

Johann Wolfgang Goethe-Universität Frankfurt am Main
Studiengang Diplom-Psychologie
Diplomarbeit

***Personalauswahlmethode und Praxisorientierung der Ausbildung als
Bedingungsfaktoren für den Ausbildungserfolg von
Verwaltungsangestellten. Eine Evaluationsstudie.***

vorgelegt von:
Matthias Mayer

Betreuender Gutachter:
Herr Prof. Dr. Helfried Moosbrugger

Zweiter Gutachter:
Herr Prof. Dr. Marin Ignatov

Wiesbaden, den 29.10.96

INHALTSVERZEICHNIS

	Seite
Danksagung	
1 Kurzfassung.....	01
2 Einleitung.....	02
3 Problemstellung.....	04
3.1 Einstellungsgespräch und Test als Methoden der Personalauswahl.....	04
3.1.1 Das Einstellungsgespräch.....	04
3.1.2 Der Test.....	08
3.2 Das Ausbildungssystem der gewerblichen Berufsgenossenschaften.....	10
3.3 Die Praxis der Auswahl geeigneter Anwärter in den BGen.....	11
3.4 Praxisorientierung der Ausbildung.....	13
3.5 Leistungsdruck und Einstellung zur Arbeit als Eigenschaften der Organisation.....	14
4 Methode.....	16
4.1 Unabhängige Variablen.....	17
4.1.1 Operationalisierung der Variablen Selektion.....	17
4.1.2 Operationalisierung der Variablen Praxisorientierung...18	18
4.1.3 Operationalisierung der Variablen Leistungsdruck.....18	18
4.1.4 Operationalisierung der Variablen Einstellung zur Arbeit.....	18
4.2 Abhängige Variablen.....	19
4.3 Hypothesen.....	21
4.4 Design.....	22
4.5 Stichprobe.....	23
4.6 Untersuchungsablauf.....	24
4.6.1 Schriftliche Befragung der Berufsgenossenschaften.....26	26
4.6.2 Erfassung der abhängigen Variablen.....	26
4.6.3 Zweite schriftliche Befragung.....	27
4.6.3.1 Entwicklung der Instrumente zur Messung der Variablen Leistungsdruck, Einstellung zur Arbeit und Praxisorientierung.....	27
4.6.3.2 Konstruktion der Leistungsdruck-Skala.....	28
4.6.3.3 Konstruktion der Skala Einstellung zur Arbeit..29	29
4.6.3.4 Praxisorientierung.....	30
4.6.4 Telefonische Befragung der Berufsgenossenschaften.....31	31

4.7	Verwendete Methoden und statistische Programmpakete.....	32
5	Ergebnisse.....	32
5.1	Ergebnisse der schriftlichen Befragung der Berufsgenossenschaften.....	32
5.2	Auswertung des Notenarchivs der Jahre 1994/95.....	34
5.2.1	Noten und durchgefallene Anwärter.....	34
5.2.2	Ausbildungsabbrecher.....	36
5.3	Ergebnisse der telefonischen Befragung der Berufsgenossenschaften.....	37
5.4	Erfassung der Variablen Einstellung zur Arbeit, Leistungsdruck und Praxisorientierung.....	40
5.5	Ergebnisse zur Testkonstruktion.....	41
5.5.1	Stichprobe.....	41
5.5.2	Skala zur Messung der Variablen Leistungsdruck.....	42
5.5.3	Skala zu Messung der Variablen Einstellung zur Arbeit..	44
5.5.4	Messung der Praxisorientierung.....	45
5.5.5	Ergebnisse der explorativen Fragen.....	46
5.6	Berechnung der BG-Kennwerte.....	47
5.6.1	Selektion.....	48
5.6.2	Schriftliche Durchschnittsnote.....	50
5.6.3	Leistungsdruck.....	51
5.6.4	Praxisorientierung.....	52
5.6.5	Einstellung zur Arbeit.....	54
5.7	Prüfung der Hypothesen.....	55
5.7.1	Überprüfung der Hypothese 1.....	55
5.7.2	Überprüfung der Hypothese 2.....	59
5.7.3	Überprüfung der Hypothese 3.....	60
5.7.4	Überprüfung der Hypothese 4.....	62
6	Diskussion.....	67
7	Zusammenfassung.....	79
	Literaturverzeichnis	

Anhang

Danksagung

Diese Diplomarbeit hätte nicht ohne die Unterstützung des Hauptverbands der gewerblichen Berufsgenossenschaften entstehen können. Besonderen Dank schulde ich dem Leiter der "Berufsgenossenschaftlichen Akademie für Arbeitssicherheit und Verwaltung" (BGA), Herrn Assessor Baucke, der mir half, bei der Planung und Durchführung der Untersuchung kleinere und größere Organisationsprobleme zu meistern. Dank an Frau Reich aus dem Prüfungsamt der BGA, ohne deren bereitwillige Hilfe ich sicher ratlos in Aktenbergen verloren gegangen wäre. Auch all jenen sei an dieser Stelle Dank ausgesprochen, die meine mündlichen oder schriftlichen Fragen geduldig beantworteten.

1 Kurzfassung

In einer Evaluation unterschiedlicher Auswahl- und Ausbildungskonzepte der gewerblichen Berufsgenossenschaften (BGen) wurden die Hypothesen geprüft, daß durch Testverfahren auswählende BGen bessere Ausbildungsleistungen erzielen als nur durch Interview auswählende, praxisorientierte Ausbildungen zu schlechteren Prüfungsleistungen führen, zwischen der Ausbildungsleistung und dem Leistungsdruck sowie der Einstellung zur Arbeit ein Zusammenhang besteht und diese Variablen sich daher zur Erhöhung des varianzaufklärenden Werts des zweifach gestuften Faktors Selektion eignen. Für die Wirkung des Faktors Selektion wurde die Praxisorientierung als Moderatorvariable vermutet. Es gelang, 75% aller BGen zu erfassen. Als unabhängige Variable wurde in einem ex-post-facto Design der Faktor Selektion auf den Stufen Interview und Test aufgefaßt. In einem einfaktoriellen univariaten kovarianzanalytischen Modell dienten Leistungsdruck, Einstellung zur Arbeit und Praxisorientierung als Kovariaten. Für jede BG ist die Selektionsmethode und die Praxis der Interviewführung durch schriftliche bzw. telefonische Befragung ermittelt worden. Der Ausbildungserfolg wurde durch die Auswertung eines zentralen Notenarchivs erhoben. Durch einen Fragebogen an die 1994/95 geprüften Personen wurden Leistungsdruck, Praxisorientierung und Einstellung zur Arbeit als Kennwerte für die BG erfaßt. Testverfahren und Interview als Stufen des Faktors Selektion eigneten sich nicht zur Vorhersage des Ausbildungserfolgs. Bei den BGen, die durch Test auswählten, wurde eine deutlich größere Variabilität der Noten nachgewiesen als bei den interviewenden BGen. Eine praxisorientierte Ausbildung ging mit schlechten Prüfungsnoten einher ($r = -.13$). Es bestand keine Wechselwirkung zwischen Praxisorientierung und Auswahlmethode. Die Aufklärung der Notenvarianz durch die Auswahlmethode konnte durch die Kovariaten nur von 2% auf 3% erhöht werden. Mit Einstellung zur Arbeit und Praxisorientierung als Kovariaten konnten insgesamt 17% Varianz aufgeklärt werden. Die Qualität der Messung der Variablen, Gründe für die große Variabilität der Noten unter der Testbedingung sowie die Frage nach der Relevanz der Befunde zur Praxisorientierung der Ausbildung wurden diskutiert.

2 Einleitung

Die Frage der Auswahl geeigneter Mitarbeiter ist in der gesamten Wirtschaft wie auch in der öffentlichen Verwaltung von großer Bedeutung. Für viele Organisationen ist es eine existentielle Frage, die richtigen Personen am richtigen Platz einzusetzen. Im Bereich der öffentlichen Verwaltung geht es, da diese nicht im direkten Wettbewerb steht, seltener um die Existenz der Organisation, aber ihr effizientes Funktionieren ist ebenfalls abhängig von einem erfolgreichen Personalmanagement. Gerade in Zeiten leerer staatlicher Kassen geraten Behörden und Ämter in den Blick der kritischen Öffentlichkeit. Diese Öffentlichkeit unterscheidet dabei nicht, ob es sich um kommunale-, Landes- oder Bundesbehörden, mittelbare oder unmittelbare Verwaltung handelt oder ob diese Verwaltung steuer- oder beitragsfinanziert ist. Insofern befinden sich die gewerblichen Berufsgenossenschaften, die als Körperschaften des öffentlichen Rechts mit Selbstverwaltung eine gewisse Unabhängigkeit genießen, neben Renten- und Krankenversicherung in ähnlicher Situation wie die übrige staatliche Verwaltung.

Die Rekrutierung des Nachwuchses und seine Ausbildung ist eine der zentralen Zukunftsaufgaben jeder Organisation. Außer dieser eher strategischen Betrachtungsweise liegt es auch im kurzfristigen finanziellen Interesse der Organisation, im Bereich der Berufsausbildung effektiv zu arbeiten. Die Einstellung eines Anwärters für den gehobenen Verwaltungsdienst bedeutet für eine Berufsgenossenschaft nicht unerhebliche Kosten. Vor diesem Hintergrund stellt sich die Frage nach der Praxis der Gewinnung neuen Personals und dessen sorgfältiger Ausbildung.

Die gewerblichen Berufsgenossenschaften (im folgenden BGen genannt) haben bei aller Ähnlichkeit der Arbeitsplätze, Arbeitsaufgaben, Organisationsstrukturen und maßgeblicher Rechtsvorschriften unterschiedliche Konzepte der Auswahl geeigneter Kandidaten für die gehobene Verwaltungslaufbahn und der Gestaltung der Ausbildung. Die vorliegende Arbeit beschäftigt sich mit der Aufklärung und Bewertung der in den Berufsgenossenschaften praktizierten Verfahren zur Auswahl und Ausbildung der Anwärter für die gehobene Verwaltungslaufbahn.

Bei der Auswahl geeigneter Bewerber um einen Ausbildungsplatz kann man im wesentlichen zwei Ansätze unterscheiden. Zum einen werden Bewerber nach einer Vorauswahl aufgrund der eingereichten Bewerbungsunterlagen (Bewerbungsschreiben, Lebenslauf, Zeugnis) zu einem persönlichen Gespräch eingeladen, nach dem entschieden wird, ob es zu einer Einstellung kommt oder nicht. Die andere Möglichkeit besteht darin, nach der Auswertung der Bewerbungsunterlagen mit Testverfahren Daten zu erheben, die zu direkt vergleichbaren Ergebnissen der Bewerber führen. Die Bewerber mit den besten Ergebnissen werden zu einem Gespräch eingeladen, aufgrund dessen dann die endgültige Entscheidung über die Einstellung fällt.

Hinsichtlich der Durchführung der praktischen Ausbildung in den einzelnen berufsgenossenschaftlichen Verwaltungen gibt es zwar Vereinbarungen und Rahmenvorschriften, wie z.B. die "Fortbildungs- und Prüfungsordnung der gewerblichen Berufsgenossenschaften" (FPO), man kann aber durchaus Unterschiede in der Ausbildung feststellen. Insbesondere das Ausmaß, in dem Dienstanwärter in die Sachbearbeitung eingebunden sind, ist unterschiedlich. Zum einen werden die Dienstanwärter Sachbearbeitern zur Ausbildung zugeordnet. Hierdurch sollen sie Einblick in die laufende Arbeit bekommen und in diese einbezogen werden. Zum anderen kommt es zur Bildung besonderer Ausbildungsgruppen, die teilweise in Form eines BG-internen Unterrichts unterwiesen werden und allenfalls einige Fälle aus der Praxis exemplarisch für bestimmte Rechtsprobleme zur Bearbeitung bekommen. Zwischen diesen beiden Extremen finden sich Abstufungen.

Bisher ist eine Evaluation dieser unterschiedlichen Konzepte unterblieben. Die vorliegende Untersuchung versteht sich als ein erster Ansatz einer solchen Evaluation. In vielerlei Hinsicht handelt es sich daher um eine explorative Arbeit. Es wird eine Aufklärung des Zusammenhanges zwischen der Auswahl geeigneter Bewerber, ihrer Ausbildung und der resultierenden Ausbildungsleistung anhand der Einstellungsjahrgänge 1991/92 angestrebt. Die Analyse wird dabei auf

Organisationsebene durchgeführt, da es sich um die Evaluation von Verfahren und Konzepten handelt, die als Eigenschaften von Organisationen zu betrachten sind.

3 Problemstellung

Da Ziel dieser Arbeit nicht Wissensgewinnung durch experimentelles Vorgehen ist, sondern es sich um die nachträgliche Beurteilung der Wirksamkeit bestimmter Maßnahmen handelt, sind in der Problemstellung vor allem die untersuchten empirischen Verhältnisse genau zu beschreiben (vgl. Abschnitt 3.2-3.5). Psychologische Theorien können durch diese Evaluation nicht falsifiziert werden und sind deswegen auch nicht Gegenstand dieser Arbeit. Sie bilden aber den Hintergrund, vor dem die Ergebnisse dieser Untersuchung zu beurteilen sind. Forschung wird daher insoweit referiert, als sie Grundlage dieser Arbeit ist (vgl. Abschnitt 3.1).

3.1 Einstellungsgespräch und Test als Methoden der Personalauswahl

3.1.1 Das Einstellungsgespräch

Das Interview steht als Methode der Personalauswahl seit langer Zeit in der Kritik. Schon 1915 wies Scott (zit. nach Schuler & Funke, 1993) auf die mangelhafte Reliabilität des Auswahlgesprächs hin. An dieser Kritik hat sich bis heute nichts geändert, auch nicht an der Tatsache, daß das persönliche Gespräch eines der beliebtesten Auswahlinstrumente geblieben ist.

"Bei der Einstellung neuer Mitarbeiter wird in nahezu allen Fällen neben der Durchsicht der Bewerbungsunterlagen ... ein Einstellungsgespräch geführt. Dieses Gespräch hat meist entscheidende Bedeutung, bei der Annahme oder Ablehnung des Bewerbers Dem unerschütterlichen Vertrauen von

Führungskräften oder Personalleitern in den Wert des Einstellungsgesprächs steht jedoch auf Seiten der Betriebspsychologie eine Phalanx ablehnender Meinungen gegenüber ."

(Neuberger, 1974, zit. nach Schoenborn, 1988)

Die Verbreitung des Interviews als Methode der Selektion steht im Gegensatz zu gesicherten wissenschaftlichen Erkenntnissen über seine Validität und Reliabilität. Für die prädiktive Validität (zum Begriff siehe Amelang & Bartussek, 1990, S.143), also der Gültigkeit des Schlusses vom Interviewergebnis auf zukünftige Leistungen, ergab sich metaanalytisch ein Validitätskoeffizient von $r=0.14$ (Hunter & Hunter, 1984). Eine Quelle dieser Unzuverlässigkeit liegt in der Tatsache begründet, daß es sich bei einem Einstellungsgespräch um einen klinischen und nicht um einen empirischen Prozeß handelt (Meehl, 1954 und Wiggins, 1973, zit. nach Weinert, 1992, S. 228). Dabei kommt die ganze Bandbreite der aus der sozialpsychologischen Forschung bekannten Urteilsverzerrungen zur Geltung. Hier sind insbesondere Stereotype, implizite Persönlichkeitstheorien, Vorerwartungen aufgrund der schriftlichen Bewerbungsunterlagen und Kontextfaktoren, wie z. B. die Reihenfolge der Bewerber, zu nennen.

Borkenau (1991) weist darauf hin, daß erste Eindrücke, wie sie in einem Einstellungsgespräch zustande kommen, durchaus eine gewisse Validität für sich beanspruchen können. Allerdings beschränkt sich diese auf wenige grundlegende Persönlichkeitseigenschaften wie Extraversion und Gewissenhaftigkeit. Schon bei der Intelligenz ergaben sich nur geringe Zusammenhänge zwischen Beobachterurteil und Testwerten. Die Vorhersage der Eignung im Hinblick auf ein spezifisches Berufsbild ist auf diesem Wege sicherlich nicht zuverlässig möglich.

"Viel zu oft macht man sich bei Personalentscheidungen ohne psychologische Verfahren auch heute noch 'ein Bild von der Person', kombiniert Ungenauigkeiten, Urteilsfehler und persönliche Voreingenommenheit zu ungerechten, folgenschweren und auch teuren Fehlentscheidungen."

(Schuler & Funke, 1987, S. 58)

Seine Beliebtheit verdankt das Einstellungsgespräch wohl den Vorteilen, die es gegenüber anderen Selektionsinstrumenten, insbesondere psychometrischen Tests, hat. Als solche Vorteile sind mit Schoenborn (1988) zu nennen:

- Unmittelbarkeit der Datengewinnung
- Flexibilität der Methode
- Möglichkeit, anders nicht erfaßbare Daten, z.B. über das non-verbale- und das Sozialverhalten, zu erfassen

"Wohl kaum ein anderes Instrument der Personalselektion und der Personalbewertung dürfte so universell und so häufig Anwendung finden wie das Auswahl- und Vorstellungsgespräch Damit kommt dem Auswahlgespräch eine Schlüsselfunktion unter den verschiedenen Selektionsinstrumenten zu, denn in einer Vielzahl von praktischen Selektionsprozessen werden Entscheidungen bereits zu diesem Zeitpunkt und an dieser Stelle gefällt."

(Weinert, 1992, S. 228)

Die bisherige Forschung zu Methoden der Personalselektion zeigt, daß das Einstellungsgespräch gegenüber wissenschaftlich fundierten psychometrischen Verfahren sowohl geringere Reliabilität als auch geringere Validität besitzt (siehe hierzu Scott, 1916; Anderson, 1960; Bolster & Springbett, 1961; Arvey & Campion, 1982; Carlson, 1967a, b; Carlson, 1968, Carlson, 1969; Carlson & Mayfield, 1967). Die von Hunter & Hunter (1984) berichtete geringe prädiktive Validität des Einstellungsgesprächs hat wohl zwei wesentliche Ursachen: die unzulängliche Informationsverarbeitung und die mangelnde testtheoretische Fundierung des Einstellungsgesprächs (Schuler, 1993, S. 247). Für viele Einstellungsgespräche gilt daher:

"Wenn sie nicht in eine planloses Ausfragerei ausarten, haben sie oft nicht mehr als Kaffekränzchenniveau." (Schuler & Funke, 1987, S. 60)

Im einzelnen hat die Forschung zum Einstellungsgespräch folgende Ursachen für die methodische Schwäche ergeben:

- Negative Information wird stärker gewichtet als positive.
- Die Reliabilität ist abhängig vom Grad der Strukturiertheit eines Gesprächs.
- Die Auswertungsobjektivität ist gering.
- Die Urteilsbildung erfolgt bereits in einem vergleichsweise frühen Stadium.
- Kontextfaktoren, wie die Güte des vorausgegangenen Kandidaten, beeinflussen das Urteil.
- Stereotype, Vorurteile und implizite Persönlichkeitstheorien des Beurteilers nehmen Einfluß.
- Aufgrund des Interaktionsprozesses, den ein Gespräch darstellt, findet eine gegenseitige Beeinflussung von Bewerter und Kandidat statt, die zu unterschiedlichen Gesprächsverläufen und damit zur Unvergleichbarkeit der Daten verschiedener Bewerber führt.

Dreher, Ash & Hancock (zit. nach Schuler, 1993, S. 247) wiesen 1988 darauf hin, daß es bei der Beurteilung der Validität zu methodischen Artefakten kommen kann, die zur Unterschätzung der Validität von Einstellungsgesprächen führen.

"... ,daß durch unterschiedliche Urteilstendenzen der Beurteiler die Validität von Interviews unterschätzt wird, die ja gewöhnlich auf aggregiertem Niveau berechnet wird, wodurch selbst verlässliche Beurteilerunterschiede als Fehlervarianz in die Berechnung eingehen." (Schuler, 1993, S. 247)

Es mag durchaus einzelne Personen geben, die in besonderer Weise dazu befähigt sind, vergleichsweise reliabel und valide Persönlichkeitseigenschaften im Gespräch zu erfassen. Problematisch ist hierbei allerdings der weitverbreitete Glaube, zu dieser besonders begabten Minderheit zu gehören, deren Existenz keineswegs belegt, sondern nur denkbar ist. Außerdem sind Persönlichkeitseigenschaften keineswegs gleichzusetzen mit der Beurteilung von Berufseignung.

Das Interview kann jedoch methodisch so umgestaltet werden, daß es unter Beibehaltung seiner Vorteile reliabler und valider wird. In ihrer standardisierten Form können Interviews Validitätskoeffizienten erreichen, die an die Werte von

Assessment-Centern heranreichen. Wiesner & Cronshaw (1988) berichten von Validitäten strukturierter Interviews in Höhe von $r=0.4$.

3.1.2 Der Test

Unter einem Test versteht man

".....ein wissenschaftliches Routineverfahren zur Untersuchung eines oder mehrerer empirisch abgrenzbarer Persönlichkeitsmerkmale mit dem Ziel einer möglichst quantitativen Aussage über den relativen Grad der individuellen Merkmalsausprägung." (Lienert, 1989, S.7)

Die Hauptgütekriterien eines Tests sind nach Lienert (1989, S.12) Reliabilität, Validität und Objektivität. Die Vorteile des Tests als Selektionsinstrument liegen in seiner Minimierung subjektiver Einflüsse von Seiten des Durchführenden und des Auswertenden (Objektivität), der Trennung von Durchführung und Urteilsbildung und der Standardisierung des Frageninhalts. Aufgrund der höheren Objektivität besitzen Testverfahren immer auch höhere Reliabilität, da die auf mangelnde Objektivität zurückgehende Fehlervarianz minimiert oder ganz eliminiert wird (zur Abhängigkeit von Reliabilität, Objektivität und Validität siehe Lienert, 1989, S. 40).

Tests, die nach wissenschaftlichen Standards entwickelt, durchgeführt und ausgewertet werden, versuchen, die für das Interview festgestellten Nachteile zu vermeiden. Testergebnisse verschiedener Personen sind direkt vergleichbar, da die Aufgaben für alle Bewerber dieselben sind. Unterschiedliche Ergebnisse können also, die Konstanz äußerer Bedingungen vorausgesetzt, auf die Person zurückgeführt werden. Desweiteren sind Tests im Vergleich zu anderen Selektionsinstrumenten relativ kostengünstig und einfach durchzuführen.

Voraussetzung für einen sinnvollen und effektiven Einsatz von Tests in der Leistungsprognose ist jedoch nicht nur ein nach psychodiagnostischen Gütemaßstäben konstruiertes Instrument. In der Praxis werden nicht selten valide

und reliable Testverfahren, z.B. Persönlichkeitstests, angewandt, ohne daß klar wird, wo die Beziehung zwischen dem Testergebnis und der Berufseignung besteht. Der Test ansich kann zur Diagnose von Persönlichkeitsstörungen im klinischen Bereich hochvalide, zur Diagnose von Berufseignung aber völlig ungeeignet sein. Eignung kann sinnvollerweise nur im Bezug auf eine Anforderung hin beurteilt werden. Anforderungen können wiederum nur aus einer Analyse der notwendigen Tätigkeiten in der Arbeit abgeleitet werden. Außer Personendiagnostik muß also zur Klärung der Anforderungen auch Situationsdiagnostik betrieben werden. An einer solchen planmäßigen Analyse mangelt es aber häufig.

"Eignungsdiagnostische Verfahren sollten also nur die Merkmale von Personen erfassen, die für die Tätigkeit wichtig sind, damit nicht durch eine ungesteuerte 'Durchleuchtung' 'gläserne Menschen' geschaffen werden." (Schuler & Funke, 1987)

Wo dieser Zusammenhang gesehen wird, bestehen allerdings oft stark vereinfachte Konzepte über die Art des Zusammenhangs zwischen Anforderung und Eignung. Zumeist wird ein linearer Zusammenhang angenommen. "Je mehr, desto besser.", scheint häufig das Motto zu sein. Die Idee, daß Testleistungen aufgrund kurvilineareren Zusammenhänge nicht nur zu schlecht, sondern auch zu gut sein können, ist vielen Praktikern kontraintuitiv.

