

Metaanalytische Studien
zu Intelligenz und Berufsleistung
in Deutschland

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

vorgelegt von
Jochen Kramer

aus
Tettnang

Bonn 2009

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

Zusammensetzung der Prüfungskommission:

Prof. Dr. Rainer Banse, Institut für Psychologie
(Vorsitzender)

Prof. Dr. Gerhard Blickle, Institut für Psychologie
(Betreuer und Gutachter)

Prof. Dr. Walter Neubauer, Institut für Psychologie
(Gutachter)

PD Dr. Bernd Schlöder, Institut für Psychologie
(weiteres prüfungsberechtigtes Mitglied)

Tag der mündlichen Prüfung: 18. März 2009

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn unter
http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

Zusammenfassung

Der Zusammenhang von Allgemeiner Intelligenz und beruflichem Erfolg in Deutschland wird metaanalytisch untersucht. Drei Berufserfolgs-Bereiche werden betrachtet: a) berufsbezogene Lernleistung, gemessen mit Berufsschulnoten und betrieblichen Leistungs-Beurteilungen ($k = 210$, $N = 30\,451$), b) spezifische Arbeitsleistung, gemessen mit subjektiven Fremdbeurteilungen ($k = 18$, $N = 2\,739$) und c) Karriereerfolg, der mit objektiven Einkommensmaßen ($k = 7$, $N = 1\,980$) und Beförderungs- und Positionsmaßen (berufliches Vorankommen; $k = 9$, $N = 1\,817$) operationalisiert wird.

Um alle relevanten Studien berücksichtigen zu können – auch solche, die insignifikante Korrelationen nicht berichten – wird ein Substitutionsverfahren eingeführt, das auf den berichteten insignifikanten Korrelationen basiert. Die Brauchbarkeit des Verfahrens wird im Vergleich mit anderen Substitutionsverfahren und der Exklusion der betroffenen Studien eingeschätzt. Die Messungenauigkeiten der Intelligenz- und Kriterienerhebungen sowie die Varianzeinschränkung von Intelligenz werden korrigiert. Die metaanalytische Aggregation wird mit der Methode von [Hunter und Schmidt \(2004\)](#) durchgeführt.

Mit diesem Vorgehen kann gezeigt werden, dass Allgemeine Intelligenz kriterienvalide ist in Bezug auf berufliche Lernleistung (korrigierte Validität = .62), subjektiv bewertete Arbeitsleistung (.66), das Einkommen (.35) und berufliche Vorankommen (.33). Der positive Zusammenhang zwischen Intelligenz und den Erfolgsmaßen kann über verschiedene in Deutschland verwendete Intelligenztests und Leistungs- bzw. Erfolgsindikatoren sowie unterschiedliche Berufsgruppen generalisiert werden. Da die Varianzaufklärung jeweils unter 75% liegt, ist die Höhe der Zusammenhänge nicht generalisierbar, sondern von Moderatorinflüssen abhängig. Moderatoranalysen werden durchgeführt, mit denen die Einflüsse verschiedener Testverfahren, Leistungskriterien, Berufskomplexitäten und Berufsgruppen auf Lernleistung untersucht werden können. Auch bei ihrer Berücksichtigung verbleiben mehr als 25% der Varianz zwischen den Studien ungeklärt.

Implikationen der Befunde für die weitere Forschung werden besprochen (z.B. zu welchen Testverfahren, Kriterien und Berufsgruppen Validierungsstudien fehlen) und die Bedeutung der Ergebnisse für Personalauswahl und -platzierung wird aufgezeigt.

Dank

Diese Arbeit wurde ermöglicht durch die engagierte Unterstützung vieler Personen. Mein Dank gilt zunächst allen, die unpublizierte Studien oder weitergehende Information zu publizierten Studien zur Verfügung gestellt haben.

Eva Sindern und Nike Feshareki haben die zweite Kodierung der Primärstudien durchgeführt. Das Korrekturlesen der Arbeit haben Matthias Mühlenberg, Christel Haarer und Danijela Pucko übernommen. Den Kodiererinnen und KorrekturleserInnen danke ich für ihr Durchhaltevermögen und ihr Engagement, mit denen sie diese Arbeiten erledigt haben.

Ich danke Prof. Dr. Gerhard Blickle für seine wertvollen Anregungen zur Durchführung der Arbeit und seine hilfreiche Unterstützung. Dr. Alexander Witzki, Jun.-Prof. Dr. Marc Solga und Paula Schneider danke ich für zahlreiche hilfreiche Diskussionen und Anregungen sowie ihren freundschaftlichen Zuspruch.

Ich danke Prof. Dr. Walter Neubauer für die Übernahme der Zweitbegutachtung dieser Arbeit sowie Prof. Dr. Rainer Banse und PD Dr. Bernd Schlöder für die Übernahme der Funktionen in der Prüfungskommission.

Ein besonderer Dank gilt Lydia und Alfred Kramer und von Herzen Roland Müller, die mich während der Durchführung der Arbeit in vielerlei Hinsicht geduldig unterstützt und ermuntert haben.

Inhaltsverzeichnis

Tabellenverzeichnis	8
Abbildungsverzeichnis	12
Abkürzungsverzeichnis	13
1. Einleitung	15
2. Theorie	17
2.1. Intelligenz	17
2.2. Beruflicher Erfolg	28
2.2.1. Bereiche von Berufserfolg	28
2.2.2. Indikatoren für beruflichen Erfolg	32
2.3. Intelligenz und beruflicher Erfolg	36
2.3.1. Intelligenz und berufliche Lernleistung	38
2.3.2. Intelligenz und Arbeitsleistung	46
2.3.3. Intelligenz und Karriereerfolg	54
2.3.4. Warum eine weitere Metaanalyse?	56
3. Methode	58
3.1. Literaturrecherche	58
3.2. Studieninklusion und -exklusion	59
3.2.1. Allgemeine Intelligenz	59
3.2.2. Erfolgskriterien	76
3.2.3. Stichprobe und Zeitabstand zwischen Intelligenz- und Kriterienerhebung	79
3.2.4. Effektgrößen	80
3.2.5. Studienqualität	82
3.3. Statistische Unabhängigkeit	82
3.4. Kodierung	83
3.4.1. Allgemeine Informationen	83
3.4.2. Prädiktor	85
3.4.3. Kriterium	86
3.5. Kumulierung abhängiger Werte	87
3.6. Artefaktkorrekturen	90
3.6.1. Unreliabilität der Prädiktoren	90
3.6.2. Unreliabilität der Kriterien	92
3.6.3. Varianzeinschränkung	95

3.7.	Fehlende Primärstudien-Daten	98
3.7.1.	file-drawer-Analysen und Funnelplots	99
3.7.2.	Substitutionsverfahren	100
3.8.	Aggregation der Studieneffekte und Homogenitätstests	104
4.	Metaanalytische Ergebnisse	110
4.1.	Intelligenz und berufsbezogene Lernleistung	111
4.1.1.	Beschreibung der inkludierten Studien	111
4.1.2.	Metaanalytische Ergebnisse bei Inklusion aller Lernleistungsstudien	118
4.1.3.	Sensitivitätsanalysen	120
4.1.4.	Moderatoranalysen	132
4.2.	Intelligenz und Arbeitsleistung	151
4.2.1.	Beschreibung der inkludierten Studien	151
4.2.2.	Metaanalytische Ergebnisse bei Inklusion aller Arbeitsleistungsstudien	156
4.2.3.	Sensitivitätsanalysen	157
4.3.	Intelligenz und Karriereerfolg	160
4.3.1.	Beschreibung der inkludierten Studien	163
4.3.2.	Metaanalytische Ergebnisse für die Indikatorgruppen Einkommen und berufliches Vorankommen	169
4.3.3.	Sensitivitätsanalysen	173
5.	Diskussion	177
5.1.	Verhältnissubstitution	177
5.2.	Berufsbezogene Lernleistung	180
5.2.1.	Validität im Hinblick auf verschiedene Lernleistungskriterien	182
5.3.	Subjektiv bewertete Arbeitsleistung	183
5.4.	Karriereerfolg	185
5.4.1.	Einkommen	185
5.4.2.	Berufliches Vorankommen	186
5.5.	Berufskomplexität	187
5.6.	Berufsgruppen	189
5.7.	Intelligenztests	190
5.8.	Generalisierbarkeit	192
5.9.	Weiterer Forschungsbedarf	194
5.10.	Bedeutung für die Personalauswahl	196
	Literaturverzeichnis	198
	Anhang	221
A.	Formeln	221
A.1.	Korrektur punktbiserialer Korrelationen	221

A.2. Kummulierung abhängiger Werte	221
A.3. Berechnungen der bare-bone Analysen	222
A.4. Operationale Validität und vollständig korrigierte Analysen	223
A.4.1. Berechnungen auf Studienebene	223
A.4.2. Berechnungen über alle Studien hinweg	224
A.5. Symmetriekennwerte für Funnelplots	225
B. Kodieranweisung	226
B.1. Studien-Kodierbogen	226
B.2. BIS-Kodierbogen	237
C. Kodierungen	239
C.1. Quellen und Studiennummern	239
C.2. Übersicht der Metaanalysen pro Studie	249
C.3. BIS-Kodierung	254
C.4. ISCO-88-Kodierung	261
D. Gewichte ausgewählter Variablen an den Analysen	267
E. Inkludierte Studien und Ergebnisse auf Studienebene	275

Tabellenverzeichnis

2.1.	Übersicht über die Inhalte des Theorieteils	18
2.2.	Beispiele für Berufserfolgs-Indikatoren getrennt für verschiedene Erfolgsbereiche, Datenarten und Datenquellen	33
2.3.	Metaanalytische Befunde aus den USA zur prädiktiven Validität von Intelligenz in Bezug auf Lernleistung	39
2.4.	Metaanalytische Befunde aus Europa zur prädiktiven Validität von Intelligenz in Bezug auf Lernleistung	43
2.5.	Metaanalytische Befunde aus den USA zur prädiktiven Validität von Intelligenz in Bezug auf Arbeitsleistung	48
2.6.	Metaanalytische Befunde aus Europa zur prädiktiven Validität von Intelligenz in Bezug auf Arbeitsleistung	52
2.7.	Metaanalytische Befunde zur prädiktiven Validität von Intelligenz in Bezug auf Karriereerfolg	55
3.1.	Begriffe zur Suche nach geeigneten Studien	59
3.2.	Anerkannte Intelligenz-Verfahren	61
3.3.	Angaben zu Dimensionen und Verfahren der subjektiven Leistungsbeurteilungen	78
3.4.	Anzahl der inkludierten Studien mit verschiedenen Berufserfolgs-Indikatoren	80
3.5.	ISCO-88 Berufshauptgruppen und Skill Level	84
3.6.	Beurteilerübereinstimmung	88
3.7.	Schutz- und Prüfmechanismen gegen Verzerrungen	99
3.8.	VW- und Substitutions-Werte	105
4.1.	Gewicht kategorialer Moderatorvariablen an der Metaanalyse zu berufsbezogener Lernleistung	112
4.2.	Lernleistung: Inkludierte Intelligenztestversionen	115
4.3.	Geschätztes Gewicht der BIS-Zellen an der Metaanalyse zur berufsbezogenen Lernleistung bei Inklusion aller Studien	116
4.4.	Lernleistung: bare-bone Analyse	120
4.5.	Lernleistung: korrigierte Validität	122
4.6.	Lernleistung: bare-bone Sensitivitätsanalyse publizierter und nicht-publizierter Studien	123
4.7.	Lernleistung: korrigierte Validitäten publizierter und nicht-publizierter Studien	124
4.8.	Lernleistung: bare-bone Analyse zu verschiedenen Quellen	126
4.9.	Lernleistung: korrigierte Validitäten für verschiedene Quellen	127

4.10. Lernleistung: bare-bone Analyse zum Einfluss des BET-Manuals	128
4.11. Lernleistung: korrigierte Validitäten zum Einfluss des BET-Manuals	129
4.12. Lernleistung: bare-bone Analyse zum Einfluss der Veröffentlichungen vor 1945	130
4.13. Lernleistung: korrigierte Validitäten zum Einfluss der Veröffentlichungen vor 1945	131
4.14. Lernleistung: bare-bone Sensitivitätsanalyse zum Einfluss konkurrender Validierungen	132
4.15. Lernleistung: korrigierte Validitäten zum Einfluss konkurrender Validität	133
4.16. Lernleistung: bare-bone-Analysen getrennt für Skill Level 2 und 3	134
4.17. Lernleistung: korrigierte Validitäten getrennt für Skill Level 2 und 3	135
4.18. Lernleistung: bare-bone-Analyse zum Vergleich geringer und hoher Bil- dungsniveaus	136
4.19. Lernleistung: korrigierte Validitäten getrennt für geringe und hohe Bil- dungsniveaus	137
4.20. Lernleistung: bare-bone-Analyse getrennt für ISCO-88 Berufs-Hauptgrup- pen	138
4.21. Lernleistung: korrigierte Validitäten getrennt für ISCO-88 Berufs-Haupt- gruppen	139
4.22. Lernleistung: bare-bone Moderatoranalyse verschiedener Intelligenztests .	141
4.23. Lernleistung: operationale Validitäten verschiedener Intelligenztests	144
4.24. Lernleistung: vollständig korrigierte Validitäten verschiedener Intelligenz- tests	145
4.25. Lernleistung: bare-bone Metaanalysen von Berufsschulnoten und betrieb- lichen Beurteilungen	147
4.26. Lernleistung: korrigierte Validitäten getrennt für Berufsschulnoten und betriebliche Beurteilungen	148
4.27. Lernleistung: bare-bone Metaanalysen für Aus- und Weiterbildungsleis- tung	149
4.28. Lernleistung: korrigierte Validitäten getrennt für Aus- und Weiterbildungs- leistung	150
4.29. Gewicht kategorialer Moderatorvariablen an den Gesamtmetaanalysen zu Arbeitsleistung	152
4.30. Geschätztes Gewicht der BIS-Zellen an der Metaanalyse zur Arbeitsleis- tung bei Inklusion aller Studien	154
4.31. Arbeitsleistung: bare-bone Metaanalyse	157
4.32. Arbeitsleistung: korrigierte Validitäten	158
4.33. Arbeitsleistung: bare-bone Sensitivitätsanalysen	159
4.34. Arbeitsleistung: korrigierte Validitäten der Sensitivitätsanalysen	162
4.35. Gewicht kategorialer Moderatorvariablen an den Metaanalysen zum Ein- kommen und beruflichen Vorankommen	164
4.36. Geschätztes Gewicht der BIS-Zellen an der Metaanalyse zum Einkommen	166

4.37. Geschätztes Gewicht der BIS-Zellen an den Metaanalysen zum beruflichen Vorankommen bei Inklusion aller Studien	169
4.38. Karriereerfolg: bare-bone Metaanalysen zum Einkommen und beruflichem Vorankommen	171
4.39. Karriereerfolg: korrigierte Validitäten für Einkommen und berufliches Vorankommen	172
4.40. Berufliches Vorankommen: bare-bone Analyse bei Exklusion von Nicht-Pearson-Korrelationen	174
4.41. Berufliches Vorankommen: korrigierte Validitäten bei Exklusion von Nicht-Pearson-Korrelationen	176
C.1. Zuordnung der Quellen und Stichproben zu den Studiennummern	239
C.2. Zuordnung der Studiennummern zu den inkludierten Quellen und Stichproben	244
C.3. Übersicht über die Metaanalysen pro Studie	249
C.4. Einordnung der Intelligenzaufgaben in die BIS-Klassifikation	254
C.5. ISCO-88-kodierte Berufe der Studien zur berufsbezogenen Lernleistung	261
C.6. ISCO-88-kodierte Berufe der Studien zur Arbeitsleistung	266
C.7. ISCO-88-kodierte Berufe der Studien zum Karriereerfolg	266
D.1. Lernleistung: bare-bone Gewichte ausgewählter Variablen an den Sensitivitäts- bzw. Moderatoranalysen	268
D.2. Arbeitsleistung: bare-bone Gewichte ausgewählter Variablen an den Sensitivitäts- bzw. Moderatoranalysen	272
D.3. Einkommen: bare-bone Gewichte ausgewählter Variablen an den Sensitivitätsanalysen	273
D.4. Berufliches Vorankommen: bare-bone Gewichte ausgewählter Variablen an den Sensitivitätsanalysen	274
E.1. Studienebene: Lernleistung	275
E.2. Studienebene: Lernleistung, Studien ohne Mittelwert-Aggregate	279
E.3. Studienebene: Lernleistung, Studien mit Mittelwert-Aggregaten	282
E.4. Studienebene: Lernleistung, publizierte Studien	283
E.5. Studienebene: Lernleistung unpubliziert	286
E.6. Studienebene: Lernleistung, Testmanuale	287
E.7. Studienebene: Lernleistung, Artikel in Zeitschriften / Herausgeberwerken	288
E.8. Studienebene: Lernleistung, Monographien / Dissertationen	289
E.9. Studienebene: Lernleistung ohne BET-Manual	290
E.10. Studienebene: Lernleistung, nur BET-Manual	293
E.11. Studienebene: Lernleistung vor 1945	294
E.12. Studienebene: Lernleistung nach 1945	295
E.13. Studienebene: Lernleistung, prädiktive Validität	298
E.14. Studienebene: Lernleistung, konkurrente Validität	302
E.15. Studienebene: Lernleistung, Skill Level 2	302

E.16.Studienebene: Lernleistung, Skill Level 3	305
E.17.Studienebene: Lernleistung, max. Hauptschul-Abschluss	306
E.18.Studienebene: Lernleistung, (Fach-)Hochschulreife	306
E.19.Studienebene: Lernleistung, Militär (HG 0)	306
E.20.Studienebene: Lernleistung, Techniker und gleichrangige (HG 3)	307
E.21.Studienebene: Lernleistung, Bürokräfte (HG 4)	308
E.22.Studienebene: Lernleistung, Handwerker (HG 7)	308
E.23.Studienebene: Lernleistung, Maschinenbediener (HG 8)	310
E.24.Studienebene: Lernleistung, IST	310
E.25.Studienebene: Lernleistung, WIT	311
E.26.Studienebene: Lernleistung, BIS	312
E.27.Studienebene: Lernleistung, LPS/PSB	313
E.28.Studienebene: Lernleistung, Matrizentests	313
E.29.Studienebene: Lernleistung, GATB-basierte Verfahren	314
E.30.Studienebene: Lernleistung, andere Intelligenztests	315
E.31.Studienebene: Lernleistung, Berufsschulnoten	316
E.32.Studienebene: Lernleistung, betriebliche Beurteilungen	318
E.33.Studienebene: Lernleistung, Ausbildung	319
E.34.Studienebene: Lernleistung, Weiterbildung	322
E.35.Studienebene: Arbeitsleistung	323
E.36.Studienebene: Arbeitsleistung, publizierte Studien	323
E.37.Studienebene: Arbeitsleistung, prädiktive Validität	324
E.38.Studienebene: Arbeitsleistung, nur Pearson-Korrelationen	324
E.39.Studienebene: Arbeitsleistung ohne Studie 116.12	325
E.40.Studienebene: Einkommen	325
E.41.Studienebene: berufliches Vorankommen	325
E.42.Studienebene: Berufliches Vorankommen, nur Pearson-Korrelationen	326

Abbildungsverzeichnis

2.1. Das Berliner Intelligenzstruktur-Modell	25
2.2. Intelligenz und Beruf	44
2.3. Pfadmodell zum Zusammenhang zwischen Intelligenz und Arbeitsleistung	53
3.1. Verteilungen der Verhältniswerte	103
4.1. Zeitabstand zwischen Intelligenz- und Kriterienerhebung der prädiktiven Studien zur Lernleistung	117
4.2. Baumdiagramm der Lernleistungsstudien	119
4.3. Kumulatives Baumdiagramm der Lernleistungs-Metaanalyse	121
4.4. Funnelplot der Studien zu berufsbezogener Lernleistung	123
4.5. Lernleistung: vollständig korrigierte Validitäten für verschiedene Berufsgruppen	142
4.6. Lernleistung: operationale Validitäten verschiedener Intelligenztests . . .	143
4.7. Vollständig korrigierte Validitäten bei unterschiedlichen Lernleistungskriterien	149
4.8. Zeitabstand zwischen Intelligenz- und Kriterienerhebung der Studien zur Arbeitsleistung	155
4.9. Baumdiagramm der Arbeitsleistungs-Studien	156
4.10. Kumulatives Baumdiagramm der Arbeitsleistungs-Metaanalyse	158
4.11. Funnelplot der Studien zu Arbeitsleistung	161
4.12. Baumdiagramme der Studien zum Einkommen und beruflichen Vorankommen	167
4.13. Kumulative Baumdiagramme der Metaanalysen zum Karriereerfolg . . .	170
4.14. Validität von Intelligenz in Bezug auf vier berufliche Erfolgskriterien . . .	171
4.15. Funnelplots zum Karriereerfolg	175

Häufig verwendete Abkürzungen

AZUBI-BK	Arbeitsproben zur berufsbezogenen Intelligenz für Büro- und kaufmännische Tätigkeiten
AZUBI-TH	Arbeitsproben zur berufsbezogenen Intelligenz für technische und handwerkliche Tätigkeiten
B	BIS-Operation Bearbeitungsgeschwindigkeit
BET	Berufseignungstest
BIS	Berliner Intelligenzstruktur(-Test)
CE	Coefficient of Equivalence
CES	Coefficient of Equivalence and Stability
CFT	Grundintelligenztest
CI	Konfidenzintervall
CS	Coefficient of Stability
CSC	Cross-Situational Consistency
CV	Glaubwürdigkeitsintervall
DGP	Deutsche Gesellschaft für Personalwesen e.V.
DOT	Dictionary of Occupational Titles
E	BIS-Operation Einfallsreichtum
EUB	Testserie für Hauptschüler
F	BIS-Inhalt <i>anschauungsgebundenes Denken</i>
FM	Modell fester Effekte
FRT	Figure Reasoning Test
g	Generalfaktor Allgemeiner Intelligenz
g_c	kristalliner Intelligenzfaktor
g_f	fluider Intelligenzfaktor
GATB	General Aptitude Test Battery
GMA	General Mental Ability
HG	ISCO Berufs-Hauptgruppe
ISCO	International Standard Classification of Occupations
IST	Intelligenz-Struktur-Test
k	Studienanzahl
K	BIS-Operation Verarbeitungskapazität
LPS	Leistungsprüfsystem
M	BIS-Operation Merkfähigkeit
N	Gesamt-Stichproben-Größe
N	BIS-Inhalt <i>zahlengebundenes Denken</i>
PJAHR	Jahr der Publikation bzw. Durchführung einer Studie
Pop.	Population

PSB	Prüfsystem für Schul- und Bildungsberatung
$\hat{\rho}$	geschätzte wahre Korrelation in der Population
\bar{r}	n -gewichtete mittlere Korrelation
\bar{r}_{insig}	Mittel der berichteten insignifikanten Korrelationen
r_{xx}	Reliabilität von Intelligenzmessungen
r_{yy}	Kriterien-Reliabilität
RM	Modell zufallsvariabler Effekte
r_k	kritischer Wert einer Korrelation
SIS	Situational Specificity
SL	ISCO Skill Level
SP	Stichprobe
SPM(PLS)	Standard Progressive Matrices (Plus)
TRA	Transportability
u oder u_x	Wert der beobachteten Varianzeinschränkung (< 1) oder -vergrößerung (> 1)
V	BIS-Inhalt <i>sprachgebundenes Denken</i>
VE	Anteil der durch die Artefakte aufgeklärten Varianz an der beobachteten Varianz
VW	Verhältniswert
WIT	WILDE Intelligenztest
WPT	Wonderlic Personnel Test

1. Einleitung

Um Personalauswahl-Entscheidungen treffen zu können, sind Eignungsurteile zu fällen, die „ausschließlich auf größtmögliche prognostische Validität“ (Jäger, 1970, S. 641) abzielen. Der Nutzen der Personalauswahl hängt maßgeblich von der prognostischen Validität der Auswahlinstrumente ab (Görlich & Schuler, 2006). Dementsprechend hält es Hunter (1986) für zentral, nicht nur zu wissen, ob Auswahlinstrumente valide sind oder nicht, sondern auch das genaue Ausmaß ihrer Validität zu kennen. In mehreren Metaanalysen haben Hunter und Schmidt die Validitäten der gängigsten Auswahlinstrumente untersucht und festgestellt, dass die Validität von Intelligenz die der anderen Verfahren (z.B. unstrukturierte Vorstellungsgespräche, Probezeit, Überprüfung von Referenzen, Assessment-Center) übertrifft. Nahezu gleich groß ist die Validität kostenintensiverer Verfahren (Arbeitsproben und strukturierte Vorstellungsgespräche; F. L. Schmidt & Hunter, 1998a, 1998b). Als Validitätskriterien wurden Indikatoren der *training performance* und *job performance* verwendet. Die Studien, auf denen diese metaanalytischen Ergebnisse basieren, stammen überwiegend aus den USA. Da in Deutschland andere Auswahlinstrumente (z.B. deutschsprachige Intelligenztests, vgl. Kapitel 3.2.1) und Erfolgsindikatoren (Salgado, Anderson, Moscoso, Bertua & de Fruyt, 2003a) verwendet werden sowie ein anderes organisationales Umfeld herrscht (Salgado & Anderson, 2002, vgl. auch Kapitel 2.3), ist es fraglich, ob die Befunde aus den USA auch in Deutschland Gültigkeit besitzen. Für den Prädiktor *Intelligenz* soll diese Frage in dieser Arbeit metaanalytisch beantwortet werden.

Im Theorieteil werden die zentralen Konstrukte *Intelligenz* (Kapitel 2.1) und *beruflicher Erfolg* (Kapitel 2.2) definiert. Es werden Rahmenmodelle gewählt, die dabei helfen zu entscheiden, welche Studien als zum Gegenstandsbereich gehörend angesehen werden können. Mit welchen Theorien sich der Zusammenhang zwischen Allgemeiner Intelligenz und beruflicher Leistung bzw. beruflichem Erfolg erklären lässt und welche metaanalytischen Befunde dazu bereits vorliegen, wird in Kapitel 2.3 thematisiert. Dabei wird versucht, einen umfassenden Überblick über den Gegenstandsbereich zu geben, wie dies z.B. Cooper (1998) fordert: „Introductions are typically short in primary research reports. . . In research syntheses, introductions should be considerably more detailed. Synthesists should attempt to present a complete overview of the research question“ (S. 159). Aus diesem Überblick werden die Forschungsfragen abgeleitet (Kapitel 2.3.4).

Das methodische Vorgehen wird in Kapitel 3 geschildert. Bei einer Metaanalyse handelt es sich um einen Prozess, der in mehreren aufeinander aufbauenden Schritten abläuft. Die metaanalytische Integration voneinander unabhängiger Studieneffekte ist der letzte Schritt, dem mehrere vorbereitende auf Studienebene vorausgehen. Während dieses Prozesses sind zahlreiche Entscheidungen zu treffen. Einige wesentliche Fragen, die unterschiedlich beantwortet werden können, sind z.B.: Welche Inklusions- bzw. Exklusionskriterien werden angelegt (Kapitel 3.2)? Wie wird damit umgegangen, wenn eine

Studie mehrere abhängige Ergebnisse berichtet (Kapitel 3.5)? Welche Artefaktkorrekturen werden vorgenommen (Kapitel 3.6)? Wie wird mit fehlenden Primärstudien-Daten umgegangen (Kapitel 3.7)? Das Ergebnis der Metaanalyse wird maßgeblich davon mitbestimmt, welche Antworten auf diese Fragen gegeben werden. Eine umfassende Darstellung und Begründung der methodischen Vorgehensweise ist deshalb unerlässlich.

Die Ergebnisse der Metaanalysen werden getrennt für drei berufliche Erfolgsbereiche im Ergebnisteil dargestellt (Kapitel 4): die Validität von Intelligenz in Bezug auf berufliche Lernleistung (Kapitel 4.1), subjektiv bewertete Arbeitsleistung (Kapitel 4.2) und Karriereerfolg (Kapitel 4.3). Jedes Teilkapitel umfasst eine Beschreibung der inkludierten Studien, die Ergebnisse ihrer metaanalytischen Aggregation und eine Abschätzung der Belastbarkeit der Ergebnisse mit Hilfe von Sensitivitätsanalysen. Für den Bereich der Lernleistung wird zudem der Moderatoreinfluss verschiedener Intelligenztests, Kriterien, Berufskomplexitäten und -gruppen untersucht. Ergebnisse auf Primärstudien-Ebene – z.B. mit welchem Gewicht welche Effektstärken aus welchen Studien berücksichtigt wurden – können in tabellarischer Form Anhang E entnommen werden.

Die Diskussion beschäftigt sich zunächst mit einer kritischen Betrachtung des im Methodenteil vorgestellten Substitutionsverfahrens (Kapitel 5.1). Anschließend werden die Ergebnisse getrennt für die Kriterienbereiche *Lernleistung* (Kapitel 5.2), *Arbeitsleistung* (Kapitel 5.3) und *Karriereerfolg* (Kapitel 5.4) diskutiert, sowie die Einflüsse der Moderatoren *Berufskomplexität* (Kapitel 5.5), *Berufsgruppe* (Kapitel 5.6) und *Intelligenztest* (Kapitel 5.7) besprochen. Dabei wird thematisiert, welche Generalisierungen die Ergebnisse erlauben und wie Unterschiede zu bereits vorliegenden metaanalytischen Befunden erklärt werden können. Eine allgemeine Einschätzung der Generalisierbarkeit der Ergebnisse enthält Kapitel 5.8. Die Grenzen der durchgeführten Studien und den weiteren Forschungsbedarf thematisiert Kapitel 5.9. Die Bedeutung für die Personalauswahl ist Gegenstand des Kapitels 5.10.

2. Theorie

Eine Übersicht über die Inhalte des Theorieteils gibt Tabelle 2.1. Zunächst werden zentrale Charakteristika von Intelligenz definiert und ein Überblick über die historische Entwicklung des Intelligenzkonstruktes gegeben. Mit zunehmender Vielfalt an Intelligenzmodellen wurden die Intelligenzkonstrukte komplexer, sie erheben den Anspruch modellübergreifende Aspekte der älteren Intelligenzkonstrukte herauszuarbeiten (Rahmenmodelle). Einige Rahmenmodelle werden – getrennt für verschiedene Vorgehensweisen bei ihrer Entwicklung – vorgestellt. Schließlich wird eine Arbeitsdefinition von Intelligenz gewählt, die der Metaanalyse zu Grunde gelegt wird.

Anschließend wird eine Übersicht über mögliche Kriterien beruflichen Erfolgs gegeben. Dies geschieht in Anlehnung an eine Systematik von Dette, Abele und Renner (2004). Es werden zunächst drei Erfolgsbereiche (Laufbahn, spezifische Arbeit und berufsbezogene Lernleistung) unterschieden und darauf folgend mögliche Indikatoren für jeden Erfolgsbereich vorgestellt. Dabei werden zwei Datenarten unterschieden: neutrale Kennzahlen und Vergleiche mit Bezugsstandards.

Im dritten Teilkapitel werden a) die Konstrukte *Intelligenz* und *beruflicher Erfolg* miteinander verglichen, b) mögliche kulturelle Einflussfaktoren besprochen, c) theoretische Begründungen und bereits vorliegende metaanalytische Befunde für den Zusammenhang von Intelligenz mit berufsbezogener Lernleistung, Arbeitsleistung und Karriere dargestellt. Die vorgestellten metaanalytischen Befunde stammen aus den USA, Europa und Deutschland. Schließlich werden d) Gründe für eine weitere Metaanalyse benannt und die Fragestellungen der Arbeit spezifiziert.

2.1. Intelligenz

Es gibt eine Menge verschiedener Definitionen von Intelligenz. Sternberg (1987) meint sogar „there seem to be almost as many definitions of intelligence as there were experts asked to define it“ (S. 376). Ein Konsens darüber lässt sich bis heute nicht erzielen, obwohl es z.B. 1921 und 1981 im Rahmen von Expertenbefragungen versucht worden ist (Sternberg & Berg, 1986; zitiert nach Süß, 2001). Es soll hier kein umfassender Überblick über diese verschiedenen Definitionen gegeben werden; dies tun z.B. Sternberg und Detterman (1986), Rösing (2004, betrachtet auch Alltagsdefinitionen und Definitionen aus anderen Kulturkreisen) oder Legg und Hutter (2007, berücksichtigen auch künstliche Intelligenz). Vielmehr sollen zunächst einige für *Intelligenz* charakteristische Aspekte verdeutlicht und anschließend verschiedene Intelligenztheorien vergleichend betrachtet werden. Zum Schluss des Kapitels ist eine Arbeitsdefinition von Intelligenz zu wählen, die den Metaanalysen zu Grunde gelegt werden kann.

Legg und Hutter (2007) betrachten 70 verschiedene Intelligenzdefinitionen und arbeiten drei häufig genannte Aspekte heraus: a) die Fähigkeit eines Individuums mit

Tabelle 2.1

Übersicht über die Inhalte des Theorieteils

Kapitel / Inhalt	Seite
Intelligenz	17
▶ Bestimmungsstücke von Intelligenz / Definitionsversuche	17
▶ historische Entwicklung: • Galton	20
• Binet	
• Spearman	
• Thurstone	
• Cattell	
• Vernon	
▶ Rahmenmodelle von Intelligenz	22
▷ Faktoren-Eklektizismus	23
• French	
▷ Content-Definitionen	23
• Guilford	
• Guttman	
▷ metaanalytische Integration	24
• Carroll	
▷ Aufgaben-Integration	24
• Jäger (Berliner Intelligenzstruktur-Modell)	
▷ theoriegeleitete Modelle	26
• Sternberg	
• Amthauer (Hierarchisches Protomodell)	
▶ Arbeitsdefinition von Intelligenz	27
Beruflicher Erfolg	28
▶ Bereiche von Berufserfolg	28
▷ Laufbahn / Karriere	
▷ Arbeitsleistung	
▷ berufsbezogene Lernleistung	
▷ Zusammenfassung: Definitionen	31
▶ Indikatoren von Berufserfolg	32
▷ neutrale Kennzahlen zu Karriereerfolg, Arbeits- und Lernleistung	33
▷ Vergleiche mit Bezugsstandards zu Karriereerf., Arbeits- und Lernleist.	34
Intelligenz und beruflicher Erfolg	36
▶ Vergleich von Intelligenz- und Erfolgskonstrukten	36
▶ Vergleich kultureller Einflüsse aus den USA, Europa und Deutschland	37
▶ Intelligenz und berufsbezogene Lernleistung	38
▷ Theorie	
▷ metaanalytische Befunde (USA, Europa, Deutschland)	
▶ Intelligenz und Arbeitsleistung	46
▷ Theorie	
▷ metaanalytische Befunde (USA, Europa, Deutschland)	
▶ Intelligenz und Karriere	54
▷ Theorie	
▷ metaanalytische Befunde (USA)	
▶ Gründe für eine weitere Metaanalyse und Fragestellungen der Arbeit	56

seiner Umwelt oder seinen Umwelten zu interagieren, b) die Fähigkeit, Ziele erfolgreich zu verfolgen und c) die Fähigkeit, sich erfolgreich auf verschiedene Zielsetzungen und Umwelten einzustellen.

Punkt b) ist für diese Arbeit von besonderer Bedeutung, da die Erreichung *beruflicher Ziele* thematisiert wird (beruflicher Erfolg). Die Relevanz erfolgreicher Zielerreichung wurde bereits 1905 von [Binet und Simon](#) betont, die Intelligenz als die erfolgreiche Bewältigung einer aktuellen Situation bezeichneten. Später z.B. von [Wechsler \(1956\)](#), der in ihr die Fähigkeit eines Individuums sah, „zweckvoll zu handeln, vernünftig zu denken und sich mit seiner Umgebung wirkungsvoll auseinander zu setzen“ (S. 13), oder von [Hofstätter \(1957\)](#), der Intelligenz definierte als „das Ensemble von Fähigkeiten, das den innerhalb einer bestimmten Kultur Erfolgreichen gemeinsam ist“ (zitiert nach [Nettelstroth, 2003, S. 27](#)).

Dem entsprechend wurde Intelligenz mit zahlreichen Variablen in Beziehung gesetzt, die als Indikatoren eines erfolgreichen Lebens angesehen werden können (z.B. akademische Leistung, gesundheitsbezogenes Verhalten, soziale Leistung, Arbeitsleistung, Kreativität; vgl. [Brand, 1987](#); [Jensen, 1998](#); [Kuncel, Hezlett & Ones, 2004](#); [F. L. Schmidt, 2002](#)).

Ein Aspekt in dem sich die Definitionen unterscheiden ist ihre Breite. Ein Konsens besteht darin, dass der Begriff *Intelligenz* ausschließlich *kognitive* Fähigkeiten umfasst. Eine Ausweitung auf andere Fähigkeiten wird abgelehnt ([Schuler, 2002b](#)). Er wird deshalb als *General Mental Ability* (GMA) oder Allgemeine Mentale Fähigkeit (AMF) bezeichnet ([F. L. Schmidt & Hunter, 1998a, 1998b](#)). Nach [Brocke und Beauducel \(2001\)](#) umfassen weite Definitionen die akademische, praktische, soziale Intelligenz, Lernfähigkeit, Kreativität und komplexes Problemlösen, während enge Definitionen ausschließlich die akademische Intelligenz umfassen. Diese akademische Intelligenz „manifestiert sich vor allem in Fertigkeiten, die in der Schule und in allgemeinen akademischen Bildungseinrichtungen vermittelt und gefördert werden“ ([Brocke & Beauducel, 2001, S. 26-27](#)).

Selbst die Wahl des vergleichsweise engen Fokus auf *Akademische Intelligenz* führt jedoch nicht dazu, dass *eine* allgemeingültige Definition gefunden werden kann, da es sich bei Intelligenz um ein hypothetisches Konstrukt handelt. [Brocke und Beauducel \(2001\)](#) verdeutlichen die Konsequenz dieses Sachverhalts:

Man bezeichnet Konstrukte dieser Art (hypothetische Konstrukte) wegen ihrer prinzipiellen Erweiterungsfähigkeit und permanenten Erweiterung im Zuge der theoretischen Entwicklung auch als *offene* Konstrukte ... Eine einzelne Definition könnte die Breite des Bedeutungsspektrums eines offenen Konstrukts in der Regel nicht erfassen. Für Konstrukte gibt es deshalb auch keinen „Einheitstest“, etwa zur Erfassung *der* Intelligenz, noch ist es sinnvoll, einen solchen Test anzustreben. (S. 13, Hervorhebungen im Original)

Danach gefragt, was Intelligenz bedeutet, antworten z.B. [Eysenck und Eysenck \(1985\)](#) mit drei Definitionen: Intelligenz A sei das biologische Substrat mentaler Fähigkeit, die Neuroanatomie und Physiologie des Gehirns, Intelligenz B die Manifestation von Intelligenz A und all jenes, was ihren Ausdruck im Alltagsverhalten beeinflusst, Intelligenz C schließlich sei das Leistungsniveau in psychometrischen kognitiven Fähigkeitstests.

Die Frage nach dem biologischen Substrat von Intelligenz (A) soll in dieser Arbeit nicht vertieft werden. Die Definition A macht aber deutlich, dass Intelligenz als zeitlich stabiles Fähigkeitskonstrukt (Süß, 2001) bzw. überdauernde Disposition (Brocke & Beauducel, 2001) angesehen werden kann. Von größerer Bedeutung für diese Arbeit sind die Intelligenzdefinitionen B und C, da nach dem Zusammenhang von psychometrischen Testscores (C) auf erfolgreiches Verhalten im realen Berufsleben (also eine Manifestation von Intelligenz, B) gefragt wird.

Wie hat die Entwicklung des *offenen* Konstruktes Intelligenz bisher ausgesehen? Die maßgeblichen Konzepte sollen anhand ihrer historischen Entwicklung dargestellt werden. Ausführlichere Darstellungen der Entwicklung geben z.B. Carroll (1982, 1993), Jensen (1987) sowie J. Funke und Vaterrodt-Plünnecke (1998).

Der Beginn der Beschäftigung mit Intelligenz im engen Sinne, d.h. der Akademischen Intelligenz, wird meist an Galton festgemacht. Er interessierte sich für mentale Fähigkeiten von Individuen, ging davon aus, dass diese in der Bevölkerung normal verteilt sind und mittels einfacher sensorischer und motorischer Aufgaben gemessen werden können (Galton, 1883; zitiert nach Carroll, 1993, S. 31). Binet und Simon vertraten im Gegensatz zu Galton die Ansicht, Intelligenz zeige sich in komplexeren Denkvorgängen wie Urteilsfähigkeit und nicht in grundlegenden Wahrnehmungsleistungen. Sie entwickelten einen Intelligenztest (Binet & Simon, 1905), dessen Aufgabentypen zum Teil bis heute verwendet werden (z.B. in der *Stanford-Binet Intelligence Scale*; Terman & Merrill, 1960), und der deshalb als *erster* Intelligenztest angesehen wird (Holling, Preckel & Vock, 2004). Charakteristika, die man auch in modernen Intelligenztests findet, sind aufsteigende Aufgabenschwierigkeiten, die Vorgabe verbaler, numerischer und räumlicher Aufgaben sowie die Möglichkeit, das mentale Alter einer Person zu bestimmen. Unterschiede zu modernen Verfahren liegen in seiner aus heutiger Sicht unzureichenden Homogenität und der Individualtestung (Carroll, 1993). Die meisten jüngeren Intelligenztests lassen sich auch in Gruppen durchführen. Tests, die sich auch für Gruppen eignen, wurden in Anlehnung an Binet von Otis (1918) zur Testung von amerikanischen Rekruten entwickelt: Army Alpha und – als nonverbaler Test – Army Beta. Eine Weiterentwicklung der Tests von Otis ist der Wonderlic Personnel Test (WPT; Wonderlic & Hovland, 1939), der in einer jüngeren Auflage auch in der vorliegenden Arbeit berücksichtigt wird (vgl. Kapitel 3.2.1).

Eine maßgebliche Weiterentwicklung erfuhr die Intelligenzforschung durch Spearman (1904). Er beobachtete hohe positive Korrelationen (.40 bis .83) zwischen der Leistung von Schülern¹ in verschiedenen Schulfächern, insbesondere nach Korrektur der Messfehler. Er nahm an, dass diese Korrelationen durch die Allgemeine Intelligenz der Schüler zu Stande kommen. Die Allgemeine Intelligenz betrachtete er als Generalfaktor (*g*), die der Leistung in allen Fächern zu Grunde liegt, dabei aber pro Fach von unterschiedlich großer Bedeutung ist. Somit weisen die Anforderungen, die in den einzelnen Fächern gestellt werden, eine unterschiedlich große *g-Sättigung* auf. Die Varianz, die nicht durch *g* erklärt werden konnte, führte Spearman zunächst ausschließlich auf aufgabenspezifische

¹Das in diesem Text verwendete generische Maskulinum bezeichnet gleichermaßen weibliche und männliche Personen.

sche Fähigkeiten zurück (s -Faktoren). Die s -Faktoren selbst korrelierten zum Teil auch positiv miteinander, was Spearman schließlich veranlasste das Wirksamwerden „spezieller Generalfaktoren“ (Holling et al., 2004, S. 18) zu postulieren. Verfahren, die explizit g operationalisieren sind z.B. Matrizen-tests (vgl. Kapitel 3.2.1) und die Intelligenztests nach Wechsler (Holling et al., 2004). Während Binet bemüht war, einen Test für Intelligenz zu entwickeln, der die Vielfalt kognitiver Leistungen berücksichtigt, war Spearmans Ziel die Reduktion dieser Vielfalt auf ihre gemeinsame Essenz (Brody, 2000).

Ebenfalls ein Modell, das im Widerspruch zu Spearmans stand, postulierte Thurstone (Holling et al., 2004). Er vertrat die Ansicht, dass die Leistungen in verschiedenen Bereichen nicht nur durch einen Faktor g und aufgabenspezifische Faktoren erklärt werden können, sondern dass verschiedene „gleichberechtigte“ Faktoren für die Leistungen verantwortlich sind. Diese sogenannten Primär- oder Gruppenfaktoren stehen für grundlegende Fähigkeiten z.B. des verbalen Verständnisses, der Wortflüssigkeit, des schlussfolgernden Denkens, des räumlichen Vorstellungsvermögens, der Merkfähigkeit, der Rechenfähigkeit und der Wahrnehmungsgeschwindigkeit. Das Modell von Thurstone in seiner ursprünglichen Form sieht keinen g -Faktor vor und damit auch keine Möglichkeit, *einen* Wert für die Intelligenz zu bestimmen. Die einzelnen Primärfaktoren sah Thurstone als unkorreliert an. Er verwendete dem entsprechend die Extraktion mit dem Ziel der *simple structure* (Brody, 2000). Später ließ er jedoch Korrelationen zwischen den Primärfaktoren zu. Sie liegen bei etwa .35 (Amelang & Bartussek, 1997) und wurden als Indiz gewertet, dass ein g -Faktor höherer Ordnung extrahiert werden kann. Testverfahren, die sich auf Thurstone berufen und mehrere Primärfaktoren erfassen sind z.B. die älteren Versionen des Intelligenz-Struktur-Tests (Amthauer, 1953a, 1955, 1973) und des Wilde-Intelligenz-Tests (Jäger & Althoff, 1983, 1994).

Obwohl die beiden Konzepte von Spearman und Thurstone ursprünglich unvereinbar nebeneinander standen, ist eine Integration in hierarchische Modelle möglich (Brody, 2000). Dies deutet sich bereits darin an, dass Spearman die spezifischen Generalfaktoren ergänzte, die den Thurstone'schen Primärfaktoren entgegen kommen, und Thurstone die Berechnung von g nicht mehr ablehnte. Als Ursache der unterschiedlichen Intelligenztheorien wurden die verschiedenen faktorenanalytischen Methoden ausgemacht, die letztlich in unterschiedlichen Datensätzen begründet sind (Brody, 2000). Gleichwohl lebt die Diskussion bis heute in der Frage nach der Bedeutung von g bzw. der Bedeutung spezifischer Faktoren fort (vgl. Reeve & Hakel, 2002). Je nach Ansicht werden unterschiedliche faktorenanalytische Modelle gewählt.

Cattell (1957) stellte ein hierarchisches Modell vor, das korrelierte Faktoren I. Ordnung, die Thurstones Primärfaktoren vergleichbar sind, um zwei unabhängige Faktoren II. Ordnung ergänzt, die g ersetzen. Dabei handelt es sich um fluide Intelligenz (g_f) und kristalline Intelligenz (g_c). Während g_f als grundlegende mentale Fähigkeit angesehen wird, kristallisiert sich g_c auf Grund von Lernerfahrungen heraus und zwar um so mehr, je mehr g_f zum Lernen investiert wird (Investmenttheorie). Es wurde vermutet, dass g_f genetisch bedingt ist, was sich aber nicht bestätigen ließ (Horn, 1998, nach W. Johnson & Bouchard, 2005). Auch die Investmenttheorie ließ sich empirisch nicht ausreichend belegen (Holling et al., 2004). Darüber hinaus sind beide Faktoren mit ca. .50 mitein-

ander korreliert. Dieser Befund veranlasste Cattell dazu, einen Faktor III. Ordnung zu postulieren, den er mit $g_{f(h)}$ bezeichnete, um ihn von Spearman's g abzugrenzen und um einen größeren Zusammenhang mit g_f als mit g_c zu verdeutlichen. Faktorenanalytisch lässt sich dieser Faktor nicht nachweisen, weil er allein auf *einer* Korrelation beruht, nämlich der zwischen g_f und g_c (Holling et al., 2004).

Gleichwohl ist die Unterscheidung von g_f und g_c bis heute von Relevanz. Dies liegt auch an der unterschiedlichen Entwicklung beider Faktoren mit zunehmendem Lebensalter. Während g_f mit zunehmendem Lebensalter abnimmt, bleibt g_c konstant (Brody, 2000). Darüber hinaus nimmt die Bedeutung von g_c für die Arbeitsleistung mit zunehmendem Alter zu (Kanfer & Ackerman, 2004). Cattell entwickelte Testverfahren, mit denen g_f möglichst rein erfasst werden sollte, die *Culture Faire Intelligence Tests*. Sie bestehen ausschließlich aus figuralen Items (vgl. Kapitel 3.2.1).

Vernon (1950) stellt wie Cattell ein hierarchisches Modell vor, das durch zwei Faktoren gekennzeichnet ist, die er als $k:m$ und $v:ed$ bezeichnet. Auf $k:m$ laden mechanisch/räumliche Fähigkeiten, auf $v:ed$ verbale und andere kulturell erworbene Fähigkeiten, wie numerisches schlussfolgerndes Denken. Diesen Faktoren übergeordnet ist g . g in Vernons Modell lässt sich nachweisen, da es zuerst extrahiert wird, $k:m$ und $v:ed$ werden dann zur Erklärung der Residualvarianz herangezogen. Dies ist ein Unterschied zu Cattells Modell, in dem ausgehend von einer Korrelation zwischen beiden Faktoren II. Ebene ein Faktor III. Ebene ($g_{f(h)}$) postuliert wird. Gustafsson und Undheim (1996, nach W. Johnson & Bouchard, 2005) konnten zeigen, dass sich beide Modellierungen ineinander überführen lassen.

