

Zur Treffsicherheit der DGP-Prognosen in den neuen Ländern: Bewährungskontrolle, Vorhersage des Studienerfolgs im ersten Studienabschnitt an der Fachhochschule für öffentliche Verwaltung in einem neuen Bundesland

Martin Kersting

Um die Frage beantworten zu können, ob die in den alten Ländern entwickelten und erprobten *DGP*-Eignungsuntersuchungen auch in den neuen Ländern eine Prognose des Ausbildungserfolgs gewährleisten, wurden in einer ersten Bewährungsprobe die schriftlichen Noten (Klausurergebnisse) aus dem ersten Studienabschnitt an einer Fachhochschule für öffentliche Verwaltung (FHöV) den *DGP*-Empfehlungsgraden gegenübergestellt. Die Stichprobe besteht aus 40 Regierungsinspektoranwärterinnen und 38 -anwärtern (insgesamt 78 Personen) einer Verwaltung **aus den neuen Bundesländern**, die nicht näher genannt sein möchte. Aus Datenschutzgründen fehlen wichtige Informationen wie Altersangaben, Ergebnisse für einzelne *DGP*-Tests (z.B. Diktat) oder Schulnoten einzelner Fächer -was sich zuungunsten der "Analysetiefe" auswirken muß.

1. Kriterium

Vorrangiges Ziel und Anspruch der *DGP*-Eignungsuntersuchungen ist die Vorhersage des Ausbildungserfolgs. Die Verwaltungen sollen mit Hilfe der *DGP* davor bewahrt werden, Kandidatinnen und Kandidaten einzustellen, die den späteren Anforderungen der Ausbildung nicht genügen.

Als Kriterium für die Treffsicherheit der *DGP* Verfahren müssen daher zunächst die Leistungen in der Ausbildung gelten. Für die *DGP*-getesteten Anwärterinnen und Anwärter aus den neuen Ländern stehen zur Zeit erst die Leistungen aus dem ersten Studienabschnitt fest. Grundlage der vorliegenden Bewährungsprobe sind die Ergebnisse der Klausuren in sechs Prüfungsfächern ("Staar", "ALVR", "BGB", "ÖDR", "ÖBWL", "VWL"). Die Benotung erfolgt nach Punkten von "0" ("ungenügend) bis "15" ("sehr gut"). ((1) "sehr gut" (14-15 Punkte), (2) "gut" (11-13 Punkte), (3) "befriedigend" (8-10 Punkte) (4) "ausreichend" (5-7 Punkte), (5) "mangelhaft" (2-4 Punkte) und (6) "ungenügend" (0-1 Punkt)).

Obwohl die Vergleichbarkeit von Noten unterschiedlicher Fächer grundsätzlich in Frage gestellt werden muß, wird im folgenden der Einfachheit und der Übersichtlichkeit halber dennoch vom Mittelwert aller sechs Klausurnoten ausgegangen. Da von niemandem ein Mittelwert von "sehr gut" oder von "ungenügend" über alle sechs Fächer hinweg erreicht wird und auch die Durchschnittsnote "gut" nur äußerst selten (vier mal) vertreten ist, lassen sich de facto nur drei Kategorien unterscheiden, nämlich die guten Anwärter(innen), die im Durchschnitt über alle sechs Klausuren eine Durchschnittsnote von "befriedigend" oder "gut" erreichen, die "fragwürdigen" mit der Note "ausreichend" und die schlechten, die im Durchschnitt über alle sechs Klausuren nicht über ein "mangelhaft" hinauskommen. Die Tabelle 1 zeigt die absoluten und relativen Häufigkeiten, mit der diese Kategorien besetzt sind.

Notendurchschnitt über 6 Klausuren d. FHöV	Häufigkeit	in Prozent
"8 --- 15 Punkte" ("befriedigend" und "gut")	41	52,6 %
"5 --- 7 Punkte" ("ausreichend")	33	42,3 %
"0 --- 4 Punkte" ("mangelhaft")	4	5,1 %
Tab. 1: Notendurchschnitt FHöV: 3 Kategorien, Definition und Häufigkeiten	78	100 %

2. Prädiktoren

2.1 Empfehlungsgrade der *DGP*

Die zur Personalauswahl und Erfolgsprognose eingesetzten Eignungsuntersuchungen der *DGP* bestehen aus einem schriftlichen und einem mündlichen Teil (Test und Verhalten). Das Ergebnis der Leistungs- und der Verhaltensbeurteilung wird unabhängig voneinander bestimmt und in Form eines "Empfehlungsgrades" festgelegt. Anschließend wird aus diesen beiden Beurteilungen ein "Gesamtempfehlungsgrad" gebildet.