Wenn der Zusammenhang zwischen Arbeitsleistung und Testergebnis unaufgeklärt bleibt, kommt es auch gelegentlich zur undifferenzierten Anwendung verschiedenster Testverfahren, die teilweise nicht einmal zu Zwecken der Eignungsdiagnostik konstruiert wurden, sondern ihren Ursprung im klinischen Bereich haben. Zumeist handelt es sich dabei um Persönlichkeitstests. Diese Praxis ist mit ethischen Problemen behaftet. In einer repräsentativen Umfrage gaben 31.2% der in der ABO-Psychologie tätigen Psychologen an, Persönlichkeitsstrukturtests durchzuführen und 23.4% nutzten klinische Tests (Schorr, 1991). Hunter & Hunter (1984) geben die durchschnittliche Validität kognitiver Fähigkeitstests mit $r=0.54$ an. Persönlichkeitstests schneiden dagegen deutlich schlechter ($r=0.27$ nach Hunter & Hirsh, 1987, zit. nach Schuler & Funke, 1987) ab.

Da psychologisch und insbesondere testtheoretisch nicht ausgebildete Personen zu Fehl- oder Überinterpretation von Testergebnissen neigen und aufgrund zweifelhafter Auswertungen und Deutungen weitreichende Entscheidungen getroffen werden, sollten Tests nur von Diplom-Psychologen oder unter ihrer Aufsicht durchgeführt und ausgewertet werden (Schuler & Funke, 1987). Tatsächlich werden jedoch in Deutschland rund 90% der psychologischen Verfahren von Nicht-Psychologen angewandt (Schmidt & Noe, 1986, S.71-115, zit. nach Schuler & Funke, 1987).

"... werden ... in vielen Beratungsfirmen ... Tests häufig von Personen angewandt, die keinerlei Vor- und Ausbildung im Gebrauch von Tests und in deren Interpretation haben, und aufgrund von Testresultaten werden Aussagen gemacht und Urteile gefällt, die der Testentwickler in seinem Manual nicht beansprucht bzw. nicht zulassen würde." (Weinert, 1992, S. 235)

Wohl nicht zuletzt aufgrund dieser Umstände rangieren Testverfahren bei den getesteten Personen in der Beliebtheit weit hinter anderen Instrumenten der Selektion. Schuler & Funke (1987) berichten, daß 30% von 385 Studenten der Wirtschaftswissenschaften von denen bereits 70% an Eignungstests teilgenommen hatten, psychologische Eignungstests generell als negativ beurteilten, 44% empfanden die Testsituation als unangenehm und 55% waren der Meinungen, daß ihre Fähigkeiten dabei nicht erfaßt wurden. Die Hälfte wollte an diesen Verfahren freiwillig nicht noch einmal teilnehmen.

3.2 Das Ausbildungssystem der gewerblichen BGen

Die BGen als Körperschaften des öffentlichen Rechts mit Selbstverwaltung bilden eigenverantwortlich Anwärter für den gehobenen Verwaltungsdienst aus. Die Ausbildung findet grundsätzlich innerhalb der berufsgenossenschaftlichen Verwaltung statt. Da aber die angewandten Rechtsvorschriften und auch die Verwaltungspraxis innerhalb dieses Zweiges der Sozialversicherung gleichartig sind,

haben die BGen unter ihrer Dachorganisation, dem "Hauptverband der gewerblichen BGen, e.V." den theoretischen, d.h. vor allem den juristischen, Teil der Ausbildung aus den einzelnen Verwaltungen ausgegliedert. Dieser Teil der Ausbildung wird in Unterrichtsblöcken an der "Berufsgenossenschaftlichen Akademie für Arbeitssicherheit und Verwaltung" (im folgenden BGA genannt) in Hennef/Sieg durchgeführt. An dieser Akademie werden auch die Abschlußprüfungen in schriftlicher und mündlicher Form vorgenommen. Die theoretische Ausbildung ist für die Anwärter aller BGen gleich organisiert, während die Auswahlprozedur, die sie als Bewerber durchlaufen mußten, und ihre praktische Ausbildung in der Verwaltung unterschiedlich ist.

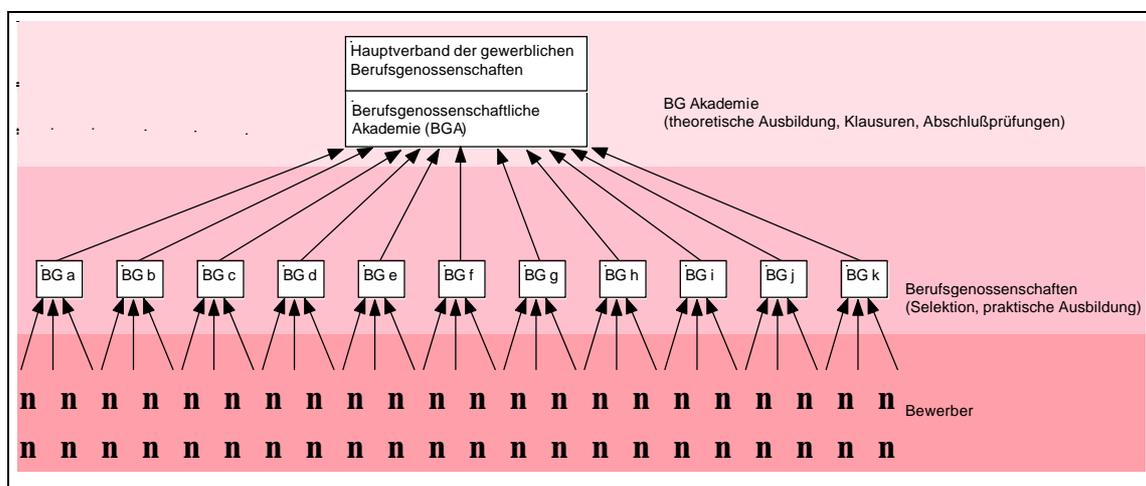


Abbildung 1: Ausbildungswesen der gewerblichen BGen

3.3 Die Praxis der Auswahl geeigneter Anwärter in den BGen

Unter Selektions- oder Auswahlmethode wird in dieser Arbeit das in der Berufsgenossenschaft praktizierte Verfahren zur differentiellen Diagnose der Berufseignung von Bewerbern für die gehobene Verwaltungslaufbahn verstanden. Um die gehobene Verwaltungslaufbahn im öffentlichen Dienst einschlagen zu können, benötigt ein Bewerber grundsätzlich die allgemeine oder fachgebundene Hochschulreife. Außer von Bewerbern mit Hochschulreife kann auch von Angehörigen des mittleren Dienstes nach entsprechender berufspraktischer Tätigkeit die gehobene Laufbahn eingeschlagen werden. Man kann also zwischen

internen und externen Bewerbern um die gehobene Verwaltungslaufbahn bei den BGen unterscheiden. An diese Unterscheidung sind Folgen geknüpft. So führt interne Rekrutierung nach den Befunden von Wanous (1992) und Premack & Wanous (1985) zu einer geringeren Fluktuation, was darauf hinweist, daß interne und externe Rekrutierung unterschiedliche "Typen" neuer Mitarbeiter erzeugen (Moser, 1995). Durch eine der Ausbildung vorausgehende praktische Tätigkeit in der Organisation kommt es nach Moser zu einer realistischen Einschätzung der Tätigkeit, die vor Enttäuschungen bewahre.

Interne Bewerber werden aber in der vorliegenden Arbeit nicht berücksichtigt, da sie in der Regel nicht das normale Auswahlverfahren durchlaufen, sondern sich über die berufspraktische Tätigkeit qualifizieren. Sie nehmen auch insofern eine Sonderstellung ein, als ihre Fortbildung nur zwei Jahre dauert, während die eingestellten Anwärter für die gehobene Laufbahn drei Jahre ausgebildet werden. Zur Frage, wie sich die Rekrutierungsquelle auf die berufliche Leistung auswirkt, siehe Moser (1992).

Bei den Bewerbern, die sich nicht über eine berufspraktische Tätigkeit qualifizieren, handelt es sich um Abiturienten, welche sich entweder direkt von der Schule oder nach ihrem Wehr- oder Ersatzdienst bewerben. Außerdem finden sich unter den Bewerbern Studienabbrecher und ehemalige Zeitsoldaten in nennenswertem Umfang. Dementsprechend liegen zum Zeitpunkt der Bewerbung außer den Schulzeugnissen in der Mehrheit der Fälle keine aussagekräftigen Unterlagen vor, aufgrund derer sich die Eignung beurteilen ließe. Die Schulnoten werden daher gerne zur Beurteilung der Eignung herangezogen. Diese Praxis kann sich auf den von Baron-Boldt und Schuler 1988 gefundenen Validitätskoeffizienten der Schulnoten von $r=0.373$ für die Ausbildungserfolgsprognose stützen.

Außer der Auswertung der übliche Bewerbungsunterlagen werden von den BGen Einstellungsgespräche geführt. Entweder geschieht dies, nachdem zuvor ein Eignungstest durchgeführt wurde, oder alleine das Gesprächsergebnis entscheidet über die Einstellung des Bewerbers.

3.4 Praxisorientierung der Ausbildung

Unter "Praxisorientierung der Ausbildung" wird in dieser Untersuchung das Ausmaß der Einbeziehung der Anwärter in die Sachbearbeitung innerhalb ihrer ausbildenden BG verstanden. Die Auszubildenden aller BGen verbringen etwa ein Drittel ihrer Ausbildungszeit an der BGA unter vergleichbaren Bedingungen. Hier werden sie in theoretischen Fächern wie "Grundzüge des bürgerlichen Rechts" oder "Finanzwesen der gewerblichen BGen" unterrichtet. Die restlichen zwei Drittel der Zeit werden sie in ihrer BG praktisch ausgebildet. Die 34 unterschiedlichen BGen unterscheiden sich in der Art und Weise, in der diese zwei Drittel der Zeit organisiert werden. Teilweise arbeiten die Auszubildenden unter Aufsicht und unter Anleitung erfahrener Sachbearbeiter an Aktenfällen. Manchmal wird ihnen sogar ein kleines Sachgebiet zugeteilt, oder sie übernehmen die Vertretung erkrankter oder beurlaubter Sachbearbeiter. In anderen BGen werden die Auszubildenden in einer Gruppe zusammengefaßt und ähnlich wie an der Akademie unterrichtet bzw. dazu angehalten, sich im Eigenstudium mit den schriftlichen Lehrmaterialien auseinanderzusetzen. Teilweise werden sogar besondere Seminare, etwa zur mentalen Prüfungsvorbereitung, durchgeführt.

Insbesondere kleinere Verwaltungen integrieren die Anwärter mehr in die praktische Arbeit, da hier die personellen Ressourcen zur Unterhaltung einer eigenen Ausbildungsabteilung nicht zur Verfügung stehen, bzw. aufgrund der geringen Anzahl von Dienstanwärttern ein aufwendiges Ausbildungswesen, wie es in den großen BGen existiert, nicht ökonomisch ist. Die Anwärter auf die gehobene Laufbahn werden an der BGA also unter vergleichbaren Bedingungen von denselben Personen theoretisch unterrichtet und schließlich geprüft. Die praktische Tätigkeit in ihren Verwaltungen sind jedoch unterschiedlich.

Da die Abschlußprüfung an der BGA ihren Schwerpunkt auf das Abfragen juristischen Wissens legt und insbesondere die Fähigkeit zur Rechtsanwendung prüft, fordert sie die eher theoretische und abstrahierende Fähigkeit, Sachverhalte unter Gesetzesregelungen zu subsumieren. Diese Fähigkeiten werden bei den weniger praktisch und eher an Lehrheften und Klausurfällen ausgebildeten

Dienstanwärtlern besser trainiert sein als bei den stark in die praktische Sachbearbeitung eingebundenen.

Da eine praxisorientierte Ausbildung eher die Fähigkeit zu konkretem praxisrelevantem Handeln trainiert, zum Beispiel die Bedienung des berufsgenossenschaftlichen Computersystems oder das richtige Ausfüllen der in der Verwaltung genutzten Formulare, wird eine solche Ausbildung auch andere Anforderungen an die Anwärter stellen als eine eher theoretische.

Unterschiedliche Auswahlinstrumente könnten aufgrund ihrer Konzeption jeweils andere Anwärter auswählen. Mittels Testverfahren werden zumeist kognitive Leistungen, wie Intelligenz und Konzentrationsfähigkeit, gemessen. Interviews dagegen, besonders in Form freier Gespräche, bei denen man sich "ein Bild" vom Bewerber machen möchte, wählen eher nach sozialer Angepaßtheit, Fähigkeit zur Selbstdarstellung und ähnlichen Merkmalen aus. Es kann also vermutet werden, daß zwischen der Auswahlmethode, die unterschiedliche Fähigkeiten diagnostiziert, und der Praxisorientierung der Ausbildung, die verschiedene Anforderungen stellt, eine Wechselwirkung besteht.

3.5 Leistungsdruck und Einstellung zur Arbeit als Eigenschaften der Organisation

Die Ausbildungsleistung ist wie alles menschliche Verhalten Ergebnis eines komplexen Zusammenwirkens persönlicher Eigenschaften und Umweltbedingungen. Dem tragen Organisationen dadurch Rechnung, daß sie auf dem Wege der Selektion versuchen, Individuen aufzunehmen, die erfolversprechende Eigenschaften haben und diese in einer Umwelt zu sozialisieren, die eine optimale Entwicklung der Leistungsfähigkeit gewährleistet. Daher wird in dieser Evaluation nicht nur die Selektion als Voraussetzung für die Ausbildungsleistung untersucht, also die Frage, inwieweit es den BGen gelingt, geeignete Individuen auszuwählen, sondern auch Umweltbedingungen miteinbezogen.

Der schon besprochene Aspekt der Praxisorientierung der Ausbildung ist ein solcher Umweltfaktor. Um eine Reduzierung der Fehlervarianz innerhalb eines varianzanalytischen Modells zu erreichen, wurden zwei weitere Variablen, nämlich der erzeugte Leistungsdruck und die Einstellung zur Arbeit, berücksichtigt. Dabei sind diese beide Größen geeignet, sowohl Individuen als auch Organisationen zu beschreiben. Leistungsdruck und Einstellung zur Arbeit sind ohne ein empfindendes Individuum nicht denkbar. So wie es diese Empfindung innerhalb von Individuen gibt, hat diese auch eine äußere Quelle: die Organisation. Bedingungen innerhalb der Umwelt "BG" sind also geeignet, in Individuen Leistungsdruck und Einstellung zur Arbeit in unterschiedlichen Ausprägungen hervorzurufen. Insofern kann man diese beiden Variablen als Eigenschaften einer Organisation begreifen und zu ihrer Beschreibung verwenden.

Motivationale Variablen haben im Rahmen der Varianzaufklärung von Leistungen erhebliche Bedeutung. Woodworth (zit. nach Heckhausen, 1988, S.455) drückte dies 1918 so aus:

".....when it is running by it's own drive it can run freely and effectively." .

Gegenstand dieser Arbeit ist aber nicht die nähere Aufklärung der Bedeutung motivationaler Variablen für die berufsgenossenschaftliche Ausbildungspraxis. Soweit die Einstellung zur Arbeit oder Leistungsdruck hier thematisiert werden, geschieht dies nicht aus einem inhaltlichen Interesse am Konstrukt heraus, sondern um den Zusammenhang zwischen dem Ausbildungserfolg und der Auswahlmethode bzw. Praxisorientierung der Ausbildung zu klären. Um die bei Feldstudien meist geringe interne Validität zu erhöhen, wird versucht, den Anteil der Varianz zu kontrollieren, der auf die Einstellung zur Arbeit und den Leistungsdruck zurückgeht. *Unter Einstellung zur Arbeit wird in dieser Untersuchung der Grad der positiven Einstellung zur Ausbildung verstanden.* Mit Brandstätter (1993, S.226) wird Einstellung zur Arbeit als Synonym für Arbeitszufriedenheit aufgefaßt.

Die ungünstige Entwicklung am Arbeitsmarkt führt zunehmend dazu, daß Anwärter nach bestandener Prüfung an der BGA nicht von ihrer ausbildenden BG in den

Verwaltungsdienst übernommen werden. Da die berufsgenossenschaftliche Ausbildung aus formalen Gründen nicht zu einem allgemein anerkannten Abschluß führt, ist der Arbeitsmarkt dieser Absolventen auf die BGen, allenfalls auf andere Träger der gesetzlichen Unfallversicherung, wie Gemeindeunfallversicherungsverbände, beschränkt. Nicht selten wird die Übernahme in den berufsgenossenschaftlichen Verwaltungsdienst vom Erreichen einer bestimmten Abschlußnote in der Prüfung an der BGA abhängig gemacht. Man verspricht sich davon eine Steigerung der Leistungen der Anwärter. Die Angst vor einer ungewissen Zukunft, vor der vergebenen Mühe einer dreijährigen Ausbildung, an deren Ende dann keine Übernahme in den berufsgenossenschaftlichen Dienst und darüberhinaus eine sehr eingeschränkte Möglichkeit steht, sich andernorts zu bewerben, wird hierbei als Quelle der Motivation instrumentalisiert. Auch hier lassen sich zwischen den BGen Unterschiede ausmachen. Wenn einige BGen auch keine Übernahmegarantie aussprechen, kann man sich als Dienstanwärter relativ sicher sein, nach erfolgreichem Bestehen der Abschlußprüfung in den Dienst übernommen zu werden. Andere BGen halten sich diese Entscheidung bis zuletzt offen oder informieren sogar darüber, daß nicht alle Anwärter übernommen werden können. *Unter Leistungsdruck wird in dieser Untersuchung daher das Ausmaß verstanden, in dem die ausbildende BG ihren Dienstanwärtern das Bewußtsein vermittelt, das Erreichen einer bestimmten Abschlußnote sei für das weitere Fortkommen oder für die Übernahme in den Dienst notwendig.*

4 Methode

Soweit in dieser Untersuchung Daten auf Personenebene erhoben wurden, geschah dieses mit dem Ziel, Eigenschaften der Organisation, der diese Personen angehören, zu messen. Dieser Weg über die Berechnung von Kennwerten für Organisationen wurde gewählt, weil bei einer Analyse der Daten auf Personenebene, die Konzepte von BGen mit vielen Auszubildenden stärker gewichtet würden als die von BGen mit wenigen Auszubildenden. Da Gegenstand dieser Evaluation die von den BGen praktizierten Selektions- und Ausbildungskonzepte sind, deren Güte aber nicht von der Zahl der Anwärter

abhängig ist (auch kleine BGen können sorgfältig auswählen und ausbilden), muß die Analyse auf Organisationsebene stattfinden.

4.1 Unabhängige Variablen

Willkürlich manipulierbare Größen konnte es unter den Bedingungen, unter denen diese Untersuchung stattfand, nicht geben. Die Anwärter bewarben sich Jahre vor dieser Untersuchung um ihre Ausbildungsplätze und durchliefen die Selektionsschritte, die zur Anstellung notwendig waren. Sie führten ihre Ausbildung durch und schlossen sie an der BGA mit der Prüfung zum gehobenen Dienst ab. Dies alles geschah, bevor diese Untersuchung durchgeführt wurde. Es fand also eine Selbstselektion der Personen auf die Bedingungen statt. Dasselbe gilt für den zweiten Teil der Untersuchung, in dem Daten per Fragebogen erhoben wurden.

In der Datenanalyse des Notenarchivs wird die berufsgenossenschaftliche Zugehörigkeit der Anwärter und damit verbunden die Selektionsmethode, mit deren Hilfe sie als geeignet ausgewählt wurden, als unabhängige Variable im Sinne eines Prädiktors betrachtet. Im varianzanalytischen Modell werden die Selektionsmethode, die Praxisorientierung der Ausbildung, der erzeugte Leistungsdruck sowie die Einstellung zur Arbeit als unabhängige Variablen betrachtet.

4.1.1 Operationalisierung der Variablen Selektion

Der Faktor Selektion wurde durch die beiden Stufen Test und Interview operationalisiert. *Unter die Test-Bedingung wurden dabei, abweichend von der in Abschnitt 3.1.2 zitierten Definition Lienerts, alle Verfahren subsumiert, die in standardisierter Form zur Messung von Eignung durchgeführt wurden.* Es kann sich dabei um Aufsätze, selbst konstruierte Tests, psychodiagnostische Verfahren oder Assessment-Center-ähnliche Verfahren handeln. Ihre Wissenschaftlichkeit wurde dabei nicht berücksichtigt. Unter der Interview-Bedingung wurden alle Verfahren zusammengefaßt, bei denen lediglich in Form eines freien Einzel- oder

Gruppengesprächs ohne die Durchführung standardisierter Verfahren die Entscheidung über die Einstellung getroffen wurde.

4.1.2 Operationalisierung der Variablen Praxisorientierung

Die Praxisorientierung der Ausbildung als Kennwert einer Organisation wurde durch die durchschnittliche Zustimmung aller im Beobachtungszeitraum von dieser BG ausgebildeten und im Rahmen dieser Untersuchung befragten Anwärter auf die Aussage "Ich war wenig in die laufende Sachbearbeitung eingebunden" operationalisiert. Diese Variable kann Werte zwischen 0 und 5 annehmen. Ein hoher Punktwert entspricht einer starken Praxisorientierung.

4.1.3 Operationalisierung der Variablen Leistungsdruck

Der erzeugte Leistungsdruck als Kennwert der Organisation wurde durch den durchschnittlichen Testwert aller im Beobachtungszeitraum von einer BG ausgebildeten und in dieser Untersuchung befragten Anwärter auf der im Rahmen dieser Untersuchung konstruierten Skala Leistungsdruck operationalisiert. Die Variable konnte Werte von 0 bis 35 annehmen, wobei ein hoher Wert einem großen Leistungsdruck entspricht.

4.1.4 Operationalisierung der Variablen Einstellung zur Arbeit

Einstellung zur Arbeit als Kennwert der Organisation wurde durch den durchschnittlichen Punktwert aller im Beobachtungszeitraum von einer BG ausgebildeten und in dieser Untersuchung befragten Anwärter auf der im Rahmen dieser Untersuchung konstruierten Skala Einstellung zur Arbeit operationalisiert. Auf

dieser Skala waren Werte von 0 bis 25 möglich. Je mehr Punkte erreicht wurden, desto positiver war die Einstellung zur Arbeit.

4.2 Abhängige Variablen

Als abhängige Variable wird die Ausbildungsleistung der Anwärter angesehen. Als Leistungsmaße kamen insbesondere in Betracht:

1. die Note der schriftlichen Prüfungen, die sich aus mehreren Klausurnoten in verschiedenen juristischen Fächern zusammensetzen
2. die Note der mündlichen Prüfungen, die die Leistung der Anwärter in einer 150 minütigen Gruppenprüfung (3Teilnehmer) wiedergibt sowie
3. die Gesamtnote, die ein gewichtetes Mittel der mündlichen (mit Faktor 2 gewichtet) und schriftlichen (mit Faktor 3 gewichtet) Note ist
4. Verhältnis der eingestellten Anwärter zu der Zahl der die Ausbildung abbrechenden Anwärter (Abbrecherquote)
5. Verhältnis der zur Prüfung gemeldeten Anwärter zu den durchgefallenen Anwärtern (Wiederholer)

Andere Maße, wie etwa die Vorgesetztenbeurteilung, sind innerhalb einer Diplomarbeit wegen des damit verbundenen Aufwands nicht nutzbar.

Aufgrund des Zustandekommens der Noten wird die schriftliche Note als Ergebnis mehrerer Prüfungsklausuren, die sich insgesamt über rund eine Woche hinziehen und von verschiedenen Personen bewertet werden, als besonders valide und reliabel betrachtet. Die mündlichen Noten erfüllen diese Kriterien in geringerem Ausmaß, da für jeden Prüfling nur etwa 50 Minuten Prüfungszeit zur Verfügung stehen, die Prüfer über die schriftliche Note der Prüflinge informiert sind und Praxis ist, daß mittels der mündlichen Note "Grenzfälle" angehoben werden, die an der Schwelle zur nächsthöheren Note stehen oder aber knapp an den mindestens zum Bestehen der Prüfung notwendigen 50 Punkten zu scheitern drohen.

Die Gesamtnote als gewichtetes Mittel der schriftlichen und mündlichen Note würde die zweifelhafte Reliabilität der mündlichen Note beinhalten und kommt daher nicht in Betracht.

Die abhängige Variable Ausbildungsleistung als Kennwert einer Organisation wurde durch den arithmetischen Mittelwert der schriftliche Note aller Anwärter einer BG in der Abschlußprüfung an der BGA in den Jahren 1994 und 1995 aufgefaßt.

Der Ausbildungserfolg, den eine BG erzielt, bildet sich jedoch nicht nur in den Noten der bis zur Prüfung gelangten Anwärter dieser BG ab, sondern auch in der Zahl der eingestellten Anwärter, die vor Erreichen der Prüfung abbrechen und der Zahl der Anwärter, die den Prüfungsanforderungen nicht genügen.