Häufig werden die Faktoren $v:ed$ und g_c gleichgesetzt: „bisher wurden – zuweilen implizit – verbale Fähigkeiten mehr oder weniger deutlich mit kristallisierter und nonverbale Fähigkeiten mit fluider Intelligenz gleichgesetzt“ (Amthauer, Brocke, Liepmann & Beauducel, 1999, S. 12). Dies ist *inhaltlich* aber unzulässig: der Unterschied zwischen Cattells Faktoren besteht in der Rolle, die erworbenes Wissen für die Fähigkeit spielt: bei g_f ist diese unbedeutend, bei g_c bedeutend. Die Faktoren Vernons sind hingegen inhaltlich verschieden, sind aber beide g_f und g_c gesättigt, nur in unterschiedlichem Maße.

g_c wird z.B. in den neueren Auflagen des Intelligenz-Struktur-Tests (IST 2000, IST 2000 R; Amthauer et al., 1999; Amthauer, Brocke, Liepmann & Beauducel, 2001) und des Wilde-Intelligenz-Tests (WIT-2; Kersting, Althoff & Jäger, 2008) mit Hilfe von Wissensaufgaben erfasst, g_f zeigt sich dann in den klassischen Intelligenzaufgaben nach Kontrolle von g_c .

Die bisher vorgestellten Modelle von Spearman, Thurstone und Cattell haben eine große Zahl weiterer Modelle angeregt, auf die hier nicht näher eingegangen werden soll, da sich die in die Metaanalyse integrierten Intelligenztests nicht ausdrücklich auf diese beziehen. Carroll (1993) listet 21 faktorenanalytische Modellbildungen auf, die in der zweiten Hälfte des 20. Jh. vorgenommen wurden. Diese Modelle unterscheiden sich sowohl in ihrer Struktur als auch in den Inhalten, die sie abdecken (Brocke & Beauducel, 2001) und sie sind beeinflusst von den untersuchten Populationen bzw. Stichproben und den verwendeten Aufgaben (Holling et al., 2004).

Die Modellvielfalt hat dazu geführt, dass übergreifende Rahmenmodelle entwickelt

wurden, mit denen modellübergreifende Aspekte von Intelligenz spezifiziert werden sollen. Zur Entwicklung der Rahmenmodelle können nach Brocke und Beauducel (2001, S. 27-32) fünf verschiedene Vorgehensweisen gewählt werden: a) der Faktoren-Eklektizismus, bei dem die Faktoren aus den verschiedenen Modellen gesammelt werden ohne eine Dimensionsreduktion anzustreben, b) die Content-Definitionen, bei der zentrale Merkmalsbereiche identifiziert werden, c) die metaanalytische Integration und d) die Aufgabenintegration. Die bisher genannten Verfahren gehen *empirisch* vor. Es sind aber auch e) primär *theoriegeleitete* Rahmenmodelle entwickelt worden. Einige bedeutsame Rahmenmodelle sollen vorgestellt werden, weil sie zum einen das Konstrukt *Intelligenz* definieren helfen; zum anderen soll der Metaanalyse ein Rahmenmodell zu Grunde gelegt werden.

Die *faktoren-eklektizistische Vorgehensweise* wurde z.B. zur Entwicklung des *Kit of Reference Tests for Cognitive Factors* von French, Ekstrom und Price (1963) angewandt. Dazu bündelten sie die „Markiertests der empirisch bestgesicherten Intelligenzfaktoren“ (Jäger, Süß & Beauducel, 1997, S. 11) zu einer Testbatterie. Der Vorteil dieses Verfahrens ist, dass die Zusammenstellung nicht von Modell- oder Methodenentscheidungen abhängt und deshalb sehr transparent ist. Das Ergebnis ist dafür abhängig von den zu einem bestimmten Zeitpunkt vorliegenden Modellen und es führt zu einer Ungleichgewichtung einzelner Fähigkeitsbereiche (Brocke & Beauducel, 2001). So konnten z.B. Süß, Oberauer und Jäger (1988) zeigen, dass 63% aller Aufgaben aus den French-Tests Verarbeitungskapazität mit figuralem Material erfassen oder verbalem Einfallsreichtum zugerechnet werden können, während andere Operationen und numerisches Aufgabenmaterial unterrepräsentiert sind.

Im Gegensatz zu den Modellen von Cattell und Vernon stellt Guilford (1967) ein Modell vor, das nicht hierarchisch aufgebaut ist, sondern danach fragt, welche Inhalte ein Intelligenzkonstrukt abdecken kann (*Content-Definition*). Von einer deskriptiven Analyse bestehender Intelligenztest-Aufgaben ausgehend, beschreibt Guilford drei Dimensionen: die Operationsdimension (Wissen, Gedächtnis, divergente Produktion, convergente Produktion und Evaluation), die Inhaltsdimension (figural, symbolisch, semantisch und Verhalten) und die Ergebnisdimension (Einheiten, Klassen, Relationen, Systeme, Transformationen und Implikationen). Diese drei Dimensionen ordnet er zu einem Quadermodell an. Jede Operations-Inhalts-Ergebnis-Kombination steht für eine mentale Fähigkeit. Insgesamt unterscheidet er 120, durch die später erfolgte Aufteilung der figuralen Inhalte in auditorische und visuelle sogar 150 (Guilford, 1977). Durch ein spezielles Rotationsverfahren (*targeted rotation*) versuchte Guilford diese Faktorenstruktur empirisch zu belegen, was nicht zufriedenstellend gelang (Horn & Knapp, 1973; nach Brody, 2000). Schließlich extrahierte Guilford (1981) Faktoren höherer Ordnung und überführte sein Modell in ein hierarchisches, das über den 150 Faktoren I. Ordnung 85 II. Ordnung und 16 III. Ordnung aufwies. Teilaspekte des Modells, die empirisch überprüft wurden, wiesen keinen zufriedenstellenden Fit auf (Brody, 2000). Das Modell hat jedoch heuristischen Wert (Jäger et al., 1997), indem es umfassend die gesamte Landschaft kognitiver Fähigkeiten abzudecken versucht, inklusive kreativer Leistungen.

Ein Modell, das ebenfalls der Ordnung des Inhaltsraumes mentaler Fähigkeiten dient,

stammt von Guttman (1965) sowie Guttman und Levy (1991). Sie unterscheiden vier Dimensionen. Die Dimensionen *Material* (verbal, numerisch, figural) und *Aufgabe* (Regelinferenz, Regelanwendung, Lernen) werden kombiniert als *Radexmodell* bezeichnet. Dieses kann um die Dimension *Expression* (schriftliche, manuelle oder mündliche Bearbeitung) zu einem *Zylindrexmodell* erweitert werden. Die vierte Dimension betrifft die *Lösung* der Aufgaben (richtig oder falsch).

Die Modelle von Guilford und Guttman haben den Vorteil aufzuzeigen, welche Fähigkeitsbereiche Intelligenz zugeordnet werden können und welche davon wie gut durch existierende Testbatterien abgedeckt werden. Weil nicht alle Fähigkeitsbereiche abgedeckt werden können, ist es nicht möglich, sie empirisch abzusichern (Brocke & Beauducel, 2001).

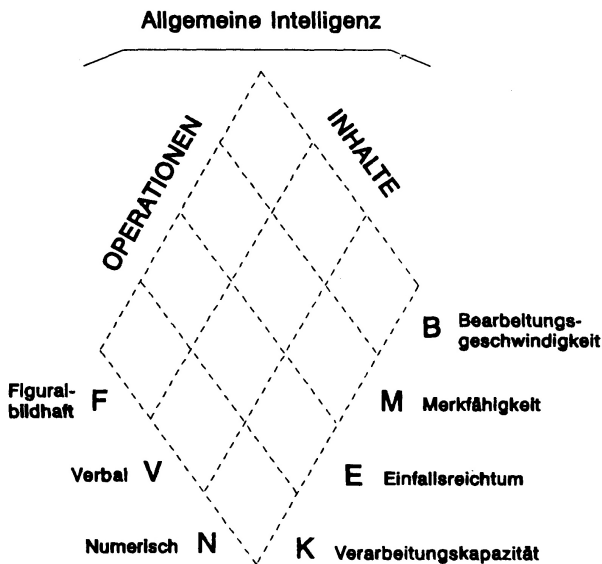
Carroll (1993) hat eine Vielzahl verfügbarer Datensätze zu den verschiedenen Intelligenzmodellen einer Metaanalyse unterzogen. Seine Lösung enthält drei Faktor-Hierarchiestufen, die er als *Strata* bezeichnet. Im dritten, dem obersten Stratum findet sich als einziger Faktor g , im zweiten Stratum sind es folgende acht Faktoren, die in der Rangfolge ihrer g -Sättigung genannt werden: g_f lädt am stärksten auf g , gefolgt von g_c , Gedächtnis und Lernen, visueller Wahrnehmung, auditiver Wahrnehmung, Abruffähigkeit (enthält Aspekte von Kreativität), kognitiver Geschwindigkeit (broad cognitive speediness) und Verarbeitungsgeschwindigkeit (processing speed, i.S.v. Reaktionszeiten). Auf dem ersten Stratum unterscheidet Carroll 80 verschiedene Faktoren (Brody, 2000; Brocke & Beauducel, 2001). Um zu dieser hierarchischen Struktur zu gelangen, hat Carroll für die einzelnen Datensätze unterschiedliche Verfahren (z.B. unterschiedliche Rotationstechniken) angewandt. Dabei musste er „subjektive, von inhaltlichen Erwägungen geleitete Entscheidungen“ (Brocke & Beauducel, 2001, S. 30) treffen, was die Verbindlichkeit seiner Lösung einschränkt.

Eine *Integration auf Aufgabenebene* wurde von Jäger (1982) angestrebt. Er geht von den Grundannahmen aus, dass a) alle mentalen Fähigkeiten an Intelligenzleistungen beteiligt sind, b) sich die Aufgaben und die zu ihrer Lösung erforderlichen Fähigkeiten verschiedenen Modalitäten zuordnen lassen und c) die mentalen Fähigkeiten in eine Hierarchie gebracht werden können, deren Strata sich durch unterschiedliche Generalität auszeichnen.

Zunächst sichtete Jäger 2000 Aufgabengruppen verschiedener Intelligenz- und Kreativitätstests. Aus ökonomischen Gründen wurden diese auf 191 Aufgabengruppen reduziert, die 98 verschiedenen Aufgabentypen zugeordnet werden konnten. Bei dieser Reduktion wurde auf die „(1) Erhaltung der Vielfalt der Aufgabenstellungen und (2) Beibehaltung von Markiervariablen für die Hauptkomponenten der konkurrierenden Strukturmodelle“ (Jäger et al., 1997, S. 8-9) geachtet. Die Aufgabengruppen wurden in einem nächsten Schritt faktoren- und clusteranalytisch gruppiert. Dabei konnten vier Operationen unterschieden werden: *Bearbeitungsgeschwindigkeit* (B), *Gedächtnis* (G), *Einfallsreichtum* (E) und *Verarbeitungskapazität* (K). *Gedächtnis* umfasste kurz- und langfristige Behaltensleistungen. Später wurden die langfristigen exkludiert und die verbleibenden kurzfristigen zur Facette *Merkfähigkeit* (M) zusammengefasst.

Die Aufteilung in inhaltlich spezifische Faktoren wurde auf Grund ihrer Bedeutung

Abbildung 2.1
Das Berliner Intelligenzstruktur-Modell



Quelle: Jäger et al. (1997, S. 5)

in mehreren Intelligenzstruktur-Konzepten ergänzt. Die Inhalte sind *sprachgebundenes Denken* (V), *zahlengebundenes Denken* (N) und *anschauungsgebundenes Denken* (F). Dadurch ergab sich ein Facettenmodell der Intelligenz, das Berliner Intelligenzstruktur-Modell (BIS-Modell, Abbildung 2.1), das mit den beiden Dimensionen *Operationen* und *Inhalte* dem Radex-Modell von Guttman ähnelt. Ebenfalls entwickelt wurde eine Testbatterie, die alle zwölf Operations-Inhalts-Strukturpaare abdeckt und zur Erfassung der vier operationshomogenen Faktoren (B, M, E, K), der drei inhalts-homogenen Faktoren (V, N, F) und dem übergeordneten Faktor *g* genutzt werden kann. Die Faktoren sind in Anhang B.2 näher beschrieben.

Die Allgemeine Intelligenz bzw. *g* wird im BIS-Modell als Integral der kognitiven Fähigkeiten angesehen, die vom BIS-Modell abgedeckt werden (Jäger et al., 1997). Es handelt sich um ein deskriptives Modell das „keine Annahmen über funktionale oder kausale Beziehungen zwischen seinen Komponenten“ (Jäger et al., 1997, S. 5) macht. Lediglich die Annahme eines moderat positiven Zusammenhangs zwischen den sieben Faktoren I. Ordnung ist Voraussetzung für das Wirken *eines* Faktors *g* II. Ordnung.

Zur Konkretisierung der genannten Definition *Allgemeiner Intelligenz* ist es erforderlich, die Fähigkeiten zu betrachten, die vom BIS-Modell abgedeckt bzw. nicht mehr abgedeckt werden: Die Aufgaben der Operation B sind einfach strukturiert. Ihre Lösung erfordert Schnelligkeit und Konzentration. Die Aufgabentypen aus der Galton-Cattell-Tradition (Lohman, 2000) lassen sich hier zuordnen. Die Aufgaben der Operation K sind komplexer. Um sie zu lösen ist schlussfolgerndes „formallogisch exaktes Denken“ (Jäger et al., 1997, S. 6) erforderlich. Dies entspricht Aufgabentypen in der Art, die Binet in seinem Test einsetzt. Die Operation K wird für besonders wichtig erachtet, da sie in einem „nahe-

zu perfekten“ (S. 144) Zusammenhang mit der Kapazität zum simultanen Speichern und Verarbeiten sowie Koordinieren von Informationen steht (Arbeitsgedächtnis-Kapazität, Süß, 2004). Die Operation M umfasst Aufgaben, die das Speichern und Reproduzieren von Informationen verlangen. Das Reproduzieren erfolgt *kurz* nach dem Abspeichern der Information. Langfristige Behaltensleistungen, z.B. die Reproduktion gespeicherter Informationen nach einer Unterbrechung, umfasst das Modell nicht. Schließlich umfasst das BIS-Modell die Operation E, die eine „zentrale Komponente“ (Jäger et al., 1997, S. 7) von Kreativität darstellt. Weitere Kreativitätsaspekte, z.B. Originalität und Ideenflüssigkeit werden nicht berücksichtigt.

Die Vorteile dieses Rahmenmodells – im Vergleich zu den anderen genannten – sind folgende: Im Gegensatz zum Faktoren-Eklektizismus findet eine Dimensionsreduktion statt, eine *Intelligenzstruktur* wird postuliert. Im Gegensatz zu den Modellen von Guilford und Guttman, die auf Content-Definitionen beruhen, liegen für alle BIS-Modell-Struktur Aufgaben vor. Das BIS-Modell konnte sich deshalb empirisch bewähren (Brocke & Beauducel, 2001).

Aber auch die Verbindlichkeit des BIS-Modells gilt nicht uneingeschränkt. Jäger et al. (1997) verweisen darauf, dass es sowohl für inhaltliche als auch strukturelle Erweiterungen offen ist. Brocke und Beauducel (2001) kritisieren die wenig transparente Aufgabenselektion im ersten Reduktionsschritt und machen darauf aufmerksam, dass die Modellstruktur zumindest zum Teil *deduktiv* von vordefinierten Modellannahmen ausgehend entwickelt wurde und nicht rein induktiv aus dem Datenmaterial. Dies betrifft die Modellierung der Inhaltsfacetten und die Inklusion von Kreativität. Diese Einschränkungen führen dazu, dass bestimmte Bereiche, die in anderen Modellen *Intelligenz* zugerechnet werden, nicht integrierbar sind. Dies sind z.B. langfristige Behaltensleistungen und Wissenstests.

Primär *theoriegeleitete Intelligenzmodelle* sind z.B. das Pyramidenmodell von Horn (1983), das *Evolutionsmodell der Intelligenzstrukturforschung* von Sternberg und Powell (1982) und das *Hierarchische Protomodell der Intelligenzstrukturforschung* (HPI) von Amthauer et al. (1999, 2001). Das Pyramidenmodell wird von Horn genutzt, um das von ihm entwickelte Leistungsprüfsystem (vgl. Kapitel 3.2.1) theoretisch einzuordnen. Im Pyramidenmodell unterscheidet Horn an der Basis sechs Operations- und sechs Inhaltsfaktoren, die zu 36 Fähigkeiten kombiniert werden können. Jede Fähigkeit kann theoretisch mit sechs Aufgabentypen erfasst werden. Vertikal sind sechs Hierarchieebenen angeordnet, mit zunehmend komplexen Faktoren. Auf der fünften Faktorebene bilden sie einen Faktor *Generelle Intelligenz g*, über dem auf der höchsten sechsten Ebene ein Faktor *Generelle Tüchtigkeit* postuliert wird. Das Evolutionsmodell von Sternberg beschäftigt sich mit der Struktur von Intelligenz (Struktur-Frage) und nicht mit den inhaltlichen Aspekten (Kontent-Frage; Brocke & Beauducel, 2001). Es ist deshalb für die vorliegende Arbeit weniger relevant und soll nicht näher ausgeführt werden.

Konvergenzen in Bezug auf Kontent *und* Struktur der Intelligenzmodelle sind im HPI zusammengefasst. Übereinstimmende Inhalte der meisten Intelligenzmodelle sind demnach die „Fähigkeit zum schlußfolgernden Denken (Reasoning), verbale, numerische und räumlich-figurale Fähigkeiten, Ideenflüssigkeit bzw. Kreativität und, mit Einschränkun-

gen, ... die Merkfähigkeit und wahrnehmungsbezogene Fähigkeiten“ (Amthauer et al., 2001, S. 9). Diese Inhalte sind auch Teil des BIS-Modells, wobei in Bezug auf Intelligenz nicht der Aspekt der Ideenflüssigkeit, sondern der des Einfallsreichtums thematisiert wird. Auf struktureller Ebene sehen Amthauer et al. Konvergenzen in der hierarchischen Anordnung der Fähigkeitskonstrukte auf mindestens zwei Ebenen und verweisen in diesem Zusammenhang beispielhaft auf das BIS-Modell.

Das BIS-Modell erscheint auf Grund der genannten Vorteile gegenüber den anderen empirisch gewonnenen Rahmenmodellen als das geeignetste für die Literaturrecherche und Strukturierung der Metaanalyse. Ferner wird direkt auf das BIS-Modell Bezug genommen in Intelligenztest-Manualen (z.B. Amthauer et al., 2001; Jäger & Althoff, 1983; Schuler & Klingner, 2004) und Studien zur Intelligenz (z.B. Knebelau, 2002). Dies gilt im Besonderen für den deutschsprachigen Kulturraum, der hier im Fokus des Interesses steht. Aber auch international ist das BIS-Modell anerkannt (vgl. Carroll, 1993).

Welche Konsequenzen ergeben sich aus der Wahl des BIS-Modells als Rahmenmodell für die Definition von Intelligenz? Jäger et al. (1997) verstehen unter „g‘ ein integriertes System separierbarer kognitiver Fähigkeiten ..., am besten erfaßbar durch eine möglichst repräsentative Stichprobe des ‚Universums‘ kognitiver Prozesse“ (S. 16). Diese Definition ist insofern weit gefasst, da sie einzig den gemeinsamen Nenner der verschiedenen Intelligenzdefinitionen benennt, nämlich, dass es sich um *kognitive Fähigkeiten* handelt. Unmittelbar ergänzt wird ein Hinweis auf die Messung des Konstruktes. Dadurch wird noch einmal deutlich, dass über die Messung eher ein Konsens herzustellen ist als über die Definition von Intelligenz (Goldstein, Zedeck & Goldstein, 2002).

Entgegen der Definition von Jäger et al. (1997) sollen auch Operationalisierungen von Intelligenz berücksichtigt werden, die nicht das Universum kognitiver Prozesse abzubilden versuchen, sondern sich auf einzelne Facetten beschränken, wenn diese a) Teil des BIS-Modells sind und b) die Allgemeine Intelligenz und nicht spezifische Fähigkeiten operationalisieren. Dies ist z.B. bei Matrizen tests wie den Culture Faire Intelligence Tests der Fall (vgl. Kapitel 3.2.1), der g_f nach Cattell operationalisiert, die als wesentliche Komponente von Intelligenz angesehen werden kann (Beauducel & Kersting, 2002). Eine entsprechend allgemeine Definition gibt am besten wieder, was unter Intelligenz verstanden werden soll:

The ability most studied is general cognitive ability, which measures the ability to learn (Hunter, 1996). Psychologists often call this ability *intelligence*, although psychologists use this word somewhat differently from nonpsychologists. Psychologists do not mean genetic potential, but rather the level of ability developed at the time that an ability test is taken. Because studies show that there is usually little change in ability during the adult years, industrial psychologists use the word *intelligence* for the level of ability developed when a person reaches his or her adult years. (Hunter & Schmidt, 1996, S. 453-454)

Um das deutlich zu machen, wird häufig statt Intelligenz der Begriff *general mental ability* (GMA, bzw. *Allgemeine Mentale Fähigkeiten*, AMF) verwendet, die hier synonym mit Intelligenz benutzt werden. Im Unterschied dazu wird die Bezeichnung g für

faktorenanalytisch gewonnene Intelligenzwerte genutzt (*g*-Faktor; vgl. Ree & Carretta, 2002).

2.2. Beruflicher Erfolg

Wie Intelligenz wird auch Berufserfolg wissenschaftlich unterschiedlich definiert und gemessen: „Es gibt in der Literatur keine allgemein akzeptierte Vorgehensweise“ (Abele-Brehm & Stief, 2004, S. 7). Berufserfolg ist ein hypothetisches Konstrukt, das nur indirekt erfasst werden kann (Marcus & Schuler, 2006). Eine Operationalisierung ist erforderlich. Definition und Operationalisierung hängen unmittelbar zusammen: „In keinem Fall ergibt sich die Beurteilungseinheit von selbst, immer steht eine ‚Theorie‘ des Beurteilers dahinter, welche Beobachtungseinheiten als Indikatoren des Erfolgs aufgefasst werden können“ (Schuler, 2004b, S. 5). Dette et al. (2004) verweisen darauf, dass es gängige Praxis sei, Erfolg anhand der in der jeweiligen Studie verwendeten Operationalisierungen zu definieren. Sie argumentieren weiter, dass es problematisch sei, die Fragen danach, was Erfolg sei und wie er gemessen werde, zu vermengen. Ziel dieses Kapitels ist es deshalb zu spezifizieren, welche Bereiche *Berufserfolg* umfasst, und Indikatoren zu benennen, mit denen sie gemessen werden können.

2.2.1. Bereiche von Berufserfolg

Eine weite Fassung des Begriffs schließt den Bereich des *Laufbahnerfolgs* mit ein. Dazu gehören unstrittig der berufliche Aufstieg in Form einer linearen Karriere, häufig gemessen z.B. an der Gehaltsentwicklung oder der hierarchischen Position. Dette et al. (2004) weisen jedoch darauf hin, dass der Begriff *Laufbahn* „im Gegensatz zu Karriere auch Berufswege bezeichnet, die keinen kontinuierlichen beruflichen Aufstieg beinhalten und somit dem englischen Begriff ‚career‘ besser entspricht“ (S. 171). Laufbahnerfolg kann somit auch andere Konzepte umfassen (etwa eine Einmalentscheidung, z.B. für einen Beruf oder spiralförmige Entwicklungen) und sich an anderen Werten als dem beruflichen Vorankommen ausrichten, z.B. an Sicherheit, Freiheit, Exzellenz, *work-life-balance* (Moser, 2004a).

Unterschiedliche Auffassungen gibt es dazu, ob sich Laufbahnerfolg auch auf den Bereich subjektiver Leistungsfaktoren wie Gesundheit, Wohlbefinden, organisationalem Selbstwert und dergleichen erstreckt. Die Bedeutung dieser Faktoren wird hauptsächlich unter dem Label der *Laufbahn-Zufriedenheit* diskutiert. Bruggemann, Großkurth und Ulrich (1975) argumentieren, Zufriedenheit beziehe sich nicht auf Leistung oder Erfolg, sondern auf das Arbeitsumfeld. Dette et al. (2004) entgegnen jedoch, dass berufliche Zufriedenheit Teil von Laufbahnerfolg ist, da sie ein berufliches Ziel vieler Personen darstellt und hoch mit Erfolgsmaßen korreliert. Ng, Eby, Sorensen und Feldman (2005) beziffern den Zusammenhang von Zufriedenheit mit Einkommen z.B. auf $\hat{\rho} = .30$ und vertreten wie Judge, Higgins, Thoresen und Barrick (1999) den Standpunkt, Zufriedenheit ergänze als intrinsisches Maß die unstrittigen extrinsischen Laufbahn-Erfolgsmaße.

Enger gefasst wird Berufserfolg dagegen, wenn er auf berufliche *Leistung* beschränkt wird. Berufliche Leistung lässt sich nach Marcus und Schuler (2006) als „Beitrag zu den

Zielen einer Organisation definieren“ (S. 434). Die berufliche Leistung der Beschäftigten sichert das Überleben der Organisation. Dieser Bereich soll [Dette et al. \(2004\)](#) folgend als Bereich der *Arbeit* bzw. Arbeitsleistung (*job performance*) bezeichnet werden.

Dabei können wiederum verschiedene Leistungsbereiche berücksichtigt werden. Viswesvaran, Ones und Schmidt (1996) unterscheiden zehn Dimensionen, die Leistungsbeurteilungen zugrunde liegen: generelle Arbeitsleistung (*overall job performance*), Produktivität (*job performance or productivity*, bezieht sich auf die Quantität der Arbeitsleistung), Qualität (*quality*), Führungsverhalten (*leadership*), Kommunikationskompetenz (*communication competence*), administrative Kompetenz (*administrative competence*), der betriebene Arbeitsaufwand (*effort*), interpersonale Kompetenz (*interpersonal competence*), Fachwissen (*job knowledge*) und die Einhaltung bzw. Akzeptanz von Regeln und Vorschriften (*compliance with or acceptance of authority*).

Eine andere Unterscheidung nimmt [Schuler \(2004b\)](#) vor, indem er die Leistungsmaße verschiedenen Beschreibungsebenen zuordnet: Auf der *Ergebnisebene* finden sich Kriterien, die die Arbeitsergebnisse quantifizieren, auf der *Verhaltensebene* wird das Verhalten der Beschäftigten im Vergleich zu einem „Idealverhalten“ betrachtet und auf der *Eigenschaftsebene* das Potential von Personen eingeschätzt. Die letzte Einschätzung ist eher auf die Zukunft gerichtet, während Ergebniskriterien vergangenheitsorientiert sind ([Marcus & Schuler, 2006](#)). Ferner können nach Schuler die Anzahl durchgeführter Tätigkeiten, der Grad der Zielerreichung und der Kenntnisstand einer Person als Leistungsmaße betrachtet werden.

[N. Schmitt, Cortina, Ingerick und Wiechmann \(2003\)](#) unterscheiden a) die aufgabenbezogene Leistung (*task performance*). Sie umfasst die Qualität, Quantität und Vollständigkeit mit der die Kernaufgaben erledigt werden. Also die Aufgaben, die formal zum Arbeitsinhalt gehören, für die man „bezahlt wird“. b) Als kontextuelle Leistung (*contextual performance*) bezeichnen sie die Aspekte der Leistung, die dem Arbeitsumfeld zugute kommen: Kooperativität, Zuverlässigkeit, die Einhaltung von Regeln und Vorschriften sowie die Unterstützung und Motivierung anderer (vgl. [Scotter & Motowidlo, 1996](#)). Dieser Aspekt von Leistung ist deshalb von Bedeutung, weil Arbeitsleistung „nur durch das organisierte Zusammenwirken mehrerer Personen, d.h. in einem arbeitsteiligen Prozess effizient erbracht werden [kann]“ ([K.-H. Schmidt & Kleinbeck, 2004](#), S. 891). Schließlich deckt c) adaptive Leistung (*adaptive performance*) die Anpassung an Veränderungen, Krisenbewältigung und Innovationen ab (vgl. [Pulakos, Arad, Donovan & Plamondon, 2000](#)). [Hunter und Schmidt \(1996\)](#) weisen darauf hin, dass die Bereiche nicht unabhängig voneinander beurteilt werden:

Most people confuse two different dimensions of personnel evaluation: (a) performance (i.e., productivity) and (b) citizenship behavior (i.e., social behavior at work) ... When supervisors evaluate workers, they rely heavily on factors other than the worker's productivity. In particular, they give substantial weight to the social behavior of the worker ([Orr, Sackett & Mercer, 1989](#)). However, this is usually not a conscious and articulated process; rather, it takes place implicitly. (S. 447)

Ein weiterer Bereich beruflicher Leistung ist die berufliche *Lernleistung*. Lernen umfasst im Allgemeinen „jede relativ *überdauernde Änderung* der *Verhaltensmöglichkeiten* eines Menschen, die nicht auf altersbedingte Veränderungen zurückgeführt werden kann“ (Bredenkamp, 1998, S. 22) oder „vermittelte, informationsverarbeitende Vorgänge, die *verhaltensregulierenden Gedächtnisbesitz* aufbauen, welcher eine effektivere Auseinandersetzung mit Anforderungen ermöglicht“ (Hacker & Skell, 1993, S. 18). Lernen kann absichtsvoll oder nicht absichtsvoll und mehr oder weniger bewusst stattfinden; der Lernprozess kann in eigens dafür initiierten Maßnahmen gefördert werden oder mehr oder weniger beiläufig beim Problemlösen, im Laufe der individuellen Entwicklung, der Ansammlung von Erfahrung, der Sozialisation (Greif & Kluge, 2004), der „Zusammenarbeit und Kommunikation mit anderen Menschen“ (Hacker & Skell, 1993, S. 18). Eine wichtige Rolle kommt der Selbstorganisation und -regulation zu, auch bei fremd bestimmtem Lernen (z.B. beim Frontalunterricht; Greif & Kluge, 2004).

Eine engere Definition wählen Hacker und Skell (1993) für individuelles Lernen in Organisationen, indem sie unbeabsichtigte Lernprozesse ausgrenzen: „Im Berufsleben herrscht ein Lernen vor, das mit der Absicht der Aneignung von möglichst optimalen Grundlagen für Leistungen und Leistungsverbesserungen einhergeht, das heißt ein bewußtes und zielgerichtetes Lernen“ (S. 17). Lernen wird als Mittel zur Leistungserbringung bzw. -steigerung angesehen.

Mögliche berufliche Lerninhalte sind nach Greif und Kluge (2004) a) fachliche und überfachliche Qualifikationen, also Fähigkeiten, Fertigkeiten, Kenntnisse, Motivation, Einstellungen und Interessen, b) Aspekte der Sozialisation bzw. Enkulturation wie Werte, Normen, Regeln, Traditionen, Überzeugungen, kommunizierte Erfahrungen sowie c) die Bewältigung von persönlichen und beruflichen Entwicklungsaufgaben. Erstere werden häufig in Form von verschiedenen Bildungsmaßnahmen vermittelt, b) und c) im Rahmen von Sozialisations- und beruflichen Entwicklungsprogrammen. Die Lernprogramme werden nach Greif und Kluge häufig als Teil umfassenderer Maßnahmen der Personal- bzw. Organisationsentwicklung oder des Wissens- bzw. Erfahrungsmanagements durchgeführt.

In Organisationen wird *Lernen* häufig unter den Begriff der *Bildung* gefasst. Mit *Bildung* kann sowohl der *Prozess* als auch das *Ergebnis* mehrjährigen Lernens bezeichnet werden; er umfasst schulische und nachfolgende außerschulische Lernaktivitäten (Greif & Kluge, 2004). Unterschieden werden Ausbildung, Fortbildung und Weiterbildung. „In der Ausbildung erwerben Jugendliche die berufliche Erstqualifikation für ministeriell anerkannte Ausbildungsberufe“ (Rüttinger & Klein-Moddenborg, 1989, S. 687). Fortbildungen und Weiterbildungen setzen nach Rüttinger und Klein-Moddenborg dagegen eine berufliche Vorbildung voraus. Nach §1 (4) des Berufsbildungsgesetzes (BBiG, 2005) soll „die berufliche Fortbildung . . . es ermöglichen, die berufliche Handlungsfähigkeit zu erhalten und anzupassen oder zu erweitern und beruflich aufzusteigen“. Einer engeren Definition folgend umfassen Fortbildungen nur den Erhalt und die Anpassung der Handlungsfähigkeit, während Bildungsmaßnahmen, die der Erweiterung der Fähigkeiten und dem beruflichen Aufstieg dienen, als Weiterbildungen bezeichnet werden (z.B. Greif & Kluge, 2004). Pawlowsky und Bäumer (1996) spezifizieren demgemäß eine Anpassungs-

Weiterbildung und eine Aufstiegs-Weiterbildung.

Die Berücksichtigung beruflicher Lernleistung als Indikator für beruflichen Erfolg ist nicht unumstritten. Sie erscheint manchen Kritikern ein zu akademisches Kriterium zu sein, die hohe Korrelation komme auf Grund gemeinsamer Methodenvarianz zustande (behavioristische Kritik; vgl. Hunter, 1986). Andererseits nimmt die Bedeutung von Fachwissen für die Arbeitsleistung immer mehr zu und damit die Bedeutung der beruflichen Lernleistung (vgl. Campion in Morgeson et al., 2007). Als Indiz dafür kann auch die Zunahme von Bildungsmaßnahmen gewertet werden, die zusätzlich zur beruflichen Erstausbildung besucht werden (Greif & Kluge, 2004). Darüber hinaus kann als gesichert gelten, dass berufliche Lernleistung mit Arbeitsleistung positiv korreliert (mit einem n -gewichteten \bar{r} von .16 laut Samson, Graue, Weinstein & Wallberg, 1984).

Lernleistung unterscheidet sich nicht qualitativ von Arbeitsleistung, sondern lediglich quantitativ (Kuncel et al., 2004): a) auch Arbeitsleistung umfasst das Erlernen neuen Wissens, b) beide Leistungsbereiche erfordern den Umgang mit (mehr oder weniger) komplexen oder praktischen Aufgaben und c) die Leistung in beiden Bereichen hängt zumindest teilweise von dem Grad bereits erworbenen Wissens und bereits erworbener Fähigkeiten ab. Die Annahme, akademische Aufgaben seien immer eindeutig definiert, hätten eine einzige korrekte Antwort und seien in sich abgeschlossen, weisen Kuncel et al. zurück. Jäger (1986) weist in Bezug auf Ausbildungserfolg darauf hin: „seine Abwertung als ‚Ersatz‘-Kriterium [für Arbeitsleistung, Anm. des Verfassers] ist weder theoretisch noch praktisch zu rechtfertigen. ‚Endgültige‘ Kriterien sind eine Utopie“ (S. 279).

Nicht im Fokus des Interesses dieser Arbeit stehen *Studienleistungen*, die auch als berufliche Lernleistungen anzusehen sind. Während eines Studiums an einer Fachhochschule oder Universität steht die Theorievermittlung im Vordergrund, betriebliche Praxiserfahrungen werden i.d.R. in Form von relativ kurzen Praktika gesammelt. Die Leistung in der betrieblichen Praxis geht i.d.R. *nicht* in die Abschlussnote mit ein. Die klassische Ausbildung findet in Deutschland *dual* statt, d.h. in Form einer betrieblichen *und* schulischen Ausbildung. Auszubildende sind bei einem Unternehmen angestellt, von dem sie bezahlt werden. Dies ist bei Studierenden nicht der Fall. Einen Überblick über die prädiktive Validität von Intelligenz in Bezug auf Studienleistungen gibt z.B. Süß (2001).

Schließlich gibt es Mischformen, z.B. duale Studiengänge an einer Berufsakademie oder die Ausbildung zum Beamten im gehobenen Dienst in der öffentlichen Verwaltung des Bundes und der Länder, die in Form eines Studiums an Fachhochschulen für öffentliche Verwaltung stattfindet (Bundesagentur für Arbeit, 2008a). Dieses Studium qualifiziert die Absolventen spezifisch zur Ausübung einer Tätigkeit beim Arbeitgeber Bund oder den Ländern, bei denen die Studierenden auch angestellt sind und von denen sie ein Gehalt beziehen. Dieses Studium hat deshalb mehr mit einer Ausbildung gemein als mit einem klassischen Studium an einer Universität (M. Kersting, persönliche Mitteilung, 23.5.2005). Sie wurde deshalb in der Metaanalyse berücksichtigt.

Zusammenfassend werden unter *Berufserfolg* alle Aspekte gefasst, die eine Wertung darüber erlauben, wie erfolgreich Personen bei der Ausübung beruflicher Tätigkeiten sind. Dies schließt die berufliche Entwicklung von der Erstausbildung bis zum Ende der Berufslaufbahn mit ein. Berufserfolg umfasst *Laufbahnerfolg*, also die Erfolgsaspekte,

die besonders aus der Perspektive des Beurteilten relevant sind (z.B. Einkommen, erreichte Position), ebenso wie *Arbeitsleistung*, die Erfolgsaspekte umfasst, die zunächst besonders aus der Perspektive der Organisation von Bedeutung sind (z.B. subjektive Leistungsbeurteilungen durch Vorgesetzte). Berufsbezogene *Lernleistung* meint die Leistungsaspekte, die den Erwerb neuer beruflicher Kenntnisse und Fertigkeiten in Aus- und Weiterbildung betreffen (Berufsschulnoten, betriebliche Abwärtsbeurteilungen).

Im Gegensatz zur Definition von Rüttinger und Klein-Moddenborg (1989) wird Ausbildung nicht auf Jugendliche und nicht auf ministeriell anerkannte Ausbildungsberufe begrenzt, sondern schließt Erwachsene ebenso mit ein, wie Anlern Tätigkeiten und Umschulungen, wenn sie ohne berufliche Vorbildung erworben werden können. *Weiterbildungen* dagegen setzen eine berufliche Vorbildung voraus und bauen auf dieser auf. Sie führen zu einem Abschluss, der als Spezialisierung der beruflichen Erstausbildung angesehen werden kann. Nicht mit betrachtet wird Studienerfolg.

2.2.2. Indikatoren für beruflichen Erfolg

Die Messung der drei Bereiche beruflichen Erfolgs *Laufbahn*, *Arbeit* und *Lernen* ist nicht anhand eines einfachen, ultimativen Kriteriums möglich, da Leistungskriterien immer defizient und kontaminiert sind (Marcus & Schuler, 2006). Dieses Kriterienproblem ist „ein Kernproblem eignungsdiagnostischer Untersuchungen“ (Brambring, 1983, S. 423).

Dette et al. (2004) systematisieren die möglichen Berufserfolgs-Indikatoren wie folgt: Sie unterscheiden die drei Datenarten *neutrale Kennzahlen*, *Vergleiche mit Bezugsstandards* und *Maße der Zufriedenheit*, die verschiedenen Datenquellen entstammen können. Neutrale Kennzahlen können Dokumenten entnommen oder durch Selbstberichte erfasst werden, Vergleiche mit Bezugsstandards von den zu Beurteilenden selbst oder externen Beobachtern (Fremdbeurteilern) stammen. Als Quelle für Zufriedenheitseinschätzungen kommen wiederum Selbstberichte in Frage.

Andere Unterscheidungen sind objektive vs. subjektive bzw. externe vs. interne Maße (Dette et al., 2004). Interne Daten sind solche, die in Form von Selbstberichten gewonnen werden, externe stammen aus anderen Quellen. Diese Unterscheidung legen auch van Maanen und Schein (1977) sowie Gattiker und Larwood (1986) ihrer Definition von subjektiven (internen) und objektiven (externen) Maßen zugrunde. Die Unterscheidung kann jedoch auch testtheoretisch verstanden werden: subjektive Maße umfassen dann alle Beurteilungen, sowohl (externe) Fremd-, als auch (interne) Selbstbeurteilungen. Objektive Maße bezeichnen dementsprechend die direkt beobachtbaren, zählbaren Größen, unabhängig davon, ob sie Dokumenten oder Selbstberichten entstammen.

Im Folgenden wird diese testtheoretische Unterscheidung zu Grunde gelegt, wenn von objektiven bzw. subjektiven Maßen die Rede ist. Anderenfalls wird von externen bzw. internen Maßen gesprochen. Berufliche Zufriedenheit steht nicht im Zentrum des Interesses dieser Arbeit, deshalb wird auf eine ausführlichere Darstellung dieses Erfolgsbereichs verzichtet. Eine Übersicht über Zufriedenheitsindikatoren geben Judge, Cable, Boudreau und Bretz (1995). Über die in dieser Arbeit betrachteten Erfolgsbereiche und ihre Indikatoren gibt Tabelle 2.2 einen Überblick.

Tabelle 2.2

Beispiele für Berufserfolgs-Indikatoren getrennt für verschiedene Erfolgsbereiche, Datenarten und Datenquellen

Datenart	Datenquelle		
	Akten/Dokumente	Fremdbeurteilung	Selbstbericht
Erfolgsbereich <i>Lernen</i>			
neutrale Kennzahlen	Anzahl Prüfungswiederholungen		Ausbildungsdauer
Vergleich mit Bezugsstandards		Benotungen	Leistungseinschätzung
Erfolgsbereich <i>Arbeit</i>			
neutrale Kennzahlen	Stückzahlen		Stückzahlen
Vergleich mit Bezugsstandards		Leistungsbeurteilung	Leistungseinschätzung
Erfolgsbereich <i>Laufbahn</i>			
neutrale Kennzahlen	Gehalt		Gehalt
Vergleich mit Bezugsstandards	Beförderungswürdigkeit	Erfolg im Vgl. zu Kollegen	

Anmerkungen. nach Dette et al. (2004, S. 174)

Neutrale Kennzahlen

Welche Maße stehen zur Operationalisierung zur Verfügung? Als neutrale Kennzahlen für *Laufbahn*-Erfolg können die Gehälter bzw. Löhne in verschiedenen Ausprägungen verwendet werden, z.B. das Bruttojahres- oder Monatsgehalt ggf. inklusive Sonderzulagen, Boni oder Aktienoptionen (Judge et al., 1995). Außer dem statischen Gehalt zu einem bestimmten Zeitpunkt können auch dynamische Maße verwendet werden, die die Gehaltsentwicklung in Bezug auf eine bestimmte Zeitspanne (Wayne, Liden, Kraimer & Graf, 1999) oder in Bezug auf bestimmte Karriereschritte (z.B. seit dem ersten Gehalt; O'Reilly & Chatman, 1994) quantifizieren.

Ein weiterer Indikator ist die Positionshöhe, die ebenfalls statisch oder dynamisch berücksichtigt werden kann. Statisch sind Fragen nach der erreichten Hierarchiestufe, der Anzahl der Hierarchiestufen über der erreichten oder der relativen Hierarchiestufe, bei der die erstgenannten miteinander in Beziehung gesetzt werden (Dette et al., 2004). Um die Vergleichbarkeit zwischen verschiedenen Organisationen zu erhöhen, werden auch mehrstufige Punkteskalen eingesetzt, anhand derer die Positionshöhe eingeschätzt wird (z.B. Schneer & Reitman, 1995; Tharenou, Latimer & Conroy, 1994). Das einfachste dynamische Maß ist die Gesamtzahl der Beförderungen, die auch ins Verhältnis zu den Arbeitsjahren gesetzt werden kann (Dette et al., 2004). Ferner können Verweildauer bzw.

Aufstiegsgeschwindigkeit verwendet werden (z.B. Burchard, 2000; Moser, 2004a). Von welcher Qualität die Beförderungen waren bleibt bei diesen Maßen allerdings unberücksichtigt.

Häufig werden die neutralen Kennzahlen zu Laufbahnerfolg Selbstberichten entnommen, da sie einen relativ langen Zeitraum abdecken und ggf. verschiedene Arbeitgeber betreffen. Eine weitere neutrale Kennzahl ist die Anzahl der unterstellten Mitarbeiter (z.B. Tharenou et al., 1994). Bei Berufseinsteigern, die ihre Ausbildung gerade abgeschlossen haben, bietet es sich an, die Anzahl der angebotenen Arbeitsstellen zu protokollieren und diese evtl. an der Anzahl der durchgeführten Bewerbungen oder Einstellungsinterviews zu relativieren (O'Reilly & Chatman, 1994). Schließlich können noch Berufsprestige-Ratings zu den neutralen Kennzahlen gezählt werden, wenn sie auf gesellschaftsweiten elaborierten Prestigeindizes basieren (Dette et al., 2004). In den USA findet z.B. der Index of Social Position von Hollingshead (1975) Verwendung (z.B. bei Judge et al., 1999). Im deutschen Sprachraum wird die Magnitude-Prestige-Skala (Wegener, 1988) eingesetzt.

Die neutralen Kennzahlen, die im Bereich der *Arbeitsleistung* verwendet werden, umfassen Maße, die die Arbeitsquantität bzw. Produktivität beziffern (z.B. Stückzahlen) ebenso wie quantifizierbare Maße der Arbeitsqualität (z.B. Fehlerraten; Bommer, Johnson, Rich, Podsakoff & MacKenzie, 1995). Neutrale Kennzahlen zur berufsbezogenen *Lernleistung* sind z.B. die Ausbildungsdauer oder die Anzahl der Wiederholungen, bis zum Bestehen des Ausbildungsabschlusses.

Generell haben die neutralen Kennzahlen den Nachteil, nicht für jeden Beruf und jede Branche die gleiche Aussagekraft zu besitzen. Sie lassen sich nicht oder nur schwer über verschiedene Organisationen hinweg generalisieren (Dette et al., 2004). Ebenso werden nicht quantifizierbare Aspekte der Arbeitsqualität nicht berücksichtigt, die für die Arbeitsleistung sehr bedeutsam sein können. Dies verdeutlichen Hunter und Schmidt (1996) am Beispiel von Verhören durch Polizisten: maßgeblich für die Aufklärung von Verbrechen und für Fahndungserfolge sind nicht in erster Linie die Anzahl der Verhöre, die ein Polizist bewältigt, sondern deren Qualität.

Vergleiche mit Bezugsstandards durch subjektive Leistungsbeurteilungen

Als Bezugsstandards kommen organisationale oder individuelle Ziele, andere Personen, die sich in ähnlicher beruflicher Lage befinden, oder eine gesellschaftliche Norm in Frage (Dette et al., 2004).

Vergleiche mit Bezugsstandards können in Form von Fremd- oder Selbstbeurteilungen durchgeführt werden. Von Fremdbeurteilungen zur Einschätzung des *Laufbahn-Erfolgs* wird vergleichsweise selten berichtet. Dette et al. (2004) vermuten als Ursache die erforderliche organisations- bzw. arbeitsplatzübergreifende Perspektive, die von Kollegen und Vorgesetzten i.d.R. nicht eingenommen werden kann. Sie berichten von Einschätzungen des Arbeitsmarktwertes oder der Beförderungswürdigkeit. Auch verschiedene Maße der Arbeitsleistung finden Verwendung, z.B. formale Arbeitsbeurteilungen oder Einschätzungen von Fähigkeiten, dem Fachwissen, der Kooperativität oder Qualität der Arbeit im Rahmen von Führungskräfte-Befragungen (Dette et al., 2004).

Sehr häufig werden Fremdbeurteilungen dagegen zur Einschätzung der *Arbeitsleistung* eingesetzt. Dabei kommen verschiedene Urteilsquellen in Betracht. Häufigste Quelle ist der oder die Vorgesetzte einer Person (Abwärtsbeurteilung), zu dessen Aufgabenbereich häufig die Beurteilung von Mitarbeitern gehört. Vorgesetzten fehlt jedoch oft die Möglichkeit zur direkten Beobachtung von Mitarbeitern und unter Umständen die zur Bewertung erforderliche Fachkenntnis. Aus diesen Gründen kann es sinnvoll sein, Kollegenurteile einzuholen (Seitwärtsbeurteilung). Vorgesetzte selbst können zudem durch ihre Mitarbeitenden beurteilt werden (Aufwärtsbeurteilung), z.B. in Bezug auf ihr Führungsverhalten. Generell kommen alle Personen aus dem Arbeitsumfeld als Urteilsquellen in Betracht, z.B. auch Kunden oder Lieferanten. Die Beurteilungen von verschiedenen Urteilsquellen bewerten unterschiedliche Aspekte der Leistung (Harris & Schaubroeck, 1988); sie sind aus verschiedenen Perspektiven verfasst. In 360°-Feedback-Verfahren können die genannten Urteilsquellen miteinander kombiniert werden, um ein möglichst umfassendes Bild des Leistungsstandes zu erhalten (Scherer, 2004).

Selbstbeurteilungen der Arbeitsleistung finden seltener Verwendung. Sie unterliegen stärkeren Urteilstendenzen (z.B. der Mildtendenz) und ihre Akkuratheit kann nicht bestimmt werden. Eine Diskussion der Güte von Selbstbeurteilungen findet man bei Moser (2004b).

Häufig werden die Beurteilenden gebeten, die Leistung *global* mit einem Item einzuschätzen. Dabei bleibt es ihnen überlassen, wieviele und welche Aspekte von Leistung sie ihrem Urteil zu Grunde legen, wie sie diese gewichten und zu einem Gesamturteil zusammenfügen. Das lässt sich besser nachvollziehen, wenn einzelne Leistungsdimensionen explizit bewertet und zu einem *compositen* Maß verrechnet werden (Bommer et al., 1995). Auch diese sind jedoch von Urteilstendenzen, z.B. der Tendenz zur Mitte, Mildtendenzen oder dem Halo-Effekt betroffen (Schuler, 2004a). Eine Übersicht darüber, welche Leistungsdimensionen bewertet werden können, wurde in Kapitel 2.2.1 aufgeführt.