Die Leistungsbeurteilung basiert auf verschiedenen, standardisierten schriftlichen Anforderungen an das sprachliche, rechnerische und logische Denken sowie auf Überprüfungen der Rechtschreibkenntnisse, der Allgemeinbildung und der Effektivität des Arbeitsverhaltens (Konzentration und Sorgfalt). Die Verhaltensbeurteilungen beruhen auf Beobachtungen der Diplompsycholog(inn)en in zwei Stationen: Erstens einem Bewerbungsgespräch und zweitens einer Diskussion in einer Gruppe von vier bis sieben Bewerber(inne)n. Die *DGP*-Psycholog(inne)n sind speziell geschult und beurteilen unabhängig solche Verhaltensdimensionen, die sich in Anforderungsanalysen als berufsrelevant erwiesen haben (z.B. Kooperationsbereitschaft, Sachbezogenheit, sprachlicher Ausdruck, Berufsmotivation).

Insgesamt vergibt die *DGP* fünf verschiedene Empfehlungsgrade, von "5" "entspricht den Anforderungen" über drei Zwischenstufen bis hin zur "1" "entspricht nicht den Anforderungen".

Die Häufigkeit, mit der die einzelnen Empfehlungsgrade in der vorliegenden Stichprobe vertreten sind, veranschaulicht ein grundsätzliches Problem von Bewährungsproben. Fast 85 % der eingestellten Bewerber(innen) haben einen Empfehlungsgrad von "4" oder besser erzielt. Dies ist verständlich, schließlich stellen die Verwaltungen überwiegend nur erfolversprechende Kandidat(inn)en ein, und nur bei "5er" und "4er" besteht nach den Erfahrungen der *DGP* eine relativ sichere Aussicht, die Prüfungen erfolgreich absolvieren zu können. Die Einstellung von "3er" Kandidat(inn)en ist laut *DGP*-Empfehlungen mit einem deutlichen Risiko verbunden, hier ist ein Ausbildungserfolg weder prinzipiell auszuschließen noch ohne weiteres vorauszusetzen, man sollte diese Kandidat(innen) entsprechend nicht einstellen, wenn man sich über den Ausbildungserfolg ganz sicher sein will. Von der Einstellung von "1er" oder "2er" Kandidat(inn)en rät die *DGP* dringend ab. In der vorliegenden Stichprobe gab es nur eine Person mit dem Gesamtempfehlungsgrad "2" und niemanden mit einem Empfehlungsgrad von "1". Da die Verwaltung diesen Empfehlungen weitgehend nachgekommen ist, beschränkt sich die Bewährungsprobe über-

wiegend auf "gute" Teilnehmer(innen). In der Begrifflichkeit der Statistik spricht man in diesen Fällen von einer "abgeschnittenen" Verteilung und meint damit, daß ein Teil der Bewerber(inn)en nicht in die Analyse miteingeht: nämlich die meisten Kandidat(inn)en mit einem Ergebnis von "3" und schlechter. Da diese Personen die Ausbildung erst gar nicht begonnen haben, kann die *DGP* auch nicht nachweisen, daß diese Personen zu Recht von der Ausbildung ausgeschlossen wurden. Dies ist ein Grund dafür, daß die berichteten Zusammenhänge zwischen den Ergebnissen der *DGP*-Eignungsuntersuchungen und den schriftlichen Leistungen in den Klausuren den tatsächlich bestehenden Zusammenhang **unterschätzen**. Dies gilt, wie Althoff (1984) gezeigt hat, insbesondere deshalb, weil bei den "schlechteren" Empfehlungsgraden wahrscheinlich ein engerer Zusammenhang besteht als bei den "besseren" Empfehlungsgraden (kurvilinearer Zusammenhang).

