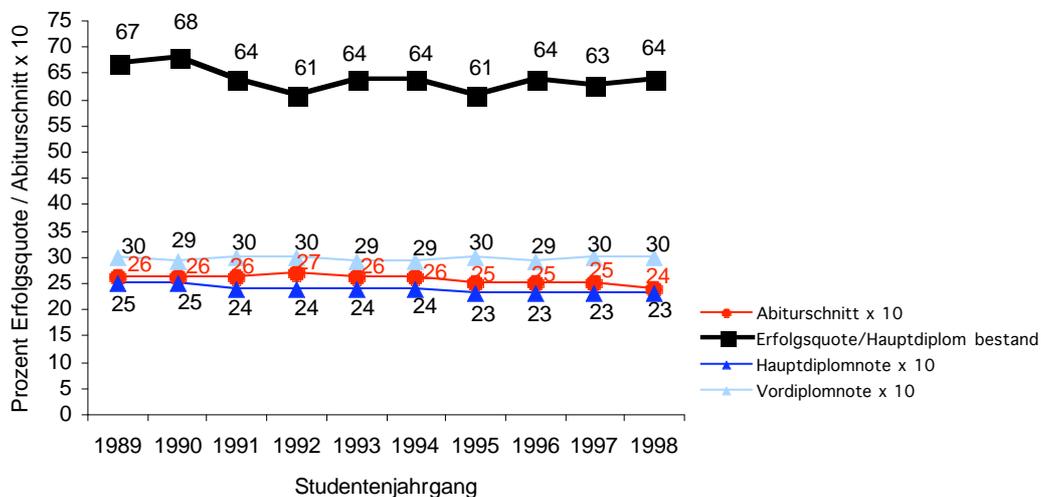


Abbildung 47: Gegenüberstellung der jährlichen Erfolgsquote (Hauptdiplom bestanden) und mittleren Abiturnote der Studentenjahrgänge 1989-98 an den Universitäten der Bundeswehr

### Gegenüberstellung von Studienerfolg und schulischer Leistungsfähigkeit pro Studienfach

Nachstehend werden die Ergebnisse der Gegenüberstellung von Studienerfolg und schulischer Leistungsfähigkeit differenziert nach den verschiedenen Studiengängen beschrieben.

Der Übersicht halber sind in Tabelle 74 noch einmal die einzelnen und für die untersuchten Jahrgänge 1989 bis 1998 relevanten Studiengänge, nach drei Feldern untergliedert, dargestellt.

Tabelle 74:

*Übersicht der Studiengänge an den Universitäten der Bundeswehr in Hamburg und München, untergliedert nach Studienfeldern (Bezugszeitraum: 1989-1998).*

<b>Universität der Bundeswehr Hamburg</b>	<b>Universität der Bundeswehr München</b>
<b>Geisteswissenschaften</b>	
Pädagogik	Pädagogik
Geschichte (ab 1990)	Staats- und Sozialwissenschaften
Politikwissenschaften (ab 1990)	Sportwissenschaften (ab 1991)
<b>Wirtschaftswissenschaften</b>	
Betriebswirtschaftslehre	Wirtschafts- & Organisationswissenschaften
Volkswirtschaftslehre	
<b>Ingenieurwissenschaften</b>	
Elektrotechnik	Elektrotechnik
Maschinenbau	Bauingenieurwesen
Wirtschaftsingenieurwesen	Informatik
	Vermessungswesen
	Luft- und Raumfahrttechnik

Wie in Abbildung 48 ersichtlich, finden sich keine Besonderheiten bezüglich bestimmter Studienfelder. Anders sieht es für die einzelnen Studienfächer aus. Die Erfolgsquoten sowohl der Sport- als auch der Wirtschaftsingenieurwesen-Studenten ragen mit 80 und 85 Prozent deutlich hervor, gefolgt von Fächern wie Geodäsie, Politik und Pädagogik mit 73 bzw. 72 Prozent. Die höchsten Abbruchquoten verzeichnen die Fächer Maschinenbauwesen und Betriebswirtschaftslehre. Es lässt sich aufgrund der Ergebnisse auf einen hohen Einfluss weniger der Abiturnote als vielmehr der Wahl des Studienfaches schließen. Vergleicht man die Variation der Schulabschlussnoten pro Studienfach, lässt sich kein oder nur ein geringer Einfluss auf den Studienerfolg vermuten. Im nächsten Abschnitt werden diese Annahmen näher untersucht.

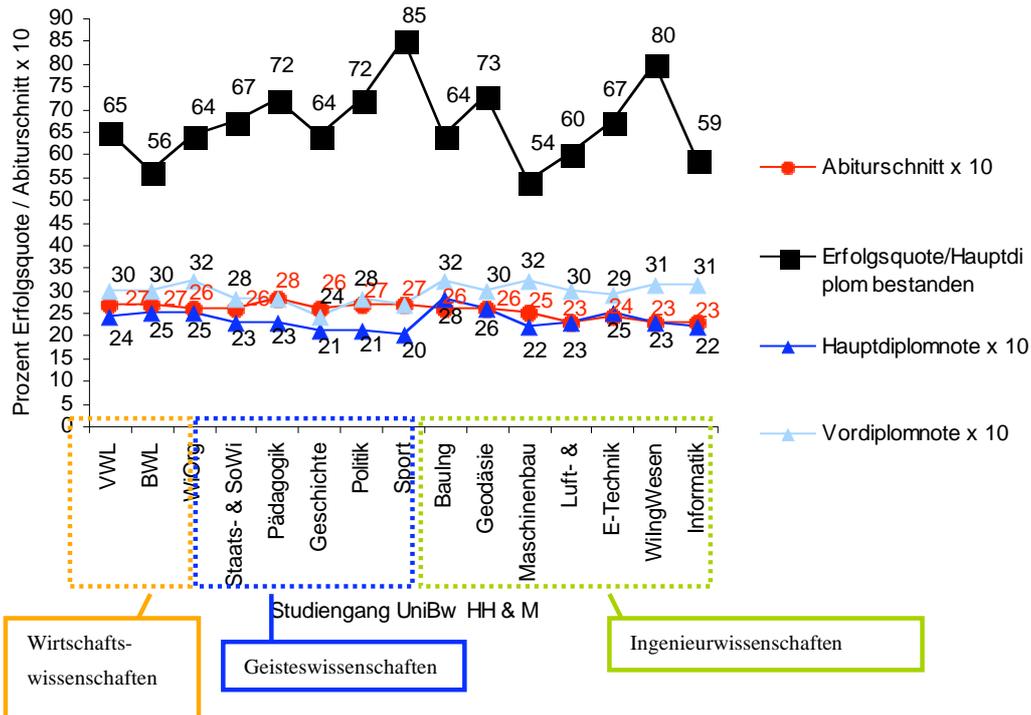


Abbildung 48: Gegenüberstellung von Studienerfolg (Hauptdiplom bestanden) und mittlerer Abiturnote der einzelnen Studiengänge an den Universitäten der Bundeswehr in Hamburg und München

### Zusammenhang zwischen schulischer Leistungsfähigkeit und Studienerfolg pro Studienfach

Der Zusammenhang von durchschnittlicher Abiturnote mit der Hauptdiplomnote fällt mit  $r = .27$  trotz geringer Koeffizienten hochsignifikant aus, was aufgrund des hohen  $N$  von 6990 zu erwarten ist. Dagegen wirkt sich die Tatsache einer in der OPZ vorselektierten Stichprobe wiederum vermindernd auf die Höhe des Zusammenhangs zwischen Abiturnote und Studienerfolg aus. Es ist von einer so genannten „Range Restriction“ auszugehen, welche sich zusätzlich mindernd auf diesen Zusammenhang auswirkt.

Wie im theoretischen Teil der vorliegenden Arbeit bereits ausführlich beschrieben, kann die Abiturnote dessen ungeachtet in Verbindung mit anderen Prädiktoren einen wertvollen inkrementellen Vorhersagebeitrag des Studienerfolges liefern. Vergleicht man die Korrelationskoeffizienten mit denen in der Forschung gefundenen, liegen sie jedoch noch deutlich unter dem Mittel von  $r = .39$ . Betrachtet man den Einfluss separiert nach Studienfach, werden deutlich höhere Koeffizienten erreicht und bestätigen somit die Vermutungen aus dem vorherigen Abschnitt. Abbildung 49 illustriert, dass je nach Studienfach die Höhe des Zusammenhangs zwischen Abitur und Hauptdiplomnote stark variiert. Bis auf die Korrelation

zwischen Sport und Studienerfolg erweisen sich die Zusammenhänge der Studiengänge mit dem Studienerfolg als statistisch sehr signifikant.

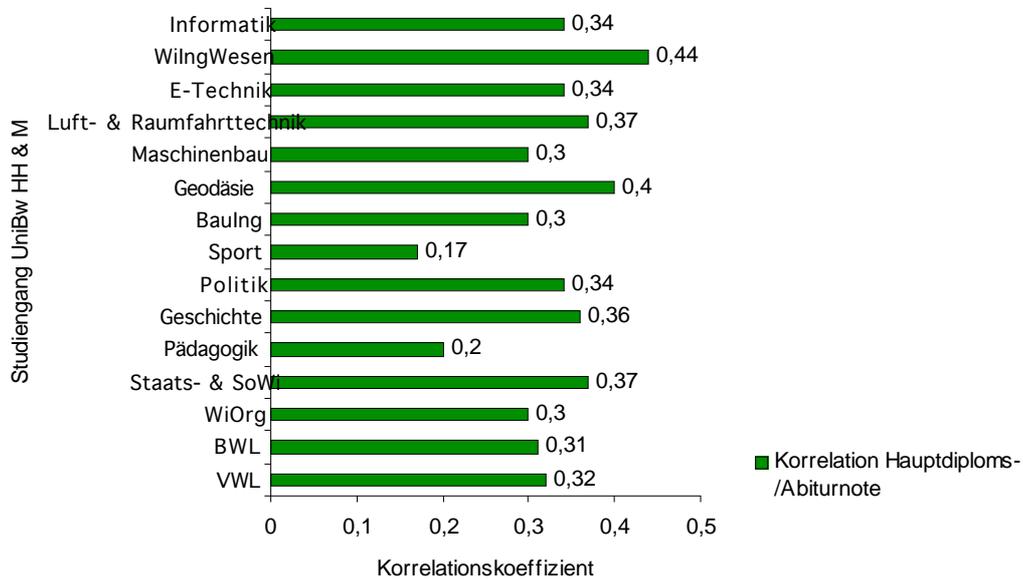


Abbildung 49: Zusammenhang zwischen schulischer Leistungsfähigkeit (Abiturnote) und Studienerfolg (Hauptdiplomsnote) innerhalb der einzelnen Studiengänge an den Universitäten der Bundeswehr in Hamburg und München. (WiOrg = Wirtschafts- und Organisationswissenschaften, WiIngWesen = Wirtschaftsingenieurwesen, E-Technik = Elektrotechnik.)

### Zusätzliche Berechnungen

In der Abbildung 50 sind zusätzlich die nach den drei TSK Heer, Luftwaffe und Marine differenzierten Abbrecherquoten der Studentenjahrgänge 1989 bis 1998 aufgeführt. Der sehr signifikante Ausreißer der TSK Marine im Jahre 1992 lässt sich aus den Daten heraus nicht erklären. Darüber hinausgehende Erklärungen wären an dieser Stelle reine Spekulation und sollen dem an der UniBw angesiedelten ZHQ überlassen werden.

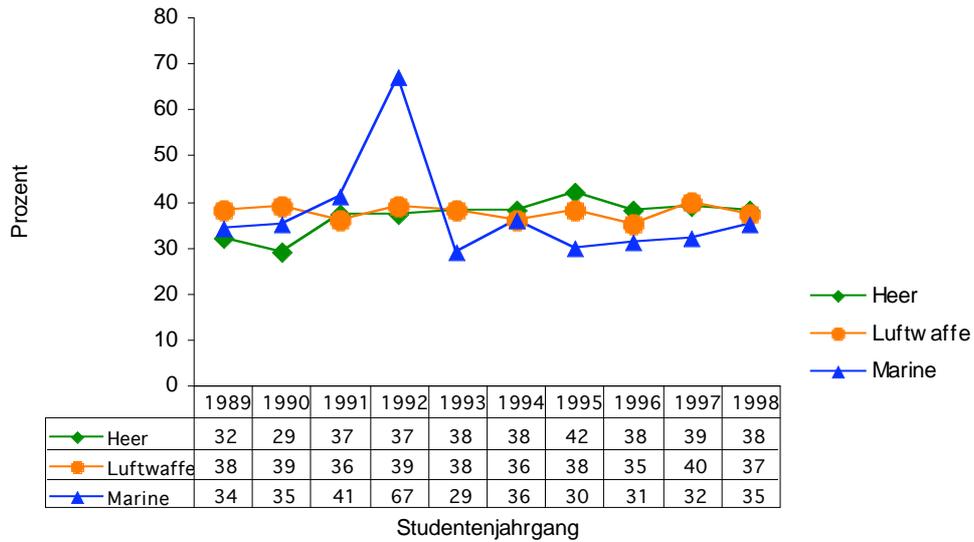


Abbildung 50: Gegenüberstellung der Studienabbrecherquoten der Studentenjahrgänge 1989-98 pro Teilstreitkraft

Wie in Abbildung 51 ersichtlich, weichen die Abbruchquoten der beiden Universitäten der Bundeswehr besonders bedeutsam in den Jahren 1995 bis 1996 zugunsten der UniBw München voneinander ab. Betrachtet man die Universitäten über die zehn Studentenjahrgänge hinweg, unterscheiden sich diese jedoch nicht hinsichtlich der Abbruchquote (Mittelwert jeweils 37,4 Prozent).

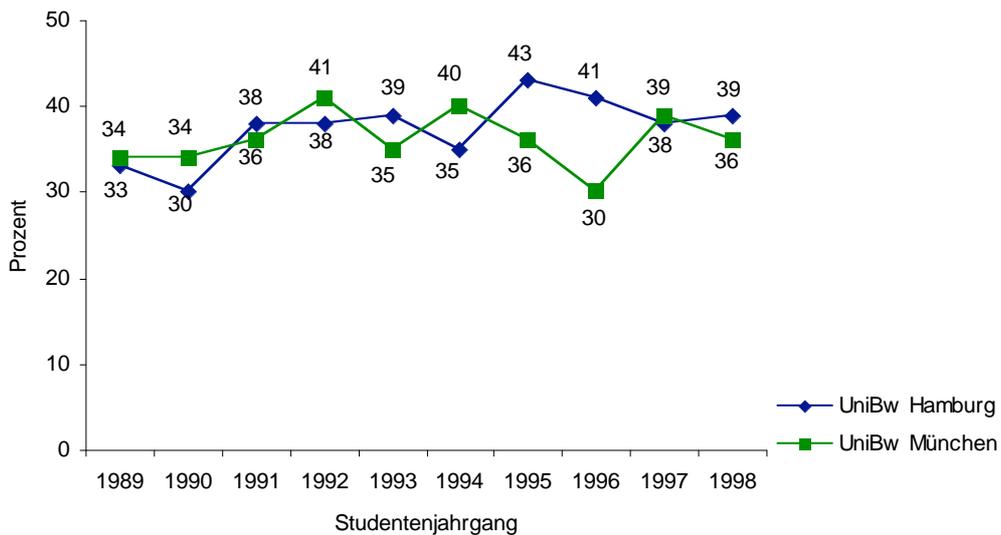


Abbildung 51: Gegenüberstellung der Studienabbrecherquoten der Studentenjahrgänge 1989-98 pro UniBw Hamburg und München

Zusammenfassend bestätigt die Studie den sowohl in der Fachliteratur beschriebenen als auch in dieser Arbeit bestätigten Einfluss der Abiturnote auf den Studienerfolg. In Bezug auf die UniBw fällt dieser unter Berücksichtigung der vorselektierten Stichprobe jedoch etwas geringfügiger aus. Untermauert wird ferner das erwartete Ergebnis einer studienfachabhängigen Vorhersagekraft.

#### 11.4 Einfluss moderierender Hintergrundvariablen

Im Folgenden werden die Hypothesen 6 bis 10 getestet, d.h. mögliche Einflüsse der Moderatorvariablen Geschlecht, TSK, Studienort, Trimesterzahl bis zum Vordiplom und soziokultureller Hintergrund auf das normalverteilte Kriterium Hauptdiplomnote.

Entsprechend der oben ermittelten theoretischen Verteilungsform wird der Einfluss der multi-kategorialen Moderatorvariablen TSK, Trimesterzahl und Bildungshintergrund über die einfaktorielle Varianzanalyse, also den Vergleich von mehr als zwei unabhängigen Stichproben geprüft.

Zur Untersuchung der Hypothese 7 liefert die Varianzanalyse ein leicht signifikantes Ergebnis ( $p < .05$ ). Das bedeutet, dass sich die TSK hinsichtlich ihres Studienerfolgs lediglich marginal unterscheiden. Der Duncan-Test liefert auf dem Niveau von  $p = .05$  nur eine homogene Untergruppe und verzeichnet infolgedessen keinen signifikanten Unterschied zwischen den Gruppen. Bezieht man statt des Hauptdiploms das Kriterium Vordiplomnote ein, findet sich ebenso wenig ein signifikanter Unterschied der beiden unabhängigen Stichproben.

Anders als in Bezug auf den Studienerfolg verhält sich der Einfluss der TSK auf den OPZ-Gesamteignungsgrad. Über 60 Prozent der TSK Heer angehörigen Testpersonen sind demnach den als geeignet bewerteten Offizierbewerbern zuzuordnen, 20 Prozent der Luftwaffe und 32 Prozent der TSK Marine. Dagegen wurde der größte Marinebewerberanteil als gut geeignet (60 Prozent), gefolgt von Luftwaffe (57 Prozent) und schließlich dem Heer mit nur 33 Prozent Anteil geeigneter Anwärter. Circa acht Prozent der Marinebewerber wurden als besonders geeignete Offizieranwärter eingestellt, nur drei Prozent der Luftwaffenbewerber sowie sieben von 175 Heeresangehörigen (entspricht vier Prozent innerhalb der TSK). Bei den besten Offizieranwärtern handelt es sich also meistens um zukünftige Marineoffiziere, bei den als gut geeignet bewerteten handelt es sich meist um Marine- oder Luftwaffenbewerber.

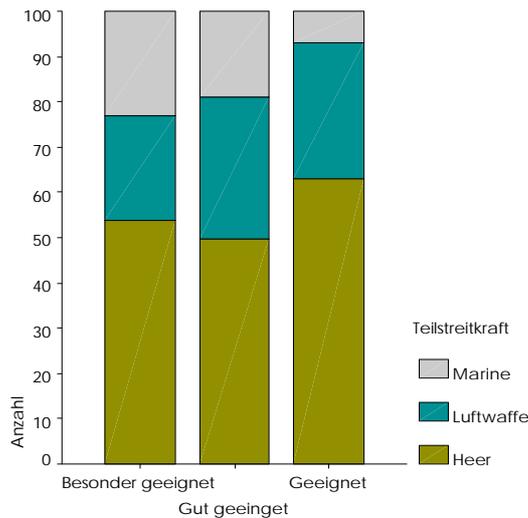


Abbildung 52: Häufigkeitsverteilung des OPZ-Gesamt-eignungsgrads in Abhängigkeit der Teilstreitkraft in der Gesamtstichprobe

Zur Überprüfung der Hypothese 9 wird der Einfluss der Trimesteranzahl bis zum Vordiplom auf den Studienerfolg untersucht. Bezogen auf das Kriterium Hauptdiplomnote ergibt sich ein signifikantes Ergebnis mit  $p = .04$ , für das Kriterium Vordiplomnote sogar ein höchst signifikantes Ergebnis mit  $p < .01$ . Im letzteren Fall liefert der post hoc angewendete Duncan-Test drei homogene Untergruppen, von denen die eine aus drei bis vier studierten Trimestern bis zum Vordiplom besteht, die zweite aus fünf und die dritte aus sechs mit kontinuierlich schlechter werdendem Vordiplomnotenschnitt. Die Studienzeit bis zum Vordiplom wirkt sich also im zunehmenden Verlauf negativ auf das Vordiplomergebnis aus.

Hypothese 10 testet den Einfluss der Moderatorvariablen soziokultureller Bildungshintergrund auf den Studienerfolg. Ausgedrückt in Form des Vaterberufs (einzig im Personalsystem vorhandene Kennzahl) sollen die Hauptdiplommittelwerte auf signifikante Unterschiede getestet werden. Auch wenn die Varianzanalyse mit  $p = .11$  kein signifikantes Ergebnis zwischen den Gruppen liefert, sind Tendenzen entsprechend der Tabelle 75 gemäß der Nullhypothese abzuleiten. Beamten- und Soldatenkinder, die sich nach dem Abitur für ein Studium an den UniBw entschließen, unterscheiden sich beispielsweise um über eine halbe Examensnote von Arbeiterkindern, erlangen also ein deutlich besseres Examen.

Tabelle 75:  
Mittelwertvergleich der Hauptdiplomnoten in Abhängigkeit des väterlichen Berufs

Beruf Vater	N	Mittelwert Hauptdiplomnote	F-Wert	Signifikanz
Arbeitsloser	3	1,56		
Beamter	7	1,78		
Soldat	6	1,93	1,844	.112
Angestellter	53	2,10		
Selbständig	14	2,24		
Arbeiter	15	2,33		

Signifikanz (2-seitig) .05

Im Falle der dichotomen Merkmale Geschlecht und Universität (Hamburg versus München) wird der t-Test bei unabhängigen Stichproben herangezogen, deren Ergebnisse der Tabelle 76 und Tabelle 77 zu entnehmen sind. Die Überprüfung der Hypothese 6, ob Männer und Frauen sich signifikant bezüglich der Hauptdiplomnote unterscheiden, führt zu keinem signifikanten Ergebnis hinsichtlich eines Kriteriumswertunterschieds. Bezogen auf das Geschlecht erweisen sich in Bezug auf den OPZ-Gesamteignungsgrad die weiblichen Offizierbewerber außerdem als häufiger (52 Prozent) gut geeignet als ihre männlichen Mitstreiter (37 Prozent), die zu 60 Prozent mit geeignet beurteilt wurden.

Tabelle 76:  
Mittelwerte für das Kriterium Studienerfolg in Form der Hauptdiplomnote (falls schon vorhanden) zwischen männlichen und weiblichen Studenten

Geschlecht	N	Mittelwert Hauptdiplomnote	Signifikanz
Männlich	135	2,06	Varianz sind gleich ,942
Weiblich	13	2,07	Varianzen sind ungleich ,932

Signifikanz (2-seitig) .05

Ebenso wenig wie das Geschlecht beeinflusst der Studienort die Hauptdiplomnote in die eine oder andere Richtung, was zu einer Beibehaltung der Nullhypothese 8 führt.

Tabelle 77:  
Mittelwerte für das Kriterium Studienerfolg in Form der Hauptdiplomnote (falls schon vorhanden) zwischen Studierenden der Universität der Bundeswehr in Hamburg und der in München

Studienort	N	Mittelwert Hauptdiplomnote	Signifikanz
Hamburg	92	2,07	Varianz sind gleich ,760

---

München	56	2,04	Varianzen sind ungleich	,751
---------	----	------	-------------------------	------

---

Signifikanz (2-seitig) .05

## 11.5 Prozessevaluation

Die vorliegende Arbeit strebt neben den vorgebrachten Grob- und Feinzielen auch eine nachhaltige Wirkung für die Praxis der Offizierauswahl an. In diesem Zusammenhang wichtig zu untersuchende Aspekte stellen demnach auch die Akzeptanz, Augenscheinvalidität des Verfahrens dar, die mit Hilfe eines Fragebogens und zahlreicher Feedbackgespräche mit Testpersonen ( $n = 312$ ) sowie Unterhaltungen mit Lehrpersonal und militärischen wie zivilen Vorgesetzten bis hin zur Universitätsleitung untersucht werden. Die Akzeptanz ist ein häufig erwähntes Kriterium für Zulassungsinstrumente. So werden aus Sicht der Bewerber Studierfähigkeitstests trotz ihrer beunruhigenden, unangenehmen und unsympathischen Wirkung noch vor Auswahlgesprächen und Schulnoten sowie Studierfähigkeitstests noch vor Intelligenztests bevorzugt (Hell & Schuler, 2005). Akzeptanzbekundungen variieren im Laufe der Zeit. Gehörte in den neunziger Jahren noch die Abiturnote zu dem akzeptiertesten Prädiktor, nehmen heute der Studierfähigkeitstest, das Auswahlgespräch und die Schulnoten die Plätze eins bis drei ein.

### 11.5.1 Biographischer Fragebogen

Ohne schriftliche Einwilligung zur freiwilligen Untersuchung und Verarbeitung der Testdaten erfolgt keine Testung. Der in Anhang B abgebildete biographische Fragebogen dient zum einen der Unterstützung der Testleiter, zum anderen liefert er über die Bereitstellung relevanter biographischer und studien- wie bundeswehrbezogener Rahmendaten sowohl dem Probanden Hintergrundinformationen als auch dem Psychologen eine erste Grundlage für spätere Berechnungen. Darüber hinaus bildet die Einverständniserklärung die datenschutzrechtliche Belehrung und sichert der Testperson die notwendige Anonymität der persönlichen Daten zu. Die Projektleitung (Autorin der vorliegenden Arbeit bzw. Vertreterin der Gruppe Wehrpsychologie im Streitkräfteamt) gewährleistet auf diese Weise strengste Einhaltung der Sicherheitsauflagen für anonymisierte Qualitätssicherung und sichert die unwiderrufliche Löschung jeglichen Personenbezugs nach erfolgter Zusammenführung der Test- mit den Lebenslaufdaten aus den bundeswehreigenen Personalführungssystemen PERFIS, BISON und SASPF zu. Die freiwilligen Testpersonen der beiden Pilotphasen an den UniBw unterzeichneten ausnahmslos die Einverständniserklärung des biographischen Fragebogens.