Zur subjektiven Leistungsbeurteilung können zudem verschiedene Beurteilungsverfahren eingesetzt werden. Eine grundlegende Unterscheidung treffen Bommer et al. (1995) indem sie Beurteilungen, die auf Basis von absoluten Vergleichen mit Leistungsstandards zustande kamen, von solchen abgrenzen, denen relative Vergleiche zwischen Personen zugrunde liegen. Eine Unterscheidung in Bezug auf die Systematik der Beurteilungsdurchführung trifft Schuler (2004b): Der unsystematischen freien Eindrucksschilderung stehen systematische Vorgehensweisen gegenüber. Systematisch ist eine Beurteilung dann, wenn sie a) geplant durchgeführt wurde, b) auf Anforderungsanalysen basiert, um einen hohen Bezug zur Tätigkeit der beurteilten Personen sicherzustellen, c) standardisierte Fragen zumindest aber geregelt geführte Gespräche und d) skalierte Antwortformate eingesetzt werden. Skalierte Antwortformate umfassen nach Marcus und Schuler (2006) Verfahren zur Einstufung von Personen auf verschiedenen Leistungsdimensionen, Kennzeichnungs- sowie Auswahlverfahren, die nicht nach Dimensionen geordnet sind, Verfahren, die sich an Zielerreichungsgraden orientieren und Rangordnungsverfahren, denen ein Vergleich mehrerer Personen zugrunde liegt.

Schließlich ist die Unterscheidung zwischen Einschätzungen der maximalen und ty-

pischen Leistung erwähnenswert (Sackett, 2007). Die maximale Leistung wird in einem vergleichsweise kurzen Zeitraum erfasst, wobei die zu Beurteilenden wissen, dass sie beurteilt werden und das bestmögliche Ergebnis erzielen sollen (z.B. in Assessment-Centern). Die typische Arbeitsleistung spiegelt hingegen die Leistung über einen längeren Zeitraum im Arbeitsalltag wider.

Im Bereich beruflicher *Lernleistung* spielen Fremdbeurteilungen die Hauptrolle. Dies gilt insbesondere bei der Bewertung von Ausbildungs- und Weiterbildungsleistungen. In der Regel werden die Beurteilungen in Deutschland in Form von Noten- oder Punkteskalen vergeben, wie sie auch in den allgemeinbildenden Schulen üblich sind (Tent, 2001). Bewertet werden insbesondere die erworbenen Fachkenntnisse aber auch praktische Fertigkeiten sowie die Anwendung der erworbenen Kenntnisse und Fertigkeiten in der beruflichen Praxis. Bei betrieblichen Beurteilungen können die Beurteilungsverfahren zum Einsatz kommen, die bei Einschätzung der Arbeitsleistung von ausgebildeten Mitarbeitenden benutzt werden (s.o.).

2.3. Intelligenz und beruflicher Erfolg

Betrachtet man die Definitionen und Operationalisierungen von Intelligenz und Berufserfolg (Kapitel 2.1), überraschen hohe Korrelationen zwischen beiden Konstrukten nicht: Intelligenz wird mit Hilfe von Leistungsmaßen operationalisiert und muss sich in Bezug auf Leistungsmaße bewähren (Süß, 2001). Bei Intelligenz und Berufserfolg handelt es sich um hypothetische Konstrukte, die nicht direkt erfassbar sind, sondern nur durch Situations-Reaktions-Beobachtungen indirekt wahrgenommen werden können (Brocke & Beauducel, 2001). Diese Beobachtungen lassen sich in Intelligenztestungen anstellen, wo Antworten (Reaktionen) auf Fragen (Situationen) gegeben werden und im Berufsleben einer Person, wo sie sich unter verschiedenen Bedingungen (Situationen) mehr oder weniger erfolgreich verhält (Reaktion).

Insofern ist es keineswegs eindeutig, ob eine Situation Teil des Prädiktors *Intelligenz* ist oder ein Validitätskriterium darstellt. Eine mathematische Textaufgabe kann z.B. Teil eines Intelligenztests oder Teil der schriftlichen Abschlussprüfung eines kaufmännischen Ausbildungsberufes sein. Manche Intelligenztests, die für bestimmte Berufsgruppen entwickelt wurden, enthalten simulationsorientierte Aufgaben, die in ähnlicher Form auch im Berufsalltag vorkommen (vgl. Kapitel 3.2.1). Die Korrelation zwischen Intelligenz und Berufserfolgs-Indikatoren wird deshalb von verschiedenen Autoren als Methodenartefakt angesehen und nicht als Beleg für eine hohe Validität, sondern als Retestrelabilität betrachtet (Schmitt in Morgeson et al., 2007).

Allerdings unterscheidet sich die Intelligenz-Testsituation von einer Alltagssituation auch in mehrerer Hinsicht. Die Aufgaben eines Intelligenztests zeichnen sich durch folgende Charakteristika aus (Brocke & Beauducel, 2001): a) sie werden nicht von den Probanden selbst erstellt, b) sie sind inhaltlich für die Probanden nicht von großem Interesse, u.a. deshalb, weil sie c) „von den allgemeinen Erfahrungen mehr oder weniger abgehoben sind“ (Brocke & Beauducel, 2001, S. 27), d) sie weisen einen hohen Grad an Strukturiertheit auf, f) haben i.d.R. nur *eine* richtige Lösung, die g) i.d.R. nur auf

einem Lösungsweg erreicht wird; schließlich sind h) alle erforderlichen Informationen bekannt, die zur Lösung benötigt werden. Die Aufgaben des Alltags sind dagegen wesentlich komplexer und inhaltlich relevant. Diese Unterschiede sind zugleich Kritikpunkte an Intelligenztests und haben zur Entwicklung alternativer Verfahren z.B. zur Erfassung komplexen Problemlösens geführt (Kersting, 1998).

Ein hoher Zusammenhang zwischen Intelligenz und Berufserfolg lässt sich aber nicht nur durch die ähnliche Operationalisierung erklären, sondern auch auf Basis von Theorien, die nach dem Zustandekommen von Berufserfolg fragen. Diese Zusammenhänge werden in den folgenden Teilkapiteln getrennt für die drei Erfolgsbereiche berufsbezogene Lernleistung, Arbeitsleistung und Karriereerfolg dargestellt.

Die empirischen Befunde dazu stammen bis zur Jahrtausendwende nahezu ausschließlich aus den USA. Erst in den letzten Jahren wurden auch Metaanalysen mit Studien aus verschiedenen europäischen Ländern (aus der früheren EU-15 und Skandinavien) und für einzelne europäische Länder – darunter Deutschland – durchgeführt. Diese Replikationen sind erforderlich, weil sich die verwendeten Leistungskriterien und Intelligenzverfahren, das organisationale Umfeld und das Verständnis von Personalauswahl zwischen verschiedenen Ländern unterscheiden (Bertua, Anderson & Salgado, 2005; Salgado et al., 2003a). Zum Beispiel werden Intelligenztests in Europa insgesamt bei der Personalauswahl stärker genutzt als in den USA, während sie in Deutschland seltener eingesetzt werden (Ryan, MacFarland, Baron & Page, 1999; Salgado & Anderson, 2002). Aus diesen Gründen wird diskutiert, ob sich die Befunde zur Validität von Intelligenz aus den USA auf andere Länder und Kulturen übertragen lassen.

Im Vergleich zu den USA könnte sich verringernd auf die Validität von Intelligenztests auswirken, dass in Europa bei Leistungsbeurteilungen Arbeitsengagement und Kooperativität (*contextual performance*) eine größere Rolle spielen als in den USA. Contextual performance hängt jedoch weniger stark mit Intelligenz zusammen (Borman, Penner, Allen & Motowidlo, 2001). Zudem betonen Salgado et al. (2003a), dass die Leistungskriterien in Europa weniger reliabel erfasst würden und dass höhere Arbeitslosenquoten zu anderen Selektionsquoten sowie zu stärkeren Varianzeinschränkungen führen könnten. Ausbildungsleistungen werden in den USA stärker mit Wissenstests erfasst und weniger stark mit Leistungsbeurteilungen als in Europa. Schließlich ist man in den USA eher an maximaler Arbeitsleistung und in Europa eher an typischer Arbeitsleistung interessiert (Ackerman, 1994; Ackerman & Humphreys, 1991). Beide Unterschiede lassen eher eine geringere Validität von Intelligenz in Europa vermuten, da die in Europa häufiger verwendeten Kriterien typischer Arbeitsleistung in stärkerem Maße neben der Allgemeinen Intelligenz von zusätzlichen Faktoren beeinflusst werden.

Als Gründe für eine möglicherweise höhere berufskriterienbezogene Validität von Intelligenztests in Europa als in den USA führen Salgado et al. (2003a) auch die größere Sprachvielfalt an, die das Arbeiten in Europa möglicherweise insgesamt komplexer mache. Dadurch könnten die Anforderungen an die verbale Intelligenz der Arbeitnehmer höher sein. Weitere Unterschiede betreffen die Organisationen, von denen Validierungsstudien vorliegen. Diese sind in den USA im Durchschnitt größer (Bertua et al., 2005). Die Organisationsgröße kann relevant sein, da sich die Personalauswahl-Praxis von

großen Organisationen – die z.B. über eigene Human-Research-Abteilungen verfügen – von der kleiner Organisationen unterscheidet (Ryan et al., 1999). Auch das organisationale Umfeld sei in Europa ein anderes. Neben der unterschiedlichen Arbeitsmarktsituation unterscheidet es sich vor allem durch andere gesetzliche Regelungen, z.B. den stärkeren Schutz der Privatsphäre der Arbeitnehmer, der sich auch auf das Verständnis und die Durchführung von Personalauswahl auswirkt (Levy-Leboyer, 1994). Schließlich werden in den USA andere Intelligenztests als in Europa eingesetzt: besonders häufig wird das Omnibusverfahren *Wonderlic Personnel Test* (WPT; Wonderlic & Hovland, 1939; Wonderlic Inc., 2002) verwendet (F. L. Schmidt & Hunter, 2004). In Deutschland liegen hauptsächlich Studien mit Intelligenzstrukturtests vor (vgl. Kapitel 4).

2.3.1. Intelligenz und berufliche Lernleistung

Weshalb GMA für Lernleistung eine hohe Bedeutung hat, erklären Hunter und Schmidt (1996) mit der klassischen Lerntheorie von Thorndike: „Edward Thorndike’s many studies of educational, vocational, and industrial training lead him to conclude, that the main determinant of individual differences in job performance is individual differences in learning“ (S. 459). Die Aneignung neuen Wissens kann in formalen Programmen stattfinden – dies erfordert die aktive Aufnahme des Wissens – oder *on the job* – dies erfordert zweierlei: 1. die Selektion bedeutsamen Wissens und 2. die Strukturierung und Speicherung des Wissens in einer Form, die dem späteren Abruf dienlich ist (Hunter, 1986). Intelligenten Personen gelingt dies leichter, sie lernen mehr, schneller und sie reagieren auf neue Herausforderungen, bei denen Erfahrungswissen nicht ausreicht, innovativer als weniger intelligente Personen. Entsprechend definieren Kuncel et al. (2004) „*g*, in part, as an ability or capacity to learn and acquire new knowledge and skill“ (S. 149).

Auf Lernerfolg haben neben kognitiven Fähigkeiten auch andere Variablen einen Einfluss, z.B. Persönlichkeitseigenschaften (Gewissenhaftigkeit, locus of control, Leistungsmotivation, Ängstlichkeit), das Alter, Situationsvariablen (Klima, Unterstützung durch Vorgesetzte und Kollegen), Berufs- und Karriere-Variablen (Involvement, berufliches, organisationales Commitment, Karriereplanung; vgl. Colquitt, LePine & Noe, 2000) sowie motivationale Variablen, Interessen und Werte (Reeve & Hakel, 2002). Intelligenz ist somit notwendig, aber nicht hinreichend für beruflichen Lernerfolg.

In verschiedenen Metaanalysen werden deutliche positive Zusammenhänge zwischen Intelligenz und beruflicher Lernleistung gefunden. Die Daten, die den Analysen zu Grunde liegen, stammen bis zur Jahrtausendwende überwiegend aus den USA. Eine tabellarische Übersicht darüber gibt Campion in Morgeson et al. (2007). Auf dieser Übersicht basiert Tabelle 2.3: sie gibt einen Überblick über die untersuchten Berufsgruppen, die Anzahl der inkludierten Studien (k), die Stichprobengröße und die mittleren beobachteten (\bar{r}) und korrigierten ($\hat{\rho}$) Effekte. Weitere Ergebnisse und Angaben zum methodischen Vorgehen werden ebenfalls gemacht. Auf sie wird in späteren Kapiteln Bezug genommen.

Neuere Metaanalysen untersuchen die Validität von Intelligenz in westeuropäischen Ländern, drei davon berichten Ergebnisse aus Deutschland (Hülshager, Maier & Stumpp, 2007; Hülshager, Maier, Stumpp & Muck, 2006; Salgado & Anderson, 2003). Die Ergebnisse liegen z.T. getrennt für verschiedene Berufsgruppen vor. In einigen Analysen sind

Tabelle 2.3

Metaanalytische Befunde aus den USA zur prädiktiven Validität von Intelligenz in Bezug auf Lernleistung

Quelle	Berufsgruppe	Prädiktor	Methode ^a	r_{xx}	r_{yy}	u_x	k	N	\bar{r}	$\hat{\rho}$	CV	VE	Gen.
F. L. Schmidt, Gast-Rosenberg und Hunter (1980)	Programmierer	GMA	HS interaktiv	.90	.60	.60	9	1 635	.64	.91	.70	41	TRA
Pearlman, Schmidt und Hunter (1980)	Bürotätigkeiten	GMA	HS interaktiv	.80	.80	.59	65	32 157	.44	.71	.56	62	TRA
F. L. Schmidt, Hunter und Pearlman (1981)	Bürotätigkeiten	GMA	HS interaktiv	–	.80	.59	61	31 535	.43	.70	o.A.	o.A.	o.A.
Hunter und Hunter (1984), Reanalyse von Ghiselli (1973)	Manager Bürotätigkeiten Sicherheitskräfte Dienstleistungen Kraftfahrzeugmechaniker Handel und Handwerk Industrie	GMA	HS interaktiv	–	.81	.60	o.A.	500-999 > 10 000 1 000 – 4 999 1 000 – 4 999 1 000 – 4 999 > 10 000 5 000 – 9 999	.30 .47 .42 .42 .18 .41 .38	.51 .71 .87 .66 .37 .65 .61	o.A.	o.A.	o.A.
Hunter und Hunter (1984); Hunter (1986)	Komplexitätslevel 1 <i>Set-up/Precision</i> Komplexitätslevel 2 <i>Synthesize/Coordinate</i> Komplexitätslevel 3 <i>Analyze/Compile</i> Komplexitätslevel 4 <i>Copy/Compare</i> Militär	GMA	HS interaktiv	–	.81	.60	4 24 54 8 828	235 1 863 3 823 575 472 539	o.A.	.65 .50 .57 .54 .62	o.A.	o.A.	o.A.
Hirsh, Northrop und Schmidt (1986)	Strafverfolgung	Gedächtnis numerisches Denken schlussfolgerndes Denken räumliches/ mechanisches Denken verbales Denken	HS interaktiv	.80	.73	.59	6 9 24 13 26	801 1 206 4 374 1 422 3 943	.22 .34 .33 .28 .37	.41 .63 .61 .50 .64	.41 .38 .47 .50 .37	100 48 75 100 45	CSC TRA CSC CSC TRA

Tabelle 2.3
(Fortsetzung)

Quelle	Berufsgruppe	Prädiktor	Methode ^a	r_{xx}	r_{yy}	u_x	k	N	\bar{r}	$\hat{\rho}$	CV	VE	Gen.
Hartigan und Wigdor (1989)	Komplexitätslevel 1 <i>Set-up/Precision</i>	GMA	HS interaktiv	o.A.	.80	–	o.A.	64	.54	.60	o.A.	o.A.	o.A.
	Komplexitätslevel 2 <i>Synthesize/Coordinate</i>							347	.30	.33			
	Komplexitätslevel 3 <i>Analyze/Compile</i>							3 169	.36	.40			
	Komplexitätslevel 4 <i>Copy/Compare</i>							106	.00	.00			
Levine, Spector, Menon, Narayanon und Cannon-Bowers (1996)	Handwerker der Versorgungswirtschaft	GMA	Raju, Burke, Normand und Langlois (1991)	.80	.76	.59	52	5 872	.38	.67	o.A.	86	o.A.
Barrett, Polomsky und McDaniel (1999)	Feuerwehrleute	GMA	HS interaktiv (op. Val.)	–	–	–	14	2 007	.50	.77	.73	o.A.	o.A.

Anmerkungen. Die Angaben sind für jede Quelle nur einmal gemacht, es sei denn sie ändern sich. o.A. = ohne Angabe. ^aHS = Hunter und Schmidt, op. Val = operationale Validität. r_{xx} = Prädiktorreliabilität wenn korrigiert, r_{yy} = Kriterienreliabilität wenn korrigiert, u_x Varianzeinschränkung im Prädiktor wenn korrigiert, k = Anzahl der Studien, N = Stichprobengröße gesamt, \bar{r} = beobachtete Korrelation, $\hat{\rho}$ = korrigierte Korrelation, CV = untere Grenze des 90% Glaubwürdigkeits-Intervalls, VE = Anteil der durch die Artefakte aufgeklärten Varianz an der beobachteten Varianz von \bar{r} bzw. $\hat{\rho}$ in %, Gen. = Generalisierung: TRA = transportability, CSC = cross-situational consistency. Die Zusammenstellung der Tabelle erfolgte auf Basis von Campion in Morgeson et al. (2007).

die Berufe verschiedenen Komplexitätsniveaus zugeordnet. Den verschiedenen Analysen der Arbeitsgruppe um Hunter und Schmidt liegen z.T. dieselben Daten zu Grunde. Das gilt auch für die vier Veröffentlichungen von Salgado et al. und die Veröffentlichungen von Hülshager et al. Eine Übersicht über die Befunde gibt Tabelle 2.4. Auch sie enthält Angaben zur Methode und Ergebnisse, auf die erst in späteren Kapiteln Bezug genommen wird. Zunächst interessieren die mittleren Effekte \bar{r} und $\hat{\rho}$.

Mit einer Ausnahme (Berufsgruppe *copy/compare* aus der Analyse von Hartigan & Wigdor, 1989), die eine Korrelation von .00 berichtet, werden ausschließlich positive Korrelationen beobachtet, die in zwei Fällen klein sind ($.10 \leq \bar{r} < .30$; Cohen, 1992), in elf Fällen mittelgroß ($.30 \leq \bar{r} < .50$) und in 23 Fällen groß ($.50 \leq \bar{r}$): In fünf Fällen werden sie nicht berichtet. Die korrigierten Korrelationen sind, ebenfalls von der genannten Ausnahme abgesehen, in zehn Fällen mittelgroß ($.30 \leq \hat{\rho} < .50$) und in den restlichen 31 Fällen groß ($.50 \leq \hat{\rho}$).

Die Ergebnisse aus den verschiedenen Metaanalysen variieren aus verschiedenen Gründen: die inkludierten Studien unterscheiden sich z.B. hinsichtlich der Berufsgruppen, und es werden verschiedene statistische Vorgehensweisen gewählt (vgl. Kapitel 3). Als repräsentative Größe für die prädiktive berufsübergreifende Validität in den USA wird von F. L. Schmidt und Hunter (1998a) $\hat{\rho} = .56$ angegeben. Der Wert basiert auf zwei Metaanalysen von Hunter und Mitarbeitern (Hunter, 1980; Hunter & Hunter, 1984). Die geschätzte Validität aus Europa entspricht mit .54 nahezu den amerikanischen Befunden (Salgado & Anderson, 2003).

Uneinheitlich sind die Ergebnisse zur Validität in Deutschland: Salgado und Anderson (2003) errechnen $\hat{\rho} = .63$ auf Basis von 26 Studien; das Ergebnis von Hülshager et al. (2006) liegt bei $\hat{\rho} = .47$ auf Basis von 90 Studien. Diese Unterschiede sind einerseits auf die unterschiedliche Datenbasis zurückzuführen und andererseits auf verschiedene statistische Vorgehensweisen, auf die in Kapitel 3 näher eingegangen wird. Vorläufig soll die Feststellung genügen, dass die empirischen Ergebnisse einen *positiven* Zusammenhang zwischen Intelligenz und berufsbezogener Lernleistung in den USA, Europa und Deutschland belegen.

Dieser Befund ist generalisierbar über verschiedene Intelligenztests, Organisationen, Berufe, Messzeitpunkte etc. hinweg: generell ist der Zusammenhang zwischen Intelligenz und Lernleistung *positiv*, da die unteren Grenzen der Glaubwürdigkeitsintervalle (CV; nicht zu verwechseln mit Konfidenzintervallen [CI]; vgl. Kapitel 3.8) über Null liegen. Die Höhe des Zusammenhangs kann jedoch nicht verallgemeinert werden, da die Anteile der durch die Artefakte aufgeklärten Varianzen (VE) unter 75% liegen. Diese Art der Generalisierung bezeichnen Kemery, Mosholder und Roth (1987) als Transportabilität (*transportability*, TRA; vgl. Kapitel 3.8).

Ein Moderator, der die Höhe der Validität zwischen Intelligenz und berufsbezogener Lernleistung beeinflusst, ist die Berufskomplexität. Eine komplexe Tätigkeit zeichnet sich dadurch aus, dass sie zu ihrer erfolgreichen Bewältigung den Erwerb und die Verarbeitung umfangreicher vielfach vernetzter Informationen erfordert (Gottfredson, 2003). Eine hohe Komplexität ist nicht nur in akademischen Berufen anzutreffen (Kuncel et al., 2004).

Die durchschnittliche Intelligenz von Berufstätigen in verschiedenen Berufen unterscheidet sich. Mit zunehmender Berufskomplexität steigt die mittlere Intelligenz und ihre Varianz nimmt ab. Dies wird z.B. anhand der Normwerte des *Army General Classification Tests* deutlich (Abbildung 2.2, Seite 44). Für die Entwicklung verschiedener Berufe ist neben den unterschiedlichen Gegenstandsbereichen die kognitive Komplexität der Anforderungen maßgeblich: ein Beruf umfasst ähnlich komplexe Tätigkeiten (Gottfredson, 2003).

Gemäß der Bedeutung, die Berufskomplexität für die Klassifikation der Berufe hat, spielt sie in Berufssystematiken eine entscheidende Rolle. Das gilt insbesondere für das *Dictionary of Occupational Titles* (DOT; U. S. Department of Labor, 1991), das in den USA zur Systematisierung der Berufe verwendet wurde². Im DOT werden berufliche Tätigkeiten daraufhin eingeschätzt, welche Anforderungen sie an den Umgang mit Daten, Personen und Dingen stellen (U. S. Department of Labor, 1991). Zur Operationalisierung von Komplexität verwenden Hunter und Hunter (1984) und Hartigan und Wigdor (1989) die Daten- und Dingedimensionen des DOT. Sie unterscheiden die Komplexitätslevel 1 *set-up/precision* (höchste Komplexität), 2 *synthesize/coordinate*, 3 *analyze/compile*, 4 *copy/compare* und 5 *feeding/offbearing* (geringste Komplexität). Auch Salgado et al. (2003a) operationalisieren Komplexität mit den DOT-Dimensionen, verwenden aber nur die Komplexitätslevel 2 (hoch) bis 4 (gering).

Die Validität von Intelligenz in Bezug auf Lernleistung unterscheidet sich zwischen Berufen unterschiedlicher Komplexität (vgl. Tabelle 2.3 und Tabelle 2.4). Sie steigt im Großen und Ganzen betrachtet bei zunehmender Berufskomplexität an, ist aber auch bei geringer Komplexität größer als Null.

Welchen Einfluss dabei unterschiedliche Varianzeinschränkung im Prädiktor spielen, lässt sich anhand der Studien aus den USA nicht einschätzen. Die Varianzeinschränkungen werden nicht berichtet und nicht korrigiert (Hartigan & Wigdor, 1989) oder nur über alle Komplexitätslevels hinweg berichtet und korrigiert (Hunter, 1986). Im Gegensatz dazu differenzieren Salgado et al. (2003b) zwischen den Komplexitätsleveln und zeigen, dass die Varianzeinschränkung im Prädiktor mit zunehmendem Komplexitätsniveau ebenfalls zunimmt (vgl. Tabelle 2.4): Das Verhältnis der Intelligenz-Standardabweichung der varianzeingeschränkten Stichproben zur Standardabweichung der nicht eingeschränkten Population (u -Wert) beträgt .83 bei niedriger, .67 bei mittlerer und .43 bei hoher Komplexität. Je geringer der u -Wert, desto höher die Varianzeinschränkung.

In Deutschland ist die Varianzeinschränkung von Intelligenz zwischen den Bildungsniveaus im Vergleich zu anderen europäischen Staaten und den USA besonders groß (Schuler & Höft, 2006). Die Intelligenzwerte von Hauptschülern, Realschülern und Gymnasiasten liegen jeweils ca. 1 SD auseinander und sie verfügen über eine unterschiedliche Varianz: je höher das Bildungsniveau, desto geringer fällt die Streuung der Intelligenzwerte aus (Schuler, 2002a). Da das Bildungsniveau ausschlaggebend für die Komplexität des Berufes ist, der erlernt wird, sollten die Varianzunterschiede zwischen den verschiedenen komplexen Berufen mit deutschen Studien deutlicher ausfallen als in internationalen Analysen.

²Inzwischen ist es von einer überarbeiteten Fassung, dem O*Net, abgelöst (Peterson et al., 2001).

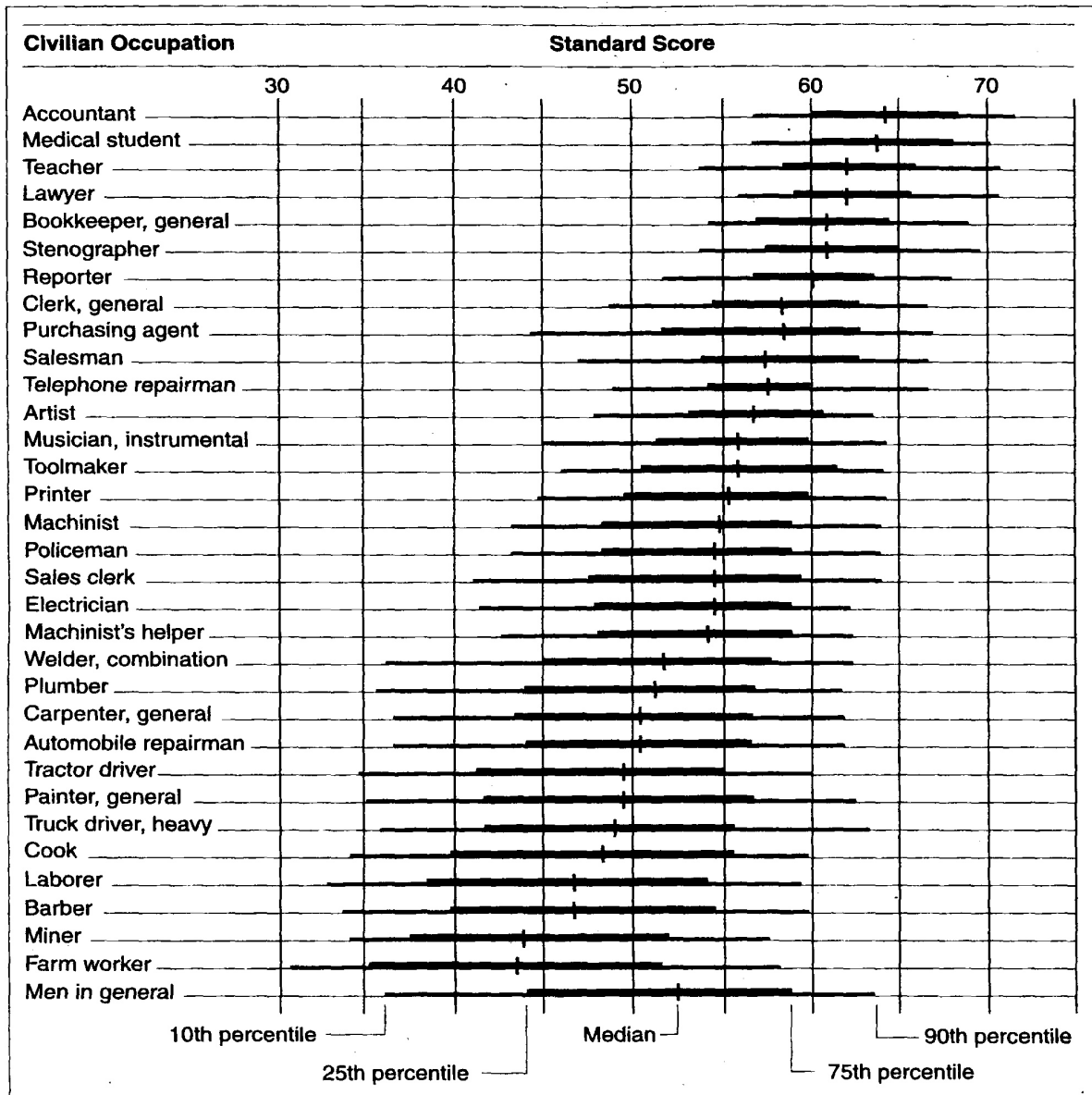
Tabelle 2.4

Metaanalytische Befunde aus Europa zur prädiktiven Validität von Intelligenz in Bezug auf Lernleistung

Quelle	Berufskomplexität ^a	Land	Methode ^b	r_{xx}	r_{yy}	u_x	k	N	\bar{r}	$\hat{\rho}$	CV	VE	Gen.
Salgado und Anderson (2002)		UK	HS interaktiv (op. Val.)	.85	.80	.61	61	20 305	.34	.56	.46	87	CSC
		E				.80	25	2 405	.35	.47	.25	47	TRA
Salgado et al. (2003a)		Europa ^c	HS interaktiv (op. Val.)	.83	.56	.67	97	16 065	.28	.54	.29	47	TRA
Salgado et al. (2003b)	hohe Komplexität mittlere Komplexität geringe Komplexität	Europa ^c	HS interaktiv (op. Val.)	.83	.56	.43	13	2 619	.29	.74	.74	100	CSC
						.67	35	4 304	.29	.53	.24	40	TRA
						.83	21	4 731	.23	.36	.18	34	TRA
Salgado und Anderson (2003)		F	HS interaktiv (op. Val.)	.83	.56	–	22	5 796	.28	.38	.26	59	TRA
		B, NL				.50	8	706	.27	.65	.48	85	CSC
		E				.74	7	712	.35	.61	.41	59	TRA
		UK				.63	59	11 128	.29	.58	.29	41	TRA
		D				.68	26	4 645	.34	.63	.42	51	TRA
Bertua et al. (2005)		UK	HS interaktiv (op. Val.)	.85	.80	.60	53	17 982	.29	.50	.33	64	TRA
Hülshager et al. (2007)	Gesamt hohe Komplexität mittlere Komplexität geringe Komplexität	D	Raju et al. (1991) (op. Val.)	.91	.80	.69	90	11 969	.31	.47	.27	34	TRA
						.67	6	1 089	.19	.30	.18	58	TRA
						.67	35	5 510	.29	.45	.25	33	TRA
						.68	45	4 931	.35	.52	.36	45	TRA

Anmerkungen. Die Angaben sind für jede Quelle nur einmal gemacht, es sei denn sie ändern sich. ^aDie Berufskomplexität ist nur angegeben, wenn Berufe unterschiedlicher Komplexität getrennt berücksichtigt werden. ^bHS = Hunter und Schmidt, op. Val = operationale Validität. ^cBelgien, Deutschland, Frankreich, Großbritannien, Irland, Niederlande, Portugal, skandinavische Länder, Spanien. r_{xx} = Prädiktorreliabilität wenn korrigiert, r_{yy} = Kriterienreliabilität wenn korrigiert, u_x Varianzeinschränkung im Prädiktor wenn korrigiert, k = Anzahl der Studien, N = Stichprobengröße gesamt, \bar{r} = beobachtete Korrelation, $\hat{\rho}$ = korrigierte Korrelation, CV = untere Grenze des 90% Glaubwürdigkeits-Intervalls, VE = Anteil der durch die Artefakte aufgeklärten Varianz an der beobachteten Varianz von \bar{r} bzw. $\hat{\rho}$ in %, Gen. = Generalisierung: TRA = transportability, CSC = cross-situational consistency.

Abbildung 2.2
Intelligenz und Beruf



Quelle: Cronbach, 1960, aus Schuler & Höft, 2006, S. 108

Metaanalytische Befunde dazu liegen von Hülshager et al. (2007) vor. Im Gegensatz zu den bisher besprochenen Befunden verwenden Hülshager et al. die Berufskodierung der *International Standard Classification of Occupations* (ISCO-88; International Labour Office, 1990) in der europäischen Adaptation (ISCO-88 COM; Elias & Birch, 1994).

Auch die ISCO-88-Kodierung beinhaltet verschiedene Komplexitätszuordnungen, die *Skill Level* der *International Standard Classification of Education* (ISCED). Sie geben an, welches Bildungsniveau für die Erlernung einer Tätigkeit vorausgesetzt wird (International Labour Office, 1990; vgl. Kapitel 3.4): Wissenschaftliche Berufe erhalten den Skill Level 4, Techniker und gleichrangige nichttechnische Berufe den Skill Level 3, Berufe die eine Ausbildung oder längere Anlernzeit voraussetzen den Skill Level 2 und Hilfsarbeits-Tätigkeiten, die keine Lernvoraussetzungen stellen, den Skill Level 1. Der Skill Level 2 ist sehr breit, er umfasst alle Berufe von Anlerntätigkeiten über alle dualen Ausbildungsberufe bis zum Meister (Geis, 2006). Soldaten, Angehörige gesetzgebender Körperschaften und leitende Verwaltungsbedienstete werden keinem Skill Level zugeordnet (vgl. Tabelle 3.5). Hülshager et al. haben diese Berufsgruppen – unter Berücksichtigung der Bildungsniveaus – wenn möglich mit dem höchsten Skill Level versehen.

Die beobachteten Korrelationen nehmen in der Studie von Hülshager et al. (2007) mit zunehmendem Komplexitätsniveau ab ($\hat{r} = .19$ hohe, $.29$ mittlere und $.35$ geringe Komplexität). Der Unterschied ist aber nicht auf die Varianzeinschränkungen zurückzuführen. Diese sind minimal: das Verhältnis der varianzeingeschränkten Standardabweichung zur Populationsstandardabweichung beträgt $u = .68$ für geringe Komplexität vs. $u = .67$ für mittlere Komplexität; zur hohen Komplexität liegen keine Varianzeinschränkungs-Daten vor. Zu ihrer Korrektur wird der mittlere u -Wert von allen Studien ($.67$) verwendet. Im Gegensatz zu den Befunden aus den USA und Europa stellen Hülshager et al. (2007) eine Verringerung der korrigierten Validität mit zunehmender Berufskomplexität fest ($.52$ bis $.30$).

Auf den ersten Blick sind die Befunde zur Validität von Intelligenz in Bezug auf Lernleistung bei verschiedenen komplexen Berufen somit widersprüchlich: in Deutschland nimmt die Validität mit zunehmender Komplexität ab und in Europa zu. In den USA nimmt die Komplexität zu, allerdings nicht kontinuierlich. Die Befunde aus den USA und Deutschland ähneln sich, wenn man nur die mittleren Komplexitätslevel 2 und 3 betrachtet, zu denen in beiden Ländern der Großteil der Studien durchgeführt wurde: die Validität bei höherer Komplexität ist geringer. Die Ergebnisse zu den Berufen höchster Komplexität müssen zurückhaltend interpretiert werden, da die Anzahl der Studien gering ist ($k = 4$ bei Hunter et al., $k = 6$ bei Hülshager et al.). Bestehende Unterschiede zwischen den Befunden aus Deutschland, Europa und den USA können möglicherweise auf die Operationalisierungen zurückzuführen sein. Die DOT-Operationalisierung von Komplexität ist detaillierter als die vergleichsweise globale Zuordnung der ISCO. Eine DOT-Klassifizierung der deutschen Ausbildungsberufe liegt meines Wissens nicht vor.

Zusammenfassend kann festgehalten werden, dass bisherige Befunde auf Basis von 26 bzw. 90 Studien aus Deutschland hohe korrigierte Validitäten zwischen Intelligenz und berufsbezogener Lernleistung finden. Die Höhe der Validität wird in beiden Analysen unterschiedlich eingeschätzt ($\hat{\rho} = .63$ vs $.47$). Sie ist zudem abhängig von der Berufs-

komplexität, wobei beide Arbeitsgruppen zu unterschiedlichen Ergebnissen in Bezug auf die Richtung des Einflusses der Komplexität kommen.

2.3.2. Intelligenz und Arbeitsleistung

Hunter und Schmidt konnten zeigen, dass für Arbeitsleistung berufsübergreifend eine Fähigkeit von besonderer Bedeutung ist: GMA (s.u.; F. L. Schmidt & Hunter, 1998a). Reeve und Hakel (2002) schlussfolgern deshalb: „If I-O psychologists could know only one thing about individuals, we would be best served by choosing to know *g*“ (S. 51). Außer dem *g*-Faktor sind je nach Aufgabenfeld *spezifische* kognitive Fähigkeiten erforderlich, von denen die erfolgreiche Bewältigung der Aufgaben abhängt (Lawshe 1952, 1975, 1984; nach Hunter & Schmidt, 1996). Um sie zu messen, sind Testbatterien erforderlich, die in Form von Intelligenzstruktur-Tests vorliegen.

Ferner gilt auch für Arbeitsleistung: eine Aufgabe kognitiv bewältigen zu können, ist notwendig, aber nicht ausreichend für Erfolg. Die *nicht* kognitiven Fähigkeiten, die im Zusammenhang mit Lernerfolg genannt wurden, sind ebenfalls wichtig (s.o). Dies gilt insbesondere in Berufen mit geringer Komplexität (s.u.) und für Aufgaben, die dem umfeldbezogenen Leistungsbereich zugeordnet werden können (Borman, Hansen & Hedge, 1997; Campbell et al., 1990; zitiert nach Reeve & Hakel, 2002, S. 53). Kognitive Fähigkeiten hängen besonders mit aufgabenbezogener und adaptiver Leistung zusammen (Hunter & Schmidt, 1996).

Auf welche Art und Weise GMA Einfluss auf die Erfolgskriterien nimmt, beschäftigt Hunter (1986). Er geht von folgenden Annahmen aus: Die erfolgreiche Aneignung von Fachwissen wird als eine Voraussetzung für hohe Arbeitsleistung (*work sample performance*) angesehen. Fachwissen fungiert somit als Mediator zwischen Intelligenz und Leistung. Neben diesem indirekten medierten gibt es auch einen vergleichsweise kleinen direkten Einfluss von Intelligenz auf *work sample performance* (vgl. Abbildung 2.3, Seite 53): Die allgemeinen mentalen Fähigkeiten sind für die Arbeitsleistung wesentlich, da von ihnen abhängt, ob eine Aufgabe kognitiv bewältigt (O'Reilly & Chatman, 1994) und gelerntes Wissen in erfolgreiches Verhalten umgesetzt werden kann (Hunter & Schmidt, 1996); intelligente Arbeitskräfte erledigen kognitiv anspruchsvolle Aufgaben schneller, können besser Prioritäten setzen, sind innovativer und passen sich schneller an neue Situationen an (Hunter, 1986).

Subjektive Fremdbeurteilungen von Leistung sind von nicht arbeitsbezogenen Faktoren beeinflusst, die die Wahrnehmung einer Person betreffen, z.B. dem äußeren Erscheinungsbild. Ferner von idiosynkratischen Faktoren, die beurteilerspezifisch wirken, z.B. der Ähnlichkeit der Persönlichkeiten von Beurteilten und Beurteilenden. Beurteilungen durch Einzelpersonen sind von beiden Fehlern behaftet. Bei Beurteilungen durch mehrere Personen wird davon ausgegangen, dass sich idiosyncratische Fehler ausmitteln. Die allgemeinen nicht arbeitsbezogenen verbleiben (Hunter & Schmidt, 1996). Hirsh et al. (1986) meinen, dass bei sehr spezialisierten Tätigkeiten, die unabhängiges Arbeiten erfordern, die Leistung nur schwer eingeschätzt werden kann.

Aus diesen Gründen ist es sinnvoll, die Indikatoren *Arbeitsproben* (*job performance*) und *Fremdbeurteilungen* (*supervisor ratings*) getrennt zu betrachten: Für Arbeitsproben

gelten die genannten Annahmen (direkter Einfluss von GMA *und* indirekter Einfluss, mediiert über Fachwissen), Fremdbeurteilungen sind Einschätzungen sowohl des Fachwissens als auch der objektiven Arbeitsleistung.

Wie hoch ist die Validität von GMA in Bezug auf Arbeitsleistung? Dazu liegen zahlreiche Studien vor, die in mehreren Metaanalysen aggregiert wurden. Die Metaanalysen stammen überwiegend aus den USA. Zusammenstellungen dieser Befunde sind bei Bobko, Roth und Potosky (1999) und Campion in Morgeson et al. (2007) zu finden. Sie sind – um einige Angaben erweitert – in Tabelle 2.5 einzusehen. Zu westeuropäischen Ländern liegen ebenfalls Befunde vor, die in Tabelle 2.6 zusammengestellt sind. Zwei Studien berichten Ergebnisse für deutsche Stichproben (Hülsheger et al., 2007; Salgado & Anderson, 2003). Als Kriterien werden überwiegend subjektive Leistungsbeurteilungen benutzt, in einigen Fällen Arbeitsproben. Den verschiedenen Analysen von Hunter und Schmidt liegen z.T. dieselben Daten zu Grunde, ebenso verhält es sich bei den vier Veröffentlichungen von Salgado et al.

20 beobachtete Korrelationen werden nicht berichtet, von den 48 berichteten beobachteten Korrelationen sind alle positiv, vier unter .10, 36 klein und acht mittelgroß (nach Cohen, 1992). Die wahren Werte werden in drei Studien nicht geschätzt, in den restlichen 59 Studien ist $\hat{\rho}$ in 17 Fällen klein, in 15 Fällen mittelgroß und in 27 Fällen groß. Da GMA in den meisten Studien zuerst erhoben wurde und die Leistungskriterien später, wird damit auch empirisch deutlich, dass GMA ein valider *Prädiktor* für Arbeitsleistung ist.

Die Ergebnisse aus den verschiedenen Metaanalysen variieren aus verschiedenen Gründen: die inkludierten Studien unterscheiden sich (z.B. hinsichtlich der Berufsgruppen) und es werden verschiedene statistische Vorgehensweisen gewählt. Als repräsentative Größe für die prädiktive Validität in den USA wird von F. L. Schmidt und Hunter (1998a) $\hat{\rho} = .51$ angegeben, die auf zwei Metaanalysen von Hunter et al. basiert (Hunter, 1980; Hunter & Hunter, 1984).

Eine weitere berufsübergreifende Metaanalyse mit großem N aus den USA stammt von (Hartigan & Wigdor, 1989). Sie kommen zu einer deutlich geringeren Validität (.22). Die Studie weist jedoch methodische Mängel auf, auf die in Kapitel 3.6.2 näher eingegangen wird.

Auch die Annahmen zum Wirkungsmodell von Intelligenz (s.o. und Abbildung 2.3) konnten Hunter und Schmidt empirisch belegen, indem sie Daten aus 14 Studien reanalysierten (Hunter, 1986; Hunter & Schmidt, 1996). Bei zivilen Berufen korreliert GMA (bei korrigierter Varianzeinschränkung im Prädiktor) mit Fachwissen zu $\hat{\rho} = .80$, mit objektiver Arbeitsleistung insgesamt (d.h. über den direkten und indirekten Einfluss) zu $\hat{\rho} = .75$ und mit subjektiven Leistungseinschätzungen durch Vorgesetzte insgesamt zu $\hat{\rho} = .47$. Bei militärischen Berufen sind diese Einflüsse geringer: Die Korrelationen von GMA mit Fachwissen sowie objektiven und subjektiven Arbeitsleistungsmaßen liegen in militärischen Stichproben bei $\hat{\rho} = .63$, $.53$ bzw. $.24$.

Akzeptiert man die Validität von .51 als Schätzer der tatsächlichen Validität für die USA (F. L. Schmidt & Hunter, 1998a, 1998b), fällt auf, dass die geschätzte Validität aus Europa ($\hat{\rho} = .62$; Salgado & Anderson, 2003) deutlich höher liegt. Das gilt auch für die

Tabelle 2.5

Metaanalytische Befunde aus den USA zur prädiktiven Validität von Intelligenz in Bezug auf Arbeitsleistung

Quelle	Berufsgruppe	Prädiktor/Kriterium ^a	Methode ^b	r_{xx}	r_{yy}	u_x	k	N	\bar{r}	$\hat{\rho}$	CV	VE	Gen.	
F. L. Schmidt, Hunter, Pearlman und Shane (1979)	Bürotätigkeiten: EDV, Buchhaltung		HS nicht interaktiv	.80	.60	.60	58	5 433	o.A.	.49	.22	57	TRA	
							65	3 986		.61	.15	43	TRA	
							75	5 143		.64	.35	66	TRA	
Hunter (1980)	zivile Berufe	Abwärtsbeurteilung Arbeitsproben	HS	-	-	o.A.	o.A.	1 790	o.A.	.47	o.A.	o.A.	o.A.	
								1 790		.75				
	militärische Berufe	Abwärtsbeurteilung Arbeitsproben							1 474		.24			
									1 474		.53			
Pearlman et al. (1980)	Bürotätigkeiten		HS interaktiv	.80	.60	.59	194	17 539	.26	.52	.21	51	TRA	
F. L. Schmidt et al. (1980)	Programmierer		HS interaktiv	.90	.60	.60	42	1 299	.38	.73	.39	63	TRA	
F. L. Schmidt, Hunter und Caplan (1981)	Maschinenbediener Petroleumindustrie		HS interaktiv	.76	.70	.84	16	1 486	.18	.26	.01	39	TRA	
							13	821	.21	.30	.06	48	TRA	
F. L. Schmidt, Hunter und Pearlman (1981)	Bürotätigkeiten		HS interaktiv	-	.60	.59	144	10 564	.24	.50	o.A.	o.A.	o.A.	
N. Schmitt, Gooding, Noe und Kirsch (1984)	Sechs Berufsgruppen	Abwärtsbeurteilung Arbeitsproben	HS Barebone	-	-	-	25	3 597	.22	-	o.A.	o.A.	o.A.	
							3	1 793	.43	-				
Hunter und Hunter (1984), Reanalyse von Ghiselli (1973)	Manager Bürotätigkeiten Verkäufer Sicherheitskräfte Dienstleistungen Kraftfahrzeugmechaniker Handel und Handwerk Industrie Verkäufer (sales clerk)		HS interaktiv	-	.60	.67	o.A.	> 10 000	.27	.53	o.A.	o.A.	o.A.	
								> 10 000	.28	.54				
								5 000 – 9 999	.19	.61				
								1 000 – 4 999	.22	.42				
								1 000 – 4 999	.27	.48				
								1 000 – 4 999	.16	.28				
								> 10 000	.25	.46				
								> 10 000	.20	.37				
	.27													

Tabelle 2.5
(Fortsetzung)

Quelle	Berufsgruppe	Prädiktor/Kriterium ^a	Methode ^b	r_{xx}	r_{yy}	u_x	k	N	\bar{r}	$\hat{\rho}$	CV	VE	Gen.
Hunter und Hunter (1984); Hunter (1986)	Komplexitätslevel 1 <i>Set-up/Precision</i>	Abwärtsbeurteilung	HS interaktiv	–	.60	.67	17	1 114	o.A.	.56	o.A.	o.A.	o.A.
	Komplexitätslevel 2 <i>Synthesize/Coordinate</i>	Abwärtsbeurteilung					36	2 455	o.A.	.58			
	Komplexitätslevel 3 <i>Analyze/Compile</i>	Abwärtsbeurteilung					151	12 933	o.A.	.51			
	Komplexitätslevel 4 <i>Copy/Compare</i>	Abwärtsbeurteilung					201	14 403	o.A.	.40			
	Komplexitätslevel 5 <i>Feeding/Offbearing</i>	Abwärtsbeurteilung					20	1 219	o.A.	.23			
	Einstiegspositionen zivile Berufe	Abwärtsbeurteilung Arbeitsproben					425	32 124	.29	.53			
	militärische Berufe	Arbeitsproben					o.A.	1 790	o.A.	.75			
						o.A.	1 474	o.A.	.53				
Hirsh et al. (1986)	Strafverfolgung	Gedächtnis	HS interaktiv	.76	.73	.61	25	3 028	.05	.10	-.07	64	SIS
		numerisches Denken					8	1 188	.13	.26	.03	44	TRA
		schlussfolgerndes Denken					29	3 175	.08	.17	.08	90	CSC
		räumliches/ mechanisches Denken					29	3 536	.09	.17	.17	100	CSC
		verbales Denken					18	2 207	.09	.18	.00	63	TRA
Hartigan und Wigdor (1989)	Gesamt	Arbeitsleistung und Lernleistung	HS interaktiv	o.A.	.80	–	264	38 521	.20	.22	o.A.	o.A.	o.A.
	Komplexitätslevel 1 <i>Set-up/Precision</i>	Arbeitsleistung					o.A.	3 900	.15	.17			
	Komplexitätslevel 2 <i>Synthesize/Coordinate</i>	Arbeitsleistung					o.A.	630	.25	.28			
	Komplexitätslevel 3 <i>Analyze/Compile</i>	Arbeitsleistung					o.A.	19 206	.21	.23			
	Komplexitätslevel 4 <i>Copy/Compare</i>	Arbeitsleistung					o.A.	10 862	.18	.20			
	Komplexitätslevel 5 <i>Feeding/Offbearing</i>	Arbeitsleistung					o.A.	200	.19	.21			
Waldman und Avolio (1991)	zehn Berufsgruppen		mittleres r	–	–	–	1	21 547	.21	–	o.A.	o.A.	o.A.
Levine et al. (1996)	Handwerker der Versorgungswirtschaft		Raju et al. (1991)	.80	.76	.59	149	12 504	.25	.43	o.A.	62	o.A.