Die nachstehende Tabelle erläutert noch einmal die Bedeutung der *DGP*-Eignungsgrade, aufgeführt sind zusätzlich die Häufigkeiten, mit der sich die ausgewählten Bewerber(innen) des Einstellungsjahrgangs 1991 auf die einzelnen Empfehlungsgrade verteilen. Die Empfehlungsgrade "3" und "2" wurden zusammengefasst, da -wie gesagt- nur eine Person mit dem Empfehlungsgrad "2" ausgestattet war.

Gesamtempfehlungsgrad	Häufigkeit	in Prozent
"5" ("entspricht den Anforderungen")	20	25,6 %
"4" ("entspricht weitgehend den Anforderungen")	46	58,9 %
"3" und einmal "2" ("Risiko")	12	15,5 %
Tab. 2: Gesamtempfehlungsgrad: Zusammenfassung zu drei Kategorien	78	100 %

2.1.1 Kreuztabelle

Einen Einblick in die Vorhersagegenauigkeit der *DGP*-Empfehlungsgrade erhält man, wenn man -wie hier in Tabelle 3- die Verteilung des Prädiktors (die *DGP*-Empfehlungsgrade) der Verteilung des Kriteriums (dem Notendurchschnitt des ersten Studienabschnitts an der FHöV) gegenüberstellt.

Für jede "Zelle" der Tabelle sind drei unterschiedliche Angaben verzeichnet, die am Beispiel der ersten Zelle ("*DGP*-Empfehlungsgrad "5" und dem Klausurdurchschnitt "befriedigend und gut") kurz erläutert werden sollen. Die Angabe in der ersten Reihe dieser Zelle bezieht sich auf die absolute Stichprobengröße von 78 Regierungsinspektorwärter(innen). Die erste Zahl ist eine absolute Nennung, die in Klammern aufgeführte Zahl der entsprechende prozentuale Wert. Also: *16 der insgesamt 78 Teilnehmer(innen) haben einen DGP- Empfehlungsgrad von "5" erzielt und in den Klausuren des ersten Studienabschnitts eine Durchschnittsnote von "befriedigend oder gut" erzielt. Das sind 20,5 % der Gesamtstichprobe von 78 Personen.* Die Zahl darunter, rechts neben der Klausurdurchschnittsnote bezieht sich nur auf die Personen, die in der Klausur eine Durchschnittsnote von "befriedigend oder besser" erzielt haben: *39 % derjenigen, die eine Durchschnittsnote von "befriedigend oder besser" erzielt haben, erreichen im*

a) Zusammenhang zwischen dem Notendurchschnitt FHöV und den DGP-Empfehlungsgraden

N = 78	DGP-Gesamtempfehlungsgrad		
	"5"	"4"	"3" oder "2"
Klausuren FHöV			
"befriedigend" und besser	16 (20,5 %)	21 (26,9 %)	4 (5,1 %)
	39 %	51,2 %	9,8 %
	80 %	45,7 %	33,3 %
"ausreichend"	4 (5,1 %)	22 (28,2 %)	7 (9,0 %)
	12,1 %	66,7 %	21,2 %
	20 %	47,8 %	58,3 %
"mangelhaft"		3 (3,8 %)	1 (1,3)
		75 %	25 %
		6,5 %	8,3 %
Tab. 3: Kreuztabelle	20 (100 %)	46 (100 %)	12 (100 %)

DGP-Eignungsverfahren einen Empfehlungsgrad von "5". Die letzte, grau-hinterlegte Angabe pro Zelle bezieht sich nur auf die Teilnehmer(innen) mit einem DGP-Ergebnis von "5": 80 % derjenigen, die im DGP-Verfahren den Empfehlungsgrad "5" erzielen, schließen die Klausuren der FHöV im Durchschnitt mit einer Note von "befriedigend" oder besser ab.