## 11.5.2 Akzeptanzfragebogen

In den Forschungsfragestellungen aus Kapitel 10 wurde anhand der Prozess- und Postevaluation nach möglichen Schlüssen zur Akzeptanz und Qualität des neuen Verfahrens gefragt. In den Anhängen x bis x sind sowohl der Akzeptanzfragebogen als auch alle nach Kategorien geordneten offenen Antworten auf die offene Frage „gibt es etwas, dass Sie uns darüber hinaus gerne mitteilen wollen?“ abgebildet. Der zur Evaluation entwickelte und im Anschluss an die Testung und das Feedbackgespräch ausgehändigte und von nahezu allen ( $n = 294$ ) Testpersonen ausgefüllte Fragebogen gliedert sich in die drei Teile „Durchführung und Organisation der Pilotphasen“, „Inhalt und Qualität der gesamten Testbatterie“ und „Akzeptanz der Persönlichkeitsfragebögen“. Die Beurteilung des zweiten Teils ist in Antwortalternativen für den allgemeinen Teil I, den allgemeinen Teil II und den spezifischen Teil des Reihungstests differenziert. Neben Aussagen zum Interview wurde die Akzeptanz über Fragen zu Layout, Computertestung, Zeitfaktor, Inhalte des allgemeinen, spezifischen und Persönlichkeitsteils, über generelle Aussagen zur Studienberatung OPZ, über Personal und äußere Bedingungen des Projektes untersucht. Alle geschlossenen Fragen sind auf der den Testpersonen bereits vertrauten Siebener-Skala von 1 „ich stimme gar nicht zu“ bis 7 „ich stimme vollkommen zu“ zu bewerten.

Mit Durchführung und Organisation der Pilotphasen waren die Mehrheit der Testpersonen sehr zufrieden. Die Abbildung 53 und Abbildung 54 dokumentieren, dass sich die Teilnehmer vollständig informiert fühlten und zum größten Teil eine computergestützte Testung am Laptop als anwenderfreundlich wahrnehmen und beurteilen.

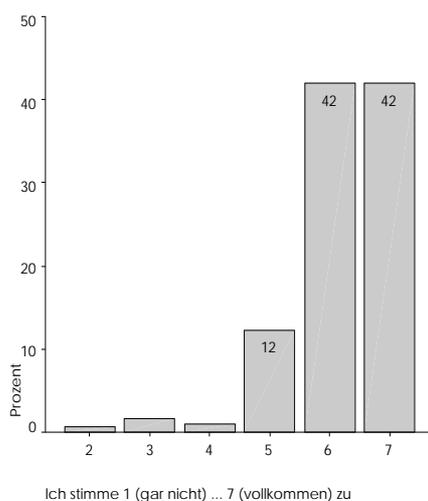


Abbildung 53: Häufigkeitsverteilung der Aussage „Ich wurde über den Hintergrund der Testung ausreichend informiert.“

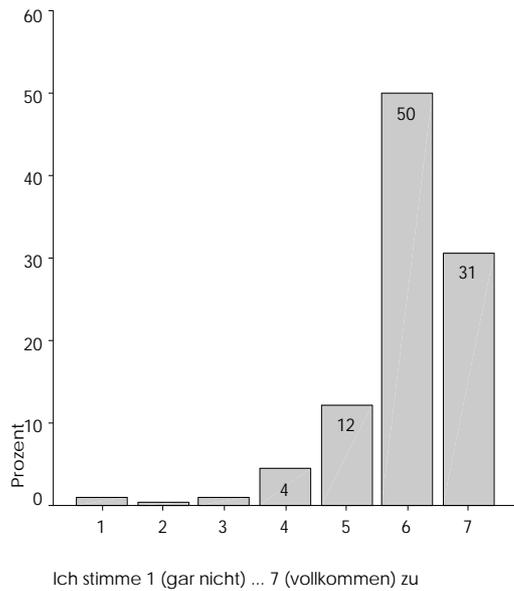


Abbildung 54: Häufigkeitsverteilung der Aussage „Die Testung am Laptop ist einfach und anwenderfreundlich.“

Abbildung 55 spiegelt die Augenscheidvalidität des neuen Verfahrens hinsichtlich seiner Fähigkeit wider, die Potentiale des jeweiligen Kandidaten treffend zu ermitteln. Lediglich 12 Prozent schreiben ihm eine mittelmäßige bis gar nicht vorhandene Konstruktvalidität zu.

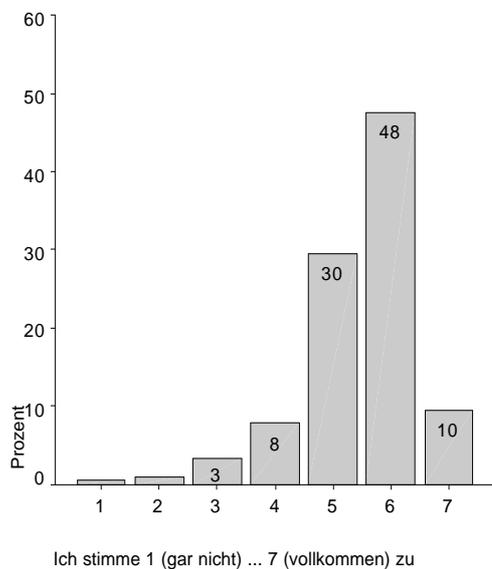


Abbildung 55: Häufigkeitsverteilung der Aussage „Meine Potentiale wurden durch die Testung treffen ermittelt.“

Als ähnlich gut erweist sich die Face Validity des Verfahrens bezogen auf die Fähigkeit, Studieneignung messen zu können. 68 Prozent der Befragten stimmen dieser Aussage stark bis vollkommen zu; nur 12 Prozent stimmen schwach bis gar nicht zu. Wie in Abbildung 57 veranschaulicht halten ebenso 68 Prozent der Testpersonen die Fachspezifität des Verfahrens als valide.

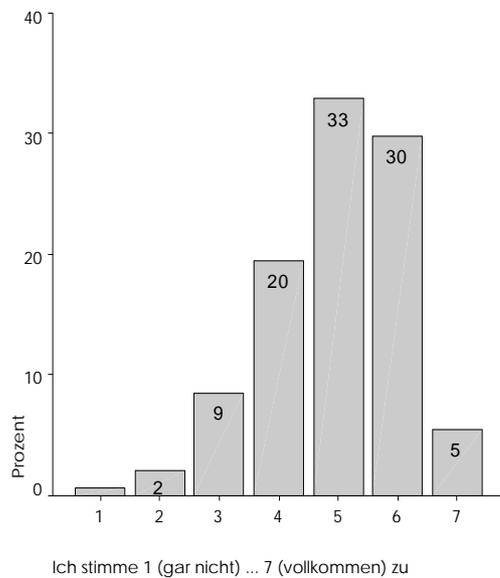


Abbildung 56: Häufigkeitsverteilung der Aussage „Mit diesem Testverfahren hat man ein geeignetes Instrument zur Feststellung der Studieneignung entwickelt.“

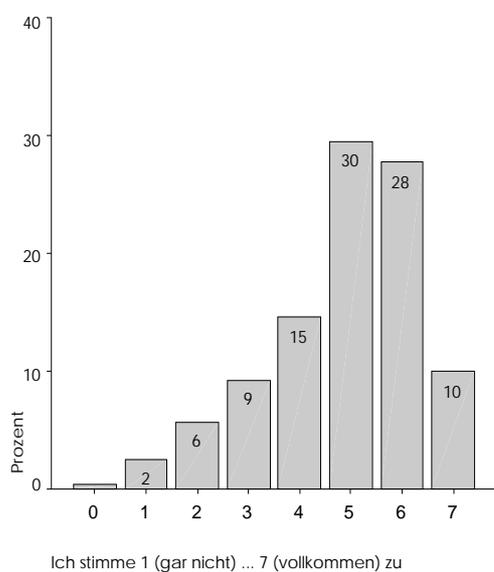


Abbildung 57: Häufigkeitsverteilung der Aussage „Mit diesen Aufgaben kann man eine Studieneignung für meinen Fachbereich prüfen.“

Ziel ist es, ein optimiertes Verfahren zu entwickeln, das sich positiv von der früheren bzw. derzeitigen Studienberatung unterscheidet. Abbildung 58 bildet ab, dass über 80 Prozent der Testpersonen die neu entwickelte Testbatterie als standardisierter und aussagekräftiger betrachtet.

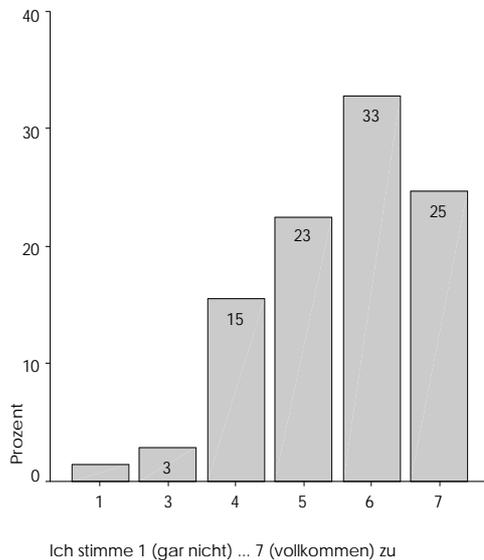


Abbildung 58: Häufigkeitsverteilung der Aussage „Die studienfeldspezifische Testbatterie ist standardisierter und aussagekräftiger als die bisherige OPZ-Studieneignungsprüfung.“

Abbildung 59 dokumentiert Fachspezifität und fachliche Relevanz des spezifischen im Vergleich zu den beiden allgemeinen Teilen der Testbatterie. Beim Vergleich der allgemeinen Teile ist die Tatsache hervorzuheben, dass dem Persönlichkeitsfragebogen mehr Studenten die Ermittlung studienrelevanter Fähigkeiten zuschreiben als dem allgemeinen Intelligenzteil zur Messung allgemeiner akademischer Intelligenz.

### Allgemeiner Teil I

### Allgemeiner Teil II

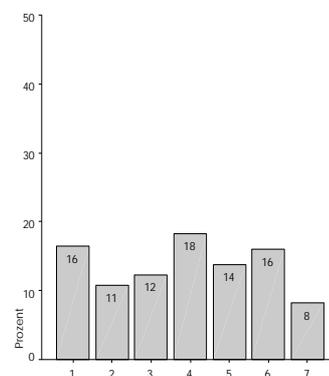
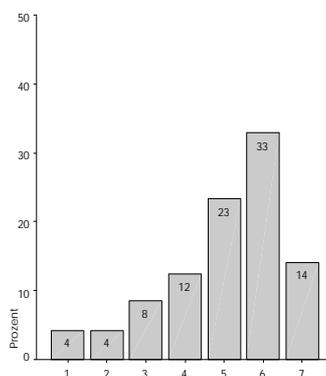
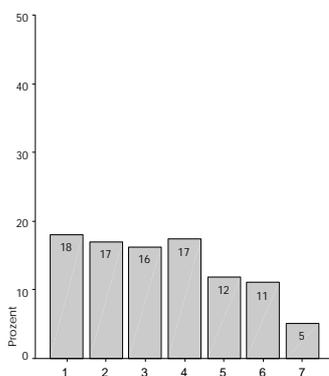


Abbildung 59: Häufigkeitsverteilungen im allgemeinen Teil I, im spezifischen Teil und im allgemeinen Teil II zur Aussage „Die Aufgaben erinnerten mich an wichtige Inhalte bzw. an verlangte Fähigkeiten meines Studiums“ auf einer Skala von 1 („ich stimme gar nicht zu“) bis 7 („ich stimme vollkommen zu“)

Die Rolle und Wichtigkeit von Persönlichkeitseigenschaften wird auch mittels Abbildung 60 dokumentiert. Über die Hälfte (51 Prozent) der Probanden weist weichen Faktoren einen grundlegenden Einfluss auf den Studienerfolg zu. Gleichzeitig wird die Möglichkeit sozial erwünschter Antwortgebung von 32 Prozent der Befragten als wahrscheinlich (starke bis vollkommene Zustimmung der Aussage) empfunden (vgl. Abbildung 61).

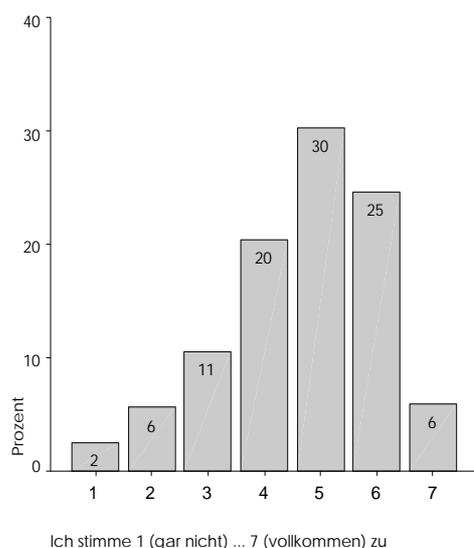


Abbildung 60: Fragen zur Persönlichkeit bilden grundlegende Eigenschaften für ein erfolgreiches Studieren ab

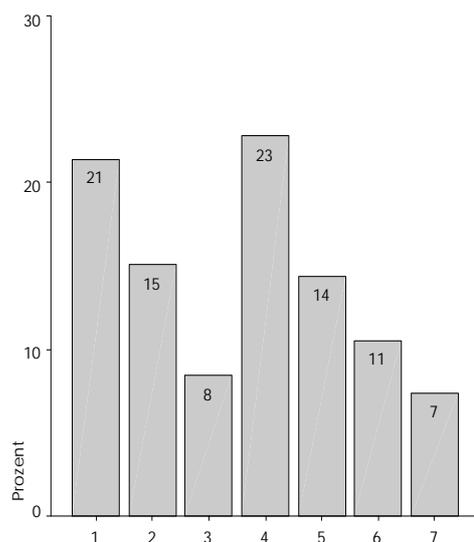


Abbildung 61: Häufigkeitsverteilung der Aussage „Ich frage mich, wer diese Fragen als Bewerber ehrlich beantworten würde“

Eine in Tabelle 78 ausgeführte Zufallsauswahl aus den offenen Antworten auf die Frage nach allgemeinem Feedback dokumentiert die hohe Akzeptanz der neuen Verfahren und Kritikpunkte an der damaligen bzw. noch laufenden Studienberatung der OPZ.

Tabelle 78:

*Zufallsauswahl der offenen Antworten auf die Frage „Gibt es etwas, dass Sie uns darüber hinaus gerne mitteilen wollen?“*

#### **Aufgaben / Inhalte spezifischer Teil**

Gewichtung im Ingenieur-Testteil ist unausgewogen. Matheteil zu lang dafür physikalisches Verständnis zu kurz.

BWL-Fragenteil hat hohes Niveau - garantiert gute Auswahl; Technik-Teil sehr bzw. zu anspruchsvoll, da im Studium eher reproduzieren als kreative Eigenleistung gefragt ist.

Die fachliche Anforderungen im spezifischen Teil BWL/VWL sind recht hoch - erfordern einige Erfahrung mit Wirtschaftszeitungen. Ansonsten recht fein! - Hier werden Fähigkeiten wie bei der PISA-Studie abverlangt - deswegen für dt. Abiturienten sehr schwer - Keine reine Mathematik und weniger Logikfragen - sehr gut.

#### **Aufgaben / Inhalte Persönlichkeitsteil**

Persönlichkeitsprofil entspricht meinen Beurteilungen "super".

Guter Persönlichkeitstest.

Fragen in Teil III zu allgemein formuliert, jeder assoziiert andere Situationen.

Die Fragen vom Teil 3 waren zu allgemein und mir weitestgehend unmöglich generell zu beantworten. Es wäre besser, die Testperson mit konkreten Beispielsituationen zu konfrontieren.

Ich hätte gerne einen Ausdruck über mein Persönlichkeitsprofil mitgenommen!

#### **Aussagen zur bisherigen Studienberatung**

Die Studienberatung an der OPZ war alles andere als eine Beratung. Ich wurde nur über studienspezifisches Wissen abgefragt und anschließend ausgezählt, weil ich dieses Detailwissen nicht besaß.

"Alte" Studienberatung habe ich als sehr subjektiv empfunden (2-3 Fragen und dann wurde geurteilt), deshalb ist der "neue" Test wesentlich besser!

Scheint ein großer Fortschritt zum alten Verfahren zu sein. Bei diesen wurde mir ein Scheitern aufgrund mangelnden Interesses prognostiziert, sollte es aber trotzdem probieren und dann möglicherweise durchfallen. Der Studiertest war bisher meiner Meinung nach ein absolutes nichts sagendes Testverfahren.

Ich frage mich manchmal ernsthaft, in wieweit die (bisherigen) Verfahren zur Feststellung der Studieneignung und der Persönlichkeitsmerkmale tauglich sind. Ich (Leutnant des Heeres, Geschichte 2003) frage mich häufig, wie es dazu kommen konnte, dass so genannte Kameraden einen Studienplatz erhalten haben, die weder intellektuell noch charakterlich zum Offizierberuf geeignet sind. Möglicherweise wäre es hilfreich, das Testverfahren um praktische Anteile zu erweitern. - Z.B. um einen Hindernisparcour, der in kleinen Trupps überwunden werden muss. - Aber das ist ja nur ein Vorschlag und richtet sich nicht primär an den psychologischen Dienst Bw. Es wäre aber evt. geboten, dass Sie Ihren Einfluss in Ihrer Richtung geltend machen.

#### **Aussagen zur Studierfähigkeitstestung**

Das Bestehen des Studierens hängt bei den meisten Offizieren nicht vom Intellekt ab, eher scheitert es meist am Grund der Faulheit.

Die Eignung zum Kampftruppenoffizier sollte unabhängig von der Studierfähigkeit geprüft werden und im Vordergrund stehen.

Studienerfolge regelmäßig überprüfen; Studienabbrecher (vorsätzliche ~) 12 Jahre in der Bundeswehr halten und nicht vorzeitig entlassen.

Ich hoffe, dass die erreichten Ergebnisse langfristig zu einer objektiveren Studienplanung führen. Und nicht im Endeffekt doch der Bedarf entscheidet.

Test als Ersatz für Studiengespräch, da Test spezifischer.

Ablauf der Studienberatung: Sinnvoll wäre eine vorherige Studienberatung (vor Testen in der OPZ), z.B. für Schüler ein Jahr vor dem Abschluss, damit den Schülern bewusst ist, was sie im Studium erwartet und ob eventuelle Wissenslücken aufgearbeitet werden müssten. Frage: Warum müssen Offiziere studieren.

#### **Personal / Ausstattung**

Sehr nette Mitarbeiter.

War gut; sehr interessant, sehr nette Leitung.

Arbeitsplatz vorne war unruhig, zu laut.

Ich empfinde es als eine Auswertung und Beratung zur Verbesserung der individuellen Fähigkeiten.

Positiv: angenehmes Arbeitsklima, nette Mitarbeiter.

Hinsichtlich der Inhalte der fachspezifischen Teile wird das Testanspruchsniveau als sehr hoch empfunden. Der Persönlichkeitsteil kam trotz seines großen Umfangs nach der mehrstündigen Testung überraschend positiv an. Der Wunsch nach Veränderung oder Verbesserung der bisherigen und als eher subjektiv empfundenen Studienberatung wird sowohl in den geschlossenen als auch offenen Antworten deutlich.

Nicht nur die Ergebnisse des zur Selbstevaluation eingesetzten anonymen Akzeptanzfragebogens, sondern auch das mündlich übermittelte äußerst positive Feedback der Leitung Studentenbereich, der Professoren und wissenschaftlichen Mitarbeiter und der Testpersonen können das Projekt als sehr erfolgreich bewerten lassen. Bereits vor der Pilotphase begegnete man der Projektleitung nicht nur mit großem Entgegenkommen bei der Durchführung der Erprobungsphase sondern auch mit dem Wunsch, das Projekt weg von einer reinen Statusdiagnostik hin zu einer längsschnitthaften Prozessdiagnostik auszudehnen. Die zusätzliche Ermittlung von Persönlichkeitseigenschaften und Arbeitshaltungen neben kognitiven Faktoren stieß aufgrund der besonderen Studienbedingungen einer Campus-Universität auf hohen Zuspruch.

### 11.5.3 Feedbackgespräch

Noch deutlicher wird die Sichtweise zur eigenen damaligen Studienberatung versus neuen Testbatterie während der im Durchschnitt dreißig Minuten andauernden Feedbackgespräche, in denen neben der Testergebnisrückmeldung auch nach den gemachten Erfahrungen mit der

Studienberatung OPZ gefragt oder unaufgefordert von Seiten der Testpersonen angesprochen wurde. Bis auf wenige Ausnahmen (circa zehn Prozent) derer Studenten, die sich entweder gar nicht äußerten oder zufrieden mit der Studienfachpassung sind, wurde durchweg die auf Wissensfragen und allgemeinen Testergebnissen beruhende Studienplatzvergabe kritisiert. Im Rahmen der Pilotphase stellte sich darüber hinaus bei Ermittlung der Kriteriumsvalidität (50 Personen unterzogen sich auch Tests fremder Fachbereiche) heraus, dass Offizieranwärter doch für den ursprünglich gewählten aber von Seiten Studienberatung nicht empfohlenen Studiengang geeignet gewesen wären! Dies spricht eindeutig gegen die in der Beratung üblicherweise verwendeten Wissensfragen und untermauert die im Bericht dargestellten Forschungsergebnisse. Bei einer überdurchschnittlichen Zahl von Offizieren stellte sich darüber hinaus heraus, dass die Frustration über die Zuordnung zu einem nicht gewünschten Fach das erfolgreiche Studieren stark behindert ist. Die Kritik vieler Testpersonen richtete sich während des Feedbackgesprächs auch gegen die in ihren Augen generell zu wünschen lassende intellektuelle und sogar kritisch gesehene charakterliche Eignung der eigenen Kameraden, die eine Verlagerung zukünftiger Beratung hin zur Berücksichtigung relevanter Persönlichkeitseigenschaften und fachspezifischer Eignung und allgemeiner Intelligenz unterstützen.

## D DISKUSSION

Der Gruppe Wehrpsychologie im Streitkräfteamt wurde im Herbst 2004 das Langzeitprojekt „Entwicklung geeigneter Prädiktoren zur Vorhersage der Studieneignung an den Universitäten der Bundeswehr“ auf Weisung des Referates PSZ III 6 im BMVg als Konsequenz der Evaluierung der OPZ übertragen. Übergeordnetes Ziel ist die Senkung der Abbruchquoten an den Universitäten der Bundeswehr und damit die Einsparung erheblicher Kosten. Im Schnitt brechen 30 Prozent das Studium ab und scheiden kurze Zeit später aus der Bundeswehr aus. Direkte Konsequenz ist die Optimierung der Studienberatung an der OPZ und demzufolge eine nachdrücklich verbesserte Passung des einzelnen Studenten auf das jeweilige Studienfach.

Zentrales Ziel der Inauguraldissertation ist im Rahmen der Personalpsychologischen Qualitätssicherung der Streitkräfte die Optimierung der Studienberatung im Eignungsfeststellungsverfahren der OPZ mittels einer studienfachspezifischen Reihungstestbatterie mit Modellcharakter, deren Evaluation und deren nähere Betrachtung hinsichtlich kognitiver wie nicht-kognitiver und moderierender Einflussfaktoren.

Es wurde eine umfangreiche Reihungstestbatterie (14 allgemeine und studienfeldspezifische Studierfähigkeits- und Persönlichkeitstests) für die Studienfelder Geistes-, Ingenieur- und Wirtschaftswissenschaften sowie Informatik zusammengestellt, neu entwickelt (Textverständnistest I und II) und weiter entwickelt (Allgemeiner Teil II). Dabei entstanden ein eigens konstruierter Textverständnistest für die Geisteswissenschaften sowie der neue Persönlichkeitstest Studieneignung. Alle vorhandenen Papier-Bleistift und neuen Tests wurden auf die eine mobile CAT-Anlage programmiert und überführt, vor Ort erprobt sowie in zwei dreiwöchigen Pilotphasen an den beiden UniBw evaluiert (Stichprobengröße Gesamt N = 319 / Hamburg n = 156 / München n = 163). Das Interesse, die Akzeptanz und Erwartungshaltung hierzu erwiesen sich bereits im Vorhinein nicht nur von Seiten der beiden Präsidenten der UniBw sondern auch seitens der Leitung der Studentenbereiche sowie des zivilen wissenschaftlichen Personals (Dekane, Professoren, Dozenten etc.) und der studierenden Offizieranwärter und Offiziere als außerordentlich hoch. Nicht zuletzt aus diesem Grunde war es möglich, die Pilotphasen im Gegensatz zu vergleichbaren Studien in der Vergangenheit mit ausschließlich im Vorhinein geworbenen freiwilligen Testpersonen durchzuführen.

Im Folgenden werden die Ergebnisse einer methodischen und inhaltlichen Bewertung unterzogen. Daran schließt sich eine Einordnung in den Kontext bisheriger Befunde und aktuellen Theorien an. Überdies werden die Ergebnisse hinsichtlich ihres Betrags zur aktuellen Forschung bewertet und damit verbundene Fragestellungen diskutiert.

## 12. Evaluation der Ergebnisse

Im Zentrum einer Überprüfung der neuen Testbatterie Studieneignung für eine optimierte Studierendenauswahl an der OPZ stand zunächst der Nachweis der Reliabilität als notwendige Voraussetzung für das Vorhandensein von Validität. Um den Grad der Genauig- und Messfehlerfreiheit zu ermitteln, wurden eine umfangreiche Itemanalyse getätigt und Normen festgelegt. Des Weiteren wurde die Konstruktvalidität zunächst mittels des methodischen Paradigmas einer explorativen Faktorenanalyse überprüft. Daran schloss sich eine Exploration und Bestätigung der Ergebnisse mit Hilfe des nomologischen Netzwerks an. Im nächsten Schritt wurde die prognostische Validierung durchgeführt über den Vergleich neuer und alter Prädiktoren mit dem für den Studienerfolg stehenden Kriterien Vor- und Hauptdiplomnote. Um dem Einfluss verschiedenster moderierender Hintergrundvariablen zu kontrollieren, wurden schließlich dahingehend Zusammenhänge in Form von Hypothesen angenommen und inferenzstatistisch untersucht. Um auch einen augenscheinlichen Bezug des Verfahrens zu den Anforderungen eines Studiums an den UniBw und eine praktische Umsetzbarkeit zu belegen, wurden abschließend Akzeptanz, Augenscheinvalidität und gleichzeitig die Qualität der Pilotphasen per se evaluiert.

### 12.1 Diskussion der Ergebnisse aus methodischer Sicht

Die methodische Überlegenheit psychologisch gestützter Ausleseverfahren beruht sowohl auf der Anwendung wissenschaftlich fundierter Verfahren, als auch auf der Evaluation diagnostischer Entscheidungen. Einige methodische Schwierigkeiten sollen hier bezogen auf die Testgütekriterien und darüber hinaus kritisch beleuchtet werden.