Tabelle 2.5
(Fortsetzung)

Quelle	Berufsgruppe	Prädiktor/Kriterium ^a	Methode ^b	r_{xx}	r_{yy}	u_x	k	N	\bar{r}	$\hat{\rho}$	CV	VE	Gen.
U. Funke, Krauß, Schuler und Stapf (1987)	Naturwissenschaftler/ Ingenieure in F&E-Einrichtungen	Abwärts-, Seitwärtsbeurteilung und Ergebniskriterien	HS interaktiv	–	.81	–	11	949	.14	.16	.07	76	CSC
Vinchur, Schippmann, Switzer und Roth (1998)	Verkäufer	GMA <i>g</i> -Faktor	HS interaktiv	–	.52	.60	25 22	1 770 1 231	.18 .23	.31 .40	.20 ^c .27 ^c	o.A.	o.A.
Barrett et al. (1999)	Feuerwehrleute		HS interaktiv (op. Val.)	–	–	–	24	2 791	.20	.42	-.03	o.A.	SIS
Kuncel et al. (2004)	akademische Berufe		HS interaktiv	.85	.66	.76	7	598	.26	.41	.21	o.A.	TRA
Hunter, Schmidt und Le (2006), Reanalyse von Hunter und Hunter (1984)	Komplexitätslevel 1 <i>Set-up/Precision</i> Komplexitätslevel 2 <i>Synthesize/Coordinate</i> Komplexitätslevel 3 <i>Analyze/Compile</i> Komplexitätslevel 4 <i>Copy/Compare</i> Komplexitätslevel 5 <i>Feeding/Offbearing</i>		HS interaktiv, indirekte VE-Korrektur	–	.60	.67	17 36 151 201 20	1 114 2 455 12 933 14 403 1 219	o.A.	.73 .74 .66 .56 .39	o.A.	o.A.	o.A.

Anmerkungen. Die Angaben sind für jede Quelle nur einmal gemacht, es sei denn sie ändern sich. o.A. = ohne Angabe; fehlen die Angaben ganz, sind sie nicht berichtet mit Ausnahme von: ^aPrädiktoren sind nur angegeben, wenn nicht GMA untersucht wurde, Kriterien nur wenn nicht Abwärtsbeurteilungen vorliegen. ^bHS = Hunter und Schmidt, op. Val. = operationale Validität. r_{xx} = Prädiktorreliabilität wenn korrigiert, r_{yy} = Kriterienreliabilität wenn korrigiert, u_x Varianzeinschränkung im Prädiktor wenn korrigiert, k = Anzahl der Studien, N = Stichprobengröße gesamt, \bar{r} = beobachtete Korrelation, $\hat{\rho}$ = korrigierte Korrelation, CV = untere Grenze des 90% Glaubwürdigkeits-Intervalls, bei ^cdes 80%-Intervalls, VE = Anteil der durch die Artefakte aufgeklärten Varianz an der beobachteten Varianz von \bar{r} bzw. $\hat{\rho}$ in %, Gen. = Generalisierung: TRA = transportability, CSC = cross-situational consistency. Die Zusammenstellung der Tabelle erfolgte auf Basis von Bobko et al. (1999) und Campion in Morgeson et al. (2007).

Befunde von Salgado und Anderson (2003) für Deutschland ($\hat{\rho} = .68$), die allerdings nur auf wenigen Studien ($k = 8$) basieren. Die Ergebnisse aus der Metaanalyse von Hülsheger et al. (2006) entsprechen mit .53 ziemlich exakt den amerikanischen Befunden, basieren aber ebenfalls auf wenigen Studien ($k = 9$).

Die Richtung und Höhe der prädiktiven Validität kann im Falle der europaweiten Ergebnisse generalisiert werden, da die untere Grenze des Glaubwürdigkeitsintervalls (CV) über Null liegt und mehr als 75% der Varianz aufgeklärt werden können (*cross situational consistency*, CSC; Kemery et al., 1987). Dies gilt nicht für die Metaanalysen, die ausschließlich Studien aus Deutschland berücksichtigen. Hier liegt – wie bei den Ergebnissen zur Lernleistung – *transportability* (TRA) vor, d.h. die Validität ist abhängig von Moderatoren, z.B. der Berufserfahrung oder der Berufskomplexität.

Betrachten wir zunächst den Moderator *Berufserfahrung*. Er wird mit der Dauer der Berufszugehörigkeit operationalisiert. Vineberg und Taylor (1972, nach F. L. Schmidt, Hunter & Outerbridge, 1986) zeigen anhand militärischer Stichproben, dass Erfahrung von null bis fünf Jahren linear mit dem erworbenen Fachwissen und der Leistung in Arbeitsproben ansteigt, bei über fünf Jahren aber keine weitere Zunahme mehr nachzuweisen ist. Colquitt et al. (2000) zeigen, dass Informationsverarbeitungskapazität besonders in frühen Lernstadien wichtig ist, bei denen noch nicht auf Erfahrung zurückgegriffen werden kann. Andererseits berichtet McDaniel (1985; nach Hunter & Schmidt, 1996), dass a) die Bedeutung von Intelligenz für die Arbeitsleistung mit zunehmender Erfahrung der Arbeitnehmer nicht ab-, sondern zunimmt und b) die Bedeutung von Erfahrungsunterschieden in Studien mit zunehmender durchschnittlicher Erfahrung der Arbeitnehmer nicht zu-, sondern abnimmt.

Das Modell zum Einfluss von Intelligenz auf Lern- und Arbeitsleistung bleibt stabil, wenn die Berufserfahrung als zusätzlicher Prädiktor mitbetrachtet wird. Dies zeigen F. L. Schmidt et al. (1986) anhand von vier militärischen Studien (Abbildung 2.3). Die korrigierten Korrelationen betragen zwischen GMA und Berufserfahrung $\hat{\rho} = .00$, GMA und Fachwissen $\hat{\rho} = .46$, GMA und objektiver Leistung in Arbeitsproben $\hat{\rho} = .38$ sowie GMA und subjektiven Abwärtsbeurteilungen $\hat{\rho} = .16$. Bei Auspartialisierung der Berufserfahrung sind die Zusammenhänge deutlicher: Intelligenz korreliert dann korrigiert mit $\hat{\rho} = .56$ mit Fachwissen, $\hat{\rho} = .46$ mit Arbeitsproben und $\hat{\rho} = .17$ mit Abwärtsbeurteilungen.

In Kapitel 2.3.1 wurde der Einfluss der *Berufskomplexität* auf die Validität von Intelligenz in Bezug auf Lernleistung gezeigt. Für das Kriterium Arbeitsleistung ist der gleiche Zusammenhang zu vermuten: mit der Berufskomplexität nimmt die Bedeutung von Intelligenz für die erfolgreiche Bewältigung der Arbeitsaufgaben zu. Gutenberg, Arvey, Osborn und Jeanneret (1983) beschreiben die Bedeutung der Anforderungsdimensionen *Entscheidungsfindung* und *Informationsverarbeitung* als Moderatoren. Sie spielen vor allem in komplexen Berufen eine wichtige Rolle. Je stärker sie zur erfolgreichen Arbeitsbewältigung erforderlich sind, desto höher ist die Validität von Intelligenztests. Die prädiktive Validität von GMA sollte entsprechend steigen. Dies gilt zumindest für korrigierte Korrelationen, bei beobachteten ist die Auswirkung unterschiedlicher Varianzeinschränkungen zu berücksichtigen.

Tabelle 2.6

Metaanalytische Befunde aus Europa zur prädiktiven Validität von Intelligenz in Bezug auf Arbeitsleistung

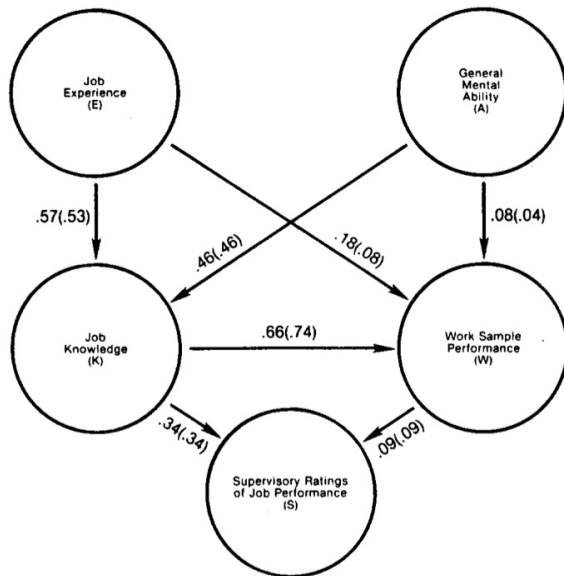
Quelle	Berufskomplexität ^a	Land	Methode ^b	r_{xx}	r_{yy}	u_x	k	N	\bar{r}	$\hat{\rho}$	CV	VE	Gen.
Salgado und Anderson (2002)		UK	HS interaktiv (op. Val.)	.85	.52	.61	45	7 283	.18	.41	.11	46	TRA
		E		.85	.52	.80	9	1 239	.36	.61	.61	100	CSC
Salgado et al. (2003a)		Europa ^c	HS interaktiv (op. Val.)	.83	.52	.62	93	9 554	.29	.62	.37	75	CSC
Salgado et al. (2003b)	hohe Komplexität mittlere Komplexität geringe Komplexität	Europa ^c	HS interaktiv (op. Val.)	.83	.52	.47	14	1 604	.23	.64	.33	66	TRA
						.69	43	4 744	.27	.53	.21	43	TRA
						.67	12	864	.25	.51	.38	87	CSC
Salgado und Anderson (2003)		F	HS interaktiv (op. Val.)	.83	.52	–	26	1 445	.48	.64	.48	62	TRA
		B, NL				.50	15	1 075	.24	.63	.20	55	TRA
		E				.74	11	1 182	.35	.64	.45	71	TRA
		UK				.63	68	7 725	.26	.56	.46	92	CSC
		D				.68	8	701	.35	.68	.35	49	TRA
Bertua et al. (2005)		UK	HS interaktiv (op. Val.)	.85	.52	.60	12	2 469	.22	.48	.17	45	TRA
Hülshager et al. (2007)		D	Raju et al. (1991) (op. Val.)	.92	.52	.85	9	746	.33	.53	.30	34	TRA

Anmerkungen. Die Angaben sind für jede Quelle nur einmal gemacht, es sei denn sie ändern sich. ^aDie Berufskomplexität ist nur angegeben, wenn Berufe unterschiedlicher Komplexität getrennt berücksichtigt werden. ^bHS = Hunter und Schmidt, op. Val = operationale Validität. ^cBelgien, Deutschland, Frankreich, Großbritannien, Irland, Niederlande, Portugal, skandinavische Länder, Spanien. r_{xx} = Prädiktorreliabilität wenn korrigiert, r_{yy} = Kriterienreliabilität wenn korrigiert, u_x Varianzeinschränkung im Prädiktor wenn korrigiert, k = Anzahl der Studien, N = Stichprobengröße gesamt, \bar{r} = beobachtete Korrelation, $\hat{\rho}$ = korrigierte Korrelation, CV = untere Grenze des 90% Glaubwürdigkeits-Intervalls, VE = Anteil der durch die Artefakte aufgeklärten Varianz an der beobachteten Varianz von \bar{r} bzw. $\hat{\rho}$ in %, Gen. = Generalisierung; TRA = transportability, CSC = cross-situational consistency.

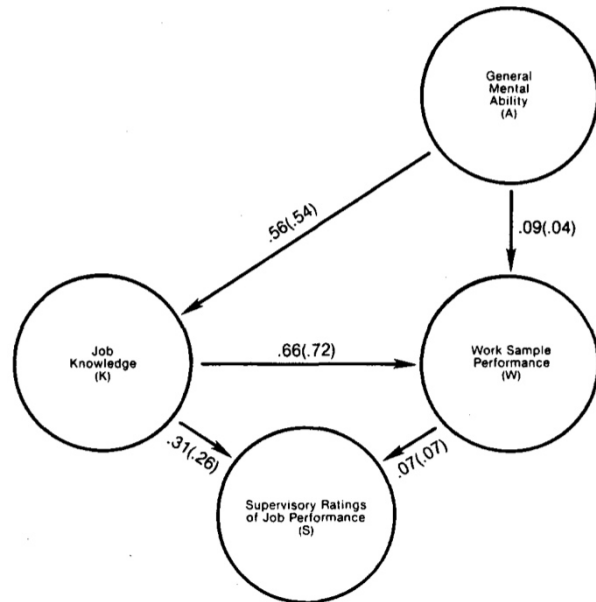
Abbildung 2.3

Pfadmodell zum Zusammenhang zwischen Intelligenz und Arbeitsleistung

a) mit Berufserfahrung



b) bei kontrollierter Berufserfahrung



Anmerkungen: Pfad-Koeffizienten von gruppierten und ungruppierten Daten (ungruppierte Daten sind in Klammern angegeben). Gruppierung nach Berufserfahrung: 1 = 1 bis 4 Monate, 2 = 5 bis 11 Monate, 3 = 12 bis 16 Monate, 4 = 17 bis 22 Monate, 5 = 23 bis 60 Monate, 6 > 61 Monate. Quelle: F. L. Schmidt et al., 1986, S. 437-438.

Hunter (1986) zeigt den vermuteten Zusammenhang empirisch: Die Validität von Intelligenz in Bezug auf Arbeitsleistung nimmt mit zunehmender Komplexität ebenfalls zu (vgl. Tabelle 2.5). Bei Hartigan und Wigdor (1989) ist der Befund uneinheitlicher: während die Validität vom vierten zum zweiten Komplexitätsniveau stetig ansteigt (.20 bis .28) weist das höchste Komplexitätsniveau die geringste Validität auf (.17). Den komplexitätsinduzierten Anstieg konnten Salgado et al. (2003a) auch für europäische Studien replizieren ($\hat{\rho}$ für gering/mittel/hoch komplexe Berufe = .51/.53/.64) sowie Bertua et al. (2005) für Studien in Großbritannien ($\hat{\rho}$ = .37 für Fahrer, .55 für Verkäufer und .69 für Manager; vgl. Tabelle 2.6). Es ließ sich also mehrfach bestätigen, was Hunter (1986) in Bezug auf die Validität von Intelligenz, gemessen mit der General Aptitude Test Battery (GATB), formulierte:

The principal finding for the prediction of job proficiency was that although the validity of the GATB tests varies across job families, it never approaches zero. The validity of cognitive ability as a predictor decreases as job complexity decreases. (S. 81)

Der Einfluss unterschiedlicher Varianzeinschränkungen lässt sich in den amerikanischen Studien nicht abschätzen, da diese gar nicht (Hartigan & Wigdor, 1989) oder mit

den gleichen Werten (Hunter & Hunter, 1984) korrigiert wurden. Die Varianzeinschränkungen für gering-, mittel- bzw. hoch-komplexe Tätigkeiten in der Studie aus Europa betragen $u = .67, .69, .47$ (Salgado et al., 2003b). Durch die deutlich höhere Varianzeinschränkung bei hoher Komplexität erreicht die korrigierte Validität in dieser Gruppe mit $\hat{\rho} = .64$ den höchsten Wert, obwohl die beobachtete Validität mit $\bar{r} = .23$ unter der der beiden geringeren Komplexitätsgruppen liegt.

2.3.3. Intelligenz und Karriereerfolg

Die Gravitationshypothese von McCormick, DeNisi und Shaw (1979) besagt, dass sich Personen im Laufe der Zeit den Tätigkeiten zuwenden, die ihren Fähigkeiten entsprechen. Intelligente Personen fühlen sich bei einfachen Tätigkeiten unterfordert und suchen anspruchsvollere, weniger intelligente Personen fühlen sich bei kognitiv anspruchsvollen Aufgaben überfordert und suchen einfachere. Diese Gravitationshypothese hat sich im Hinblick auf Intelligenz bewährt (Wilk, Desmarais & Sackett, 1995): mit der Zeit und zunehmender Berufserfahrung sortieren sich Menschen durch Selbst- und Fremdselektion den Berufen zu, die zu ihren kognitiven Fähigkeiten passen.

Für diese Mobilität sind aber nicht nur die Fähigkeiten und Leistungen der einzelnen Personen relevant (*contest-mobility*), sondern auch die Unterstützung, die ihnen zuteil wird (*sponsored-mobility*, Ng et al., 2005). Ein Zusammenhang von Intelligenz mit Karriereerfolg ist für beide Mobilitäten plausibel: Zum einen sind intelligente Personen eher in der Lage, erforderliche Fertigkeiten zu erlernen, sich notwendiges Wissen anzueignen und dadurch eine höhere Arbeitsleistung zu erbringen (*contest-mobility*), zum anderen werden sie eher unterstützt (*sponsored-mobility*).

Dies sollte insgesamt bedeuten, dass intelligente Personen schneller und höher aufsteigen, als weniger intelligente. Mit den Beförderungen und durch die Wahl komplexerer Berufe sollte auch ein höheres Einkommen verbunden sein. Drei Metaanalysen berichten Validitäten dazu (vgl. Tabelle 2.7). Die Daten stammen aus den USA. Beobachtete Korrelationen von GMA mit Kriterien beruflichen Vorankommens berichten N. Schmitt et al. (1984): $\bar{r} = .44$ mit *achievement/grades* und $.28$ mit *status change*. Die Anzahl der Studien ist mit fünf und neun gering. Umsatzerlöse bzw. das Einkommen als Kriterium korreliert bei N. Schmitt et al. (1984) zu $\bar{r} = .14$, bei Vinchur et al. (1998) zu $\bar{r} = .02$, wenn der g -Faktor betrachtet wird. Bei Summenscores beträgt die Korrelation $\bar{r} = -.02$. Die korrigierten Werte lauten $\hat{\rho} = .04$ für g und $\hat{\rho} = -.03$ für Summenscores. Ng et al. (2005) berichten eine korrigierte Korrelation zwischen GMA und Einkommen von $\hat{\rho} = .27$. Auch die genannten Analysen zum Einkommen basieren auf vergleichsweise wenig Studien (zwischen acht und 18).

Die Einschätzung der Generalisierung ist nur für die Ergebnisse möglich, die Vinchur et al. (1998) zum Einkommen berichten: selbst das vergleichsweise liberale Glaubwürdigkeitsintervall von 80% enthält Null, es ist deshalb nicht möglich die Richtung des Zusammenhangs zu generalisieren (*situational specificity*, SIS; Kemery et al., 1987).

Tabelle 2.7

Metaanalytische Befunde zur prädiktiven Validität von Intelligenz in Bezug auf Karriereerfolg

Quelle	Berufsgruppe	Prädiktor	Kriterium	Methode ^a	r_{xx}	r_{yy}	u_x	k	N	\bar{r}	$\hat{\rho}$	CV	Gen.
N. Schmitt et al. (1984)	verschiedene	GMA	Umsatzerlöse Status change	HS Barebone	–	–	–	8	12 449	.14	–	o.A.	o.A.
					–	–	–	9	21 190	.28	–		
Vinchur et al. (1998)	Verkäufer	GMA <i>g</i>	Einkommen	HS interaktiv	–	–	.60	18	1 876	-.02	-.03	-.21	SIS
					–	–	–	12	1 310	.02	.04	-.12	SIS
Ng et al. (2005)	verschiedene	GMA	Einkommen	HS interaktiv	o.A.	–	–	8	9 560	o.A.	.27	o.A.	o.A.

Anmerkungen. Die Angaben sind für jede Quelle nur einmal gemacht, es sei denn sie ändern sich. o.A. = ohne Angabe. ^aHS = Hunter und Schmidt. r_{xx} = Prädiktorreliabilität wenn korrigiert, r_{yy} = Kriterienreliabilität wenn korrigiert, u_x Varianzeinschränkung im Prädiktor wenn korrigiert, k = Anzahl der Studien, N = Stichprobengröße gesamt, \bar{r} = beobachtete Korrelation, $\hat{\rho}$ = korrigierte Korrelation, CV = untere Grenze des 80% Glaubwürdigkeits-Intervalls, VE = Anteil der durch die Artefakte aufgeklärten Varianz an der beobachteten Varianz von \bar{r} bzw. $\hat{\rho}$ in %, Gen. = Generalisierung: SIS = Situationspezifität.

2.3.4. Warum eine weitere Metaanalyse?

Aus welchen Gründen ist es sinnvoll, eine weitere Metaanalyse mit deutschsprachigen Studien durchzuführen? Das vorangehende Kapitel hat gezeigt, dass die Befunde zur Validität in Bezug auf Lernleistung uneinheitlich sind und in Bezug auf Karriereerfolg noch keine metaanalytischen Ergebnisse für Studien aus Deutschland vorliegen.

Die Befunde von [Salgado und Anderson \(2003\)](#) basieren auf wenigen Studien. Dies gilt auch für die Ergebnisse zur Arbeitsleistung von [Hülsheger et al. \(2007\)](#). Die Datenbasis der vorliegenden Studie ist größer, was zum einen daran liegt, dass seit der Veröffentlichung der bereits durchgeführten Metaanalysen neue Studien erschienen sind. In der vorliegenden Arbeit werden Studien bis zum Jahr 2006 berücksichtigt. Bedeutender ist jedoch, dass andere Inklusions- und Exklusionskriterien Verwendung finden. Darauf wird in Kapitel 3.2 näher eingegangen. Vorläufig soll ein Hinweis auf die zwei wesentlichen Unterschiede genügen: Erstens werden Studien berücksichtigt, die GMA im Rahmen von Testbatterien erheben, die auch Skalen zur Erhebung nicht kognitiver Fähigkeiten vorsehen. Dies ist der Fall bei deutschen Adaptationen der General Aptitude Test Battery (GATB; [U. S. Employment Service, 1970](#)). Die metaanalytischen Befunde aus den USA basieren zu einem großen Teil auf GATB-Resultaten: „Most of the earlier discussions have drawn on studies that use batteries like the GATB and the Armed Services Vocational Aptitude Battery.“ ([Levine et al., 1996](#), S. 2). Zweitens werden Studien inkludiert, die insignifikante Korrelationen nicht berichten. Diese wurden – wie es üblich ist – in den bereits vorliegenden Metaanalysen exkludiert. Dabei besteht jedoch die Gefahr der Überschätzung der Validität. Um die betroffenen Studien inkludieren zu können, wird eine alternative Vorgehensweise entwickelt (Kapitel 3.7.2).

Diese Entscheidungen haben dazu geführt, dass für berufsbezogene Lernleistung 210 Studien ($N = 30\,451$) berücksichtigt werden können, statt bisher 90 Studien ($N = 11\,969$), für Arbeitsleistung nunmehr 18 Studien ($N = 2\,739$, statt bisher $k = 9$, $N = 746$) und für Karriereerfolg 10 Studien ($N = 2\,260$, bisher nicht für Deutschland untersucht). Die Inklusion von deutlich mehr Studien rechtfertigt an und für sich ein Update.

Hinzu kommt die Weiterentwicklung der metaanalytischen Methode, die darin besteht, dass Varianzeinschränkungen nicht nur *direkt*, sondern auch *indirekt* korrigiert werden können. Wie in Kapitel 3.6.3 begründet wird, ist die indirekte Korrektur im vorliegenden Gegenstandsbereich angemessen. Die bisherigen Ergebnisse aus Deutschland basieren jedoch auf *direkten* Korrekturen der Varianzeinschränkung. Das kann zu einer bedeutsamen Unterschätzung von ρ führen: [Hunter et al. \(2006\)](#) zeigen, dass die Unterschätzung bei der Analyse von [Hunter und Hunter \(1984\)](#) bei 25% lag. Deshalb empfehlen [F. L. Schmidt und Raju \(2007\)](#) eine Neuberechnung von Ergebnissen, die auf direkter Varianzeinschränkungs-Korrektur basieren. Schließlich werden auch die anderen Korrekturmaße etwas anders gewählt als bei [Salgado et al.](#) und [Hülsheger et al.](#) Die Begründung, die dieser Vorgehensweise zu Grunde liegt, und die Unterschiede zu den Verfahren in den Vorläufer-Studien werden im Kapitel 3.6 dargestellt und ihre Auswirkungen in Kapitel 5 diskutiert. Mit dem erweiterten Datensatz und der in Kapitel 3 beschriebenen Methode sollen folgende Forschungsfragen für Studien aus Deutschland beantwortet werden:

1. Sind Intelligenztests kriterienvalid in Bezug auf berufliche Lernleistung, spezifische Arbeitsleistung und Karriereerfolg?
2. Moderieren unterschiedliche Erfolgsindikatoren pro Erfolgsbereich die Validität?
3. Nimmt die Validität der Intelligenztests mit zunehmender Berufskomplexität zu?
4. Moderieren die Berufsgruppe oder der verwendete Intelligenztest die Validität?

3. Methode

Die Durchführung einer Metaanalyse läuft in mehreren Schritten ab (Hunter & Schmidt, 2004): die Literaturrecherche, die Studienin- bzw Exklusion, die Kodierung, die Prüfung der statistischen Unabhängigkeit der einzelnen Studien, die Zusammenfassung voneinander abhängiger Effektgrößen zu einem Wert, die Klärung der Fragen, ob und ggf. wie Artefakte zu korrigieren sind und wie mit fehlenden Primärstudien-Daten umzugehen ist. Nach diesen vorbereitenden Schritten erfolgt die eigentliche metaanalytische Zusammenfassung der einzelnen Studienergebnisse, die mit einer Prüfung der Studienhomogenität einhergeht. Sensitivitäts- und Moderatoranalysen folgen und helfen dabei, mögliche Erklärungen für die metaanalytischen Ergebnisse zu finden.

Diese Schritte bauen aufeinander auf, können aber iterativ mehrfach durchlaufen werden (Rustenbach, 2003). Weil nachfolgende Schritte auf den Ergebnissen der vorhergehenden aufbauen, werden in diesem Kapitel für jeden vorbereitenden Schritt nicht nur die Vorgehensweisen dargestellt und begründet, sondern auch die Ergebnisse berichtet. Die metaanalytischen Ergebnisse selbst werden in Kapitel 4 dargestellt.

3.1. Literaturrecherche

Um dem Ideal der Vollerhebung aller relevanten Studien möglichst nahe zu kommen, wurden folgende acht Suchstrategien angewendet: 1. Zunächst wurde eine Recherche in den Psyn dex-Abstracts (1977-2006) nach Kombinationen von Prädiktor- und Kriterien-Stichworten durchgeführt. Tabelle 3.1 gibt einen Überblick darüber. Die Stichworte aus den Zeilen der Tabelle wurden zu sinnvollen Kombinationen zusammengefügt, z.B. zu **intelligenz* AND *beruf* AND *eignung**. Dabei symbolisieren die Asteriske, dass auch nach Worten gesucht wurde, die sich aus den Begriffen und weiteren Wortbestandteilen zusammensetzen. Eine vergleichbare Suche in PsychInfo nach entsprechenden englischen Kriterien-Stichworten erbrachte keine neuen Funde mit deutschen Stichproben. Deshalb wurden die folgenden Recherchen auf Psyn dex-Datenbanken beschränkt. 2. Die Psyn dex-Recherche wurde auf Autoren erweitert, die mehrfach zum Thema veröffentlicht haben (Althoff, Funke, Kersting, Marschner, Thielepape, Schmidt-Atzert, Schuler) und 3. auf Schlagworte, denen die gefundenen Artikel häufig zugeordnet waren (*intelligenz**, *kognitive**, *beruf**, *personal**, *praediktive-validitaet*). 4. Berücksichtigt wurden die Studien an deutschen Stichproben, die die Arbeitsgruppen um Salgado und Hülsheger in ihren Metaanalysen inkludierten (Hülsheger et al., 2006, 2007; Salgado & Anderson, 2003; Salgado et al., 2003a, 2003b). 5. Die Suche nach Testvalidierungen für Intelligenztests wurde auf zwei Arten durchgeführt: zunächst wurden die Testmanuale in den Testarchiven des Instituts für Psychologie der Universität Bonn und der zugehörigen Abteilung für Arbeits-, Organisations- und Wirtschaftspsychologie durchgesehen, sodann 6. in den Psyn dex-Abstracts nach Veröffentlichungen gesucht, die einschlägige Intelligenztests

Tabelle 3.1

Begriffe zur Suche nach geeigneten Studien

Prädiktoren	Fähigkeit, Intellekt, intellektuell, Intelligenz, IQ, kognitiv, mental
Kriterien	Arbeitsleistung, Beförderung, berufliche Entwicklung, beruflicher Erfolg, Beurteilung, Bewertung, Eignung, Einkommen, Gehalt, Karriere, Laufbahn, Lohn, Platzierung, Produktivität, Qualifikation, Tauglichkeit
weitere	Ausbildung, Beruf, Bewerber, Personal, Umschulung, (prognostische) Validität, Weiterbildung

verwendet haben. Dazu wurden die Überblickswerke von [Brickenkamp \(2002b\)](#), [Holling und Liepmann \(2004\)](#) sowie [Sarges und Wottawa \(2004\)](#) herangezogen. 7. Die Mitglieder folgender Institutionen wurden per E-Mail angeschrieben und um veröffentlichte und nicht veröffentlichte Primärstudien gebeten: die arbeits- und organisationspsychologischen Fachgruppen der deutschen psychologischen Berufsverbände (Berufsverband Deutscher Psychologinnen und Psychologen [BDP], Deutsche Gesellschaft für Psychologie [DGPs]), die Deutsche Gesellschaft für Personalführung (DGFP), die Deutsche Gesellschaft für Personalwesen (DGP), die Deutsche Gesellschaft für Luft- und Raumfahrt (DLRA), die Sparkassenakademie, die Gruppe Wehrpsychologie des Streitkräfteamtes und das Zentrum für Psychologische Information und Dokumentation (ZPID). 8. Die Literaturverzeichnisse aller geeigneten Artikel wurden nach Hinweisen auf weitere Primärstudien durchsucht.

Bei Hinweisen darauf, dass eine Quelle relevante Ergebnisse enthalten könnte, wurde diese nach statistischen Kennwerten durchsucht, die etwas über den Zusammenhang zwischen Intelligenz und den in Kapitel 2.2.2 beschriebenen Erfolgsindikatoren aussagen. Anhand verschiedener Inklusions- und Exklusionskriterien wurde entschieden, ob der Kennwert und somit die Studie berücksichtigt werden konnte.

3.2. Studieninklusion und -exklusion

Damit eine Studie berücksichtigt wurde, musste sie verschiedenen Inklusionskriterien genügen. Diese betreffen a) die Intelligenzmaße, b) die Erfolgskriterien, c) die untersuchte Stichprobe und den Zeitabstand zwischen Intelligenz- und Kriterienerhebung, d) die statistischen Zusammenhangsmaße sowie e) die Studienqualität.

3.2.1. Allgemeine Intelligenz

Inklusions- und Exklusionsregeln

Der Fokus dieser Arbeit liegt auf der höchsten Hierarchieebene des BIS-Modells: der Allgemeinen Intelligenz. In einigen Studien werden Ergebnisse berichtet, die untergeordneten Hierarchiestufen zuzuordnen sind: der Ebene einzelner Fähigkeiten bzw. Faktoren oder der Ebene einzelner Aufgabengruppen. Sie wurden dann inkludiert und zu einem

Gesamtergebnis für Allgemeine Intelligenz aggregiert (vgl. Kapitel 3.5), wenn keine oder weniger aussagekräftige Ergebnisse auf höherer Hierarchieebene berichtet wurden. Unterschiedlich aussagekräftige Hierarchieebenen können bei Studien vorliegen, die insignifikante Ergebnisse nicht berichten. Das Aggregat aus den einzelnen Fähigkeiten muss als *relevanter* Indikator für GMA gelten können und nicht lediglich eine oder wenige spezifische kognitive Fähigkeiten abdecken.

Um zu entscheiden, ob die erfassten Fähigkeiten Indikatoren für Allgemeine Intelligenz darstellen, wurden drei Regeln angewendet, die bis auf wenige Ausnahmen als Entscheidungsgrundlage ausreichen. Zunächst werden diese Regeln vorgestellt, anschließend wird auf Besonderheiten eingegangen.

Regel 1: Es wurden alle Testverfahren inkludiert, die als Operationalisierungen für Allgemeine Intelligenz anerkannt sind. Dazu wurde zum einen betrachtet, ob die Verfahren mit dem Ziel entwickelt wurden, Intelligenz zu erfassen, und welche Theorie von Intelligenz ihnen zu Grunde liegt. Zum anderen wurden die Testrezensionen in folgenden einschlägigen Überblickswerken herangezogen: Einen umfassenden Überblick über verwendete Verfahren der Personalauswahl bei Privatunternehmen, der ehemaligen Bundesanstalt für Arbeit und der Bundeswehr gibt [Brambring \(1983\)](#); jüngere Übersichten, die auch neuere Verfahren berücksichtigen, geben [Brickenkamp \(2002b\)](#), [Holling et al. \(2004\)](#) sowie [Kanning \(2002\)](#). Eine Zusammenstellung von Verfahren, die in der Wirtschaftspsychologie häufig eingesetzt werden, ist in [Sarges und Wottawa \(2004\)](#) zu finden. Eine Übersicht über die mit dieser Regel inkludierten Testverfahren gibt Tabelle 3.2.

Regel 2: Bei Testverfahren, die sich nicht eindeutig als Intelligenztests auszeichnen, wurde geprüft, ob sie sich in das BIS-Modell einordnen lassen. Dazu wurden sämtliche potenziellen Intelligenzaufgaben von zwei Kodierern unabhängig voneinander den BIS-Zellen zugeordnet. Dies erfolgte durch Inspektion der Aufgaben und den Vergleich mit den Beschreibungen der BIS-Dimensionen ([Jäger et al., 1997](#)) nach Maßgabe inhaltlicher Plausibilität. Nähere Angaben zur BIS-Kodierung können Kapitel 3.4 entnommen werden.

Für die *Inklusion* sollten idealerweise alle zwölf BIS-Structupel abgedeckt sein, mindestens aber die wichtigste Operationsklasse *Verarbeitungskapazität* ([Jäger et al., 1997](#)) und mehrere Inhaltsklassen (z.B. die numerische und verbale). Nicht inkludiert wurden Studien, die ausschließlich aus einzelnen oder wenigen Facetten bestehen und Verarbeitungskapazität nicht berücksichtigen. Ein anderes Vorgehen würde zu stark defizienten „Intelligenz“-Maßen führen. Die Berücksichtigung einzelner Intelligenzsubfacetten würde zudem bedeuten, dass die Suche nach Primärstudien auch auf Quellen ausgeweitet werden müsste, die Verfahren berichten, die einzelnen BIS-Zellen zugeordnet werden können, z.B. auf Rechenleistungstests, Konzentrationstests, etc.

Eine Ausnahme davon bilden Matrizen-tests, die gemäß der ersten Inklusionsregel (anerkannt als Intelligenzverfahren) akzeptiert wurden, obwohl sie nur aus Items bestehen, die der FK-Facette des BIS-Modells zugeordnet werden können ([Süß, 2001](#)). Dies widerspricht Regel 2. Mit einer Sensitivitätsanalyse wird geprüft, ob sie sich in Bezug auf ihre

Tabelle 3.2
Anerkannte Intelligenz-Verfahren

Testverfahren	Quelle (jüngste inkludierte Auflage)	Konstruktbezeichnung	Intelligenztheorie	Nennung in Überblickswerken
Intelligenz-Struktur-Test (IST)	Amthauer (1973)	Intelligenz-Niveau	Thurstone	Brambring (1983); Brickenkamp (2002b)
Intelligenz-Struktur-Test 2000 R (IST 2000 R)	Amthauer et al. (2001)	schlussfolgerndes Denken	Hierarchisches Protomodell der Intelligenzstrukturforschung; Cattell	Brickenkamp (2002b); Holling et al. (2004); Sarges und Wottawa (2004)
Berliner Intelligenzstruktur-Test (BIS)	Jäger et al. (1997)	Allgemeine Intelligenz, g_{BIS}	BIS-Modell	Brickenkamp (2002b); Holling et al. (2004); Sarges und Wottawa (2004)
Berliner Intelligenzstruktur-Test der Deutschen Gesellschaft für Personalwesen e.V. (BIS-r-DGP)	Deutsche Gesellschaft für Personalwesen e.V. (1996)	Allgemeine Intelligenz, g_{BIS}	BIS-Modell	Sarges und Wottawa (2004)
Wilde-Intelligenz-Test (WIT)	Jäger und Althoff (1994)	Allgemeine Intelligenz	Thurstone	Brickenkamp (2002b)
Grundintelligenztest Skala 3 (CFT 3)	Weiß (1980)	general ability, general mental ability, g	Spearman, Thurstone	Brambring (1983); Brickenkamp (2002b); Holling et al. (2004)
Prüfsystem für Schul- und Bildungsberatung (PSB)	Horn (1969)	Intelligenz-Niveau	Thurstone	Brambring (1983); Brickenkamp (2002b)
Leistungsprüfsystem (LPS)	Horn (1983)	g	Thurstone, Pyramidenmodell von Horn	Brambring (1983); Brickenkamp (2002b)
Figure Reasoning Test (FRT)	Daniels (1962a)	non-verbal intelligence		Brambring (1983); Brickenkamp (2002b)
Raven's Standard Progressive Matrices (SPM)	Raven et al. (2000)	g_f	Spearman, Cattell	Brambring (1983); Brickenkamp (2002b); Holling et al. (2004)
Raven's Standard Progressive Matrices plus (SPMPLS)	Raven et al. (2003)	g_f	Spearman, Cattell	Sarges und Wottawa (2004)
Berufseignungstest (BET)	Schmale und Schmidtke (1995b, 2001)	Allgemeine Intelligenz	GATB	Brambring (1983); Sarges und Wottawa (2004)
Testserie für Hauptschüler (EUB)	Bundesanstalt für Arbeit (1982)	intellektuelles Leistungsniveau	GATB	Brambring (1983)
AZUBI-Tests	Schuler und Klingner (2005); Görlich und Schuler (2007)	berufsbezogene Intelligenz	BIS-Modell	Sarges und Wottawa (2004, nur BK-Version)

prädiktive Validität von den anderen Intelligenztests unterscheiden (Kapitel 4.1.3).

Bei Ergebnissen auf Skalenebene führte die Verwendung des BIS-Modells als Rahmenmodell zur *Exklusion* von Konzentrationstests (z.B. des d2; [Brickenkamp, 2002a](#)), Aufgaben der Handgeschicklichkeit (z.B. der Drahtbiegeprobe, [Lienert, 1967](#)), des technischen oder mechanischen Verständnisses (z.B. des *Test of mechanical comprehension* von [Bennett & Fry, 1969](#) oder des *Mechanisch-technischen Verständnistests* von [Lienert, 1958](#)) und Wissensaufgaben, die z.B. im Erweiterungsmodul des IST 2000 R ([Amthauer et al., 2001](#)) zur Operationalisierung von kristalliner Intelligenz eingesetzt werden.

Regel 3: Auf Gesamt- oder Faktorebene wurden Kontaminationen mit Aufgabentypen, die nicht vom BIS-Modell abgedeckt werden, in begrenztem Maße hingenommen, wenn die Kontamination ein geringes Gewicht hat oder hohe Korrelationen mit anerkannten Verfahren nachgewiesen werden können. Als hinreichend hoch wird eine Korrelation von über .70 betrachtet. Ab .50 liegt nach [Cohen \(1992\)](#) ein großer Zusammenhang vor; verschiedene Intelligenztests korrelieren im Durchschnitt mit .77 ([Jensen, 1984](#)).

Obwohl Kontaminationen vorhanden sind, wurden folgende Intelligenztests *inkludiert*: Die Arbeitsproben zur berufsbezogenen Intelligenz für Büro- und kaufmännische Tätigkeiten (AZUBI-BK, [Schuler & Klingner, 2005](#)) sowie für technische und handwerkliche Tätigkeiten (AZUBI-TH, [Görlich & Schuler, 2007](#)). Sie enthalten Diktataufgaben, die Wissen über die Rechtschreibung abfragen und Postmodule, die auch Gewissenhaftigkeit erfassen. Die Intelligenzmaße beider Testverfahren korrelieren jedoch hoch mit denen anderer Intelligenztests: AZUBI-BK und IST 70: $r = .73$ ([Schuler & Klingner, 2002](#)), AZUBI-TH und IST 70 bzw. WIT: $r = .78$ bzw. $.83$ ([Görlich & Schuler, 2004](#)).

[Neumann \(1938\)](#) verwendet zur Einschätzung der kognitiven Fähigkeiten drei Aufgabengruppen, die er als Intelligenz, Aufmerksamkeit und Gedächtnis bezeichnet. Die Intelligenzgruppe besteht aus Kopfrechen-Aufgaben und der Beurteilung technischer Skizzen. Die Beurteilung technischer Skizzen konnte dem BIS-Modell nicht zugeordnet werden. Das Gewicht der Kontamination kann nicht genau bestimmt werden, beträgt aber maximal 16.7%, da die Bewertung der Skizzen „lediglich als Stütze für das bereits durch die Rechenproben gewonnene Leistungsbild des Prüflings [diente], besonders in dortigen Versagensfällen“ ([Neumann, 1938, S. 121](#)).

[Roloff \(1928\)](#) operationalisiert die kognitive Fähigkeit „geistige (intellektuelle) Betätigung“ (S. 44) durch drei Dimensionen, die jeweils aus verschiedenen Aufgabengruppen bestehen. Die Dimensionen *Arbeiten mit räumlichen Vorstellungen* und *Sprachliche Intelligenz* lassen sich dem BIS-Modell zuordnen, die Dimension *Technisches Verständnis* hingegen nicht. Das Gewicht dieser Kontamination beträgt 40% ([Roloff, 1928](#)). Die Auswirkungen der Kontaminationen der Studien von [Neumann \(1938\)](#) und [Roloff \(1928\)](#) werden sensitivitätsanalytisch geprüft (Kapitel 4.1.3).

Besonderheiten: Die unterschiedlichen Modellvorstellungen von Intelligenz, die den verschiedenen Testverfahren zu Grunde liegen, wirken sich auch in der Operationalisierung aus. Im Berufseignungstest (BET; [Schmale & Schmidtke, 1995b](#)) und der Testserie für Hauptschüler (EUB; [Bundesanstalt für Arbeit, 1982](#)), die nach der GATB entwickelt

wurden, werden nur bestimmte Skalen zu einem Wert für Allgemeine Intelligenz aggregiert. Andere Skalen, die ebenfalls vom BIS-Modell abgedeckt werden, bleiben dabei unberücksichtigt (vgl. Kapitel 3.2.1). In solchen Zweifelsfällen wurde der Operationalisierung gefolgt, die in den Testmanualen bzw. den einzelnen Studien vorgeschlagen wird, da die meisten Testanwender die in den Manualen vorgeschlagenen Durchführungsweisen wählen, um Objektivität zu gewährleisten. Im Falle des BET und der EUB entspricht dieses Vorgehen den Metaanalysen der Arbeitsgruppe von Hunter und Schmidt, deren Ergebnisse zur Validität von Intelligenztests zu einem großen Teil auf dem kognitiven Faktor der GATB beruhen (Hunter, 1980; Hunter & Hunter, 1984; F. L. Schmidt & Hunter, 1998b).

Die Begründung dafür, der Operationalisierung in den Testmanualen zu folgen, wurde auch angewendet, wenn *zusätzlich* zu den üblichen Intelligenzwerten, die auf Basis von Summen oder Mittelwerten einzelner Aufgaben gebildet werden, Ergebnisse auf Basis einer faktorenanalytischen Extraktion von g berichtet werden: Die faktorielle Lösung wurde – um ein möglichst einheitliches Vorgehen bei allen Studien zu erreichen – nur verwendet, wenn keine Summenscores vorliegen. Dies ist bei einer Studie von Schuler, Funke, Moser und Donat (1995) der Fall.

Schließlich weisen die Studien von Greif (1972) in Bezug auf die Prädiktoren Besonderheiten auf, die Auswirkungen auf die Inklusion hatten. Greif erhob fünf verschiedene Intelligenzverfahren an einer Stichprobe: IST, LPS, WIT, den Analytischen Intelligenztest (AIT) von Meili (1966) und „Markiervariablen“ der JÄGERschen Dimensionen der Intelligenz“ (später: BIS; Greif, 1972, S. 142). Das Hauptinteresse Greifs galt dabei dem IST, dem LPS und dem WIT. Diese wurden – um Übungs- und Transfereffekte auszuschließen – in zufällig permutierter Reihenfolge als erste appliziert, gefolgt von den beiden anderen Verfahren. Er merkt dazu an: „So günstig die Bedingungen für die bevorzugten Verfahren bei diesem Untersuchungsplan sind, so schwierig ist es, die Ergebnisse der benachteiligten Verfahren - des AIT und der Markiervariablen - zu interpretieren“ (S. 69). Aus diesem Grunde wurden aus Greif (1972) nur die Ergebnisse zum IST, LPS und WIT inkludiert.

Beschreibung der inkludierten Intelligenzmaße

Die folgenden Kurzbeschreibungen sollen die Charakteristika der integrierten Intelligenztests verdeutlichen und die Informationen zusammenfassen, die für die metaanalytische Integration benötigt werden. Dazu gehören insbesondere Informationen zur Normierung, Reliabilität und kriterienbezogenen Validität in Bezug auf die genannten Berufs-Erfolgskriterien, die an *nicht* varianzeingeschränkten Stichproben gewonnen wurden. Nur diese sind für die weitere Metaanalyse relevant (vgl. Kapitel 3.6). Weitere Gütemaße und detailliertere Informationen zu den Verfahren sind den Testmanualen und Rezensionen zu entnehmen, auf die jeweils verwiesen wird. Einen Kurzüberblick über die am häufigsten verwendeten Verfahren geben auch Brickenkamp (2002b), Holling et al. (2004), Kanning (2002) sowie Sarges und Wottawa (2004). Die inkludierten Verfahren werden in der Reihenfolge ihrer Häufigkeit besprochen. Die BIS-Struktur-Zuordnung, die zur Inklusions-/Exklusions-Entscheidung durchgeführt wurde, enthält Anhang C.4. Das Vor-

gehen der BIS-Kodierung selbst wird in Kapitel 3.4 erläutert.

GATB-basierte Verfahren: Der Berufseignungstest wurde von Schmale und Schmidtke (1966a) als deutsche Adaptation der General Aptitude Test Battery (GATB; U. S. Employment Service, 1970) veröffentlicht, die neueste dritte Auflage stammt von 1995. Einen Überblick über den Test gibt Schmale (2004). Grundlage für seine Konstruktion war die Faktorentheorie Thurstones. Neben fünf Primärfaktoren wird auch ein *g*-Faktor für Allgemeine Intelligenz und Lernfähigkeit mit drei Subtests (2: Körperabwicklungen, 6: Rechenaufgaben und 7: Begriffsähnlichkeit) abgedeckt. Die Skalen 1: Werkzeugvergleich, 3: Adressenvergleich, 4: Grundrechnen und 5: Figurenlesen sind ebenfalls ins BIS-Modell integrierbar, laden aber auf anderen Faktoren (optomotorische Koordination, Hand- und Fingergeschick). Sie werden nicht zur Messung von *g* verwendet und wurden deshalb nicht inkludiert. Insgesamt besteht der BET aus acht Papier-Bleistift- und neun Gerätetests, es gibt zwei parallele Testformen A und B.

Die Retest-Reliabilität liegt zwischen .88 und .93 für die einzelnen Subtests (Schmale, 2004), interne Konsistenzen sind nicht berichtet. Altersspezifische Normen ($M = 100$, $SD = 10$), liegen für 14-34jährige (erste Version) bzw. 14-38jährige (neueste Version) vor. Berufsspezifische Validierungsstudien werden laufend erweitert, in der neuesten vierten Auflage des Tabellenbandes sind Validierungsergebnisse für 70 Berufe angegeben (Schmale & Schmidtke, 2001). Ergebnisse aus älteren Testauflagen (Schmale & Schmidtke, 1966b, 1984, 1995a) oder Forschungsberichten (Schmale & Schmidtke, 1969) wurden dabei mit berücksichtigt und deshalb *nicht* zusätzlich inkludiert.

Ebenso wie der BET basiert die Testserie für Hauptschulabgänger (EUB) auf der GATB. Sie wurde von der Zentralen Arbeitsgruppe für Grundlagenarbeiten im Psychologischen Dienst (ZAP) der Bundesanstalt für Arbeit entwickelt (Engelbrecht, 1975, 1978), bis ca. 1990 eingesetzt und zum Berufswahltest (BWT; Engelbrecht, Schröder & Elgert, 1991) weiterentwickelt, der auch für Realschüler geeignet ist.

Die EUB umfasst neben Leistungstests Selbstbeurteilungen berufsrelevanter Fähigkeiten und Interessen. Die Leistungstests umfassen logisches Denken, räumliches Vorstellungsvermögen, Rechenfähigkeit, Sprachverständnis, Formauffassung, Büro routine und Hand- sowie Fingergeschicklichkeit (Engelbrecht, 1978). Die ersten vier genannten Leistungstests werden zum *Intellektuellen Leistungsniveau* zusammengefasst (Brambring, 1983) und in dieser Metaanalyse als Intelligenzmaß inkludiert.

Die durchschnittliche Retest-Reliabilität der Leistungstests wird mit $r = .76$ angegeben ($\Delta_t = 6$ Monate, $n = 140$; Engelbrecht, 1975). Normen liegen für beide Testformen, beide Geschlechter und die Altersstufen 14 und 15 Jahre in Form von Staninewerten vor ($N > 10\,000$). Die Untersuchungen zur prädiktiven Validität, die im Manual berichtet sind, konnten auf Grund mangelnder Kennwerte nicht in die Metaanalyse integriert werden (Engelbrecht, 1975, 1978), dafür wurden zwei Studien von Schuler, Barthel und Fünf gelt (1984) in die Analyse einbezogen.