Bei der Analyse der Noten bleiben all jene Anwärter, die durch die Prüfung fallen, unberücksichtigt, da deren Noten (unter 50 Punkte) grundsätzlich nicht archiviert sind. Eine Ausnahme bilden hier nur sehr wenige Personen, die aufgrund einer Einzelfallentscheidung auch mit einer knapp unter 50 Punkten liegenden schriftlichen Leistung zur mündlichen Prüfung zugelassen wurden. Insofern handelt es sich hier um eine nicht unerhebliche statistische Mortalität, da 10.9% aller zur Prüfung angemeldeten Anwärter der Jahre 1994/95 die Prüfung nicht bestanden und für diese Personen damit auch keine Noten archiviert sind. Außerdem werden einige Anwärter nie geprüft, weil sie vorzeitig ihre Ausbildung abbrechen. Dies kann aus ganz unterschiedlichen Gründen erfolgen: Schwangerschaft, Tod, schwere Krankheit, Aufnahme eines Studiums, Unzufriedenheit, Unter- und Überforderung sind hier denkbar. Zwischen der Wahrscheinlichkeit von Abbrüchen und der Selektionsmethode könnten Zusammenhänge bestehen, so daß ohne deren Berücksichtigung ein falsches Bild über die Wirksamkeit der Selektionsmethode erzeugt würde.

Daher werden auch die Variablen Abbrecherquote als Verhältnis der Zahl der Anwärter einer BG, die ihre Ausbildung vor Erreichen der Prüfung beenden, zur Zahl der eingestellten Anwärter dieser BG und Anzahl der durchgefallenen Anwärter als Verhältnis der Zahl der Anwärter einer BG, die die Prüfung an der BGA nicht

bestanden, zur Zahl der zur Prüfung angemeldeten Anwärter einer BG als Operationalisierungen des Ausbildungserfolgs betrachtet.

4.3 Hypothesen

Die Evaluation der Ausbildungs- und Selektionspraxis der BGen wird anhand der Prüfungsjahrgänge 1994/95 durch Prüfung der folgenden Hypothesen vorgenommen:

H1

BGen, die mit Testverfahren die Eignung von Bewerbern um die gehobene Verwaltungslaufbahn feststellen, erzielen im Durchschnitt eine bessere Ausbildungsleistung als BGen, die per Interview auswählen.

Diese Hypothese wird folgendermaßen operationalisiert:

H1.1

BGen, die mit Testverfahren die Eignung von Bewerbern um die gehobene Verwaltungslaufbahn feststellen, erzielen durchschnittlich bessere schriftliche Abschlußnoten als BGen, die durch Interviews auswählen.

H1.2

BGen, die mit Testverfahren die Eignung von Bewerbern um die gehobene Verwaltungslaufbahn feststellen, haben eine geringere Abbrecherrate als BGen, die durch Interviews auswählen.

H1.3

BGen, die mit Testverfahren die Eignung von Bewerbern um die gehobene Verwaltungslaufbahn feststellen, haben eine geringere Rate von Anwärtern, die den Prüfungsanforderungen nicht genügen.

H2

Je praxisorientierter eine BG ausbildet, desto schlechter ist die durchschnittliche schriftliche Note ihrer Dienstanwärter in der Abschlußprüfung.

H3

Der Zusammenhang zwischen der Auswahlmethode und der Ausbildungsleistung wird durch den Grad der Praxisorientierung der Ausbildung in der Weise moderiert, daß durch Tests ausgewählte Personen bei hoher Praxisorientierung keine besseren schriftlichen Noten erreichen als durch Interview ausgewählte, während bei geringer Praxisorientierung die durch Testverfahren ausgewählten Personen bessere schriftliche Noten erzielen.

H4

Die Variablen Praxisorientierung, Einstellung zur Arbeit und Leistungsdruck stehen mit dem Ausbildungserfolg in korrelativem Zusammenhang und eignen sich daher zur Reduzierung der Fehlervarianz in einem kovarianzanalytischen Modell.

Diese Hypothese wird folgendermaßen operationalisiert:

H4.1

Praxisorientierung der Ausbildung, Leistungsdruck und Einstellung zur Arbeit korrelieren mit der schriftlichen Note.

H4.2

Die Varianzaufklärung der schriftlichen Noten durch die beiden Stufen des Faktors Selektion wird sowohl durch die einzelne, als auch durch die kombinierte Aufnahme von Leistungsdruck, Praxisorientierung und Einstellung zur Arbeit als Kovariaten verbessert.

4.4 Design

In dieser Untersuchung sollen bereits abgeschlossene Vorgänge in ihrer Wirksamkeit evaluiert werden. Nach Wottawa & Thierau (1990) handelt es sich damit um eine sog. "summative" Evaluation. Die Noten der Anwärter erhielten diese am Ende ihrer Ausbildung und nachdem sie durch Test oder Interview ausgewählt

wurden. Zum Untersuchungszeitpunkt war ihre Ausbildung beendet, die Noten erzielt. Daher entspricht diese Untersuchung in ihrem Design einer Ex-post-facto Anordnung. Es konnten keine unabhängige Variable willkürlich manipuliert und keine Randomisierung durchgeführt werden.

Durch die Selbstselektion der Personen auf die Bedingungen (BG verbunden mit bestimmtem Selektionsmechanismus) ist nach streng wissenschaftlichen Maßstäben die Aussagefähigkeit der Ergebnisse eingeschränkt, da die Äquivalenz der Gruppen nicht durch Randomisierung oder Parallelisierung kontrolliert werden konnte. Zum Verhältnis von Evaluation und Wissenschaft, speziell aber zur Realisierbarkeit von Idealforderungen in der praktischen Durchführung von Evaluationen, wird in dieser Untersuchung mit Wottawa & Thierau (1990, S.123ff) eine maximale Annäherung an wissenschaftliche Standards angestrebt.

4.5 Stichprobe

Es wird versucht, eine Totalerhebung der innerhalb dieser Untersuchung relevanten Daten aller gewerblichen BGen durchzuführen. In § 646 Abs. I Reichsversicherungsordnung (RVO), die in diesen Monaten durch das 7. Buch des Sozialgesetzbuchs (SGB) ersetzt wurde, sind 36 BGen durchnummeriert. Nummer 13 existiert nicht mehr, Nummer 14, 16 und 20 bilden eine Verwaltungsgemeinschaft, so daß sie im Rahmen dieser Untersuchung als ein Untersuchungsobjekt zu betrachten sind. Eine BG bildete im Beobachtungszeitraum nicht aus. Die See BG gehört aufgrund organisatorischer Besonderheiten nicht in den engeren Kreis der gewerblichen BGen, sondern bildet mit der Kranken- und Rentenversicherung der Seeleute einen Versicherungsträger. Sie wird deswegen nicht in die Untersuchung aufgenommen. So verbleiben 31 Untersuchungsobjekte in der Grundgesamtheit.

Soweit der Ausbildungserfolg einer BG im Beobachtungszeitraum gemessen werden sollte, wurden aus dem Notenarchiv der BGA die Noten aller Anwärter dieser BG gemittelt und der BG als Kennwert zugeordnet. Die Anzahl der Anwärter einer BG,

die den Prüfungsanforderungen nicht genügen konnten, wurde ebenfalls durch Auswertung des Notenarchivs ermittelt. Die Anzahl der Anwärter einer BG, die vor Erreichen der Prüfung die Ausbildung abbrachen, wurde durch einen Abgleich der von dieser BG zum ersten Lehrgang an der BGA angemeldeten Anwärter mit den nach drei Jahren zur Prüfung angemeldeten Anwärtern der BG festgestellt. In all diesen Punkten handelt es sich um eine Totalerhebung.

Die Erhebung der Daten zur Praxisorientierung der Ausbildung, der Einstellung zur Arbeit und des Leistungsdrucks geschah durch Versenden von Fragebögen an die BGen. Da hier nicht sichergestellt werden konnte, daß jeder in den Jahren 1994/95 geprüfte Anwärter einen Fragebogen erhielt und jeder ausgeteilte Fragebogen ausgefüllt zurückgesandt wurde, handelt es sich insoweit um eine Stichprobe aus der Grundgesamtheit der 1994/95 geprüften Anwärter.

4.6 Untersuchungsablauf

In einer schriftlichen Befragung der BGen (vgl. Abschnitt 4.6.1), wurde die Praxis der Auswahl geeigneter Anwärter auf die gehobene Verwaltungslaufbahn aufgeklärt. Zur Erfassung der abhängigen Variablen (Note) wurden dann in einem zweiten Schritt Daten im Notenarchiv der BGA erhoben (vgl. Abschnitt 4.6.2). Eine zweite schriftliche Befragung (vgl. Abschnitt 4.6.3) richtete sich an ehemalige Anwärter. Auf Basis der in dieser Befragung erhaltenen Angaben, die mit Hilfe der entwickelten Instrumente zur Messung der Variablen Leistungsdruck, Einstellung zur Arbeit und Praxisorientierung (vgl. Abschnitte 4.6.3.1-4.6.3.4) erhoben wurden, erfolgte die Berechnung der Kennwerte für die einzelnen BGen. Außerdem wurden in einer telefonischen Befragung der in den BGen mit der Personalauswahl betrauten Mitarbeiter weitere Details der Personalauswahl erhoben sowie unklare Angaben aus der ersten schriftlichen Befragung geklärt (vgl. Abschnitt 4.6.4).

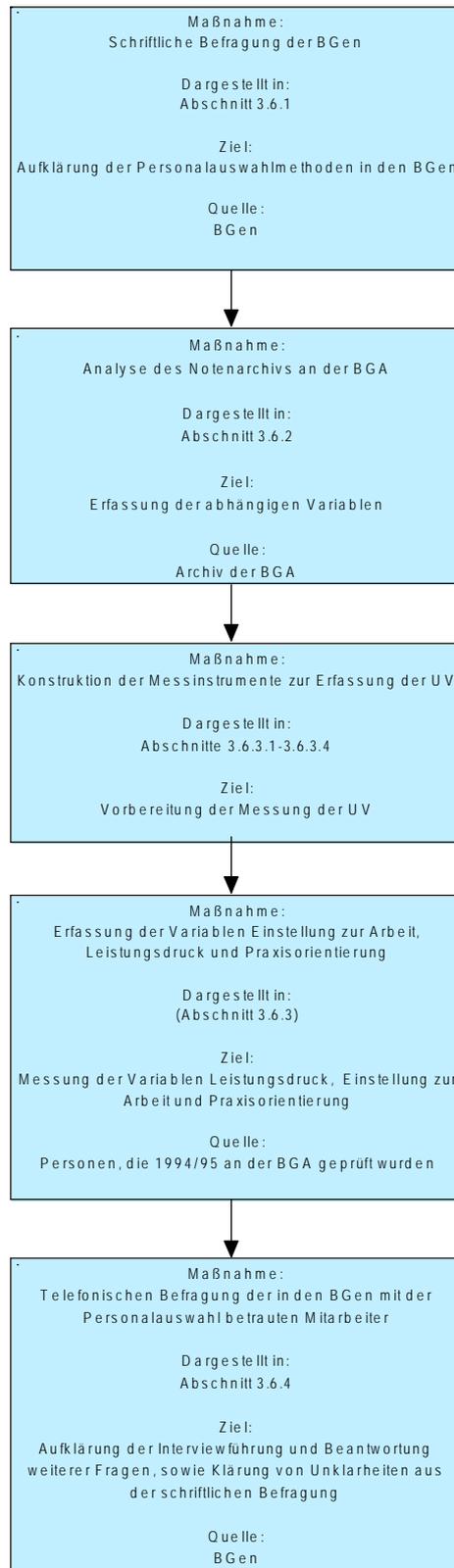


Abbildung 2: Ablauf der Datenerhebung

4.6.1 Schriftliche Befragung der BGen

Zunächst wurden die BGen schriftlich zu ihrer Einstellungspraxis befragt (Fragebogen siehe Anhang A), um einen Überblick über verschiedene Einstellungspraktiken zu bekommen. Dabei wurde festgestellt, ob per Assessment-Center, Einstellungstest oder Interview selektiert wird, welche Unterlagen zur Vorauswahl herangezogen werden, ob auch interne Bewerber berücksichtigt werden und ggf. ob diese dieselbe Auswahlprozedur durchlaufen wie externe Bewerber. Außerdem wurde um die Angabe einer Kontaktperson für die spätere telefonische Befragung und die Übersendung der Fragebögen gebeten.

4.6.2 Erfassung der abhängigen Variablen

An der BGA wurde das Notenarchiv für die Prüfungsjahrgänge 1994/95 erfaßt. Aus Datenschutzgründen wurden die Namen der Prüflinge dabei außer Acht gelassen. Aufgenommen wurde das Geschlecht, die ausbildende BG und Bezirksverwaltung, das Geburtsjahr, die mündliche und schriftliche Note, die Gesamtnote sowie das Prüfungsjahr.

Durch einen Vergleich der Anmeldeliste der BGen zum ersten Lehrgang 1991 und 1992 und der in den Jahren 1994 und 1995 erfolgten Anmeldungen zur Prüfung, wurden die Anwärter identifiziert, die zwischenzeitlich die Ausbildung abgebrochen hatten. Dieses indirekte Vorgehen war notwendig, da keine Aufzeichnungen über Personen, die vor Erreichen der Prüfung die Ausbildung abbrechen, geführt werden.

Bei diesem Namensvergleich konnten drei Ereignisse stattfinden:

1. Eine Person war zum ersten Lehrgang angemeldet, aber nach drei Jahren nicht zur Prüfung.
2. Eine Person war zur Prüfung angemeldet, aber nicht zum ersten Lehrgang.
3. Eine Person war sowohl zum ersten Lehrgang als auch nach drei Jahren zur Prüfung angemeldet.

Die dritte Möglichkeit ist der Normalfall. Man beginnt die Ausbildung mit dem Besuch des ersten Lehrgangs und schließt sie drei Jahre später mit der Prüfung ab. Fallgestaltung zwei trifft auf alle Personen zu, die sich über eine berufspraktische Tätigkeit zum Aufstieg in den gehobenen Dienst qualifiziert haben. Diese Personen werden nur zwei Jahre lang fortgebildet und nehmen nicht an den ersten Lehrgängen teil. Ihre Namen tauchen nur in den Anmeldungen zur Prüfung, nicht aber in denen zum ersten Lehrgang auf. Personen die zum ersten Lehrgang gemeldet waren, nicht aber zur Prüfung, wurden als Ausbildungsabbrecher eingestuft.

Bei diesem Vorgehen gibt es Ungenauigkeiten. So ist dem Autor z.B. ein Fall bekannt geworden, in dem eine Person sich bei zwei BGen beworben hat, von beiden eine Zusage erhielt und doppelt zum ersten Lehrgang angemeldet wurde. Es ist im Einzelfall auch nicht auszuschließen, daß Anwärter ihre Ausbildung unterbrechen, um sie wieder aufzunehmen, oder daß eingestellte Anwärter die Ausbildung nicht aufnehmen. Diese Fälle wären im eigentlichen Sinne keine Ausbildungsabbrecher. Da diese seltenen Einzelfälle sich aber zufällig auf die BGen verteilen dürften und hier nur zwischen den beiden Gruppen von BGen unterschieden wird, die per Test oder Interview auswählen, ist keine systematische Datenverzerrung zu befürchten.

4.6.3 Zweite schriftliche Befragung

4.6.3.1 Entwicklung der Instrumente zur Messung der Variablen Leistungsdruck, Einstellung zur Arbeit und Praxisorientierung

Nach der Erfassung der abhängigen Variablen wurde ein Fragebogen zur Messung der Einstellung zur Arbeit, Praxisorientierung und Leistungsdruck, die als potentielle Kovariaten in das varianzanalytische Modell eingefügt werden sollen, konstruiert.

Außerdem wurden in diesem Fragebogen Daten zur Person (Geburtsjahr, Prüfungsjahr, Geschlecht, ausbildende BG, Abschlußnote, Selektionsmethode) erhoben und zwei eher exploratorische Fragen ("Am besten fand ich in meiner Ausbildung, daß ...", " Am meisten hat mich in meiner Ausbildung gestört, daß ...") mit freiem Beantwortungsmodus gestellt. Die Frage nach der ausbildenden BG und der durchlaufenen Selektionsmethode war notwendig, weil innerhalb der gewerblichen BGen, aber auch mit den landwirtschaftlichen BGen und Gemeindeunfallversicherungsverbänden und zum Teil auch mit den Krankenkassen bzw. Rentenversicherungsträgern ein Personalfluß besteht. Nicht jeder Absolvent der BGA arbeitet später auch in der BG, die ihn ausgebildet hat.

4.6.3.2 Konstruktion der Leistungsdruck-Skala

In Absprache mit Experten (Anwärter, Sachbearbeiter) und aus eigener Erfahrung als ehemaliger Anwärter heraus wurden folgende Items zur Quantifizierung des Leistungsdrucks erstellt (invertierte Items sind durch (-) gekennzeichnet):

Item L1 "Ich ging damals davon aus, daß mein weiteres Fortkommen in der BG-Laufbahn von meiner Abschlußnote an der BGA abhängt."

Item L2 "Ich ging damals davon aus, daß meine Übernahme vom Erreichen einer bestimmten Prüfungsnote an der BGA abhängt."

Item L3 "Meine ausbildende BG legte sehr großen Wert auf gute Noten an der BGA."

Item L4 "Ich stand während meiner Ausbildung unter großem Leistungsdruck."

Item L5 "Meine Leistungen wurden von meinen Vorgesetzten genau beobachtet."

Item L6 (-) "Man hat in der Ausbildung keinerlei Druck auf mich ausgeübt."

Item L7 "Meine Übernahme in den berufsgenossenschaftlichen Dienst nach Abschluß der Fortbildung war ungewiß."

Die Antworten konnten anhand der folgenden sechsstufigen Antwortskala gegeben werden, die von "Trifft vollkommen zu" bis "Trifft überhaupt nicht zu" reichte (vgl. Abbildung 3).

Trifft überhaupt nicht zu	Trifft im wesent- lichen nicht zu	Trifft eher nicht zu	Trifft etwas zu	Trifft im wesent- lichen zu	Trifft vollkommen zu
()-----	()-----	()-----	()-----	()-----	()-----

Abbildung 3: Antwortskala

Eine Antwort im Sinne des Konstrukts wurde mit 5 Punkten, eine gegen das Konstrukt mit 0 Punkten gewertet. Je höher ein erreichter Testscore ist, desto mehr Leistungsdruck wurde empfunden. Durch die Bereitstellung mehrerer Antwortkategorien wurde anerkannt, daß viele Items nur schwer alternativ beantwortet werden können. Durch die Wahl einer geraden Anzahl von Kategorien ist andererseits eine Entscheidung forciert worden. Eine zentrale Tendenz des Urteils führt also in keinem Fall zu einem indifferenten Ergebnis.

4.6.3.3 Konstruktion der Skala Einstellung zur Arbeit

In Absprache mit Experten und aus eigener Erfahrung als ehemaliger Anwärter heraus wurden folgende Items erstellt:

Item M1 "Die Inhalte meiner Ausbildung interessierten mich."

Item M2 "Meine Ausbildung machte mir Spaß."

Item M3 "Man bemühte sich in meiner Ausbildung darum, mich zu fördern."

Item M4 "Ich ging während der Ausbildung gerne zur Arbeit in die Verwaltung."

Item M5 "Man half mir dabei, meine Leistungen zu verbessern."

Für den Beantwortungsmodus dieser Skala gilt das für die Skala Leistungsdruck ausgeführte. Auch hier wurde eine Antwort im Sinne des Konstrukts "Positive Einstellung zur Arbeit" mit 5 Punkten gewertet, so daß ein hoher Testscore eine besonders positive Einstellung zur Arbeit widerspiegelt.

4.6.3.4 Praxisorientierung

Die Konstruktion einer Skala zur Erfassung des Grades der Praxisorientierung der Ausbildung gelang in einer Vorstudie zu dieser Untersuchung nicht. Es kam keine reliable Skala zustande. Um diese Variable zu messen, wurde das Item "Ich war wenig in die laufende Sachbearbeitung eingebunden." als Operationalisierung des Konstrukts betrachtet. Eine Antwort im Sinne des Konstrukts bestand hier in einer Verneinung dieser Aussage. Zunächst sollte das Konstrukt mit drei Items abgefragt werden. Hier wurde eine Antwort im Sinne des Konstrukts "Praxisorientierung der Ausbildung" mit 5 Punkten gewertet, so daß eine besonders praxisorientierte Ausbildung sich in hohen Punktwerten niederschlagen sollte.

Nach einem einleitenden Text über den Zweck der Befragung wurden die Items aller Skalen aufgeführt, die auf einer sechsstufigen Likert-Skala zu beantworten waren. Die dritte Seite enthielt drei Ja/Nein-Fragen, eine Wahlaufgabe, zwei Fragen mit freiem Beantwortungsmodus und die Fragen zur Person. Ein Fragebogen ist als Anhang B beigefügt.

Die Fragebögen wurden zu Händen der Kontaktperson gesandt, die in der ersten schriftlichen Befragung der BGen angegeben und mit der auch das Telefoninterview geführt wurde. Diese Kontaktperson sorgte für die Weiterleitung an Sachbearbeiter, die in den Jahren 94/95 ihre Abschlußprüfung an der BGA abgelegt hatten.

Die Fragebögen erhoben an Daten zur Person die Prüfungsnote an der BGA, das Geburtsdatum, Prüfungsjahr, Geschlecht und ausbildende BG. Außerdem enthielt der Fragebogen die Skalen zur Erfassung der Konstrukte Leistungdruck, Praxisorientierung und Einstellung zur Arbeit (vgl. Anhang B). Die Fragebögen wurden zu Händen der Kontaktperson geschickt, mit der Bitte, diese an Sachbearbeiter/innen weiterzuleiten, die in den Jahren 1994/95 die Abschlußprüfung an der BGA ablegten. Diese Zeitbegrenzung wurde aus zwei Gründen vorgenommen: Erstens um die Angaben der Befragten zuverlässig zu halten, da die Skalen sich auf die Vergangenheit bezogen und damit zu rechnen war, daß bei den älteren Sachbearbeitern Gedächtniseffekte auftraten. Der

Zeitfaktor wurde so als weitere Varianzquelle ausgeschlossen. Zweitens wurde so erreicht, daß die Ergebnisse der Notenarchivanalyse mit denen der Fragebogenaktion verglichen bzw. in Zusammenhang gesetzt werden konnten, da beide Datenquellen sich somit auf dieselben Prüfungsjahrgänge bezogen.

Die Auswertung der per Post oder Fax eingehenden Fragebögen wurde zur Vermeidung von Fehlern und damit zur Sicherung der Auswertungsobjektivität mit einer Schablone vorgenommen.

4.6.4 Telefonische Befragung der BGen

In einem weiteren Schritt wurden die BGen telefonisch mit folgenden Fragen zu Details der Interviewführung befragt:

- Was genau wollen Sie mit dem Interview erfassen ?
- Worauf achten Sie bei Bewerbern besonders (Kleidung o.ä.) ?
- Worin sehen Sie die besonderen Vorteile des Bewerbungsgesprächs?
- Wie verläuft die Auswertung des Gesprächs ?
- Sind die Abläufe / Fragestellungen in den Interviews standardisiert ?
- Wurde eine Arbeits- und Anforderungsanalyse durchgeführt ?

Die Beantwortung erfolgte frei. Zu einer ökonomischen Darstellung der Antworten, wurden bei der Auswertung Kategorien gebildet, deren Besetzung im Ergebnisteil dieser Arbeit in Tabellenform dargestellt ist.

Außerdem wurden die Angaben der BGen aus der ersten schriftlichen Befragung überprüft, insbesondere die Frage ob in den Jahren 1990-92 (Einstellungsjahre der untersuchten Prüfungsjahrgänge 1994/95) die Auswahlverfahren schon so waren, wie in der schriftlichen Befragung für die Gegenwart angegeben. Desweiteren wurde die Übersendung von Fragebögen angekündigt, die an die Sachbearbeiter weitergeleitet werden sollten, die in den Jahren 1994/95 ihre Abschlußprüfung an der BGA abgelegt hatten.

4.7 Verwendete Methoden und statistische Programmpakete

Zur Prüfung der Hypothesen wurden deskriptivstatistische Kennwerte, Produkt-Moment-Korrelationen und t-Tests berechnet. Die Güte der konstruierten Skalen wurde nach der klassischen Testtheorie berechnet. Außerdem wurde eine einfaktorielle, univariate Varianzanalyse sowie eine einfaktorielle, univariate, hierarchische Kovarianzanalyse mit bis zu drei Kovariaten durchgeführt. Alle Analysen wurden mit dem Statistikprogramm SPSS für Windows Version 5.0.1 berechnet. Für die Erstellung einiger Grafiken und Abbildungen wurde die Tabellenkalkulation Lotus 1-2-3 Version 4.0.1 verwendet.