#### 12.1.1 Reliabilität

Zur Reliabilitätsuntersuchung im Rahmen der vorliegenden Arbeit lassen sich zwei hauptsächliche Kritikpunkte anmerken. Zum einen benötigt man zur Berechnung der inneren Konsistenz neben einer großen Stichprobe (ab circa 50 Personen), zum anderen eine Personen-

gruppe, die in ihrer Zusammensetzung einander möglichst ähnlich ist (Kanning, 2004). Für den Fachbereich Informatik, der lediglich an der UniBw München vertreten ist, konnte die für eine Berechnung von Cronbachs Alpha geforderte Stichprobengröße mit lediglich 30 Testpersonen nicht erfüllt werden. Da es sich bei der vorliegenden Untersuchung um eine Messung von stabil ausgeprägten Merkmalen handelt, wäre eine Berechnung der Retest-Reliabilität über die Zeit hinweg zusätzlich zu empfehlen gewesen. Aus Effizienzgründen wurde jedoch darauf verzichtet (vgl. Kriteriumsvalidität). Im vorliegenden Falle wurde bei der Normenfindung einer möglicher Überschätzung der Testergebnisse dahingehend begegnet, dass sowohl bei Schülern durchgeführte Ergebnisse der BA als auch Testscores der bereits an den UniBw studierenden Testpersonen Berücksichtigung fanden.

### 12.1.2 Konstruktvalidität

Hinsichtlich der Überprüfung der Konstruktvalidität lässt sich Folgendes festhalten: Die Hauptkomponentenanalyse wurde gewählt, um eine objektive Bestimmung der konvergenten und diskriminanten Validität zu erlauben. Speziell die Faktorenanalyse leistete einen Beitrag zur *Konstruktaufklärung*. Laut Bortz (1993) ist aber eine exakte Interpretation der Faktorenanalyse als schwierig zu bewerten. Ihre Ergebnisse sind als „hypothesengenerierend“ zu verstehen und erlauben keine direkte Überprüfung von Annahmen. Diese angesprochene Problematik lässt die Frage nach alternativen Auswertungs- bzw. Überprüfungsmethoden, wie nach weiteren varianz-, faktoren- und pfadanalytischen Verfahren laut werden. In der vorliegenden Arbeit wurde die Bestimmung des nomologischen Netzwerks gewählt.

### 12.1.3 Kriteriumsvalidität

Studienerfolg ist nur bedingt vorhersagbar, da unzählige und im zeitlichen Verlauf unablässig einfließende komplexe Wechselwirkungen zwischen Person- und Situationsmerkmalen nicht kontrolliert werden können. Hinsichtlich der Untersuchung der prognostischen Validität wäre optimalerweise etwa eine zeitlich unverfälschte Messung in der OPZ und während der Offizierausbildung bzw. zu Beginn des Studiums gewesen. Für eine vollständige Evaluation wären eine oder mehrere Kohorten von Studierenden, Abbrechern und ggf. abgelehnten Bewerbern zu begleiten gewesen. Abgesehen von der Tatsache, dass der Datenschutz eine Kontrolle von aus der Bundeswehr Ausgeschiedenen (Abbrecher) oder gar nicht Angenommenen unmöglich macht, verzögerte sich außerdem bei Beobachtung Studierender gerade aufgrund der dem Studium zeitlich voran geschalteten Truppenoffizierausbildung eine Be-

rechnung der prädiktiven Validität dann um circa sechs Jahre. Erfahrungsbedingt ändern sich in einem derart langen Zeitraum auch OPZ-Auswahlmethoden oder Studienbedingungen wie die Bachelor- und Master-Studiengangeinführung. Eine nicht zu unterschätzende Argumentationsgrundlage liefert ein einfacher Kosten-Nutzen-Vergleich, der im vorliegenden Falle über eine grobe Abschätzung des Aufwands geschah und zur Entscheidung einer zum Kriterium zeitnahen Untersuchung führte.

Ferner sollte bei einer methodischen Diskussion der Vorhersagevalidität die Fehleranfälligkeit des Kriteriums Berücksichtigung finden. Die Indikator-Kriterienproblematik wurde in Kapitel 2.3.1 bereits ausführlich erörtert. Die exakte Klärung des ultimativen Kriteriums Studienerfolg ist nicht hundertprozentig möglich. Daher orientiert man sich in der Praxis oft an einer Vielzahl weicher und harter Kriterien. Wie Thorndike bereits 1947 empirisch belegte (vgl. Kapitel 2.3), nimmt die Vorhersagegüte von den unmittelbaren zu den endgültigen Kriterien aus mehreren Gründen ab. Zum Beispiel beeinflussen viele unkontrollierte und nicht vorhersehbare Variablen aus dem persönlichen und beruflichen Bereich den Prädiktor-Kriterium-Zusammenhang. Am besten operationalisierbar und messbar sind sicherlich die Kriterien Vor- und Hauptdiplomnote, auf die sich in der vorliegenden Arbeit beschränkt werden musste. Bedeutsam wären jedoch auch übrige Kriterien wie Studienabbruch, Studienzufriedenheit, Studienfachwechsel usw. gewesen, welche sich aber aus Datenschutzgründen gar nicht oder nur höchst aufwendig sowohl mittels subjektiver als auch objektiver harter Daten messen ließen. Aus Gründen der Praktikabilität konnte nur ein realistischerweise erhebbares Kriterium befürwortet werden.

Ferner sollte die Stichprobenproblematik beachtet werden. Weder Studienabbrecher noch die für den Offizierberuf als nicht geeignet bewerteten Bewerber fanden Berücksichtigung. Dadurch, dass zum Zeitpunkt der Evaluation Studenten aus verschiedenen Trimestern zur Testung zugelassen wurden, um eine hohe Stichprobengröße pro Fachbereich zu erzielen, wird das Ergebnis gegebenenfalls trotz des Messens stabiler Merkmale geringfügig verzerrt. Des Weiteren sind aufgrund der teilstreitkraftspezifisch unterschiedlich lang andauernden Offizierausbildungszeiten (für die damalige Testgruppe Heer 3 Jahre, Luftwaffe und Marine 15 Monate) identische Lernvoraussetzungen nicht garantiert.

Zurzeit müssen, um an personen- bzw. laufbahnbezogene Daten und Beurteilungen der Offizierbewerber zu gelangen, mehrmonatige Genehmigungsverfahren in Kauf genommen werden. Darüber hinaus verlängert sich die Vorbereitungszeit einer Bewährungskontrolle innerhalb der Streitkräfte dadurch, dass Daten nicht kontinuierlich „gepflegt“ werden, vielleicht

weil es drei unabhängige und nicht immer identische Datenspeicherbanken gibt (PERFIS, BISON und SASPF). Ferner bleiben bei Erhalt der Daten meist in Schriftform oder computerisiert in Excel-Format alle unabhängigen und abhängigen Variablen solange unverwendbar, bis die jeweiligen Schlüssel über lange Wege beschafft sind.

#### 12.1.4 Weitere methodische Einflüsse und Fazit

Unter methodischen Gesichtspunkten treten ferner eine Reihe weiterer bekannter Probleme auf: Fragliche Skalendignität des Kriteriums und der diagnostischen Instrumente, Stichprobenselektion mit der Folge reduzierter Streuungen, Interaktionen mit Moderatorvariablen, unbeobachtete Situationsbedingungen etc. Unter streng methodischen Aspekten müssten Evaluationsstudien theoretisch alle Offizierbewerber und über den gesamten Zeitraum bis zum Studien- oder sogar Berufsende berücksichtigen, um unverfälschte Beziehungen zwischen Daten und Kriterien berechnen zu können. In der Praxis ist dies aber kaum umsetzbar. Außerdem variiert die Qualifikation der Offizieranwärter stark in Abhängigkeit externer Einflüsse und führt zu einer Varianzeinschränkung der Stichprobe und zu verminderten Korrelationen. Bei der Bundeswehr stellt sich einerseits das Problem, dass sich in Zeiten hoher Arbeitslosigkeit auch überdurchschnittlich viele schwächere Kandidaten minderer Qualifikation und Überzeugung bewerben. Inwiefern das auch für sehr gute Schulabgänger gilt, bedarf einer weiterführenden Untersuchung. Andererseits resultiert unabhängig von Grund- und Selektionsrate eine wesentliche Varianzeinschränkung aufgrund der als geeignet bewerteten und angenommenen bzw. der als ungeeignet klassifizierten und abgelehnten Offizierbewerber. Eine ebenfalls bedeutsame Schwierigkeit stellt sich für die Streitkräfte dann, wenn auch innerhalb der Gruppe der Angenommenen ein Selektionsprozess über die Zeit stattfindet. Sehr gute Offizieranwärter werden nach einiger Zeit sehr weit aufgestiegen oder in die freie Wirtschaft abgewandert sein und daher für Evaluationsstudien nicht mehr zur Verfügung stehen. Prinzipiell müssen eingeschränkte Stichproben nicht zwangsläufig zu einer Verzerrung der Evaluationsergebnisse führen. Sie wirken sich aber immer dann negativ auf Korrelationen zwischen Prädiktoren und Kriterien aus, wenn die Extrema der Verteilung abgeschnitten werden.

Bezüglich der Skalenniveaus wurde versucht, diese im Rahmen der Arbeit zu überprüfen, wenn nötig deren Voraussetzung statistisch abzusichern und, entsprechend dem Beispiel der Untersuchung moderierender Hintergrundvariablen, mit nonparametrischen Überprüfungstechniken zu begegnen.

Der Vorteil der vorliegenden Untersuchung war ihre Freiwilligkeit und somit anzunehmende fehlende Einflüsse des Antwortverhaltens auf die Persönlichkeitsfragen des Allgemeinen Teils II der Testbatterie Studieneignung. In der späteren Praxis ist von sozial erwünschtem oder anderweitig verzerrtem Antwortverhalten auszugehen.

Abschließend ist kritisch anzumerken, dass mit Kontrollverfahren, die auf statistischen Maßzahlen basieren, nicht zwangsläufig das ganze Spektrum möglicher Schwachstellen bzw. Beurteilungsfehler trennscharf sichtbar gemacht werden können. Überdies wird der Beurteilungsprozess neben systematischen Fehlern häufig auch durch zufallbedingte Faktoren beeinträchtigt, die schwer identifizierbar und demzufolge auch nicht korrigierbar sind.

## 12.2 Zusammenfassung und Diskussion der Ergebnisse aus empirischer und theoretischer Sicht

Die Ergebnisse werden in den nächsten Unterkapiteln nach den in Kapitel 10 aufgeführten Forschungsfragestellungen geordnet zusammengefasst und diskutiert. Ihre Diskussion geht unter anderem auf bereits bestehende empirische und im Theorieteil referierte Forschungsergebnisse ein. Es wird versucht, die zur Prognose der Studieneignung relevanten Prädiktoren zu filtern. Die Ergebnisse zur Reliabilität, Kriteriums- und Konstruktvalidität und zum Einfluss moderierender Hintergrundvariablen erheben den Anspruch, nicht nur Schwachpunkte, sondern insbesondere Verbesserungspotentiale sowohl bisheriger als auch neuer Verfahren zur Studieneignungsfeststellung angehender deutscher Offiziere aufzuzeigen und leiten über auf Implikationen für die Praxis und Forschung (Kapitel 13).

### 12.2.1 Reliabilität

Wie genau und messfehlerfrei wird das intendierte Konstrukt Studienerfolg mit Hilfe der Prädiktoren im Vergleich zur Abiturgesamt- bzw. Schulabschlussnote gemessen?

Das Fachkonzept Personalpsychologische Qualitätssicherung legt die Mindeststandards für Einführung, Nutzung und anwendungsbegleitende Kontrolle personalpsychologischer Instrumente und Verfahren im psychologischen Dienst der Bundeswehr fest (Puzicha, 1997). Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen wurden 2002 in der DIN 33430 hinterlegt. Diese Norm stellt Qualitätsstandards zu Inhalt und Ablauf von Personalauswahl- und -entwicklungsverfahren auf. Neben dieser Norm dienen auch die in den personalpsychologischen Standards der Bundeswehr festgelegten Mindest-

maße der jeweiligen Testgüte als Entscheidungsgrundlage für oder gegen einen Untertest. Die Maximierung des Nettonutzens spielt bei der Auswahl der am besten geeigneten Tests und Fragebögen dabei eine entscheidende Rolle. Reliabilitätswerte von mindestens Cronbachs Alpha = .70 sind akzeptabel, Werte größer als .85 gelten als wünschenswert aber selten realisierbar. Wirft man einen Blick auf die Reliabilitäten der einzelnen untersuchten Subtests und deren Itemkennwerte wie Schwierigkeit und Trennschärfe, schneiden folgende Untertests hinsichtlich ihrer Reliabilität zufrieden stellend und gut ab: Die Subtests Sprachensysteme, Modellanalyse, Diagramme und Tabellen und technisches Verständnis. Über eine sehr gute Reliabilität verfügen die Untertests Figuren erkennen, verbale Analogien, Formalisieren technischer Zusammenhänge, PMO, Big Five, Stressbewältigungsressourcen. Mit Sicherung einer mindestens akzeptablen Reliabilität erfüllt sich die zu fordernde Voraussetzung für das Vorhandensein einer Testvalidität.

Ein Einsatz des gesamten Untertests bzw. des Gros der vorhandenen Items ist jedoch bei einigen der Prädiktoren zumindest langfristig zu überdenken. Die Itemkennwerte der neu entwickelten Tests Textverständnis I und Textverständnis II, sowie die BA-Subtests Flussdiagramme und formale Systeme entsprechen nur zum Teil oder gar nicht den Anforderungen des Personalpsychologischen Qualitätssicherungskonzepts der Bundeswehr und genauso wenig den Bestimmungen der DIN 33430. Im Vergleich zu den erhobenen Testgütemaßen im Rahmen von Studien der BA schneiden ergeben sich unterschiedliche Ergebnisse. Der Studienberatungstest Informatik - insbesondere der Untertest „Flussdiagramme“ – erzielte bei Berechnungen der BA positiv zu bewertende Maße. Frühere Pilotphasen an den UniBw erbrachten demgegenüber nicht akzeptable Testgütewerte.

Zusammenfassend lässt sich sagen, dass mit Ausnahme des spezifischen Testteils des Fachbereichs Informatik und der beiden neu entwickelten Textverständnisaufgaben für den geisteswissenschaftlichen Anteil Genauigkeit und Messfehlerfreiheit des Verfahrens Studieneignung in ausreichendem Maße vorhanden sind.

### 12.2.2 Konstruktvalidität

Wie gut werden die intendierten Konstrukte mittels untersuchter Reihungstests bzw. mittels der einzelnen Subtests tatsächlich gemessen?

Die Reihungstestbatterie Studieneignung erweist sich mit Hilfe der beiden aufgezeigten Auswertungsmethoden entsprechend der Annahmen als hinreichend konstruktvalide.

Die Ergebnisse aus der Faktorenanalyse, deren varimax-rotierte Ladungsmatrix deutlich auf den einzelnen spezifischen Testteilen lädt, veranschaulichen, dass die spezifischen Testverfahren mit Ausnahme des Informatikteils Anteile allgemeiner Intelligenzfaktoren wie Wahrnehmungsgeschwindigkeit oder logisches Denken (gemessen im allgemeinen Teil I) – nur in ihrem Falle fachbezogen – prüfen. Die spezifischen Untertests decken zu einem gewissen Grad demnach die Inhalte der Subtests „verbale Analogien“ und „Figuren erkennen“ ab. Wie im theoretischen Teil detailliert beschrieben, nehmen die kognitiven Faktoren zwar eine herausragende Rolle beim Einfluss auf die Studierfähigkeit ein. Nicht zu vernachlässigen sind jedoch die engen Verknüpfungen zwischen Studieneignung und persönlichen Arbeitshaltungen und weiteren Persönlichkeitseigenschaften. Die Hauptkomponentenanalyse des allgemeinen Teils II deckt fünf Faktoren auf, deren erster, welcher 18 Prozent der Varianz aufklärt, wichtige Arbeitshaltungen wie Leistungsmotivation, Gewissenhaftigkeit, Selbstdisziplin, internale Kontrollüberzeugung einschließt, die bereits in viel zitierten Untersuchungen als außerordentlich relevant für ein erfolgreiches Studieren entdeckt wurden. Aufgrund dieser Untersuchungen werden im Rahmen dieser Arbeit im zweiten Faktor Forschungsergebnisse dahingehend repliziert, als dass Persönlichkeitsmerkmale wie z.B. emotionale Stabilität (geringe Ausprägung von Neurotizismus) oder geringe externale Kontrollüberzeugung oder ein hohes Engagement und eine ausgeprägte Selbstwirksamkeitsüberzeugung ausschlaggebend für ein gutes Examen aufgedeckt werden. Der dritte Faktor betont Teamfähigkeit und Bewältigungsverhalten als bedeutungsvoll für ein erfolgreiches Studieren an den Campusuniversitäten der Bundeswehr. Die hoch ladenden Haltungen Offenheit für Erfahrung und Herausforderung erklären elf Prozent der Gesamtvarianz auf und erweisen sich außerdem als für ein erfolgreiches Studieren förderliche Dispositionen. Typische „offiziernahe“ Eigenschaften wie Führungspotential oder Durchsetzungsfähigkeit klären überraschenderweise den geringsten Gesamtvarianzteil (10 Prozent) auf. Ein Studieren an den UniBw ist im starken Kontrast zum Offizierdasein in der vorher absolvierten Offizierausbildung oder in späteren Führungsverwendungen als Truppenoffizier zu sehen. Unter Umständen wirkt sich eine zu nahe Identifikation oder Fähigkeit mit dem Führungs- bzw. Offizierberuf eher als tendenziell hinderlich auf ein zielorientiertes Studieren im zivilen Universitätsalltag aus und lässt in manchen Momenten vermutlich sogar am Sinn eines Pflichtstudiums für angehende Offiziere zweifeln. Dieser Aspekt wurde auch in den bis zu 30 Minuten andauernden Feedbackgesprächen deutlich. Unterstützt werden die Resultate durch die Ergebnisse aus dem nomologischen Netzwerk. Auch diese Auswertungsmethode weist dem Verfahren mindestens befriede-

digende bis sehr gute Konstruktvalidität zu. Die Persönlichkeitsfragebögen zeigen das Vorhandensein konvergenter und diskriminanter Validität auf.

Die Ergebnisse beider Analyseformen replizieren richtungweisende und viel zitierte Untersuchungen. Wie im Kapitel 3 ausführlich diskutiert, bleibt das Gros dessen, was Studieneignung ausmacht, nicht vollends bestimmbar. Besonders schwierig wird es, wenn das Studium an sich oder unzählige innere und äußere Bedingungen, wie sie im vorliegenden Falle der Bundeswehruniversitäten vorliegen, eine Erfassung aller möglichen Einflussfaktoren unmöglich ist. Nichtsdestotrotz lässt sich im Sinne einer angestrebten und in der Praxis umsetzbaren Statusdiagnostik das Ziel einer langfristigen Orientierungs- und Entscheidungshilfe der Studienberatung im Eignungsfeststellungsverfahren der OPZ mittels einer studienfachspezifischen Reihungstestbatterie als realistisch bezeichnen. Der grundlagenforschungsorientierte Wunsch einer gänzlichen Aufklärung des Konstrukts Studieneignung widerfährt demnach mehr einem weiteren Beitrag als einer vollständigen Lösung.

Fazit der Diskussion der Konstruktvaliditätsergebnisse sollte nicht nur die Replikation vorhandener Forschungsergebnisse zum Konstrukt Studieneignung durch die neue Testbatterie bzw. die Bestätigung ihrer Konstruktvalidität sein. Gleichzeitig liefern die Resultate auch Hinweise auf die Übertragbarkeit der Ergebnisse auf die soldatische Studieneignung. Gerade wegen der gesonderten Stellung des straff organisierten Trimesterstudiums der UniBw inmitten zahlreicher und von anderen privaten wie auch öffentlichen Hochschulen zu unterscheidenden Einflussvariablen ist jedoch eine besonders starke Gewichtung von Soft Skills zu vermuten.

### 12.2.3 Kriteriumsvalidität

Wie leistungsfähig erweist sich der studienfeldspezifische Eignungstest für die Prognose späterer Studien- und Offiziereignung?

Welche Resultate liefert die empirische Untersuchung an den beiden Universitäten der Bundeswehr in Hamburg und München hinsichtlich geeigneter Prädiktoren zur Vorhersage der Studieneignung?

Den Kern von Bewährungskontrollen bildet die Ermittlung der Prognosegüte des Messverfahrens. Erst das Zusammenführen der Erkenntnisse aus der alten und neuen Studieneignungsempfehlung mit den im Studium erbrachten Leistungen bildet den Kern von Erfolgs- und Bewährungskontrollen. Wie im theoretischen Teil umfassend erörtert bestehen in Rei-

hungstests statistisch positive Zusammenhänge zwischen den zur Studieneignungsbeurteilung vorliegenden Bewertungskriterien sowie den hinzugezogenen psychodiagnostischen Erkenntnismitteln einerseits und den Leistungen im Studium oder Examen andererseits. Die Zusammenhänge zwischen Reihungstestergebnis und Hauptdiplomnote sind gemäß der Interpretationshilfen der DIN 33430 und des Personalpsychologischen Qualitätssicherungskonzepts zwischen  $r = .30$  und  $.55$  als gut bis sehr gut und unter  $.30$  als befriedigend bis schwach zu bezeichnen. Die Reihungstestbatterie Studieneignung sagt den späteren Studienerfolg mit einer Höhe von  $r = .34$  (Vordiplom) und  $r = .42$  (Hauptdiplom) signifikant und gut vorher. Die Hypothese einer besseren Prognoseleistung der Testbatterie Studieneignung gegenüber der OPZ-Studienberatung, die mit  $r = .20$  den Studienerfolg ebenfalls signifikant aber noch schwächer vorhersagt, lässt sich auf einem Signifikanzniveau von  $p < .01$  bestätigen.

Die Hypothese der Überlegenheit des spezifischen Teils gegenüber dem allgemeinen Teil findet keine Bestätigung. Bis auf den ingenieurwissenschaftlichen Testteil mit einer prädiktiven Validität in Höhe von  $r = .20$  finden sich keine signifikanten Gültigkeitswerte. Trotz ähnlicher Testbedingungen – z.B. Testung bereits studierender Offizieranwärter – und ungeachtet der Tatsache einer freiwilligen Testung lassen sich frühere Ergebnisse durch der BA oder itb Consulting nicht wiederholen.

Nicht zu unterschätzen sei der Hinweis Bohlebers (1982), dass Schlüsselqualifikationen im Sinne persönlicher und sozialer Kompetenzen dahingehend eine Rolle spielen, dass diese häufig erst im Studienverlauf ihre Entfaltung finden. Viele in der Drop-Out-Forschung als bedeutsam herausgestellte Einstellungs- und Verhaltensmuster entwickeln sich jedoch erst während des Studiums und als Ergebnis einer Interaktion von persönlichen Bedingungen und Studiensituation. Insofern ist eine Prognose im Sinne der Vorwegnahme eines Entwicklungsergebnisses kaum bzw. überhaupt nicht denkbar. Studienerfolgsprognose wird sich daher auf zweifelsfrei feststellbare Sachverhalte wie die oben aufgeführten Teilkonstrukte des Konstrukts Studieneignung stützen müssen.

Erwartungskonform mit der dritten Hypothese verhalten sich im Vergleich zu den übrigen Untertests die Persönlichkeitsfragebögen Big Five und PMO. Beide liefern in entscheidenden Dimensionen höchst aussagekräftige Koeffizienten und bestätigen das Vorhandensein prognostischer Validität. So definiert Trost bereits 1975 Studierfähigkeit als breites Spektrum individueller Voraussetzungen. Neben intellektuellen Fähigkeiten bezeichnet auch er Lern- und Arbeitstechniken, Arbeitsdisziplin, Fleiß und Durchhaltevermögen, Selbstkritik

und emotionale Stabilität als Teil des Konstrukts Studieneignung. Konegen-Grenier (2001) betont die Wichtigkeit persönlicher Arbeitshaltungen wie Selbstkontrolle, Gewissenhaftigkeit und Belastbarkeit. Brandstätter et al. (2001) weisen den Zusammenhang zwischen Höhe der Ausprägung der Persönlichkeitsmerkmale Selbstkontrolle, Gewissenhaftigkeit, Belastbarkeit und Introversion und Höhe des Studienerfolgs hinsichtlich Examensnote und Abbruchquoten nach. Heldmann (1984) zählt ferner Ausprägungen der Persönlichkeit und Engagement als erfolgsrelevant für das Studium dar. Auch Leistungsmotivation weist dabei eine mittlere Validität auf (Schuler, 2003). Gewissenhaftigkeit, Zuverlässigkeit und Leistungsmotivation werden ebenso von Hitpass (1975) als sehr bedeutungsvoll für Studienerfolg angenommen. Melter ermittelt 1977 Angstneigung, Belastbarkeit, Selbstsicherheit und Empfindsamkeit als relevante Einflussgrößen. Rindermann & Oubaid (1999) betonen außerdem die Wichtigkeit von Soft Skills für ein erfolgreiches Studieren.

In der vorliegenden Arbeit werden anhand der Prognosewerte der Testbatterie Studieneignung als maßgeblich aussagekräftig für den Studienerfolg die Persönlichkeitsfaktoren Selbstdisziplin und Gewissenhaftigkeit, Handlungsorientierung, Führungspotential und Teamfähigkeit sowie Leistungsmotivation subsumiert. Zieht man anstelle der Haupt- die Vordiplomnote als Kriterium hinzu, kristallisieren sich lediglich die Faktoren Selbstdisziplin, Gewissenhaftigkeit und Risikokompetenz heraus. Offenbar nehmen die Merkmale Handlungsorientierung, Führungs- und Teamfähigkeit sowie Leistungsmotivation mit der Dauer des Studiums und im Hauptdiplom an Bedeutung zu. Auch an dieser Stelle sei auf den Hinweis Bohlebers (s.o.) hingewiesen.