In einer Studie von Kuhl und Greve (1971) wird Intelligenz ebenfalls auf Basis der GATB operationalisiert bzw. in Anlehnung an diese entwickelt. Es handelt sich um zwölf Aufgabengruppen, die faktorenanalytisch den drei Dimensionen *Allgemeinintelligenz*,

Auffassungsgeschwindigkeit und *Sensumotorische Fertigkeiten* zugeordnet sind. Die Aufgabengruppen *Denkrechnen*, *Zahlenrechnen*, *Analogietest*, *Figurenreihen* und *Listenvergleich* laden auf dem Faktor *Allgemeinintelligenz* und werden integriert. Nicht integriert werden die Aufgabengruppen *Werkzeuge-Vergleichen*, *Figuren-Suchen*, die mit *Listenvergleich* auf dem Faktor *Auffassungsgeschwindigkeit* laden, sowie die Skalen *Drahtbiegen 1 und 2*, *Körper-Finden*, *Zeichentest* und *Technischer Bildtest*, die den Faktor *Sensumotorische Fähigkeiten* bilden. Gütemaße und Normwerte liegen für diese Adaptation der GATB nicht vor.

Intelligenz-Struktur-Tests: In 63 Studien wurde eine Version des *Intelligenz-Struktur-Tests* (IST) verwendet. Er erschien 1953 in der ersten Auflage (Amthauer, 1953a) und blieb bis zur dritten Auflage (IST70, Amthauer, 1973) inhaltlich unverändert. Der IST70 besteht aus neun Aufgabengruppen, die den vier Fähigkeitsbereichen *Sprachliche Intelligenz* (Anzahl der Aufgabengruppen: 4), *Rechnerische Intelligenz* (2), *Räumliche Vorstellung* (2) und *Merkfähigkeit* (1) zugeordnet werden (Amthauer et al., 1999). Neben der Interpretation des Gesamt-Testwertes werden auch Interpretationen auf Basis der einzelnen Fähigkeitsbereiche und Aufgabengruppen angeregt (Amthauer, 1973).

Rezensionen zum IST70 (Brocke, Beauducel & Tasche, 1998; Schaarschmidt, 1997; Schmidt-Atzert, Hommers & Hess, 1995; Schmidt-Atzert, 1997) kritisieren seine mangelnde theoretische Fundierung und unzureichende Konstruktvalidität. Statt vier unterscheidbarer Fähigkeitsbereiche werden empirisch zwei bis drei Faktoren gefunden (Amthauer et al., 1999; Greif, 1972; Schmidt-Atzert et al., 1995; Heyde, 1996; Brocke et al., 1998), von Interpretationen auf Strukturebene wird deshalb abgeraten. Eine weitere Kritik bezieht sich auf die Berechnung des Gesamtwertes als Summe aller Aufgabenergebnisse. Dadurch erhalten die einzelnen Fähigkeitsbereiche ein unterschiedliches Gewicht; der so errechnete Gesamtwert spiegelt vor allem sprachliche Intelligenz wieder. Positiv hervorgehoben wird die große Anzahl an Validierungsstudien die zum IST70 vorliegt.

Normen liegen in Form von Standardwerten ($M = 100$, $SD = 10$, $N > 15\,000$) vor. Für den Gesamtwert werden diese getrennt für verschiedene Bildungsniveaus ($N = 12\,964$) und Altersgruppen ($N = 5\,284$) berichtet (Amthauer, 1973).

Die Testhalbierungs-Reliabilität liegt laut Handanweisung für den Gesamtwert bei $r = .97$ ($N = 100$), die Retest-Reliabilität beträgt je nach Studie $r = .95$ ($N = 200$, $\Delta_t = 3\text{-}5$ Tage, Testformen A+B), $r = .91$ ($N = 128$, $\Delta_t = 1$ Jahr, Testformen A+B) bzw. $r = .94$ (N nicht angegeben, 18 Monate). Reliabilitäten auf Aufgabengruppenebene werden ebenfalls berichtet (Amthauer, 1973) müssen aber nach Schmidt-Atzert (1997) zum Teil deutlich nach unten korrigiert werden, wohingegen “der Gesamtwert ein hoch reliables Maß der allgemeinen Intelligenz darstellt“ (S. 112).

Zur prädiktiven Validität werden in der Handanweisung Korrelationen mit beruflichen (Lern-)Leistungsdaten von 13 Stichproben berichtet, die in die Metaanalyse integriert wurden. Für elf Stichproben liegen die Korrelationen auf der Aufgabengruppen-Ebene vor, für zwei Stichproben auf Ebene des Gesamtwertes.

Der IST70 wurde in 57 Studien verwendet. In vier weiteren Studien, die dem Streitkräfteamt der Bundeswehr (1996a) entstammen, kam eine unveröffentlichte IST70-Version

der Bundeswehr zum Einsatz, zu der keine näheren Angaben vorliegen.

Die Kritikpunkte, die zum IST70 angebracht wurden, wurden bei der Weiterentwicklung zum IST 2000 berücksichtigt (Amthauer et al., 1999). Die Fähigkeitsbereiche verbale, numerische und figurale Intelligenz werden jeweils mit drei Aufgabengruppen erfasst und können zu einem Gesamtwert für *Schlussfolgerndes Denken* (BIS-Operation *Verarbeitungskapazität*) verrechnet werden. Zusätzlich wird *Merkfähigkeit* mit einer verbalen und einer figuralen Aufgabengruppe erfasst, die kurzfristige Behaltensleistungen erfordern. Diese Aufgaben bilden zusammen das Grundmodul des IST 2000 und können als Kurzversion eingesetzt werden.

Zusätzlich steht ein Erweiterungsmodul zur Verfügung, um die Messung von figuraler und kristalliner Intelligenz zu ermöglichen. Das Erweiterungsmodul besteht aus einem „Test zur Erfassung des erworbenen Wissens“ (Amthauer et al., 1999, S. 16), der nicht in das BIS-Modell integriert werden kann, da er *langfristige* Behaltensleistungen umfasst. Theoretische Basis bilden die hierarchischen Strukturmodelle von Thurstone, Vernon und Cattell (vgl. Kapitel 2.1). Die Aufgabengruppen des IST 2000 können zwischen den Polen kristalline und fluide Intelligenz angeordnet werden, wobei Matrizenaufgaben dem Pol fluide Intelligenz am nächsten kommen und Wissensaufgaben dem Pol kristalline Intelligenz. Beide Aufgabengruppen werden deshalb als Markiervariablen für die beiden Pole eingesetzt. Rezensionen liegen von Kersting (1999b, 2000) vor. Diese IST-Version wurde in keiner Primärstudie verwendet, war aber Basis für die bis dato neueste, den IST 2000 R (Amthauer et al., 2001), der in zwei Studien (Amelang & Steinmayr, 2006b; Nettelstroth, 2004) eingesetzt wurde.

Der IST 2000 R unterscheidet sich vom IST 2000 durch – im Hinblick auf Trennschärfe und Reliabilität – optimierte Skalen des Grundmoduls. Der Gesamtwert des Grundmoduls wird als schlussfolgerndes Denken (fluide Intelligenz) mit Wissensanteilen (kristalline Intelligenz) interpretiert (Amthauer et al., 2001). Wie im IST 2000 ist die Erfassung der „reinen“ fluiden und kristallinen Intelligenz mit Hilfe des Erweiterungsmoduls möglich, welches stark überarbeitet wurde. Es soll in der neuen Version den Konstruktbereich besser abdecken (zu sechs Themenbereichen werden Fragen nach verbal, numerisch und figural kodiertem Wissen gestellt) und einfachere Operationalisierungen von Wissen, der „reinen“ kristallinen und „reinen“ fluiden Intelligenz ermöglichen.

Eine kritische Auseinandersetzung mit dem IST 2000 R liefert Schmidt-Atzert (2002). Er würdigt die Bemühungen der Autoren des IST, diesen zu verbessern, und spricht dem IST 2000 R einen „Spitzenplatz unter den deutschsprachigen Intelligenztests“ (S. 54) zu. Kritisiert wird die nach wie vor „unzulängliche Modellbegründung“ (S. 55), noch zu leistende Nachweise der Testgüte (Reliabilität und Validität) und das Fehlen eines Maßes für die Allgemeine Intelligenz: „Das Manual liefert keine Entscheidungshilfe dafür, ob dafür das schlussfolgernde Denken mit Wissensanteilen oder das ‚ohne Wissensanteile‘ (fluide Intelligenz) oder vielleicht sogar das Wissen der geeignete Prädiktor ist.“ (S. 56). In den inkludierten Studien wird das schlussfolgernde Denken zur Operationalisierung von Intelligenz verwendet.

Für den IST 2000 R liegen Normen auf der Basis von 3 484 Probanden vor, die 1998–2000 erhoben wurden. Berichtet werden Rohwerte und Standardwerte ($M = 100$, SD

= 10). Diese liegen weiterhin für acht Altersgruppen vor, sowohl als Gesamtwert pro Altersgruppe als auch getrennt für Gymnasiasten und Nicht-Gymnasiasten.

WILDE Intelligenztests: Der WILDE Intelligenztest (WIT) findet in 25 Studien Verwendung. Er wurde erstmalig von Jäger (1963) vorgestellt, die erste Auflage erschien jedoch erst 1983 (Jäger & Althoff, 1983). Eine zweite Auflage mit marginalen Veränderungen (Umformulierung von vier Items und Fehlerkorrekturen in den Normtabellen, vgl. Knebelau, 2002) erschien 1994. Der WIT basiert auf dem Strukturmodell von Thurstone. Die sieben Primärfähigkeiten (*space, number, verbal comprehension, word fluency, memory, reasoning* und *perceptual speed*) werden mit 15 Subtests erfasst, die alle in das BIS-Modell eingeordnet werden können (Jäger & Althoff, 1994; Knebelau, 2002). „Zur Schätzung der *Allgemeinen Intelligenz* empfiehlt sich der Einsatz von mindestens 6 Subtests. ... Die größte gemeinsame Varianz mit der Allgemeinen Intelligenz haben AL [Analogien] und BR [Buchstabenreihen]“ (Jäger & Althoff, 1994, S. 45, Hervorhebung im Original). Der WIT liegt in zwei parallelen Formen (1+2) vor, die beide in einer Lang- und Kurzversion eingesetzt werden können. Lang- und Kurzversion unterscheiden sich nicht im Hinblick auf die Aufgaben, sondern nur durch unterschiedliche Bearbeitungszeiten (Jäger & Althoff, 1994).

Eine Besonderheit des WIT ist die fehlende Normierung des Gesamtwertes. Jäger und Althoff (1994) schreiben dazu:

Der alle Subtestleistungen zusammenfassende Gesamtleistungswert wurde nicht gesondert normiert. Er wird als arithmetisches Mittel der Standardwerte aller Subtests gebildet. Da die Subtests positiv korrelieren, ist damit eine Streuungsreduktion des Gesamtleistungswertes verbunden und bei seiner Interpretation zu berücksichtigen. (S. 12)

Diese Streuungsreduktion ist bei der Korrektur der Varianzeinschränkung zu beachten (Kapitel 3.6.3).

In den Testmanualen des WIT (Jäger & Althoff, 1983, 1994) werden acht Studien aus sieben Quellen zur Kriterienvalidität mit verschiedenen Indikatoren beruflicher (Lern-) Leistung berichtet, die in der Metaanalyse inkludiert wurden (Wolff & Voullaire, 1968; Diepenbrock & Thielepape, 1973; Althoff, 1977; Lometsch, 1975; Thielepape, 1980; Althoff, 1978; Seggebruch, 1982).

Weitere Studien waren über die DGP erhältlich¹. Diese verwenden zum Teil Skalen, die nicht in die veröffentlichte WIT-Version Eingang gefunden haben: *Absurde Schlüsse* (AS), *Verschiedene Beziehungen* (VS), *Figurenmatrizen* (FZ), *Zahlenmatrizen* (ZZ) und *Ähnliche Wortbedeutungen* (ÄW; Bretz & Oldendörp, 1992; Kleinevoss, 1983; Seggebruch, 1982, 1984; Wolf, 1990a; Hollstegge, 1994). Zu dieser modifizierten WIT-Version liegen keine separaten statistischen Kennwerte vor.

Empirische Hinweise auf die Einordnung der Subtests in das BIS-Modell liefern Jäger und Althoff (1994) und Knebelau (2002).

¹Hierfür danke ich S. Jerusel und M. Kersting.

Eine stark überarbeitete Version des WIT erschien 2008 (WIT-2; Kersting et al.). Diese unterscheidet sich von der Ursprungsversion dadurch, dass fünf statt sieben Primärfaktoren nach Thurstone erhoben werden (*verbal comprehension, number, space, reasoning, memory*), wobei das schlussfolgernde Denken im Unterschied zum WIT-1 und „(anders als bei Thurstone) als eine dem verbalen, rechnerischen und räumlichen Denken übergeordnete Skala konzipiert [wird]“ (Kersting, 2006, S. 1). Neun Subtests des WIT-1 wurden in den WIT-2 übernommen und mehr oder weniger stark modifiziert. Neu konstruiert wurde der Subtest *Merkfähigkeit*, der sich dadurch von vergleichbaren Skalen anderer Tests absetzt, dass zwischen dem Lernen und Wiedergeben der Lerninhalte ein anderer Subtest zu bearbeiten ist. „Der WIT-2 erfasst die Merkfähigkeit und nicht, wie andere Tests, die unmittelbare Behaltensleistung.“ (Kersting, 2006, S. 3). Zusätzlich werden zwei weitere Dimensionen erfasst: a) *Arbeitseffizienz* in Form der Simulation der Büro-routinetätigkeit *E-Mails bearbeiten* auf Basis neuer Informationen anstatt auf Basis bereits erlernten Wissens, und b) *Wissen* aus den Bereichen Wirtschaft und Informationstechnologien. Eine Studie enthält Intelligenzdaten, die mit einer Vorläufer-Version des WIT-2 erhoben wurden (Kersting, Althoff & Jäger, 2005).

Matrizentests (CFT, FRT und Raven): Der Grundintelligenztest (CFT) ist eine Adaption des Culture Faire Intelligence Test (CFIT; Cattell, 1957). Erfasst wird der *g*-Faktor nach Spearman, und zwar sprachunabhängig. Ursprünglich war intendiert, Intelligenz kulturunabhängig zu erfassen. Dies ließ sich nicht vollständig verwirklichen, der CFT wird von seinem Autor aber als „relativ kulturfrei“, d.h. im Vergleich zu anderen Meßverfahren unabhängiger von sozialen, verbalen, herkunfts- und vorbildungsmäßigen Einflüssen“ (Weiß, 1980, S. 5) angesehen. Der CFT liegt in drei Versionen vor: Skala 1 (CFT 1; Cattell, Weiß & Osterland, 1997) für 5-9jährige Kinder, Skala 2 (CFT 20; Weiß, 1987) für Kinder ab acht Jahren und Erwachsene „mit einfacher Schulvorbildung“ (Lüken, 2002a, S. 136) sowie Skala 3 (CFT 3; Weiß, 1980), die für Jugendliche ab 14 Jahren und Erwachsene entwickelt wurde. In zehn Primärstudien kommt der CFT 3 zum Einsatz, die Skalen 1 und 2 werden in keiner Studie verwendet und sollen deshalb nicht näher beschrieben werden. Wenn im folgenden vom CFT die Rede ist, ist der CFT 3 gemeint.

Die relative Kulturunabhängigkeit des CFT wird durch die ausschließliche Vorgabe figuraler Items erreicht. Der CFT 3 besteht aus vier Aufgabengruppen, die mit einer knapp bemessenen Zeitbegrenzung zu bearbeiten sind (*speed*-Test; *power*-Testung durch Verdoppelung der Testzeit möglich). Es liegen zwei Parallelformen vor (Form A und B), die jeweils in einer Kurz- und Langversion durchgeführt werden können. Weiß (1980) extrahiert drei Faktoren, die 75% der Gesamtvarianz umfassen: *Reasoning, beziehungsstiftendes Denken* und *Interferenzneigung*. Eingeordnet in die BIS-Modellklassifikation deckt der CFT 3 somit vor allem figurale Verarbeitungskapazität ab. Besonders in der *speed*-Version dürfte er zudem auf Bearbeitungsgeschwindigkeit laden.

Normen liegen für Schüler mit mittleren und höheren Schulabschlüssen, Berufsschüler und Studierende verschiedener Fachrichtungen vor ($N > 3000$). Lüken (2002b) verweist darauf, dass die Normen auf Grund ihres Alters (ca. 35 Jahre) revisionsbedürftig sind.

Die Retestreliabilität liegt bei „großem Zeitintervall“ (Weiß, 1980) zwischen .70 und .80, die Testhalbierungs-Reliabilität beträgt je nach (nicht näher beschriebener) Stichprobe zwischen .82 und .95.

Der Figure Reasoning Test (FRT) wurde 1962 von Daniels veröffentlicht. Er misst wie der CFT den g -Faktor mit sprachfreien, aber nicht kulturunabhängigen figuralen Matrizenaufgaben. Obwohl der Test mehrere Neuauflagen erfahren hat, wurden die Normen ($M = 100$, $SD = 15$) erst in einer überarbeiteten Neuauflage erneuert und um eine Parallelförmigkeit ergänzt (Daniels, 2004).

Für die ursprüngliche Version berichtet Daniels (1962b) einen Testhalbierungs-Koeffizienten von .96. Die Retestreliabilität liegt bei .97 nach 14 Tagen und .89 nach einem Jahr. Die interne Konsistenz der überarbeiteten Version wird mit .86 (Form A) und .82 (Form B) angegeben. Der FRT wird in 14 Studien eingesetzt.

Mit den Matrizen tests von Raven wird der g -Faktor nach Spearman erfasst, sie haben zudem eine hohe Ladung auf fluider Intelligenz (Carroll, 1993).

Die Matrizen tests nach Raven liegen in drei Schwierigkeitsgraden vor (Holling & Liepmann, 2004): Coloured Progressive Matrices (CPM) stellen das einfachste Verfahren dar, das speziell für den Einsatz bei Kindern entwickelt wurde. Standard Progressive Matrices (SPM) decken den mittleren Fähigkeitsbereich ab und können ab einem Alter von sechs Jahren eingesetzt werden. Eine überarbeitete Version der SPM mit etwas höherer Itemschwierigkeit wurde unter dem Namen Standard Progressive Matrices Plus (SPMPLS; Raven et al., 2003) veröffentlicht. Die Advanced Progressive Matrices (APM) sind schließlich für überdurchschnittlich begabte Jugendliche und Erwachsene konzipiert.

In einer Studie kommt der SPMPLS zum Einsatz (Schmidt-Atzert & Eser, 2006). Darüber hinaus werden in mehreren Studien Matrizenaufgaben in der Art der Raven-Tests gestellt, ohne dass diese genauer spezifiziert werden. Die Salzgitter Service und Technik GmbH (2004) verwendet z.B. eine „interne Kurzform der Raven Progressive Matrices“. Normen für eine deutsche Stichprobe liegen vor, auch getrennt nach Bildungsniveau und Alter. Die Split-Half-Reliabilität der SPMPLS wird mit $> .90$ angegeben, Retest-Reliabilitäten liegen zwischen .83 und .93 (Schuhfried, 2004).

Leistungsprüfsystem bzw. Prüfsystem für Schul- und Bildungsberatung: Von Horn wurde 1962 das Leistungsprüfsystem (LPS) als Weiterentwicklung des Begabungstestsystems (BTS; Horn, 1956) vorgestellt. Die Tests basieren auf dem Strukturmodell und den Faktorenanalysen von Thurstone (Thurstone, 1941, nach Horn, 1969). Eine Matrizenaufgabe in der Art von Raven's Progressive Matrices wurde ergänzt.

Das LPS besteht aus 15 Aufgaben, die die sieben Primärfaktoren Thurstones erfassen. Der Test kann in einer Lang-, Standard- und Kurzform eingesetzt werden. Es existieren zwei parallele Formen A und B. Die Version von 1962 wird in fünf Studien aus vier Quellen eingesetzt (Barthel, Fünfgelt & Schuler, 1982; Deutsche Lufthansa AG, 1980; Greif, 1972; Kettel & Simmat, 1969), die zweite Auflage in zwei Studien (Dahle & Erdmann, 2001; Steinmann, 1997).

Eine Normierung des LPS von 1983 liegt für $N > 9\,000$ Personen vor, die in Bezug auf die Art der allgemeinbildenden Schule, das Alter, Geschlecht und den Beruf des Vaters

repräsentativ sind. Die Normen werden auch für verschiedene Altersgruppen getrennt berichtet.

Testhalbierungs-Reliabilitäten liegen zwischen .90 und .99 für die Untertestgruppen und zwischen .88 und .99 für Aufgabengruppen. Paralleltest-Reliabilitäten werden nicht berichtet.

Für Schüler hat Horn (1969) eine auf schulische Leistungserfassung optimierte Version des LPS von 1962 veröffentlicht (Prüfsystem für Schul- und Bildungsberatung, PSB). Sie besteht aus zehn LPS-Aufgaben, die die sieben Primärfaktoren Thurstones abdecken, und liegt ebenfalls in zwei parallelen Formen A und B vor. Die Normierung erfolgte anhand einer Stichprobe von ca. 10 000 Schülern aus Baden-Württemberg im Alter von 9 und 20 Jahren (Centilwerte). Sie ist repräsentativ im Hinblick auf väterliche Berufe, Schularten und geographische Verteilung. Psychometrische Gütemaße zum PSB werden nicht berichtet. LPS und PSB werden in 16 inkludierten Studien eingesetzt (vgl. Tabelle E.27).

Berliner Intelligenzstruktur-Tests: Der Berliner Intelligenzstruktur-Test (BIS) stimmt strukturell mit dem BIS-Modell überein, das in Kapitel 2.1 vorgestellt wurde. Ziel der BIS-Entwicklung war es, einen Test zur Verfügung zu stellen, der „eine möglichst repräsentative Stichprobe des ‚Universums‘ kognitiver Prozesse“ (Jäger et al., 1997, S. 16) darstellt, indem alle zwölf Facetten des BIS-Modells abgedeckt werden. Dadurch soll eine konstruktvalide Messung der Allgemeinen Intelligenz und der Subdimensionen von Intelligenz ermöglicht werden (Süß, 2004).

Die einzelnen Aufgabengruppen des BIS können entweder zu inhaltshomogenen Skalen gebündelt werden, die in Bezug auf die Operationen heterogen sind (*sprachgebundenes/verbales, zahlengebundenes/numerisches* oder *anschauungsgebundenes/figurales Denken*) oder zu operationshomogenen Skalen, die in Bezug auf die Inhalte heterogen sind (*Verarbeitungskapazität, Einfallsreichtum, Merkfähigkeit* und *Bearbeitungsgeschwindigkeit*; vgl. Anhang B.2). Eine Besonderheit im Vergleich zu anderen Intelligenz-Modellen und -Tests stellt dabei die explizite Berücksichtigung kreativer Fähigkeiten durch die Operation *Einfallsreichtum* dar (Jäger et al., 1997). Die gemeinsame Varianz aller Aufgabengruppen wird zur Operationalisierung der Allgemeinen Intelligenz verwendet (Jäger et al., 1997).

Eine Testung auf Basis des gesamten BIS dauert über zwei Stunden; eine Kurzform, die aus 15 Aufgaben besteht und ca. 45 Minuten in Anspruch nimmt, liegt ebenfalls vor. In beiden Formen hat die Operation *Verarbeitungskapazität* ein stärkeres Gewicht als die anderen drei Operationen jeweils haben (Jäger et al., 1997). Mit der Kurzform können die Verarbeitungskapazität und die Allgemeine Intelligenz valide erfasst werden, eine Interpretation der Kurzformen der anderen Operationsbündel oder der einzelnen Inhaltsklassen wird von Jäger et al. (1997) als nicht zulässig erachtet.

Ausgangspunkt für die Entwicklung des BIS war das BIS-Modell mit den 2 000 Aufgabentypen, die ihm zugrunde liegen. Die Aufgabentypen wurden in einem mehrstufigen Prozess auf 46 reduziert, die Optimierung des Tests anschließend in mehreren Schritten durchgeführt. So entstanden die BIS Form 1 (Jäger, Süß & Brühl, 1988), Form 2 (Süß,

Brühl & Jäger, 1989) und Form 3 (Süß & Jäger, 1990) als technische Berichte am Institut für Psychologie der FU Berlin. Die erste veröffentlichte und bisher einzige Auflage des BIS erschien mit Form 4 (Jäger et al., 1997) „trotz einiger Defizite im bisherigen Entwicklungsstadium auf Drängen zahlreicher Fachvertreter, die mit dem Instrument arbeiten möchten bzw. die bereits mit Vorformen erfolgreich praktiziert haben“ (S. 2). Der BIS wird in sechs inkludierten Studien eingesetzt (vgl. Tabelle E.26).

Zu den Defiziten des BIS zählt die erst als vorläufig zu erachtende Normierung für 16- bis 19-jährige anhand einer Stichprobe von $N = 478$ Mittelschülern und Gymnasiasten aus der deutschsprachigen Schweiz. „Ihre Schulbildung ist mit Oberschülern (Gymnasiasten) und Realschülern in der Bundesrepublik Deutschland vergleichbar“ (Jäger et al., 1997, S. 47). Der Einsatz des BIS eignet sich nach Süß (2004) bei Pbn dieser Bildungsniveaus bis zu einem Alter von 25 Jahren. Allerdings stehen für Pbn über 19 Jahren keine Altersnormen zur Verfügung. Normierte Werte sind in der Standardmetrik ($M = 100$, $SD = 10$) berichtet.

Als Reliabilitäten werden interne Konsistenzen und Split-Half-Koeffizienten angegeben. Cronbachs α liegt für die Skalen zwischen .75 und .89 – ein Wert, der die Autoren zufriedenstellt, da die Konstruktion des BIS nicht auf möglichst hohe Homogenität ausgelegt war (Jäger et al., 1997). Für aussagekräftiger werden die internen Konsistenzen der operations- bzw. inhaltshomogenen Bündel gehalten, die zwischen .87 und .95 liegen. Testhalbierungs-Koeffizienten liegen zwischen .76 und .90 auf Skalenebene, für Verarbeitungskapazität in der Kurzform (K-S) wird .73, für Allgemeine Intelligenz in der Kurzversion (AI-S) .81 berichtet. Die zeitliche Stabilität liegt zwischen .74 und .90 ($\Delta_t = 1$ Jahr).

Eine weitere Version des BIS, die in drei Studien zum Einsatz kommt, ist eine interne Version der Deutsche Gesellschaft für Personalwesen e.V. (DGP) von 1996 (BIS-r-DGP). Sie unterscheidet sich laut Kersting und Beauducel (2004) vom BIS in der Form 4 durch die Erfassung von *Einfallsmenge* statt *Einfallsreichtum*, wobei dafür weniger Aufgaben benutzt werden (fünf statt zwölf). Die Zelle EN des BIS-Modells ist nicht abgebildet. Somit besteht der BIS-r-DGP aus 38 statt 45 Aufgaben. Dafür können weitere eher simulationsorientierte Module (z.B. Allgemeinwissen, BüroROUTINETätigkeit, Diktat) integriert werden. Empirische Befunde zur Einordnung der BIS-r-DGP-Aufgabengruppen in das BIS-Modell enthält Kersting (1999a).

Die Normierung wurde ursprünglich anhand einer Stichprobe von 3 274 Personen durchgeführt. Da der BIS-r-DGP nicht vertrieben wird, „sondern die Testdurchführung, die Testauswertung und die Testinterpretation“ (Kersting & Beauducel, 2004, S. 149), kann die DGP ihre Normbasis beständig erweitern. Sie lag 2004 schon bei über 10 000 Personen, „die den Test unter den ‚Ernstfallbedingungen‘ einer Bewerbung bearbeitet haben“ (S. 154).

Die internen Konsistenzen für die einzelnen Skalen liegen zwischen .73 und .88. Die interne Konsistenz von AI beträgt .89 bei Berechnung auf Basis der Aufgaben und .94 bei Berechnung auf Basis der operations- bzw. inhaltshomogenen Aufgabenbündel (Kersting & Beauducel, 1997).

Wonderlic Personnel Test: Mit dem Wonderlic Personnel Test (WPT; [Wonderlic Inc., 2002](#)) liegt ein Verfahren vor, das sich im Vergleich zu den bereits beschriebenen Testverfahren durch seine Zeitökonomie auszeichnet. Die Messung von g nach Spearman dauert lediglich zwölf Minuten. In dieser Zeit werden bis zu 50 verbale, numerische und figurale Items in aufsteigender Schwierigkeit beantwortet. Eine Erfassung spezifischer Intelligenzfaktoren ist mit dem WPT nicht möglich. Der WPT ist in der Personalpsychologie in den USA das am weitesten verbreitete Verfahren und wird auch in der Forschung häufig eingesetzt ([Judge, Colbert & Ilies, 2004](#); [F. L. Schmidt & Hunter, 2004](#)). Es existieren zehn englischsprachige Testformen ([Wonderlic Inc., 2002](#)) und seit Januar 2007 eine revidierte Version (WPT-R; [Wonderlic Inc., 2007](#)).

Die englischsprachigen Versionen sind für verschiedene Altersgruppen, Bildungsniveaus und Berufsgruppen normiert ($M = 21$, $SD = 7.12$; [Wonderlic Inc., 2002](#)). Zu ihrer Reliabilität liegen zahlreiche Befunde vor, die [Wonderlic Inc. \(2002\)](#) zusammenfassen: Die Retestreliabilität liegt zwischen .82 und .94, wenn beide Tests kurz hintereinander vorgelegt werden, und wird mit .94 angegeben, wenn zwischen beiden Testzeitpunkten ein Jahr liegt. Die Paralleltest-Reliabilität beträgt .73 bis .95, die Testhalbierungs-Reliabilität .88 bis .94, wenn die zunehmende Itemschwierigkeit mit in Betracht gezogen wird .88 (nach Kuder-Richardson).

In vier Studien aus drei Quellen ([Blickle, Kramer & Mierke, 2005](#); [Blickle, Witzki & Schneider, 2006](#); [Marcus, Schuler, Quell & Hümpfer, 2002](#)) wird eine deutschsprachige Version des WPT eingesetzt ([Wonderlic Inc., 1996](#)). Für diese liegen keine separaten Normen und keine Gütemaße vor, die an repräsentativen Stichproben gewonnen wurden. Zwar liegen Adaptationen des WPT für elf Sprachen vor, diese befinden sich jedoch noch in einem Entwicklungsstadium. [Wonderlic Inc. \(2002\)](#) mahnen deshalb an: „interpretation of test results when using these forms should be taken with some caution“ (S. 44).

Begabungstest: Der Begabungstest (BT) ist ein Verfahren, welches vom Psychologischen Dienst der Hüttenwerk Salzgitter AG entwickelt wurde. Der Test ist nicht veröffentlicht, die Informationen wurden deshalb den Primärstudien direkt entnommen ([Ehlers, 1961](#); [Paff, 1966](#)). Der BT enthält Rechenaufgaben, Figurenanalogien und logische Denkaufgaben und erfasst damit „die geistige Beweglichkeit“ ([Paff, 1966](#), S. 24). Damit deckt er mehr Inhaltsfacetten des BIS-Modells ab (numerisch, figural) als die Matrizen tests (nur figural), aber weniger als der WPT, der auch verbale Komponenten beinhaltet. Normen zum Test sind nicht bekannt, ebenso werden keine Gütemaße für nicht varianzeingeschränkte Populationen berichtet.

Arbeitsproben zur berufsbezogenen Intelligenz: Testverfahren zur Erfassung kognitiver Fähigkeiten, die sich durch einen hohen Anforderungsbezug auszeichnen, indem sie Tests und Arbeitsproben integrieren sind die *Arbeitsproben zur berufsbezogenen Intelligenz*. Sie liegen in zwei Versionen vor: einer Arbeitsprobe für Büro- und kaufmännische Tätigkeiten (AZUBI-BK; [Schuler & Klingner, 2005](#)) und einer Arbeitsprobe für technische und handwerkliche Tätigkeiten (AZUBI-TH [Görlich & Schuler, 2007](#)).

Der AZUBI-BK besteht aus einem Grundmodul, das sprachliche und rechnerische Fähigkeiten sowie Gedächtnisleistungen mit acht arbeitsplatzbezogenen Aufgaben (z.B. *Bilanzwerte vergleichen, Telefon- und Zimmernummern merken*) abprüft. Darüber hinaus werden vier simulationsorientierte Postkorbmodule vorgelegt: *Fax vervollständigen, Emails sortieren, Adressen prüfen* und *Porto berechnen*. Die spezifischen Fähigkeitsbereiche können getrennt ausgewertet werden. Ist man nur an einer Operationalisierung von Allgemeiner Intelligenz interessiert, kann eine Kurzversion eingesetzt werden.

Der AZUBI-TH besteht aus den drei Aufgabengruppen *Rechnen, technisches Verständnis, Rechtschreibkenntnisse* und zwei Aufgabengruppen zum *räumlichen Vorstellungsvermögen*. Der AZUBI-TH liegt in zwei parallelen Testformen vor.

Beide AZUBI-Testverfahren wurden nicht für die Gesamtbevölkerung, sondern für potenzielle Bewerber auf Ausbildungsplätze normiert: insbesondere für Gymnasiasten und Berufsschüler beim AZUBI-BK ($N = 2\,135$) sowie hauptsächlich für Haupt-, Real- und Berufsschüler beim AZUBI-TH ($N = 4\,452$). Die Reliabilitäten des AZUBI-BK betragen Cronbachs $\alpha = .96$ bzw. streuungskorrigierte Retestreliabilität $r_{tt} = .93$ (Kurzmodul: $\alpha = .95$, $r_{tt} = .88$). Beim AZUBI-TH betragen beide Reliabilitäten $.92$.

Die integrierten Daten stammen aus der *Validierungsstudie 1* des AZUBI-BK (Schuler & Klingner, 2002) und einer Validierungsstudie zum AZUBI-TH, die vor ihrer Veröffentlichung von Görlich und Schuler (2004) zur Verfügung gestellt wurden.

Eigenentwicklungen: In 17 Quellen werden von den jeweiligen Autoren selbst entwickelte Testverfahren angewendet oder eigene Testbatterien mit verschiedenen Aufgabengruppen zusammengestellt, die zum Teil den bereits beschriebenen Verfahren entnommen worden sind. Diese Eigenentwicklungen und -zusammenstellungen sollen im Folgenden kurz vorgestellt werden. Dabei wird i.d.R. darauf verzichtet, die Aufgabenstellung im Detail zu beschreiben. Die Aufgabenbenennungen sind meist so aussagekräftig, dass sie für einen ersten Eindruck genügen. Die Testverfahren sind überwiegend nicht normiert. Ebenso liegen meist keine Reliabilitätskennwerte vor, die an varianz-uneingeschränkten Stichproben erhoben wurden.

Biernat und Scholz (1995) verwenden in ihrer Studie 14 Prädiktoren, die den fünf Bereichen *Kenntnisse* (Diktat, Grundrechnen, Wissensfragen), *Berufsspezifische Fähigkeiten* (FLT, DBP, Praktisches Verständnis), *Konzentrationale Belastbarkeit* (d2), *Anschauungsgebundenes Denken* und *Gedächtnis* zugeordnet werden. Die drei erstgenannten Bereiche wurden nicht inkludiert. Berücksichtigt werden die Tests, die dem Anschauungsgebundenen Denken und Gedächtnis zugeordnet werden können. Dabei handelt es sich um *Rechts-Links-Wendungen, Abwicklungen, Drehfiguren* und *Unterschiede Erkennen* für das Anschauungsgebundene Denken sowie um *Verbales, Numerisches* und *Figurales Gedächtnis*, die den entsprechenden BIS-Inhalten und der Operation *Merkfähigkeit* zugeordnet werden können. Die Tests wurden von der DGP eingesetzt, die auch an der Entwicklung von WIT und BIS beteiligt war.

Das Deutsche Zentrum für Luft- und Raumfahrt (DLR) entwickelte eine eigene Testbatterie zur Eignungsdiagnose (Damitz et al., 2000), diese umfasst einige kognitive Merkmalsbereiche, die nicht inkludiert wurden. Dies sind *Englisch* (akustisch und schriftlich),

Technik (Fliegerischer Interessen-Test, Technischer Verständnis-Test) und eine simulationsorientierte Aufgabengruppe (*Approach Control Test*). Inkludiert wurden hingegen die Merkmalsbereiche *Rechnen, Raumvorstellung, Aufmerksamkeit/Konzentration* und *Merkfähigkeit*. Der Merkmalsbereich *Rechnen* wird durch die Aufgabengruppen *Kopfrechnen* und *Rechenaufgaben* abgedeckt, der Merkmalsbereich *Raumvorstellung* durch die Aufgabengruppen *Akustische Richtungsangaben, Flugzeug-Positionstest* und *Versteckte Figuren Finden*, der Merkmalsbereich *Aufmerksamkeit/Konzentration* durch die Aufgabengruppen *Bourdon-Buchstabendurchstreichtest, Konzentrations-Belastungstest, Würfel-Summen* und *Vigilanz-Test*. Der *Vigilanz-Test* ist nicht gut ins BIS-Modell integrierbar. Am ehesten kann er der Operation B zugeordnet werden. Der Merkmalsbereich *Merkfähigkeit* schließlich wird durch die Aufgabengruppen *Clearance-Test* und *Merkfähigkeitstest* abgedeckt. Reliabilitätsangaben werden auf der Ebene der Aufgabengruppen berichtet, die kriterienbezogenen Validitätskoeffizienten auf der Ebene der Merkmalsbereiche.

In der Studie von Herzog (2004) werden zur Erfassung der *Allgemeinen Verwendungsbreite* (AVB) ein Wortbeziehungstest, ein Figurendenkttest und ein Rechentest verwendet. Die AVB kann als Intelligenzmaß angesehen werden: „Die drei Tests der AVB sind die besten Prädiktoren, welche die Intelligenz (G-Faktor) wiedergeben. Untersuchungen dazu wurden in den 70er Jahren durchgeführt, dargestellt in den Wehrpsychologischen Untersuchungen“ (Herzog, 2004, S. 46). Die prädiktive Validität der AVB wird berichtet, es ist keine Aggregation der Daten auf Einzeltest-Ebene nötig. Aus diesen Gründen wurde die AVB inkludiert, aber keine weiteren Prädiktoren (zum mechanisch-technischen Verständnis, räumlichen Vorstellungsvermögen, zur Reaktions-, Rechen- oder Sprachfertigkeit), obwohl sie sich teilweise in das BIS-Modell einordnen ließen.

Hossiep (1995) verwendete Aufgabengruppen aus mehreren Intelligenztests und stellt daraus eine eigene Testbatterie zusammen. Die Aufgabengruppen *Satzergänzen, Rechenaufgaben* und *Analogien* sind IST-Derivate, *Zahlenreihen* wurden dem Analytischen Intelligenz-Test von Meili (1966) entnommen und *Begriffliche Gemeinsamkeiten* dem BET. Die bisher genannten Aufgabengruppen wurden gegenüber ihren Originalen mit dem Ziel modifiziert, höhere Gütemaße zu erreichen. Dazu führte Hossiep zunächst eine Vorstudie durch. Weitere Aufgabengruppen wurden unverändert dem CFIT entnommen. Es handelt sich um *Schlussfolgerndes Denken, Erkennen von Gemeinsamkeiten, Beobachtungsfähigkeit* und *Entdecken von Regelwidrigkeiten*. Nicht integriert wurden die Leistungstests *Rechtschreibtest* und *Konzentrations-Leistungstest* (KLT).

Kreuscher (1987) verwendete ein Verfahren, das von Sieber in den Jahren 1966 bis 1985 entwickelt wurde. Integriert wurden zwei Skalen, die aus mehreren Aufgabengruppen bestehen: *Intelligente Anpassung*, bestehend aus den Aufgaben *Sätze ergänzen, Text-Rechenaufgaben lösen, Zahlenreihen fortführen* und *Analogien finden* sowie *Allgemeine Intelligenz*, bestehend aus 21 Matrizenaufgaben nach Raven (Kreuscher, 1987). Weitere kognitive Variablen werden nicht berichtet.

Linz (1985) verwendete einen „Schätzwert der allgemeinen bildungsunabhängigen Intelligenz“ (S. 316), von dem nur bekannt ist, dass er von der Bundesagentur für Arbeit (BA) erhoben wurde. Das Intelligenzmaß liegt in fünfstufiger ordinaler Form vor und

wird als „Intelligenzrating (C-Skala)“ (S. 316) bezeichnet.

Neumann (1938) berichtet die prädiktive Validität verschiedener *Leistungsproben*. Darunter eine, die mit „Intelligenz, Gedächtnis, Aufmerksamkeit“ (S. 9) bezeichnet wird. Der Aspekt *Intelligenz* wird mit „Kopfrechenaufgaben zur Feststellung der Kombinationsgabe“ (S. 119) und einer „Beurteilung technischer Vorlagen zur Feststellung des technischen Verständnisses“ (ebd.) erfasst. Die Aspekte *Gedächtnis* und *Aufmerksamkeit* erfassen die Behaltensleistung von sinnvollen Worttripeln (BIS-Zelle VM; Neumann, 1938). Die weiteren Leistungsproben umfassen sensumotorische Fähigkeiten und werden nicht integriert.

Beim räumlichen Vorstellungstest (RVT), den Paff (1966) einsetzt, handelt es sich um eine Aufgabengruppe, die dem Differential Aptitude Test (DAT; Bennett, Seashore & Wesman, 1952) entnommen wurde. Die Aufgaben bestehen darin, Puzzleteile zu benennen, aus denen ein bestimmtes Muster erstellt werden kann. Nähere Angaben zum DAT können Bennett, Seashore und Wesman (1956) entnommen werden.

Die ältesten inkludierten Studien stammen von Roloff (1928). Er prüft die prädiktive Validität verschiedener Variablen, die er in körperlich/sensumotorische (nicht integriert) und geistig/intellektuelle unterteilt. Zu den geistig/intellektuellen gehören die *sprachlich-logische Intelligenz* und das *Arbeiten mit räumlicher Vorstellung* (beide werden integriert) sowie das *technische Verständnis* (nicht integriert). Die sprachlich-logische Intelligenz wird mit den Aufgaben *Ergänzen von Textlücken* und *Gedächtnis für sinnvolle Zusammenhänge* operationalisiert. Räumliche Vorstellung umfasst *Rybakowfiguren*, *Vergleiche von Körperzeichnungen*, *Räumliche Konstruktionen*, *Kombinieren von Flächenelementen* und *Figurvergleiche*.

In einer Untersuchung zum beruflichen Erfolg von Mitarbeitern einer Forschungs- und Entwicklungsabteilung in einem Industriebetrieb erhoben Schuler, Funke et al. (1995) zehn kognitive Fähigkeiten anhand von z.T. modifizierten Aufgabengruppen aus unterschiedlichen Quellen. Sie prüften faktorenanalytisch, ob eine Dimensionsreduktion möglich ist. Sie berichten prädiktive Validitäten für zwei Faktoren: a) Einen Kreativitätsfaktor, der mit den Verfahren *Hypothesen aufstellen*, *Funktionsumwandlung* und *Versteckte-Figuren-Test* erfasst wird. Der Kreativitätsfaktor wurde auf Grund der geringen Korrelation mit dem Intelligenzfaktor (.15) nicht integriert (Schuler, Funke et al., 1995). b) Einen Allgemeine-Intelligenz-Faktor, der erfasst wird mit den Verfahren *Wortauswahl*, *Mängel erkennen*, *Fehler im Text*, *Bild-Zahl-Paare*, *Matrizentest*, *Aufgaben zum schlussfolgernden Nachdenken* und *Assoziationen*.

Es wurden auch Studien inkludiert, die nicht näher bezeichnete Intelligenztests verwenden. Dies war der Fall bei einer Validitätsstudie zur Prüfung der Güte von Eignungsbeurteilungen bei Bankangestellten von Schuler, Moser, Diemand und Funke (1995), die Ergebnisse auf Basis eines „kognitiven Fähigkeitstests“ (S. 50) berichten. Ebenso inkludiert wurden Studien, die von der Gruppe Wehrpsychologie des Streitkräfteamtes durchgeführt wurden und einen „konventionellen Intelligenztest des PsychDstBW [Psychologischen Dienstes der Bundeswehr]“ (Streitkräfteamt der Bundeswehr, 1996b, 2001) oder die *Allgemeine Verwendungsbreite* benutzten, die mit einem adaptiven Basisfähigkeitstest des Psychologischen Dienstes der Bundeswehr operationalisiert wurde. Letztere

besteht aus Analogie-, Matrizen- und Rechenaufgaben (Melter, 6.11.2006, persönliche Kommunikation).

In zwei Studien der Salzgitter Service und Technik GmbH (2004), werden jeweils Matrizenaufgaben nach Raven und ein Test zum Erkennen logischer Beziehungen bei sprachlichem Material verwendet. Beide Verfahren wurden integriert.

3.2.2. Erfolgskriterien

Im Gegensatz zum Intelligenzkonstrukt ist beruflicher Erfolg kein homogenes Konstrukt, das sich mit *einem* Faktor abbilden ließe (Reeve & Hakel, 2002). Aggregate von inhaltlich heterogenen Maßen können nur schwer interpretiert werden („Äpfel und Birnen“-Problem; Sharpe, 1997). Dementsprechend weisen auch Hunter und Schmidt (2004) darauf hin, dass „measures of different dependent variable constructs should ordinarily not be combined in the same meta-analysis ..., but if they are, separate meta-analyses should also be reported for each conceptually different dependent variable“ (S. 469). Auf Seite der Kriterien war deshalb eine Homogenisierung der Erfolgsindikatoren erforderlich: sie werden getrennt für die drei Erfolgsbereiche Lern-, Arbeitsleistung und Karriereerfolg betrachtet.

Berufsbezogene Lernleistung

Nach Durchsicht der berichteten Erfolgsindikatoren wurden für *berufsbezogene Lernleistung* nur Vergleiche mit Bezugsstandards aus Fremdbeurteilungen inkludiert. Dies sind Berufsschulnoten und Beurteilungen aus der betrieblichen Praxis. Neutrale Kennzahlen wurden in vergleichsweise wenigen Studien berichtet: die Dauer bis zum erfolgreichen Abschluss der Aus- oder Weiterbildung (Kreuscher, 1987; Schäfer, 1986) oder Fehlerquoten (Damitz et al., 2000). Ebenfalls nicht berücksichtigt wurden Indikatoren des Studienerfolgs (vgl. Kapitel 2.2.1).

Im Fokus dieser Arbeit befindet sich beruflicher Erfolg im Allgemeinen. Werden mehrere Leistungsdimensionen berichtet, so wurden diese nur inkludiert, wenn kein aggregiertes Ergebnis auf Gesamtleistungs-Ebene berichtet wird. Dies ist z.B. der Fall, wenn einzelne Berufsschul-Fachnoten berichtet werden, aber nicht die Gesamtnoten des Zwischen- oder Abschlusszeugnisses. Zusammenhänge mit Einzelnoten wurden ansonsten exkludiert (z.B. bei Lang, 1990; Seggebruch, 1982).

Spezifische Arbeitsleistung

Für den Erfolgsbereich *spezifische Arbeit* liegen ebenfalls ausreichend Studien vor, bei denen *typische* Arbeitsleistung anhand von Normvergleichen von Fremdurteilern beurteilt wird. Exkludiert wurden Studien mit Einschätzungen der *maximalen* Arbeitsleistung, wie sie z.B. in Assessment Centern zur Personalentwicklung erfasst werden (z.B. bei Hasselmann, 1993).

Da die Validität von Intelligenz in Bezug auf *praxisrelevante* Arbeitsleistung untersucht werden soll, wurden Kriterien exkludiert, die speziell zur Erfassung der konvergen-ten oder diskriminanten Validität von Intelligenz zusammengestellt wurden. Sie werden

nicht wegen ihrer Bedeutung für die Berufspraxis berichtet, sondern weil sie eine besonders hohe bzw. geringe Validität von Intelligenz aufweisen sollen. Dies ist z.B. der Fall bei Marschner (1964): für den Arbeitsplatz relevant sind 30 Bewährungskriterien, von denen nur sechs ausgewertet wurden, die den Dimensionen *Fachwissen* und *Denken* zugeordnet werden können (konvergente Validität). Auch in der Studie von Kersting (1999a) werden neben dienstlichen Beurteilungen – die berücksichtigt wurden – zu Forschungszwecken Einschätzungen der Intelligenz und des Problemlöseverhaltens erhoben (konvergent), sowie das Kooperationsverhalten (diskriminant). Diese wurden exkludiert.

Werden Kriterien hingegen auf Grund ihrer Bedeutung für die Berufspraxis in den Studien berücksichtigt, werden sie inkludiert, auch wenn es sich um Kriterien handelt, bei denen eine besonders hohe (z.B. Intelligenzeinschätzung) oder geringe Korrelation (z.B. Motivationseinschätzung) mit Intelligenz vermutet werden muss.

Selbstbeurteilungen korrelieren laut Harris und Schaubroeck (1988) moderat mit Abwärts- ($\hat{\rho} = .35$) und Seitwärtsbeurteilungen ($\hat{\rho} = .36$); die Korrelation wird zudem von der Berufsgruppe beeinflusst. Selbstbeurteilungen wurden nicht inkludiert, da Fremdbeurteilungen in der betrieblichen Praxis von größerer Relevanz sind und zudem der *single source bias* vermieden werden kann. Darüber hinaus umfassen (globale) Fremdbeurteilungen durch Vorgesetzte mehr Leistungsdimensionen als (globale) Selbstbeurteilungen, etwa Aspekte der adaptiven Leistung wie die Flexibilität, mit der auf Neuerungen und Veränderungen reagiert wird (Hunter & Schmidt, 1996).