5 Ergebnisse

5.1 Ergebnisse der schriftlichen Befragung der BGen

Von den 35 gewerblichen BGen wurden außer der See BG, die wegen organisatorischer Besonderheiten, welche die Vergleichbarkeit mit den anderen BGen eingeschränkt hätten, alle angeschrieben. Von den verbleibenden 34 BGen entfielen zwei, da die Lederindustrie-, Zucker- und die Papiermacher BG eine Verwaltungseinheit bilden. Eine BG bildete im Beobachtungszeitraum keine Anwärter aus. Es verblieben also 31 Untersuchungsobjekte. Vier BGen sandten den Fragebogen nicht zurück, so daß 27 Verwaltungen erfaßt werden konnten. Dies entspricht, gemessen an der Zahl der Untersuchungsobjekte, einem Rücklauf der Fragebögen von 87.1%.

Insgesamt gaben 10 BGen (37%) an, außer einem Interview auch ein oder mehrere Testverfahren einzusetzen (vgl. Tabelle 1). BG 1 & BG 33 führten in einzelnen Bezirksverwaltungen Testverfahren durch, nicht aber in der gesamten Organisation. Dabei arbeiteten (bezogen auf die Jahre 1990-92) drei BGen (11%) mit externen Stellen (Diplom-Psychologen, Personalberatungen) zusammen. 14 BGen (52%),

stellten Anwärter ausschließlich in ihren Hauptverwaltungen ein. 22 BGen (82%) bildeten auch interne Bewerber zur gehobenen Laufbahn fort. Diese internen Bewerbern mußten in nur drei BGen die gleichen Selektionsverfahren durchlaufen wie externe.

Tabelle 1: Ergebnisse der schriftlichen Befragung der BGen

	Absolutzahl	% bezogen auf 27 befragte BG'n
Einstellungstest	10 (+zwei BV'n)	37 (44)
Einstellung nur in Hauptverwaltung	14	52
externe Berater	3	11
Interview	27	100
Lebenslauf	23	85
Schulzeugnis	26	96
Bewerbungsschreiben	23	85
Empfehlung	4	15
"Vorschaltjahr"	1	0,04
Berücksichtigung interner Bewerber	22	82
gleiche Selektion für interne und externe Bewerber	3	11

Ein sog. "Vorschaltjahr", aus eignungsdiagnostischer Sicht handelt es dabei um eine Art Arbeitsprobe, wurde in einer BG durchgeführt. Die Bewerber um die gehobene Laufbahn lernen die Verwaltung zunächst von innen kennen, indem sie z.B. in der Registratur und anderen organisatorischen Einheiten zu einfachen Tätigkeiten angelernt werden. Weil diese BG innerhalb des Beobachtungszeitraums nur drei Anwärter stellte, war es sinnlos, diese Variante der Auswahl als dritte Stufe des Faktors Selektion dem Versuchsplan hinzuzufügen.

Alle BGen führten unabhängig von der Frage ob sie einen Test anwandten, ein Vorstellungsgespräch. 23 (85%) werteten den Lebenslauf aus und 26 (96%) das Schulzeugnis. Ebenfalls 23 (85%) analysierten das Bewerbungsschreiben und lediglich vier gaben an, auch auf Empfehlung einzustellen.

5.2 Auswertung des Notenarchivs der Jahre 1994/95

5.2.1 Noten und durchgefallene Anwärter

Tabelle 2 zeigt die Verteilung der Personen auf die Stufen des Faktors Selektion, die Geschlechter sowie die beiden Prüfungsjahre. Es wurden die beiden Prüfungsjahrgänge 1994/95 vollständig erfaßt. Tabelle 3 gibt die deskriptivstatistischen Kennwerte für die schriftliche, die mündliche Note und die sich daraus errechnende Gesamtnote sowie das Alter im Prüfungsjahr wieder. Die Notenpunkte beziehen sich dabei auf eine 100-Punkte-Skala, wobei die Mindestanforderung zum Bestehen der Prüfung bei 50 Punkten liegt.

Tabelle 2: Verteilung der Personen auf die Stufen des Faktors Selektion, die Geschlechter, sowie die beiden Prüfungsjahre

Value Label	Value	Frequency	Percent	Valid Percent
Interview	,00	569	71,5	71,5
Test	1.00	227	28,5	28,5
weiblich	,0	375	47,1	47,1
männlich	1.0	421	52,9	52,9
1994		404	50,8	50,8
1995		392	49,2	49,2
Valid cases	796	Missing cases	0	

Abbildung 3 zeigt die symmetrische und unimodale Verteilung der schriftlichen Noten.

Frequency	Stem &	Leaf
1.00	4 *	&
26.00	4 .	78899&
106.00	5 *	0000111222233333334444
169.00	5 .	5555556666677777888888999999
194.00	6 *	0000000011111112222223333334444444
128.00	6 .	5555556666677777888889999
65.00	7 *	0001112223344
21.00	7 .	5667&
1.00	8 *	&
1.00	Extremes	(81)

Stem width: 10
Each leaf: 5 case(s)
& denotes fractional leaves.

Abbildung 3: Stamm & Blatt-Darstellung für die schriftlichen Noten im Beobachtungszeitraum

Tabelle 3: Alter, schriftliche, mündliche Note und Gesamtnote im Beobachtungszeitraum

Variable	AGE	Alter		
Mean	25.235		S.E. Mean	.117
Std Dev	3.291		Variance	10.829
Kurtosis	2.384		S.E. Kurt	.173
Range	20.000		Minimum	21
Maximum	41			
Valid observations	-	796	Missing observations	- 0

Variable	SCHRIFTL	Note schriftlich		
Mean	61.142		S.E. Mean	.260
Std Dev	6.920		Variance	47.882
Kurtosis	-.455		S.E. Kurt	.183
Range	36.000		Minimum	45
Maximum	81			
Valid observations	-	709	Missing observations	- 87*

Variable	GESAMT	Note gesamt		
Mean	64.292		S.E. Mean	.246
Std Dev	6.545		Variance	42.834
Kurtosis	-.286		S.E. Kurt	.183
Range	32.000		Minimum	50
Maximum	82			
Valid observations	-	709	Missing observations	- 87*

Variable	MUENDL	Note mündlich		
Mean	68.959		S.E. Mean	.294
Std Dev	7.820		Variance	61.149
Kurtosis	.186		S.E. Kurt	.183
Range	49.000		Minimum	45
Maximum	94			
Valid observations	-	709	Missing observations	- 87*

* Die in den Statistiken ausgewiesenen 87 fehlenden Fälle sind die Anwärter, die durch die Abschlußprüfung gefallen sind, da keine Notenwerte für diese Fälle archiviert wurden.				

5.2.2 Ausbildungsabbrecher

Tabelle 4 setzt für jede BG die Zahl der zu den ersten Lehrgängen 1994/95 angemeldeten Anwärter zu der Zahl der ermittelten Abbrecher ins Verhältnis.

Tabelle 4: Ausbildungsabbrecher

BG	Anzahl der zum ersten Lehrgang angemeldeten Anwärter	Abbrecher	Abbrecher%
1	22	0	0
2	7	0	0
3	15	2	13,33
4	6	1	16,67
6	55	0	0
7	21	1	4,76
8	48	8	16,67
9	6	0	0
10	41	3	7,32
11	31	3	9,68
12	38	2	5,26
14,16, 20	5	0	0
15	9	1	11,11
17	5	3	60
18	40	3	7,5
19	8	0	0
21	9	2	22,22
22	34	3	8,82
23	19	2	10,53
24	13	2	15,38
25	7	0	0
26	15	2	13,33
27	13	1	7,69
28	22	2	9,09
29	52	2	3,85
30	27	0	0
31	86	7	8,14
32	6	0	0
33	51	5	9,8
35	1	0	0
36	93	10	10,75

Die Abbrecherrate bewegt sich in einem Bereich von 0 - 22.22%.

5.3 Ergebnisse der telefonischen Befragung der BGen

Von 27 BGen wurde auf die schriftliche Befragung hin ein Ansprechpartner benannt. Folgende Fragen konnten 23 dieser Personen, die alle an eignungsdiagnostischen Interviews als Beurteiler teilnahmen, gestellt werden:

- 1. Was genau wollen Sie mit dem Interview erfassen ?*
- 2. Worauf achten Sie bei Bewerbern besonders (Kleidung o.ä.) ?*
- 3. Worin sehen Sie die besonderen Vorteile des Bewerbungsgesprächs?*
- 4. Wie verläuft die Auswertung des Gesprächs ?*
- 5. Sind die Abläufe/Fragestellungen in den Interviews standardisiert ?*
- 6. Wurde eine Arbeits- und Anforderungsanalyse durchgeführt ?*

Bei den Fragen 1-4 waren Mehrfachnennungen möglich.

Es stellte sich heraus, daß Einstellungsgespräche im allgemeinen in relativ unstrukturierter Weise durchgeführt und ausgewertet werden (vgl. Tabelle 7), obwohl dem Anspruch nach eine ganze Reihe von Persönlichkeitsmerkmalen erfaßt werden soll (vgl. Tabelle 9).

Anhand welcher Indikatoren die in Tabelle 9 genannten Eigenschaften zu messen sind, blieb der Intuition der Gesprächsteilnehmer überlassen. Explizite Bewertungskriterien wurden nicht formuliert. In drei Fällen gaben die befragten Personen jedoch an, sich selbst anhand einer Checkliste Notizen über den Gesprächsverlauf zu machen. Dies galt jedoch nicht für die anderen Interviewer, die an dem Gespräch teilnahmen.

Eine Arbeits- und Anforderungsanalyse wurde nur in zwei BGen durchgeführt. Die Anforderungen der Stelle und die darauf bezogene Beurteilung der Eignung eines Bewerbers wurde deswegen in der Regel nicht aus einer planmäßigen Analyse der zu verrichtenden Tätigkeiten beurteilt, sondern aus der persönlichen Erfahrung der Interviewer in der Verwaltung abgeleitet (vgl. Tabelle 5).

Tabelle 5: Antworten auf die Frage "Wurde eine Arbeits- und Anforderungsanalyse durchgeführt ?"

	absolut	% bezogen auf 23 befragte BG'n
Ja	2	8.7%
Nein	21	91.3%

Auf die Frage nach Merkmalen oder Verhaltensweisen, die besondere Beachtung finden, gaben 38% der Befragten an, auf nichts besonders zu achten. Die äußere Erscheinung wurde von 42% als besonders beachtenswert genannt (vgl. Tabelle 6).

Tabelle 6: Antworten auf die Frage "Worauf achten Sie besonders (Kleidung o.ä.) ?"

Äußere Erscheinung	11	42,31
Händedruck	1	3,85
Gestik	1	3,85
Benehmen/soziales Verhalten/Begrüßung	3	11,54
Nichts	10	38,46

Hinsichtlich der Auswertung des Gesprächs gab es ganz unterschiedliche Praktiken, die teilweise auch innerhalb der BG unterschiedlich gehandhabt wurden. Immer jedoch fand ein freies Gespräch über die individuellen Eindrücke statt. Eine systematisierte Beurteilung anhand einer Eigenschaftsliste oder eine Trennung von persönlichem Eindruck und beobachtetem Verhalten fand nicht statt. In vier Fällen wurde eine Beurteilung erst am Ende des Tages vorgenommen, nachdem alle Bewerber interviewt waren (vgl. Tabelle 7).

Tabelle 7: Antworten auf die Frage "Wie verläuft die Auswertung des Gesprächs ?"

	absolut	% bezogen auf Gesamtzahl der Nennungen
Freies Gespräch nach jedem Bewerber	9	40,91
Freies Gespräch nach jedem Bewerber und am Ende des Tages Gesamtschau	5	22,73
Unterschiedlich	1	4,55
Nach jedem Bewerber mit Beurteilungsschema*	3	13,64
Am Ende des Tages	4	18,18

* bezieht sich nur auf die telefonisch befragte Person, nicht auf die anderen Teilnehmer des Gesprächs

Die besonderen Vorteile des Einstellungsgesprächs wurden vor allem in der Möglichkeit des persönlichen Kontakts gesehen (vgl. Tabelle 8). Ausgesprochen

genaue Vorstellungen bestehen über das Ziel des Gespräch. In Tabelle 9 werden die genannten Bereiche, über die Informationen gewonnen werden sollen, benannt.

Tabelle 8: Antworten auf die Frage "Worin sehen Sie die besonderen Vorteile des Einstellungsgesprächs?"

	absolut	% bezogen auf Gesamtzahl der Nennungen
persönlicher Eindruck	13	54,17
Lenkbarkeit/Flexibilität/auf die Person eingehen	3	12,5
direkter als Schulnoten	1	4,17
erfassen der Ausdrucksfähigkeit	2	8,33
Informationen über soziales Umfeld	1	4,17
nicht die Tagesform maßgeblich wie bei Tests	1	4,17
geringer Aufwand/ohne Spezialisten durchführbar	2	8,33
Möglichkeit, Informationen zu geben	1	4,17

Tabelle 9: Antworten auf die Frage "Was genau wollen Sie mit dem Interview erfassen, was sind Ihre Ziele im Gespräch ?"

	absolut	% bezogen auf Gesamtzahl der Nennungen
fachliche Eignung	4	8
soziale Kompetenz/Teamfähigkeit	7	14
Informationen geben	5	10
persönlichen Eindruck vom Bewerber bekommen	13	26
geistige Wendigkeit	1	2
Interessen	2	4
verbale Fähigkeiten	4	8
Selbstbewußtsein	2	4
Reaktion auf Stress	1	2
Motivation	7	14
Führungsqualitäten	1	2
soziales Engagement	1	2
soziales Umfeld des Bewerbers kennenlernen	2	4

Auf die Frage "Sind die Fragestellungen standardisiert ?" wurden keine eindeutig zustimmenden Antworten gegeben. Allerdings wiesen die meisten Interviewten darauf hin, daß die Abläufe des Gesprächs sowie die angesprochenen Themenbereiche in der Regel einem Muster folgten, dies jedoch unbewußt und ohne Planung. Es handelte sich dabei um Routinen, wie sie sich beim Wiederholen von Tätigkeiten ergeben, nicht aber um eine geplante Standardisierung mit dem Ziel der systematischen Gewinnung vergleichbarer Daten.

5.4 Erfassung der Variablen Einstellung zur Arbeit, Leistungsdruck und Praxisorientierung

Der in Abschnitt 4.6.3 beschriebene Fragebogen wurde von 156 Personen ausgefüllt zurückgesandt, was bei etwa 500 verschickten Bögen einer Rücklaufquote von 31% entspricht.

Tabelle 10: Verhältnis von beschäftigten Dienstanwärtinnen in den Jahren 1994/95 zu den bis Erhebungsende eingegangenen Fragebögen (Rücklaufquote)

Berufsgenossenschaft	Nummer	n Anw.	n Frgbg.	Prozent
Bergbau	1	24	7	29.2
Steinbruchs	2	5	1	20.0
Keramische- und Glas	3	13	0	0.0
Gas- Und Wasserwerke	4	5	3	60.0
Maschinenbau BG	6	56	10	17.9
Norddeutsche Metall	7	22	5	22.7
Süddeutsche Metall	8	42	5	11.9
Edel- & Unedel	9	6	0	0.0
Feinmechanik & Elekt	10	39	15	38.5
Chemie	11	32	11	34.4
Holz	12	35	4	11.4
Zucker, Leder, Papiermacher	14,16,20	5	3	60.0
Druck und Papiervera	15	10	3	30.0
Leder	16	2	0	0.0
Textil & Bekleidung	17	2	0	0.0
Nahrungsmittel & Gas	18	43	12	27.9
Fleischerei	19	8	5	62.5
Bau Hamburg	21	8	1	12.5
Bau Hannover	22	32	17	53.1
Bau Wuppertal	23	16	4	25.0
Bau Frankfurt	24	13	0	0.0
Südwestliche Bau	25	7	0	0.0
Württembergische Bau	26	13	7	53.9
Bayrische Bau	27	16	5	31.3
Tiefbau	28	23	1	4.4
Großhandels- und Lag	29	52	22	42.3
Einzelhandel	30	31	0	0.0
Verwaltungs BG	31	91	1	1.1
Straßen-, U-Bahnen &	32	6	3	50.0
Fahrzeughaltung	33	54	0	0.0
Binnenschifffahrt	35	1	0	0.0
Gesundheitsdienst &	36	86	6	7.0
ohne BG-Angabe		-	5	-
Total		796	156	

n Anw. = Anzahl der Dienstanwärtinnen 1994/95
n Frgbg. = Anzahl der zurückgesandten Fragebögen
Prozent = Anzahl der erhaltenen Fragebögen bezogen auf die Anzahl der Dienstanwärtinnen 1994/95

Zur Erläuterung der Nummerierung: BG 5 hat 1994/95 keine Dienstanwärtinnen ausgebildet, BG 14,16,20 bilden eine Verwaltungsgemeinschaft, BG 13 existiert nicht, BG 34 (See-BG) wurde nicht in die Untersuchung aufgenommen.

Wie Tabelle 10 zu entnehmen ist, sind die BGen am Rücklauf der Fragebögen in sehr unterschiedlichem Ausmaß beteiligt. Zur Beurteilung der Frage, wie genau die Messung der Konstrukte Leistungsdruck, Praxisorientierung und Einstellung zur Arbeit für die untersuchte Organisation gelungen ist, muß sowohl die Absolutzahl der eingegebenen Bögen als auch die Prozentzahl bezogen auf die in den Jahren 1994/95 geprüften Anwärter der BG beachtet werden.

5.5 Ergebnisse zur Testkonstruktion

5.5.1 Stichprobe

Die Ausbildungsleistung der BGen wurde durch eine Totalerhebung der Prüfungsjahrgänge 1994/95 bestimmt (Auswertung des Notenarchivs der BGA), Leistungsdruck, Einstellung zur Arbeit und Praxisorientierung dagegen nur aufgrund der schriftlichen Befragung einer Stichprobe ($n = 156$) aus diesen Jahrgängen. Daher muß anhand einer detaillierten deskriptivstatistischen Beschreibung überprüft werden, ob und ggf. inwieweit sich die Stichprobe von der Grundgesamtheit der Prüfungsjahrgänge unterscheidet. Eine inferenzstatistische Überprüfung der gefundenen Unterschiede zwischen Stichprobe und Grundgesamtheit kann unterbleiben, da die Werte der Grundgesamtheit bekannt sind und es sich daher um den Vergleich zweier bekannter Größen handelt.

Das Alter der befragten Stichprobe entspricht sehr genau dem in der Grundgesamtheit festgestellten Wert. Bei der Geschlechtsverteilung finden sich leichte Unterschiede. Auch in der durchschnittlichen Note, die hier übrigens die Gesamtnote ist, also ein gewichtetes Mittel aus mündlicher (2/5 Gewicht) und schriftlicher Note (3/5 Gewicht), findet sich in der Stichprobe ein um (2.15 Punkte) besseres Ergebnis. Hinsichtlich der Verteilung der Personen auf die Auswahlmethode Interview oder Test weicht die Stichprobe um 1.3% von der Grundgesamtheit ab.

Tabelle 11: Vergleich der Stichprobe mit der bekannten Grundgesamtheit (Prüflinge der Jahre 1994/95)

		Grundgesamtheit	Stichprobe
Alter zum Prüfungszeitpunkt	<i>Mittelwert</i>	25.23	25.08
	<i>Standardabweichung</i>	3.29	3.54
	<i>Umfang</i>	N=796	n=144
	<i>missing</i>	0	12
Geschlecht	<i>weiblich (%)</i>	47.1%	44.9%
	<i>männlich (%)</i>	52.9%	53.2%
	<i>Umfang</i>	N=796	n=153
	<i>missing</i>	0	3
Gesamtnote*	<i>Mittelwert</i>	64.29	66.44
	<i>Standardabweichung</i>	6,54	6.99
	<i>Umfang</i>	N=709	n=127
	<i>missing</i>	87	29
Selektion	<i>Interview (%)</i>	72.1%	70.8%
	<i>Test (%)</i>	27.9%	29.2%
	<i>Umfang</i>	N=796	n=154
	<i>missing</i>	0	2

* Im Fragebogen wurde nicht nach der schriftlichen, sondern der erreichten Gesamtnote gefragt, da diese als eigentliches Ergebnis der Abschlussprüfung besser erinnerlich ist.

Soweit in den Fragebögen Daten fehlten, blieben diese auch in der Analyse unberücksichtigt. Eine Ausnahme bildeten hier drei Fälle, in denen ein Item der Skala Leistungsdruck oder Einstellung zur Arbeit fehlte. Hier wurde, um die Rücklaufquote nicht weiter zu mindern, die fehlende Antwort durch den Mittelwert der für die Skala von der Person gegebenen Antworten ersetzt.

5.5.2 Skala zur Messung der Variablen Leistungsdruck

Die Kennwerte aller konstruierten Skalen wurden nach der klassischen Testtheorie berechnet. Die Leistungsdruck-Skala erreichte mit ihren sieben Items eine Reliabilität (Cronbachs α) von 0.67. Die Schwierigkeit der Items rangierte in einem Bereich von 0.28 bis 0.61, die Trennschärfe von 0.24 bis 0.54.

Tabelle 12: Kennwerte der Skala Leistungsdruck

Item	Trennschärfe	Varianz	Schwierigkeit
L1	0.24	2.06	0.5
L2	0.47	2.64	0.28
L3	0.54	1.67	0.61
L4	0.35	1.7	0.49
L5	0.32	1.74	0.58
L6	0.44	1.81	0.38
L7	0.33	3.48	0.32

Reliability Coefficients 7 items

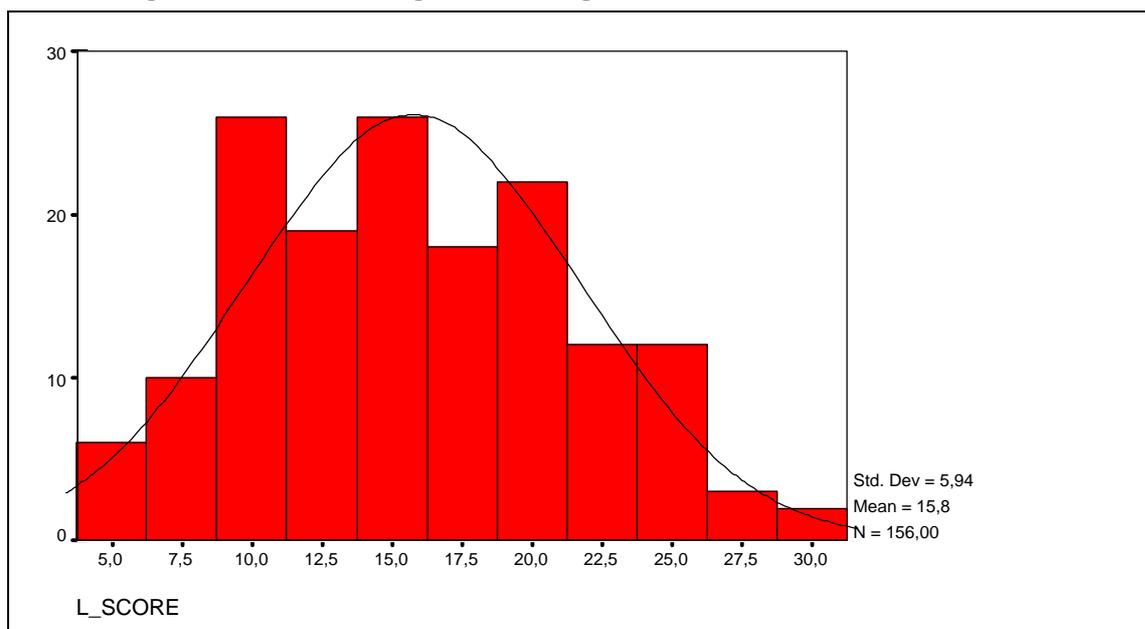
Alpha = .6676

Die Antwortverteilung entspricht nicht der Normalverteilungsannahme (vgl. Tabelle 13, Abbildung 5). Im Anhang C beigefügt sind die vollständigen Skalenkennwerte und Statistiken für die einzelnen Items, berechnet nach der klassischen Testtheorie.