Das aufgestellte multiple Modell unterstützt die Bedeutsamkeit weicher Faktoren im Hinblick auf die Prognose der Hauptdiplomnote und stützt die Wichtigkeit der Abiturnote zur Vorhersage der Vordiplomnote. Letzteres lässt sich durch die Nähe der Abiturnote zu weniger anwendungs- als vielmehr grundlagenbezogenen Vordiplomsinhalten erklären. Dagegen subsumiert ein erfolgreiches Hauptdiplom auch Aspekte, die erst während des Studiums erworben oder bedeutender werden. So sprechen Rindermann und Oubaid (1999) Studieneingangstests die Funktion einer aktuellen Objektivierung von Schulleistungen oder Ergänzung von Fähigkeiten zu, die durch die Abiturnote nicht oder nicht ausreichend abgedeckt werden. Insgesamt betrachtet werden die im theoretischen Teil aufgeführten Forschungsergebnisse zur Prädiktorkraft der Abiturnote repliziert. Zusätzlich sei an dieser Stelle auf das Exkurskapitel und unabhängig von der Datengrundlage der vorliegenden Arbeit zu sehenden Untersuchung Prädiktor Abiturnote an den UniBw in Kapitel 11.3 hingewiesen. Die Studien bestäti-

gen den sowohl in der Fachliteratur beschriebenen als auch in dieser Arbeit bestätigten Einfluss der Abiturnote auf den Studienerfolg. In Bezug auf die UniBw fällt dieser unter Berücksichtigung der vorselektierten Stichprobe jedoch etwas geringfügiger aus.

Wie im Theoretischen Teil zur inkrementellen Validität aufgeführt, können in Kombination mit dem Schulnotenmittelwert allgemeine und spezifische Studieneignungstests die inkrementelle Validität auf bis zu  $R = .49$  erhöhen (Kuncel, Hezlett & Ones, 2001; Mouw & Khanna, 1993; Robins, Lauver, Le, Davis, Langley & Carlstrom, 2004). Im Falle der Testbatterie Studieneignung und unter Hinzunahme von Persönlichkeitsfaktoren mittels PMO und Big Five erhöht sich die inkrementelle Testgültigkeit des wirtschaftswissenschaftlichen Teils sogar auf  $R = .68$  (Kriterium Hauptdiplomnote), die des geisteswissenschaftlichen Teils auf  $R = .60$ . Die Hypothese, nach der eine Erhöhung der Validität durch Hinzunahme von Persönlichkeitstests aus dem allgemeinen Teil II zu erwarten ist, findet ihre Bestätigung für die Studienfelder Geistes- und Wirtschaftswissenschaften. Die Ergebnisse für den Ingenieurwissenschaftstest sind nicht signifikant, die Statistik für den Informatiktestteil lässt sich wegen fehlender Korrelationen nicht berechnet werden. Bei Hinzunahme der Abiturnote steigt die Validität jedoch nicht signifikant verschieden vom oben zitierten Ergebnis an.

Zusammengefasst lässt sich die Testbatterie Studieneignung zwar als prognostisch gut valide bezeichnen, ist jedoch insbesondere in Bezug auf den Mehrwert durch die nur schwach bis kaum prädiktiv validen fachspezifischen Untertests als lediglich befriedigend vorhersagegütig anzusehen. Wie der bereits ermittelte hohe Einfluss von Persönlichkeitsfaktoren auf das Overallergebnis vermuten lässt, begründen zu einem großen Teil die ermittelten weichen Faktoren die ermittelte Vorhersagegüte.

#### 12.2.4 Moderierende Hintergrundvariablen

Welche Schlüsse lassen Prozess- und Postevaluation hinsichtlich möglicher interner und externer Einflüsse zu?

Bestehen Unterschiede in der Studienerfolgs-Vorhersageleistung für verschiedene Teilgruppen oder aufgrund anderer interner und externer Einflüsse?

Militäraffine wie andere Variablen moderieren möglicherweise die Studienerfolgsvorhersageleistung der neuen Testbatterie Studieneignung. Erwartungsgemäß unterscheidet sich der Studienerfolg der drei TSK kaum bis gar nicht. Anders verhält sich der Einfluss auf das Kriterium OPZ-Gesamteignungsgrad, wo relativ auf die TSK bezogen die Mehrheit der als be-

sonders geeignet Beurteilungen Marinesoldaten zu Eigen ist. Bei den besten Offizieranwärtern handelt es sich relativ gesehen also meistens um zukünftige Marineoffiziere, bei denen als gut geeignet bewerteten um Marine- oder Luftwaffenangehörige. Der größte Anteil der Heeresoffiziere wird als lediglich „geeignet“ eingestellt.

Entgegen der Hypothese unterscheiden sich Studenten mit unterschiedlich langer Trimesterzahl bis zum Vordiplom hinsichtlich des Studienerfolgs. Die Studienzeit bis zum Vordiplom wirkt sich im zunehmenden Verlauf negativ auf das Vordiplomergebnis aus. Je kürzer – desto besser, je länger – desto schlechter! Damit scheint das „Drei-Jahre-Plus“-Studienmodell der UniBw bestätigt, wenn man die Note als alleiniges Kriterium für Studienerfolg zugrunde legt. Es ist demnach fraglich, ob ein genehmigter Antrag auf Studienzeitverlängerung auch zu einer effektiven Nutzung der Lernzeit führt. Dass Studenten unter den Bedingungen von Regelstudienzeit und Höchststudiendauer ein wissenschaftliches Studium angemessen absolvieren und auch mit guten Noten abschließen können, ist somit belegt. Wer länger studiert, geht – statistisch gesehen – das Risiko eines schlechteren Abschlusses ein. Unbeantwortet bleibt jedoch die Frage, ob es einem länger studierenden Offizieranwärter eher die Chance ermöglicht, fachübergreifende, persönlichkeitsbildende Angebote wahrzunehmen.

Der Offizierberuf mit seinem Studienangebot ist für Studienberechtigte aus einem eher bildungsfernen Familienmilieu attraktiv und ließ aufgrund der dadurch bedingten erschwerten Lernvoraussetzungen ein schwächerer Studienerfolg erwarten. Hinsichtlich des soziokulturellen Bildungshintergrunds ist entgegen der Hypothese kein signifikanter Einfluss auf die Studienleistungen ermittelt worden. Eine mögliche Erklärung könnte der Campusalltag und das Studieren bei vollem Gehalt sein, in dem keiner der vor Ort Studierenden durch schwierige Lern- oder Lebensbedingungen eine Benachteiligung widerfährt. Der Berufsweg mit integriertem Studium bietet einer bildungsfern-aufstiegsorientierten Klientel günstige Bedingungen und Chancen, in eine gegenüber dem Elternhaus vergleichbar andere Bildungs- und Lebenswelt zu wechseln.

Hypothesenkonform verhalten sich weder die Variable Geschlecht noch der Studienort als einflussreich auf den Studienerfolg. Jedoch schneiden weibliche Bewerber im ACBw der OPZ häufiger besser bzw. gut geeignet (52 Prozent) ab als ihre männlichen Mitstreiter (37 Prozent), die zu 60 Prozent mit geeignet beurteilt werden. Die Dominanz der weiblichen Offizierbewerber im AC Bw könnte auch darauf hinweisen, dass Frauen soziale Anforderungen in der Auswahluntersuchung besser bewältigen als Männer. Aus den Geschlechtseffekten lässt sich jedoch nicht zwingend eine generelle Überlegenheit von Bewerberinnen ableiten.

Gerade im Rahmen der Auswahl von weiblichen Offizieranwärtern, welche erst ab Juli 2001 uneingeschränkt für alle Laufbahngruppen möglich ist, ist eine Selbstselektion als wahrscheinlich anzunehmen. Hinzu kommt möglicherweise die besonders hohe Motivation, sich gegenüber dem anderen Geschlecht in einer „Männerbastion“ behaupten zu wollen.

In dem Zusammenhang soll noch einmal auf das Exkurskapitel „Untersuchung Prädiktor Abiturnote an den UniBw“ hingewiesen werden. Untermuert wird dort der Einfluss des Studienfachs auf den Studienerfolg bzw. auf die Drop-Out-Quote.

### 12.2.5 Evaluation

Wie steht es mit der Akzeptanz der praktizierten und untersuchten Auswahlverfahren bei Offizieranwärtern?

Wie leistungsfähig erweisen sich Qualität, Objektivität, Effizienz und Umsetzbarkeit bestehender und untersuchter Prädiktoren?

Neben der Erfüllung aufgestellter Grob- und Feinziele strebt die vorliegende Arbeit außerdem eine nachhaltige Wirkung für die Praxis der Offizierauswahl an. Mit Hilfe eines Akzeptanzfragebogens und 20- bis 30-minütiger Feedbackgespräche mit nahezu allen Testpersonen kann eine Evaluation des neuen Studieneignungstests und der Pilotphasen die Ergebnisse weiter verdichten. Besprechungen mit Lehrpersonal und militärischen sowie zivilen Vorgesetzten bis hin zur Universitätsleitung erhärten die Schlüsse. Die als außerordentlich hoch ermittelte Akzeptanz und Augenscheinvalidität der vorgestellten Reihungstests und Hinweise auf die besondere Stellung der privaten Bundeswehruniversitäten bezüglich weicher zu messender Faktoren sowie nicht zuletzt aus Kosten-Nutzen-Gesichtspunkten und eigenem vorhandenem Know-How lassen trotz oder gerade wegen kritisch zu sehender Einzeltestaufgaben eine Ausweitung des Projektes hin zu einem kontinuierlich Aufgaben produzierenden Langzeitprojekt befürworten.

Die Rolle und Wichtigkeit von Persönlichkeitseigenschaften und Arbeitshaltungen neben kognitiven Faktoren wird auch von Seiten der Testpersonen sowohl im Fragebogen als auch im späteren Feedbackgespräch in den Vordergrund gerückt. Die in vielen Fällen bitteren Erfahrungen der Studienfachzuweisung entgegen des eigentlichen Studienfachwunsches zogen und ziehen persönlich bedeutsame Auswirkungen auf die Zeit während und nach der Bundeswehr mit sich. Umso herber wirkte letztlich ein in der fachspezifisch anspruchsvollen Testbatterie Studieneignung gemachtes sehr gutes Ergebnis zur Prognose des Studienfachs,

dass im OPZ-Bewerbungsverfahren als Erststudienwunsch benannt, deren Eignung jedoch aus Sicht des OPZ-Studienberaters nicht gesehen wurde.

Vor einer vorschnellen Verurteilung der Studienberatung OPZ muss an dieser Stelle allerdings gewarnt werden. Dem beratenden - meist militärischen - Personal fehlt die fundierte Grundlage, überhaupt eine im Rahmen einer sequentiellen Vorgehensweise Erfolg versprechende Beratung und Studienempfehlung innerhalb der vorgesehen zehn Minuten zu tätigen. Damit gemeint ist die notwendige, aus den Reihungstests erschließbare, Daten- und Interviewbasis.

### 13. Implikationen für Praxis und Forschung

Wo liegen Ansatzpunkte für eine Optimierung des entwickelten studienfeldspezifischen Eignungstest?

Die vorangegangenen Ausführungen zeigen, dass neben wesentlichen Vorteilen der Einsatz von studienfeldspezifischen Reihungstests mit erheblichen Problemen verbunden sein kann. Die aus der inhaltlichen und statistischen Überprüfung resultierenden Optimierungsvorschläge sollen die Grundlage für die in den nächsten Kapiteln aufgeführten Implikationen für Forschung und Praxis sein.

#### 13.1.1 Methodische Implikationen

Unter methodischen Gesichtspunkten werden folgende Implikationen für die Praxis vorgeschlagen: In Kapitel 12.1.2 wurde auf die als schwierig zu bezeichnende Interpretation der als hypothesengenerierend zu verstehenden Faktorenanalyse kritisiert. Die Interpretation zur faktorenanalytisch ermittelten Konstruktvalidität in dieser Arbeit soll daher als Überlegung verstanden werden, die es zu überprüfen gilt. In weiteren Untersuchungen sollte geklärt werden, ob die Übungen wirklich die Bewertungen bestimmen. Eine Alternative zur explorativen Faktorenanalyse stellt die bereits beschriebene konfirmatorische Faktorenanalyse mit ihrem hypothesenprüfenden Charakter dar.

Voraussetzung für eine aussagekräftige Evaluation ist, befriedigende Maße zu finden, die als Kriterien der Validierung von Auswahlverfahren dienen (Problematik der Indikator-Kriteriumsbeziehungen). Es wurde in dieser Arbeit ferner aufgezeigt, dass die Vor- bzw. Hauptdiplomnote um weitere Kriterien ergänzt werden sollte, da die Auswahl nur eines Kriteriums möglicherweise die Ursache des nicht optimalen Abschneidens des Studieneignungs-

tests bezüglich der Kriteriumsvalidität sein könnte. Stattdessen sollten in zukünftigen Bewährungskontrollen alle organisatorischen und rechtlichen Bedingungen durch das BMVG bereitgestellt werden, um problemlos eine Validierungsuntersuchung durch *multiple* Kriterien zu ermöglichen. Wünschenswert für eine qualitativ hochwertige Evaluation wären außerdem gut vernetzte Expertensysteme in computerisierter Form (zum Beispiel SASPF) für eine schnellere und effizientere Abwicklung.

Viele der geschilderten methodischen Probleme (Stichprobenselektion mit der Folge reduzierter Streuungen, fragliche Skalendignität des Kriteriums oder der diagnostischen Instrumente usw.) lassen sich mit Hilfe neuerer Methoden der Datenanalyse lösen. Sie erlauben den Verzicht auf meist nicht gegebene Voraussetzungen hinsichtlich der metrischen Eigenschaften des Datenmaterials (Skalenniveau) und eröffnen neue Möglichkeiten zur Integration der diagnostischen Einzelinformationen. Dabei sind vor allem die auf Häufigkeitsanalysen aufbauenden Verfahren wie die Konfigurationsfrequenzanalyse (Krauth & Lienert, 1973) oder die in jüngster Zeit mehrfach erfolgreich im diagnostischen Feld eingesetzten Methoden der Hypothesen-Agglutination (HYPAG) (Piepersjohanns, 1983. Wottawa, 1983 und Orgass, 1984) von Bedeutung, die unter anderem für eine gezielte Verbesserung von Komponenten der eignungsdiagnostischen Prozedur genutzt werden können (vgl. Wottawa & Hossiep, 1987).

Vor Entwicklung studienfeldspezifischer Auswahlverfahren müssen Anforderungen spezifischer Studienfelder und -fächer, Studienerfolgskriterien und prognostisch relevante kognitive und nicht kognitive Eingangsmerkmale bestimmt werden (Amelang, 1997). Zielsetzung der von Professor Kastner durchgeführten und dieser Arbeit zugrunde liegenden theoretischen Anforderungsanalyse war die verbesserte Informationsgrundlage für die OPZ-Studienberatung. Sie erhebt ausdrücklich nicht den Anspruch einer Realitätsabbildung und kann vielmehr als Inhaltsanalyse schriftlich festgehaltener Curricula angesehen werden. Im Rahmen der vorliegenden Arbeit wurde sich außerdem auf Aussagen von Professoren sowie Ergebnissen aus einschlägigen Forschungsstudien gestützt. Für eine langfristig sinnvolle Auswahl zukünftig an den UniBw Studierender ist eine standardisierte Befragung der Hochschullehrer und Studenten verschiedener Fachrichtungen mit voran geschalteten qualitativen Interviews zu empfehlen.

Um Einflussfaktoren auf das Antwortverhalten wie sozial erwünschte Beantwortung von Persönlichkeitsfragebögen zu begegnen, lässt sich zum einen auf die Verwendung geeigneter objektiver Persönlichkeitstests oder alternativ auf das engere Eingebundensein des multimo-

dalen Interviews und weiterer AC Bw-Bausteine verweisen. Gruppeninteraktionen und Bewältigung von situativen Aufgaben wie Rollenspiel, Kurzvortrag usw. ließen Ableitungen hinsichtlich Studienmotivation, Organisationsvermögen, Handlungsorientierung, Führungspotential und weitere die allgemeine Studierfähigkeit betreffende Arbeitshaltungen und Persönlichkeitseigenschaften zu.

Bei genügend großer fachspezifischer Stichprobengröße ließen sich zukünftig mittels Diskriminanzanalyse Gruppen überprüfen und differente Faktormuster zwischen den Analysegruppen ausmachen.

Hit- und Base-Rate-Problematiken wurden aufgrund der uneindeutigen Festlegung der Basisrate (Anteil geeigneter Bewerber) vernachlässigt, könnten aber insbesondere im Hinblick auf die angestrebte sequentielle Vorgehensweise wertvolle Ergebnisse liefern. Zu untersuchen wäre, ob auch Verfahren mit geringem inkrementellen Validitätsgewinn statt einer Verschlechterung im Vergleich zur Zufallsauswahl zu einer Erhöhung der Validität beitragen (Schönemann & Thompson, 1996; Schönemann, 1997).

### 13.1.2 Fazit für die Praxis

Betrachtet man die Ausbildung von Studierenden als mehrgliedrige Wertkette, die von den Bewerbern über die Studienanfänger und Absolventen bis zum reüssierenden Berufsanfänger reicht, so wird die Bedeutsamkeit von aussagefähigen Zulassungsverfahren besonders augenscheinlich, da sie den Beginn der Kette bilden und somit ihre Aussagekraft alle nachfolgenden Prozesse entscheidend beeinflusst. (Hell, Workshop der Österreichischen Forschungsgemeinschaft, 10.-11.03.2006, S. 5).

Die außerordentlich große Akzeptanz und Face-Validität der vorgestellten Reihungstests und Hinweise auf die besondere Stellung der privaten Bundeswehruniversitäten bezüglich weicher zu messender Faktoren sowie nicht zuletzt aus Kosten-Nutzen-Gesichtspunkten und eigenem vorhandenem Know-How lassen trotz oder gerade wegen kritisch zu sehender Einzeltestaufgaben eine Ausweitung des Projektes hin zu einem kontinuierlich Aufgaben produzierenden Langzeitprojekt befürworten!

Die Empfehlung eines eingangs beschriebenen multimodalen Verfahrens zur Erzielung einer erhöhten Diagnosesicherheit muss mit den Anforderungen in der OPZ verbunden werden! Aus Kosten-Nutzen-Gesichtspunkten ist ein *sequentielles Modell* mit mehreren Stufen bzw. vorgegebenen Quoten im Sinne des zukünftig in der OPZ eingesetzten Expertensystems

OBESY zu empfehlen. So kann beispielsweise im ersten Schritt die Zulassung eines Teils der Studienbewerber aufgrund ihrer herausragenden Abiturnote und die weitere Zulassung eines Teils nach Kombination von Abiturdurchschnittsnote und studienfeldbezogenen Fähigkeitstest stehen. In einem letzten Schritt würden von den verbleibenden Personen unklarer Prognose (in der Studienberatung der OPZ gerne als „Grauzone“ bezeichnet) mittels eines Auswahl- bzw. Studienberatungsgesprächs ausgewählt. Denkbar wäre langfristig auch ein parallel laufender weiterer Ausbau der Zusammenarbeit mit Fachhochschulen, öffentlichen Universitäten oder privater Hochschulen. Frühere Untersuchungen ergaben große Unterschiede zwischen Studenten an öffentlichen versus militärischen Hochschulen.

Aus dem Gesagten lassen sich Unzufriedenheit im Studium, Misserfolge, Leistungsdefizite und Studienabbruch zum einen auf verspätete Selektion mit Folge einer mangelnden motivationalen und leistungsmäßigen Passung als Voraussetzung für erfolgreiches Absolvieren eines Hochschulstudiums zurückführen. Zum anderen kann eine viel versprechende Selektion auch im Laufe des Studiums durch äußere Bedingungen fehlgeleitet werden und bei absenter äußerer Hilfestellung des Studierenden zu Misserfolg im Studium bis hin zum Studienabbruch führen. Umfangreiche Untersuchungen des HIS (2002) bestätigen die Annahme, dass eine genaue Abgrenzung entlang der Abbruchgründe und Risikofaktoren trotz vorheriger Identifizierung abbruchgefährdeter Studenten beinahe unmöglich erscheint.

Ziel von Studierendenauswahl muss demzufolge die Herstellung einer größtmöglichen Passgenauigkeit zwischen dem Anforderungs- und inhaltlichen Profil des Studiengangs und dem Fähigkeits- und Interessenprofil der ausgewählten Studierenden sein. Da Schüler nur schwierig eigene Fähig- und Fertigkeiten mit den Studienanforderungen abgleichen können, rückt die Bedeutsamkeit förderlicher Studienberatung neben dem eigentlichen Auswahlprozess in den Vordergrund. Eine eingehende Auseinandersetzung mit den kognitiven Fähigkeiten, Interessen und Bedürfnissen einerseits und Anforderungen und Rahmenbedingungen des Studienfachs und der Hochschule andererseits kann bereits präventiv Orientierung geben und Rüstzeug für ein aussichtsreiches Studieren sein.

Die Ergebnisse lassen eine Entscheidung für oder gegen ein Studium auf Basis psychometrischer Verfahren in Kombination mit Beratungssystemen wahrscheinlich werden. „Insgesamt ist es notwendig, die Beratung und Auswahl von Studienbewerbern von vorneherein als umfassenden Prozess zu betrachten und professionell zu gestalten“ (Kersting, 2005, S. 150). Der gesamte Auswahlprozess von geeigneten Studierenden lässt sich als ganzheitliches System mit einem breiten Angebotsspektrum wie Information, Eignungsdiagnostik, Coaching, Tu-

tor-, Mentorenprogramme, Beratung studentischer Arbeits- oder Wohngruppen beschreiben. Diesem muss die Klärung der Anforderungen mittels Anforderungsanalysen vorangestellt und sein Prozesscharakter impliziert werden. Studienberatung ist von Studierendenauswahl nicht zwingend zu trennen; im Falle der Offizierauswahl in der Bundeswehr sogar gleichzusetzen (vgl. Kapitel 6.2.2). Weil sich ein umfassendes Beratungskonzept für die Bundeswehr von denen öffentlicher oder übriger privater Hochschulen stark unterscheidet, verlangt es eine differenzierte Auseinandersetzung mit den streitkräftetypischen Gegebenheiten. Die verstärkten oder neuen Anforderungen, die aus der Transformation der Bundeswehr für das militärische Personal erwachsen, erfordern das Überdenken des bisherigen Verständnisses von Personalauswahl und -entwicklung. Die Bundeswehr konkurriert am Arbeitsmarkt bei der Gewinnung von jährlich weit über 20 000 Soldaten auf Zeit mit der Wirtschaft und weiteren Bereichen des öffentlichen Dienstes. Im zunehmenden Wettbewerb um qualifiziertes Personal muss sie sich als potenzieller Arbeitgeber erfolgreich positionieren. Es gilt, die Möglichkeiten und Erfordernisse der Streitkräfte mit den Vorstellungen und Wünschen der Zielgruppen weitestgehend in Einklang zu bringen. Zur erfolgreichen Personalgewinnung bedarf es daher nicht nur einer zielgruppengerechten Kommunikation unter Herausstellung der Besonderheiten des soldatischen Dienstes, sondern auch eines attraktiven Angebots wie die des bezahlten Hochschulstudiums für angehende Offiziere. Angesichts der demographischen Entwicklung bzw. der in den kommenden Jahren sinkenden Geburtsjahrgangsstärken reduziert sich die Anzahl derjenigen, die für die ewige Personalgewinnung in Betracht kommen. Verschärfend kommt hinzu, dass der Anteil an Grundwehrdienstleistenden im Zusammenhang mit den Umfangsreduzierungen abgesenkt wird, wodurch das Potenzial für die Binnengewinnung deutlich zurückgeht. Gleichzeitig ist abzusehen, dass sich die körperliche Leistungsfähigkeit der Jugend dramatisch verschlechtert. Der Fürsorge der Vorgesetzten ist es zuzusprechen, nachfolgenden Soldaten die Chance einzuräumen, den erforderlichen Erfahrungsschatz zu erwerben.

### 13.1.3 Ausblick

Infolge der Ergebnisse der PISA-Studie (Programme for International Student Assessment) der OECD (Organization for Economic Co-operation and Development) überwiegt in der öffentlichen Diskussion die negative Kritik über das deutsche Bildungssystem. Gleichzeitig ist im Zuge der Diskussion um die Stärkung der internationalen Wettbewerbsfähigkeit deutscher Hochschulen eine politische Debatte um den Hochschulzugang neu entbrannt. Aus

einer Identitätskrise erwuchs der Neuanfang. Als Folge dieser Effizienz-Debatte bietet die Novellierung der rechtlichen Rahmenbedingungen den Hochschulen die Möglichkeit, ab dem Wintersemester 2005/2006 sechzig Prozent der zur Verfügung stehenden Studienplätze in den Numerus Clausus-Fächern autonom zu vergeben. Der deutsche Wissenschaftsrat systematisierte 2004 noch einmal in seinen Empfehlungen zur Reform des Hochschulzugangs die Argumente für eine Übertragung des Zulassungsrechtes auf die Hochschulen. Der Akt der Zulassung selbst soll demzufolge stärker als „ein Mittel zur Förderung von Profilbildung und Wettbewerb“ (Wissenschaftsrat, 2004) verstanden wird. Über die Verfahren in der Bandbreite von schriftlichen Tests bis hin zu Auswahlgesprächen mit Professoren sollen die Hochschulen ausdrücklich selbst entscheiden. Mit der stärkeren Eigenverantwortung werden den Universitäten auch vielfältige Verfahrensfestlegungen und Organisationsentscheidungen abverlangt. Jene machen allerdings bisher von den neuen Möglichkeiten kaum Gebrauch. Offenbar wirkt eine so geringe Quote weder profilbildend noch den Verwaltungsaufwand bei gleich bleibenden finanziellen Ressourcen rechtfertigend. Die Hochschulen beginnen erst sehr zögerlich und sukzessive, präzisere und spezifischere Anforderungen bei der Studierendenauswahl zu formulieren als die Durchschnittsnote des Abiturs. Gelten „cum tempore“ und akademische Freiheit also auch noch in Zeiten der Novellierung? Die Umsetzung des neuen Hochschulrahmengesetzes ist nicht umstandslos möglich, weil die Entwicklung von validen Zulassungstests mit fachbezogenen Komponenten mehrere Jahre in Anspruch nimmt. Zudem müssen solche Verfahren „gerichtsfest“ sein. Es gehört insofern noch immer zu den zentralen Aufgaben im deutschen Hochschulsystem, dass die deutschen Universitäten eine passgenaue Auswahl der Studierenden, die ihnen noch vor wenigen Jahren rechtlich verwehrt gewesen ist, erst nach und nach einführen. Ein besserer Erkenntnisstand über eine Vielzahl möglicher Effekte verschiedener Auswahlstrategien ist sodann erst in einigen Jahren zu erwarten.