Bei den inkludierten Fremdurteilen handelt es sich i.d.R. um Abwärtsbeurteilungen, die in einem Fall mit Seitwärtsbeurteilungen (Kersting, 1999a) und in einem weiteren Fall mit Seitwärts- und Aufwärtsbeurteilungen (Blickle et al., 2005) kombiniert sind. In der Studie von Greif (1972) besteht das Kriterium ausschließlich aus Seitwärtsbeurteilungen. Da Abwärts- und Seitwärtsbeurteilungen hoch miteinander korreliert sind ($\hat{\rho} = .62$ laut der Metaanalyse von Harris & Schaubroeck, 1988) scheint eine Zusammenfassung legitim, auch wenn in weiteren Metaanalysen gezeigt wurde, dass die Höhe der Übereinstimmung zwischen Abwärts- und Seitwärtsbeurteilungen von den Urteilsdimensionen abhängt (vgl. dazu Viswesvaran, Schmidt & Ones, 2002). Entscheidend ist, dass es sich um Fremdbeurteilungen von Beurteilern handelt, die Einblick in die Arbeitsweise der Beurteilten haben.

Aggregierte Beurteilungen mehrerer Leistungsdimensionen wurden gegenüber Single-Item-Globalurteilen bevorzugt, wenn beide berichtet werden (vgl. Kapitel 2.2.2). Dies betrifft z.B. die Studie von Schuler, Funke et al. (1995): die globale Einschätzung wird dort zur Validierung der umfassenden Leistungsbeurteilung benutzt und wurde deshalb nicht inkludiert. Die Studien von Althoff (1968) berichten ausschließlich Globalurteile; sie wurden berücksichtigt.

Neutrale Kennzahlen, die dem Erfolgsbereich *spezifische Arbeit* zugerechnet werden können, werden nur selten berichtet, z.B. Verkaufserlöse bei Kreuscher (1987) oder Produktionsdaten bei Wiegand (1977). Sie wurden exkludiert. Auf diese Art und Weise wurde eine Homogenisierung im Hinblick auf die Datenart und Datenquelle erreicht. Die Studien sind jedoch *inhaltlich* heterogen (Tabelle 3.3): Neben Erfolgsmaßen, die der Kern-, umfeldbezogenen oder adaptiven Leistung zugeordnet werden können (z.B. bei

Tabelle 3.3

Angaben zu Dimensionen und Verfahren der subjektiven Leistungsbeurteilungen

Quelle	Inhalt der Ratings	Anzahl Rater	Anzahl Items	w_{BB} (%)
Althoff (1968)	Globalurteil	1	1	5.0
Brandstätter (1970)	mittlere Gesamtbeurteilung verschiedener Dienstleistungszeugnisse	o. A.	o. A.	3.7
Dreyer und Grabitz (1971)	Führungseigenschaften, Zusammenarbeit mit Kollegen, Leistungsumfang, Qualität, Arbeitseinstellung, geistige Beweglichkeit/Umstellungsfähigkeit	2-4	o. A.	3.8
Amthauer (1973)	Mittel betrieblicher Beurteilungen	o. A.	o. A.	29.8
Schuler, Funke et al. (1995)	16 Verhaltens-, 15 Eigenschafts- und 9 Ergebnis-Kriterien	1	40	5.3
Lometsch (1975)	Erfahrung, sprachlicher Ausdruck, Sorgfalt, Leistungsmotivation, vorhandenes Potential, wahrscheinliche Endposition			2.6
Hossiep (1995)	Leistungs-, Sozialverhalten, Leistungsergebnis, Entwicklungspotential	1	13	33.3
Kersting (1999a)	Auffassungsgabe, Merkfähigkeit, Denk- und Urteilsfähigkeit, geistige Beweglichkeit, Entscheidungsfähigkeit, Aufgeschlossenheit, Einfallsreichtum, Selbständigkeit, konzeptionelles Arbeiten, organisatorische Fähigkeiten, schriftliche und mündliche Ausdrucksfähigkeit, Kontaktfähigkeit, Verhandlungsgeschick, Durchsetzungsvermögen, Kritikfähigkeit, Kooperationsfähigkeit, Führungsfähigkeit, Zuverlässigkeit, Belastbarkeit	1-2	20	3.6
Paff (1963)	Gesamtleistung, Intelligenz, Aufmerksamkeit	3	3	1.1
Blickle et al. (2005)	Produkt aus Qualität und Wichtigkeit für Kernleistung (Qualität, Quantität, umfassende Aufgabenerledigung), umfeldbezogene Leistung (Kooperativität, Zuverlässigkeit, Motivierung anderer, Regeleinhaltung) und adaptive Leistung (Anpassung, Krisenbewältigung, Innovation)	$M = 1.3$	20	4.0
Schmitz (2006)	Kundenorientierung, Kooperation/Teamfähigkeit, Planung/Organisation, Fachkompetenz, Qualitätsorientierung/Arbeitsorgfalt/Selbstkontrolle, soziale Belastbarkeit, Initiative/Erfolgsorientierung	o. A.	o. A.	1.5
Marcus et al. (2002)	Kernleistung (Qualität, Quantität) und umfeldbezogene Leistung (Engagement, Gewissenhaftigkeit, Kooperation)	o. A.	12	6.4

Anmerkungen. w_{BB} = Gewicht der Quelle in der bare-bone-Analyse zur Arbeitsleistung. o.A. = ohne Angabe.

Blickle et al., 2005), werden auch die Leistungsmotivation (z.B. bei Lometsch, 1975) und die zugeschriebene Intelligenz bzw. Facetten von Intelligenz eingeschätzt (z.B. bei Paff, 1963). In einigen Studien werden die Urteilsdimensionen nicht benannt (Brandstätter, 1970; Amthauer, 1973) oder es werden Globalurteile eingeholt (Althoff, 1968).

Globale Laufbahn

Ein in Bezug auf Datenart und -quelle insgesamt heterogenes Bild ergibt sich bei den Indikatoren der *globalen Laufbahn* (Tabelle 3.4), die zwei Bereichen zugeordnet wurden. Der erste Bereich enthält *Einkommensindikatoren*: das Einkommen zu einem bestimmten Zeitpunkt (aktuell, zu einem bestimmten Stichtag oder zu einem bestimmten Ereignis; Amelang & Steinmayr, 2006b; Blickle et al., 2006; Dahle & Erdmann, 2001; Lometsch, 1975; Nettelstroth, 2004), die Einkommensentwicklung (Nettelstroth, 2004) oder die tarifliche Eingruppierung (Hossiep, 1995; Steinmann, 1997). Diese Daten stammen aus Akten (Hossiep, 1995; Lometsch, 1975; Steinmann, 1997) oder Selbstberichten (Amelang & Steinmayr, 2006b; Blickle et al., 2006; Dahle & Erdmann, 2001; Nettelstroth, 2004).

Der zweite Bereich enthält Beförderung- und Positionsindikatoren. Er wird als *berufliches Vorankommen* bezeichnet und enthält im einzelnen: a) Neutrale Kennzahlen, die aktenbasiert sind: die Anzahl der erhaltenen Beförderungen und der Rangplatz für eine bevorstehende Beförderung (Lometsch, 1975), die Funktionsstufe (Hossiep, 1995), Führungsverantwortung und Teilnahme an einem Lehrgang für Führungskräfte (Steinmann, 1997) sowie den beruflichen Status (Linz, 1985). b) Neutrale Kennzahlen, die Fremdeinschätzungen entstammen, umfassen Position und Mitarbeiteranzahl (Kreuscher, 1987). c) Selbstberichtete neutrale Kennzahlen sind die Anzahl der Beförderungen (Nettelstroth, 2004) und Angaben zur Leitungsebene und Führungsverantwortung (Blickle et al., 2006). d) Auch Vergleiche mit Bezugsstandards werden inkludiert: berufliches Ansehen (Amelang & Steinmayr, 2006b) und Berufsstatus (Blickle et al., 2005). Sie stammen größtenteils aus Fremdbeurteilungen. In einer Studie (Blickle et al., 2006) werden Selbsteinschätzungen der Hierarchiestufe als Kriterium verwendet.

Kersting (1999a) weist darauf hin, dass die Interpretation der Korrelation zwischen Intelligenz und Laufbahnstatus als Maß für *prädiktive Validität* nicht zulässig ist, wenn das berufliche Vorankommen direkt vom Abschneiden in Intelligenztests abhängt. Dies ist in der Studie von Kersting (1999a) der Fall; diese Korrelation wurde deshalb exkludiert. Bei den anderen Studien gibt es keine Hinweise darauf, dass zur Entscheidung über das berufliche Vorankommen Intelligenztestungen durchgeführt wurden.

3.2.3. Stichprobe und Zeitabstand zwischen Intelligenz- und Kriterienerhebung

Folgende Kriterien sind in Bezug auf die untersuchte Stichprobe ausschlaggebend für die Inklusion einer Studie: a) Es handelt sich um Daten aus dem früheren Deutschen Reich oder der Bundesrepublik Deutschland. b) Die beurteilte Tätigkeit ist beruflicher Art, d.h. die untersuchten Personen sind Sich-Bewerbende, Auszubildende, Trainees oder Berufstätige. Studierende wurden nicht mit aufgenommen, um eine Vergleichbarkeit mit bereits durchgeführten Metaanalysen zu erreichen (vgl. Kapitel 2.2.1). Ferner wurden

Tabelle 3.4

Anzahl der inkludierten Studien mit verschiedenen Berufserfolgs-Indikatoren

Datenart	Datenquelle		
	Akten/Dokumente	Fremdbeurteilung	Selbstbericht
Erfolgsbereich <i>Lernen</i>			
Vergleich mit Bezugsstandards		210 Benotungen oder betriebl. Beurteilungen	
Erfolgsbereich <i>Arbeit</i>			
Vergleich mit Bezugsstandards		18 Abwärts- und Seitwärtsbeurteilungen	
Erfolgsbereich <i>Karriere</i>			
neutrale Kennzahlen	3 Einkommen 4 Vorankommen	1 Vorankommen	4 Einkommen 2 Vorankommen
Vergleich mit Bezugsstandards		2 Vorankommen	1 Vorankommen

die Studien einer Quelle (Kirsch, 1961) exkludiert, die Korrelationen zwischen Intelligenz und dem Erfolg in einer Pilotenausbildung berichten, da es sich möglicherweise um Hobbypiloten handelt. Weitere Einschränkungen in Bezug auf die Stichprobe wurden nicht vorgenommen.

Von besonderer Bedeutung bei eignungsdiagnostischen Verfahren sind *prädiktive* Validitäten, die auch den meisten gefundenen Studien zu Grunde liegen (207 von 227). In 13 Studien werden konkurrente Validitäten berichtet und in zwei Studien retrograde. Einige Studien berichten unterschiedliche Validitätskoeffizienten. In diesen Fällen werden prädiktive gegenüber konkurrenten bevorzugt und konkurrente gegenüber retrograden. Liegen prädiktive Validitäten zu mehreren Messzeitpunkten vor, die an einer Stichprobe erhoben wurden, werden diese zusammengefasst.

3.2.4. Effektgrößen

Zur metaanalytischen Integration eignen sich prinzipiell unterschiedliche Effektgrößen, die quantitative Angaben über die Relation von Intelligenz und Leistungskriterien machen. Das verwendete Programm von F. L. Schmidt und Le (2005) erlaubt die Berechnung entweder auf der Basis von d -Werten oder „Pearson r s“ (Hunter & Schmidt, 2004, S. 195). In der Fragestellung dieser Arbeit wird ein Zusammenhang postuliert, dies spricht für die Verwendung der Korrelation. Sie ist angemessen, wenn von einem linearen Zusammenhang zwischen Intelligenz und Leistung ausgegangen werden kann, was laut Jensen (1998) der Fall ist.

Pearson r s sind nach Hunter und Schmidt Pearson Produkt-Moment-Koeffizienten kontinuierlicher Variablen oder ordinaler Variablen (Spearman's ρ). Andere Korrelati-

onsmaße weisen höhere Standardfehler auf, die zu einer Überschätzung von \widehat{SD}_ρ führen. Hunter und Schmidt (2004) raten deshalb – wenn möglich – von ihrer Verwendung ab und schlagen vor, betroffene Studien zu exkludieren. Dies betrifft im vorliegenden Fall zwei Studien, die t -Werte (Steinmann, 1997) bzw. Kontingenzkoeffizienten (Dreyer & Grabitz, 1971) berichten. Beide Studien werden inkludiert; der Einfluss der Inklusion wird sensitivitätsanalytisch geprüft (Kapitel 4.2.3).

Abgesehen von diesen zwei Studien werden Pearson r s mitgeteilt, die i.d.R. unkorrigiert sind. In zwei Studien (Schuler, Funke et al., 1995; Schuler, Moser et al., 1995) werden für das Kriterium messfehlerkorrigierte Korrelationen berichtet. Diese Korrekturen werden rückgängig gemacht und die unkorrigierten Korrelationen zur weiteren Berechnung verwendet.

Sollen t -Werte berücksichtigt werden, empfehlen Hunter und Schmidt (2004) folgendes Vorgehen: Zunächst werden die t -Werte nach Glass, McGaw und Smith (1984) in punktbiseriale Korrelationen transformiert. Die punktbiseriale Korrelation wird allerdings beeinflusst von ungleichen Gruppengrößen. Die Korrektur dieses statistischen Artefaktes erfolgt nach Hunter und Schmidt (2004) mit Formel A.1. Diese Korrektur wirkt sich vergrößernd auf die Standardfehler aus, was nicht erwünscht ist. Deshalb ist zusätzlich eine nachträgliche Korrektur der Standardfehler erforderlich, die über eine Reduktion von n gemäß Formel A.2 erreicht wird. Auf diese Weise können Ergebnisse von Steinmann (1997) berücksichtigt werden.

Kontingenzkoeffizienten C werden von Dreyer und Grabitz (1971) berichtet. Sie entsprechen laut Kendall und Stuart (nach Glass et al., 1984) ungefähr der Produkt-Moment-Korrelation. Ihr Maximalwert (C_{\max}) liegt jedoch nicht bei 1, sondern – abhängig von der Anzahl der Kategorien der zugrunde liegenden Kreuztabellen – darunter. C_{\max} beträgt bei Dreyer .89.

Eine Korrektur dafür wird von Dreyer und Grabitz durchgeführt. Dabei gilt $C_{\text{korrr}} > C$. Diese Korrektur ist nicht unproblematisch, da auch andere Korrelationsmaße nur unter bestimmten Bedingungen den Wert ± 1 annehmen können. Darauf weist Bortz (2005) in Bezug auf den Φ -Koeffizienten hin, in den C transformiert werden kann (Diehl & Arbinger, 2001). Wirtz und Caspar (2002) halten das Assoziationsmaß V von Cramer für „besser geeignet zum Vergleich mit anderen Korrelationskoeffizienten“ (S. 55). Es resultiert in deutlich geringeren Korrelationen als C bzw. C_{korrr} . Auf Grund dieser uneindeutigen Sachlage wird der „mittlere“ Weg gewählt, indem das unkorrigierte C , wie von Glass et al. (1984) vorgeschlagen, inkludiert wird.

In der ältesten inkludierten Quelle (Roloff, 1928) werden Rangkorrelationen nach Spearman berichtet, die das Vorliegen von Ties *nicht* berücksichtigen. Dies ist nach Diehl und Kohr (2004) nicht zulässig. Da die Rohdaten in der Quelle enthalten sind, konnten die Rangkorrelationen mit Berücksichtigung der Ties neu berechnet werden (Diehl & Kohr, 2004). Nach diesen Transformationen liegen ausschließlich korrelative Effektgrößen vor.

3.2.5. Studienqualität

Es findet vor der Kodierung keine Exklusion in Bezug auf die Qualität der Primärstudien statt, da nahezu ausschließlich Korrelationsstudien inkludiert werden, die generell im Vergleich zum wissenschaftlichen „Goldstandard“ randomisierter und kontrollierter Experimente weniger valide Aussagen erlauben (Rustenbach, 2003).

Die Qualität der einzelnen Studien unterscheidet sich im Hinblick auf die Vollständigkeit und Detailliertheit der Berichterstattung, die Stichprobengrößen, die Varianzeinschränkungen in Prädiktor (Vorauswahl) und Kriterium (z.B. drop outs) sowie die Güte der verwendeten Prädiktoren und Kriterien. Letzteres umfasst die psychometrischen Gütekriterien, die Aktualität und Bevölkerungsrepräsentativität der Normen sowie die Relevanz der verwendeten Maße.

Die Stichprobengröße, Varianzeinschränkung im Prädiktor sowie die Messfehler in Prädiktor und Kriterium wurden in den Analysen berücksichtigt, indem sie zur Gewichtung der Studien herangezogen wurden: je akkurater und weniger von diesen Artefakten beeinflusst eine Studie ist, desto größer ist ihr Gewicht (Kapitel 3.8). Eine Studie, die in Kersting (1999c), Beauducel und Kersting (2002) sowie Kersting (2004) veröffentlicht wurde, weist eine sehr geringe Intelligenzvarianz auf. Diese Studie konnte nicht berücksichtigt werden, da die Varianzeinschränkung über dem Maximum liegt, das korrigiert werden kann (vgl. Kapitel 3.6.3).

Fehlende Kennwerte, die zur Berechnung der Metaanalyse erforderlich sind, wurden – wenn möglich – berechnet (z.B. aggregierte Korrelationen, wenn mehrere abhängige Werte berichtet werden, vgl. Kapitel 3.5) oder geschätzt (nicht signifikante Korrelationen, die nicht berichtet werden, vgl. Kapitel 3.7; Reliabilitäten von Prädiktoren und Kriterien sowie Varianzeinschränkungen im Prädiktor, vgl. Kapitel 3.6).

Die Einflüsse unzureichender Aktualität und Akkuratheit der Normdaten von Intelligenz, unzureichender Relevanz der Intelligenz- und Kriterienmaße sowie der Einfluss von Varianzeinschränkungen im Kriterium können nicht quantifiziert werden.

3.3. Statistische Unabhängigkeit

In vielen Quellen wird mehr als eine Effektstärke zwischen Intelligenz und Leistung berichtet. Wenn diese von *einer* Stichprobe stammen, sind sie zu einem Gesamteffekt zu kumulieren (vgl. Kapitel 3.5), anderenfalls sollten sie nach Hunter und Schmidt (2004) als getrennte Studien in der Metaanalyse berücksichtigt werden. Der Forderung, nur einen Kennwert pro Quelle zuzulassen, erteilen sie eine Absage: „Statistically, outcome values for nonoverlapping groups have the same properties as values from different studies“ (S. 440).

In einigen Studien werden Effektstärken für Teilstichproben berichtet. Diese werden dann als getrennte Studien behandelt, wenn Moderator- oder Sensitivitätsanalysen eine unterschiedliche Behandlung der Teilstichproben erforderlich machen. Dies ist z.B. der Fall, wenn die Teilstichproben unterschiedliche Berufsgruppen darstellen (z.B. bei A. A. Schmitt, 2006). Eine getrennte Berücksichtigung beider Geschlechter ist dagegen nicht intendiert, deshalb werden keine getrennten Effektstärken für Männer und Frauen

berücksichtigt, wenn die Effektstärke auch für beide Geschlechter zusammen vorliegt (wie in Brandstätter, 1970).

Neben der Aufteilung einer Quelle in mehrere Studien ist auch der umgekehrte Fall erforderlich: die Zusammenfassung mehrerer Quellen zu einer Studie, wenn die in den verschiedenen Quellen berichteten Werte aus denselben Stichproben stammen. Dadurch wird die mehrfache Berücksichtigung ein und derselben Stichprobe vermieden (*duplication bias*). Um dies zu gewährleisten, wurden die Stichprobenbeschreibungen der Studien miteinander verglichen, die vom selben Autor veröffentlicht oder mit veröffentlicht wurden. Tabelle C.1 gibt darüber Auskunft, welche Studien aus mehreren Quellen stammen.

3.4. Kodierung

Alle Studien wurden von zwei unabhängigen Beurteilern kodiert. Ein Beurteiler war der Autor dieser Arbeit, die zweite Beurteilung wurde von einer studentischen Hilfskraft der Abteilung Arbeits-, Organisations- und Wirtschaftspsychologie an der Universität Bonn vorgenommen. Kodiert wurde in Microsoft Excel. Das der Kodierung zu Grunde liegende Schema enthält Anhang B.1. Im Folgenden werden die kodierten Variablen getrennt für allgemeine Studieninformationen, Prädiktor-, Kriterieninformationen und Informationen zu dem Zusammenhang zwischen beiden kurz dargestellt und Besonderheiten besprochen. Anschließend wird die Beurteilerübereinstimmung quantifiziert.

3.4.1. Allgemeine Informationen

Die kodierten allgemeinen Informationen wurden pro Studie kodiert. Sie umfassen eine fortlaufende Studiennummer, die Autoren, das Veröffentlichungsjahr bzw. das Durchführungsjahr unveröffentlichter Studien (PJAHR), die Art der Veröffentlichung (PART: 1 *nicht veröffentlicht*, 2 *Dissertation*, 3 *Zeitschriftenartikel*, 4 *Monographie ohne Testmanuale*, 5 *Beitrag in Herausgeberwerk*, 6 *Testmanual*, 9 *unklar*), das Geschlecht der Stichprobe (SEX: 1 *weiblich*, 2 *männlich*, 3 *gemischt*, 9 *unklar*), der Anteil der Frauen an der Stichprobe in % (WEIBL), Mittel und Standardabweichung des Alters der Stichprobe zum Zeitpunkt der Prädiktorerhebung (MAGEX, SDAGEX), das Bildungsniveau, den Beruf wie bezeichnet (BERUF), die ISCO-88-Kodierung des Berufes (ISCO88) und den ISCO-88-Skill-Level des Berufes (ISCOSL), die Berufserfahrung (ERF; nicht bei Studien zur berufsbezogenen Lernleistung) und das Design der Studie (VTYP: Längsschnitt vs. Querschnitt, mit oder ohne Einfluss der Intelligenzdaten auf das Kriterium).

Die Kodierung der Berufe hat das Hinzuziehen von Berufssystematiken erfordert. Genutzt wurde dazu eine deutsche Adaptation der ISCO-88 des Zentrums für Umfragen, Methoden und Analysen, Mannheim (ZUMA; Geis, 2006) sowie – wenn erforderlich – die Originalversion der ISCO-88 (International Labour Office, 1990). Berufsbezeichnungen, die heute nicht mehr geläufig sind (z.B. Bohrer) wurden zunächst mit Hilfe der BERUFENET-Datenbank der Bundesagentur für Arbeit (2008b) nach der *Klassifizierung der Berufe - Systematisches und alphabetisches Verzeichnis der Berufsbenennungen* kodiert (KldB88; Statistisches Bundesamt, 1988). Die KldB88-Kodes wurden dann mit

Tabelle 3.5
ISCO-88 Berufshauptgruppen und Skill Level

Hauptgruppen	Skill Level
1 Angehörige gesetzgebender Körperschaften und leitende Verwaltungsbedienstete	–
2 Wissenschaftler	4
3 Techniker und gleichrangige nichttechnische Berufe	3
4 Bürokräfte, kaufmännische Angestellte	2
5 Dienstleistungsberufe, Verkäufer in Geschäften und auf Märkten	2
6 Fachkräfte in der Landwirtschaft und Fischerei	2
7 Handwerks- und verwandte Berufe	2
8 Anlagen- und Maschinenbediener sowie Montierer	2
9 Hilfsarbeitskräfte	1
0 Streitkräfte	–

Quelle. Geis (2006, S. 34-49).

Hilfe eines Umsteigeschlüssels von Geis (2006) in ISCO-88-Kodes überführt. Dabei war zu beachten, dass dem Umsteigeschlüssel eine neuere KldB-Version von 1992 zu Grunde liegt. Die Kodierung war deshalb zusätzlich mit weiteren Kodieranweisungen von Geis (2006) und des [International Labour Office \(1990\)](#) abzusichern.

Der ISCO-Kode besteht aus vier Stellen. Die erste Stelle gibt die Berufshauptgruppe an, die drei folgenden Stellen spezifizieren die Berufsgruppen in Sub-Hauptgruppen (2. Stelle), Minorgruppen (3. Stelle) und Unitgruppen (4. Stelle) weiter. Kodiert wurden, wenn möglich, alle vier Stellen. Wenn dies nicht möglich war, wurden die Stellen der unklaren Unit-, Minor- und Sub-Haupt-Gruppen mit Null aufgefüllt. War auch die Berufshauptgruppe nicht klar zuzuordnen, wurde kein ISCO-Kode vergeben. Alle Berechnungen (auch die Beurteilerübereinstimmungen) basieren auf den einstelligen Berufshauptgruppen. Unterschieden werden zehn Berufshauptgruppen, die vier Skill Levels zugeordnet werden können (Tabelle 3.5). Die vollständige Kodierung der Berufe enthalten die Tabellen C.5 bis C.7.

Ein Vorteil der ISCO-88-Kodierung gegenüber der KldB-Kodierung besteht in dieser Operationalisierung von Berufskomplexität. Skills werden in der ISCO-88 definiert als „ability to carry out the tasks and duties of a given job“ ([International Labour Office, 1990](#), S. 2), Skill Levels als „function of the complexity and range of the tasks and duties involved“ (S. 2). Da die ISCO-Kodierung zur internationalen Verwendung entwickelt wurde, werden vier breite Skill Levels auf Basis der *International Standard Classification of Education* (ISCED; UNESCO 1976 nach [International Labour Office, 1990](#)) verwendet:

- (a) *The first ISCO skill level* was defined with reference to ISCED category 1, comprising primary education which generally begins at the age of 5, 6 or 7 and lasts about five years.
- (b) *The second ISCO skill level* was defined with reference to ISCED categories 2 and 3, comprising first and second stages of secondary edu-

cation. The first stage begins at the age of 11 or 12 and lasts about three years, while the second stage begins at the age of 14 or 15 and also lasts about three years. A period of on-the-job training and experience may be necessary, sometimes formalised in apprentice-ships. This period may supplement the formal training or replace it partly or, in some cases, wholly.

- (c) *The third ISCO skill level* was defined with reference to ISCED category 5, (category 4 in ISCED has been deliberately left without content) comprising education which begins at the age of 17 or 18, lasts about four years, and leads to an award not equivalent to a first university degree.
- (d) *The fourth ISCO skill level* was defined with reference to ISCED category 6 and 7, comprising education which also begins at the age of 17 or 18, lasts about three, four or more years, and leads to a university or postgraduate university degree, or the equivalent. (S. 2-3)

3.4.2. Prädiktor

Die kodierten Informationen zum Prädiktor wurden pro berichteter Korrelation kodiert. Dies sind: die Gruppe des Intelligenztests (IQ: 1 *IST*, 2 *WIT*, 3 *BIS*, 6 *LPS/PSB*, 7 *Matrizentests*, 8 *GATB-basierte Verfahren*, 9 *andere*; die Kategorien 4 *BOMAT* und 5 *HAWIE* wurden nicht benötigt), Mittelwert (PM) und Standardabweichung des Prädiktors in der Stichprobe (PSD), Informationen zur Varianzeinschränkung im Prädiktor, z.B. die Selektionsquote (PVE), Angaben zu Korrelationen zwischen den Prädiktoren (PIK) und zur Reliabilität der Prädiktoren, die auf der Stichprobe basiert (PREL).

Zur Entscheidung, ob einzelne Testverfahren als Teil der Operationalisierung von Intelligenz angesehen werden können oder nicht, wurden darüber hinaus die Aufgaben- gruppen der verschiedenen Intelligenztests den Structupeln des BIS-Modells zugeordnet (vgl. Kodierbogen in Anhang B.2). Diese Zuordnung erfolgte theoriegeleitet, d.h. durch Vergleiche der Testbeschreibungen bzw. -items mit den Beschreibungen der BIS-Facetten (Jäger et al., 1997). Sofern vorhanden, wurden empirische Ergebnisse zum Zusammenhang von Testaufgaben mit BIS-Facetten ergänzend berücksichtigt (z.B. Kersting, 1999a; Knebelau, 2002; Jäger & Althoff, 1994). Dennoch haben die getroffenen Zuordnungen (Tabelle C.4) hypothetischen Charakter.

In einigen Fällen war keine eindeutige Zuordnung möglich, z.B. wenn eine Aufgaben- gruppe aus verschiedenen BIS-Inhalten besteht. In diesen Fällen wurde eine Mehrfach- zuordnung vorgenommen. Die Abgrenzung der Geschwindigkeitsoperation (B) von den anderen drei BIS-Operationen, insbesondere von Verarbeitungskapazität (K) kann letztlich nur empirisch geklärt werden. In Bezug auf den Faktor *Broad Visual Perception* diskutiert das z.B. Carroll (1993).

Zur Zuordnung von B und K wurde deshalb ergänzend eine Unterscheidung eingeführt, die Süß, Oberauer, Wittmann, Wilhelm und Schulze (2002) zwischen Aufgaben treffen,

die die Arbeitsgedächtnis-Kapazität messen (Bearbeitungsgeschwindigkeit) und solchen, die schlussfolgerndes Denken messen (Verarbeitungskapazität):

In our view, “reasoning” refers to a relatively complex cognitive process consisting of an organized sequence of steps that must be appropriately scheduled in order to reach a goal. . . . Working”=memory tasks, in contrast, require only simple cognitive operations. The sequence of operations is highly restricted by the instructions and the task display; only a minimum amount of freedom remains for participants to decide by themselves what to do next. (S. 267)

Die Geschwindigkeitsfacette (B) wurde kodiert, wenn es um reine Routineaufgaben geht, bei denen kaum Anforderungen i.S.d. Verarbeitungskapazität (K) gestellt sind. Sobald z.B. bei numerischen Aufgaben mehrstufiges mathematisches Regelwissen angewendet werden kann, wurde die Aufgabe der K-Facette des BIS-Modells zugeordnet.

Zur Berechnung der Beurteilerübereinstimmung wurden alle acht möglichen Inhaltskombinationen und alle 16 möglichen Operationskombinationen in nominal skalierte Variablen überführt. Die Kodierung gelang mit hoher Übereinstimmung zwischen den Beurteilern: κ betrug für die Inhalte .98 und für die Operationen .97. Eine Übersicht über die Zuordnung der inkludierten Aufgaben zu den BIS-Strukturpeln gibt Tabelle C.4.

3.4.3. Kriterium

Zum Kriterium wurden pro berichteter Korrelation zunächst der Erfolgsbereich der Kriterienmessung erfasst (KINHALT: 1 *Arbeitsleistung oder Karriereerfolg*, 2 *Ausbildungsleistung*, 3 *Weiterbildungsleistung*, 8 *anderes, Kombination*, 9 *unklar*). Die kodierten Erfolgsindikatoren (KMM) spezifizieren die Kriterien, die verwendet wurden: subjektive Indikatoren sind 0 *schulische Abwärtsbeurteilung*, 1 *betriebliche Abwärtsbeurteilung*, 2 *Seitwärtsbeurteilung* und 3 *andere subjektive Leistungsbeurteilung*, objektive Indikatoren sind 4 *Fehlzeiten*, 5 *Produktionsdaten*, 6 *Einkommen*, 7 *Beförderungen* und 8 *andere objektive Maße*. Weitere Kriterienvariablen, die kodiert wurden, sind Mittel und Standardabweichung des Kriteriums (KM, KSD), Informationen zur Varianzeinschränkung im Kriterium, z.B. die *drop out*-Rate (KVE), Informationen zur Korrelation zwischen verschiedenen Kriterien (KIK) und Informationen zur Reliabilität der Kriterien, die auf der Stichprobe basieren (KREL).

Zusammenhang zwischen Prädiktor und Kriterium

Schließlich wurden zum Zusammenhang zwischen Prädiktor und Kriterium pro berichteter Korrelation kodiert: der Zeitabstand zwischen Prädiktor- und Kriterienerhebung in Wochen (TDIFF) sowie ein Konfidenzrating dazu (TDIFFKR, s.u.), die beobachtete Validität (V) und die ihr zu Grunde liegende Stichprobengröße (VN), der Typ der beobachteten Validität (VKTYP: Pearson r_s , nicht-Pearson r_s und andere Effektmaße vgl. Anhang B.1) mit näheren Angaben beim Vorliegen korrigierter Effektmaße (KTINFO).

Bei der ersten Sichtung einiger Studien durch den Autor dieser Arbeit wurde bereits deutlich, dass insbesondere die Angaben zur Zeitdifferenz zwischen Prädiktor- und Kriterienerhebung häufig vage oder sogar widersprüchlich angegeben sind. Aus diesem Grund wurde die Kodierung der Zeitdifferenz mit einem vierstufigen Konfidenzrating versehen (TDIFFKR: 1 *unsicher*, 2 *wahrscheinlich*, 3 *ziemlich sicher*, 4 *sicher*) und nur dann ausgewertet, wenn eine ziemlich sichere oder sichere Einschätzung möglich war.

Beurteilerübereinstimmung

Zur Einschätzung der Beurteilerübereinstimmung wurde für dichotome und polytome Daten Cohens κ berechnet. Die Beurteilerübereinstimmung ordinaler, intervallskalierter und absoluter Daten wird mit ICC_{unjust} angegeben (Wirtz & Caspar, 2002). ICC_{unjust} setzt voraus, dass alle Beurteilungseinheiten von allen Beurteilern kodiert werden. Dies ist im vorliegenden Fall nicht gegeben, da fehlende Werte i.d.R. nicht kodiert wurden. Dadurch ist es möglich, dass ein Beurteiler einen Wert kodiert hat, ein anderer Beurteiler nicht. Deshalb wurden zusätzlich zu ICC_{unjust} auch κ -Koeffizienten berechnet, die angeben, wie groß die Übereinstimmung darüber ist, ob eine Beurteilungseinheit kodiert oder nicht kodiert wurde (Entdeckungsübereinstimmung). Die ICC_{unjust} -Koeffizienten wurden an den zugehörigen κ s relativiert.

Die genannten Gütemaße sind unter .40 inakzeptabel, zwischen .40 und .60 mäßig, zwischen .60 und .75 gut und ab .75 exzellent (Orwin, 1994, nach Rustenbach, 2003). Rustenbach führt weiter aus: „In der Integrationsforschung sind generell sehr hohe Interraterreliabilitäten anzustreben: Ein exzellenter Koeffizient erscheint notwendig“ (S. 51).

Die erzielten Beurteilerübereinstimmungen liegen zwischen .722 bei der Einschätzung der Berufserfahrung und 1.000 beim Publikationsjahr (Tabelle 3.6). Die mittlere Beurteilerübereinstimmung beträgt .933. Dies kann als sehr gute Übereinstimmung angesehen werden. Bei nicht übereinstimmenden Kodierungen wurde versucht, einen Konsens zwischen den Beurteilern zu erzielen, was in allen Fällen möglich war. Bei intervallskalierten Variablen liegt κ jeweils unter ICC_{unjust} . Das bedeutet, dass die Schwierigkeit bei der Kodierung eher darin bestand, die zu kodierenden Werte zu entdecken als sie fehlerfrei zu übertragen.

3.5. Kumulierung abhängiger Werte

In zahlreichen Studien wird nicht nur *eine* Effektstärke berichtet, sondern mehrere statistisch abhängige, d.h. an derselben Stichprobe erhobene. In vielen Fällen ließ sich durch die Anwendung der Inklusions- bzw. Exklusionsregeln bestimmen, welche Effektstärke zu verwenden ist (vgl. Kapitel 3.2).

Es gibt aber Fälle, in denen mehrere Effektstärken simultan berücksichtigt werden mussten. Dies war der Fall, wenn für eine Stichprobe die Ergebnisse für mehrere Intelligenztests, mehrere Intelligenzfacetten oder mehrere Leistungskriterien vorlagen. Auch bei Messwiederholungen war eine Aggregation möglich.

Tabelle 3.6
Beurteilerübereinstimmung

Variable (Bezeichnung in der Kodieranweisung)	Cohens κ	ICC _{unjust}	$\kappa * ICC_{unjust}$
Studienebene	$n = 229$		
Publikationsjahr (PJAHR)	–	1.000	1.000
Publikationsart (PART)	.964	–	.964
Geschlecht (SEX)	.940	–	.940
Frauenanteil (WEIBL)	.923	1.000	.923
Mittleres Alter (MAGEX)	.932	1.000	.932
SD Alter (SDAGEX)	.923	1.000	.923
Bildungsniveau bis Hauptschulabschluss (SBNIV1)	.920	1.000	.920
Bildungsniveau mittlere Reife (SBNIV2)	.928	1.000	.928
Bildungsniveau (Fach-)Hochschulreife (SBNIV3)	.928	1.000	.928
ISCO88 Hauptgruppe (aus ISCO88)	.985	.999	.984
ISCO88 Skill Level (ISCOSL)	.968	–	.968
Erfahrung (ERF)	.722	1.000	.722
Design (VTYP)	.820	–	.820
Korrelationsebene	$n = 3\,287$		
Intelligenztest (IQ)	.890	–	.890
Mittel der Intelligenzwerte (PM)	.977	1.000	.977
SD Intelligenz (PSD)	.981	1.000	.981
Erfolgsbereich (KINHALT)	.982	–	.982
Erfolgsindikator (KMM)	.912	–	.912
Mittel der Kriterien (KM)	.995	.996	.991
SD der Kriterien (KSD)	.975	1.000	.975
Zeitabstand Prädiktor-Kriterium (TDIFF)	.870	.928	.807
Validitätskoeffizient (V)	.986	.985	.971
n (VN)	.908	.999	.907
Art des Validitätskoeffizienten (VKTYP)	.969	–	.969
Ebene der Aufgabengruppen der Intelligenztests	$n = 301$		
BIS-Operationen	.969	–	.969
BIS-Inhalte	.981	–	.981
M	.934	.994	.933

Anmerkung. Variablen, für die Cohens κ und ICC_{unjust} angegeben sind, sind intervallskaliert. κ bezieht sich in diesen Fällen auf die dichotomisierten Variablen: kodiert vs. nicht kodiert (Entdeckungsübereinstimmung).

In jeder Metaanalyse konnte pro Studie nur *eine* Effektstärke berücksichtigt werden, um unterschiedliche Gewichtungen zu vermeiden, die allein auf Grund einer unterschiedlichen Anzahl berichteter Effektgrößen zu Stande kommen würden (Hunter & Schmidt, 2004). Deshalb war es erforderlich, abhängige Ergebnisse in einer Studie zu einem Gesamtwert zu aggregieren. Die Aggregation ist dann zulässig, wenn die einzelnen Maße als Teilaspekte eines Konstruktes angesehen werden können, sie also nicht zu heterogen sind. Diese Anforderung kann in Bezug auf die Intelligenzmaße als gegeben erachtet werden: Intelligenztests sowie Aufgabengruppen aus Intelligenztests sind untereinander hoch korreliert (vgl. Kapitel 2.1). Anders liegt der Fall bei verschiedenen Leistungsmaßen (vgl. Kapitel 2.2), die deshalb die Exklusionen bestimmter Leistungsaspekte (vgl. Kapitel 3.2.2) bzw. die getrennte Betrachtung in verschiedenen Metaanalysen erforderlich machen.

Die Kumulierung kann auf zwei Arten erfolgen. Eine Möglichkeit stellt die einfache Mittelung der Korrelationen dar. Sie liefert jedoch nur dann eine angemessene Schätzung der Gesamtkorrelation, wenn die einzelnen Variablen (also die einzelnen Prädiktoren bzw. die einzelnen Kriterien) *nicht* miteinander korreliert sind (Hunter & Schmidt, 2004). Davon kann – wie eben ausgeführt – i.d.R. nicht ausgegangen werden, da unter diesen Bedingungen eine Aggregation nicht angezeigt wäre. Eine konstruktvalidere Aggregation erreicht man mittels linearer Verbünde (*linear composites*; Hunter & Schmidt, 2004), die die Korrelationen zwischen den einzelnen Variablen berücksichtigen (Formel A.3; Hunter & Schmidt, 2004).

Sind die einzelnen Korrelationen zwischen den Prädiktoren bzw. Kriterien nicht bekannt, können sie evtl. anderen Quellen entnommen werden. Für die *Prädiktoren* kann dazu auf andere Studien, die dieselben Prädiktoren verwenden, sowie die Manuale der Intelligenztests zurückgegriffen werden (Hunter & Schmidt, 2004). Beim Vorliegen mehrerer *Kriterien*, deren Interkorrelationen unbekannt sind, wurde auf die Schätzung der Korrelation durch andere Quellen verzichtet, da die Kriterien weitaus heterogener und weniger standardisiert sind als die Prädiktoren.

Ist die Bildung linearer Verbünde nicht möglich, wird die Aggregation i.d.R. durch einfache Mittelung der Korrelationen durchgeführt (z.B. bei Hülshager et al., 2007). Alternativ dazu können die betroffenen Studien exkludiert werden (dies tun z.B. F. L. Schmidt, Hunter & Caplan, 1981). Bei der einfachen Mittelung wird eine Unterschätzung des mittleren Effektes in Kauf genommen (Hunter & Schmidt, 2004), bei der Exklusion entfernt man sich vom Ideal der Vollerhebung aller relevanten Studien. In dieser Arbeit wurden die Studien deshalb mit einfacher Mittelung berücksichtigt und ihr Einfluss sensitivitätsanalytisch durch Exklusion geprüft.

Bei der Aggregatbildung stellte sich auch die Frage, mit welcher Gewichtung die einzelnen berichteten Korrelationen berücksichtigt werden sollen. Auf Prädiktorseite wurden die Gewichte so gewählt, dass a) Subfacetten von Intelligenztests die Gewichte erhalten, die im Testmanual bei Berechnung des GMA-Maßes vorgesehen sind und b) die GMA-Maße mehrerer Intelligenztests gleichgewichtig berücksichtigt wurden. Bei den Kriterien wurde eine Gleichgewichtung verschiedener Urteilsquellen (z.B. Berufsschulnoten und betriebliche Beurteilungen) angestrebt.

Die Gewichtung bei der Aggregatbildung über lineare Verbünde erfolgte gemäß Formel A.4, bei der Mittelung wurde das gewichtete Mittel gebildet. Welche Art der Aggregation pro Studie durchgeführt wurde und ob dabei Gewichtungen vorgenommen wurden, kann für die einzelnen Metaanalysen den Tabellen E.1 bis E.42 entnommen werden.

3.6. Artefaktkorrekturen

Eine Besonderheit der Vorgehensweise von Hunter und Schmidt (2004) stellen Artefaktkorrekturen dar. Artefakte sind Mängel von Studien, die sich auf die Ergebnisse auswirken: „We refer to study imperfections as ‘artifacts’ to remind ourselves that errors in study results produced by study imperfections are artifactual or man-made errors and not properties of nature“ (Hunter & Schmidt, 2004, S. 33). Hunter und Schmidt unterscheiden elf mögliche Artefakte: Stichprobenfehler, Unreliabilität in Prädiktor oder Kriterium, Dichotomisierung kontinuierlicher Prädiktoren oder Kriterien, Varianzunterschiede zwischen der Stichprobe und der Population im Prädiktor oder Kriterium (i.d.R. handelt es sich um *Varianzeinschränkungen*), imperfekte Konstruktvaliditäten von Prädiktor oder Kriterium, Report- bzw. Transkriptfehler und der Einfluss nicht berücksichtigter Drittvariablen. Die Artefakte führen zu einer Verringerung der beobachteten Effekte im Vergleich zum wahren Effekt auf Konstruktebene, mit Ausnahme der Varianzvergrößerungen, die zu einer Vergrößerung der beobachteten Effekte führen. Die Einflüsse der Artefakte können korrigiert werden, wenn sie in ihrer Höhe bekannt sind oder geschätzt werden können. Dies ist in der vorliegenden Arbeit für den Stichprobenfehler, die Unreliabilitäten in Prädiktor und Kriterium sowie die Varianzeinschränkung im Prädiktor der Fall.

Der Stichprobenfehler wird bei der Vorgehensweise nach Hunter und Schmidt (2004) durch die Gewichtung der Studien mit n korrigiert. Bei aggregierten Effektstärken, die auf unterschiedlichen n basieren, wird das maximale n als Gewichtungsfaktor verwendet. Dies ist erforderlich, um Studien, die mehrere Korrelationen berichten, nicht zu benachteiligen. Eine Mittelung der Stichprobengrößen hätte zur Folge, dass die betroffene Studie mit einem kleineren n berücksichtigt wird als beim Vorliegen nur einer Korrelation mit großem n .

3.6.1. Unreliabilität der Prädiktoren

Die Korrektur der Unreliabilität geht zurück auf Spearman (1904; nach Ree & Carretta, 2002). Sie wird vorgenommen, indem die unkorrigierte Korrelation durch die Wurzel aus der Reliabilität dividiert wird (Hunter & Schmidt, 2004).

Eine vollständige Korrektur der Unreliabilität der Prädiktoren umfasst die drei Fehlerquellen *specific error*, *random response error* und *transient error* (Hunter & Schmidt, 2004). Die ersten beiden Fehlerquellen können mit einem Paralleltest-Design (*coefficient of equivalence*, CE; Cronbach, 1947), die letzten beiden mit einem Retest-Design (*coefficient of stability*, CS) quantifiziert werden. Zur vollständigen Reliabilitätskorrektur werden Reliabilitätswerte aus einem kombinierten Paralleltest-Retest-Design (*coefficient*

of equivalence and stability, CES) benötigt, die nur selten berichtet werden (Hunter & Schmidt, 2004). Im Rahmen dieser Arbeit werden sie in zwei Quellen genannt (Amthauer, 1953b; Marschner, 1983b). Von der Verwendung von CS- oder CE-Reliabilitäten raten Hunter und Schmidt (2004) ab: „Use of inappropriate types of reliability estimates leads to incomplete corrections for measurement error“ (S. 138).

Laut F. L. Schmidt, Le und Ilies (2003) ist es jedoch für kognitive Leistungstests möglich, CES auf Basis von CE nach der Formel $CES = CE - .05$ zu bestimmen. Dieses Vorgehen zur Ermittlung von CES empfehlen auch Hunter und Schmidt (2004): „CE estimates can be adjusted or corrected for transient error by subtracting .04 or .05 from these reliability estimates“ (S. 102-103). Für CE können Testhalbierungs-, Paralleltest-Reliabilitäten und interne Konsistenzen verwendet werden. Cronbachs α wird z.B. von F. L. Schmidt et al. (2003) genutzt. Wenn möglich wird α den Testmanualen entnommen (Amthauer, 1973; Amthauer et al., 2001; Jäger et al., 1997; Horn, 1962; Schuler & Klingner, 2005; Jäger & Althoff, 1983), anderenfalls werden Testhalbierungs-Reliabilitäten verwendet (Weiß, 1980) oder es wird auf weitere Literatur zurückgegriffen (Brocke et al., 1998; Gaschok, 2002; Engelbrecht, 1975; Kersting & Beauducel, 1997; Kersting et al., 2005; Nettelstroth, 2003; Schmidt-Atzert et al., 1995; Schuler & Klingner, 2002; Tent, 1969; Thielepape & Kersting, 2005b).

Zur Einschätzung der *Reliabilität von Prädiktorverbänden* ist ausschlaggebend, ob die prädiktorspezifischen Varianzanteile als Fehlervarianz anzusehen sind. Wird dies bejaht, entspricht die Reliabilität des Prädiktorverbundes Cronbachs α der Einzelprädiktoren (Hunter & Schmidt, 2004). Handelt es sich nicht um Fehlervarianz, sondern um unterschiedliche relevante Aspekte des „wahren“ Prädiktors, kann die Reliabilität des Prädiktorverbundes mit einem Verfahren von Mosier (1943, nach Hunter & Schmidt, 2004) errechnet werden (Formel A.5). Ausschlaggebend für die Entscheidung, ob die spezifischen Faktoren Fehlervarianz oder relevante Anteile am Prädiktor enthalten, soll laut Hunter und Schmidt (2004) die Definition des Konstruktes sein:

For example, verbal ability may be *defined* as what different measures of verbal ability have in common, thus implying that specific factor variance is error variance. . . . In most cases that we have encountered, specific factor variance should be treated as measurement error. (S. 439, Hervorhebung im Original).

Was für die Erhebung von Intelligenzfaktoren wie der verbalen Intelligenz gilt, gilt gleichermaßen für die Erfassung des Konstrukts *Allgemeine Intelligenz* durch verschiedene Instrumente. Da jeder Intelligenztest für sich genommen beansprucht „Intelligenz“ vollständig zu erfassen, sollten nach Hunter und Schmidt (2004) die testspezifischen Faktoren als Fehlervarianz angesehen werden und die Reliabilität des Prädiktorverbundes mittels Cronbachs α ermittelt werden. Dies führt jedoch dazu, dass der Prädiktorverbund eine geringere Reliabilität aufweist als die Einzelprädiktoren, wenn die Korrelation zwischen beiden Subprädiktoren nicht ausgesprochen hoch ist. So etwa berichtet Weiß (1969) Korrelationen zwischen CFT-, IST-Werten und Berufsschulnoten. Die Reliabilitäten (CE) der Prädiktoren betragen .95 (CFT) und .91 (IST). Da beide zu .69 korrelieren,

beträgt Cronbachs α des Prädiktorverbundes lediglich .82. Legt man die durchschnittliche Korrelation zwischen verschiedenen Intelligenzmaßen zugrunde (.77 nach Jensen, 1984), beträgt sie .87. Obwohl also zwei Messungen vorliegen, wird die Reliabilität geringer eingeschätzt als beim Vorliegen nur eines der beiden Prädiktormäße. Folgt man der Idee, dass das Konstrukt durch die Ermittlung von zwei Maßen messgenauer erfasst wird, sollte die Formel nach Mosier angewendet werden. Im Beispiel aus Weiß (1969) erhält man dadurch eine Reliabilität des Verbundes von .96. Da dies zudem die konservativere Vorgehensweise ist, wird dieses Verfahren angewendet.