Tabelle 13: Test auf Normalverteilung für die Antwortverteilung der Leistungsdruck-Skala

	Statistic	df	Significance
K-S (Lilliefors)	.0854	156	.0073

Abbildung 5: Antwortverteilung der Leistungsdruck-Skala in Relation zur Normalverteilung



5.5.3 Skala zu Messung der Variablen Einstellung zur Arbeit

Die Skala Einstellung zur Arbeit erreichte mit ihren 5 Items für Cronbachs α einen Wert von 0.65. Die Schwierigkeit der Items lag in einem Bereich von 0.23 bis 0.44. Die Trennschärfe variierte von 0.32 bis 0.50. Die Antworten sind unimodal, linksschief verteilt und erfüllen nicht die Normalverteilungsannahme (vgl. Tabelle 15, Abbildung 6).

Tabelle 14: Kennwerte der Skala Einstellung zur Arbeit

Item	Trennschärfe	Varianz	Schwierigkeit
M1	0.32	0.48	0.23
M2	0.33	0.71	0.25
M3	0.5	1.59	0.4
M4	0.4	1.29	0.26
M5	0.47	1.85	0.44

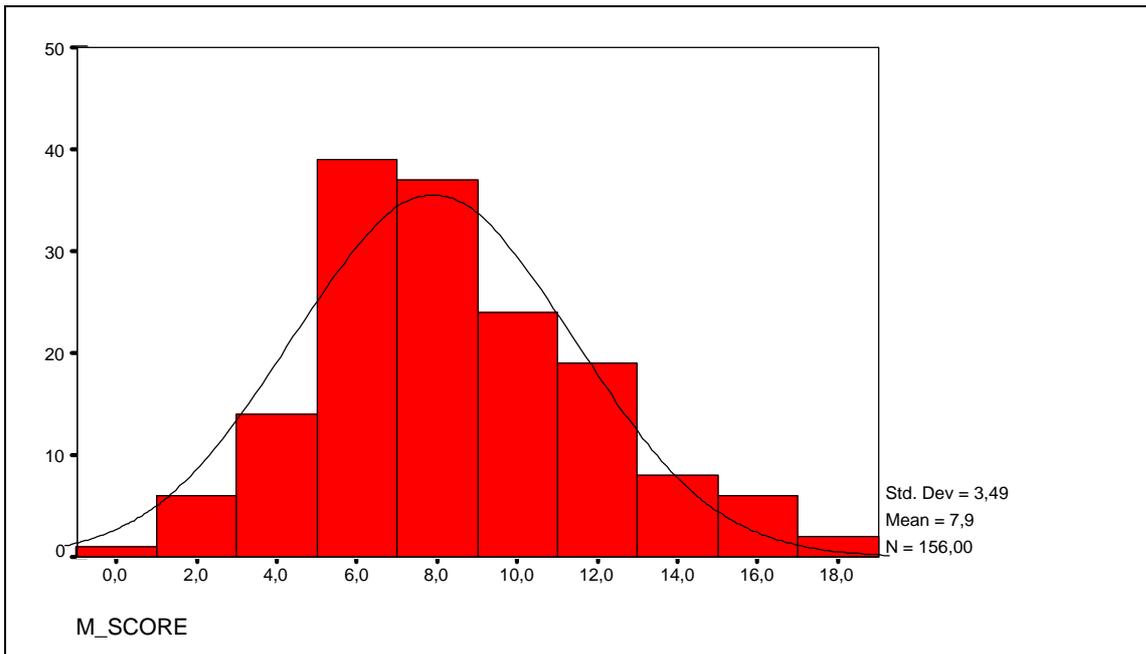
Reliability Coefficients 5 items
Alpha = .6433

Im Anhang C beigefügt sind die vollständigen Skalenkennwerte und Statistiken für die einzelnen Items, berechnet nach der klassischen Testtheorie.

Tabelle 15: Test auf Normalverteilung für die Antwortverteilung der Skala Einstellung zur Arbeit

	Statistic	df	Significance
K-S (Lilliefors)	.1101	156	.0001

Abbildung 6: Antwortverteilung der Skala Einstellung zur Arbeit in Relation zur Normalverteilung



5.5.4 Messung der Praxisorientierung

Die Antworten auf die Frage nach der Praxisorientierung verteilen sich rechtsschief und bimodal. Sie entsprechen nicht der Normalverteilungsannahme (vgl. Tabelle 16, Abbildung 7).

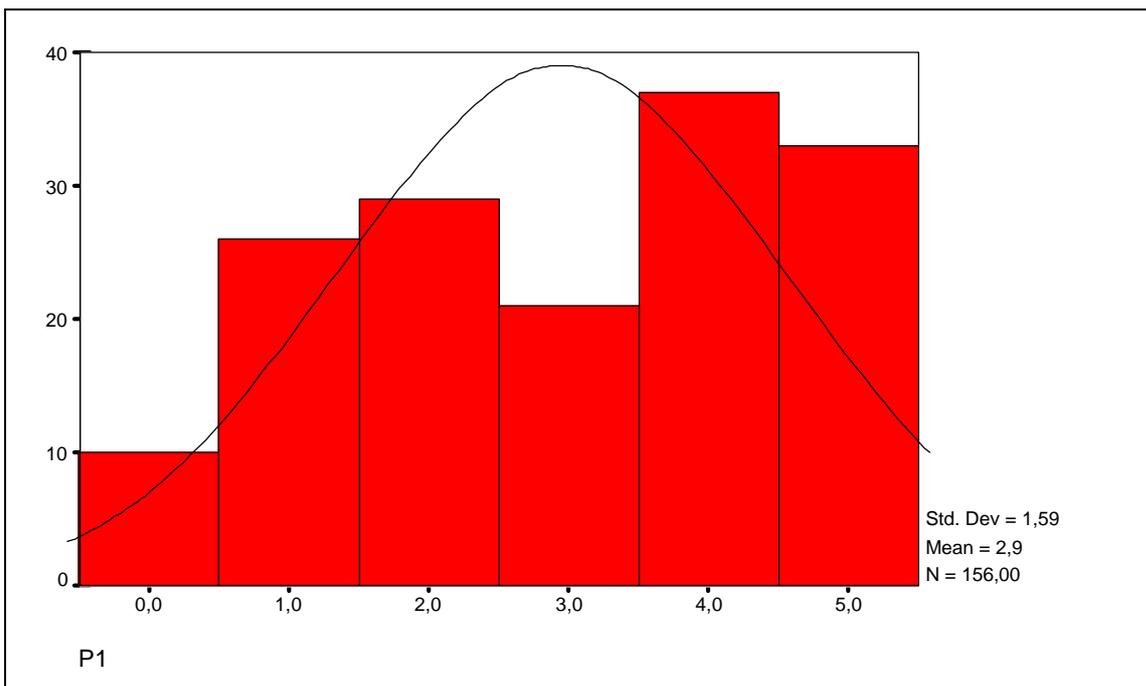


Abbildung 7: Antwortverteilung der Frage nach der Praxisorientierung

Tabelle 16: Antwortverteilung der Frage nach der Praxisorientierung der Ausbildung

Value Label		Value	Frequency	Percent	Valid Percent	Cum Percent	
		.0	10	6.4	6.4	6.4	
		1.0	26	16.7	16.7	23.1	
		2.0	29	18.6	18.6	41.7	
		3.0	21	13.5	13.5	55.1	
		4.0	37	23.7	23.7	78.8	
		5.0	33	21.2	21.2	100.0	
		-----		-----		-----	
Total			156	100.0	100.0		
Mean	2.949	Std err	.127	Median	3.000		
Mode	4.000	Std dev	1.589	Variance	2.526		
Kurtosis	-1.185	S E Kurt	.386	Skewness	-.237		
S E Skew	.194	Range	5.000	Minimum	.000		
Maximum	5.000						
Valid cases	156	Missing cases	0				

In Abbildung 7 sind klar zwei lokale Maxima bei 2 und 4 Punkten zu erkennen. Angesichts dieser bimodalen Antwortverteilung kann keine Normalverteilung vorliegen.

5.5.5 Ergebnisse der explorativen Fragen

Auf die Ergänzungsfragen "Am besten fand ich in meiner Ausbildung, daß ..." und "Am meisten hat mich in meiner Ausbildung gestört, daß ..." gaben 33 befragte Personen keine Antwort. Die Fähigkeit der Dozenten wurde in 34 Fällen bemängelt. Sonst wurde vor allem beklagt, daß die Ausbildung nicht zum Abschluß eines Diplom-Verwaltungswirts führt, wie dies z.B. bei den Krankenkassen der Fall ist. Als angenehm wurde der Wechsel zwischen Praxisphasen und den Lehrgängen an der BGA empfunden, obwohl teilweise auch die langen Anfahrtswege als belastend angesehen wurden. Der Umstand, daß die in den drei Jahren der Ausbildung gezeigten Leistungen nicht in die Abschlußnote einfließen, wurde ebenso wie die gute Bezahlung thematisiert.

5.6 Berechnung der BG-Kennwerte

Die Analyse der Daten erfolgt auf Organisationsebene und nicht auf der Ebene von Individuen, da in dieser Untersuchung Konzepte zur Auswahl und Ausbildung evaluiert werden sollen, die als Eigenschaften von Organisationen aufzufassen sind. Zunächst werden daher aus den Daten Kennwerte für die Organisation berechnet. Die Variation des Maßes für den Ausbildungserfolg, operationalisiert durch den arithmetischen Mittelwert der Noten aller Anwärter einer BG innerhalb des Beobachtungszeitraums, soll dabei auf den Faktor Selektion auf den beiden Stufen Interview und Test zurückgeführt werden. Vermutete Einflüsse relevanter anderer Größen, wie Praxisorientierung der Ausbildung, subjektiv empfundener Leistungsdruck und Einstellung zur Arbeit, werden durch die Einfügung dieser Variablen als Kovariaten in das varianzanalytische Modell herauspartialisiert. Nach Bortz (1993, S. 341) kann so die Fehlervarianz vermindert werden.

Als Kriterium für die Bedeutsamkeit der gefundenen Effekte kommt in dieser Untersuchung die statistische Signifikanz nicht in Betracht. Dies würde in die Irre führen, da bei der Signifikanztestung implizit von einer unbekanntem, im Prinzip unendlich großen Grundgesamtheit ausgegangen wird, aus der eine Zufallsstichprobe des Umfangs n gezogen wurde. Da diese Zufallsstichprobe in Abhängigkeit von ihrer Größe die Verhältnisse in der Grundgesamtheit nur annähernd darstellt, wird normalerweise eine Signifikanztestung der gefundenen Unterschiede vorgenommen.

Aus zwei Gründen werden Signifikanzen hier nicht als sinnvoll interpretierbare Größen angesehen:

Erstens ist in dieser empirischen Untersuchung die Grundgesamtheit durch die Anzahl der gewerblichen BGen vorgegeben. Von den 36 in § 646 Abs. I RVO durchnummerierten verbleiben aus bereits dargelegten Gründen 31 Untersuchungsobjekte als Grundgesamtheit. Nach Aufspaltung einer der beiden BGen, die wegen der abweichenden Auswahlpraxis in einzelnen Bezirksverwaltungen hinsichtlich des Faktors Selektion nicht eindeutig zuzuordnen waren, erhöht sich die Zahl der Untersuchungsobjekte auf 33. Von dieser sehr

begrenzten Grundgesamtheit konnten 24 Untersuchungsobjekte, von denen alle relevanten Variablen (Auswahlverfahren, Praxisorientierung, Leistungsdruck, Note und Einstellung zur Arbeit) bekannt waren, in die Analyse mit aufgenommen werden, was einer Erfassung von 73% der Grundgesamtheit entspricht. Da also die Stichprobe 3/4 der Grundgesamtheit umfaßt, führte eine Signifikanztestung zur Unterschätzung der gefundenen Effekte. Dies kann als Vorteil dieser Untersuchung betrachtet werden, da somit nicht die Gefahr besteht, Überlegungen zur Relevanz der Befunde durch formale Betrachtungen über die statistische Signifikanz zu ersetzen.

Zweitens ist Ziel dieser Untersuchung eine Aufklärung der Wirkung verschiedener Auswahlmethoden und Ausbildungskonzepte. Die Wirkung dieser Variablen auf die Ausbildungsleistung wird anhand der Prüfungsjahrgänge 1994/95 evaluiert. Eine Verallgemeinerung der Befunde über den Beobachtungszeitraum hinaus und auf BGen, die an dieser Untersuchung nicht teilnahmen, wird nicht angestrebt.

5.6.1 Selektion

Aus der ersten schriftlichen Befragung und der fernmündlichen Nachfrage ist für jede BG ermittelt worden, ob sie im Beobachtungszeitraum Bewerber für die gehobene Verwaltungslaufbahn durch einen Einstellungstest (in Kombination mit einem Interview) oder durch ein Interview auswählte. Die Ergebnisse sind in Tabelle 17 dargestellt. Problematisch war die Einordnung von BG 1 und BG 33, die unterschiedliche Auswahlmethoden praktizierten. Man hätte prüfen können, auf welchem Wege die Mehrheit der Anwärter der BG ausgewählt werden, um die gesamte Organisation entsprechend einzuordnen. Da jedoch die Note als Kennwert der Organisation auf Basis aller Anwärter, ob getestet oder interviewt, berechnet wurde, würde eine eventuelle Abhängigkeit zwischen Selektionsmethode und Noten so verwischt. Es käme zu einer Konfundierung der beiden Stufen des Faktors Selektion. Das wurde durch die Auftrennung der BG zur Analyse auf Organisationsebene verhindert. Dabei wurde die Bezirksverwaltung, die von der Selektionspraxis des Rests der BG abwich, als eigener Datensatz behandelt.

BG-Kennwerte und auch die durchschnittliche schriftliche Note als Operationalisierung des Ausbildungserfolgs wurden für die Bezirksverwaltung, die von der sonst praktizierten Auswahlmethode abwich, getrennt berechnet. Für BG 33 war eine Auftrennung nicht notwendig, da von dieser BG kein ausgefüllter Fragebogen vorlag.

Tabelle 17: Verteilung der Kennwerte für die Variable Selektion auf die verschiedenen BGen

Variable	Label	Selekt
BG	01 Bergbau	0 (1 in BV 2)
BG	02 Steinbruch	1
BG	03 Keramische- und Glas	1
BG	04 Gas- und Wasserwerke	0
BG	06 Maschinenbau	0
BG	07 Norddeutsche Metall	1
BG	08 Süddeutsche Metall	0
BG	09 Edel- und Unedel	0
BG	10 Feinmechanik und Ele	0
BG	11 Chemie	0
BG	12 Holz	0
BG	14,16,20 Zucker,Papiermacher,Leder	0
BG	15 Druck- und Papierver	0
BG	17 Textil und Bekleidun	1
BG	18 Nahrungsmittel und G	0
BG	19 Fleischerei	0
BG	21 Bau Hamburg	0
BG	22 Bau Hannover	1
BG	23 Bau Wuppertal	0
BG	24 Bau Frankfurt	1
BG	25 Südwestliche Bau	0
BG	26 Württembergische Bau	1
BG	27 Bayrische Bau	1
BG	28 Tiefbau	0
BG	29 Großhandels- und Lag	0
BG	30 Einzelhandel	0
BG	31 VBG	1
BG	32 Straßen-, U-Bahnen u	0
BG	33 Fahrzeughaltung	0 (1 in BV 5)
BG	35 Binnenschifffahrt	1
BG	36 BGW	0

Total Cases = 31
0 = Interview (n=21)
1 = Test (n=10)

Zur Erläuterung der Nummerierung: BG 5 hat 1994/95 keine Dienstanwärter ausgebildet, BG 14,16,20 bilden eine Verwaltungsgemeinschaft, BG 13 existiert nicht, BG 34 (See-BG) wurde nicht in die Untersuchung aufgenommen.

5.6.2 Schriftliche Durchschnittsnote

Aus dem Notenarchiv der BGA wurde für die Jahre 1994/95 für jede BG die durchschnittliche schriftliche Note der Anwärter in der Abschlußprüfung berechnet. Dabei ergaben sich folgende Werte:

Tabelle 18: Durchschnittliche schriftliche Noten der Anwärter verschiedener BGen im Beobachtungszeitraum

Variable	Value	Label	Mean	Std Dev	Cases
Summaries of SCHRIFTL Note schriftlich					
By levels of BG					
For Entire Population			61.1425	6.9197	709
BG	1	Bergbau (Interview)	61.8182	7.1947	11
BG	1	Bergbau (Test)	61.2500	8.0014	12
BG	2	Steinbruchs	60.6000	8.4143	5
BG	3	Keramische- und Glas	58.8182	8.9310	11
BG	4	Gas- Und Wasserwerke	61.2500	5.3774	4
BG	6	Maschinenbau	62.3404	6.3768	47
BG	7	Norddeutsche Metall	61.5263	6.1497	19
BG	8	Süddeutsche Metall	60.6486	6.5371	37
BG	9	Edel- & Unedel	62.3333	9.8116	6
BG	10	Feinmechanik & Elekt	60.2895	6.4887	38
BG	11	Chemie	60.3793	7.6036	29
BG	12	Holz	61.5000	5.6682	32
BG	14,16,20	Zucker,Papiermacher,Leder	54.0000	4.0000	3
BG	15	Druck und Papiervera	58.6000	4.2999	10
BG	16	Leder	57.5000	3.5355	2
BG	17	Textil & Bekleidung	66.0000	9.8995	2
BG	18	Nahrungsmittel & Gas	60.4286	7.9605	35
BG	19	Fleischerei	66.4286	7.1147	7
BG	21	Bau Hamburg	60.2857	5.4989	7
BG	22	Bau Hannover	62.4333	5.7216	30
BG	23	Bau Wuppertal	61.1333	7.0596	15
BG	24	Bau Frankfurt	58.4000	3.8644	10
BG	25	Südwestliche Bau	60.1429	7.6033	7
BG	26	Württembergische Bau	63.3333	4.0751	12
BG	27	Bayrische Bau	61.4667	6.8647	15
BG	28	Tiefbau	60.4286	7.7044	21
BG	29	Großhandels- und Lag	64.3125	7.2528	48
BG	30	Einzelhandel	62.3571	5.8577	28
BG	31	Verwaltungs BG	60.0988	7.4357	81
BG	32	Straßen-, U-Bahnen &	57.8333	7.3052	6
BG	33	Fahrzeughaltung (Interv.)	60.5000	7.5515	40
BG	33	Fahrzeughaltung (Test)	55.4000	1.8166	5
BG	35	Binnenschifffahrt	49.0000	.	1
BG	36	Gesundheitsdienst &	61.3151	6.8837	73
Total Cases = 796					
Missing Cases = 87 or 10.9 Pct (durchgefallene Anwärter)					

Die durchschnittlich von den BGen im Beobachtungszeitraum erzielten schriftlichen Noten haben einen Range von 66,42 bis 49,00 Punkten. Dabei beruhen die Durchschnittswerte auf sehr unterschiedlich großen Populationen. BG 35 bildete im

Beobachtungszeitraum nur einen Anwärter aus, während BG 31 allein 81 Prüflinge stellte. Die Mittelwerte resultieren mithin auf sehr unterschiedlich großen Gruppen.

5.6.3 Leistungsdruck

Tabelle 19: Verteilung der Kennwerte für die Variable Leistungsdruck auf die verschiedenen BGen

Summaries of		L_SCORE (Leistungsdruck)			
By levels of		BG			
Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			15.7086	5.8658	151
BG	1	(Interview)	19.6667	3.5119	3
BG	1	(Test)	15.0000	4.2426	4
BG	2		10.0000	.	1
BG	3		.	.	0
BG	4		12.3333	4.1633	3
BG	6		25.6000	3.2042	10
BG	7		18.2000	5.4498	5
BG	8		16.0000	2.8284	5
BG	9		.	.	0
BG	10		17.6667	5.5635	15
BG	11		17.0000	6.8411	11
BG	12		21.2500	3.8622	4
BG	14,16,20		11.0000	6.9282	3
BG	15		13.0000	4.5826	3
BG	17		.	.	0
BG	18		13.5833	5.3336	12
BG	19		14.0000	2.5495	5
BG	21		21.0000	.	1
BG	22		16.7059	4.4828	17
BG	23		16.0000	3.4641	4
BG	24		.	.	0
BG	25		.	.	0
BG	26		15.0000	3.8730	7
BG	27		9.4000	3.1305	5
BG	28		14.0000	.	1
BG	29		11.6364	4.8260	22
BG	30		.	.	0
BG	31		24.0000	.	1
BG	32		13.6667	5.5076	3
BG	33	(Interview)	.	.	0
BG	33	(Test)	.	.	0
BG	34		.	.	0
BG	35		.	.	0
BG	36		13.5000	5.2440	6

Missing Cases = 5 or 3.2 Pct

Zur Erläuterung der Nummerierung: BG 5 hat 1994/95 keine Dienstanwärter ausgebildet, BG 14,16,20 bilden eine Verwaltungsgemeinschaft, BG 13 existiert nicht, BG 34 (See-BG) wurde nicht in die Untersuchung aufgenommen.

Tabelle 19 zeigt, wie sich die Skalenwerte für den Leistungsdruck auf die verschiedenen BGen verteilen.

Die fünf missing cases in Tabelle 19 sind Fälle, in denen die Frage nach der ausbildenden BG nicht beantwortet wurde. Diese Fälle konnten auch hinsichtlich der anderen durch den Fragebogen erhobenen Variablen nicht zugeordnet werden. Die durchschnittlich von den BGen erreichten Skalenwerte variieren um einen Mittelwert von 15.71 mit einem Minimum von 9.4 und einem Maximum von 25.6.

5.6.4 Praxisorientierung

Aus Tabelle 20 ist die Verteilung der Skalenwerte für den Leistungsdruck auf die verschiedenen BGen zu sehen.

Tabelle 20: Verteilung der Kennwerte für die Variable Praxisorientierung auf die verschiedenen BGen

Summaries of		P1	(Praxisorientierung)		
By levels of		BG			
Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			2.9272	1.6004	151
BG	1	(Interview)	4.3333	1.1547	3
BG	1	(Test)	1.7500	1.5000	4
BG	2		.0000	.	1
BG	3		.	.	0
BG	4		2.6667	1.1547	3
BG	6		3.4000	1.5776	10
BG	7		3.2000	1.3038	5
BG	8		2.2000	1.3038	5
BG	9		.	.	0
BG	10		3.6000	1.6818	15
BG	11		3.0909	1.5783	11
BG	12		3.5000	.5774	4
BG	14,16,20		4.6667	0.5774	3
BG	15		3.0000	1.0000	3
BG	17		.	.	0
BG	18		2.0000	1.1282	12
BG	19		3.0000	1.5811	5
BG	21		4.0000	.	1
BG	22		1.2941	1.3117	17
BG	23		2.7500	2.3094	4
BG	24		.	.	0
BG	25		.	.	0
BG	26		4.1429	1.0690	7
BG	27		3.8000	.4472	5
BG	28		5.0000	.	1
BG	29		3.1818	1.7630	22
BG	30		.	.	0
BG	31		5.0000	.	1
BG	32		2.3333	0.5774	3
BG	33	(Interview)	.	.	0
BG	33	(Test)	.	.	0
BG	34		.	.	0
BG	36		2.8333	.9832	6
Missing Cases = 5 or 3.2 Pct					
Zur Erläuterung der Nummerierung: BG 5 hat 1994/95 keine					

5.6.5 Einstellung zur Arbeit

Die Werte für die Variable Einstellung zur Arbeit verteilen sich wie in Tabelle 21 dargestellt.

Tabelle 21: Verteilung der Kennwerte für die Einstellung zur Arbeit auf die verschiedenen BGen

Summaries of		M_SCORE (Einstellung zur Arbeit)			
By levels of		BG			
Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			7.7947	3.4415	151
BG	1	(Interview)	9.6667	6.4291	3
BG	1	(Test)	10.0000	5.7735	4
BG	2		7.0000	.	1
BG	3		.	.	0
BG	4		10.0000	1.7321	3
BG	6		5.8000	2.2509	10
BG	7		8.2000	1.9235	5
BG	8		8.6000	4.5607	5
BG	9		.	.	0
BG	10		6.8667	3.6029	15
BG	11		7.2727	4.1735	11
BG	12		9.5000	2.6458	4
BG	14,16,20		10.0000	3.4621	3
BG	15		11.6667	3.7859	3
BG	17		.	.	0
BG	18		8.4167	3.9648	12
BG	19		5.4000	3.2094	5
BG	21		3.0000	.	1
BG	22		7.7647	2.5867	17
BG	23		7.0000	3.1623	4
BG	24		.	.	0
BG	25		.	.	0
BG	26		6.4286	1.7182	7
BG	27		8.2000	.4472	5
BG	28		7.0000	.	1
BG	29		8.2727	3.4666	22
BG	30		.	.	0
BG	31		5.0000	.	1
BG	32		9.3333	3.0551	3
BG	33	(Interview)	.	.	0
BG	33	(Test)	.	.	0
BG	35		.	.	0
BG	36		7.3333	4.5461	6

Missing Cases = 5 or 3.2 Pct

Die Werte bewegen sich zwischen einem Minimum von 3,0 und einem Maximum von 15 Punkten.