Mit der Einigung über die Änderung des Hochschulrahmengesetzes streben die Forschungsanstalten eine Abgrenzung untereinander und Steigerung der eigenen Attraktivität an. Der Wettbewerb erhält durch zwei externe Entwicklungen zusätzliche Dynamik. Zum einen durch die wachsende Zahl privater Hochschulen; zum anderen durch die wachsende Internationalisierung bzw. Europäisierung des Hochschulraums, der die Mobilitätsoptionen für Mittel, Personal und Studierende stetig ausweitet.

Unterschiedlichste Reformwerke werden derzeit an Universitäten umgesetzt. So führt man kürzere, verschulte Bachelor- und Master-Studiengänge ein, erhebt Gebühren für Langzeit-

studenten und ein Zweitstudium, Globalhaushalte organisieren die Mittel neu, und Professoren erhalten Verträge mit Leistungszulagen. Evaluation von Forschung und Lehre sichert darüber hinaus die kontinuierliche Qualitätssicherung. Obendrein soll eine Exzellenzinitiative Eliteuniversitäten schaffen und Wettbewerb unter den Hochschulen und angehenden Spitzenhochschulen forcieren. Die deutsche Alma Mater wird immer weniger begriffen als Träger eines gesellschaftlichen Bildungsauftrages, sondern vielmehr als Quasi-Unternehmen (Bultmann, 1996), die in Konkurrenz miteinander auf einem Markt für ökonomisch nachgefragte Qualifikationen und Forschungsdienstleistungen eigene Profile entwickeln. Der Chef des Centrums für Hochschulentwicklung (CHE), Detlef Müller-Böling (2001) formuliert dies Folgendermaßen: „Grundlage für Wettbewerb ist die Handlungsfreiheit der Anbieter und Nachfrager von Leistungen. Planung und Vertrieb von Studienangeboten gehören insofern in die autonome Entscheidungskompetenz jeder einzelnen Hochschule.“ Kein gewerbliches Unternehmen lässt sich seine Mitarbeiter per Gesetz oder Rechtsanspruch zuweisen. In dem Maße wie Hochschulen als miteinander konkurrierende Betriebe begriffen werden dürfen diese sich folglich ihre Studierenden auch in Übereinstimmung mit ihrem Angebotsprofil aussuchen. Jedoch tritt die in diesem Modell beschriebene Freiheit der Nachfrager hinter die vom CHE verwendete rhetorische Figur der doppelten Freiheit. „Mit dem Recht der Studieninteressenten auf freie Wahl einer Hochschule korrespondiert das Recht der Hochschulen zur Auswahl der Studierenden“ (Müller-Böling 2001). In Wirklichkeit scheinen sich beide Freiheiten gegenseitig eher aufzuheben. Auch wenn sich Studienplatzaspiranten dort bewerben können, wo sie wollen, verlieren sie gleichzeitig den grundsätzlichen Rechtsanspruch auf Studienfach und -ort eigener Wahl.

Vielerorts wächst überdies die Befürchtung, soziale Selektivität würde durch die neuartigen Auswahlverfahren verstärkt. Nach Angaben der Hochschulrektorenkonferenz (HRK) gibt es in Deutschland derzeit in etwa 9 300 grundständige Studiengänge. Das gegenwärtige System einer staatlich-bürokratischen Studienplatzverwaltung und Qualitätssicherung dürfte weitaus transparenter sein als ein angestrebter Zustand, in dem auf einem deregulierten Bildungsmarkt diese Studiengänge nach je individuellen Marketingstrategien um durchschnittlich 350.000 Studienplatzbewerber pro Jahr (2002) als Kunden konkurrieren. Für die Schulabgänger ist dies eine Situation objektiver Überforderung und vollständiger Undurchschaubarkeit. Es gibt keinerlei Kriterien, nach denen sie die ihnen präsentierten Informationen bewerten und womit sie was vergleichen können. Zu erwarten ist, dass zur Koordination von Angebot und Nachfrage eine eigenständige professionelle Beratungsindustrie entsteht, die

Hochschulaufnahmeprüfungen trainiert. In den USA kosten dementsprechende sechswöchige Kurse im Schnitt 800 bis 900 Dollar (Hoffacker 2004). Agenturen, die sich verpflichten, eine erfolgreiche Studienbewerbung einzufädeln, erhalten schon mal ca. 30 000 Dollar pro Klient. Insgesamt geben US-amerikanische Eltern jährlich 250 Millionen Dollar für Vorbereitungskurse aus. In den USA ist es zudem üblich, sich an bis zu zehn Hochschulen gleichzeitig zu bewerben (Hoffacker 2004). Das ist insofern ökonomisch rational gehandelt als mit der Anzahl der Bewerbungen auch die Wahrscheinlichkeit steigt, im gewünschten Studienfach einen Platz zu bekommen. Auch hierzulande werden vermutlich Mehrfachbewerbungen die Regel sein, deren Kosten (Reisen, Unterkunft) zuzüglich der Inanspruchnahme von Beraterfirmen privat aufgebracht werden müssen. Mit der Höhe der jeweiligen Familieneinkommen wachsen auch die Studienchancen der Kinder.

Werden komplexer werdende zeitliche Planungsfaktoren eingerechnet, reduziert sich der zur Verfügung stehende Zeitrahmen noch weiter. Eine "ordentliche" Studierendenauswahl wird realistisch betrachtet immer schwieriger umzusetzen, wenn nicht alle Ausgewählten ihrer Zusage auch nachkommen und frei gebliebene Plätze nachbesetzt werden müssen. Nach aktuellen Erkenntnissen entscheiden sich durchschnittlich sechzig Prozent der erfolgreichen Mehrfachbewerber für eine andere Hochschule. Verzögerungseffekte sind die Folge, da so eine Mehrbelastung im Administrations- und ggf. auch im Auswahlverfahren entsteht.

Der profilbildende Wettbewerb unter den Hochschulen ist als „Prozess dynamischer Wechselwirkungen“ (Trost & Haase, 2005) zwischen Eigendarstellung der Universität und interessiertem Aspiranten zu verstehen, der sich hinsichtlich seines Qualifikationsprofils bestmöglich preisgeben möchte. Dabei spielt zukünftig das Internet als Bewerbungsportal eine herausragende Rolle, in dem sowohl umfangreiche in Studierendenmarketing eingebundene Hintergrundinformationen über Universität, Studium und Studiengänge als auch die Möglichkeit von Self-Assessment (z.B. Rheinisch-Westfälische Technische Hochschule Aachen, Universität Freiburg) mit E-Learning-Elementen offeriert wird (Schüpbach & Klauer, 2006). Unterhaltung suggerierend lassen sich Auskünfte zur Passung des persönlichen Fähigkeitsprofils mit den Studienanforderungen eruieren. „Im Self-Assessment lässt sich die Angemessenheit der Erwartungen bezüglich des Studiums mittels gezielter Überprüfung der Informiertheit sowie der Erhebung der Gründe für die Studienfachwahl klären“ (Schüpbach, Pixner & Zapf, 2005). Eine einträgliche Selbstselektion vermindert den anschließenden Selektionsaufwand durch die Hochschulen dadurch, dass sich im Idealfall nur geeignete Kandidaten tatsächlich bewerben. In Online-Assessment integriertes Marketing stellt ein nicht unwichti-

ges Stellglied der angestrebten konkurrenzorientierten Profilbildung der Fächer und einen entscheidenden Schachzug im Wettbewerb unter den Hochschulen dar. Hornke & Schilling (2004) machen jedoch darauf aufmerksam, dass den meisten Erstsemestern Angaben zu den Forschungsbemühungen der Universitäten unwesentlich erscheinen. Die Perspektive potentieller Bewerber wird vielfach vernachlässigt. Wichtig hingehend sind den Bewerbern transparente und klar verständliche Hintergründe über und die Akzeptanz von Studieninhalten, Zulassungbedingungen und Auswahlverfahren (Hornke & Zimmerhofer, 2006).

Die Bewerberperspektive wird in den nächsten Jahren wissenschaftlich weiter beleuchtet werden müssen. Zukünftig vollzieht sich an universitätsinternen Beratungsstellen ein Themenwechsel hin zur Hilfestellung bei Prüfungsangst, diffuser Gefühlsarmut, Kontaktproblemen bis hin zu ernsthaften Erkrankungen wie Zwängen, Ess- und Persönlichkeitsstörungen. Das Studentenwerk verzeichnete in den Jahren 2004 auf 2005 einen Anstieg des Erstkontakts mit intrauniversitären Beratungsstellen um 9 000 auf 73 000 (Süddeutsche.de, 2006). Laut Studentenwerk verbergen sich hinter vorangestellten Studienberatungsproblematiken oft individuelle Konflikte und Identitätskrisen, die offenbar durch die großen Umwälzungen in der deutschen Hochschullandschaft bedingt sind. So erweisen sich die neuen Bachelor- und Masterstudiengänge als wesentlich verschulter, Studiengebühren belasten zusätzlich und die Arbeitsmarktlage lässt Konkurrenzdruck und Zukunftsangst anwachsen. Aus individueller Perspektive ist ein Studienabbruch dann kritisch, wenn Selbstzweifel und Unsicherheit bei der Neuorientierung überwiegen. Informations- und Kenntnisdefizite, unzureichende Voraussetzungen und sachfremde Hochschulwahlmotive sowie Zweifel an der eigenen Eignung erhöhen das Risiko falscher Erwartungen und Orientierungen, die wiederum einen Studienabbruch begünstigende Leistungs- und Studienmotivationsprobleme nach sich ziehen (Heublein et al., 2003). Dahingehend fanden Schüpbach, Pixner & Zapf (2006) heraus, dass das Assessment von studien Erfolgskritischen Handlungskompetenzen zu messbaren Verringerungen verlaufsbezogener Studienprobleme führen kann. Studienabbruch, Studienfachwechsel, Unzufriedenheit mit dem Studium und Probleme bei dessen Bewältigung sollten sich – wie in der vorliegenden Arbeit herausgearbeitet – langfristig aufgrund optimierter Studienberatung und Passung des einzelnen auf das richtige Studienfach verringern.

Es ist davon auszugehen, dass mit der Einführung von Bachelor- und Masterprogrammen auch an den UniBw eine Vielfalt von Studienangeboten, die das traditionelle Spektrum von Abschlüssen zahlenmäßig um ein vielfaches übertrifft, resultieren wird. Allein die Zentrale

Evaluations- und Akkreditierungsagentur der Universität Hannover zertifiziert hat in den letzten Jahren circa 300 Bachelor- und Master-Abschlüsse. Künftig wird es ein enormes Angebot an stärker spezialisierten und diversifizierten Abschlüssen geben, die auf spezifische Berufsfelder vorbereiten. Diese Entwicklung gilt auch für die Universitäten der Bundeswehr, die schon heute Kontakt zu privaten Universitäten aufnehmen. Letztere wählen bereits jetzt ihre Studenten mit Reihungstests aus. Das Studienangebot wird sich mit der Einführung von Bachelor- und Master-Abschlüssen daran orientieren müssen, um weiterhin attraktiv für Bewerber sein zu können, und um ein eigenes Angebotsprofil in Deutschland zu entwickeln. Damit ergibt sich als neue Aufgabe die Feststellung der Eignung von Studierenden für spezifische Bachelor- und Master-Programme.

Auf den ersten Blick vielleicht indiskutabel erscheinend – auf den zweiten Blick aber durchaus in vielerlei Hinsicht nützlich und umsetzbar – ist die Idee, die Studierendenauswahl in eine freiwillige Beratung von interessierten Abiturienten an verschiedenen Orten Deutschlands (beispielsweise in Kreiswehrrersatzämtern oder Zentren für Nachwuchsgewinnung) zu verwirklichen. So könnten nicht nur aufgrund von Rationalisierungsmaßnahmen ungenutzte aber neuwertige CAT-Anlagen genutzt, sondern herausragende Abiturienten, die den Weg sonst nie zur Bundeswehr gefunden hätten, geworben und gleichzeitig die Offizierbewerberprüfzentrale entlastet werden.

Ein in Folge dessen generellen Imagegewinns der Bundeswehr sowie des Psychologischen Dienstes liegt auf der Hand. Dem Kampf um die besten Studierenden wird auch hinsichtlich des demographischen Wandels die Bundeswehr frühzeitig berücksichtigen müssen (vgl. Kapitel 13.1.2).

Aufgrund der sehr großen Anklang findenden ausschließlich computergestützt laufenden Testbatterie mit Modellcharakterfunktion gelingt es der Bundeswehr, schon heute eine Vorreiterrolle im Bereich der Studienerfolgsprognose zu übernehmen. Abzuleitendes mittel- und langfristiges Ziel stellt die kontinuierliche Weiterentwicklung des Verfahrens dar. Aus Kosten-Nutzen-Gesichtspunkten ist eine stetige Eigenproduktion von Testaufgaben der einzelnen studienfachspezifischen Beratungstests auch im Hinblick sich verändernder Fächerinhalte durch die Einführung von Bachelor- und Masterstudiengängen anzustreben. Aus diesem Grund und infolge der bisher gemachten Erfahrungen erwuchs eine diesjährig beginnende Längsschnittstudie zwischen der UniBw Hamburg, der Offizierbewerberprüfzentrale und dem Streitkräfteamt unter der organisatorischen Leitung der Gruppe Wehrpsychologie im Streitkräfteamt. Diese soll die kontinuierliche Sicherstellung studienfeldbezogener Testauf-

gaben gewährleisten. Bei der Entwicklung neuer Aufgaben sollten die Vorerfahrungen aus den Modellversuchen genutzt werden, um Testkonzepte weiter- oder gänzlich neu zu entwickeln. Die neuen Aufgaben können dabei dieselben Strukturmerkmale wie die ursprünglichen aufweisen. Ihre Bewältigung fordert somit stets gleichartige kognitive Prozesse. Gleichsam werden Persönlichkeitsfragebögen stetig aktuellen Anforderungen angepasst. Mit Hilfe objektiver Persönlichkeitstests bzw. experimentalpsychologischer Verhaltensdiagnostik würden sogar für Studieneignung gewichtige Merkmale unverfälscht erfassbar (Kubinger, 2006). Da sowohl Unternehmensfokus als auch Zweckrichtung und Prozessintegration in den Personalbeschaffungsprozess variabel sind, erweist sich eine kombinierte Form von internet-basierter Personal- und Studierendenauswahl mit Personalmarketingaspekten gerade für die Offizierauswahl der Bundeswehr als zielführend.

## E ABBILDUNGSVERZEICHNIS

<i>Abbildung 1:</i> Konstrukte in einem räumlichen Netz (Brocke & Beauducel, 2001) .....	28
<i>Abbildung 2:</i> Die drei Eigenschaften des Kriteriums .....	37
<i>Abbildung 3:</i> Probleme der Indikator-Kriteriumsbeziehungen (nach Wottawa & Hossiep, 1987) ..	37
<i>Abbildung 4:</i> Modell der Eignung für verschieden Studienfelder (Troost, 2003, S. 13).....	43
<i>Abbildung 5:</i> Zusammenhang zwischen Bewerbermerkmalen und Studienleistung.....	45
<i>Abbildung 6:</i> Bedingungsmodell des Studienerfolgs (Rindermann & Oubaid, 1999, S. 176) .....	47
<i>Abbildung 7:</i> Soziologisches Modell des Drop-Out-Prozesses nach Spady (1970, S. 79) .....	49
<i>Abbildung 8:</i> Tintos konzeptionelles Studienabbruchmodell (Tinto, 1975, S. 95).....	50
<i>Abbildung 9:</i> Phasen der Eignungsfeststellung von Bewerbern - Idealtypische Darstellung .....	57
<i>Abbildung 10:</i> Zusammenhang von Eignungsmerkmalen, Bewährungswahrscheinlichkeit und Eignungsgrad im Eignungsfeststellungsverfahren der Offizierberberprüfzentrale .....	68
<i>Abbildung 11:</i> Erwartete Prüfergebnisse bezogen auf die Bw-Noten-Verteilung .....	69
<i>Abbildung 12:</i> Nutzen und positive Effekte optimierter standardisierter Studieneignungstestbatterien.....	92
<i>Abbildung 13:</i> Ablauforganisation Pilotphase I (Universität der Bundeswehr Hamburg) Projekt Studieneignung.....	93
<i>Abbildung 14:</i> Ablauforganisation Pilotphase II (Universität der Bundeswehr München) Projekt Studieneignung.....	94
<i>Abbildung 15:</i> Anforderungsdimensionen aufbereitet und gewichtet in Anlehnung an Kastner (2004).....	96
<i>Abbildung 16:</i> Fünfteilung der Validität. Die Ecken des Fünfecks zielen jeweils auf eine der Validitätsformen nach Barell (1992) .....	107
<i>Abbildung 17:</i> Das nomologische Netzwert. Die Pfeile stellen die Beziehung zwischen Test bzw. Indikator und Konstrukten sowie jeweils untereinander dar.....	108
<i>Abbildung 18:</i> Zugehörigkeit der Gesamtstichprobe der Pilotphasen I und II zu Studienfeldern und Studienfächern an den UniBw und der FHS Bw. Angaben in Prozent (WirtschaftsWi: Wirtschaftswissenschaften; FHS: Fachhochschule der Bundeswehr München; E-Technik: Elektrotechnik; LRT: Luft- und Raumfahrttechnik) .....	115
<i>Abbildung 19:</i> Vorhersage des Kriteriumswertes (abhängige Variable) über verschiedene Prädiktorwerte (unabhängige Variable) zur Erfassung der prognostischen Validität im AC Bw und im optimierten Studieneignungstest .....	118

<i>Abbildung 20:</i> Durchschnittliches Persönlichkeitsprofil eines Offizieranwärters .....	121
<i>Abbildung 21:</i> Histogramme der Gesamtindexverteilung (Allgemeiner Teil I und spezifischer Teil) mit Normalverteilungskurve .....	122
<i>Abbildung 22:</i> Verteilung der Prädiktoren (Noten gemäß OPZ-Notenverteilung) im Allgemeinen Teil I der Testbatterie Studieneignung. Anmerkung: MW = Mittelwert .....	122
<i>Abbildung 23:</i> Verteilung der Prädiktoren (Noten gemäß OPZ-Notenverteilung) im Allgemeinen Teil I der Testbatterie Studieneignung .....	123
<i>Abbildung 24:</i> Histogramm des Gesamtindex der Testbatterie Studieneignung (Allgemeiner Teil I und spezifischer Teil) bezogen auf Studenten der Geisteswissenschaften mit Normalverteilungskurve (n = 102) .....	124
<i>Abbildung 25:</i> Histogramm des Gesamtindex der Testbatterie Studieneignung (Allgemeiner Teil I und spezifischer Teil) bezogen auf Studenten der Wirtschaftswissenschaften mit Normalverteilungskurve (n = 86) .....	124
<i>Abbildung 26:</i> Histogramm des Gesamtindex der Testbatterie Studieneignung (Allgemeiner Teil I und spezifischer Teil) bezogen auf Studenten der Ingenieurwissenschaften mit Normalverteilungskurve (n = 51) .....	125
<i>Abbildung 27:</i> Histogramm des Gesamtindex der Testbatterie Studieneignung (Allgemeiner Teil I und spezifischer Teil) bezogen auf Studenten der Informatik mit Normalverteilungskurve (n = 65).....	125
<i>Abbildung 28:</i> Histogramm vorhandener Schulnotenmittelwerte (N = 301) der Testpersonen aus den Pilotphasen I und II mit Normalverteilungskurve.....	126
<i>Abbildung 29:</i> Häufigkeitsverteilung Studienempfehlung aller Testpersonen (N = 312) aus den Pilotphasen I und II.....	127
<i>Abbildung 30:</i> Häufigkeitsverteilung des OPZ-Gesamteignungsgrad aller Testpersonen (N = 312) aus den Pilotphasen I und II .....	128
<i>Abbildung 31:</i> Histogramm der Vordiplomnoten der Testpersonen aus den Pilotphasen I und II an den UniBw mit Normalverteilungskurve (n = 308).....	128
<i>Abbildung 32:</i> Häufigkeits- und Normalverteilung der Vordiplomnoten der Testpersonen (n = 308) aus den Pilotphasen I und II an den UniBw .....	129
<i>Abbildung 33:</i> Häufigkeitsverteilung Zugehörigkeit Teilstreitkraft aller Testpersonen (N = 312) aus den Pilotphasen I und II .....	129
<i>Abbildung 34:</i> Häufigkeitsverteilung der Anzahl Trimester bis Vordiplom aller Testpersonen (N = 312) aus den Pilotphasen I und II.....	130
<i>Abbildung 35:</i> Beispielaufgabe Untertests Figurenerkennen .....	133
<i>Abbildung 36:</i> Beispielaufgabe des ADAPT-Untertests Verbale Analogien.....	133

<i>Abbildung 37:</i> Beispielaufgabe des Untertests Sprachensysteme.....	140
<i>Abbildung 38:</i> Beispielaufgabe des Untertests Textverständnis I (bezogen auf den zugrunde liegenden Text) .....	141
<i>Abbildung 39:</i> Beispielaufgabe des Untertests Modellanalyse .....	146
<i>Abbildung 40:</i> Beispielaufgabe des Untertests Diagramme und Tabellen .....	147
<i>Abbildung 41:</i> Beispielaufgabe des Untertests Formalisieren technischer Zusammenhänge .....	152
<i>Abbildung 42:</i> Beispielaufgabe des Untertests Technisches Verständnis .....	152
<i>Abbildung 43:</i> Beispielaufgabe des Untertests Flussdiagramme .....	157
<i>Abbildung 44:</i> Beispielaufgabe des Untertests Formale Systeme .....	158
<i>Abbildung 45:</i> Grafische Gegenüberstellung des inkrementellen Validitätswachstums der Regressionsanalyse zur Überprüfung verschiedener Prädiktorkombinationen auf den Studienerfolg in Form der Vordiplomnote. GeiWi: Geisteswissenschaften; WiWi: Wirtschaftswissenschaften; IngWi: Ingenieurwissenschaften; AI: Allgemeiner Teil I; Sp: Spezifischer Teil; AII: relevante Persönlichkeitsmerkmale aus Allgemeinem Teil II; Abi: Abiturnote .....	188
<i>Abbildung 46:</i> Grafische Gegenüberstellung des inkrementellen Validitätswachstums der Regressionsanalyse zur Überprüfung verschiedener Prädiktorkombinationen auf den Studienerfolg in Form der Hauptdiplomnote. GeiWi: Geisteswissenschaften; WiWi: Wirtschaftswissenschaften; IngWi: Ingenieurwissenschaften; AI: Allgemeiner Teil I; Sp: Spezifischer Teil; AII: relevante Persönlichkeitsmerkmale aus Allgemeinem Teil II; Abi: Abiturnote .....	189
<i>Abbildung 47:</i> Gegenüberstellung der jährlichen Erfolgsquote (Hauptdiplom bestanden) und mittleren Abiturnote der Studentenhjahrgänge 1989-98 an den Universitäten der Bundeswehr .....	191
<i>Abbildung 48:</i> Gegenüberstellung von Studienerfolg (Hauptdiplom bestanden) und mittlerer Abiturnote der einzelnen Studiengänge an den Universitäten der Bundeswehr in Hamburg und München .....	193
<i>Abbildung 49:</i> Zusammenhang zwischen schulischer Leistungsfähigkeit (Abiturnote) und Studienerfolg (Hauptdiplomsnote) innerhalb der einzelnen Studiengänge an den Universitäten der Bundeswehr in Hamburg und München. (WiOrg = Wirtschafts- und Organisationswissenschaften, WiIngWesen = Wirtschaftsingenieurwesen, E-Technik = Elektrotechnik.).....	194
<i>Abbildung 50:</i> Gegenüberstellung der Studienabbrecherquoten der Studentenhjahrgänge 1989-98 pro Teilstreitkraft .....	195
<i>Abbildung 51:</i> Gegenüberstellung der Studienabbrecherquoten der Studentenhjahrgänge 1989-98 pro UniBw Hamburg und München .....	195

<i>Abbildung 52:</i> Häufigkeitsverteilung des OPZ-Gesamt-eignungsgrads in Abhängigkeit der Teilstreitkraft in der Gesamtstichprobe .....	197
<i>Abbildung 53:</i> Häufigkeitsverteilung der Aussage „Ich wurde über den Hintergrund der Testung ausreichend informiert.“ .....	200
<i>Abbildung 54:</i> Häufigkeitsverteilung der Aussage „Die Testung am Laptop ist einfach und anwenderfreundlich.“ .....	201
<i>Abbildung 55:</i> Häufigkeitsverteilung der Aussage „Meine Potenziale wurden durch die Testung treffen ermittelt.“ .....	201
<i>Abbildung 56:</i> Häufigkeitsverteilung der Aussage „Mit diesem Testverfahren hat man ein geeignetes Instrument zur Feststellung der Studieneignung entwickelt.“ .....	202
<i>Abbildung 57:</i> Häufigkeitsverteilung der Aussage „Mit diesen Aufgaben kann man eine Studieneignung für meinen Fachbereich prüfen.“ .....	202
<i>Abbildung 58:</i> Häufigkeitsverteilung der Aussage „Die studienfeldspezifische Testbatterie ist standardisierter und aussagekräftiger als die bisherige OPZ-Studieneignungsprüfung.“ .....	203
<i>Abbildung 59:</i> Häufigkeitsverteilungen im allgemeinen Teil I, im spezifischen Teil und im allgemeinen Teil II zur Aussage „Die Aufgaben erinnerten mich an wichtige Inhalte bzw. an verlangte Fähigkeiten meines Studiums“ auf einer Skala von 1 („ich stimme gar nicht zu“) bis 7 („ich stimme vollkommen zu“) .....	204
<i>Abbildung 60:</i> Fragen zur Persönlichkeit bilden grundlegende Eigenschaften für ein erfolgreiches Studieren ab .....	204
<i>Abbildung 61:</i> Häufigkeitsverteilung der Aussage „Ich frage mich, wer diese Fragen als Bewerber ehrlich beantworten würde“ .....	204