Um die Vorhersagegenauigkeit der Eignungsuntersuchungen der DGP zu bestimmen, ist natürlich diese zuletzt erläuterte Zahl in den grau-hinterlegten Feldern am interessantesten. Die einstellende Behörde will ja wissen, inwieweit sie sich darauf verlassen kann, daß ein(e) Bewerber(in), die von der DGP als "den Anforderungen entsprechend" klassifiziert wurde, auch tatsächlich den Anforderungen der Ausbildung genügt. Die Tabelle belegt die hohe Treffsicherheit der DGP -Empfehlungsgrade auch in den neuen Ländern: Von den Bewerber(innen), die von der DGP als "den Anforderungen entsprechend" (Empfehlungsgrad "5") beurteilt werden, gibt es kein(e)n, der den Anforderungen der schriftlichen Klausuren dann entgegen der Vorhersage nicht genügen würde. Alle diese Bewerber(innen) bestehen die Klausuren mit mindestens ausreichend, 80 % gehören erwartungsgemäß in die Gruppe der Jahrgangsbesten. Der Empfehlungsgrad "5" der DGP impliziert also quasi eine Garantie für ein erfolgreiches Abschneiden bei den schriftlichen Prüfungen. Eine solche Garantie gibt es natürlich für Bewerber(innen) mit "leichten Einschränkungen" (Empfehlungsgrad "4") nicht. Dennoch entsprechen die meisten Teilnehmer(innen) mit diesem Empfehlungsgrad noch den Anforderungen (93,5 %). Lediglich 6,5 % der Bewerber(innen) mit diesem Empfehlungsgrad erweisen sich erwartungswidrig in den Klausuren als "mangelhaft". Diese Personen bezeichnet man in der Statistik als "falsch-positiv", da sie von der DGP **fälschlicherweise** als **positive** Bewerber(innen) einklassifiziert wurden und in Wirklichkeit **negativ** abgeschnitten haben. Von dieser "falschen" Vorhersage sind drei der insgesamt 78 Personen (oder 3,8 %) betroffen. Bezieht man die Anzahl der drei Personen, die von einer

falschen Vorhersage betroffen sind, einmal auf die Gruppe, die von der *DGP* insgesamt empfohlen wurde (also auf die 20 "fünfer" und die 46 "vierer" Kandidat(inn)en), so kann man formulieren, daß die Fehlerquote der *DGP* in Bezug auf eine Auswahl der Erfolgreichen in den neuen Ländern bei drei von 66 oder bei 4,5 % liegt- ein beachtlich geringer Fehlerwert! Die Fehlerquote der *DGP* in Bezug auf die "Aussortierung" von Ausbildungsversagern kann anhand der vorliegenden Daten -wie bereits erwähnt wurde- nicht bestimmt werden, da die von der *DGP* als ungeeignet eingestuften Bewerber(innen) ja gar nicht erst an der Ausbildung teilgenommen haben, sondern schon vorher (im Vortest oder im Haupttest) "aussortiert" wurden.

Der Empfehlungsgrad "3" ("entspricht nur teilweise den Anforderungen") signalisiert entsprechend der *DGP*-Definition ein gewisses Risiko in bezug auf den Ausbildungserfolg. Die Chance, daß solche Bewerber/ (innen) im Durchschnitt eine Note von "befriedigend" oder besser erzielen, beträgt gerade einmal 1:3, also nur 33,3 % dieser Bewerber(innen) liegen in der Spitzengruppe der Azubis.

Insgesamt vermitteln die Ergebnisse einen guten Eindruck von der prognostischen Validität der *DGP*- Empfehlungsgrade. Wichtig wäre es natürlich zu prüfen, ob dieser Zusammenhang für alle einzelnen Tests (z.B. Diktat, Aufgabenheft "Tabellen und Statistiken" usw.) gilt oder auf einige bestimmte Aufgaben beschränkt ist. Leider sind im vorliegenden Datensatz nur die Gesamtwerte verzeichnet, so daß diese tiefergehenden Analysen zur Zeit nicht möglich sind.