Die Reliabilität eines linearen Verbundes kann nur bestimmt werden, wenn die Korrelationen zwischen den Subprädiktoren bekannt sind. Sie wurden den Testmanualen oder anderen Studien entnommen, die die Tests verwendet haben. Für die Berechnung nach Mosier sind zusätzlich die Reliabilitäten der Subprädiktoren erforderlich.

Die Berechnung erfolgte mit einer modifizierten Excel-Tabelle, die in der Software von F. L. Schmidt und Le (2005) enthalten ist. Die Modifikation war erforderlich, da in dieser Softwareversion bei einer Formel Zähler und Nenner vertauscht sind². Eine weitere Einschränkung der Software besteht darin, dass die Reliabilitäten nach Mosier nur für maximal zehn Prädiktoren berechnet werden können. Liegen mehr als zehn Prädiktoren vor, wurden eigene Exceltabellen erstellt. Dies war auch erforderlich, wenn die einzelnen Subprädiktoren mit unterschiedlichem Gewicht berücksichtigt wurden. Die Gewichtung bei der Berechnung der Reliabilität erfolgte nach Formel A.6.

Ließ sich die CES-Reliabilität des Prädiktorverbundes ausgehend von Cronbachs α nicht bestimmen, wurde das Mittel der Reliabilitätsverteilung der einzelnen Testverfahren verwendet, zu denen Reliabilitätswerte vorliegen. Dies war möglich für die AZUBI-Tests (mittleres $r_{xx} = .89$), die BIS-Verfahren (.87), die IST-Verfahren (.87), die Matrixentests (.89) sowie für die Versionen des WIT (.91). PSB und LPS wurden auf Grund ihrer nahen Verwandtschaft (vgl. Kapitel 3.2.1) zusammengefasst; für ihre Studien wurde die Prädiktorreliabilität des LPS zu Grunde gelegt (.94), da keine Informationen zum PSB vorhanden sind. Keine CE(S)-Reliabilitäten sind für die Verfahren berichtet, die auf Basis der GATB entwickelt wurden (EUB, BET). Auch für die Primärstudien, die Prädiktoren verwenden, die in der Restkategorie zusammengefasst wurden, sind keine CE/CES-Reliabilitäten bekannt, mit Ausnahme der AZUBI-Tests. Konnte die Reliabilität mit den dargestellten Methoden nicht ermittelt werden, wurde sie mit dem Mittel aus den genannten Reliabilitätsverteilungen geschätzt (mittleres $r_{xx} = .895$).

3.6.2. Unreliabilität der Kriterien

Berufsbezogene Lernleistung: Für das Kriterium *berufsbezogene Lernleistung* liegen i.d.R. Beurteilungen in Form von Berufsschulnoten und Beurteilungen aus den Ausbildungsbetrieben vor. Reliabilitäten für beides werden in den wenigsten Primärstudien berichtet, sodass auf einschlägige weitere Literatur zurückgegriffen werden musste.

Mit der Metaanalyse von Baron-Boldt und Mitarbeitern (Baron-Boldt, Funke & Schuler, 1989; Baron-Boldt, Schuler & Funke, 1988) liegt eine Artefaktverteilung für die Un-

²Für den Hinweis darauf bedanke ich mich bei U. Hülsheger.

reliabilität von Ausbildungsleistung vor, die ausschließlich auf deutschen Stichproben beruht ($k = 15$, mittleres $r_{yy} = .642$). Dieser Wert kann als Reliabilität der Ausbildungsleistung insgesamt angesehen werden. Die Reliabilität einzelner Fachnoten dürfte niedriger liegen und von zahlreichen Faktoren abhängen. Einen Überblick zu dieser Problematik im Hinblick auf die Retestrelabilität allgemeinbildender Schulnoten gibt [Tent \(2001\)](#). Da die CE(S)-Reliabilität einzelner Fachnoten nicht bekannt ist, wurde für die gesamte Berufsschulleistung die Reliabilität von $.642$ angenommen, auch wenn das Kriterium in Form eines linearen Verbundes aus Teilleistungen geringerer Breite zusammengesetzt ist.

Liegen mehrere Messungen der beruflichen Schulleistung in der gesamten Breite vor, z.B. in Form von Zwischen- und Abschlussprüfungen, wurden beide als gleichermaßen relevanter Teil des Ausbildungserfolges angesehen. Sie sind nicht deckungsgleich, da sie zu unterschiedlichen Zeitpunkten erfasst wurden und Leistung über die Zeit hinweg variabel ist. Deshalb wurde in den Fällen, die beide Prüfungsleistungen getrennt berücksichtigen, wenn möglich die Reliabilität des Kriterienverbundes nach Mosier berechnet (vgl. Kapitel 3.6.1).

Zur Schätzung der Reliabilität *betrieblicher* Aus- und Weiterbildungsbeurteilungen wurde auf die Reliabilität von Abwärtsbeurteilungen zurückgegriffen. Sie beträgt nach [Rothstein \(1990\)](#) und [Viswesvaran et al. \(1996\)](#) im Mittel $.52$ (s.u.). Beim Vorliegen von Berufsschulnoten *und* betrieblichen Beurteilungen wurde ebenfalls davon ausgegangen, dass unterschiedliche relevante Anteile des konzeptuellen Kriteriums erfasst werden. Die spezifische Varianz der Teilkriterien wurde *nicht* als Messfehler betrachtet und die Reliabilität des Verbundes beider Teilkriterien wird – wenn möglich – nach Mosier bestimmt.

Die Reliabilitäten der verschiedenen Lernleistungskriterien wurden an Stichproben von Stelleninhabern ermittelt, nicht an Bewerberstichproben, d.h. sie sind – wie von [Hunter et al. \(2006\)](#) gefordert – varianzeingeschränkt.

Subjektive Leistungsbeurteilungen: Auch für Leistungsbeurteilungen lassen sich die Reliabilitätsarten CE und CS unterscheiden, die beide korrigiert wurden (CES). Dazu sind *Interrater*- und keine *Intrarater*-Reliabilitäten zu verwenden ([Viswesvaran et al., 2002](#)): „Inter-rater reliability is the CES (delayed parallel forms reliability) for ratings“ ([Hunter & Schmidt, 2004](#), S. 101).

Die mittlere Reliabilität beruflicher Leistungsbeurteilungen durch Vorgesetzte wird in einschlägigen Metaanalysen häufig mit $.60$ beziffert. Sieben der 19 Metaanalysen aus den USA verwenden diesen Wert (Tabelle 2.5). Der Wert geht samt zugehöriger hypothetischer Artefaktverteilung zurück auf [F. L. Schmidt und Hunter \(1977\)](#). Er wird in einer Studie von [Rothstein \(1990\)](#) als angemessene Schätzung der Beurteilungsreliabilität bestätigt, wenn die Beurteilung durch Vorgesetzte erfolgt, welche die zu Beurteilenden lange kennen: „My findings suggest that this [value of $.60$] is an accurate one when the rater has had long exposure to the ratee and is a conservative (i.e., high) estimate otherwise“ ([Rothstein, 1990](#), S. 326).

Die Ergebnisse von [Rothstein \(1990\)](#) basieren auf Interraterreliabilitäten eines Beurteilungsinventars, das neun Leistungsdimensionen umfasst. Beurteilt wurden 9 975 Mit-

arbeiter in 79 Organisationen. Die mittlere Reliabilität wird von Rothstein mit .48 für *duty ratings* bzw. .52 für *ability ratings* angegeben. Die Höhe der Reliabilität ist dabei abhängig von der Dauer, mit der sich Beurteiler und Beurteilte kennen. Kennen sie sich weniger als einen Monat, liegt die von Rothstein regressionsanalytisch geschätzte Reliabilität lediglich bei .35 (*duty rating*) bzw. .39 (*ability rating*). Sie steigt bis auf .55 (*duty rating*) bzw. .59 (*ability rating*) an, wenn sich Beurteiler und Beurteilte 235 Monate kennen.

Viswesvaran et al. (1996) betrachten die Interraterreliabilität getrennt nach Leistungsdimensionen. Die globale Beurteilung der Leistung durch Vorgesetzte gelingt mit einer Reliabilität von $\bar{r} = .52$, einzelne Leistungsaspekte erreichen im Mittel Reliabilitäten von .55 (zwischen .47 für *interpersonale Kompetenz* und .63 für *Arbeitsqualität*). Die Beurteilung der Leistung durch Kollegen gelingt insgesamt betrachtet mit etwas geringerer Reliabilität ($.34 \leq \bar{r} \leq .71$; im Mittel beträgt sie .43, bei einem Globalitem als Kriterium .42).

Neben der Unterscheidung nach Leistungsdimensionen, Urteilsquellen und der Dauer, mit der sich Beurteiler und Beurteilter kennen, sind andere Faktoren denkbar, welche die Reliabilität beeinflussen und zu denen keine metaanalytisch ermittelten Schätzwerte vorliegen. So thematisieren Marcus und Schuler (2006) die Reliabilität verschiedener Ratingformate und die Frage, ob die Beurteiler geschult sind. F. L. Schmidt, Ones und Hunter (1992) betrachten die Abhängigkeit der Reliabilität von der Anzahl der Raters. Conway und Huffcutt (1997) finden Reliabilitätsunterschiede abhängig davon, ob Personen mit oder ohne Führungsverantwortung Beurteilungen vornehmen. Ferner ist die Reliabilität abhängig von der Anzahl der Items pro Rater (Greguras & Robie, 1998).

Die Reliabilität von Fremdbeurteilungen hängt somit von zahlreichen Faktoren ab, deren Ausprägungen i.d.R. nicht bekannt sind. Manche Autoren schlagen deshalb die Verwendung konservativer Korrekturwerte vor, um Überkorrekturen zu vermeiden: Hartigan und Wigdor (1989) sehen z.B. den Korrekturwert von F. L. Schmidt und Hunter (1977, basierend auf einer Reliabilität von .60) als zu liberal an und rechnen mit einer Reliabilität von .80. Dieser Wert wurde jedoch nicht empirisch ermittelt. Zudem warnen Viswesvaran et al. (1996) vor der Verwendung möglichst konservativer Werte, da sie zu einer größeren Verzerrung führen als die Verwendung mittlerer Werte, die auf der Basis von Metaanalysen gewonnen wurden. Deshalb wurde in dieser Arbeit die mittlere Reliabilität der Beurteilung mit der mittleren Reliabilität nach Rothstein (1990) geschätzt (.52), die von Viswesvaran et al. (2002) und Salgado et al. (2003a) repliziert wurde. Sie wurde in allen bisherigen Metaanalysen aus Europa ebenfalls verwendet (Tabelle 2.6), ebenso in den Studien aus den USA von Vinchur et al. (1998). Die anderen Studien aus den USA nutzen höhere Reliabilitäten, die zwischen .60 und .81 liegen (Tabelle 2.5).

Dieser Wert spiegelt nach Rothstein (1990) und Viswesvaran et al. (2002) die Reliabilität einer Leistungsbeurteilung wieder, die zu einem Messzeitpunkt erhoben wurde und auf einer oder mehreren Dimensionen beruht. Liegen Ergebnisse für mehrere Messzeitpunkte oder von mehreren unabhängigen Beurteilern vor, die zu linearen Verbänden verrechnet werden können, sollte die Reliabilität des Verbundes nach Mosier berechnet werden. Die Berechnung nach Mosier sollte zumindest dann durchgeführt werden, wenn

davon ausgegangen werden kann, dass raterspezifische Varianz keinen Messfehler darstellt, sondern verschiedene Beurteilungen unterschiedliche relevante Anteile am Kriterium erfassen. Dies wäre nur für eine Studie von Paff (1966) möglich gewesen. Deshalb wurde die Reliabilität von Verbänden, die auf mehreren Leistungsbeurteilungen basieren, *nicht* berechnet, und der Wert von .52 wurde unabhängig davon verwendet, wie viele Messwiederholungen vorliegen, wie lange sich Beurteilende und Beurteilte kennen, ob es sich um Abwärts- oder Seitwärtsbeurteilungen handelt, ob die Globalleistung oder einzelne Leistungsdimensionen eingeschätzt werden, usf.

Auch die Reliabilitäten der subjektiven Leistungsbeurteilungen entstammen Stelleninhaber-Stichproben, die im Vergleich zu Bewerber-Stichproben varianzeingeschränkt sind. Diese varianzeingeschränkten Werte werden zur Korrektur benötigt (Hunter et al., 2006).

Karriereerfolg: Die Einkommensindikatoren werden ausschließlich und die Indikatoren des beruflichen Vorankommens werden überwiegend mit neutralen Kennzahlen operationalisiert („objektive“ Maße). Deshalb wurde *keine* Messfehlerkorrektur für die Metaanalysen des Karriereerfolgs-Bereichs vorgenommen. Dies entspricht dem Vorgehen in den bisher durchgeführten Metaanalyse zum Karriereerfolg (vgl. Tabelle 2.7).

3.6.3. Varianzeinschränkung

Die Verzerrungen der Korrelationen auf Grund von Varianzeinschränkungen oder -vergrößerungen können mit dem Verfahren von Hunter und Schmidt (2004) korrigiert werden, wenn ihre Höhe bekannt ist. Im Folgenden wird zunächst auf die Berechnung der Varianzeinschränkungen (bzw. -vergrößerungen) eingegangen, anschließend wird ihre Korrektur besprochen.

Als Maß für die *beobachteten* Varianzunterschiede im Prädiktor wird der Quotient aus den Standardabweichungen der Stichprobe (SP) und Population (Pop.) gebildet: $u_x = SD_{SP}/SD_{Pop.}$. Die Standardabweichung der Stichproben-Intelligenzwerte wurde den Primärstudien entnommen, wenn sie berichtet werden oder aus korrigierten Werten berechnet werden können. Die Normwerte stammen i.d.R. aus den Testmanualen (vgl. Kapitel 3.2.1) oder sie werden in den Primärstudien berichtet.

Eine Besonderheit in Bezug auf die Normierung weist der WIT aus: der Wert für Allgemeine Intelligenz wurde nicht standardisiert, sondern lediglich die Werte der einzelnen Aufgabengruppen. Da diese positiv miteinander korreliert sind, resultiert dies in einer Streuungsreduktion der Allgemeinen Intelligenz. Als Schätzwert für die Streuung der Gesamtpopulation verwendet Greif (1972) deshalb die Streuung einer Stichprobe von 17jährigen, die Jäger und Todt (1964) berichten. Dazu heißt es: „In ihrer Streuung weicht die allgemeine Intelligenz unserer Gruppe mit $s = 6,932$ nur unerheblich von einer repräsentativen Stichprobe ab“ (Jäger & Todt, 1964, S. 4). Die mit diesem Wert ermittelten Varianzeinschränkungen für WIT-Studien ($u_x: M = .763, SD = .137$) unterscheiden sich von den u_x -Werten der Studien, die andere Testverfahren einsetzen ($M = .780, SD = .176$), auch auf einem konservativen α -Level von 20% (Diehl & Arbinger, 2001) nicht

signifikant ($U = 129$, $N = 39$, $p > .20$, zweiseitig). Dabei spielen auch unterschiedliche Komplexitätsniveaus keine Rolle (s.u.).

Die „Norm“ des WIT, auf die Bezug genommen wird, ist über 40 Jahre alt. Ein Mangel an aktuellen aussagekräftigen Normen für jüngere Primärstudien ist jedoch auch für andere Testverfahren zu beklagen (vgl. Kapitel 3.2.1). Das Alter der IQ-Normen wirkt sich besonders auf die mittlere Intelligenz der Bevölkerung aus, die um ca. 0.33 IQ-Punkte pro Jahr ansteigt (Flynn-Effekt; Flynn, 1987). Zur Erklärung des Effektes haben Dickens und Flynn (2001) drei Modelle entwickelt, die nach Rowe und Rodgers (2002) implizieren, dass auch die IQ-Varianz zunehmen müsste, was sich allerdings empirisch nicht zeigen lässt: „However, the historical rise in IQ mean level has not been accompanied by substantial variance changes“ (Rowe & Rodgers, 2002, S. 759). Eine Verwendung älterer Normen auch bei jüngeren Primärstudien scheint demnach zur Ermittlung der Varianzeinschränkung zulässig.

Ferner ist zu berücksichtigen, dass sich die Normen aus den Manualen i.d.R. auf die Gesamtbevölkerung beziehen, die interessierende Population aber *nicht* die Gesamtbevölkerung ist, sondern die *Population der Erwerbstätigen*. Sackett und Ostgaard (1994) können mit Hilfe der Daten aus dem WPT-Manual von 1983 zeigen, dass die Streuung in der Gesamtbevölkerung (8.07 gemessen mit dem WPT) größer ist als in der Population der Erwerbstätigen (7.40). Sie zeigen ferner, dass der Unterschied zwischen beiden Gruppen mit zunehmendem Komplexitätsniveau zunimmt. Sie schlagen deshalb vor, die Norm-SDs für die Studien, die nicht das unterste Komplexitätsniveau betreffen, um liberale 10% zu reduzieren oder (konservativ) die untersten um 3% und die anderen um 20%. Da in dieser Arbeit keine Studien zum untersten Komplexitätsniveau vorliegen, wären alle Norm-SDs um 20% zu verringern. Dagegen erachten Hunter et al. (2006) den Unterschied zwischen Gesamtbevölkerungs- und Erwerbstätigen-SD als minimal.

Auf Basis dieser Überlegungen wurde der „mittlere Weg“ gewählt: die Reduktion der Norm-SDs um 10%. Dies betrifft alle Studien mit Ausnahme der AZUBI-Tests, deren Normstichproben nicht für die Gesamtbevölkerung, sondern für potentielle Auszubildende repräsentativ sind (Kapitel 3.2.1).

Konnte die Varianzeinschränkung im Prädiktor für eine Studie nicht ermittelt werden, wurde sie mittels der u_x -Werte der anderen Studien geschätzt. Dabei war zu berücksichtigen, dass die Varianzeinschränkungen vom Komplexitätsniveau des Berufes abhängig sein können: Schuler und Höft (2006) postulieren, dass mit zunehmendem Komplexitätsniveau eines Berufes die Varianzeinschränkung zunimmt. Dies konnte in den vorliegenden Studien *nicht* nachgewiesen werden. Allerdings liegen nur für die beiden mittleren *skill levels* 2 und 3 ausreichend Daten vor. Es fehlen Angaben zu sehr einfachen und sehr komplexen Berufen. Der Unterschied zwischen den mittleren Standardabweichungen beider Gruppen geht in die erwartete Richtung (Skill Level 2: $M = 0.791$, $SD = 0.144$; Skill Level 3: $M = 0.752$, $SD = 0.200$). Er ist aber auch auf einem konservativen α -Level von 20% (Diehl & Arbinger, 2001) nicht signifikant ($t_{37} = -0.28$, $U = 119$, $N = 33$, $p > .20$, einseitig). Die Teststärke dieses Vergleichs ist auf Grund der geringen Stichprobengröße allerdings gering ($1 - \beta = .41$ für den t -Test).

Zur Schätzung nicht berichteter Varianzeinschränkungen wurden die Mittel der Skill-

Level-Analysen herangezogen. Studien, die den Skill Levels 1 und 2 zugeordnet werden konnten, wurden mit dem Mittel des Skill Levels 2 substituiert. Studien, die den Skill Levels 3 und 4 zugeordnet werden konnten, wurden mit dem Mittel des Skill Levels 3 substituiert. In allen anderen Fällen wurde das Mittel der Varianzeinschränkungen über alle Studien hinweg ($M = 0.776$) verwendet. Die korrigierten u_x -Werte pro Studie können den Tabellen E.1 bis E.42 entnommen werden.

Auch für die Leistungskriterien ist damit zu rechnen, dass Varianzeinschränkungen wirken, da nicht erfolgreiche Personen ausscheiden (Awiszus, 1967; Marschner, 1978; Thielepape, 1980; Weiß, 1969). Aber auch das Gegenteil ist der Fall: bei Kersting (1999a) konnten die Arbeitsleistungen der Besten nicht erfasst werden, da sie zum Zeitpunkt der Kriterienerhebung an einer Weiterbildung teilgenommen hatten. Roloff (1928) hingegen wählt die Studienteilnehmer bewusst so aus, dass das gesamte Kriterienspektrum abgedeckt ist.

Es ist davon auszugehen, dass die Kriterien auch in den anderen Studien i.d.R. varianzeingeschränkt sind. Die wenigen berichteten Daten reichten nicht aus, um die Höhe der Varianzeinschränkung verlässlich zu bestimmen und zu korrigieren. Dies ist durchaus üblich: in den bisher durchgeführten Metaanalysen wurde ebenfalls auf die Korrektur der Varianzeinschränkung im Kriterium verzichtet. Damit wird eine Unterschätzung der wahren Korrelationen zwischen Intelligenz und den Leistungskriterien in Kauf genommen (Hunter & Schmidt, 2004).

Außer der Frage, wie groß die zu korrigierenden Varianzeinschränkungen bzw. -vergrößerungen sind, war zu klären, auf Welche Art und Weise sie wirken: direkt oder indirekt (Hunter & Schmidt, 2004). Eine *direkte* Varianzeinschränkung im Prädiktor liegt dann vor, wenn die Selektion der Pbn ausschließlich und unmittelbar anhand der Intelligenzwerte vorgenommen wird. Dann hängt die Selektion ausschließlich vom wahren Intelligenzwert einer Person und dem Messfehler ab (Hunter & Schmidt, 2004). Dies ist für Personalauswahl-Situationen sehr untypisch. Es würde bedeuten, dass alle Bewerber zu Intelligenztests eingeladen würden und dann die intelligentesten ausgewählt würden, ohne andere Personalauswahl-Instrumente wie Interviews, Sichtung von Bewerbungsmappen etc. einzusetzen. Ist der beobachtete Intelligenzwert einer Person nicht mit der Selektionsvariable identisch, ist die Varianzeinschränkung *indirekt*. Das gilt nach Hunter und Schmidt (2004) eindeutig für die Fälle, in denen die Prädiktorvariable nicht Teil der Selektionsvariable ist, und ebenso für die Fälle, in denen die Selektionsvariable Intelligenz unter anderen Variablen enthält.

Ursprünglich war mit der Methode von Hunter und Schmidt nur die Korrektur direkter Varianzeinschränkungen möglich, inzwischen sind indirekte Korrekturen möglich (Hunter & Schmidt, 2004). Die Verwendung der direkten Korrektur beim Vorliegen indirekter Varianzeinschränkung führt zu einer Unterschätzung der wahren Korrelation um ca. 25% (Hunter et al., 2006). Zur Korrektur der indirekten Varianzeinschränkung werden nicht die beobachteten Werte (u_x) verwendet, sondern die Varianzeinschränkung der wahren Prädiktorwerte (u_T). Diese lassen sich aus u_x und den Messfehlern der nicht varianzeingeschränkten Population (r_{xxa}) nach Formel A.18 berechnen. Dabei gilt $u_T < u_x$.

Die indirekte Varianzkorrektur setzt Varianzeinschränkungen voraus, die einen Maximalwert nicht überschreiten: Die u_x -Werte müssen mindestens so groß sein wie $\sqrt{1 - r_{xx_a}}$ (Hunter & Schmidt, 2004). Dies ist bei einer Studie nicht der Fall (Kersting, 1999c): u_x liegt dort mit .304 unter $\sqrt{1 - .865} = .367$; die Studie wurde exkludiert.

3.7. Fehlende Primärstudien-Daten

Das Ideal einer Metanalyse ist die vollständige Inklusion der Ergebnisse aus allen Stichproben, die aus der Population gezogen worden sind, auf die sie generalisiert werden soll. Aus mehreren Gründen (s.u.) können jedoch i.d.R. nicht alle relevanten Ergebnisse berücksichtigt werden. Dies kann zu Validitätseinschränkungen führen (Rustenbach, 2003), wenn die Daten nicht zufällig fehlen (Lüdtke, Robitzsch, Trautwein & Köller, 2007; Rubin, 1976).

Welche Ursachen für fehlende Werte in Frage kommen und welche Verzerrungen der metaanalytischen Ergebnisse durch sie zu befürchten sind, wird z.B. von Rustenbach (2003) und Hopewell, Clarke und Mallett (2005) besprochen. Am bedeutendsten dürfte die Verfügbarkeits- oder Erhältlichkeitsverzerrung sein (*availability bias, retrieval bias, selection bias*), die zustande kommt, wenn Studien mit einem erwünschten (i.d.R. also signifikanten) Ergebnis mit höherer Wahrscheinlichkeit inkludiert werden als Studien mit unerwünschtem (i.d.R. nicht signifikantem) Ergebnis. Da publizierte Studien i.d.R. leichter verfügbar sind als nicht publizierte, wird in diesem Zusammenhang auch von Veröffentlichungsverzerrung (*publication bias*) gesprochen. Ob sich veröffentlichte von nicht veröffentlichten Studien unterscheiden, ist mit Moderatoranalysen prüfbar, wenn auch ausreichend nicht veröffentlichte Studien bekannt sind.

Ein detaillierteres Bild erhält man, wenn man einzelne Veröffentlichungsquellen unterscheidet, z.B. Veröffentlichungen in Zeitschriftenbeiträgen, Beiträgen in Herausgeberwerken, Dissertationen, Diplomarbeiten, grauer Literatur etc. Sind die berichteten Studieneffekte mit den Veröffentlichungsquellen konfundiert, liegt eine Quellenverzerrung vor (*source bias*, Hunter & Schmidt, 2004). Darüber hinaus wird verschiedentlich angenommen, dass die Veröffentlichungsquellen etwas über die Qualität der Studien aussagen. „Diese a priori Behauptung ist jedoch bis heute empirisch keineswegs eindeutig belegt“ (Rustenbach, 2003, S. 37).

Weitere mögliche Verzerrungen werden z.B. von Hopewell et al. (2005) benannt. Sie werden kurz dargestellt und bezüglich ihrer Bedeutung für die vorliegende Arbeit eingeschätzt. Der *time-lag bias* besteht darin, dass Studien, die signifikante Ergebnisse zeigen, früher beendet werden als Studien, die (noch) ohne signifikantes Ergebnis sind. Es ist möglich, dass dies in den Primärstudien der Fall ist, die zu Forschungszwecken durchgeführt wurden. Administrative Studien sollten nicht vom time-lag bias betroffen sein. Zum *language bias* kommt es, wenn nur Studien inkludiert werden, die in bestimmten Sprachen veröffentlicht wurden. Dies dürfte in der vorliegenden Studie nur von geringer Relevanz sein, da nur Studien in Betracht kommen, die sich auf den deutschsprachigen Kulturraum beziehen. Diese sollten überwiegend auf Deutsch veröffentlicht sein. (Tatsächlich trifft dies auf alle inkludierten Quellen zu, mit Ausnahme von Marcus, 2002,

Tabelle 3.7

Schutz- und Prüfmechanismen gegen Verzerrungen

Verzerrungen	Schutz- und Prüfmechanismen
Publikation	file-drawer-Analyse, Funnelpplot, Moderatoranalyse veröffentlichter vs. nicht veröffentlichter Studien
Quellen	Moderatoranalyse zu verschiedenen Quellen
Duplikate	Zusammenfassung von Duplikaten zu einer Studie
Report	Sensitivitätsanalysen, Exklusion von Studien mit nicht berichteten Daten
Zitation	Literaturrecherche umfasst neben der Suche nach Zitationen sechs weitere Strategien

die in Englisch verfasst ist.) Der *duplication bias* bezieht sich auf die Gefahr, dass statistisch signifikante Ergebnisse mehrfach publiziert werden. Dies lässt sich auch an einigen vorliegenden Studien nachvollziehen (z.B. Althoff, 1975, 1977). Um dieser Verzerrung zu begegnen, wurden Duplikate nur einmal in die Analyse aufgenommen (vgl. Kapitel 3.3). Der *citation bias* (signifikante Ergebnisse führen zu größerer Anzahl an Zitationen als insignifikante) hat möglicherweise einen Einfluss, indem sich abhängig von der Zitationshäufigkeit die Wahrscheinlichkeit ändert, eine Studie aufzufinden.

Die genannten Verzerrungen können die Ergebnisse gesamter Studien betreffen. Zum Teil wirken sie auch auf weiteren Ebenen, die Sutton und Pigott (2005) beschreiben: a) Effektstärken zu bestimmten Subgruppen bzw. Moderatoren werden nicht berichtet. Dazu können alle Verzerrungen führen, mit Ausnahme des language bias. b) Einzelne Effektstärken innerhalb einer Studie werden nicht berichtet, z.B. insignifikante Ergebnisse (*report bias*). Die Effektstärken selbst werden berichtet, es fehlen aber andere Werte, die für die Berechnung der Metaanalyse notwendig sind. Das ist bei der Validitätsgeneralisierung nach Hunter und Schmidt (2004) insbesondere die Stichprobengröße, die zwingend erforderlich ist. Sie ist in der vorliegenden Arbeit in allen Fällen angegeben. Andere erforderliche Kennwerte, z.B. die Höhe der Varianzeinschränkung, sind zumindest für die bare-bone Metaanalyse nicht zwingend erforderlich und können ggf. geschätzt werden.

Welche Verzerrungen in der vorliegenden Metaanalyse bedeutsam sein können und wie versucht wird, ihren Einfluss zu begrenzen bzw. zu quantifizieren, fasst Tabelle 3.7 zusammen. Die Schutzmechanismen zur Verfügbarkeitsverzerrung und dem report bias werden im folgenden näher beschrieben, die anderen finden sich in den entsprechenden Methodenkapiteln dieser Arbeit.

3.7.1. file-drawer-Analysen und Funnelpplots zur Verfügbarkeitsverzerrung

Um abzuschätzen, welchen Einfluss mögliche Verfügbarkeitsverzerrungen hätten, schlagen Hunter und Schmidt (2004) file-drawer-Analysen vor. Dazu wird berechnet, wie viele Studien mit einem Nulleffekt nötig wären, um die resultierende Effektstärke \bar{r} unter einen kritischen Wert zu senken. Als kritischer Wert wird in dieser Arbeit .10 verwendet, ab dem nach Cohen (1992) von einer *kleinen* Korrelation gesprochen werden kann. Die Berechnung der Anzahl erforderlicher Nulleffekt-Studien erfolgt nach Hunter und Schmidt

(2004) gemäß der Formel $fs_k = k(.10/\bar{r} - 1)$. Der Kennwert wird häufig in Metaanalysen angegeben und z.B. als *failsafe N* bezeichnet. Sterne, Becker und Egger (2005) empfehlen, ihm keine große Bedeutung zu schenken, da „there is no clear-cut and justifiable statistical criterion for what constitutes a ‘large’ failsafe N“ (S. 123).

Als weitere Methode werden Funnelplots angewandt, die eine grafische Einschätzung der Verfügbarkeitsverzerrung ermöglichen. Begründen sie den Verdacht, dass eine Verzerrung vorliegt, kann die Höhe der Verzerrung mit der *trim-and-fill*-Methode von Duval (2005) eingeschätzt werden. Das Vorgehen wird in Kapitel 4.1.3 näher erläutert.

3.7.2. Substitutionsverfahren

Problematisch ist der Umgang mit Studien, die insignifikante Korrelationen nicht berichten, da einerseits eine Vollerhebung aller relevanten Daten anzustreben ist, diese andererseits aber nicht in ihrer genauen Höhe bekannt sind. Eine mögliche Vorgehensweise in diesem Fall ist es, Studien mit fehlenden Werten zu exkludieren. Weil damit systematisch Studien mit insignifikanten Ergebnissen exkludiert werden, führt dies zu einer Verzerrung, wenn die Effektstärke in der Population signifikant ist. Davon ist im vorliegenden Fall auszugehen. Darüber hinaus verringert sich die Teststärke der Metaanalyse und möglicherweise die Repräsentativität der inkludierten Studien (Rustenbach, 2003).

Weiterhin wird die in den fehlenden Werten enthaltene Information, dass die Effekte nicht die Signifikanzgrenze erreichen, nicht genutzt. Im Gegensatz zu einer Effektstärke, über die nichts bekannt ist, außer dass sie erhoben wurde, liegt die relevante Information zumindest in *dichotomer* Form vor. Um diese Information zu nutzen, können nicht berichtete Werte mit Schätzwerten substituiert werden. Eine Möglichkeit ist die Substitution mit Null. Dies wird z.B. von U. Funke et al. (1987) angewendet. Die Nullsubstitution führt allerdings bei positiven Populationseffekten zu einer Unterschätzung derselben (Rustenbach, 2003). Die Ergebnisse der Metaanalysen mit Nullsubstitution sind somit als mögliche Untergrenze der geschätzten Populationseffekte anzusehen, obgleich es theoretisch möglich wäre, dass sie sogar negative Werte annehmen. Bei – wie im vorliegenden Fall – zu erwartenden deutlich positiven Effekten ist dies unwahrscheinlich.

Andere Substitutionsverfahren sind die Substitution mit dem Mittelwert \bar{r} oder Regressionsverfahren, bei denen ausgehend von der Ausprägung von Drittvariablen die fehlenden Werte geschätzt werden. Regressionsverfahren können nur bei zufällig fehlenden Werten angewendet werden (Rustenbach, 2003). Beide Verfahren können bei großen Stichproben zu Schätzungen führen, die über dem maximal möglichen kritischen Wert (r_k) liegen.

Da die bisher dargestellten Substitutionsverfahren zu Über- oder Unterschätzungen führen, wird in dieser Arbeit eine andere Substitution gewählt und den genannten Verfahren gegenüber gestellt: als Schätzwert für die *nicht berichteten insignifikanten* Korrelationen ($r_{n.ber.}$) wird der Mittelwert der *bekanntesten insignifikanten* Korrelationen (\bar{r}_{insig}) gewählt. Dass diese Schätzung unter bestimmten Bedingungen zu einem unverzerrten \bar{r} bzw. ρ führt, ist die Grundannahme des hier vertretenen Substitutionsverfahrens. Die Bedingungen, unter denen diese Substitution nicht zu Verzerrungen führt, sind: a.) gleiche

Stichprobengrößen, b) Ziehung aller möglichen Stichproben, c) eine homogene Population, d) gleiches Signifikanzniveau, und e) es dürfen nur einige beliebige insignifikante Korrelationen fehlen.

Allerdings sind nicht alle der genannten Bedingungen realistisch. So wird nur eine kleine Auswahl aller möglichen Stichproben gezogen. Dies ist eine generelle Schwierigkeit von Metaanalysen, die dazu führt, dass ρ nur geschätzt werden kann, trifft aber insignifikante Korrelationen besonders, weil diese evtl. mit geringerer Wahrscheinlichkeit veröffentlicht bzw. zugänglich gemacht werden (Verfügbarkeitsverzerrung). Das heißt, die mittlere beobachtete insignifikante Korrelation ist ebenfalls nur eine Schätzung der tatsächlichen mittleren insignifikanten Korrelation aller möglichen Stichproben. Diese Schätzung gelingt umso besser, je mehr insignifikante Korrelationen berichtet sind. Deshalb wurde die Bedingung aufgestellt, dass nur *einige* Korrelationen fehlen dürfen.

Die fehlenden Korrelationen sollen zudem *beliebig* fehlen, d.h. sich in ihrer Höhe nicht systematisch von den berichteten unterscheiden. Ein systematischer Unterschied läge z.B. vor, wenn insignifikante negative Korrelationen eher nicht berichtet werden als insignifikante positive Korrelationen.

Auch die Voraussetzung, es mit *einer* homogenen Population zu tun zu haben, ist in der Realität nicht unbedingt gegeben. Das bedeutet, dass sich die nicht berichteten Korrelationen auch in allen anderen relevanten Variablen nicht von den berichteten unterscheiden dürfen. Bei den vorhandenen Daten zu berufsbezogener Lernleistung lässt sich z.B. zeigen, dass die nicht berichteten Korrelationen vor allem Studien entstammen, die eine Version des WIT nutzen, die Berechnung von \bar{r}_{insig} aber überwiegend auf Studien basiert, die andere Intelligenztests nutzen (vgl. Tabelle D.1). Eine naheliegende Forderung ist deshalb, \bar{r}_{insig} für jede Subpopulation getrennt zu bestimmen. Dies bedeutet allerdings einen sehr hohen Aufwand, der – in Anbetracht des relativ geringen Gewichts der nicht berichteten insignifikanten Korrelationen in den vorliegenden Analysen, vgl. Tabellen 4.1, 4.29 und 4.35 – kaum gerechtfertigt erscheint. Zum anderen ist zunächst davon auszugehen, dass das RM gilt, d.h. jede einzelne Studie einer eigenen Population entstammen kann. Aus diesen Gründen wird auf eine getrennte Analyse für Subgruppen verzichtet. Dies scheint zulässig, da bei der Studieninklusion darauf geachtet wird, dass die Studien hinreichend homogen sind um in *einer* Metaanalyse zusammengefasst werden zu können und Aussagen nur auf Ebene aller Studien insgesamt, nicht auf Einzelstudien-Ebene getroffen werden sollen.

Eine weitere Voraussetzung ist in der Realität nicht unbedingt gegeben: dasselbe Signifikanzniveau und dieselbe Seitigkeit der Signifikanztestung. Diese beiden Größen und der Stichprobenumfang bestimmen den kritischen Wert r_k . Liegen verschiedene Signifikanzniveaus bzw. Testseitigkeiten vor, kann die Schätzung von \bar{r}_{insig} getrennt für unterschiedliche Signifikanzniveaus bzw. Seitigkeiten der Testung vorgenommen werden.

Das Niveau und die Seitigkeit der Signifikanztestung sind in den vorliegenden Quellen, die insignifikante Korrelationen nicht berichten (Althoff, 1974, 1975; Balck, Leins & Schröder, 1979; Biernat & Scholz, 1995; Jungkunz, 1995; Kleinevoss, 1983; Lometsch, 1975; Nettelstroth, 2004; J. U. Schmidt, 1987; Schwadorf, 2003; Seggebruch, 1982, 1984; Wolf, 1990a) nicht eindeutig angegeben. Die Fragestellungen sind allerdings einseitig und

sollten demnach statistisch auch so getestet worden sein. Davon ausgehend sind i.d.R. Korrelationen ab dem 5%-Niveau berichtet. J. U. Schmidt (1987) gibt abweichend davon eine Korrelation an, die selbst auf dem 10%-Niveau (einseitig) nicht signifikant ist. Möglicherweise handelt es sich dabei um eine fehlerhafte Darstellung. Da die Signifikanztestung auf dem 5%-Niveau die übliche ist (Diehl & Arbinger, 2001), wird davon ausgegangen, dass alle betroffenen Studien die Fragestellung einseitig auf dem 5%-Niveau getestet haben.

Somit werden Signifikanzniveau und die Testseitigkeit im vorliegenden Falle als konstant erachtet und r_k hängt infolgedessen ausschließlich von n ab. Um dieser Abhängigkeit Rechnung zu tragen, ist eine Relativierung der r_{insig} -Werte an r_k erforderlich: die Substituierung erfolgt deshalb durch die Quotienten der insignifikanten berichteten Korrelationen (r_{insig}) zu ihren kritischen Werten (r_k). Diese Quotienten spiegeln das Verhältnis wieder, in dem die insignifikanten Korrelationen zum kritischen Wert stehen. Sie sollen deshalb als Verhältniswerte (VW) bezeichnet werden: $VW = r_{\text{insig}}/r_k$.

Die VW-Werte sind negativ bei negativen Korrelationen und positiv bei positiven. Ein VW von .50 bedeutet, dass die insignifikante Korrelation halb so groß ist, wie die gerade signifikante Korrelation r_k , ein VW von -.50 bedeutet, dass die insignifikante Korrelation halb so groß ist, wie die gerade signifikante und ein anderes Vorzeichen aufweist. Der Range von VW liegt (näherungsweise) bei ± 1 . Höhere Werte sind nicht möglich, da sie nur von signifikanten Korrelationen erzielt werden können.

Die Verteilung einer großen Anzahl von VW-Werten sollte bei *positivem* signifikantem $\hat{\rho}$ von -1 bis +1 stetig ansteigen, der Modalwert nahe bei +1 liegen. Bei *negativem* signifikantem $\hat{\rho}$ sollte die Verteilung der VW-Werte von -1 bis +1 stetig abnehmen, der Modalwert nahe bei -1 liegen. Bei nicht signifikanten Korrelationen sollte der Modalwert der VW-Verteilung dem VW-Wert von $\hat{\rho}$ entsprechen.

Der mittlere VW-Wert (M_{VW}) wird als Schätzer für die Studien verwendet, die die insignifikanten Effekte nicht angegeben haben. Da die Stichprobengröße in allen Fällen bekannt ist, können für jede Studie r_k errechnet und mit M_{VW} die fehlenden $r_{\text{n.ber.}}$ geschätzt werden: $\hat{r}_{\text{n.ber.}} = r_k M_{VW}$.

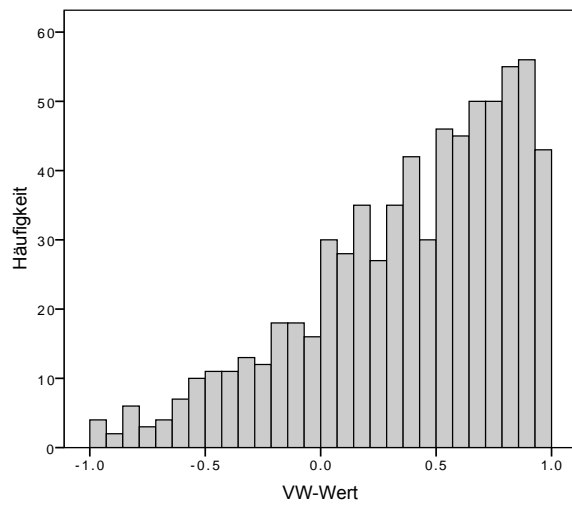
Dieses Substitutionsverfahren kann nur dann angewendet werden, wenn bekannt ist, dass die fehlenden Werte erhoben wurden und nur auf Grund ihrer Insignifikanz nicht berichtet sind. Bei Lometsch (1975) war dies nicht eindeutig ersichtlich. Aus der Darstellung der Studie im Manual des WIT (Jäger & Althoff, 1994) wurde geschlossen, dass die WIT-Skalen, zu denen keine Ergebnisse berichtet werden, *nicht* erhoben wurden. Die fehlenden Werte aus Lometsch (1975) werden deshalb auch nicht substituiert.

Die Verteilungen der VW-Werte (Abbildung 3.1) entsprechen für Lernleistungs- und Arbeitsleistungs-Studien im Großen und Ganzen den oben formulierten Verteilungsannahmen. M_{VW} liegt für Lernleistung bei .39 ($SD = .45$, $n = 707$) und für Arbeitsleistung bei .34 ($SD = .50$, $n = 31$). Die Verteilung der VW-Werte der Karriereerfolgs-Studien entspricht den Annahmen nicht. Die Verteilung beruht allerdings auf wenigen VW-Werten ($n = 23$, $M = .39$, $SD = .38$).

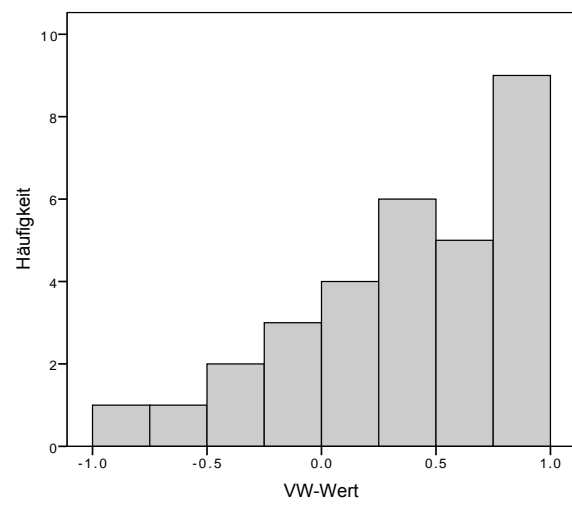
Die mittleren VW-Werte und die sich aus ihnen ergebenden Substitute pro Studie und Metaanalyse sind in Tabelle 3.8 abgebildet. Aus dieser Tabelle ist auch ersichtlich, dass

Abbildung 3.1
Verteilungen der Verhältniswerte

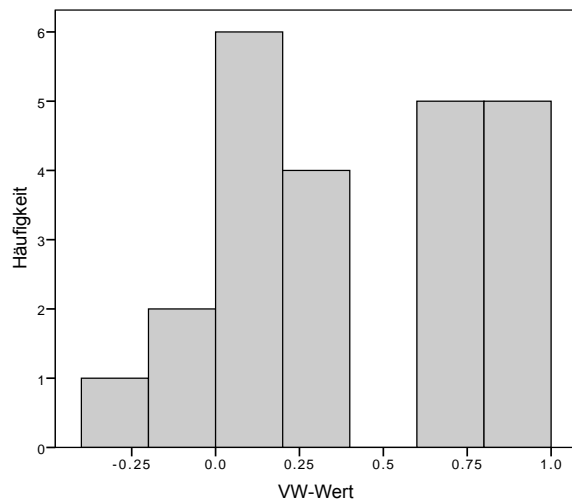
a) Lernleistung



b) Arbeitsleistung



c) Karriereerfolg



die Höhe der Substitute innerhalb einer Metaanalyse ausschließlich von n abhängig sind: je kleiner die Stichprobe, desto größer das verwendete Substitut und desto kleiner das Gewicht der Studie in der Metaanalyse. Möglicherweise führt dieser Sachverhalt dazu, dass die Varianz überschätzt wird, die durch den Stichprobenfehler verursacht wird. Um das zu prüfen werden auch andere Substitutionsverfahren durchgeführt (s.u.).

Da der mittlere VW-Wert von den Studien abhängig ist, die in einer Metaanalyse inkludiert werden und die insignifikante Korrelationen berichten, können sich die Substitutionswerte *einer* Studie zwischen verschiedenen Metaanalysen unterscheiden. Es ist deshalb erforderlich, die Substitution für jede betroffene Studie *und* jede Metaanalyse, die die Studie berücksichtigt, getrennt durchzuführen.

Zusammenfassend wird bei der VW-Substitution wie folgt vorgegangen:

1. Berechnung von r_k pro kodierter Korrelation
2. Berechnung von VW pro berichteter insignifikanter Korrelation
3. Berechnung von M_{VW} pro Metaanalyse
4. Berechnung des Substitutionswertes für jede insignifikante nicht-berichtete Korrelation pro Metaanalyse ($r_k M_{VW}$)

Dieses Verfahren wird in der vorliegenden Arbeit bei allen Metaanalysen angewendet, die Studien inkludieren, die nicht-berichtete insignifikante Korrelationen enthalten, auch wenn das Gewicht dieser Studien an der Metaanalyse vergleichsweise gering ist. Dann ist zwar nur ein minimaler Effekt der Substitutionsverfahren zu erwarten; Ziel ist es jedoch, die Auswirkungen des hier vorgestellten Substitutionsverfahrens auf die verschiedenen durchzuführenden Metaanalysen zu prüfen.

Es wird erwartet, dass sich die mittleren Effekte bei VW-Substitution zwischen denen bei Null-Substitution und der Substitution mit r_k befinden. Die alternative Vorgehensweise der Exklusion kann dagegen zu mittleren Effekten führen, die über dem möglichen Maximum bei Substitution mit r_k liegen.

Da die Höhe der Substitutionswerte innerhalb einer Analyse ausschließlich von n abhängig ist, wird damit gerechnet, dass der Anteil der durch den Stichprobenfehler aufklärten Varianz im Vergleich zu den anderen Substitutions- bzw. Exklusionsverfahren erhöht ist. Welche Auswirkungen das Substitutionsverfahren schließlich auf die anderen Kennwerte der Metaanalysen hat, wird ebenfalls beobachtet. Ob sich das Verfahren zur Substitution in den vorliegenden Analysen bewährt, wird in Kapitel 5.1 diskutiert.

3.8. Aggregation der Studieneffekte und Homogenitätstests

In den Kapiteln 3.1 bis 3.7 wurden die vorbereitenden Schritte der Metaanalyse dargestellt: die Literaturrecherche, Studienin- bzw. exklusion, Kodierung, Prüfung der statistischen Unabhängigkeit der einzelnen Studien, Zusammenfassung voneinander abhängiger Effektgrößen zu einem Wert, Klärung der Frage, welche Artefakte wie korrigiert werden und der Umgang mit fehlenden Primärstudien-Daten. Nach diesen Schritten kann die Aggregation der unterschiedlichen Studieneffekte durchgeführt werden.