5.7 Prüfung der Hypothesen

5.7.1 Überprüfung der H1

Die Überprüfung der Hypothese 1 erfolgt durch drei Operationalisierungen:

1. Vergleich der durchschnittlich von den Anwärtern der BGen erzielten schriftlichen Noten auf beiden Stufen des Faktors Selektion zur Prüfung der **H1.1**:

BGen, die mit Testverfahren die Eignung von Bewerbern um die gehobene Verwaltungslaufbahn feststellen, erzielen durchschnittlich bessere schriftliche Abschlußnoten als BGen, die durch Interviews auswählen.

Tabelle 22: Durchschnittliche schriftliche Note auf beiden Stufen des Faktors Selektion

Variable	Number of Cases	Mean	SD	SE of Mean

SCHRIFTL				
SELEKT 0	21	60.8195	2.224	.485
SELEKT 1.	12	59.7500	4.292	1.239

Mean Difference = 1.0695				
Levene's Test for Equality of Variances: F= 3.023 P= .092				
t-test for Equality of Means				
Variances	t-value	df	2-Tail Sig	95% CI for Diff
Equal	.95	31	.351	(-1.233; 3.372)
Unequal	.80	14.45	.435	(-1.785; 3.924)

Die durch Interview auswählenden BGen erzielen im Mittel eine schriftliche Note von 60.82 Punkten gegenüber einem Mittel von 59,75 Punkten für die BGen die durch Testverfahren auswählten. Die Ausbildungsleistung für die BGen unter der Interviewbedingung liegt damit um 1.07 Punkte (~1/3 der Standardabweichung) höher als die der durch Test ausgewählten. Auffallend ist die für die BGen unter der Testbedingung deutlich höhere Standardabweichung von 4.29 zu 2.23.

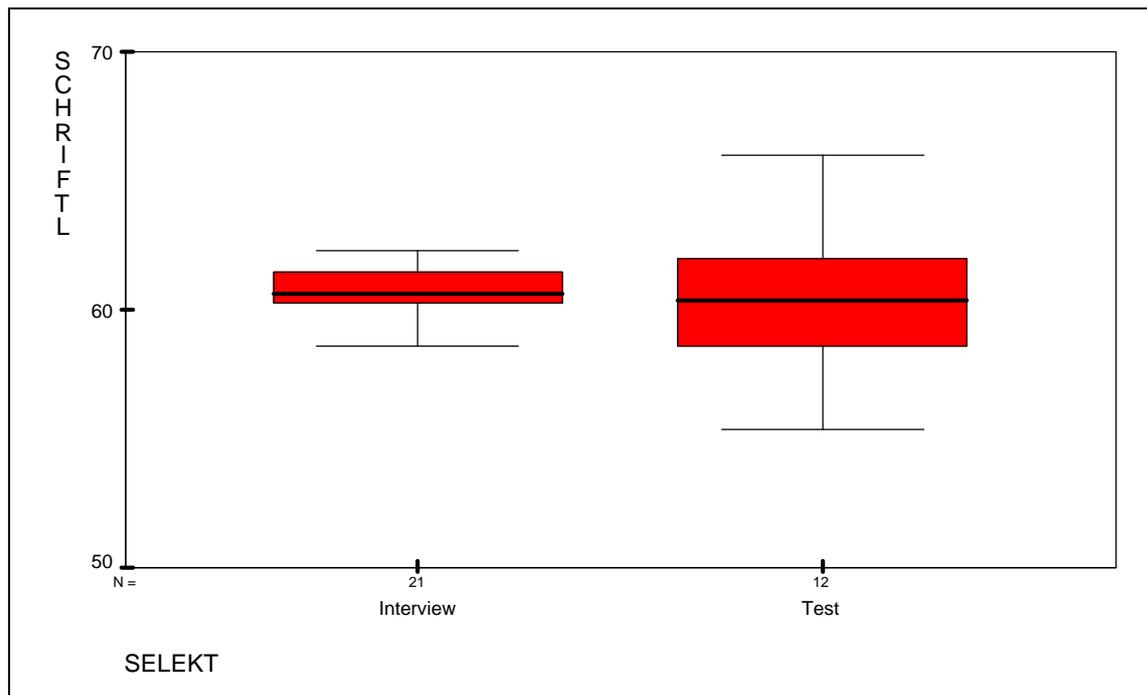


Abbildung 8: Box-Plot der schriftlichen Note für die BGen auf beiden Stufen des Faktors Selektion

2. Vergleich der Abbrecherquote der BGen auf beiden Stufen des Faktors Selektion zur Prüfung der **H1.2**:

BGen, die mit Testverfahren die Eignung von Bewerbern um die gehobene Verwaltungslaufbahn feststellen, haben eine geringere Abbrecherrate als BGen, die durch Interviews auswählen.

Tabelle 23: Durchschnittliche Abbrecherquoten auf den beiden Stufen des Faktors Selektion

Variable	Number of Cases*	Mean	SD	SE of Mean

ABBR_RAT	Abbrecherrate			
SELEKT 0	19	.0598	.060	.014
SELEKT 1.	10	.1253	.175	.055

Mean Difference = -.0655				
Levene's Test for Equality of Variances: F= 3.051 P= .092				
t-test for Equality of Means				95%
Variances	t-value	df	2-Tail Sig	SE of Diff

Equal	-1.49	27	.147	.044
Unequal	-1.15	10.11	.278	.057

* Bei den beiden Berufsgenossenschaften (BG 1 & BG 33), die teilweise Tests durchführten und teilweise Interviews, konnte die Anzahl der Abbrecher nicht, wie dies bei den Noten möglich war, eindeutig den Stufen des Faktors Selektion zugeordnet werden. Sie wurden daher in dieser Analyse nicht aufgenommen				

Die Wahrscheinlichkeit, die Ausbildung abzubrechen, liegt danach bei den BGen, die per Test auswählen ($p = 0.13$), etwa doppelt so hoch als bei den durch Interview auswählenden ($p = 0.06$). Auch hier ist die Standardabweichung der Testgruppe (0.18) deutlich höher als die der Interviewgruppe (0.6).

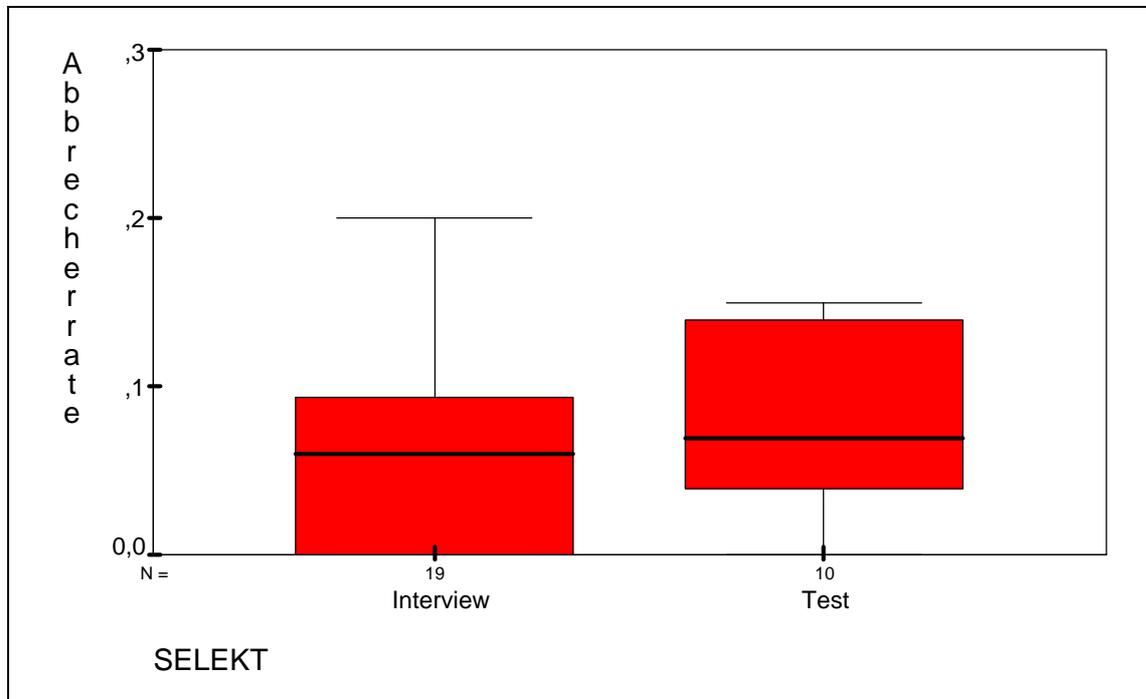


Abbildung 9: Box-Plot der Abbrecherquote der BGen auf beiden Stufen des Faktors Selektion

3. Vergleich der Anzahl der Anwärter, die den Prüfungsanforderungen nicht genügten auf den beiden Stufen des Faktors Selektion zur Prüfung der **H1.3**:
BGen, die mit Testverfahren die Eignung von Bewerbern um die gehobene Verwaltungslaufbahn feststellen, haben eine geringere Rate von Anwärtern, die den Prüfungsanforderungen nicht genügen.

Unter der Testbedingung ist die Rate der Anwärter, die den Prüfungsanforderungen nicht genügen, um 1.48% größer als unter der Interviewbedingung. Dieser Unterschied beträgt nur etwa 1/3 Standardabweichungen.

Tabelle 24: Quote der durchgefallenen Anwärter auf den beiden Stufen des Faktors Selektion

Variable	Number of Cases*	Mean	SD	SE of Mean

WDH_RATE	Wiederholerrate			
SELEKT 0	19	.0407	.045	.010
SELEKT 1.	10	.0495	.060	.019

Mean Difference = -.0089				
Levene's Test for Equality of Variances: F= 1.212 P= .281				
t-test for Equality of Means				
				95%
Variances	t-value	df	2-Tail Sig	SE of Diff

Equal	-.45	27	.654	.020
Unequal	-.41	14.50	.685	.021

95% CI for Diff				
(-.049; .031)				
(-.055; .037)				

* Bei den beiden Berufsgenossenschaften (BG 1 & BG 33), die teilweise durch Test, teilweise durch Interview ausgewählt, konnte die Anzahl der Abbrecher nicht, wie dies bei den Noten möglich war, eindeutig den Stufen des Faktors Selektion zugeordnet werden. Sie wurden daher in dieser Analyse nicht mit aufgenommen				

Bei keiner der Operationalisierungen des Ausbildungserfolgs zeigte sich ein hypothesenkonformes Ergebnis. Alle Unterschiede zwischen den beiden Stufen des Faktors Selektion

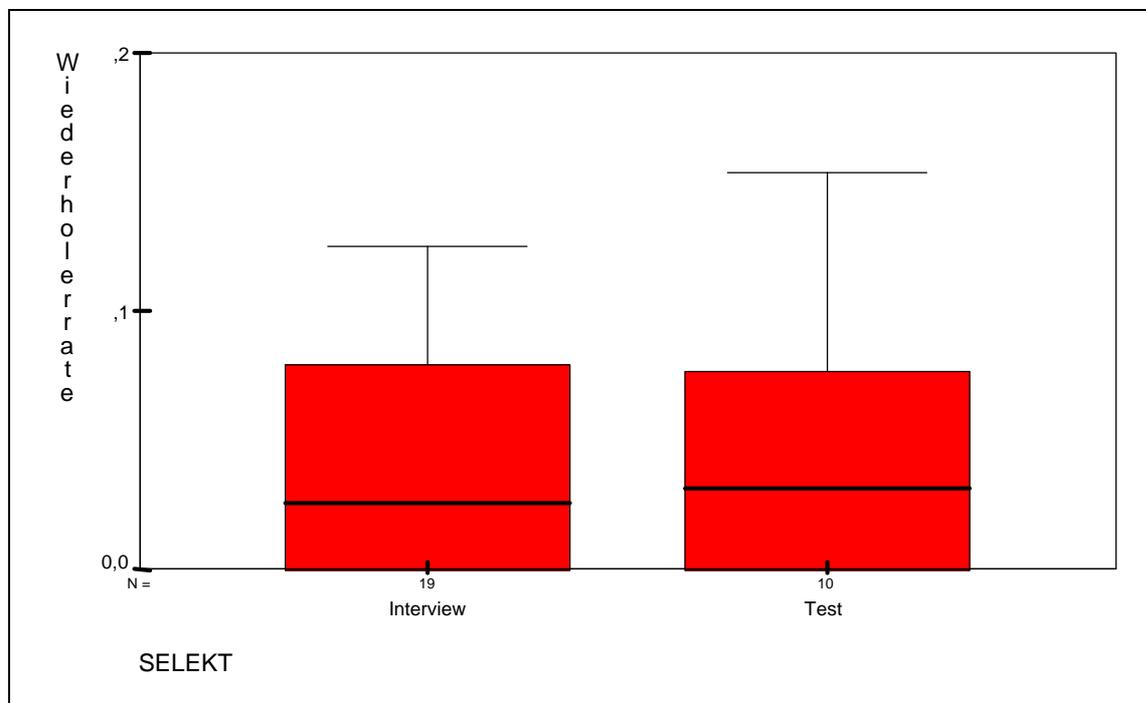


Abbildung 10: Box-Plot der Wiederholerrate für die BGen auf beiden Stufen des Faktors Selektion

5.7.2 Überprüfung der H2

Durchgeführt wurde ein t-Test für die schriftlichen Noten der beiden durch Medianhalbierung der Variable Praxisorientierung gewonnenen Gruppen von BGen zur Prüfung der **H2**:

Je praxisorientierter eine BG ausbildet, desto schlechter ist die durchschnittliche schriftliche Note ihrer Dienstanwärter in der Abschlußprüfung.

Für die am Median der Praxisorientierung getrennten Gruppen erreichen die BGen, die eher praxisorientiert ausbilden, eine um 0.38 Punkte schlechtere schriftliche Note als die wenig praxisorientiert ausbildenden BGen.

Tabelle 25: Notenunterschiede zwischen den durch Medianhalbierung anhand des P-Wertes gewonnenen beiden Gruppen von BGen

t-tests for independent samples of P_SCORE						
Variable		Number of Cases	Mean	SD	SE of Mean	

SCHRIFTL						
P_SCORE >= 3.135		12	60.6200	2.016	.608	
P_SCORE < 3.135		12	61.0045	2.203	.664	

Mean Difference = -.3845						
Levene's Test for Equality of Variances: F= .016 P= .901						
t-test for Equality of Means						95%
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff	

Equal	-.43	20	.674	.900	(-2.263; 1.494)	
Unequal	-.43	19.84	.674	.900	(-2.263; 1.494)	

Je praxisorientierter eine BG ausbildet, desto mehr Anwärter brechen die Ausbildung ab bzw. müssen die Abschlußprüfung an der BGA wiederholen. Die bei der Prüfung erzielte Note wird mit größerer Praxisorientierung der Ausbildung tendenziell schlechter. Die hypothetisch behauptete Beziehung zwischen Ausbildungserfolg und Praxisorientierung findet sich für alle Operationalisierungen des Ausbildungserfolgs.

Tabelle 26: Korrelation zwischen der Praxisorientierung der Ausbildung und den Operationalisierungen des Ausbildungserfolgs

	ABBR_RAT	SCHRIFTL	WDH_RATE
P_SCORE	,1078 (,24) P= ,616	-,1308 (,24) P= ,542	,2488 (,24) P= ,241
P_SCORE = Praxisorientierung			
ABBR_RATE = Abbrecherrate			
WDH_RATE = Wiederholerrate			
SCHRIFTL = schriftliche Note			

5.7.3 Überprüfung der H3

Prüfung der **H3**:

Der Zusammenhang zwischen der Auswahlmethode und der Ausbildungsleistung wird durch den Grad der Praxisorientierung der Ausbildung in der Weise moderiert, daß durch Tests ausgewählte Personen bei hoher Praxisorientierung keine besseren schriftlichen Noten erreichen als durch Interview ausgewählte, während bei geringer Praxisorientierung die durch Testverfahren ausgewählten Personen bessere schriftliche Noten erzielen.

Es wird die intervallskalierte Variable Praxisorientierung durch Medianhalbierung ($Md = 3.135$) dichotomisiert. Die sich ergebenden vier Kombinationen von Praxisorientierung und Auswahlmethode werden hinsichtlich der erreichten Mittelwerte der schriftliche Note verglichen (vgl. Tabelle 27).

Tabelle 27: Mittelwert der schriftlichen Note bei verschiedenen Kombinationen von Praxisorientierung und Selektion

	Interview	Test
praxisorientiert ($P \geq 3.135$)	60.44 (n = 8)	61.28 (n = 4)
wenig praxisorientiert ($P < 3.135$)	60.89 (n = 9)	61.43 (n = 3)

In Abbildung 11 zeigt sich eine tendenziell bessere Leistung für die Stufe "Test" des Faktors Selektion. Dies widerspricht den bereits dargelegten gegenteiligen Befunden nicht, da sich Abbildung 11 und Tabelle 27 nur auf die BGen beziehen, von denen

beantwortete Fragebögen vorlagen. Da die Praxisorientierung durch Fragebogen erhoben wurde, konnten in diese Analyse nur die Datensätze aufgenommen werden, für die Note, Auswahlmethode und Praxisorientierung bekannt waren. Dies war für 24 von insgesamt 33 Fällen gegeben.

Die in Tabelle 27 dargestellten Mittelwertunterschiede sind in Abbildung 11 graphisch veranschaulicht.

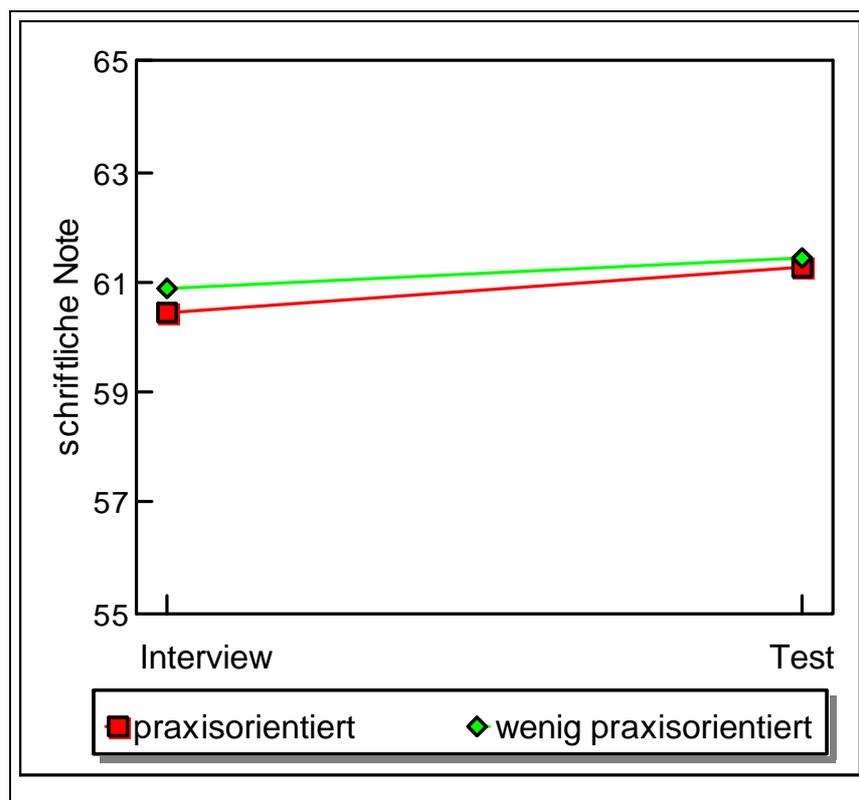


Abbildung 11: Zusammenhang zwischen der schriftlichen Note und dem Faktor Selektion auf den beiden durch Medianhalbierung gewonnenen Stufen der Variablen Praxisorientierung

Der in Hypothese 3 behauptete Moderatoreffekt der Praxisorientierung auf die Wirksamkeit der Stufen des Faktors Selektion, läßt sich in den Daten nicht finden. Es besteht eine schwache der Hypothese widersprechende Beziehung.

5.7.4 Überprüfung der Hypothese 4

Berechnet wurden Produkt-Moment-Korrelationskoeffizienten zur Prüfung der **H4.1**: *Praxisorientierung der Ausbildung, Leistungsdruck und Einstellung zur Arbeit korrelieren mit der schriftlichen Note.*

Nach Bortz (1993, S. 341) sind Kovariaten in einem varianzanalytischen Modell nur dann in der Lage, zur Reduzierung der Fehlervarianz beizutragen, wenn sie mit dem Kriterium korrelieren. Bevor das endgültige Modell erstellt wird, muß dieser Zusammenhang überprüft werden.

Tabelle 28: Korrelationen zwischen dem Kriterium und den Kovariaten

	L_SCORE	M_SCORE	P_SCORE
SCHRIFTL	,1462	-,3473	-,1308
	(24)	(24)	(24)
	P= ,495	P= ,096	P= ,542
M_SCORE	= Einstellung zur Arbeit		
L_SCORE	= Leistungsdruck		
P_SCORE	= Praxisorientierung		
SCHRIFTL	= schriftliche Note		

Wie in Hypothese 4 formuliert weisen Leistungsdruck, Praxisorientierung und Einstellung zur Arbeit eine Korrelation mit dem Kriterium auf. Leistungsdruck und Praxisorientierung korrelieren nur geringfügig mit der Ausbildungsleistung. Die Signifikanz kann aus schon genannten Gründen nicht als Ausschlußkriterium angesehen werden.

Zur Prüfung der Varianzaufklärung unter Berücksichtigung von Leistungsdruck, Einstellung zur Arbeit sowie der Praxisorientierung, wird in einer Kovarianzanalyse sukzessive das Inkrement der multiplen Bestimmtheit errechnet, das jede Kovariate erbringt. Hypothese 4 ist, soweit Korrelationen mit dem Kriterium vorhergesagt werden, als bestätigt anzusehen. Ob diese Hypothese auch hinsichtlich der **H4.2**: *Die Varianzaufklärung der schriftlichen Noten durch die beiden Stufen des Faktors Selektion wird durch Aufnahme von Leistungsdruck, Praxisorientierung und Einstellung zur Arbeit einzeln oder in Kombination miteinander verbessert.* zutrifft, wird nun überprüft.

Dazu wird zunächst eine einfaktorielle Varianzanalyse über die durchschnittliche schriftliche Abschlußnote der Anwärter der BGen mit dem zweifach gestuften Faktor Selektion gerechnet (vgl. Tabelle 29).

Wie aus Tabelle 29 zu ersehen, ist die varianzaufklärende Leistung des Faktors Selektion ist praktisch Null. In Hypothese 4 wird eine bessere Varianzerklärung durch die Berücksichtigung der Kovariaten vorhergesagt. Da nicht für alle in Tabelle 29 berücksichtigten BGen Kovariate erhoben werden konnten, reduziert sich bei einer Kovarianzanalyse das Sample von 33 auf 24 BGen. Der Erklärungszuwachs ist also nicht durch den Vergleich von Tabelle 29 mit dem Ergebnis der Kovarianzanalyse zu ermitteln, da die Effekte durch die Kovariaten konfundiert wären mit eventuellen Effekten die auf das Ausscheiden der BGen ohne erfaßte Kovariate zurückgehen. Tabelle 30 zeigt die Ergebnisse der einfaktoriellen Varianzanalyse für die 24 BGen, für die Praxisorientierung, Leistungsdruck und Einstellung zur Arbeit bekannt sind.

Tabelle 29: Ergebnisse der einfaktoriellen Varianzanalyse (Faktor Selektion/2fach gestuft)

SCHRIFTL					
by SELEKT					
UNIQUE sums of squares					
All effects entered simultaneously					
Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	8.735	1	8.735	.898	.351
SELEKT	8.735	1	8.735	.898	.351
Explained	8.735	1	8.735	.898	.351
Residual	301.534	31	9.727		
Total	310.269	32	9.696		
33 cases were processed.					
0 cases (.0 pct) were missing.					
SCHRIFTL = schriftliche Note					
SELEKT = Auswahlverfahren					
				R ² = 0.028	

Wie Tabelle 30 zu entnehmen ist, zeigt sich hier im Vergleich zu Tabelle 29 kein wesentlich anderes Ergebnis. Durch das Ausscheiden der BGen von denen keine beantworteten Fragebögen vorlagen, hat sich die durch den Faktor Selektion erklärte Varianz um 0.4% verringert.