2.2 Schulnoten

Während in Fachkreisen die Unzulänglichkeiten von Schulnoten als Auslese-kriterium hinreichend bekannt sind (z.B. Althoff, 1986; Athanasiadis, 1974; Goldman & Slaughter, 1976, Junghans & Althoff, 1978), gelten Schulnoten bei einigen (von empirischen Begründungszwängen offensichtlich befreiten) Verwaltungspraktiker(innen) nach wie vor als Alternative zu psychologischen Eignungsuntersuchungen. Gerade die Finanznot, die organisatorische Überlastung und die Fremdheit der Testverfahren nähren in den neuen Ländern die Stimmen für eine finanziell und aufwandsmäßig kaum belastende Personalauswahl nach Zeugnissen. Vor dem Hintergrund der bekannten Subjektivität und Unvergleichbarkeit von Schulnoten gilt es hier allerdings zu fragen, ob solche -nicht sachlichen, sondern:- vordergründig pragmatischen Argumente die Personalauswahl mit ihren weitreichenden Konsequenzen für die Organisation und für die bewerbende Person leiten sollten?

Auf Wunsch der betroffenen Verwaltung hat die *DGP* eine gewichtete Durchschnittsschulnote des Schulabschlusses in die Analyse miteinbezogen. In die Berechnung der Note gingen die Fächer "Deutsch", "Fremdsprachen", "Naturwissenschaften" und "Mathematik" zu jeweils 25 % ein.

Es muß hier allerdings explizit betont werden, daß ein oft gewünschter direkter "empirischer" Vergleich zwischen Schulnoten und psychologischen Eignungsuntersuchungen mit dem gegebenen Datensatz prinzipiell unmöglich ist, da ja nur diejenigen Personen zur Ausbildung zugelassen wurden, die aufgrund der *DGP*-Empfehlungen in Frage kamen. Dieser Punkt soll später noch

erläutert werden, zunächst zu den Schulnoten und dem Zusammenhang zwischen Schulnoten und dem Ausbildungserfolg.

Notendurchschnitt FH-Reife / Abitur, gewichtet	Häufigkeit	in Prozent
"1,0 --- 1,49" ("sehr gut")	20	26,4 %
"1,5 --- 2,49" ("gut")	41	53,9 %
"2,5 --- 3,5" ("befriedigend")	15	19,7 %
Tab. 4: Notendurchschnitt: drei Kategorien Definition und Häufigkeiten	76	100 %

Für die vorliegende Analyse wurden drei Kategorien von Zeugnisnoten berücksichtigt. Tabelle 4 zeigt die Häufigkeiten, mit der sich die ausgewählten Bewerber(innen) des Einstellungsjahrgangs 1991 auf die einzelnen Zeugnisnoten verteilen. Für zwei Bewerberinnen konnten der Notendurchschnitt nicht ermittelt werden, daher bezieht sich die Aufstellung lediglich auf 76 Teilnehmer(innen). Man sieht, daß über 80 % der Bewerberinnen und Bewerber einen Notendurchschnitt von "gut" oder "sehr gut" aufweisen.

2.2.1 Kreuztabelle

Tabelle 5 zeigt den Zusammenhang zwischen dem Notendurchschnitt FHöV und dem Notendurchschnitt FH-Reife/Abitur. Die Tabelle zeigt -und das war für uns überraschend,- daß durchaus ein Zusammenhang zwischen den durchschnittlichen Schulabschlußnoten der ausgewählten Anwärterinnen und Anwärter und deren Abschneiden im ersten Studienabschnitt besteht. So hat in der Gruppe der von der Schule mit "sehr gut" bewerteten Anwärter(innen) kein(e) einzige(r) einen Klausurendurchschnitt von "mangelhaft".

Dies bedeutet aber genau genommen nur folgendes: Wenn jemand einen Schulnotendurchschnitt von "sehr gut" aufweist und diese Person **zusätzlich** von der *DGP* mit einem Empfehlungsgrad von nicht schlechter als "3" eingestuft wurde, dann (und nur dann) liegt sie in der Gruppe derjenigen, deren Klausurendurchschnitt besser als "mangelhaft" ist. Anhand der vorliegenden Tabelle kann man nämlich nichts über das Abschneiden der großen Zahl von Bewerber(innen) mit einem Schulnotendurchschnitt von "sehr gut" aussagen, die aufgrund ihrer schlechten Ergebnisse in den *DGP*-Eignungsuntersuchungen bei der Auswahl ausgeschieden sind und gar nicht erst die Ausbildung begonnen haben. Berücksichtigt man, (1.) daß nur einer von acht Bewerberinnen oder Bewerbern den *DGP* -Anforderungen entspricht und daß es (2.) kaum einen Zusammenhang von Schulnoten und den Ergebnissen von psychologischen Eignungsuntersuchungen gibt und nimmt ferner einmal an, daß (3.) ca. 25 % aller Bewerberinnen und Bewerber einen Notendurchschnitt von "sehr gut" aufweisen, so muß man davon ausgehen, daß in dem ursprünglichen Bewerberpool von etwas über 600 Personen mindestens 150 mit einer Durchschnittsnote von "sehr gut" ausgestattet waren.