## F TABELLENVERZEICHNIS

Tabelle 1: <i>Arten und Charakteristika von Studierfähigkeitstests</i> .....	14
Tabelle 2: <i>Schematischer Überblick über Prädiktoren zur Vorhersage des Konstrukts Studieneignung</i> .....	35
Tabelle 3: <i>Schematischer Aufbau der sequentiellen Prozedur des Abitur-Test-Interview-Modells (ATIM) als Beispiel für ein sequentielles Zulassungsmodell mit vorgegebenen Quoten nach Rindermann &amp; Oubaid (1999)</i> .....	44
Tabelle 4: <i>Ermittlung der Studienabbruchquoten deutscher Hochschulen in Prozent. Quellen: Griesbach et al., 1977, S 167; Reissert &amp; Marciszewski, 1987, S. 3; Griesbach et al. 1992, S. 7; Lewin et al. 1995, S. 2</i> .....	53
Tabelle 5: <i>Studienabbruchquoten für deutsche Studierende an Universitäten nach Fächergruppen und ausgewählten Studienbereichen in Prozent (HIS- Studienabbruchuntersuchung, 2005, S. 19)</i> .....	53
Tabelle 6: <i>Studienbereichen in Prozent (HIS-Studienabbruchuntersuchung Studienabbruchquoten für deutsche Studierende an Fachhochschulen nach Fächergruppen und ausgewählten, 2005, S. 23)</i> .....	54
Tabelle 7: <i>Auswahl an Methoden der Anforderungsanalyse mit ihren Vor- und Nachteilen (Reimann)</i> .....	59
Tabelle 8: <i>Die 12 bei der Offizierbewerberprüfzentrale gemessenen Eignungsmerkmale und ihre Definition</i> .....	65
Tabelle 9: <i>Studieneingangsbeurteilungsergebnisse der Offizierbewerberprüfzentrale im Überblick</i> .....	71
Tabelle 10: <i>Überblick über Beurteilungen der Offizierbewerber im Assessment Center</i> .....	72
Tabelle 11: <i>Erfassung der Anforderungen in den jeweiligen Übungen im AC Bw. Die grauen Kästchen stellen die gemessenen Dimensionen dar. BF: Biographischer Fragebogen, Psych.: Psychische, Sprache: Sprachliche Ausdrucksfähigkeit und Argumentationsvermögen, MathKT: Mathematischer Kenntnistest, Schwarz: In die Studienberatung einfließende Ergebnisse</i> .....	74
Tabelle 12: <i>Übersicht der an den beiden Universitäten und einer Fachhochschule der Bundeswehr zu studierende Studienfächer</i> .....	79
Tabelle 13: <i>Studienabbruchquoten für deutsche Studierende an Universitäten und Fachhochschulen nach an den UniBw vorhandenen Fächergruppen und</i>	

ausgewählten Studienbereichen in Prozent (BMVg, 2006; HIS-Studienabbruchuntersuchung, 2005, S. 19) .....	82
Tabelle 14: Übersicht Testbatterie Studieneignung unterteilt nach Studienfelder der Universitäten der Bundeswehr. ' = Minutenangabe .....	97
Tabelle 15: Zu messende fächerübergreifende Basiskompetenzen gemäß Anforderungsprofil des Kastner-Abschlussberichts und in den Pilotphasen an den UniBw angewandte Prädiktoren.....	98
Tabelle 16: Moderatorvariablen in der Studieneignungsvorhersage der Bundeswehr.....	118
Tabelle 17: Überblick Testbatterie Studieneignung über Testergebnisverteilung aller Probanden .....	119
Tabelle 18: Die Hypothesen der Evaluation des Eignungsfeststellungsverfahrens an der Offizierbewerberprüfzentrale der Bundeswehr mit statistischer Formulierung.....	131
Tabelle 19: Übersicht über die internen Konsistenzkoeffizienten für den allgemeinen Teil der Testbatterie Geisteswissenschaften bezogen auf beide Pilotphasen ohne Untertest Figurenreihen.....	134
Tabelle 20: Kennwerte des Tests Figuren erkennen bezogen auf Pilotphase I.....	135
Tabelle 21: Cronbachs Alpha des Tests Verbale Analogien bezogen auf Pilotphase I.....	135
Tabelle 22: Kennwerte des Tests Figuren erkennen bezogen auf Pilotphase II .....	136
Tabelle 23: Cronbachs Alpha des Tests Verbale Analogien bezogen auf Pilotphase II .....	136
Tabelle 24: Kennwerte des Tests Figuren erkennen bezogen auf Pilotphase I und II.....	137
Tabelle 25: Kennwerte des Tests Figurenreihen bezogen auf Pilotphase I und II .....	137
Tabelle 26: Kennwerte des Tests Verbale Analogien bezogen auf Pilotphase I und II.....	138
Tabelle 27: Normen der Tests Figuren erkennen und Verbale Analogien .....	138
Tabelle 28: Übersicht über die internen Konsistenzkoeffizienten für den Spezifischen Teil der Testbatterie Geisteswissenschaften bezogen auf beide Pilotphasen .....	141
Tabelle 29: Kennwerte des Tests Sprachensysteme bezogen auf Pilotphase I.....	142
Tabelle 30: Kennwerte des Tests Textverständnis I bezogen auf Pilotphase I.....	142
Tabelle 31: Kennwerte des Tests Textverständnis II bezogen auf Pilotphase I.....	143
Tabelle 32: Kennwerte des Tests Sprachensysteme bezogen auf Pilotphase II .....	143
Tabelle 33: Kennwerte des Tests Textverständnis I bezogen auf Pilotphase II .....	144
Tabelle 34: Kennwerte des Tests Textverständnis II bezogen auf Pilotphase II.....	144
Tabelle 35: Kennwerte des Tests Sprachensysteme bezogen auf Pilotphase I und II.....	144
Tabelle 36: Kennwerte des Tests Textverständnis I bezogen auf Pilotphase I und II .....	145
Tabelle 37: Kennwerte des Tests Textverständnis II bezogen auf Pilotphase I und II .....	145
Tabelle 38: Übersicht über die internen Konsistenzkoeffizienten für den Spezifischen Teil der Testbatterie Wirtschaftswissenschaften bezogen auf beide Pilotphasen.....	147

Tabelle 39: Kennwerte des Tests Modellanalyse bezogen auf die Pilotphase I.....	147
Tabelle 40: Kennwerte des Tests Diagramme und Tabellen bezogen auf die Pilotphase I.....	148
Tabelle 41: Kennwerte des Tests Modellanalyse bezogen auf die Pilotphase II .....	149
Tabelle 42: Kennwerte des Tests Diagramme und Tabellen bezogen auf die Pilotphase II .....	149
Tabelle 43: Kennwerte des Tests Modellanalyse bezogen auf die Pilotphase I und II.....	150
Tabelle 44: Kennwerte des Tests Diagramme und Tabellen bezogen auf die Pilotphase I und II .	151
Tabelle 45: Übersicht über die internen Konsistenzkoeffizienten für den Spezifischen Teil der Testbatterie Wirtschaftswissenschaften bezogen auf beide Pilotphasen.....	153
Tabelle 46: Kennwerte des Tests Formalisieren technischer Zusammenhänge bezogen auf die Pilotphase I .....	153
Tabelle 47: Kennwerte des Tests technisches Verständnis bezogen auf die Pilotphase I .....	154
Tabelle 48: Kennwerte des Tests Formalisieren technischer Zusammenhänge bezogen auf die Pilotphase II .....	154
Tabelle 49: Kennwerte des Tests technisches Verständnis bezogen auf die Pilotphase II.....	155
Tabelle 50: Kennwerte des Tests Formalisieren technischer Zusammenhänge bezogen auf die Pilotphase I und II .....	156
Tabelle 51: Kennwerte des Tests technisches Verständnis bezogen auf die Pilotphase I und II ....	156
Tabelle 52: Übersicht über die internen Konsistenzkoeffizienten für den Spezifischen Teil der Testbatterie Informatik .....	158
Tabelle 53: Kennwerte des Tests Flussdiagramme .....	159
Tabelle 54: Kennwerte des Tests Formale Systeme .....	159
Tabelle 55: Normen der Tests Sprachensysteme und Textverständnis I und II.....	160
Tabelle 56: Normen der Tests Modellanalyse und Diagramme und Tabellen .....	162
Tabelle 57: Normen der Tests Formalisieren technischer Zusammenhänge und Technisches Verständnis .....	163
Tabelle 58: Normen der Tests Flussdiagramme und Formale Systeme .....	165
Tabelle 59: Übersicht aller Skalen des Allgemeinen Teils II (Persönlichkeitsfragebögen).....	166
Tabelle 60: Übersicht über die internen Konsistenzkoeffizienten für den Persönlichkeitstest PMO über beide Pilotphasen hinweg .....	174
Tabelle 61: Übersicht über die internen Konsistenzkoeffizienten für Persönlichkeitstest BIG V über beide Pilotphasen hinweg .....	175
Tabelle 62: Übersicht über die internen Konsistenzkoeffizienten für Persönlichkeitstest Stressbewältigungsressourcen über beide Pilotphasen hinweg.....	175
Tabelle 63: Übersicht über die internen Konsistenzkoeffizienten für den allgemeinen Teil insgesamt über beide Pilotphasen hinweg ohne Subtest Figurenreihen .....	176

Tabelle 64: Übersicht über die internen Konsistenzkoeffizienten für den allgemeinen Teil II über beide Pilotphasen hinweg .....	176
Tabelle 65: Übersicht über die internen Konsistenzkoeffizienten für den gesamten spezifischen Teil über beide Pilotphasen hinweg ohne Subtests Textverständnis I und II und Flussdiagramme .....	177
Tabelle 66: Allgemeiner Teil I und spezifischer Teil der Testbatterie Studieneignung: Erklärter Prozentanteil der vier Faktoren an der Gesamtvarianz. ....	178
Tabelle 67: Ergebnisse der Hauptkomponentenanalyse (varimax-rotierte Komponentenmatrix). Die Variablen sind den Testteilen zugeordnet. ....	178
Tabelle 68: Allgemeiner Teil II der Testbatterie Studieneignung: Erklärter Prozentanteil der vier Faktoren an der Gesamtvarianz. ....	179
Tabelle 69: Ergebnisse der Hauptkomponentenanalyse (varimax-rotierte Komponentenmatrix). Die Variablen sind den Testteilen zugeordnet. ....	180
Tabelle 70: Ergebnisse der Korrelationsanalyse zur Überprüfung der prognostischen Validität und Gegenüberstellung früherer OPZ-Studieneignungsbeurteilung versus neuer Gesamtestbatterie Studieneignung (studienfachübergreifend). ....	182
Tabelle 71: Vergleich der Rangkorrelationskoeffizienten der Prädiktoren aus dem Allgemeinen Teil II nach Spearman mit Vor- und Hauptdiplomnote als Kriterien für Studienerfolg. ....	186
Tabelle 72: Inkrementelle Validitäten - Ergebnisse der Regressionsanalyse zur Überprüfung des Einflusses verschiedener Prädiktorkombinationen auf den Studienerfolg in Form der Vordiplomnote .....	187
Tabelle 73: Inkrementelle Validitäten - Ergebnisse der Regressionsanalyse zur Überprüfung des Einflusses verschiedener Prädiktorkombinationen auf den Studienerfolg in Form der Hauptdiplomnote. ....	189
Tabelle 74: Übersicht der Studiengänge an den Universitäten der Bundeswehr in Hamburg und München, untergliedert nach Studienfeldern (Bezugszeitraum: 1989-1998). ....	192
Tabelle 75: Mittelwertvergleich der Hauptdiplomnoten in Abhängigkeit des väterlichen Berufs..	198
Tabelle 76: Mittelwerte für das Kriterium Studienerfolg in Form der Hauptdiplomnote (falls schon vorhanden) zwischen männlichen und weiblichen Studenten .....	198
Tabelle 77: Mittelwerte für das Kriterium Studienerfolg in Form der Hauptdiplomnote (falls schon vorhanden) zwischen Studierenden der Universität der Bundeswehr in Hamburg und der in München .....	198
Tabelle 78: Zufallsauswahl der offenen Antworten auf die Frage „Gibt es etwas, dass Sie uns darüber hinaus gerne mitteilen wollen?“ .....	205

## G LITERATURVERZEICHNIS

15. KOMMANDEURTAGUNG ZMILDBW. (1998). *Führung, Ausbildung und Erziehung der Studierenden Offiziere an den Universitäten der Bundeswehr*. Bonn: BMVg.
- ABELE-BREHM, A. E. & STIEF, M. (2004). Die Prognose des Berufserfolgs von Hochschulabsolventinnen und -absolventen. *Zeitschrift für Arbeits- und Organisationspsychologie*, 48, 4-16.
- AMELANG, M. (1997). Differentielle Aspekte der Hochschulzulassung: Probleme, Befunde, Lösungen. In Th. Hermann (Hrsg.), *Hochschulentwicklung – Aufgaben und Chancen* (88-105). Heidelberg: Asanger.
- AMELANG, M. (1980). Tests als Instrumente der Hochschulzulassung. Anmerkungen zu einer vielschichten Problematik. In U. Göbel (Hrsg.), *Tests - Wegweiser für die Bildungs- und Berufsplanung*. Köln, Deutscher Institutsverlag, 125-163.
- AMELANG, M. (1976). Erfassung einiger Kriterien des Studienerfolgs in mehreren Fachrichtungen mit Hilfe von Leistungs- und Persönlichkeitstests. *Psychologie in Erziehung und Unterricht*, 23, 259-273.
- AMELANG, M. & BARTUSSEK, (1997). *Hochschulzulassung und Studieneignung: Studienfeldbezogene Verfahren zur Feststellung der Eignung für Numerus-clausus- und andere Fächer*. Göttingen, Zürich: Vandenhoeck & Ruprecht.
- AMELANG, M. & HOPPENSACK, T. (1976). *Persönlichkeitsstruktur und Hochschulbesuch I. Merkmalsveränderungen während des Studiums bei Studierenden verschiedener Fachrichtungen*. Arbeiten aus dem Psychologischen Institut der Universität Hamburg.
- AMELANG, M. & ZIELINSKI, W. (1997). *Psychologische Diagnostik und Intervention*. Heidelberg: Springer.
- ARNHOLD, N. & HACHMEISTER, C.-D. (2004). *Leitfaden für die Gestaltung von Auswahlverfahren an Hochschulen*. Arbeitspapier Nr. 52, Gütersloh.
- ASSOULINE, M. & MEIER, E. I. (1987). Metaanalysis of the relationship between congruence and wellbeing measures. *Journal of Vocational Behavior*, 31, 319-332.
- BARELL, G. (1992). *Bewährungskontrollen von Assessment Centern mittels testtheoretischer Validitätsmodelle*. Bern: Lang.

- BARON-BOLDT, J. ET AL. (1989). Prognostische Validität von Schulnoten. Eine Metaanalyse der Prognose des Studien- und Ausbildungserfolgs. In S. Jäger et al. (Hrsg.), *Tests und Trends 7. Jahrbuch der Pädagogischen Diagnostik*. Beltz: Weinheim.
- BARON-BOLDT, J., SCHULER, H. & FUNKE, U. (1988). Prädiktive Validität von Schulabschlussnoten: Eine Metaanalyse. *Zeitschrift für Pädagogische Psychologie*, 2 (2), 79-90.
- BARTONE, P. T. (1999). Personality Hardiness as a Predictor of Officer Cadet Leadership Performance. Paper presented at the RTO HFM Workshop on „Officer Selection“, held in Monterey, USA, 9-11, November 1999, published in RTO MP-55.
- BAUMERT, J. & WATERMANN, R. (2000). Institutionelle und regionale Variabilität und die Sicherung gemeinsamer Standards in der gymnasialen Oberstufe. In J. Baumert, W. Bos & R. Lehmann (Hrsg.), *TIMSS/III. Band 2: Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe* (317-372). Opladen: Leske & Budrich.
- BECKER, B.J. (1990). Coaching fort he scholastic aptitude test: Further synthesis and appraisal. *Review of Educational Research*, 60 (3), 373-417.
- BECKER, P. (2003). Persönlichkeitsfragebogen. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (332-337). Weinheim: Beltz/PVU.
- BERGMANN, C. BRANDSTÄTTER & EDER, F. (1994). *Studienberatungstests für Maturanten. Abschlussbericht. Universität. Linz: Institut für Pädagogik und Psychologie.* (Forschungsbericht im Auftrag des BM für Wissenschaft und Forschung).
- BERMGANN, C. (1992). Schulisch-berufliche Interessen als Determinanten der Studien- bzw. Berufswahl und -bewältigung: Eine Überprüfung des Modells von Holland. In A. Krapp & M. Prenzel (Hrsg.), *Interesse, Lernen, Leistung: Neuere Ansätze einer pädagogisch-psychologischen Interessenforschung* (195-220). Münster: Aschendorff.
- BETZ, D. (1974). Rhythmische Schwankungen als Fehler in der Notegebung bei mündlichen Prüfungen. *Psychologie in Erziehung und Unterricht*, 21, 1-14.
- BIÄSCH, H. & ROHR, A. (1968). *Beiträge zur Talentforschung. Eine Studie über Studenten an der ETH*. Bern/Stuttgart.
- BLUM, M. & NAYLOR, J. C. (1968). *Industrial Psychology: Its Theoretical and Social Foundation*. New York: Harper and Row.
- BMVG FÜ S I 4. (2006). *Streitkräftegemeinsames Anforderungsprofil für Offiziere der Bundeswehr*. Unveröffentlicht. BMVg

- BOHLEBER, W. (1982). Spätadoleszente Entwicklungsprozesse. Ihre Bedeutung für Diagnostik und psychotherapeutische Behandlung von Studenten. In: E. Krejci & W. Bohleber (Hrsg.): *Spätadoleszente Konflikte. Indikation und Anwendung psychoanalytischer Verfahren bei Studenten*. Göttingen: Verlag für Medizinische Psychologie.
- BONNEMANN, A. & POSNER, C. (2002). *Studieren an den UniBw. Vergleichende Ergebnisse aus der Studentenforschung 1999*. Hamburg: Zentrum für Hochschulforschung und Qualitätssicherung.
- BORKENAU, P. & OSTENDORF, F. (1993). *NEO-Fünf-Faktoren Inventar (NEO-FFI) nach Costa und McCrae. Handanweisung*. Göttingen: Hogrefe.
- BRANDSTÄTTER, H. FARTHOFER, A. & GRILLICH, L. (2001). Die Stabilität der Studienwahl als Funktion von Interessenkongruenz, Selbstkontrolle und intellektueller Leistungsfähigkeit. *Psychologie in Erziehung und Unterricht*, 48, 200-218.
- BROCKE, B. (1990). Struktur und Prozeß. In W. Sarges (Hrsg.), *Management-Diagnostik*. Göttingen: Hogrefe.
- BROWN, W.F. & HOLTZMANN, W.H. (1955). *A Study Attitudes Questionnaire For Predicting Academic Success*, Philadelphia.
- BUNDESAGENTUR FÜR ARBEIT. (2004). Studienfeldbezogene Beratungs-Testserie – Handanweisungen. Nürnberg.
- BUNDESMINISTERIUM DER VERTEIDIGUNG. (1970). *Weissbuch zur Sicherheit der Bundesrepublik Deutschland und zur Lage der Bundeswehr*. Bonn. BMVg.
- BUNESMINISTERIUM DER VERTEIDIGUNG. (1996a). *Bestimmungen zur Methodik der Eignungsfeststellung bei Offizierbewerbern*. Bonn: BMVg - PSZ III 4 - Az 66-10-03.
- BUNESMINISTERIUM DER VERTEIDIGUNG. (1997). *Fachkonzept „Personalpsychologische Qualitätssicherung“*. Bonn: BMVg - PSZ III 6 - Az 66-01-10.
- BUNESMINISTERIUM DER VERTEIDIGUNG. (1998). *Aufnahmebestimmungen für Offizierbewerber*. Bonn: BMVg - PSZ III 3 - Az 16-20-10-01.
- BUNDESMINISTERIUM DER VERTEIDIGUNG. (1999). *Personalpsychologische Qualitätssicherung im Psychologischen Dienst der Bundeswehr - Bewährungskontrollen*. Bonn: BMVg - PSZ III 4 - Az 66-60-15.
- BUNDESMINISTERIUM DER VERTEIDIGUNG. (2005). *PSZ I 1 Personalinformation 2005*. Bonn: BMVg.
- BUNDESMINISTERIUM DER VERTEIDIGUNG. (2006a). *Das Weißbuch 2006 zur Sicherheitspolitik Deutschlands und zur Zukunft der Bundeswehr*. Berlin: BMVg.

- BUNDESMINISTERIUM DER VERTEIDIGUNG. (2006b). *PSZ I 1 Personalinformation 2006*. Bonn. BMVg.
- BUNDESZENTRALE FÜR POLITISCHE BILDUNG. (2002). *Aus Politik und Zeitgeschichte*. Bonn: Bundeszentrale für politische Bildung.
- BUNGARD, W. (1987). Zur Problematik von Reaktivitätseffekten bei der Durchführung eines Assessment Centers. In H. Schuler & Stehle (Hrsg.), *Assessment Center als Methode der Personalentwicklung* (99-125). Stuttgart : Verlag für Angewandte Psychologie.
- BULTMANN, T. (1996). Die standortgerechte Dienstleistungshochschule. In: *Prokla 104, Zeitschrift für kritische Sozialwissenschaft, "Universität"*, 26. (3), 329-355.
- CAMARA, W. J. & KIMMEL, E. W. (2005). *Choosing students: Higher education admissions tools for the 21st century*. Mahwah, NJ: Lawrence Erlbaum.
- CAMPBELL, D. T. & FISKE, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 81, 81-105.
- CAMPBELL, J. P., DUNETTE, M. D., LAWLER, E. E. & WEICK, K. E., JR. (1970). *Managerial behavior, performance, and effectiveness*. New York: McGraw-Hill.
- CAMPION, M. A., CAMPION, J. E. & HUDSON, J. P. (1994). Structured interviewing: A note on incremental validity and alternative question types. *Journal of Applied Psychology*, 6, 998-1102.
- CAPLAN, R. D. (1987). Person-environment fit theory and organizations: Commensurate dimensions, time perspectives, and mechanisms. *Journal of Vocational Behavior* 31 (3), 248-267.
- CARTER, R.S. (1953). Non-intellectual variables involved in teachers marks. *Journal of Educational Research*, 47.
- CATTELL, R. B. (1987). *Intelligence: Its structure, growth, and action*. Amsterdam: Elsevier.
- CATTELL, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.
- CATTELL, R. B. (1963). Theory of fluid and crystallized intelligence: a critical experiment. *Journal of Educational psychology*, 54, 1-22.
- CATTELL, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, 38, 426-506.
- CHUR, D. (2002). Entwicklungslinien der Studierendenberatung: Von der traditionellen Studienberatung zum Kompetenz-Center für (Aus-) Bildungsqualität. In: Nestmann, F. & Engel, F. (Hrsg.): *Die Zukunft der Beratung*. Tübingen: dgvt-Verlag

- CRAWLEY, B., PINDER R. & HERRIOT, P. (1990). Assessment center dimensions, personality, and aptitudes. *Journal of Occupational Psychology*, 63, 211-216.
- CRONBACH, L. J. (1951). Coefficient Alpha And The Internal Structure of Tests. *Psychometrika*, 16, 297-334.
- CRONBACH, L. J. (1947). Test "Reliability": Its Meaning and Determination. *Psychometrika*, 12, 1-16.
- DÄUMLING, A.M. (1968). *Psychologische Probleme des Hochschulexamens*, Jahrbuch für Psychologie, Psychotherapie und medizinische Anthropologie, 16, 104-115.
- DE RAAD, B. & PERUGINI, M. (2002). *Big Five Assessment*. Bern: Hogrefe & Huber Publishers.
- DEIDESHEIMER KREIS. (AMELANG, M., BARTUSSEK, D., BRACKMANN, H.-J., EGLI, H., HAA-SE, K., HINRICHSSEN, K., KLAUER, K. J., KORNADT, H.-J., MICHEL, L. & TROST, G.) (1997). *Hochschulzulassung und Studieneignungstests. Studienfeldbezogene Verfahren zur Feststellung der Eignung für Numerus clausus und andere Studiengänge*. Göttingen/Zürich: Vandenhoeck & Ruprecht.
- DEIDESHEIMER KREIS. (AMELANG, M., BARTUSSEK, D., BRACKMANN, H.-J., EGLI, H., HAA-SE, K., HINRICHSSEN, K., KLAUER, K. J., KORNADT, H.-J., MICHEL, L. & TROST, G.) (1997). *Hochschulzulassung und Studieneignungstests. Studienfeldbezogene Verfahren zur Feststellung der Eignung für Numerus-Clausus- und andere Studiengänge*. Göttingen: Vandenhoeck & Ruprecht.
- DEIDESHEIMER KREIS. (AMELANG, M., BARTUSSEK, D., BRACKMANN, H.-J., EGLI, H., HAA-SE, K., HINRICHSSEN, K., KLAUER, K. J., KORNADT, H.-J., MICHEL, L. & TROST, G.) (1993). *Feststellung der Studieneignung im Rahmen der Hochschulzulassung*. Bonn: Bericht für das Bundesministerium für Bildung und Wissenschaft.
- DEUTSCHE GESELLSCHAFT FÜR PSYCHOLOGIE (DGPs). (2005). Stellungnahme der Deutschen Gesellschaft für Psychologie e. V. (DGPs) zur Auswahl von Studierenden durch die Hochschulen (vom 22. November 2004). *Psychologische Rundschau*, 56, 153-154.
- DEUTSCHE GESELLSCHAFT FÜR PSYCHOLOGIE E. V. (2005). Stellungnahme der Deutschen Gesellschaft für Psychologie e. V. (DGPs). Zur Auswahl von Studierenden durch die Hochschulen (vom 22. November 2004). *Psychologische Rundschau*, 56 (2), 153-154.