Tabelle 30: Ergebnisse der einfaktoriellen Varianzanalyse unter Ausschluß der BGen, von denen kein beantworteter Fragebogen vorlag

SCHRIFTL					
by SELEKT					
UNIQUE sums of squares					
All effects entered simultaneously					
Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	2.164	1	2.164	.535	.472
SELEKT	2.164	1	2.164	.535	.472
Explained	2.164	1	2.164	.535	.472
Residual	88.893	22	4.041		
Total	91.057	23	3.959		

24 cases were processed.
0 cases (.0 pct) were missing.

R²= 0.024

Der durch die Berücksichtigung von Kovariaten erreichte Erklärungszuwachs kann nun durch einen Vergleich von Tabelle 30 mit Tabelle 31 bestimmt werden. Danach gewinnt der Faktor Selektion 1% Erklärungswert, wenn man als Kovariate Leistungsdruck einfügt. Praxisorientierung, Einstellung zur Arbeit oder aber Leistungsdruck in Kombination mit einer oder beiden anderen Kovariaten führt dagegen nicht zu einer Verbesserung der Varianzaufklärung durch den Faktor Selektion. Im Gegenteil verliert der Faktor Selektion bei Aufnahme der Kovariaten Einstellung zur Arbeit alleine oder in Kombination mit einer der beiden anderen Kovariaten sogar an Erklärungswert. Den größten Beitrag zur Klärung der Varianz in den Noten leistet der Faktor Selektion in Kombination mit Leistungsdruck ($R^2_{\text{Selekt, Note}} = 0.03$).

Tabelle 31: Durch den Faktor Selektion unter Berücksichtigung verschiedener Kovariaten aufgeklärte Notenvarianz (vollständige Ergebnisse aller Kombinationen im Anhang D)

Faktor	Kovariate1	Kovariate2	Kovariate2	Kovariate2
		L	M	P
Selektion mit	L	$R^2_{\text{Selekt}} = 0.03$ $R^2_{\text{L-Wert}} = 0.02$ $R^2_{\text{Gesamt}} = 0.05$	$R^2_{\text{Selekt}} = 0.01$ $R^2_{\text{M-Wert}} = 0.09$ $R^2_{\text{L-Wert}} = 0.00$ $R^2_{\text{Gesamt}} = 0.14$	$R^2_{\text{Selekt}} = 0.02$ $R^2_{\text{P-Wert}} = 0.03$ $R^2_{\text{L-Wert}} = 0.04$ $R^2_{\text{Gesamt}} = 0.07$
Selektion mit	M		$R^2_{\text{Selekt}} = 0.01$ $R^2_{\text{M-Wert}} = 0.11$ $R^2_{\text{Gesamt}} = 0.14$	$R^2_{\text{Selekt}} = 0.01$ $R^2_{\text{M-Wert}} = 0.13$ $R^2_{\text{P-Wert}} = 0.03$ $R^2_{\text{Gesamt}} = 0.17$
Selektion mit	P			$R^2_{\text{Selekt}} = 0.02$ $R^2_{\text{P-Wert}} = 0.01$ $R^2_{\text{Gesamt}} = 0.03$

L = Leistungsdruck

M = Einstellung zur Arbeit

P = Praxisorientierung

Da $R^2_{\text{Selekt, M, Note}}$ und $R^2_{\text{Selekt, M, L, Note}}$ jeweils mehr Varianz erklären als aufgrund der Summe der von den einzelnen Prädiktoren erklärten Varianz zu erwarten wäre, scheint hier ein sog. Suppressoreffekt zu bestehen (Moosbrugger & Klutky, 1987, S. 87). Die Effekte sind jedoch sehr schwach und die als Suppressoren in Frage kommenden Variablen Einstellung zur Arbeit und Leistungsdruck korrelieren, wenn auch schwach, mit dem Kriterium, so daß insoweit das Vorliegen einer echten Suppressorvariable zweifelhaft ist.

In der Kovarianzanalyse wurden auch 4 BGen berücksichtigt, deren Kennwerte sich lediglich auf Basis von weniger als drei beantworteten Fragebögen errechnete (vgl. Tabellen 19, 20, 21). Da hier eine Kennzeichnung der Organisation lediglich aus der Sicht eines oder zweier Anwärter vorgenommen wurde, ist die Beziehung zwischen den Kennwerten als Prädiktoren und der schriftlichen Note als Kriterium besonders unscharf. Um die Stabilität der Effekte beurteilen zu können, wäre es wichtig, Richtung und Ausmaß der Ergebnisveränderung bei Eliminierung dieser "schwachen" Messungen feststellen zu können. Daher wird die Kovarianzanalyse mit dem Faktor Selektion und der Kovariaten, die nach Tabelle 31 die größte Reduzierung der Fehlervarianz bewirkt, ohne die BGen mit weniger als drei beantworteten Fragebögen nochmals durchgeführt. Das untersuchte Sample reduziert sich in dieser Analyse von 24 auf 20 BGen.

Tabelle 32: Kovarianzanalyse unter Ausschluß der BGen, bei denen die Kennwerte nur auf Basis von weniger als zwei beantworteten Fragebögen errechnet wurden

SCHRIFTL						
by SELEKT						
with P_SCORE						
M_SCORE						
UNIQUE sums of squares						
All effects entered simultaneously						
Source of Variation	Sum of Squares	DF		Mean Square	F	Sig of F
Covariates	28.691	2	R ² = 0.32	14.346	4.059	.038
P_SCORE	3.508	1	R ² = 0.04	3.508	.993	.334
M_SCORE	26.190	1	R ² = 0.29	26.190	7.411	.015
Main Effects	2.258	1		2.258	.639	.436
SELEKT	2.258	1	R ² = 0.03	2.258	.639	.436
Explained	33.092	3	R ² = 0.37	11.031	3.121	.055
Residual	56.546	16		3.534		
Total	89.638	19		4.718		
20 cases were processed.						
0 cases (.0 pct) were missing.						
SELEKT = Auswahlverfahren						
M_SCORE = Einstellung zur Arbeit						
P_SCORE = Praxisorientierung						

Beschränkt man sich in der Analyse auf die BGen, deren Kennwerte auf Basis von mehr als drei Fragebögen erfaßt wurden, zeigt sich eine Erhöhung der Varianzaufklärung, die man als Stabilisierung der gefundenen Effekte beschreiben kann. Der Faktor Selektion gewinnt dabei jedoch nicht an Erklärungswert.

Tabelle 33: Varianzaufklärung des Faktors Selektion und der Kovariaten Einstellung zur Arbeit bei Berücksichtigung aller Fälle und nach Ausschluß "schwacher" Fälle, bei denen eine zuverlässige Berechnung der in die Analyse eingehenden BG-Kennwerte nicht möglich war

Varianzaufklärung durch:	Auf Basis aller Fälle	Ohne BG'n, deren Kennwerte nur aufgrund eines Fragebogens berechnet wurden
Selektion	1 %	3 %
Motivation	13 %	29 %
Praxisorientierung	3%	4%

6 Diskussion

Die Praxis der Auswahl geeigneter Bewerber entspricht bei den befragten BGen in den Jahren 1991/92 (Einstellungsjahrgänge der Prüfungsjahrgänge 1994/95) den allgemein üblichen Verfahren. Die Auswertung von Lebenslauf, Zeugnis, Bewerbungsschreiben sind auch hier wie in der Wirtschaft und der sonstigen staatlichen Verwaltung üblich. Etwa ein Drittel der befragten BGen führten im Rahmen der Eignungsdiagnostik auch Testverfahren durch (vgl. Tab. 2). Jedoch nur bei drei BGen wurden hierbei Experten hinzugezogen.

Der in Hypothese 1 angenommene Unterschied zwischen der Ausbildungsleistung von BGen, die durch Test auswählten, gegenüber denen, die durch Interview auswählten, fand sich nicht. Hinsichtlich der durchschnittlichen schriftlichen Note auf beiden Stufen des Faktors Selektion bestand kein praktisch bedeutsamer Unterschied (vgl. Tab. 22, 29). Für die Quote der Abbrecher fand sich zwar ein Unterschied, aber zugunsten der Interviewbedingung, was der Hypothese widersprach (vgl. Tab. 23). BGen, die durch Tests auswählten, hatten dabei mit 12% eine doppelt so hohe Abbrecherrate als interviewende BGen. Die Quote der Anwärter, die den Prüfungsanforderungen nicht genügten, unterschied sich tendenziell ebenfalls entgegen der angenommenen Richtung (vgl. Tab. 24). Die in Hypothese 1 angenommene bessere Ausbildungsleistung von BGen, die durch Testverfahren auswählten, konnte somit in den Daten nicht bestätigt werden. Die durch ein Testverfahren auswählenden BGen schnitten sogar tendenziell schlechter ab als die ausschließlich durch ein Eignungsgespräch auswählenden (vgl. Tab. 22). Man kann davon ausgehen, daß kein bedeutsamer Unterschied in der Vorhersageleistung der Noten zwischen den beiden Stufen des Faktors Selektion besteht. Dies steht im Gegensatz zu gesicherten Forschungserkenntnissen über die prädiktive Validität von Test und Interview (Scott, 1916; Anderson, 1960; Bolster & Springbett, 1961; Arvey & Campion, 1982; Carlson, 1967a, b; Carlson, 1968, Carlson, 1969; Carlson & Mayfield, 1967).

Dieser Befund kann folgende Ursachen haben:

1. Im Bereich der gewerblichen BGen werden Interviews besonders sorgfältig und auf hohem methodischen Niveau durchgeführt, so daß sie in ihrer prädiktiven Validität an die für Tests berichteten Leistungen heranreichen.
2. Die durchgeführten Tests sind schlecht konstruiert und/oder schlecht durchgeführt und/oder schlecht ausgewertet und/oder zur Vorhersage beruflicher Leistung nicht geeignet.

Aufgrund der Ergebnisse der telefonischen Befragung der an den Einstellungsgesprächen der BGen teilnehmenden Personen, kann die erste mögliche Ursache ausgeschlossen werden. Die Interviews im berufsgenossenschaftlichen Bereich werden keineswegs besser oder strukturierter durchgeführt als in der freien Wirtschaft oder der sonstigen Verwaltung. Ganz im Gegenteil weisen sie gerade die Mängel auf, die ursächlich für die geringe Vorhersageleistung dieses Instruments sind: Mangelnde Vergleichbarkeit der gewonnenen Daten wegen fehlender Standardisierung, Auswertung teilweise erst nachdem weitere Bewerber gesprochen wurden, keine Trennung von Beobachtung und Bewertung, Auswertung im freien Gespräch ohne Zuhilfenahme expliziter Bewertungskriterien, sondern auf Basis des Gesamteindrucks (vgl. Tab. 7), verschwommene Vorstellung von dem, was Berufseignung ausmacht. Eignungen werden ohne direkten Bezug zu Anforderungen, die aus einer planmäßigen Analyse der Arbeitstätigkeit abgeleitet wurden, beurteilt (vgl. Tab. 5).

Die mit den Befunden am ehesten zu vereinbarende Erklärung ist, daß die verwendeten Testverfahren psychodiagnostischen Anforderungen nicht genügen. Da der wissenschaftliche Charakter des Testverfahrens in dieser Untersuchung nicht vorausgesetzt wurde, deutet sich hier eine der Ursachen für das Fehlen eines Mittelwertunterschieds an. Nur drei BGen arbeiten mit Experten zusammen oder führen Tests durch, die sich in Verwaltung und Wirtschaft zu Zwecken der Berufseignungsdiagnostik bewährt haben (vgl. Tab. 1). Eine solch unterschiedliche Praxis führt zu einer sehr heterogenen Gruppe. Diese Heterogenität zeigt sich in Abbildung 8 deutlich. Während dort die interviewenden BGen sich in einem eng umschriebenen Bereich der Notenskala bewegen, finden sich die testenden BGen über einen weiten Bereich verstreut. Die Standardabweichung der Note ist für die

testenden BGen mehr als doppelt so groß (2.2 zu 4.3) als die der interviewenden. Dies bedeutet inhaltlich, daß durch Tests sowohl sehr gute als auch unterdurchschnittliche Ergebnisse in der Vorhersage des Ausbildungserfolgs erreicht werden. Diese heben sich im Mittel gegenseitig auf, so daß der eigentliche Unterschied zwischen der Test- und der Interviewbedingung in der Variationsbreite der Ergebnisse zu sehen ist.

Die in Hypothese 2 angenommene negative Korrelation zwischen der Praxisorientierung der Ausbildung und der Ausbildungsleistung, operationalisiert durch die schriftliche Note, bestand tendenziell. Es ergab sich ein Korrelationskoeffizient von $r_{\text{Praxisorientierung, Note}} = -0.13$ (vgl. Tab. 26). Dieser Zusammenhang ist weder besonders deutlich noch signifikant, und es kann sich daher um ein Zufallsergebnis handeln. Auch der Mittelwertunterschied in den schriftlichen Noten zwischen beiden durch Medianhalbierung gewonnenen Stufen der Variablen Praxisorientierung sind im Verhältnis zur Standardabweichung gering (vgl. Tab. 27). In keinem Fall sind an diesen Befund größere Interpretationen zu knüpfen. Dazu bedürfte es einer weiteren Aufklärung des Zusammenhangs zwischen Praxisorientierung und Ausbildungsleistung. Bei einer solchen weiteren Aufklärung wäre es von besonderer Bedeutung, das Konzept der Praxisorientierung besser meßbar zu machen, als dies in der vorliegenden Untersuchung gelungen ist.

Der in Hypothese 3 vermutete Moderatoreffekt der Variablen Praxisorientierung auf den Faktor Selektion konnte zwar tendenziell in den Daten nachgewiesen werden (vgl. Tab. 27, Abb. 11). Im Hinblick auf die gefundenen Unterschiede, die sich im Bereich von Zehntel Notenpunkten bewegen, verbietet sich jedoch jede inhaltliche Interpretation. Angesichts der nur unzureichend genauen Messung der Praxisorientierung, können solch feine Unterschiede nicht als gesicherte Befunde angesehen werden.

Die in Hypothese 4 postulierten Korrelationen zwischen den Kovariaten und dem Kriterium bestanden (vgl. Tab. 28). Allerdings sind die gefundenen Zusammenhänge ausgesprochen schwach. Inhaltlich überraschend ist die negative Korrelation zwischen der Einstellung zur Arbeit und der schriftlichen Note. Paradoxerweise ist

gerade in den BGen, deren Anwärter eine besonders positive Einstellung zur Arbeit haben, die durchschnittliche schriftliche Note als Operationalisierung des Ausbildungserfolgs niedrig. Man macht die Ausbildung zwar gerne, aber schlecht. Der Leistungsdruck korreliert schwach positiv mit dem Ausbildungserfolg. Denkbar wäre daher, daß ohne Leistungsdruck höhere Zufriedenheit herrscht, aber auch schlechtere Noten erzielt werden, während hoher Leistungsdruck zwar bessere Leistungen erzeugt, aber auch Unzufriedenheit.

Die in Hypothese 4 behauptete Reduzierung der Fehlervarianz durch die Aufnahme von Kovariaten in das einfaktorielle varianzanalytische Modell zeigte sich in einer Erhöhung der durch den Faktor Selektion erklärten Varianz um 1% auf 3% (vgl. Tab 30, 31). Das gesamte Modell erklärte dabei maximal 17% Varianz, wobei dies hauptsächlich auf die Regression mit den Kovariaten zurückzuführen ist. Formal ist auch der zweite Teil der Hypothese 4 bestätigt, da das Ausmaß der Verbesserung der Varianzaufklärung nicht spezifiziert ist. Angesichts der Geringfügigkeit der Verbesserung und der in nicht-experimentellen Designs meist erheblichen Störvarianz ist eine inhaltliche Interpretation eines so geringen Zuwachses nicht angebracht.

Zwischen der varianzaufklärenden Kapazität des Faktors Selektion in einer einfachen Varianzanalyse und einer Kovarianzanalyse mit bis zu drei Kovariaten besteht praktisch kein Unterschied. Das zeigt, daß maßgeblich für diesen fehlenden Zusammenhang nicht ein "Rauschen" in den Daten, also eine ungenaue Messung, ist. Sofern zwischen zwei Variablen kein Zusammenhang besteht, kann er auch nicht durch genauere Messung oder die Kontrolle von Störvariablen hergestellt werden. Der Einsatz der Kovariaten erzielte durchaus eine Wirkung, wie an der Verbesserung der erklärten Gesamtvarianz vom $R^2_{\text{Selekt, Note}} = 0.02$ auf $R^2_{\text{M-Wert, P-Wert, Selekt, Note}} = 0.17$ zu erkennen ist (vgl. Tab. 31). Da sich jedoch an der durch den Faktor Selektion erklärten Varianz dabei praktisch nichts änderte, weist dies darauf hin, daß nicht Meßungenauigkeiten ursächlich für diese geringe Varianzaufklärung sind. Auch bei der Eliminierung der BGen, deren Kennwerte sich nur auf Basis von weniger als 3 beantworteten Fragebögen errechneten, zeigte sich zwar eine

Verbesserung der durch die Kovariaten aufgeklärten Varianz, aber für den Faktor Selektion ergab sich praktisch keine Veränderung.

Diese Untersuchung fand nicht unter Laborbedingungen statt, so daß die Erfordernisse wissenschaftlicher Genauigkeit, besonders die für die Aufklärung von Kausalitäten notwendige Kontrolle über Variablen und Versuchsbedingungen, nicht gegeben waren. Diese Situation ist typisch für Evaluationsvorhaben, die den Anforderungen wissenschaftlicher Forschung in der Regel nicht entsprechen können. Wottawa & Thierau (1990, S. 16) weisen darauf hin,

" ... daß man bei allen Evaluationsprojekten immer Kritikpunkte finden wird, da die ideale Konzeption eines solchen Vorhabens selbst bei unbegrenzten Ressourcen ... niemals realisierbar ist und einschränkende Setzungen durch den Verantwortlichen eines solchen Projekts erfordert."

Einige dieser einschränkenden Setzungen, zum Beispiel die Operationalisierung des Ausbildungserfolgs durch die Noten der Abschlußprüfung, und ihrer Konsequenzen sollen im folgenden besprochen werden. Man kann davon ausgehen, daß keine deterministische Beziehung zwischen der Abschlußnote und der Qualität eines Sachbearbeiters besteht. Die Noten beziehen sich auf den rein fachlichen Aspekt der Arbeit. Andere Fähigkeiten, die zum erfolgreichen Bewältigen der Aufgabe ebenso wichtig sein können, wie z.B. soziale Kompetenz, bilden sich in den Noten nicht ab. Durch die weiteren Operationalisierungen (Ausbildungsabbrecher, Wiederholer) wurde eine allzu eingeschränkte Auffassung des Ausbildungserfolgs vermieden. Andere Maße für den Ausbildungserfolg, wie Vorgesetztenbeurteilungen, waren im Rahmen dieser Studie nicht zu erhalten. Der Vorteil der Noten lag in ihrer leichten Erreichbarkeit, da sie zentral archiviert sind. Daher konnten zwei komplette Prüfungsjahrgänge mit fast 800 Personen erfaßt werden.

Auch das Fehlen willkürlich manipulierter unabhängiger Variablen und der Randomisierung schränkt die Möglichkeit zu kausalen Interpretationen der Ergebnisse ein. Für diese Untersuchung kann auch nicht ausgeschlossen werden, daß durch die Selbstselektion der Anwärter auf die BGen Gruppenunterschiede

zustandegekommen sind, die auf keine der erfaßten Variablen zurückgehen. Allerdings spricht nicht sehr viel für diese Vermutung, da von Bewerberseite keine Unterschiede zwischen den verschiedenen BGen, die ohnehin nicht besonders bekannt in der Öffentlichkeit sind, wahrgenommen werden. Wäre ein Vergleich zwischen Unternehmen der freien Wirtschaft und Organisationen der öffentlichen Verwaltung angestrebt worden, spräche sicher vieles für Unterschiede in den Charakteristiken der Teilnehmer beider Ausbildungsgänge. Da in der vorliegenden Untersuchung jedoch lediglich Organisationen verglichen wurden, die allesamt zur öffentlichen Verwaltung, und zwar zum Zweig der Sozialversicherung, genauer der gesetzlichen Unfallversicherung, gehören und alle nach denselben Vorschriften anstellen und bezahlen, kann man sagen, daß sich für den Bewerber die verschiedenen BGen lediglich in ihrem Namen unterscheiden. Es handelt sich für den Bewerber nicht um die Wahl verschiedener Berufe.

Die angestrebte Totalerhebung aller gewerblichen BGen wurde nicht erreicht. Vier BGen antworteten schon auf das erste Anschreiben nicht. Es gelang jedoch herauszufinden, welche Selektionsmethode diese BGen durchführten, so daß zumindest für die Beurteilung des Effekts des Faktors Selektion auf die Ausbildungsleistung die angestrebte Totalerhebung erreicht werden konnte. Die 27 erfaßten BGen erhielten Fragebögen zugesandt. Von 23 BGen wurden ausgefüllte Fragebögen zurückgesandt (vgl. Tab. 10). Die 31 zu untersuchende Einheiten (vgl. Abschnitt 5.1) erhöhten sich auf 33 durch die Aufgliederung einer von zwei BGen, die den Stufen des Faktors Selektion aufgrund der abweichenden Auswahlmethoden einzelner Bezirksverwaltungen nicht klar zuzuordnen waren (vgl. Abschnitt 5.5.6), auf 33. Davon konnten 24 in die Kovarianzanalyse einbezogen werden, was einer Erfassung der Grundgesamtheit der gewerblichen BGen von 73% entspricht. Direkte Gültigkeit können die Ergebnisse aufgrund des deskriptiven Charakters dieser Untersuchung nur für die erfaßten BGen haben. Angesichts der grundsätzlich gleichen Verwaltungsstrukturen und der nicht wesentlich anderen Ausbildungskonzeptionen läßt sich aber kein Grund dafür erkennen, warum die Ergebnisse nicht auch auf die fehlenden BGen übertragbar sein sollten.

Fraglich ist, ob die 156 Personen, die den Fragebogen zur Erfassung der Praxisorientierung, des Leistungsdrucks und der Einstellung zur Arbeit ausfüllten als für die Grundgesamtheit der Absolventen der BGA in den Jahren 1994/95 repräsentativ zu betrachten sind. Nur unter dieser Voraussetzung wäre der Schritt der Kennzeichnung der BGen mit den durchschnittlichen Testwerten ihrer ehemaligen Anwärter gerechtfertigt. Die Stichprobe war zum Prüfungszeitpunkt knapp zwei Monate jünger als die Grundgesamtheit, hatte 2.2% weniger Frauen, war 2.15 Notenpunkte (bezogen auf eine 100 Punkte Skala) besser, und durch Interview ausgewählte Personen waren weniger vertreten (vgl. Tabelle 11).

Da Anwärter mit schlechten Noten mit höherer Wahrscheinlichkeit nicht in den Dienst übernommen werden als solche mit besseren, überrascht der gefundene Notenunterschied nicht. Den geringeren Frauenanteil könnte man durch persönliche Gründe, insbesondere Heirat und Schwangerschaft, erklären. Der Altersunterschied sowie der um 0.4% erhöhte Anteil der durch Interview ausgewählten Personen ist geringfügig und bedarf daher keiner weiteren Aufklärung. Man kann feststellen, daß die gefundenen Abweichungen der Stichprobenwerte von denen in der Grundgesamtheit keinen Anlaß zur Vermutung bieten, es handle sich bei der Stichprobe um eine Extremgruppe. Auch wenn es weiterhin denkbar bleibt, daß hinsichtlich anderer Merkmale keine Vergleichbarkeit besteht, darf man hier jedoch feststellen, daß es aufgrund der Daten keine begründeten Zweifel an der Repräsentativität der Stichprobe für die Prüfungsjahrgänge 1994/95 gibt.