b) Zusammenhang zwischen dem Notendurchschnitt FHöV und dem Notendurchschnitt FH-Reife/Abitur

Klausuren FHöV	Schulnotendurchschnitt FH-Reife/Abitur		
	"sehr gut"	"gut"	"befriedigend"
"befriedigend" und besser	14 (18,4 %)	18 (23,7 %)	8 (10,5 %)
	35 %	45 %	20 %
	70 %	43,9 %	53,3 %
"ausreichend"	6 (7,9 %)	21 (27,6 %)	5 (6,6 %)
	18,8 %	65,6 %	15,6 %
	30 %	51,2 %	33,3 %
"mangelhaft"		2 (2,6 %)	2 (2,6 %)
		50 %	50 %
		4,9 %	13,3 %
Tab. 5: Kreuztabelle	20 (100 %)	41 (100 %)	15 (100 %)

Die wirklich anspruchsvolle Aufgabe, nämlich unter diesen ca. 150 Personen die Geeigneten herauszufinden (8 hat die *DGP* hier entdeckt) wurde hier ja bereits von den *DGP*-Verfahren übernommen, so daß unter Umständen die Vielzahl von Ausbildungsversager(innen), die man sich eventuell aufgrund einer Personalauswahl nach Schulnoten eingehandelt hätte, schon gar nicht mehr in der Stichprobe berücksichtigt sind. Ein Beleg für die Tauglichkeit der Schulnoten wäre nur dann gegeben, wenn man nachweisen könnte, daß auch die verbleibenden 142 Personen die Ausbildung erfolgreich bestehen könnten. Stellt man sich etwa zwanzig Bewerberinnen und Bewerber vor, zehn in berufsrelevanten Dingen leistungsfähige und zehn eher leistungsschwache Bewerber(innen), die alle einen Schulnotendurchschnitt von "sehr gut" aufweisen (die einen aufgrund ihrer Leistungen, die anderen aufgrund z.B. ihrer politischen Contenance, guter Beziehungen oder ähnlichem), so würde ja nur die leistungsstarken Bewerber(innen) das *DGP*-Auswahlverfahren bestehen, zur Ausbildung zugelassen und in der Ausbildung natürlich auch gute Ergebnisse zeitigen. Unter diesen Bedingungen kann man dann nicht im nachhinein aus der falschen Behauptung "Honig saugen", für diese Vorhersage hätte man das anspruchsvolle *DGP*-Verfahren nicht gebraucht, dies hätte man aufgrund der Schulnote von "sehr gut" doch auch "vorher" schon gewußt. Da ein aussagekräftiger empirischer Vergleich auf Grundlage der gegebenen Daten prinzipiell nicht möglich ist, sollen hier noch einmal kurz die theoretischen Erwägungen angeführt werden, die gegen eine Verwendung von Zeugnisnoten für Personalentscheidungen sprechen. Dabei soll von Anfang an klargestellt werden, daß keinesfalls der Sinn und Zweck von Schulnoten insgesamt angezweifelt wird. Schulnoten mögen innerhalb des Systems "Schule" ihre Berechtigung und pädagogische Bedeutung haben. Die *DGP* wendet sich lediglich dagegen, Schulzeugnisse für Zwecke zu mißbrauchen, für die sie

nicht ausgestellt wurden. Zur Feststellung der beruflichen Eignung sollten nur solche Methoden angewandt werden, die den Kriterien der (1) Objektivität (2) Zuverlässigkeit (3) Vergleichbarkeit und (4) Berufsspezifität genügen.