- DEUTSCHES INSTITUT FÜR NORMIERUNG 33430. (2002). *Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen*. Berlin: Deutsches Institut für Normierung e.V.
- DONLON, T. F. (ED.) (1984). *The College Board Technical Handbook for the Scholastic Aptitude Test and Achievement Tests*. New York: College Entrance Examination Board.
- DURKHEIM, E. (1897). *Selbstmord*. Berlin/Neuwied.
- EDER, C. & BERGMANN, F. (1999). *Allgemeiner Interessen-Struktur-Test / Umwelt-Struktur-Test. Manual (AIST/UST) (2. Aufl.)*. Göttingen: Beltz Test GmbH.
- ENTWISTLE, N.J. & ENTWSITLE, D. (1970). The relationships between personality, study methods and academic performance“. *Britain Journal for educational Psychology*, 40, 132-143.
- ERDFELDER, E. & RUDINGER, G. (Hrsg.). (1996). *Handbuch Quantitative Methoden..* Weinheim: Psychologische Verlagsunion.
- EZEKIEL, M. & FOX, K. A. (1959). *Methods of Correlation and Regression Analysis, linear and curvilinear*. 3. ed., Wiley, New York
- FAY, E. (1982). *Der „Test für medizinische Studiengänge“ TMS*. Braunschweig, Pedersen.
- FECKER, T. (1989). *Möglichkeiten und Grenzen einer Kontrolle der Beurteilungsqualität im Assessment-Center mit Hilfe statistischer Verfahren*. Dissertation, Universität Münster.
- FESTERLING, K.-H. (1971). Akademische Abschlussprüfungen. Baden Württemberg in Wort und Zahl. *Statistische Monatshefte*, 19, 337-342.
- FISSENI, H. J. (1997). *Lehrbuch der psychologischen Diagnostik*. Göttingen: Hogrefe.
- FLANAGAN, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 1954, 51, 327-358.
- FLECKENSTEIN, B. (2002). Die Universitäten der Bundeswehr vor neuen Aufgaben. *Gewerkschaftliche Monatshefte. Bundeswehr, Gewerkschaften, Demokratie*. 53 (2/3), 142-151.
- FRIES, M. (2002). Abitur und Studienerfolg: Welchen „Wert“ hat das Abitur für ein erfolgreiches Studium? *In: Beiträge zur Hochschulforschung*, 1 (24), 30-51.
- FRUHNER, R., SCHULER, H., FUNKE, U. & MOSER, K. (1991). Einige Determinanten der Bewertung von Personalauswahlverfahren. *Zeitschrift für Arbeits- und Organisationspsychologie*, 35, 170-178.
- GAENSSLER, H. & SCHUBÖ, W. (1973). *Einfache und komplexe statistische Analyse*. München: Reinhardt (UTB 274).

- GERPOTT, T. J. & DOMSCH, M. (1984). *Studienmotivationsaspekte, Abiturleistungsniveau und Leistung im militärischen Berufsfeld*, Manuskript, Hamburg.
- GIESEN, H. & GOLD, A. (1996). Individuelle Determinanten der Studiendauer. Ergebnisse einer Längsschnittuntersuchung. In Lompscher, J. & Mandl, H. (Hrsg.), *Lehr- und Lernprobleme im Studium* (86-99). Bern: Huber.
- GIESEN, H., BÖHMECKE, W., EFFLER, M., HUMMER, A., JANSEN, R., KÖTTER, B., KÄMER, H.-J., RABENSTEIN, E. & WERNER, R. R. (HRSG.). (1981). *Vom Schüler zum Studenten. Bildungslebensläufe im Längsschnitt*. München: Reinhardt.
- GIESEN, H., GOLD, A., HUMMER, A. & JANSEN, R. (1986). *Prognose des Studienerfolgs. Ergebnisse aus Längsschnittuntersuchungen*. Frankfurt a. M.: Institut für Pädagogische Psychologie.
- GOLD A. & KLOFT, C. (1991). Der Studienabbruch: Eine Analyse von Bedingungen und Begründungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 2, 265-279.
- GOLD, A. (1988). *Studienabbruch, Abbruchneigung und Studienerfolg: Vergleichende Bedingungsanalysen des Studienverlaufs*. Frankfurt am Main: Peter Lang.
- GOLD, A. & SOUVIGNIER, E. (1997). Examensleistung und Studierenerleben bei Hochschulabsolventen. *Zeitschrift für Pädagogische Psychologie*, 11(1), 53-63.
- GOLDBERG, L. R. (1981). (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.). *Review of personality and social psychology*, 2, 141-165. Beverly Hills, CA: Sage.
- GRIESBACH, H., LEWIN, K. & SCHACHER, M. (1977). *Studienverlauf und Beschäftigungssituation von Hochschulabsolventen und Studienabbrechern*. HIS-Hochschulplanung Bd. 27.
- GUILFORD, J. P. (1967). *The nature of human intelligence*. New York, McGraw Hill.
- GUTTHRIE, G. M. & BENNETT, A. B. (1970). Individual cultural differences in implicit personality theory. *International Journal of Psychology*, 6, 305-312.
- GUTTMAN, L. (1965). A faceted definition of intelligence. In R. Eiferman (Hrsg.), *Studies in Psychology, scripta hierosolymitana* (Bd. 14). Jerusalem: The Hebrew University.
- HARDEGGER, S. C. (2002). *Persönlichkeitstests (NEO-PI-R<sup>TM</sup>) in der Selektion von Sicherheitsangestellten*. Unveröffentlichte Lizenziatsarbeit, Universität Zürich, Psychologisches Institut, Abteilung Angewandte Psychologie.
- HECKHAUSEN, H. (1989). *Motivation und Handeln* (2. Auflage). Berlin: Springer-Verlag.

- HEINE, C., BRIEDIS, K., DIDI, H.-J., HAASE, K. & TROST, G. (2006). *Auswahl- und Eignungsfeststellungsverfahren beim Hochschulzugang in Deutschland und ausgewählten Ländern. Eine Bestandsaufnahme*. Hannover: Hochschul-Informationssystem.
- HEINE, C., SPANGENBERG, H., SCHREIBER, J. & SOMMER, D. (2005). *Studienanfänger in den Wintersemestern 2003/2004 und 2004/2005. Wege zum Studium, Studien- und Berufswahl, Situation bei Studienbeginn*. HIS-Hochschulplanung Band 180. Hannover: Hochschul-Informationssystem.
- HEINE, CH., SPANGENBERG, H., SCHREIBER, J. & SOMMER, D. (2005). *Studienanfänger in den Wintersemestern 2003/04 und 2004/05. Wege zum Studium, Studien- und Hochschulwahl, Situation bei Studienbeginn*. HIS Kurzinformation A 12/2005, Hannover.
- HEISE, E., WESTERMANN, R., SPIES, K. & SCHIFFLER, A. (1997). Studieninteresse und berufliche Orientierungen als Determinanten der Studienzufriedenheit. *Zeitschrift für Pädagogische Psychologie*, 11(2), 123-132.
- HELDMANN, W.(1984). *Studierfähigkeit*. Göttingen: Otto Schwartz.
- HELL, B. (2006). *Die Prognosekraft von Verfahren der Studierendenauswahl für den Studienerfolg*. Beitrag auf dem Workshop der Österreichischen Forschungsgemeinschaft vom 10.-11.03.2006.
- HELL, B. & SCHULER, H. (2005). Verfahren der Studierendenauswahl aus Sicht der Bewerber. *Empirische Pädagogik*, 19 (4), 361-376.
- HELL, B., TRAPMANN, S., WEIGAND, S. & SCHULER, H. (2007). Die Validität von Auswahlgesprächen im Rahmen der Hochschulzulassung – eine Metaanalyse. *Psychologische Rundschau*, 58 (2), 93-102. Hogrefe. Göttingen.
- HELL, B., TRAPMANN, S., WEIGAND, S.,HIRN, J. O. & SCHULER, H. (2005). Die Validität von PRädiktoren des Studienerfolgs - eine Metaanalyse. *Vortrag auf der 4. Tagung der Fachgruppe Arbeits- und Organisationspsychologie der Deutschen Gesellschaft für Psychologie*, Bonn.
- HINNEBERG, H. (2003). *Abiturnote und Studienerfolg*. Zentrale Studienberatung Universität Ulm.
- HINRICHS, J. R. & HAANPERÄ, S. (1976). *A technical research report on management assessment in IBM World Trade Corporation. Personnel Research Study No. 18*. IBM Europe.
- HITPASS, J. (1967). *Das Studienschicksal des Immatrikulationsjahrgangs 1958*, Gütersloh.
- HOCHSCHULINFORMATIONSSYSTEM (HIS). (2002). *HIS-Ergebnisspiegel 2002*.

- HOCHSCHULVERBAND. (1984). *Thesen zur Studierfähigkeit und zum Hochschulzugang*. Göttingen: Otto Schwartz.
- HÖDL, E. (2002). *Hochschulzugang in Europa*. Wien/Köln: Böhlau.
- HOFFACKER, W. (2004): Zu Nutzen und Kosten einer Auswahl der Studierenden durch die Hochschulen. In: *Das Hochschulwesen*, 1, 2-11.
- HÖPPEL, D. & MOSER, K. (1993). Die Prognostizierbarkeit von Studiennoten und Studierendauer durch Schulabschlussnoten. *Zeitschrift für Pädagogische Psychologie*, 7 (1), 25-32.
- HORNKE, L. & ZIMMERHOFER, A. (2005). Profilbildung von Hochschulen und Studierenden als Aufgabe der Studierendenauswahl. *Psychologische Rundschau*, 53 (2), 146-148.
- HORST, P. (1971). *Messung und Vorhersage*. Weinheim.
- HOSSIEP, R. (1996). Psychologische Tests: die vernachlässigte Dimension in Assessment Centern. In W. Sarges (Hrsg.). *Weiterentwicklungen der Assessment Center-Methode*. Göttingen, Verlag für Angewandte Psychologie.
- HÜLSHEGER, U. R., SPINATH, F. M., KÜPPERS, A. & ETZEL, S. (2004). Experimentelle Untersuchung zweier Methoden zur Reduzierung Sozialer Erwünschtheit in einem computergestützten eignungsdiagnostischen Testverfahren. *Zeitschrift für Personalpsychologie*, 3 (1), 24-33. Hogrefe. Göttingen.
- HÜLSHEGER, U. R., SPINATH, F. M., KÜPPERS, A. & ETZEL, S. (2004). Experimentelle Untersuchung zweier Methoden zur Reduzierung Sozialer Erwünschtheit in einem computergestützten eignungsdiagnostischen Testverfahren. *Zeitschrift für Personalpsychologie*, 3 (1), 24-33.
- HUNTER, J.E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, 29, 340-362.
- INGENKAMP, K. (1986). Zur Diskussion über die Leistungen unserer Berufs- und Studienanfänger. Eine kritische Bestandsaufnahme der Untersuchungen und Stellungnahmen. *Zeitschrift für Pädagogik*, 32 (1), 1-29.
- INSPEKTEUR DES HEERES. (1994). *Weisung: Anforderungen an den Offizier des Heeres*. Bonn: BMVg.
- INSPEKTEUR DES HEERES. (1997). *Ausbildung zum Offizier des Truppendienstes. Anlagen*. Bonn: BMVg.
- JAHNKE, J. (1971). *Empirische Untersuchungen über das Arbeitsverhalten im akademischen Studium, Meisenheim am Glan*.

- JERUSALEM, M. (1990). *Persönliche Ressourcen, Vulnerabilität und Stresserleben*. Göttingen: Hogrefe.
- JERUSALEM, M. (1991). Allgemeine Selbstwirksamkeit und differentielle Stressprozesse. *Psychologische Beiträge*, 33, 388-406.
- JERUSALEM, M. & SCHWARZER, R. (1999). *Allgemeine Selbstwirksamkeit* Online. Available: [http://www.Fu-berlin.de/gesund/skalen/allg.../hauptteil\\_allgemeine\\_selbstwirksamkeit](http://www.Fu-berlin.de/gesund/skalen/allg.../hauptteil_allgemeine_selbstwirksamkeit)
- JESERICH, W. (1981). *Mitarbeiter auswählen und fördern: Assessment-Center-Verfahren*. München: Hanser.
- JOHN, O. P. & SRIVASTAVA, S. (1999). The Big-Five trait taxonomy: History measurement, and theoretical perspectives. In L. Pervin & O. P. John (Eds.) *Handbook of personality theory and research* (2<sup>nd</sup> Ed.) (pp. 102-109). New York: Guilford Press.
- KANNING, U. P. (2004). Entwicklungslinien der Studierendenberatung: Von der traditionellen Studienberatung zum Kompetenz-Center für (Aus-) Bildungsqualität. In: Nestmann, F. & Engel, F. (Hrsg.): *Die Zukunft der Beratung*. Tübingen: dgvt-Verlag.
- KANNING, U. P. (2004). *Standards der Personaldiagnostik*. Berlin: Beuth.
- KASTNER, M. (2004). Unveröffentlichter Abschlussbericht Beratung und Umsetzungsförderung bei der Gestaltung der Studienberatung in der OPZ der Bundeswehr und die Entwicklung eines sequenziellen Entscheidungsmodells.
- KATH, G., OEHLER, CH. & REICHWEIN, R. (1966). *Studienweg und Studienerfolg Studien und Berichte des Instituts für Bildungsforschung in der Max-Planck-Gesellschaft*, Bd. 6, Berlin.
- KERSTING, M. & BEAUDUCEL, A. (2007). Start-P. (In Vorbereitung).
- KERSTING M., GEHRING F. (2006). *University Citizenship Behavior – die andere Seite des Studienerfolgs*. Symposium auf dem 45. Kongress der Deutschen Gesellschaft für Psychologie e.V. (DGPs).
- KERSTING, M. (2005). Beratung und Auswahl von Studienbewerbern: Ziele und Methoden. *Psychologische Rundschau*, 56, 149-150.
- KERSTING, M. & GEHRING, F. (2006). University citizenship behavior. *Vortrag auf dem Psychologienkongress 2006 der Deutschen Gesellschaft für Psychologie e.V.*
- KLAGES, L. (1926). *Die Grundlagen der Charakterkunde*. Leipzig: Barth.
- KLEINMANN, M. (1997). *Assessment Center: Stand der Forschung - Konsequenzen für die Praxis*. Göttingen: Verlag für Angewandte Psychologie.

- KLEINMANN, M. (1993). Are assessment center rating dimensions transparent for participants? Consequences for criterion and construct validity. *Journal of Applied Psychology*, 78, 988-993.
- KLIMOSKI, R. J. & STRICKLAND, W. J. (1977). Assessment Centers: Valid or merely precient. *Personnel Psychology*, 30, 353-363.
- KLIMOSKI, R. J. & BRICKNER, M. (1987). Why do assessment centers work? The puzzle of Assessment-Center validity. *Personnel Psychology*, 40, 243-260.
- KNEBEL, H. (1995). Ungereimtes aus der Führungspraxis. *Personal*, 47, (5), 348-353.
- KOBASA, S. C., MADDI, S. R. & KAHN, S. (1982). Hardiness and health: a prospective study. *Journal of Personality and Social Psychology*, 42 (1), 168-177.
- KOBRIN, J. L. & SCHMIDT, A. E. (2005). *The research behind the new SAT*. New York: The College Entrance Examination Board.
- KOHLMANN, C. W. (1990). *Stressbewältigung und Persönlichkeit: flexibles versus rigides Copingverhalten und seine Auswirkungen auf Angsterleben und physiologische Belastungsreaktionen*. Dissertation, Johannes Gutenberg-Universität Mainz, Fachbereich 12 Sozialwissenschaften. Bern: Huber.
- KÖLLER, O. & BAUMERT, J. (2002). Das Abitur – immer noch ein gültiger Indikator für die Studierfähigkeit? *Aus Politik und Zeitgeschichte Band 6*.
- KOMPA, A. (1989). *Assessment Center: Bestandaufnahme und Kritik*. (3. Auflage). München: Hampp.
- KONEGEN-GRENIER, C. (2001). *Studierfähigkeit und Hochschulzugang*. Köln: (gedruckt. Monographie).
- KONRADT, U. & SARGES, W. (HRSG.). (2003). *E-Recruitment und E-Assessment. Rekrutierung, Auswahl und Entwicklung von Personal im Intra- und Internet*. Göttingen: Hogrefe.
- KRAMPEN, G. (1991). *FKK-Fragebogen zu Kompetenz- und Kontrollüberzeugungen*. Handanweisung. Göttingen: Hogrefe.
- KRAMPEN, G. (1981). *IPC-Fragebogen zu Kontrollüberzeugungen*. Handanweisung. Göttingen: Hogrefe.
- KREIS WISSENSCHAFTLICHER BERATER. (1977). *Empfehlungen zur Frage der Verwendung psychologischer Tests im Rahmen einer Neuregelung des Hochschulzuges*. Saarbrücken.

- KREX, L. & STORM, G. (2005). Studienerfolg an den Universitäten der Bundeswehr. Unveröffentlichter Ergebnisbericht. Bonn: BMVg, PSZ III 6.
- KUBINGER, K. D. (2005). Psychologische Diagnostik der Eignung für das Psychologiestudium. *Psychologische Rundschau*, 56 (2), 133-135.
- KUBINGER, K. D. & PROYER, R. (2004). Gütekriterien. In K. Westhoff, L. J. Hellfritsch, L. F. Hornke, K. D. Kubinger, F. Lang, H. Moosbrugger, A. Püschel und G. Reimann. (Hrsg.), *Grundwissen für die berufsbezogene Eignungsbeurteilung nach DIN 33430*. S. 186-194. Lengerich: Pabst.
- KUBINGER, K.D. & EBENHÖH, J. (1996). *Manual Arbeitshaltungen*. Frankfurt: Swets Test Services.
- KUNCCEL, N. R., HEZLETT, S.A. & ONES, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, 86, 148-161.
- LANDECK, K-J. (1987). *Studienverhalten und Studienerfolg an der Uni Bw*.
- LATTMANN, C. (HRSG.). (1989). *Das Assessment-Center-Verfahren der Eignungsbeurteilung. Sein Aufbau, seine Anwendung und sein Aussagegehalt*. Heidelberg: Physika Verlag.
- LAZARUS, R. S. & FOLKMANN, S. (1984). *Stress, Appraisal and Coping*. New York: Springer.
- LAZARUS, R. S. & LAUNIER, R. (1978). Stress related transactions between person and environment. In L. A. Pervin & M. Lewis (Eds.), *Perspectives in international psychology* (pp. 287-327). New York: Plenum.
- LEWIN, K., HEUBLEIN, U., SOMMER, D. & CORDIER, H. (1995). Studienabbruch – Gründe und anschließende Tätigkeiten. Ergebnisse einer bundesweiten Befragung im Studienjahr 1993/94. In: *Hochschul-Informationen-System A1/95*. Hannover.
- LIENERT, G. A. (1969). *Testaufbau und Testanalyse* (3., erg. Aufl.). Weinheim: Beltz.
- LIENERT, G. A. & RAATZ, U. (1989). *Testaufbau und Testanalyse*. München: Psychologie Verlags Union.
- LISSMANN, U. (1976). *Prognostische Möglichkeiten von Einzelnoten und ihrer Gewichtung im Abiturzeugnis*. Landau, Zentrum für empirische pädagogische Forschung.
- MADDI, S. R. (2001). Hardi-Survey-III-R On-line. Available: <http://www.Hardinessinstitute.com>

- MADDI, S. R. (1999). Comments and Trends in Hardiness Research and Theorizing. *Consulting Psychology Journal*, 51 (2), 67-71.
- MADDI, S. R., KOBASA, S. C. & HOOVER, M. (1979). An Alienation Test. *Journal of Humanistic Psychology*, 19 (4), 73-76.
- MCCRAE R. R. & COSTA, P. T., (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (pp. 139-153). New York: Guilford Press.
- 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*, 4, 18-35.
- MELTER, A. (1977). *Persönlichkeitsmerkmale und Leistungen in einem Hochschuleingangstest. Eine differentialpsychologische Untersuchung im Rahmen der Erprobung eines Deutschen Hochschuleingangstests für das Studienfeld Medizin*. Dissertation. Mannheim.
- MICHEL, L. (1977). *Hochschuleignungstest für das Studienfeld Medizin*. Bundesministerium für Bildung und Wissenschaft (Hrsg.). Schriftenreihe Hochschule Nr. 25. Bonn.
- MICKO, H. C. & NIPPERT, K. (1979). *Entwicklung eines aktuellen Struktur- und Ablaufplans für die wissenschaftliche Untersuchung des Studienabbruchs und Fachwechsels unter den Bedingungen des Modellversuchs Studienberatung, Niedersachsen*. Hannover.
- MOUW, J. T. & KHANNA, R. K. (1993). Prediction of academic success: A review of the literature and some recommendations. *College Student Journal*, 27, 328-336.
- MÜLLER, F. H. (2001). *Bedingungen und Auswirkungen des Studieninteresses. Eine Studentenbefragung an der Universität der Bundeswehr München*. Hochschulkurier Nr. 12 / Oktober 2001. Universität der Bundeswehr München.
- MÜLLER-BÖLING, D. (2001). *Für eine nachfrageorientierte Steuerung des Studienangebots an Hochschulen – Vorschläge zur Ablösung der Kapazitätsverordnung (Ms.)*. Berlin.
- NEIDIG, R. F., MARTIN, J. C. & YATES, R. E. (1978). *The FBI's Management Aptitude Program assessment center: Research report no. 1 (TM-78-3)*. Washington, DC: U.S. Civil Service Commission. (NTIS No. PB-286 181).
- NEIDIG, R. F. UND NEIDIG, P. J. (1982). Multiple Assessment-Center exercises and job relatedness. *Journal of Applied Psychology*, 69 (1), 182-86.
- NEISSER, U. (1976). *Cognition and Reality*. San Francisco: Freeman.

- NORMAN, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, 66, 574-583.
- NOWACK, K. M. (1989). Coping Style, Cognitive Hardiness, and Health Status. *Journal of Behavioral Medicine*, 12 (2), 145-158.
- OBERMANN, C. (2002). *Assessment Center. Entwicklung, Durchführung, Trends*. (2. Überarb. Aufl.). Wiesbaden: Gabler.
- OSTENDORF, F. (1990). *Sprache und Persönlichkeitsstruktur: Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit*. Regensburg: S. Roderer-Verlag.
- OSWALD, F. L., SCHMITT, N., KIM, B. H., RAMSAY, L. J. & GILLESPIE, M. A. (2004). Developing a biodata measure and situational judgement inventory as predictors of college student performance. *Journal of Applied Psychology*, 89, 187-207.
- PARSONS, F. (1909). *Choosing a vocation*. Boston: Houghton Mifflin
- PREISER, G. (1975). Zur „Objektivität“ mündlicher Prüfungen im Fach Psychologie. *Psychologische Rundschau*, 26, 256-281.
- PURSELL, E. D., CAMPION, M. A. & GAYLORD, S. R. (1980). Structured interviewing: Avoiding section problems. *Personnel Journal*, 11, 907-012.
- PUZICHA, K. J. (1997). Fachkonzept der Personalpsychologischen Qualitätssicherung. In: Arbeitspsychologischer Dienst der Bundeswehr (Hrsg.), *Neue Wege in der Personalpsychologie*. Bonn: BMVg – P II 4. 1, 7-44.
- REIMANN, G. (2004). Arbeits- und Anforderungsanalyse. In K. Westhoff, L. Hellfritsch, L. F. Hornke, K. D. Kubinger, F. Lang, H. Moosbrugger, A. Püschel, G. Reimann (Hrsg.). *Grundwissen für die berufsbezogene Eignungsdiagnostik nach DIN 33430*. Lengerich: Pabst.
- REISSERT, R. & MARCISZEWSKI, B. (1987). *Studienverlauf und Berufseintritt – Ergebnisse einer Befragung von Hochschulabsolventen und Studienabbrechern des Studienjahres 1984*. HIS-Hochschulplanung 61. Hannover.
- RHEINBERG, F. (2000). *Motivation* (3. Auflage). Stuttgart: Kohlhammer. (zur Anschaffung empfohlen).
- RINDERMANN, H. (2000). Für ein bundesweites Auswahlverfahren von Studienanfängern über Fähigkeitsmessung. *Psychologische Rundschau*, 56 (2), 127-129.

- RINDERMANN, H. & OUBAID, V. (1999). Auswahl von Studienanfängern durch Universitäten – Kriterien, Verfahren und Prognostizierbarkeit des Studienerfolgs. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 20 (3), 1999, 172-191.
- Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R. & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, 130, 261-288.
- ROLFS, H. (2001). *Berufliche Interessen*. Göttingen: Hogrefe
- ROLFS, H. & SCHULER, H. (2000). *Berufliche Interessenskongruenz und das Erleben im Studium*. *Zeitschrift für Arbeits- und Organisationspsychologie*, 46, 3, 137-149.
- ROSENSTIEL, VON L. (1980). *Grundlagen der Organisationspsychologie*. Stuttgart: Poeschel.
- ROST, J. (1996). *Lehrbuch Testtheorie, Testkonstruktion*. Bern: Huber.
- ROTHER, C. (1981). *Abiturientenberatung und weiterer Bildungsweg*. Hannover, HIS-GmbH.
- ROTTER, J. B. (1975). Some problems and misconceptions related to construct of internal versus external control of reinforcement. *Journal of consulting clinical Psychology*, 43, 56-67.
- ROTTER, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 80 (1), Whole No. 609.
- RYAN, R. M. & LAGUARDIA, J. G. (1999). Achievement motivation within a pressured society: Intrinsic and extrinsic motivations to learn and the politics of school reform. In T. Urland (Ed.). *Advances in motivation and achievement*. (Vol. 11, pp. 45-85). Greenwich, CT: JAI Press.
- SAENGER-CEHA, M. M. TH. (1970). *Psychological and social factors in Student drop-out*, Amsterdam 1970.
- SAGIE, A. & MAGNEZY, R. (1997). Assessor type, number of distinguishable dimension categories and Assessment-Center constructs validity. *Journal of Occupational and Organisational Psychology*, 70, 103-108.
- SCHARNHORST, G. VON. In Heinz G. Nitschke Die Preussischen Militärreformen 1807-1813. Die Tätigkeit der Militärreorganisationskommission und ihre Auswirkungen auf die preussische Armee.
- SCHIEFELE, U. & URHAHNE, D. (2000). Motivationale und volitionale Bedingungen der Studienleistung. In U. Schiefele & K.-P. Wild (Hrsg.), *Interesse und Lernmotivation* (183-206). Münster: Waxmann.