Problematisch bei der Messung von Organisationseigenschaften per Befragung einzelner Organisationsmitglieder ist, daß Organisation und Person Quellen der Varianz sind. Da jede Person nur eine Organisation kennt und daher auch nur eine beurteilen kann, ist ein individueller Score nicht eindeutig auf Organisation oder Person zurückzuführen. Durch die Aggregation der Daten über mehrere Beurteiler kann diese Personenvarianz eliminiert werden. Wenn die befragten Mitglieder einer Organisation systematisch höhere oder niedrigere Scores erzielen als die Mitglieder anderer Organisationen, muß dies auf organisationale Bedingungen (Selektion, Rahmenbedingungen der Ausbildung, Betriebsklima, etc.) zurückgeführt werden. Dies gilt zumindest dann, wenn für jede Organisation genügend Aussagen zur

Verfügung stehen. Leider ist dies in der vorliegenden Untersuchung wegen des geringen Rücklaufs der Fragebögen nicht immer gesichert. Da auch nur eine begrenzte Anzahl von Prüflingen der Jahrgänge 94/95 zur Verfügung stehen, einzelne BGen haben in diesem Zeitraum sogar überhaupt nicht ausgebildet, kann auch durch Versendung einer größeren Zahl von Fragebögen der Rücklauf (vgl. Tab. 10) nicht beliebig erhöht werden. Die Kennwerte der Organisation konnten teilweise nur auf Basis von einem Fragebogen ermittelt werden. Andere Organisationen mußten aus der Analyse herausgelassen werden, weil kein beantworteter Fragebogen vorlag (vgl. Tab. 10). Aus diesen Gründen sind die Ergebnisse der Varianzanalyse nur eingeschränkt aussagefähig, sofern Feststellungen über die erfaßten BGen hinaus getroffen werden sollen.

Differenziert zu sehen ist auch der Versuch, die BGen durch die Berechnung von Kennwerten zu bezeichnen. Dies stellt sicherlich eine Abstraktion von der Realität dar. BGen sind keine monolithischen Blöcke, sondern Organisationen, die sich - zumindest gilt das für die größeren BGen - mit ihren Untergliederungen, den Bezirksverwaltungen, über das ganze Bundesgebiet erstrecken. Manche BGen gliedern sich in über 10 Bezirksverwaltungen. Natürlich kann die Ausbildung in der Bezirksverwaltung München in ihrer Realität ganz anders aussehen als die in der Bezirksverwaltung Hamburg, obwohl beide zur selben BG gehören. Auch hängt, wie schon in der Einleitung angesprochen, Leistungsdruck, Praxisorientierung und Einstellung zur Arbeit von individuellen Charakteristiken ab. Gerade diese sollen aber durch die Berechnung von Kennwerten für die Organisation eliminiert werden, um so Charakteristiken der Organisation klarer hervorzubringen. Für die bundesweit arbeitenden BGen gilt, daß die Organisationskultur sich auch in Richtlinien und Konzepten der Ausbildung abbildet, die innerhalb einer BG bundesweit gelten. Die Anwärter werden keineswegs ausschließlich in einer Bezirksverwaltung praktisch ausgebildet, sondern lernen andere Bezirksverwaltungen, mindestens jedoch die Hauptverwaltung der eigenen BG kennen, in der sie Teile ihrer Ausbildung in zentralen Abteilungen absolvieren müssen. Bei über 50% der befragten BGen ist diese Frage unproblematisch, weil dort Anwärter ausschließlich in ihrer Hauptverwaltung eingestellt werden, so daß hier eine große Ähnlichkeit der Ausbildungsbedingungen für alle Anwärter der gesamten BG bestand.

Von zentraler Bedeutung für diese Untersuchung ist auch die Qualität der Messung der Variablen Leistungsdruck, Einstellung zur Arbeit und Praxisorientierung. Mangels geeigneter standardisierter Verfahren zur Messung dieser Konstrukte, wurden eigene Verfahren entwickelt. Dabei wurde bewußt versucht, die Zahl der Items gering zu halten, um die Rücklaufquote der Fragebögen nicht zu vermindern. Je genauer eine Variable gemessen werden kann, desto klarer sind Zusammenhänge mit anderen Variablen analysierbar. Inwieweit dies in der vorliegenden Untersuchung gelungen ist, muß daher diskutiert werden. Die Antwortverteilungen der Skalen zur Messung der Einstellung zur Arbeit und des Leistungsdrucks sowie die Antwortverteilung auf die Frage nach der Praxisorientierung der Ausbildung, sind nicht normalverteilt (vgl. Tab. 13, 15 & Abb. 5, 6, 7). Bei Leistungsdruck und Einstellung zur Arbeit liegt, im Gegensatz zur Praxisorientierung, allerdings eine unimodale und annähernd symmetrische Verteilung vor. Nach Lienert (1989, S. 176) kann eine anormale Antwortverteilung durch eine heterogene Analysestichprobe, mangelhafte Konstruktion des Tests oder aufgrund einer anormalen Verteilung des Merkmals in der Population verursacht sein.

Sicher ist ein Mangel der Skalen, daß sie wenige Items enthalten. So sind die Reliabilitäten von $\alpha = 0.67$ für die Skala Leistungsdruck mit 7 Items und $\alpha = 0.64$ für die Skala Einstellung zur Arbeit mit 5 Items als hoch anzusehen. Hinsichtlich der Iteminterkorrelationen (vgl. Anhang C), der Varianz der Items und ihrer Trennschärfe (vgl. Tab. 12, 14, 16) ergibt sich kein Anhalt für Konstruktionsmängel. Außer der gefundenen Rohwertverteilung gibt es keine begründeten Hinweise dafür, daß die untersuchten Merkmale in der Population anormal verteilt sein könnten. Eine heterogene Analysenstichprobe als Ursache ist besonders für die bimodale Verteilungsform der Antworten auf die Frage nach der Praxisorientierung (vgl. Abb. 7) denkbar. Hier bietet sich als Interpretation der bimodalen Antwortverteilung an, daß zwei Extremgruppen sich überlagernde Normalverteilungen erzeugt haben.

Aufgrund der in dieser Untersuchung gewonnen Daten läßt sich letztlich nicht klären, ob die Verteilung durch die Zusammensetzung der untersuchten Stichprobe, Mängel des Meßinstruments oder auch durch beides zugleich zustande kam. Die

Normalverteilung der Testrohwerte ist auch keine notwendige Voraussetzung für die Verwendbarkeit der Testwerte, sondern vor allem im Hinblick auf eine spätere Standardisierung eines Testverfahrens erwünscht. Für die weitere Analyse sind die mittels der konstruierten Skalen gewonnenen Daten aber unabhängig von einer Standardisierung zu gebrauchen.

Testtheoretisch problematisch ist die Tatsache, daß die benutzten Testverfahren zur Messung des Leistungsdrucks und der Einstellung zur Arbeit an denselben Personen konstruiert wurden mit denen auch die Hypothesenprüfung stattfand. Die nach der klassischen Testtheorie berechneten Gütekriterien, insbesondere die Reliabilitätskoeffizienten, sind Größen, die auf korrelativen Beziehungen beruhen. Die Korrelation aber ist es, die " ... untrennbar mit der Merkmals- bzw. Testwertverteilung in der jeweiligen Population zusammenhängt." (Moosbrugger, 1990/91, S. 112). Damit bezeichnen nach der klassischen Testtheorie Testwerte sie nicht nur das konstruierte Testverfahren, sondern ebenso die Population an der das Verfahren erprobt wurde. Da die konstruierten Verfahren jedoch ausschließlich zum Zweck dieser Untersuchung konstruiert wurden. Vergleiche mit anderen Populationen werden nicht angestrebt. Denkbar ist jedoch, daß für die Gesamtheit der 1994/95 geprüften Anwärter die Testwerte anders aussähen als sie sich für die Subgruppe der Personen die den Fragebogen zurücksandten zeigten. In Tabelle 11 wurden für einige der zentralen Variablen Vergleiche zwischen der Stichprobe und der Grundgesamtheit vorgenommen, ohne daß sich wesentliche Unterschiede zeigten.

Viele Praktiker werden gegen die Operationalisierung der abhängigen Variablen Ausbildungserfolg einwenden, daß durch die Selektionsmethode nicht die schriftliche Note des Anwärters vorhergesagt werden soll, sondern die Berufseignung. Aus den Telefoninterviews ergab sich, daß als Ziel des Einstellungsgesprächs, das einschließlich der testenden alle untersuchten BGen durchführten (Tab. 1), zumeist auch Aspekte wie die "Passung" des Bewerbers in die Organisation, seine sozialen Fähigkeiten u.ä. erfaßt werden sollten (vgl. Tab. 6, 9). Inwieweit dieser Anspruch verwirklicht werden konnte, war nicht Inhalt dieser Untersuchung. Dazu hätten aufwendige Datenerhebungen in den BGen, Vorgesetzten- und Mitarbeiterbeurteilungen etc. erhoben werden müssen, was aus

Kosten- und Zeitgründen nicht möglich war. Das finale Kriterium für die Beurteilung der Wirksamkeit einer Selektionsmethode ist der Berufserfolg. Dieser kann abschließend im Prinzip erst nach der Pension des Mitarbeiters beurteilt werden.

"Kein Evaluationsprojekt kann die 'endgültigen' Folgen einer Maßnahme bewerten; es ist immer notwendig, Zwischenziele festzulegen und die Optimierung des Verhaltens an einem solchen, letztlich willkürlich gesetzten Zwischenstadium auszurichten." (Wottawa & Thierau, 1990, S. 15)

Ein Zusammenhang von Abschlußnote an der BGA zu späterem Berufserfolg, also die Validität der Noten, dürfte nicht nur dem Anspruch der BGA nach, sondern auch der Forschungslage zur prädiktiven Validität von Schul- und Studiennoten entsprechend, gegeben sein (Baron-Boldt & Schuler, 1988). Die Praxis, aufgrund der Abschlußnote an der BGA auch über die Übernahme der Anwärter zu entscheiden, manchmal sogar das Erreichen konkreter Notenwerte als für die Einstellung notwendig zu bezeichnen oder die Dauer bis zur ersten Beförderung von der Note abhängig zu machen, unterstellt diesen Zusammenhang zwischen Berufseignung und Abschlußnote implizit bereits. Wenn aufgrund der Selektionsmethode noch nicht einmal der Ausbildungserfolg, operationalisiert durch die schriftliche Abschlußnote, zuverlässig über einen Zeitraum von drei Jahren vorhergesagt werden kann, ist es ausgesprochen unwahrscheinlich, daß dies beim Berufserfolg, dessen Vorhersage über eine größere zeitliche Distanz erfolgen müßte, möglich ist. Wie sich aus bei der Analyse des Notenarchivs zeigte, ist die mündliche Note durchschnittlich um 7,82 Punkte besser als die schriftliche (vgl. Tab. 3). Dies weist auf die Tendenz hin, mittels der mündlichen Note "Härtefälle" in der Notengebung zu vermeiden. Da solche Erwägungen bei der schriftlichen Note keine Rolle spielen, ist sie, wie bereits in Abschnitt 4.2 erörtert, als reliablere Größe anzusehen.

Angesichts der vorliegenden Befunde kann gesagt werden, daß die Praxis der Personalselektion im Bereich der gewerblichen BGen in den Einstellungsjahren 1991/92 suboptimal war. Sofern Interviews geführt wurden, geschah dies in einer Form, die keine vergleichbaren Daten für die verschiedenen Bewerber lieferte. Es

handelte sich praktisch ausschließlich um sog. "Freie Gespräche", deren Verlauf nur grob strukturiert war. Zum Teil wurden die Vergleiche zwischen verschiedenen Bewerbern erst am Ende des Tages oder nachdem sich alle Bewerber vorgestellt hatten, vorgenommen, so daß Reihenfolgeeffekte, implizite Persönlichkeitstheorien, Stereotype und Gedächtniseffekte in ihrer Wirkung voll zur Entfaltung kommen konnten.

Aber auch die Testpraxis beseitigte diese Mängel nicht. In der Regel handelte es sich um sog. "hausinterne" Tests, die von Laien konstruiert, durchgeführt und ausgewertet wurden und deren Objektivität, Reliabilität und Validität keiner Überprüfung unterzogen wurden. Daher ist es nicht verwunderlich, wenn sich entgegen gesicherter Untersuchungsbefunde hier keine bessere prädiktive Validität ergibt. Zwischen den verschiedenen angewandten Testverfahren gab es hinsichtlich der Güte der Vorhersage des Ausbildungserfolgs große Unterschiede.

Praxisorientierte Ausbildungen gingen tendenziell mit schlechteren Noten in der Abschlußprüfung an der BGA einher. Inwiefern durch eine solche Ausbildung aber Kompetenzen gestärkt werden, die für die Tätigkeit als Sachbearbeiter wichtig sind, konnte hier nicht geklärt werden. Eine Bewertung der Konsequenzen praxisorientierter Ausbildungskonzepte wurde hier nur im Hinblick auf die Leistungen in der Abschlußprüfung 1994/95 an der BGA vorgenommen.

Abschließend bleibt festzuhalten, daß die Durchführung von Testverfahren bei den BGen in den Jahren 1991/92 im Vergleich zu Interviews nicht zu den erhofften besseren Ausbildungsleistungen führten. Soweit sie sich seitdem nicht wesentlich verändert haben, sind sie deutlich zu verbessern. Die Durchführung von Einstellungsgesprächen, aber auch die Konstruktion, Durchführung und Auswertung von Testverfahren, sind als nicht optimal zu bezeichnen.

7. Zusammenfassung

Die gewerblichen Berufsgenossenschaften (BGen) unterscheiden sich im Hinblick auf die Selektion geeigneter Bewerber für den gehobenen Verwaltungsdienst. Neben der Auswertung der üblichen Bewerbungsunterlagen kommen bei der Auswahl Einstellungsgespräche und Testverfahren zum Einsatz. Außer in der Methode der Selektion geeigneter Bewerber unterscheiden sich die BGen auch in der Gestaltung der praktischen Ausbildung in der Verwaltung. Dagegen findet die theoretische Ausbildung für die Auszubildenden aller BGen einheitlich in Form von Blockseminaren an der "Berufsgenossenschaftlichen Akademie für Arbeitssicherheit und Verwaltung" (BGA) statt. Dort werden auch die Abschlußprüfungen vorgenommen.

Anhand der 1991/92 eingestellten und 1994/95 geprüften Dienstanwärter wurden die Hypothesen geprüft, daß durch Testverfahren auswählende BGen bessere Ausbildungserfolge erreichen als durch Interview auswählende sowie die Hypothese, daß eher praktisch ausgebildete Anwärter schlechtere Prüfungsleistungen erbringen als weniger praktisch ausgebildete. Außerdem wurde eine Moderatorwirkung der Praxisorientierung auf die beiden Stufen des Faktors Selektion angenommen sowie eine Fehlervarianz reduzierende Wirkung von Leistungsdruck, Einstellung zur Arbeit und Praxisorientierung als Kovariaten in einem varianzanalytischen Modell.

Als Operationalisierungen des Ausbildungserfolgs werden dabei die durchschnittlichen schriftlichen Noten der Dienstanwärter jeder BG in den Jahren 1994/95 angesehen. Außerdem werden die Quote der Ausbildungsabbrecher und der durch die Prüfung gefallenen Anwärter als Indikatoren des Ausbildungserfolgs erhoben. Leistungsdruck und Einstellung zur Arbeit werden zum Zweck der Reduzierung der Fehlervarianz per Fragebogen erfaßt und mit der Praxisorientierung in ein einfaktorielles kovarianzanalytisches Modell zur Erklärung der Notenvarianz eingefügt. Zur Messung der Variablen Leistungsdruck und Einstellung zur Arbeit wurden Skalen entwickelt. Etwa 500 Fragebögen sind zusammen mit exploratorischen Fragen und einer Frage zur Messung der Praxisorientierung der Ausbildung an die BGen verschickt und dort an

Sachbearbeiter weitergeleitet worden, die im Beobachtungszeitraum an der BGA geprüft wurden. Auf Basis des Rücklaufs von 156 beantworteten Fragebögen wurden Kennwerte für jede BG berechnet, die den Ausbildungserfolg, die Abbrecherquote, die Quote der durchgefallenen Anwärter, den Leistungsdruck, die Praxisorientierung der Ausbildung und die Einstellung zur Arbeit quantifizieren.

Das Einstellungsgespräch als Methode der Personalselektion weist in der Praxis der BGen die aus der Literatur bekannten methodischen Mängel auf, welche zu den geringen Validitätswerten dieses Instruments führen. Freie Gesprächsführung führt zu nicht vergleichbaren Ergebnissen, fehlende Trennung von Bewertung und Beobachtung zu selbsterfüllenden Prophezeiungen statt objektiver Diagnostik. Allerdings konnte demgegenüber die Gruppe der BGen, die durch Testverfahren ausgewählt, keinen Vorsprung in der Validität vorweisen.

Der angenommene Unterschied zwischen der Ausbildungsleistung von BGen, die durch Test ausgewählt, gegenüber denen, die durch Interview ausgewählt, fand sich nicht. Die schriftliche Note auf beiden Stufen des Faktors Selektion unterschied sich nicht bedeutend. Für die Quote der Abbrecher fand sich zwar ein Unterschied, aber zugunsten der Interviewbedingung, was der Hypothese widersprach. Die Quote der Anwärter, die den Prüfungsanforderungen nicht genügten, unterschied sich tendenziell ebenfalls entgegen der angenommenen Richtung. Die Erklärung ist, daß die verwendeten Testverfahren psychodiagnostischen Anforderungen nicht genügen. Die Standardabweichung der Note ist für die testenden BGen mehr als doppelt so groß als die der interviewenden. Der eigentliche Unterschied zwischen der Test- und der Interviewbedingung liegt daher in der Variationsbreite der Ergebnisse.

Die angenommene negative Korrelation zwischen der Praxisorientierung der Ausbildung und der Ausbildungsleistung, operationalisiert durch die schriftliche Note, bestand tendenziell. Es ergab sich ein Korrelationskoeffizient zwischen Praxisorientierung und Note von $r = -0.13$. Der vermutete Moderatoreffekt der Variablen Praxisorientierung auf den Faktor Selektion konnte tendenziell in den Daten nachgewiesen werden .

Es bestanden schwache Korrelationen zwischen den Kovariaten und dem Kriterium. Eine Reduzierung der Fehlervarianz durch die Aufnahme von Kovariaten in das einfaktorielle varianzanalytische Modell bestand nur um 1%. Das gesamte Modell erklärte maximal 17% Varianz.

Die für die Aufklärung von Kausalitäten notwendige Kontrolle über Variablen und Versuchsbedingungen waren in der vorliegenden Untersuchung nicht gegeben. Diese Situation ist typisch für Evaluationsvorhaben, die den Anforderungen wissenschaftlicher Forschung in der Regel nicht entsprechen können.

Hiermit erkläre ich, daß ich diese Arbeit selbstständig verfaßt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Wiesbaden, den 29.10.1996

7 Literaturverzeichnis

- Amelang, M. & Bartussek, D.**, Differentielle Psychologie und Persönlichkeitsforschung (3. Aufl.). Stuttgart Berlin Köln: Kohlhammer, 1990.
- Anderson, C.W.**, The relation between speaking times and decision in the employment interview. *Journal of Applied Psychology*, 1960, 44, 267-268.
- Arvey, R.D. & Campion, J.E.**, The employment interview: A summary and review of recent research. *Personnel Psychology*, 1982, 35, 281-322.
- Baron-Boldt, J. & Schuler, H. & Funke, U.**, Prädiktive Validität von Schulabschlußnoten: Eine Metaanalyse. *Zeitschrift für Pädagogische Psychologie*, 1988, 2, 79-90.
- Borkenau, P.**, Sind erste Eindrücke valide, und wenn ja, warum? In: Schuler, H. & Funke, U. (Hrsg.), *Eignungsdiagnostik in Forschung und Praxis*. Göttingen: Hogrefe, 1991.
- Bolster, B.L. & Springbett, B.M.**, The reaction of interviewers to favourable and unfavorable information. *Journal of Applied Psychology*, 1961, 45, 97-103.
- Bortz, J.**, *Statistik für Sozialwissenschaftler* (4. Aufl.). Berlin Heidelberg New York Tokyo: Springer, 1993.
- Brandstätter, H.**, Persönliche Verhaltens- und Leistungsbedingungen. In: Schuler, H. (Hrsg.), *Lehrbuch Organisationspsychologie*. Bern Stuttgart Toronto: Huber, 1993.
- Carlson, R.E.**, Selection interview decisions: The relative influence of appearance and factual written information on an interviewer's final rating. *Journal of Applied Psychology*, 1967 (a), 51, No. 6, 461-468.
- Carlson, R.E.**, Selection interview decisions: The effect of interviewer experience, relative quota situation, and applicant sample on interviewer decisions. *Personnel Psychology*, 1967 (b), 20, 259-280.
- Carlson, R.E.**, Employment decisions: Effect of mode of applicant presentation on some outcome measures. *Personnel Psychology*, 1968, 21, 193-207.
- Carlson, R.E.**, Relative influence of a photograph vs. factual written information on an interviewer's employment decision. *Personnel Psychology*, 1969, 22, 45-56.
- Carlson, R.E. & Mayfield, E.**, Evaluating interview and employment application data. *Personnel Psychology*, 1967, 20, 441-460.

- Cascio, W.F.**, Personnel Selection in the Context of Human Resource Management. In: Schuler, H. & Funke, U. (Hrsg.): Eignungsdiagnostik in Forschung und Praxis. Göttingen: Hogrefe, 1991.
- Dreher, G.F., Ash, R.A. & Hancock, P.**, The role of traditional research design in underestimating the validity of the employment interview. *Personnel Psychology*, 1988, 41, 315-327.
- Heckhausen, H.**, Motivation und Handeln. Berlin Heidelberg New York London Paris Tokyo Hong Kong: Springer, 1988.
- Hunter, J.E. & Hirsh, H.R.**, Applications of meta-analysis. In: Cooper, C.L. & Robertson, I.T. (Eds.), *International Review of industrial and organizational psychology* (pp. 321-357). New York: Wiley, 1987.
- Hunter, J.E. & Hunter, R.F.**, Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 1984, 96, 72-98.
- Lienert, G.A.**, Testaufbau und Testanalyse (4. Aufl.). München Weinheim: PVU, 1989.
- Meehl, P.E.**, Clinical vs. statistical prediction. Minneapolis, Minn.: University of Minnesota Press, 1954.
- Moosbrugger, H.**, Arbeiten aus dem Institut für Psychologie an der Johann Wolfgang Goethe-Universität Frankfurt am Main, Testkonstruktion und Testtheorie. 1990/91.
- Moosbrugger, H. & Klutky, N.**, Regressions- und Varianzanalysen auf der Basis des Allgemeinen Linearen Modells. In: Pawlik, K. (Hrsg.), *Methoden der Psychologie*. Band 6, Bern Stuttgart Toronto: Huber, 1987.
- Moser, K.**, Personalmarketing. Eine Einführung. München: Quintessenz, 1992.
- Moser, K.**, Vergleich unterschiedlicher Wege der Gewinnung neuer Mitarbeiter. *Zeitschrift für Arbeits- und Organisationspsychologie*, 1995, 39, (N.F.13), 3, S. 105-114
- Neuberger, O.**, Das Einstellungsgespräch. *Psychologie und Praxis*, 1974, 18, 8-22.
- Premack, S.L. & Wanous, J.P.**, (1985). A meta-analysis of realistic job preview experiments. *Journal of Applied Psychology*, 70, 706-790.
- Schmidt, N. & Noe, R.A.**, Personnel selection and equal employment opportunity. In: Cooper, A. & Robertson, I.T. (Hrsg.), *International review of industrial and organizational psychology*. 71-115, Cichester, 1986.
- Schoenborn, C.**, Informationssuche und Informationsgewichtung im Kontext der Personalselektion. Dissertation an der Universität Mannheim, Fakultät für Sozialwissenschaften, 1988.

Schorr, A., Diagnostische Praxis in der Arbeits- und Organisationspsychologie, Teilergebnisse aus einer repräsentativen Umfrage zur diagnostischen Praxis. In: Schuler & Funke (Hrsg.): Eignungsdiagnostik in Forschung und Praxis, Göttingen: Hogrefe, 1991.

Schuler, H. & Funke, U., Tests sollen keine Tortur sein. Psychologie heute, 1987, 14 (9), 58-65

Schuler, H. & Funke, U., Diagnose beruflicher Eignung und Leistung. In: Schuler, H. (Hrsg.), Lehrbuch Organisationspsychologie, Bern Stuttgart Toronto: Huber, 1993.

Schuler, H., (Hrsg.), Lehrbuch Organisationspsychologie. Bern Stuttgart Toronto: Huber, 1993.

Scott, W.D., Selection of employees by means of quantitative determinations. Annals of the american academy of political and social science, 65, 1916.

Wanous, J.P. , Organizational entry. Reading, MA: Addison Wesley, 1992.

Weinert, A.B., Lehrbuch der Organisationspsychologie. PVU, Weinheim, 3. Aufl. 1992

Wiesner, W.H. & Cronshaw, S.F., A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview. Journal of occupational Psychology, 1988, 61, 275-290.

Wiggins, J.S., Personality and prediction: Principles of personality assessment. Reading, Mass.: Addison-Wesley, 1973.

Wottawa, H. & Thierau, H., Evaluation, Bern Stuttgart Toronto: Huber, 1990.