Während die *DGP*-Verfahren diese Kriterien erfüllen, haben Althoff (1986) sowie Arndt und Althoff (1986) wiederholt festgestellt, daß dies für Schulnoten keinesfalls gilt. Schulnoten sind weder zwischen verschiedenen Schultypen noch zwischen Schulen des gleichen Schultyps vergleichbar (dies gilt auch für Universitätsbeurteilungen, siehe beispielsweise Bee & Dolton, 1985). Selbst innerhalb einer Schule wird durch unterschiedliche Lehrer(innen) und durch unterschiedliche Niveaus bestimmter Klassenzusammensetzungen eine vergleichbare Bewertung vereitelt. Ja sogar ein und dieselbe Lehrperson beurteilt gleiche Leistungen zu verschiedenen Zeitpunkten abweichend (siehe z. B. Betz, 1974). Ein grundsätzliches Problem besteht auch darin, daß Lehrpersonen oft nicht den Mut aufbringen, Leistungs- und Verhaltensmängel bzw. Kenntnislücken von Schüler(innen) auch in entsprechenden Beurteilungen zum Ausdruck zu bringen. In der vorliegenden Untersuchung war die schlechteste Note ein "befriedigend", eine Note, die allein vom semantischen Gehalt her keine negative Einstufung vermuten läßt.

All diese Probleme sind -wie gesagt- seit langem bekannt. Neu ist, daß seit der Vereinigung Bewerber(innen) mit Zeugnissen aus grundsätzlich verschiedenen Schultypen ("DDR" versus "BRD") auf dem Personalmarkt sind. Die Gleichsetzung von Schulnoten unterschiedlicher Systeme scheint noch abwegiger als die Gleichsetzung von Noten innerhalb eines Systems. Auch vor diesem Hintergrund wäre der Rückschritt zur Verwendung von Schulnoten für eignungsdiagnostische Fragestellungen besonders abenteuerlich. Gerade jetzt, in der Zeit des Übergangs, da ein Teil der Bewerber(innen) Zeugnisse aus Schulen der ehemaligen "DDR" vorlegt, ein anderer Teil (die jüngeren Bewerber(innen)) aber schon in gewandelten Schulen unterrichtet und benotet wurde, scheint es unumgänglich, eine neue Meßlatte anzulegen, die allen die gleichen Chancen einräumt bzw. für alle gleich hoch hängt.

Da zwischen Schulnoten und *DGP*-Ergebnissen schließlich kaum ein Zusammenhang besteht (das heißt, wer gute Zeugnisnoten hat, schneidet nicht unbedingt im *DGP*-Verfahren gut ab und umgekehrt, siehe Arndt und Althoff, 1986), so würde dies für den Fall, daß Schulnoten tatsächlich einmal (in vergangener oder zukünftiger Zeit) den Ausbildungserfolg prognostizieren, ohnehin nur bedeuten, daß man mit einer Kombination von Schulnoten und *DGP*-Empfehlungsgraden die beste Prognose treffen könnte. Wenn nämlich zwei Prädiktoren unabhängig voneinander gute Vorhersagen leisten, so bedeutet dies, daß sie jeweils einen anderen wichtigen Aspekt diagnostizieren. Die Vorhersage wird dann natürlich noch gesteigert, wenn man beide Aspekte gleichzeitig berücksichtigt.

3. Fazit

In dem vorliegenden Text wurde der Zusammenhang zwischen den Personaleinstellungsempfehlungen der *DGP* und der durchschnittlichen Note bei schriftlichen Klausuren im Anschluß an den ersten Studienabschnitt an der FHöV theoretisch erörtert und empirisch geprüft. Die Ergebnisse für die insgesamt 78 Regierungsinspektoranwärter(innen) belegen deutlich die hohe Treffsicherheit der "Empfehlungsgrade" der *DGP* **auch in den neuen Ländern**. Bezogen auf diejenigen 66 Personen, die von der *DGP* als "*den Anforderungen* zumindest *weitgehend entsprechend*" (Empfehlungsgrad "4" und "5") beurteilt worden waren, liegt die Fehlerquote in bezug auf eine Auswahl der Erfolgreichen bei drei von 66 oder bei **4,5 %!** Dieser Wert ist erfreulich niedrig und dürfte kaum durch eine andere Vorgehensweise zu unterbieten sein. Er belegt, daß die Eignungsuntersuchungen der *DGP* auch in den fünf neuen Ländern psychologisch und ökonomisch wertvolle Informationen über den voraussichtlichen Ausbildungserfolg liefern.