- SCHIEFELE, U., KRAPP, A. WILD, & WINTELER, A. (1993). Der Fragebogen zum Studieninteresse (FSI). *Diagnostica*, 39 (4), 335-351.
- SCHIEFELE, U., STREBLOW, L., ERMGASSEN, U. & MOSCHNER, B. (2003). *Lernmotivation und Lernstrategien als Bedingungen der Studienleistung*.
- SCHIEFELE, U., KRAPP, A. & SCHREYER, I. (1993). Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25 (2), 120-148.
- SCHMIDT, F. L., LAW, K., HUNTER, J. E., ROTHESTEIN, H. R., PEARLMAN, K. & MCDANIEL, M. (1993). Refinements in validity generalization methods: Implications for the situational specificity hypothesis. *Journal of Applied Psychology*, 78, 3-12.
- SCHMIDT-ATZERT, L. (2005). Prädiktion von Studienerfolg bei Psychologiestudenten. *Psychologische Rundschau*, 56, 131-133.
- SCHNEIDER, B. (1987). Vorbereitung auf Intelligenztests: Eine Gefahr für die Eignungsdiagnostik? In R. Horn, K. Ingenkamp & R. S. Jäger (Hrsg.), *Tests und Trends 6* (Jahrbuch der pädagogischen Diagnostik, 3-25. München: Psychologie-Verlags-Union.
- SCHOLZ, G. (1994). *Das Assessment Center. Konstruktvalidität und Dynamisierung*. Stuttgart: Verlag für Angewandte Psychologie.
- SCHOLZ, G. & SCHULER, H. (1993). Das nomologische Netzwerk des Assessment Centers: Eine Metaanalyse. *Zeitschrift für Arbeits- und Organisationspsychologie*, 27, 33-44.
- SCHÖNEMANN, H. (1997). Some new results on hit rates and base rates in mental testing. *Chinese Journal of Psychology*, 39 (2), 137-192.
- SCHÖNEMANN, P. H. & THOMPSON, W. W. (1996). Hit-rate bias in mental testing. *Cahiers de Psychologie Cognitive-Current Psychology of Cognition*, 15 (1), 3-28.
- SCHULER, H. (2003). *Eignungsdiagnostische Methoden zur Auswahl von Studierenden durch die Hochschulen*. Vortrag auf der Gebit-Tagung m 13.03.2003 in Hohenheim (nicht veröffentlicht).
- SCHULER, H. (1995). *Lehrbuch Organisationspsychologie* (2., korrigierte Aufl.). Bern: Verlag Heinz Huber.
- SCHULER, H. (1998). Interviews. In S. Greif, H. Holling & N. Nicholson (Hrsg.), *Arbeits- und Organisationspsychologie* (260-265). München: Psychologie-Verlags-Union.
- SCHULER, H. & FUNKE, U. (1989). Berufseignungsdiagnostik. In E. Roth (Hrsg.), *Enzyklopädie der Psychologie. Themenbereich D, Serie III, Band 3: Organisationspsychologie* (S. 281-320). Göttingen: Hogrefe.

- SCHULER, H. (HRSG.). (2001). *Lehrbuch der Personalpsychologie*. Göttingen: Hogrefe.
- SCHÜPBACH H & KLAUER CH. (2006). Studierendenauswahl: Beiträge der Psychologie zu einem universitären Pilotprojekt. *Psychologische Rundschau*, 56 (2), 137-138
- SCHÜPBACH, H., KRAUSE, A. & PHILIPP, A. (2006). *Emotionsregulation bei Lehrkräften und Schulentwicklung*. Vortrag auf der 2. Fachtagung Psychologie der Tätigkeit in Dresden.
- SCHÜPBACH, H., PIXNER, J., & ZAPF, S. (2006). *Handlungskompetenz im Hochschulstudium. Bildung und Erziehung*, 59 (2), 147-166.
- SCHWARZER, R. (1994). Optimistische Kompetenzerwartung: Zur Erfassung einer personellen Bewältigungsressource. *Diagnostica*, 40, 105-123.
- SCHWARZER, R. (HRSG.). (1990). *Gesundheitspsychologie. Ein Lehrbuch*. Göttingen: Hogrefe.
- SKAGER, R. W., KLEIN, S. P., SCHULTZ, C. B. (1967). The Prediction of Academic and Artistic Achievement at a School Of Design. *Journal of Educational Measurement* 4 (2), 105–117.
- SPADY, W. G. (1970). Dropouts from Higher Education: An Interdisciplinary Review and Synthesis, *Interchange*, 1 (1), 64-85.
- SPIES, K., WESTERMANN, R., HEISE, E. & HAGEN, M. (1998). Zur Abhängigkeit der Studienzufriedenheit von Diskrepanzen zwischen Fähigkeiten und Anforderungen. *Psychologie in Erziehung und Unterricht*, 45, (1) 36-52.
- SPOKANE, A. R. (1985). *A review of research on person-environment congruence in Holland's theory of careers*. *Journal of Vocational Behavior*, 26, 306-343.
- STEBLER, P. (2000). *Studienerfolg und Studienzufriedenheit*. Dissertation: Universität Freiburg.
- STEYER, R, YOUSFI, S. & WÜRFEL, K. (2005). Prädiktion von Studienerfolg: Der Zusammenhang zwischen Schul- und Studiennoten im Diplomstudiengang Psychologie. *Psychologische Rundschau*, 56, 129-131.
- STORM, E. G. (1998). *Adaptive Testverfahren – Hintergründe, Testökonomie und Akzeptanz. Vortrag und Handout auf dem Kongress für Arbeits-, Betriebs- und Organisationspsychologie der Sektion des Berufsverbandes Deutscher Psychologen*. Viernheim.
- STRÖHLEIN, G. (1983). *Bedingungen des Studienabbruchs. Eine Längsschnittuntersuchung bei Studenten ingenieurwissenschaftlicher Fakultäten*. Bern: Lang.
- SÜDDEUTSCHE.DE. (2006). *Studenten in der Krise – Vom Hörsaal auf die Psycho-Couch*. URL: /jobkarriere/berufsstudium/artikel/189/88101/article.html.

- SUPER, D. E., SHARISHEVSKY, R., MATLLIN, N. & JORDAAN, J. P. (1963). *Career Development: Self-Concept Theory*. New York: College Entrance Examination Board.
- THORNDIKE, R. L. (1947). *Research problems and techniques*. New York: Wiley.
- THORNTON, G. C. & BYHAM, W. C. (1982). *Assessment-Center and managerial performance*. New York: Academic Press.
- THURSTONE, L. L. (1938). Primary mental abilities. *Psychometric Monographs*, 1.
- TINTO, V. (1982). Defining dropout: A matter of perspective. In: E. Pascarella (Hg.): *Studying student attrition*. San Francisco.
- TINTO, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research* 45, 1975, 89-125.
- TRAVERS, R.M. & GRONLUND, N.E. (1950). The meaning of marks. *Journal of higher Education*, 21, 369-374.
- TROST, G. (1975). *Vorhersage des Studienerfolgs*. Braunschweig, Westermann.
- TROST, G. (1984). Was ist Eignung zum Mathematikstudium? In Reiß, M. & Steiner, H.-G. (Hrsg.), *Mathematik-Kenntnisse - Leistungsmessung -Studierfähigkeit*. Köln: Audis Verlag Deubner.
- TROST, G. (1985). Pädagogische Diagnostik beim Hochschulzugang, dargestellt am Beispiel der Zulassung zu medizinischen Studiengängen. In R. S. Jäger, R. Horn und K.-H. Ingenkamp (Hrsg.), *Test und Trends 4*. Weinheim, Beltz, 41 - 83.
- TROST, G. (2003). *Deutsche und internationale Studierfähigkeitstests. Arten, Brauchbarkeit, Handhabung. Dokumentationen & Materialien*, Band 51. Bonn: Deutscher Akademischer Austauschdienst.
- TROST, G. & HAASE, K. (2005) *Hochschulzulassung: Auswahlmodelle für die Zukunft. Eine Entscheidungshilfe für die Hochschulen*. Schriftenreihe der Landesstiftung Baden Württemberg / Stifterverband für die Deutsche Wissenschaft.
- TROST, G. & VON HAYN, S. (2001). *Auswahlgespräche mit Studienbewerbern. Handreichung für die Hochschulen*. Bonn: ITB GmbH.
- TROST, G. & BICKEL, H. (1979). *Studierfähigkeit und Studienerfolg*. München.
- TROST, G. KLIEME, E. & NAUELS, H.-U. (1997). Prognostische Validität des Tests für medizinische Studiengänge (TMS). In T. Herrmann (Hrsg.), *Hochschulentwicklung - Aufgaben und Chancen* (57-78). Heidelberg: Asanger.
- VERNON, P. E. (1965). *Ability factors and environmental influences*. *American Psychologist*, 20, 723-733.

- WAGEMANN, C.-H. (1987). Die Schnellen und die Superschnellen. *Hochschulausbildung*, 5 (2), 115-122.
- WEINGARDT, E. (1989). Untersuchungen über Korrelationen zwischen Reifeprüfungsnoten und Erfolg auf der Universität. In K. Ingenkamp (Hrsg.), *Die Fragwürdigkeit der Zensurengebung* (S. 252-255). Weinheim: Beltz.
- WESTERMANN, R., HEISE, E., SPIES, K. & TRAUTWEIN, U. (1996). Identifikation und Erfassung von Komponenten der Studienzufriedenheit. *Psychologie in Erziehung und Unterricht*, 43, 1-22.
- WHETTON, C., McDONALD, A. & NEWTON, P. (2001). Aptitude testing for university entrance. *Paper presented at 27<sup>th</sup> Annual Conference of the International Association for educational Assessment*, Rio de Janeiro, Brazil 6-11 May 2001.
- WIDEN, H. & CARLSTEDT, L. (1996). Officer selection in the Swedish armed forces. In Centro de Psicologia Aplicada do Exército, Lisboa (Hrsg.), *Proceedings of the 32nd International Applied Military Psychology symposium*, (p. 47-59). Lisboa.
- WIDMAIER, H.P., JERMANN, M. & REICHHOLD, F. (1969). *Analysen im Hochschulbereich. Bildung in neuer Sicht*. Reihe A, Nr. 19.
- WILCKE, B. A. (1976). *Studienmotivation und Studienverhalten*. Göttingen.
- WINTELER, A. & SISTERWALD, W. (1987). Entwicklung und Überprüfung eines Fragebogens zum Studieninteresse (FSI). *Hochschulausbildung*, 5, 223-242.
- WINTELER, A., SIERWALD, W. & SCHIEFELE, U. (1988). Interesse, Leistung und Wissen: Die Erfassung von Studieninteresse und seine Bedeutung für Studienleistung und fachbezogenes Wissen. *Empirische Pädagogik*, 2, 227-250.
- WINTELER, A., SCHIEFELE, U. & KRAPP, A. (1986). Studieninteresse und fachbezogene Wissensstruktur. *Psychologie in Erziehung und Unterricht*, 35, 106-118.
- WISSENSCHAFTSRAT. (2004). *Empfehlungen zur Reform des Hochschulzugangs*. Köln: Wissenschaftsrat.
- WOLTER, A. (2005). Profilbildung und universitäre Weiterbildung. In: W. Jütte & K. Weber (Hrsg.): *Kontexte wissenschaftlicher Weiterbildung. Entstehung und Dynamik von Weiterbildung im universitären Raum* (S. 93-111). Münster, New York, München, Berlin: Waxmann.
- WOTTAWA, H. & HOSSIEP, R. (1987). *Grundlagen psychologischer Diagnostik*. Göttingen: Hogrefe.

- WOTTAWA, H. & THIERAU, H. (1998). *Lehrbuch Evaluation. (2. vollst. überarb. Aufl.)*. Bern: Huber.
- ZIMMERHOFER, A. (2003). *Neukonstruktion und erste Erprobung eines webbasierten Self-Assessments zur Feststellung der Studieneignung für die Fächer Elektrotechnik, Technische Informatik sowie Informatik an der RWTH Aachen*. Diplomarbeit. RWTH Aachen.
- ZIMMERHOFER, A., HEUKAMP, V. K. & HORNKE, L. F. (2006). Ein Schritt zur fundierten Studienfachwahl - Webbasierte Self-Assessments in der Praxis. *Report Psychologie, 31* (2). 62-72.
- ZUCKERMAN, M. (1979). *Sensation seeking: Beyond the optimal level of arousal*. Hillsdale: N.J.: Erlbaum.

H ANHANG

## Anhang A: Zeitlicher Ablauf des Eignungsfeststellungsverfahrens der OPZ

### 1. Prüftag

Prüfgruppe Uhrzeit	1	2	3	4	5	6	7	8	9
6.20	Begrüßung / Einweisung durch Persönlicher Berater Aufsatz								
7.00	CAT				ARZT	GSV			
8.00									
9.00	INT	PersBer	PersBer	PersBer	CAT	Arzt	AHA	MAT	CAT
10.00	PersBer	INT	AHA	Arzt			PersBer		
11.00	AHA	AHA	INT		AHA	PersBer	MAT	StudBer	PersBer
12.00	Mittagspause								
13.00	MAT	MAT	Arzt	INT	PersBer	MAT	StudBer	Arzt	MAT
				PersBer	INT				
14.00	StudBer	StudBer	MAT	PersBer	AHA	PersBer	PersBer	INT	
15.00					PersBer				
<i>Anmerkungen:</i> CAT = Computer-Assisted-Testing, GSV = Gruppensituationsverfahren, INT = ADAPT-Intelligenztest, PersBer = Personalberatung, AHA = Arbeitshaltungstest, StudBer = Studienberatung; MAT = Mathematischer Kenntnistest.									

## 2. Prüftag

Prüfgruppe Uhrzeit	1	2	3	4	5	6	7	8	9
7.00	<b>GSV</b>					<b>StudBer</b>	<b>CAT</b>		<b>Arzt</b>
8.00									
9.00									
10.00	<b>Arzt</b>	<b>MAT</b>	<b>StudBer</b>	<b>MAT</b>	<b>CAT</b>	<b>INT</b>	<b>PersBer</b>	<b>StudBer</b>	
11.00		<b>StudBer</b>		<b>AHA</b>		<b>StudBer</b>	<b>INT</b>		<b>Arzt</b>
12.00		<b>Mittagspause</b>							
12.30	Beobachterkonferenz / Aktenvervollständigung der Prüfgruppen Ab 12:45 <b>PHYSICAL FITNESS TEST</b>  Anschließend: Ergebniseröffnung								
15.00	Informationsvortrag für Sanitätsoffizierbewerber								
<i>Anmerkungen:</i> CAT = Computer-Assisted-Testing, GSV = Gruppensituationsverfahren, INT = ADAPT-Intelligenztest, PersBer = Personalberatung, AHA = Arbeitshaltungstest, StudBer = Studienberatung; MAT = Mathematischer Kenntnistest.									

## 3. Prüftag

8.00 - 8.45	Einführungsvortrag des Dezernatsleiters Einplanung/Einstellung zum weiteren Ablauf des Annahmeverfahrens
8.45 - 12.00	Einplanungsvorträge der Sachgebietsleiter Anschließend: Einplanung
12.00	<b>Mittagspause</b>
12.30 - 15.30	Einplanung

## Anhang B: Biographischer Fragebogen

### **SEHR GEEHRTE TEILNEHMERIN, SEHR GEEHRTER TEILNEHMER!**

Vielen Dank für Ihre Bereitschaft, an der Entwicklung neuer Studienberatungstests für die Bundeswehr im Auftrag des BMVg - PSZ III 6 - als Proband teilzunehmen. Wir möchten Sie bitten, den nachfolgenden Fragebogen vollständig zu bearbeiten. Bitte füllen Sie die Felder in DRUCKBUCHSTABEN aus und kreuzen Sie die zutreffende Antwort an.

Das Dezernat Personalpsychologie der Gruppe Wehrpsychologie gehört zu den wenigen Dienststellen der Bundeswehr, die „besonders schützenswerte Daten“ unter strengsten Sicherheitsauflagen für anonymisierte Qualitätssicherung verwenden darf. Dies wird im Rahmen des IT-Sicherheitskonzeptes und in § 9 des Bundesdatenschutzgesetzes sowie dazu erlassenen Durchführungsbestimmungen für den PsychDstBw erschöpfend geregelt.

Für die Zusammenführung erhobener und schon vorhandener Testdaten (zum Beispiel aus der OPZ) benötigen wir zunächst Schlüsselvariablen wie Ihren Namen und Ihre Personenkennziffer. Diese werden unmittelbar nach Abgleich der verschiedenen Quellen unwiderruflich vernichtet und gelöscht. Der Datensatz ist ab diesem Zeitpunkt irreversibel anonymisiert und kann infolgedessen nicht mehr einzelnen Personen zugeordnet werden!

 **Die Angabe Ihrer persönlichen Daten hat demnach für Sie keinerlei Auswirkung und findet keine weitere Verwendung!**

Natürlich teilen wir Ihnen persönlich gerne Ihr Testergebnis mit. Wenden Sie sich dazu bitte vertrauensvoll an den/die anwesende/n Diplom-Psychologen/in. Eine gute Möglichkeit bietet das Feedbackgespräch im Anschluss der Testung.

Es wird nun um Ihre schriftliche Einwilligung gebeten, dass Sie freiwillig an der oben beschriebenen Untersuchung teilnehmen und mit der Verarbeitung der Daten einverstanden sind.

**Hiermit gebe ich meine Einwilligung zur Nutzung meiner personenbezogenen Daten.**

**Datum:** \_\_\_\_\_ **Unterschrift:** \_\_\_\_\_

Teilstreitkräfte:

 Heer Luftwaffe Marine

Schulbildung:  Fachhochschulreife  allgemeine Hochschulreife  sonstiges .....

Abiturnote: \_\_\_\_\_ (mit einer Dezimalstelle)

Hochschule:  UniBwMünchen

 FH München UniBwHamburg

Studentenjahrgang: \_\_\_\_\_

Studienfeld:  Geisteswissenschaften (PÄD; SPO; SOWI; POL; GESCH)

Ingenieurwissenschaften (MB; LRT; ET; GEO; BAUV; RI; ETTI)

Wirtschaftswissenschaften (WIW; BWL; BW; WOW; VWL; WI)

Informatik (INF; EIT)

Studienfach:

1  Bauingenieurwesen & Umwelttechnik

12  Staats- & Sozialwissenschaften

2  BWL

13  Geodäsie & Geoinformation (Vermessungstechnik)

3  Elektrotechnik

14  VWL

4  Geschichtswissenschaften

15  Wirtschaftsingenieurwesen

5  Informatik

16  Wirtschafts- & Organisationswissenschaften

6  Luft- & Raumfahrttechnik

17  Wirtschaftsinformatik

7  Maschinenbau

18  Elektro- & Informationstechnik

8  Pädagogik

19  Maschinenbau (FH)

9  Politikwissenschaften

20  Betriebswirtschaft (FH)

10  Rechnergestützte Ingenieurwissenschaften

21  Elektrotechnik & Technische Informatik (FH)

11  Sportwissenschaften

Falls Ihr Studienfach von Ihrem ersten Studienwunsch der OPZ-Bewerbung abweicht, geben Sie bitte zusätzlich Ihren damaligen "Wunschstudiengang" an. Tragen Sie dazu die entsprechende Nummer (siehe oben) in das Feld ein.

Studienwechsel:  ja wann: \_\_\_\_\_

nein

Studienfach vor Wechsel:  
(siehe Nummerierung oben)

Studienabbruch:  ja  nein

Note: \_\_\_\_\_ Vordiplom: \_\_\_\_\_

Hauptdiplom: \_\_\_\_\_

Gewünschter Einstellungs-  
termin (GEWET) in:  
Einstellungstag Offizierausbildung \_\_\_\_\_

**Falls Sie sich bei den folgenden Fragen nicht sicher sind, lassen Sie bitte die Felder frei!**

Abschlussnote Offizierprüfung: \_\_\_\_\_

Prüfungsart/-bezeichnung: \_\_\_\_\_

OPZ-Gesamteignungsgrad: A (Bewährung ist sehr wahrscheinlich.) M (Bewährung ist wahrscheinlich.)  
X (Bewährung ist noch wahrscheinlich.) U (Bewährung ist unwahrscheinlich.)

Beurteilung der Studien- 1/A (Empfohlen.) 2/B (Mit Einschränkung empfohlen.)  
eignung: 3/C (Nicht empfohlen.)



## BEFRAGUNG ZUR AKZEPTANZ DES TESTVERFAHRENS

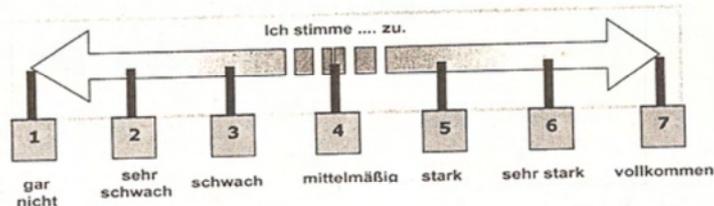
Sehr geehrte Testperson,  
 der psychologische Dienst der Bundeswehr möchte die Qualität seiner Arbeit kontinuierlich verbessern. Damit wir diesem Anspruch gerecht werden können, haben wir einen Feedbackbogen konzipiert, der uns zielgerichtete Anregungen zu Themen wie Durchführung & Organisation, Qualität und Inhalt der durchgeführten Pilotphase geben soll.

Wir würden uns freuen, wenn Sie an den Verbesserungsbestrebungen der Gruppe Wehrpsychologie Dezernat Personalpsychologie konstruktiv mitwirken. Daher möchten wir Sie bitten, den anliegenden Fragebogen im Anschluss an die Testung bzw. nach Ihrem Feedbackgespräch auszufüllen. Sie benötigen dazu lediglich fünf Minuten. Haben Sie auch hier keine Bedenken, Ihre Meinung offen und ehrlich wiederzugeben. Wir sichern Ihnen selbstverständlich eine anonyme und vertrauliche Durchführung und Auswertung dieser Befragung zu. Daher sollten Sie den ausgefüllten Fragebogen in die Box neben der Tür einwerfen.

Durch Ihre Teilnahme an dieser Befragung tragen Sie maßgeblich zum Erfolg unserer Arbeit bei!

Bitte behalten Sie bei der Beantwortung der Fragen im Hinterkopf, dass die spätere Zielgruppe Offizierbewerber d.h. also i.d.R. Abiturienten sind. Bitte versetzen Sie sich in diese Situation zurück und geben Sie uns Ihre Einschätzung an.

Benutzen Sie dafür die Ihnen nun schon bekannte Skala von *stimme gar nicht zu* (1) bis *stimme vollkommen zu* (7).



Durchführung und Organisation der Pilotphase	Ihre Bewertung
Ich wurde über den Hintergrund der Testung ausreichend informiert.	<input type="checkbox"/>
Ich wurde über den Ablauf der Testung ausreichend informiert.	<input type="checkbox"/>
Bei Fragen bestand stets die Möglichkeit, sich an Fachpersonal zu wenden.	<input type="checkbox"/>
Das Projektpersonal begegnete mir stets freundlich.	<input type="checkbox"/>
Die Testung am Laptop war ein ungewohntes und unangenehmes Gefühl.	<input type="checkbox"/>
Die Testung am Laptop ist einfach und anwenderfreundlich.	<input type="checkbox"/>
Wenn ich die Wahl hätte, zöge Ich Papier-Bleistift-Tests einer Computertestung vor.	<input type="checkbox"/>
Meine Potenziale wurden durch die Testung treffend ermittelt.	<input type="checkbox"/>
Die Darbietung am Computer wirkt gegenüber Papier-Bleistift-Verfahren ansprechender.	<input type="checkbox"/>
Das Feedbackgespräch empfand ich als aussagekräftig.	<input type="checkbox"/>
Die Dauer des Feedbackgesprächs hat meinem Informationsbedürfnis entsprochen.	<input type="checkbox"/>
Während des Feedbackgesprächs konnte ich Fragen meinerseits klären.	<input type="checkbox"/>
Ich profitiere von der Teilnahme im Hinblick auf die Zeit nach der Bundeswehr.	<input type="checkbox"/>
Ich denke, dass man mit diesem Testverfahren ein geeignetes Instrument zur Feststellung der Studieneignung entwickelt hat.	<input type="checkbox"/>
Die studienfeldspezifische Testbatterie empfinde ich als standardisierter und aussagekräftiger als die bisherige OPZ-Studieneignungsprüfung.	<input type="checkbox"/>

**Inhalt und Qualität der gesamten Testbatterie**

	<i>Test -Teil I (allgemein)</i>	<i>Test -Teil II (spezifisch)</i>	<i>Test -Teil III (Persönlichkeits- profil)</i>
Beim Lesen der Instruktion sind mir Unklarheiten aufgefallen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich hätte gerne zwischendurch zur Instruktion zurückgeblättert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich habe die Instruktionen immer schnell weitergeblättert, ohne Sie zu lesen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich habe mich gefragt, wozu man ein Beispiel zur Bedienung der Tastatur eingebaut hat.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mir war die Bedienungsinstruktion zu lang.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich musste mir die Bearbeitungsbeispiele mehrmals durchlesen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich fühlte mich stark unter Zeitdruck.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich hatte nach der Bearbeitung der Aufgaben noch Zeit.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich habe aus Zeitgründen nicht alle Aufgaben bearbeiten können.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich musste mich anstrengen, um die Schrift lesen zu können.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mich haben die Hintergrundbilder bzw. Grafiken abgelenkt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Texte und Grafiken waren für mich so angeordnet, dass ich alles gut lesen konnte.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Aufgaben zur Studieneignung empfand ich als praxisnah.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Aufgaben haben mich an wichtige Inhalte bzw. an verlangte Fähigkeiten meines Studiums erinnert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Fähigkeiten, die ich im Studium anwenden muss, werden mit den Aufgaben geprüft bzw. zumindest angeregt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Als Offizierbewerber hätte ich Probleme mit den Aufgaben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich denke mit diesen Aufgaben kann man eine Studieneignung für meinen Fachbereich prüfen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Nur bezogen auf den Teil III (Persönlichkeitsprofil)**

	<i>Ihre Bewertung</i>
Mir waren die Fragen zu persönlich.	<input type="checkbox"/>
Ich hätte gerne einige Fragen unbeantwortet gelassen.	<input type="checkbox"/>
Ich frage mich wer diese Fragen als Bewerber ehrlich beantworten würde.	<input type="checkbox"/>
Beim Bearbeiten der Fragen ist die Richtung, in die die Frage zielt, sehr offensichtlich.	<input type="checkbox"/>
Ich habe mir am Ende nur noch gewünscht, endlich fertig zu sein.	<input type="checkbox"/>
Fragen zur Persönlichkeit bilden grundlegende Eigenschaften für ein erfolgreiches Studieren ab.	<input type="checkbox"/>
Ich habe mich gefragt, was die Fragen mit meinem Studium zu tun haben.	<input type="checkbox"/>

**FÜR IHR ENGAGEMENT UND IHRE MITARBEIT DANKEN WIR IHNEN!**