Die Entscheidung über den Einsatz oder den Verzicht von Eignungsuntersuchungen sollte sich überwiegend an solchen empirischen Bewährungskontrollen orientieren, deren Prinzip hier vorgestellt wurde. Die Personalauswahl darf sich weder an kurzfristigen Sparvorhaben noch am "*Evidenzgefühl*" (Hörmann, 1964) einzelner orientieren, die der Meinung sind, man könnte auch mit selbstgestrickten Verfahren oder mit Schulzeugnissen die geeigneten Personen herausfinden.

4. Literatur

- Althoff, K. (1971). Die Voraussage des Berufserfolges von Regierungsinspektoren und Revierförstern - zwei Bewährungskontrollen. Diagnostica, 17, 146-180.
- Althoff, K. (1973). Untersuchungen zum Einfluß von Berufsausbildungen und Berufstätigkeiten auf die Ergebnisse einer Eignungstestbatterie. Diagnostica, 19, 143-180.
- Althoff, K. (1984). Zur prognostischen Validität von Intelligenz- und Leistungstests im Rahmen der Eignungsdiagnostik. Psychologie und Praxis - Zeitschrift für Arbeits- und Organisationspsychologie, 28, 144-148.
- Althoff, K. (1986). Zur Aussagekraft von Schulzeugnissen im Rahmen der Eignungsdiagnostik. Psychologie und Praxis - Zeitschrift für Arbeits- und Organisationspsychologie, 30, 77-85.
- Arndt, R. & Althoff, K. (1986). Vorauslese nach Zeugnisnoten? Bericht über die Analyse der Beziehungen zwischen Zeugnisnoten und Testergebnissen bei einer nicht nach Zeugnisnoten vorausgelesenen Bewerberstichprobe. DGP Informationen, 46, 49-65.
- Athanasiadis, T. (1974). Zum System der Leistungsbewertung mit Zensuren. Ein Beitrag zur Quantifizierung von Leistungsvergleichen. Statistische Studien, 7.
- Bee, M. & Dolton, P. (1985). Degree class and pass rates: an inter-university comparison. Higher Education Review, 17, 45-62.
- Betz, D. (1974). Rhythmische Schwankungen als Fehler in der Notengebung bei mündlichen Prüfungen. Psychologie in Erziehung und Unterricht, 21, 1-14.
- Brandstätter, H. (1970). Leistungsprognose und Erfolgskontrolle. Bern: Huber.
- Goldman, R. & Slaughter, E. (1976). Why College Grade Point Average is difficult to predict. Journal of Educational Psychology, 68, 9-14.
- Hörmann, H. (1964). Aussagemöglichkeiten psychologischer Diagnostik. Zeitschrift für Experimentelle und Angewandte Psychologie, 11, 353-390.
- Jäger, A.O. (1970). Personalauslese. In K.Gottschaldt, Ph. Lersch, F.Sander & H.Thomae (Hrsg.), Handbuch der Psychologie (Bd. 9, Betriebspsychologie, S. 613-667). Göttingen: Hogrefe.
- Kleinevoss, R. & Sonnenberg, H.G. (1987). Die prognostische Validität einer eignungsdiagnostischen Prozedur - dargestellt mit einem konfiguralen Diskriminanzmodell. Zeitschrift für Arbeits- und Organisationspsychologie, 31, 15-21.
- Junghans & Althoff, K. (1978). Sind die Zeugnisnoten der Bewerber für den mittleren und gehobenen Verwaltungsdienst als Zulassungskriterium für die Teilnahme an einer psychologischen Eignungsuntersuchung geeignet? DGP Informationen, 16.
- Wolf, B. (1990). Zusammenhänge zwischen Ergebnissen von eignungsdiagnostischen Verfahren und Ausbildungserfolg - Dargestellt am Beispiel der Auswahl von Bewerbern für den gehobenen Polizeivollzugsdienst (Diplomarbeit an der Georg-August Universität zu Göttingen).