Tabelle 3.8

VW- und Substitutions-Werte pro Metaanalyse und Studie, die nicht-berichtete Korrelationen enthält

Metaanalyse	VW-Werte			Substitute für die nicht-berichteten Korrelationen pro Studiennummer ^a und Metaanalyse															
	<i>M</i>	<i>SD</i>	<i>n</i>	100	102	105	107	107	107	108	129	132	138	139	145	146	175	177	177
				.00	.00	.00	.01	.02	.03	.00	.00	.00	.00	.00	.00	.00	.00	.01	.02
Lernleistung																			
Lernleistung gesamt	.385	.452	707	.073	.087	.063	.085	.054	.053	.040	.066	–	.139	.059	–	.095	.053	.033	.045
WIT	.395	.438	217	.075	.089	.064	.088	.056	.054	.041	.068	–	–	–	–	–	–	–	–
BIS	.202	.635	24	–	–	–	–	–	–	–	–	–	–	–	–	.050	–	–	–
LPS/PSB	.509	.397	37	–	–	–	–	–	–	–	–	–	–	–	–	–	.070	.043	.059
GATB basiert	.479	.429	83	–	–	–	–	–	–	–	–	–	.172	–	–	–	–	–	–
andere Intelligenztests	.297	.387	33	–	–	–	–	–	–	–	–	–	–	.046	–	–	–	–	–
Schulnoten	.362	.457	544	.069	.082	.059	.080	.051	.050	.038	.062	–	–	.056	–	.090	–	.031	.042
betriebl. Beurteilungen	.349	.451	105	–	–	.057	–	–	–	–	–	–	.126	–	–	–	.048	–	–
ohne BET-Manual	.364	.458	660	.082	.069	.059	.081	.051	.050	.038	.063	–	.131	.056	–	.090	.050	.031	.042
Ausbildungsleistung	.374	.454	670	–	.085	.061	.083	.053	.051	–	.064	–	.135	.057	–	.093	.051	.032	.044
Weiterbildungsleistung	.660	.257	7	.125	–	–	–	–	–	.069	–	–	–	–	–	–	–	–	–
publiziert	.357	.457	536	.068	.081	.058	.079	.050	.049	.037	.061	–	.129	.055	–	.089	.049	.030	.042
ISCO-88-HG 3	.390	.439	234	–	.088	.064	.087	.055	.053	.041	.067	–	–	–	–	.097	–	–	–
ISCO-88-HG 4	.326	.476	90	–	–	–	–	–	–	–	–	–	.117	–	–	–	–	–	–
Skill Level 2	.384	.456	346	–	–	–	–	–	–	–	–	–	.138	.059	–	–	.053	.033	.045
Skill Level 3	.390	.439	234	–	.088	.064	.087	.055	.053	.041	.067	–	–	–	–	.097	–	–	–
ohne PJAHR < 1945	.382	.453	705	.073	.086	.062	.085	.054	.052	.040	.066	–	.138	.059	–	.095	.052	.032	.045
prädiktiv	.360	.460	597	.068	.081	.059	.080	.051	.049	.037	.062	–	.130	.055	–	.089	.049	.031	.042
Monographien	.297	.452	254	–	–	–	–	–	–	–	–	–	–	–	–	–	.041	.025	.035
Artikel	.389	.462	240	.074	.088	.063	.086	.055	.053	.04	.067	–	.140	.060	–	.096	–	–	–
Arbeitsleistung																			
Arbeitsleistung gesamt	.337	.496	31	–	–	–	–	–	–	–	–	–	.067	–	–	–	–	–	–
nur Pearson <i>rs</i>	.337	.496	31	–	–	–	–	–	–	–	–	–	.067	–	–	–	–	–	–
Karriereerfolg																			
Einkommen	.462	.348	8	–	–	–	–	–	–	–	–	–	.091	–	–	–	–	–	–
berufl. Vorankommen	.229	.340	5	–	–	–	–	–	–	–	–	–	.045	–	–	.022	–	–	–
dto. nur Pearson <i>rs</i>	.229	.340	5	–	–	–	–	–	–	–	–	–	.045	–	–	.022	–	–	–

Anmerkungen. Nicht aufgeführte Metaanalysen enthalten keine Studien mit nicht-berichteten Korrelationen. ^aDie zu den Studiennummern gehörenden Studien können Tabelle C.1 entnommen werden.

Als metaanalytisches Verfahren wird die Methode der Validitätsgeneralisierung von Hunter und Schmidt (2004) verwendet. Dies ist die in der Arbeits- und Organisationspsychologie übliche Vorgehensweise: „the work of Hunter et al. (1982) became a quasi-standard in the field of industrial and organizational (I/O) psychology“ (Schulze, 2004, S. 15), nahezu alle Metaanalysen zum vorliegenden Gegenstandsbereich verfahren nach dieser Methode (vgl. Tabellen 2.3 bis 2.7). Sie wurde erstmals von F. L. Schmidt und Hunter (1977) vorgestellt und insbesondere in den Jahren 1980 (Schmidt et al.; interaktive Methode) und 2004 (Hunter und Schmidt; indirekte Korrektur der Varianzeinschränkung) weiterentwickelt.

B. T. Johnson, Mullen und Salas (1995) unterscheiden neben dem Verfahren von Hunter und Schmidt zwei weitere Schulen metaanalytischen Denkens: Hedges und Olkin (1985) sowie Rosenthal und Rubin (z.B. Rosenthal, 1991). Weitere metaanalytische Methoden, die zum Teil den drei Schulen zugeordnet werden können (Schulze, 2004) stammen von DerSimonian und Laird (1986), Olkin und Pratt (1958) sowie Raju et al. (1991). Detaillierte Übersichten und Vergleiche der metanalytischen Modelle geben z.B. Cooper (1998), Hunt (1997), Lipsey und Wilson (2001), Rustenbach (2003), Schulze (2004) sowie Schulze, Holling und Böhning (2003). Hier soll nur ein kurzer Vergleich der Modelle vorgenommen werden, um die Besonderheiten der Vorgehensweise von Hunter und Schmidt herauszuarbeiten. Dies geschieht in Anlehnung an Schulze (2004) und betrifft a) die verwendeten Effektmaße, b) die Gewichtung der Studien, c) das generelle Integrationsmodell und d) die empfohlene Homogenitätsprüfung.

Als Effektmaß verwenden Hunter und Schmidt (2004) die Produkt-Moment-Korrelation r oder d . In den meisten anderen Verfahren werden entweder Unterschiedsmaße genutzt (d bei Hedges und Olkin) oder z -transformierte Korrelationen (Hedges und Olkin sowie Rosenthal und Rubin). Die z -Transformation vermeidet eine Überschätzung des mittleren Effektes. Dabei wird allerdings eine Unterschätzung in Kauf genommen, die nach Hunter und Schmidt (2004) größer ist als die Überschätzung bei Verwendung von r .

Die Gewichtung der Studien wird von Hunter und Schmidt (2004) anhand der Stichprobengröße n vorgenommen. Dadurch wird der Stichprobenfehler korrigiert. Bei individueller Artefaktkorrektur pro Studie wird bei der Berechnung der Studiengewichte (w) zusätzlich die Höhe der anderen Artefakte mitberücksichtigt: $w = nA^2$, wobei A für den Attenuationsfaktor steht, der dem Quotient aus unkorrigierter und korrigierter Effektstärke der Studie entspricht (Hunter & Schmidt, 2004). Das bedeutet, je größer die vorgenommenen Artefaktkorrekturen, desto geringer das Gewicht der Studie, da $A < 1$. In den anderen Verfahren wird ebenfalls die Stichprobengröße als Gewichtungsmaß zu Grunde gelegt (z.B. bei Rosenthal und Rubin oder der Aggregation von r nach Hedges und Olkin), oder die Gewichtung basiert auf dem Standardfehler der Studien (z.B. bei der Aggregation von d nach Hedges und Olkin).

Zwei generelle Integrationsmodelle lassen sich unterscheiden (Rustenbach, 2003; Schulze, 2004): das Modell fester Effekte (*fixed model*, FM) und das Modell zufallsvariabler Effekte (*random model*, RM). Beim FM wird a priori davon ausgegangen, dass allen Studien *eine* konstante Effektgröße zu Grunde liegt. Die Variabilität zwischen den beob-

achteten Effekten ist ausschließlich auf den Einfluss der Artefakte zurückzuführen. Beim RM wird hingegen a priori angenommen, dass die Effektgröße nicht konstant, sondern variabel ist. Die Variabilität zwischen den beobachteten Effekten ist – wie beim FM – auf den Einfluss der Artefakte zurückzuführen, allerdings nur zum Teil. Zusätzlich ist die Varianz innerhalb der Effektgröße (Heterogenitätsvarianz) bei der Gewichtung der Studien und bei der Schätzung des wahren Effektes zu berücksichtigen (Schulze, 2004).

Die Einordnung der Validitätsgeneralisierung nach Hunter und Schmidt (2004) in eines der beiden Integrationsmodelle ist nicht eindeutig möglich, da die Heterogenitätsvarianz (*primary second-order sampling error* bei Hunter und Schmidt) zwar berechnet, aber nicht zur Gewichtung der Studien und zur Schätzung des wahren Effektes verwendet wird (Schulze, 2004). Zudem erfolgt die Modellentscheidung i.d.R. nicht a priori, sondern in Abhängigkeit von der geschätzten Heterogenitätsvarianz. Hunter und Schmidt (2004) selbst haben ihre Vorgehensweise in früheren Veröffentlichungen als FM bezeichnet (Hunter & Schmidt, 1990), in späteren als RM (Hunter & Schmidt, 2000). Der neueren Auffassung schließt sich Schulze (2004) an, da die Effektstärke zunächst als zufallsvariabel angesehen wird. Er bezeichnet das Vorgehen als „conditionally random effects procedure“ (S. 42).

Ob eine Heterogenitätsvarianz vorhanden ist oder nicht, wird in den meisten Metaanalyseverfahren mit einem Homogenitätstest geprüft, der die Prüfgröße Q verwendet. Sie beschreibt das mit der Studienanzahl multiplizierte Verhältnis der beobachteten Varianz der Effektstärken zur Fehlervarianz. Da sich die beobachtete Varianz additiv aus der Heterogenitätsvarianz und der Fehlervarianz zusammensetzt, kann mit einem Signifikanztest von Q (χ^2 -verteilt mit $k - 1$ Freiheitsgraden) auf das Vorhandensein oder Nicht-Vorhandensein einer Heterogenitätsvarianz geschlossen werden. Hunter und Schmidt (2004) bemängeln an der Signifikanztestung von Q allerdings die Abhängigkeit der Testpower von der Anzahl der inkludierten Studien: „Because of these problems, we recommend against use of the homogeneity [Q] test“ (S. 413). Stattdessen schlagen Hunter und Schmidt vor, die Heterogenitätsvarianz ($\hat{\sigma}_\rho$) direkt zu schätzen: „to take the estimated residual standard deviation at face value. Another alternative is to use our 75% rule“ (S. 423-424). Auf die 75%-Regel wird weiter unten eingegangen.

In Ergänzung dazu empfehlen F. L. Schmidt und Hunter (1977) die Interpretation von Glaubwürdigkeitsintervallen (*credibility intervals*, CV). Sie werden mit der korrigierten Standardabweichung um den mittleren korrigierten Effekt gebildet. Sind sie sehr weit oder schließen sie Null mit ein, ist der mittlere Effekt als Mittel von mehreren Subpopulationen anzusehen (heterogener Fall, Situationsspezifität), anderenfalls kann von einer homogenen Population ausgegangen werden (Whitener, 1990). In der Regel wird ein CV von 90% angelegt. Dies ist z.B. in den bisherigen Metaanalysen zum vorliegenden Gegenstandsbereich der Fall (vgl. Tabellen 2.3 bis 2.7), mit Ausnahme von Vinchur et al. (1998), die ein liberaleres Intervall von 80% wählen. Hunter und Schmidt (2004) legen ebenfalls nahe, ein 80%iges CV anzulegen, das Software-Paket von F. L. Schmidt und Le (2005) berechnet dieses automatisch. Da es sich bei den CVs allerdings ohnehin um ein relativ liberales Maß handelt (s.u.), wird in dieser Arbeit das 90%ige Intervall verwendet.

Während die anderen metaanalytischen Schulen keine Artefaktkorrekturen vorschlagen (außer der Korrektur des Stichprobenfehlers), empfehlen [Hunter und Schmidt \(2004\)](#) die Korrektur möglichst vieler Artefakte, um den „wahren“ Effekt auf Konstruktebene beziffern zu können. Es ist möglich, eine Validitätsgeneralisierung nach Hunter und Schmidt auch ohne weitergehende Artefaktkorrekturen durchzuführen (*bare-bone meta-analysis*). Ebenso ist es möglich, aber unüblich, Artefaktkorrekturen in Metaanalysen vorzunehmen, die anderen Schulen folgen ([Schulze, 2004](#)). Die Verwendung von Artefaktkorrekturen ist nicht unumstritten (vgl. Kapitel 5). Dennoch: durch die Artefaktkorrekturen kann eine Vergleichbarkeit der Ergebnisse zwischen Studien erreicht werden, die unterschiedlich stark von ihnen beeinflusst sind. Ebenso ist es möglich, durch einen Vergleich des korrigierten mittleren Effektes mit dem theoretischen Maximum von 1 abzuschätzen, um wieviel höher eine Korrelation sein könnte. Es sind deshalb beide Maße von Interesse: die beobachteten und die artefaktkorrigierten Effekte.

Zudem wird die *operationale Validität* bestimmt, bei der der mittlere Effekt *nicht* um die Unreliabilität im Prädiktor korrigiert wird. Dieses Vorgehen ist sinnvoll, wenn die Eignung verschiedener Testverfahren z.B. zur Personalauswahl im Fokus des Interesses steht, da die Messgenauigkeit eine Eigenschaft der Testverfahren selbst ist. Korrigiert werden hingegen Varianzeinschränkungen und Unreliabilitäten im Kriterium, da für diese Artefakte die Intelligenztests nicht verantwortlich gemacht werden können. Bei der Schätzung der *Standardabweichung* der operationalen Validität wird in jedem Fall auch die Unreliabilität im Prädiktor berücksichtigt. Die Variabilität der Studieneffekte würde sonst unterschätzt ([Hunter & Schmidt, 2004](#)).

Hunter und Schmidt berechnen die vollständig korrigierten Validitäten, während die anderen Arbeitsgruppen i.d.R. operationale Validitäten angeben (Tabellen 2.3 bis 2.7). In der vorliegenden Arbeit interessiert der Zusammenhang von Intelligenz und Berufserfolg im Allgemeinen. Es wird kein Vergleich der verschiedenen Testverfahren untereinander angestrebt. Deshalb ist die vollständig korrigierte Validität von Interesse. Die operationale Validität wird zusätzlich berechnet, um einen einfacheren Vergleich mit bereits durchgeführten Metaanalysen zu ermöglichen und – im Vergleich mit den bare-bone und vollständig korrigierten Ergebnissen – die Einflüsse der Artefakte auf die mittleren Effekte besser einschätzen zu können.

[Hunter und Schmidt \(2004\)](#) nutzen Artefaktkorrekturen auch zur Klärung der Frage, ob der korrigierte Effekt über verschiedene Situationen hinweg generalisierbar ist (z.B. Organisationen und Berufsgruppen). Sie betrachten dazu die geschätzte Varianz des korrigierten Effektes, die übrig bleibt, wenn man die Varianzanteile korrigiert, die auf Artefakte zurückzuführen sind (d.h. von der beobachteten Varianz subtrahiert). Da i.d.R. nicht alle Artefakte korrigiert werden, wird die wahre Varianz überschätzt. [F. L. Schmidt und Hunter \(1977\)](#) haben deshalb die 75%-Regel empfohlen: „we were arguing that in any given meta-analysis it is probably the case that the unknown and uncorrected artifacts account for 25% of the variance. Thus, if the estimate of real variance is not at least this large, there may be no real variance“ ([Hunter & Schmidt, 2004](#), S. 146). Dieser Fall wird von Hunter und Schmidt als *Validitätsgeneralisierung* bezeichnet.

Die CVs und die 75%-Regel geben beide Antworten auf die Frage nach der Gene-

ralisierung der metaanalytischen Befunde. Allerdings sind die Antworten die sie geben unter Umständen unterschiedlich. Die CVs sind vergleichsweise liberal und legen schnell eine homogene Population nahe, während die 75%-Regel vergleichsweise konservativ ist und Validitätsgeneralisierung häufig ausschließt. Kemery et al. (1987) unterscheiden deshalb nicht zwischen Validitätsgeneralisierung und Situationsspezifität, sondern zwischen *cross-situational consistency*, *transportability* und *situational specificity*. Beim Vorliegen von *cross-situational consistency* (CSS) schließt das CV Null aus und mehr als 75% der beobachteten Varianz werden durch Artefakte erklärt. Das bedeutet, dass der geschätzte wahre Effekt über verschiedene Situationen hinweg generalisiert werden kann. Das Gegenteil ist beim Vorliegen von *situational specificity* (SIS) der Fall: das CV enthält Null und weniger als 75% der beobachteten Varianz werden durch Artefakte erklärt. Der mittlere korrigierte Effekt ist der mittlere Effekt heterogener Subpopulationen. Die Höhe und die Richtung des Effektes können nicht verallgemeinert werden. *Transportability* (TRA) schließlich liegt dann vor, wenn das CV Null ausschließt, aber weniger als 75% der Varianz durch Artefakte erklärt werden. Die Richtung des Effektes kann dann verallgemeinert werden, allerdings muss davon ausgegangen werden, dass seine Höhe von Moderatoren abhängt.

Die Akkuratheit der Schätzung des Populationseffektes wird mit Konfidenzintervallen (CIs) eingeschätzt (Whitener, 1990). Diese basieren auf dem Standardfehler, der für $SE_{\bar{r}}$ im homogenen (vgl. Anhang A, Formel A.13) und heterogenen Fall (Formel A.14) sowie für SE_{ρ} (Formel A.27) unterschiedlich berechnet wird. Angegeben werden 95%-Intervalle. Sie werden herangezogen, um zu entscheiden, ob Differenzen von Null bzw. zwischen unterschiedlichen Moderatorgruppen bedeutsam sind. Auf die Berechnung von Signifikanzen wird verzichtet, da: „the point estimate and confidence interval are more informative, and less likely to be misinterpreted, than the significance test“ (Borenstein, 2005, S. 196).

Zur Prüfung der Belastbarkeit der Befunde werden Sensitivitätsanalysen durchgeführt und zur Prüfung der Fragestellungen (Kapitel 2.3.4) Moderatoranalysen. Beide Analysen unterscheiden sich statistisch nicht voneinander, werden aber unterschiedlich bezeichnet, um die unterschiedlichen Ziele, die mit ihnen verfolgt werden, deutlich zu machen. Für beide Arten von Analysen ist eine Mindestanzahl von Studien erforderlich. Mathematisch liegt dieses Minimum bei zwei Studien, damit sie berechenbar. Pearlman et al. (1980) hingegen legen das Minimum auf sechs Studien fest. Bei zu wenigen Studien ist die Akkuratheit der Varianzschätzung nicht gewährleistet (F. L. Schmidt, Hunter & Pearlman, 1981) und die Varianzaufklärung durch Artefakte wird überschätzt (N. Schmitt et al., 1984). Dies lässt sich an den kumulativen Baumdiagrammen (Abbildungen 4.3, 4.10 und 4.13) nachvollziehen: die Ergebnisse bei Inklusion weniger Studien unterscheiden sich deutlich von denen bei Inklusion aller Studien. Schließlich sind zur Aufdeckung von Unterschieden zwischen Populationskennwerten in der Größenordnung von .20 bei angemessener Teststärke mindestens sechs Studien erforderlich (mit durchschnittlichem n von 75; Levine et al., 1996). Gelegentlich wird das minimale k auch höher als auf sechs festgelegt, z.B. auf acht von F. L. Schmidt, Hunter und Caplan (1981). In dieser Arbeit werden Metaanalysen berechnet, wenn mindestens sechs Studien vorhanden sind.

4. Metaanalytische Ergebnisse

Die metaanalytischen Ergebnisse werden getrennt für die drei Berufserfolgs-Bereiche *berufsbezogene Lernleistung*, *Arbeitsleistung* und *Karriereerfolg* berichtet. Zunächst werden jeweils die inkludierten Studien beschrieben, anschließend die Ergebnisse der metaanalytischen Aggregation. Es folgen verschiedene Sensitivitätsanalysen zur Absicherung der metaanalytischen Befunde sowie Moderatoranalysen zur Prüfung der Fragestellungen aus Kapitel 2.3.4. Die Anzahl der durchgeführten Sensitivitäts- bzw. Moderatoranalysen ist pro Erfolgsbereich unterschiedlich und hängt im Wesentlichen von der Anzahl der inkludierten Studien ab: zu Lernleistung liegen vergleichsweise viele Studien vor, die anderen zwei Erfolgsbereiche berücksichtigen vergleichsweise wenige. Für die letztgenannten sind deshalb auch nur wenige Sensitivitäts- und keine Moderatoranalysen durchführbar.

Die Ergebnisse der Sensitivitäts- und Moderatoranalysen könnten selbst wiederum durch weitere Sensitivitäts- und Moderatoranalysen abgesichert und beschrieben werden. Dieses Vorgehen führt allerdings zu einer nahezu unüberschaubaren Anzahl an Metaanalysen. Darüber hinaus erlaubt die Anzahl der inkludierten Studien keine metaanalytische Integration für die meisten Variablenkombinationen. Aus diesen Gründen wird auf Sensitivitäts- und Moderatoranalysen zweiter und höherer Ordnung weitestgehend verzichtet. Um dennoch einen Eindruck zu bekommen, welche Variablen einen Einfluss auf die einzelnen Analysen ausüben, enthält Anhang D die Ausprägungen der wichtigsten Variablen an den einzelnen Analysen.

Pro Analyse werden die verschiedenen Substitutionsverfahren insignifikanter nicht-berichteter Korrelationen angewendet und eine Analyse bei Exklusion der betroffenen Studien durchgeführt, um die Auswirkungen der in Kapitel 3.7.2 eingeführten Verhältnissubstitution beobachten zu können. Darüber hinaus wird jeweils eine Analyse ohne Studien durchgeführt, deren Effektstärke eine Mittelwert-Aggregation mehrerer Einzelkorrelationen darstellt, da diese den wahren Effekt unterschätzen (vgl. Kapitel 3.5). Enthalten die verbleibenden Studien, deren Effektstärken keine Aggregate oder lineare Verbünde darstellen, insignifikante nicht-berichtete Korrelationen, werden diese nach einer Verhältnissubstitution berücksichtigt, andere Substitutionsverfahren bzw. Exklusionen werden dafür nicht durchgeführt. *Zusammenfassungen* der wichtigsten Ergebnisse sind dem Diskussionsteil (Kapitel 5) zu entnehmen.

Die *Ergebnisse auf Studienebene* sind in Anhang E wiedergegeben: Für jede durchgeführte Metaanalyse ist aufgeführt, welche Studien inkludiert werden, ob diese nicht-berichtete Korrelationen enthalten, welche Art der Aggregation durchgeführt wurde, in welcher Höhe Artefakte korrigiert wurden, wie hoch die beobachtete und korrigierte Effektstärke ist und welches Gewicht die Studie an der Analyse hat. Die einzelnen Studien sind in der Reihenfolge, in der sie kodiert wurden, durchnummeriert. Die Zuordnung der Studiennummern zu den Quellen enthält Tabelle C.1, die Zuordnung der Quellen zu den

Studennummern Tabelle C.2. Schließlich bietet Tabelle C.3 eine Übersicht darüber, in welchen Analysen die einzelnen Primärstudien berücksichtigt sind.

4.1. Intelligenz und berufsbezogene Lernleistung

Bevor die Ergebnisse der Metaanalyse bei Inklusion aller Lernleistungs-Studien (Kapitel 4.1.2) sowie der Sensitivitäts- (Kapitel 4.1.3) und Moderatoranalysen (Kapitel 4.1.4) berichtet werden, werden die inkludierten Studien näher beschrieben (Kapitel 4.1.1). Der eilige Leser findet Zusammenfassungen der wichtigsten Ergebnisse zum Erfolgsbereich berufsbezogene Lernleistung in der Diskussion (Kapitel 5.2).

4.1.1. Beschreibung der inkludierten Studien

Zum Erfolgsbereich *berufsbezogene Lernleistung* liegen 210 Studien vor (Variable KINHALT = 2 \vee 3), die 71 Quellen entstammen. Mit Abstand am meisten Studien (70) stammen aus dem Manual des BET (Schmale & Schmidtke, 2001), das sind 33.3% der Studien. Ihr Gewicht in der bare-bone Metaanalyse beträgt 42.2%. Welchen Einfluss das BET-Manual auf das Ergebnis hat, wird in einer Sensitivitätsanalyse geprüft (Kapitel 4.1.3). Durchschnittlich berichtet eine Quelle die Ergebnisse zu 2.93 unabhängigen Stichproben (bzw. zu 1.99 unabhängigen Stichproben, wenn man das BET-Manual nicht berücksichtigt). Die Stichprobengröße beträgt $N = 30\,451$ Personen, die durchschnittliche Stichprobengröße pro Studie $\bar{n} = 145.0$ (Range: 7 bis 967, $SD = 147.3$).

Über die Geschlechterverteilung können in 88 Studien Aussagen gemacht werden: fünf der 88 Stichproben sind rein weiblich, 32 rein männlich und 51 gemischtgeschlechtlich. Die Gesamtstichprobe besteht zu 64.4% aus Personen, deren Geschlecht nicht berichtet ist. Die Personen, von denen das Geschlecht mitgeteilt wird, sind zu 22.8% weiblich und zu 77.2% männlich. Für 72 Studien ist das durchschnittliche Alter der Teilnehmer berichtet, es liegt zwischen 15.8 und 41.0 Jahren ($M = 21.1$, $SD = 6.0$), der n -gewichtete Altersdurchschnitt beträgt 20.4 Jahre. Das Alter der Auszubildenden liegt im Mittel bei 18.9 Jahren (n -gewichtete 19.5 Jahre, $SD = 3.5$ Jahre), das der Weiterbildungsteilnehmer bei 28.8 Jahren (n -gewichtete 25.4 Jahre, $SD = 6.8$ Jahre).

In 147 Studien wird keine Aussage zum Bildungsniveau der Teilnehmer getroffen, sieben Studien können dem Bildungsniveau *bis Hauptschulabschluss* zugeordnet werden, 17 dem Bildungsniveau *(Fach-)Hochschulreife*. Das Bildungsniveau der Gesamtstichprobe ist zu 69.2% unbekannt, 8.7% haben maximal den Hauptschulabschluss, 10.6% die mittlere Reife und 11.5% die (Fach-)Hochschulreife.

Von zwei Veröffentlichungen abgesehen, die Ergebnisse zu neun Stichproben beinhalten (Roloff, 1928; Neumann, 1938), wurden die Studien zwischen 1960 und 2006 veröffentlicht bzw. durchgeführt. Die Studien entstammen zum Großteil Testmanualen (Gewicht: ca. 50%), gefolgt von Zeitschriftenartikeln (ca. 25%). Dissertationen, Monographien und Beiträge in Herausgeberwerken haben ein geringeres Gewicht (0.6% bis 6.7%). 25 Studien stammen aus nicht veröffentlichten Quellen (Gewicht in der bare-bone Analyse: 11.9%). Weitere Informationen dazu und zu den Variablen, die im Folgenden beschrieben werden, enthält Tabelle 4.1.

Tabelle 4.1

Gewicht kategorialer Moderatorvariablen an der Metaanalyse zu berufsbezogener Lernleistung

Kategorie	Häufigkeit		Gewicht ^a (%)		
	<i>k</i>	(%)	bare-bone	operational	vollst. korr.
Publikationsjahr					
1920-1929	3	1.4	0.1	0.1	0.1
1930-1939	6	2.9	0.3	0.4	0.4
1940-1949	0	0.0	0.0	0.0	0.0
1950-1959	0	0.0	0.0	0.0	0.0
1960-1969	15	7.1	4.6	4.5	4.4
1970-1979	21	10.0	13.7	14.3	14.5
1980-1989	43	20.5	8.8	8.4	8.5
1990-1999	23	11.0	17.7	15.9	15.6
2000-2006 (ohne BET-Manual)	29	13.8	12.6	13.2	13.2
BET-Manual (2001)	70	33.3	42.2	43.3	43.4
Publikationsart					
Testmanuale	81	38.6	49.0	50.9	51.0
Zeitschriftenartikel	67	31.9	27.3	26.0	26.1
Dissertationen	29	13.8	6.7	6.5	6.5
Monographien (ohne Testmanuale)	6	2.9	3.6	2.9	3.0
Beiträge in Hrsg.-Werken	2	1.0	0.6	0.6	0.5
nicht veröffentlicht	25	11.9	12.9	13.1	13.0
ISCO-88 Berufs-Hauptgruppe					
0 Militär	7	3.3	4.3	3.6	3.5
1 Leitende Berufe	1	0.5	0.2	0.1	0.1
2 Wissenschaftler	2	1.0	0.7	0.6	0.6
3 Techniker und Gleichrangige	50	23.8	17.1	15.7	15.6
4 Bürokräfte	35	16.7	18.9	18.6	18.7
5 Dienstleistungsberufe	2	1.0	0.5	0.5	0.5
6 Landwirtschaft / Fischerei	0	0.0	0.0	0.0	0.0
7 Handwerk	86	41.0	40.4	42.9	42.9
8 Maschinenbediener	8	3.8	6.9	6.7	6.4
9 Hilfsarbeitskräfte	0	0.0	0.0	0.0	0.0
gemischt / unklar	19	9.0	11.0	11.4	11.8
ISCO-88 Skill Level					
ohne	8	3.8	4.5	3.6	3.5
1 (geringe Komplexität)	0	0.0	0.0	0.0	0.0
2	138	65.7	72.0	74.1	74.3
3	50	23.8	17.1	15.7	15.6
4 (hohe Komplexität)	2	1.0	0.7	0.6	0.6
gemischt / unklar	12	5.7	5.8	5.9	5.9
Bildungsniveau					
bis Hauptschulniveau	7	3.3	2.8	2.2	2.2
Mittlere-Reife-Niveau	0	0.0	0.0	0.0	0.0
(Fach-)Hochschulreife-Niveau	17	8.1	6.6	5.9	5.9
gemischt / unklar	186	88.6	90.6	91.9	91.8
Gesamt	210	100.0	100.0	100.0	100.0

Tabelle 4.1
(Fortsetzung)

Kategorie	Häufigkeit		Gewicht ^a (%)		
	<i>k</i>	(%)	bare-bone	operational	vollst. korr.
Prädiktoren					
GATB-basiert	72.7 ^b	34.6	43.1	44.0	44.1
IST	49.8 ^b	23.7	22.7	22.2	21.2
WIT	29.8 ^b	14.2	9.9	8.5	8.8
Matrizen	12.7 ^b	6.0	3.8	3.5	3.5
LPS/PSB	12.5 ^b	6.0	7.3	7.6	8.3
BIS	9.0 ^b	4.3	2.9	4.2	4.0
andere	23.5 ^b	11.2	10.4	10.0	10.0
Kriterien					
Ausbildung	190	90.5	94.3	94.1	94.3
Weiterbildung	20	9.5	5.7	5.9	5.7
Schulnoten	120.0 ^b	57.1	52.5	51.1	51.0
betriebliche Beurteilungen	20.0 ^b	9.5	5.4	5.6	5.6
unklar	70.0 ^b	33.3	42.2	43.3	43.4
Art der Validität					
prädiktiv	197	93.8	97.7	98.0	98.0
konkurrent	8	3.8	1.8	1.5	1.5
retrograd	1	0.5	0.3	0.3	0.3
unklar	4	1.9	0.1	0.2	0.2
nicht-berichtete Korrelationen enthalten					
nein	196	93.3	95.6	96.3	96.1
ja	14	6.7	4.4	3.7	3.9
Aggregationen enthalten einfache Mittelung					
nein	169	80.5	83.8	85.3	84.9
ja	41	19.5	16.2	14.7	15.1
Typ des Korrelationskoeffizienten					
Pearson-Korrelation ^c	210	100.0	100.0	100.0	100.0
andere Koeffizienten	0	0.0	0.0	0.0	0.0
Gesamt	210	100.0	100.0	100.0	100.0

Anmerkungen. ^aGewicht bei Inklusion der Studien, die die insignifikanten Werte nicht berichten; ^bMehrfachnennungen möglich: die einzelnen Prädiktoren/Kriterien werden gleichgewichtig berücksichtigt, die betroffenen Studien anteilig aufgeteilt. ^cPearson-Korrelationen für Intervalldaten (Produkt-Moment-Korrelationen) und ordinale Daten (Spearman-Rangkorrelationen).

Eine detaillierte Übersicht über die inkludierten und ISCO-88-kodierten Berufe ist Anhang C.5 zu entnehmen, die zusammengefassten Werte Tabelle 4.1. Mit 40.6% an der Gesamtstichprobe ist die handwerkliche Berufsgruppe (ISCO-88 Hauptgruppe [HG] 7) am stärksten vertreten, gefolgt von Bürotätigkeiten (HG 4, 18.9%), Technikern (HG 3, 17.1%), Maschinenbedienern (HG 8, 6.9%) und Militärangehörigen (HG 0, 4.3%). Die Studienanzahl erlaubt getrennte Moderatoranalysen für die bisher genannten Gruppen. Dies ist nicht möglich für diejenigen Berufsgruppen mit weniger als sechs Studien, auf die jeweils weniger als 1% der Stichprobe entfallen: Wissenschaftler (HG 2), Dienstleistungs- (HG 5) und leitende Berufe (HG 1). Keine Daten liegen für landwirtschaftliche und Fischerei-Berufe (HG 6) und Hilfsarbeitskräfte (HG 9) vor. 10.8% der Stichprobe lassen sich den Berufsgruppen nicht eindeutig zuordnen.

Der Großteil der Stichprobe (72.0%) besteht aus Personen mit dem Skill Level 2, das vom angelernten Hilfsarbeiter bis zum Meister reicht. 17.1% entfallen auf Skill Level 3 (Techniker und vergleichbare Berufe). Für diese beiden Gruppen sind getrennte Moderatoranalysen möglich. Personen mit dem höchsten Skill Level 4 (Wissenschaftler) machen nur 0.7% der Stichprobe aus, Personen mit dem geringsten Skill Level 1 (ungelernte Hilfsarbeiter) sind gar nicht vertreten. Die militärischen und leitenden Berufsgruppen, denen kein Skill Level zugeordnet wird, machen 4.5% der Gesamtstichprobe aus; die restlichen 5.8% lassen sich nicht eindeutig zuordnen.

Zur Operationalisierung von Intelligenz wird am häufigsten der BET verwendet: mit weiteren Intelligenzverfahren auf Basis der GATB beträgt sein Gewicht in der bare-bone Analyse 43.1% – im Gegensatz zum IST (22.7%), WIT (9.9%), dem LPS bzw. PSB (7.3%), Matrizentests (3.8%) und dem BIS (4.3%). Andere Verfahren werden in 10.4% der Studien eingesetzt. Eine Aufstellung der einzelnen Testversionen enthält Tabelle 4.2, getrennt für Inklusion (Substitution) und Exklusion von Studien mit nicht-berichteten Korrelationen. Dabei fällt auf, dass besonders die WIT-Gruppe von fehlenden Werten beeinflusst ist.

Damit haben die einzelnen BIS-Facetten für das Ergebnis der Metaanalysen zur berufsbezogenen Lernleistung folgende Bedeutung (Tabelle 4.3): Weitaus am wichtigsten ist Verarbeitungskapazität (83.2%), gefolgt von Bearbeitungsgeschwindigkeit (6.2%). Die Operationen Merkfähigkeit (5.3%) und Einfallsreichtum (1.4%) haben eine relativ geringe Bedeutung. Die drei Inhaltsfacetten sind dagegen ungefähr gleich stark ausgeprägt (36.6% verbal, 30.1% figural, 29.4% numerisch).

Auf Kriterienseite lassen sich Studien unterscheiden, die Ausbildungsleistung (bare-bone Gewicht: 94.3%) oder Weiterbildungsleistung (5.7%) thematisieren. Berufsschulnoten sind für 52.5% des bare-bone Gewichtes verantwortlich, 13.8% gehen auf den Einfluss betrieblicher Beurteilungen zurück. Bei den restlichen 42.2% der Studien ist entweder unklar, ob es sich um schulische oder betriebliche Beurteilungen handelt, oder aber beide Beurteilungen sind konfundiert.

In der Regel handelt es sich um Längsschnittstudien, die die prädiktive Validität von Intelligenz erfassen. In einer Studie, die in Kersting (1999a, 2001) veröffentlicht wurde, wurde Intelligenz 121 Monate *nach* dem Kriterium erhoben (retrograde Validität). Konkurrent validiert wurde von Greif (1972) sowie in sieben Studien aus Zeißig (1989,

Tabelle 4.2

Lernleistung: Inkludierte Intelligenztestversionen

Testversion	k bei Substitution	k bei Exklusion
GATB-basierte Verfahren	74	73
BET	71	70
EUB	2	2
Eigenentwicklung	1	1
Intelligenz-Struktur-Test (IST)	56	56
1953-1970	52	52
Bundeswehr	4	4
WILDE Intelligenztest (WIT)	34	26
1963-1994	25	23
DGP-Version	8	2
WIT-2 in Vorbereitung	1	1
Matrizentests	22	22
CFT 3	10	10
FRT	14	14
Raven	3	3
Eigenentwicklung	2	2
LPS/PSB	15	12
PSB	10	7
LPS	5	5
Berliner Intelligenzstruktur-Test (BIS)	9	8
BIS Form 4	6	5
BIS-r-DGP	3	3
andere	24	23
Begabungstest	2	2
AZUBI-BK	1	1
AZUBI-TH	1	1
RVT	1	1
Eigenentwicklung	20	19

Anmerkung. Mehrfachnennungen möglich.

Tabelle 4.3

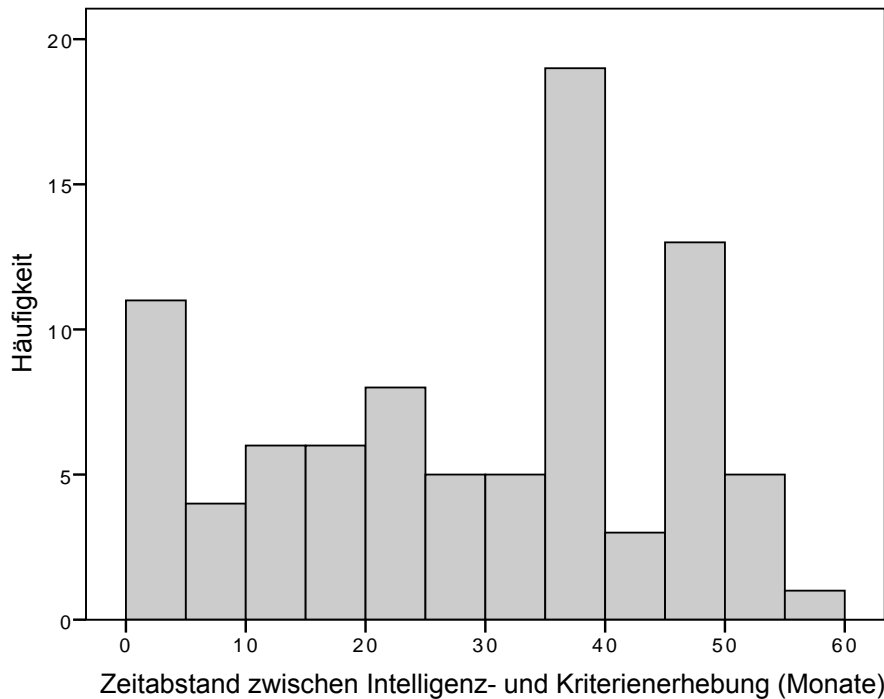
Geschätztes Gewicht der BIS-Zellen an der Metaanalyse zur berufsbezogenen Lernleistung bei Inklusion aller Studien

Operationen	Inhalte (%)			Gesamt
	verbal	numerisch	figural	
Verarbeitungskapazität				
bare-bone	29.8	26.5	26.9	83.2
operationale Validität	29.8	26.4	27.0	83.1
vollständig korrigiert	29.4	26.4	27.0	82.9
Bearbeitungsgeschwindigkeit				
bare-bone	2.6	1.4	2.2	6.2
operationale Validität	2.7	1.5	2.2	6.4
vollständig korrigiert	2.9	1.6	2.3	6.7
Einfallsreichtum				
bare-bone	0.9	0.2	0.3	1.4
operationale Validität	1.0	0.3	0.4	1.7
vollständig korrigiert	1.1	0.3	0.4	1.7
Merkfähigkeit				
bare-bone	3.3	1.3	0.7	5.3
operationale Validität	3.4	1.4	0.7	5.4
vollständig korrigiert	3.3	1.4	0.7	5.3
Gesamt				
bare-bone	36.6	29.4	30.1	96.1 ^a
operationale Validität	36.9	29.5	30.3	96.7 ^a
vollständig korrigiert	36.7	29.7	30.3	96.7 ^a

Anmerkungen. ^aAn 100% fehlende Angaben lassen sich keinen Facetten zuordnen.

Abbildung 4.1

Zeitabstand zwischen Intelligenz- und Kriterienerhebung der prädiktiven Studien zur Lernleistung



Anmerkungen. Abgebildet sind die Häufigkeiten der 86 *prädiktiven* Studien, bei denen die Zeitabstände mit *hoher Sicherheit* eingeschätzt werden konnten (Konfidenzratings 3 und 4).

Studiennr. 179.08–179.14, vgl. Anhang C.1). Bei Roloff (1928) ist unklar, ob eine konkurrente oder eine prädiktive Validierung mit geringem Zeitabstand vorliegt.

Die Zeitdifferenz zwischen Intelligenz- und Kriterienerhebung wurde in 160 Studien nicht oder nicht eindeutig berichtet (Konfidenzrating 0-3). 50 Studien geben die Zeitdifferenz exakt an (Konfidenzrating 4). Betrachtet man die prädiktiven Studien, die eine Einschätzung der Zeitdifferenz einigermaßen gesichert zulassen (Konfidenzratings 3 und 4; 86 Studien), ergibt sich die in Abbildung 4.1 wiedergegebene Verteilung. Das Mittel liegt bei 28.5 Monaten ($SD = 16.0$ Monate). Der Modalwert liegt bei 36 Monaten. Das entspricht der Dauer der meisten Ausbildungsgänge. Zwei weitere Verteilungsgipfel liegen bei ca. 48 Monaten und unter fünf Monaten. Die höchste Zeitdifferenz einer prädiktiven Studie beträgt 59 Monate.

Studien, die insignifikante Korrelationen nicht berichten, haben ein Gewicht an der bare-bone Metaanalyse von 4.4%. Die Bedeutung der insignifikanten Korrelationen innerhalb dieser Studien ist unterschiedlich (Tabelle E.1). Berücksichtigt man dies, lässt sich der Einfluss der insignifikanten Korrelationen selbst berechnen. Er beträgt 0.03% bei der bare-bone Analyse, 2.73% bei der Berechnung der operationalen Validität und 2.80% bei der vollständig korrigierten Analyse.

Eine Unterschätzung der tatsächlichen Validitäten durch die Aggregatbildung über Mittelwerte ist in 19.5% der Studien möglich. Ihr Gewicht an der bare-bone Analyse beträgt 16.2%. Alle Studien berichten Pearson-Produkt-Moment- oder Rangkorrelationen.

Eine Übersicht über die beobachteten Korrelationen aus den einzelnen Studien, ihre Konfidenzintervalle und ihr Gewicht an der bare-bone Analyse geben Abbildung 4.2 und Tabelle E.1. In ihnen wird deutlich, dass vergleichsweise akkurate Studien (enge Konfidenzintervalle) mit größerem Gewicht in die Analyse eingehen als weniger akkurate.

4.1.2. Metaanalytische Ergebnisse bei Inklusion aller Lernleistungsstudien

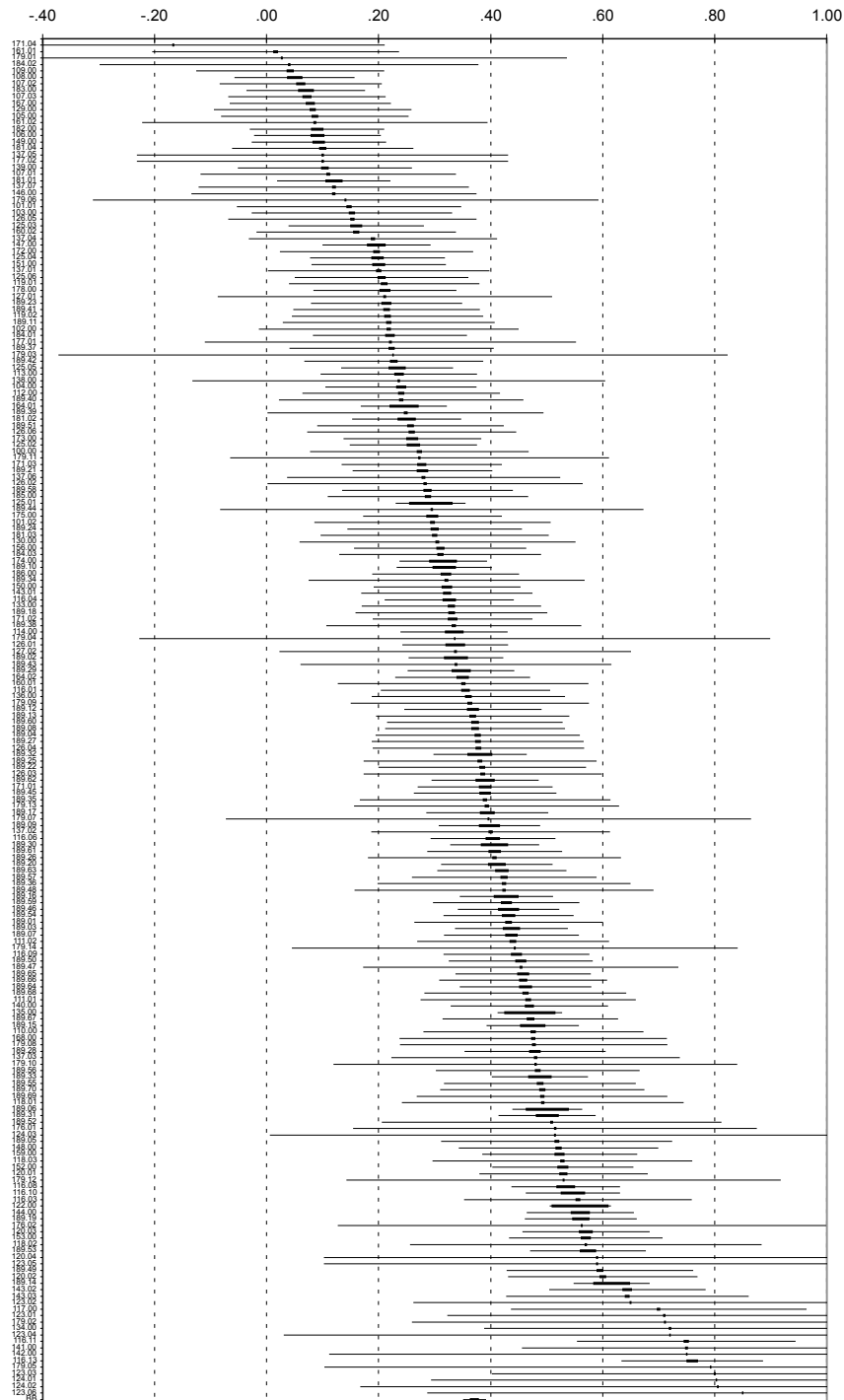
Die mittlere Validität der Studien zu berufsbezogener Lernleistung bei Korrektur der Stichprobenfehler und Verhältnissubstitution beträgt $\bar{r} = .371$. Diese ist nur geringfügig höher als bei Nullsubstitution ($\bar{r} = .369$) und geringfügig niedriger als bei Exklusion der Studien, die insignifikante Korrelationen nicht berichten ($\bar{r} = .382$). Die mittlere Validität bei der Exklusion liegt damit etwas über der Validität, die maximal möglich wäre ($\bar{r} = .375$). Diese marginalen Unterschiede sind auf das geringe Gewicht der nicht-berichteten Korrelationen an der Gesamtanalyse zurückzuführen. Auch die anderen Kennwerte unterscheiden sich bei verschiedenen Substitutionsverfahren nur unbedeutend (Tabelle 4.4). Die Varianz, die auf den Stichprobenfehler zurückzuführen ist, ist bei Verhältnissubstitution mit 22.9% erwartungsgemäß etwas höher als bei Substitution mit Null (21.9%), liegt aber unter der Varianz bei Exklusion der Studien, die insignifikante Korrelationen nicht berichten (24.1%).

Die Exklusion der Studien mit Mittelwert-Aggregaten resultiert in einem \bar{r} von .390. Nicht-berichtete Korrelationen wurden dabei verhältnissubstituiert. Der Unterschied zur Inklusion aller Studien ($\bar{r} = .371$) ist gering; die Konfidenzintervalle überlappen sich (vgl. Tabelle 4.4). In allen Fällen klärt der Stichprobenfehler weniger als 25% der beobachteten Varianz auf, es ist von Moderatoreinflüssen auszugehen. Das 95% Konfidenzintervall liegt deutlich über Null. Das faile-safe k von über 490 zeigt an, dass weit mehr Studien mit Nulleffekten erforderlich wären als inkludiert sind, um \bar{r} auf unter .10 zu drücken.

Über die korrigierten Validitäten gibt Tabelle 4.5 Auskunft. Die vollständig korrigierte Validität beträgt $\hat{\rho} = .621$ bei Verhältnissubstitution. Die Ergebnisse der anderen Substitutionsverfahren unterscheiden sich davon ebenfalls nur geringfügig und in gleicher Weise wie bei der bare-bone Analyse. Der Wert für $\hat{\rho}$ bei Exklusion liegt über dem maximal möglichen bei Substitution mit r_k . Die Varianzaufklärung durch die Artefaktkorrekturen liegt mit 25.6% deutlich unter 75%, die untere Grenze des 90% Glaubwürdigkeitsintervalls ist größer als Null (.32). Somit liegt Transportabilität vor: es kann generell von einem positiven Zusammenhang zwischen Intelligenz und berufsbezogener Lernleistung ausgegangen werden. Die operationale Validität liegt jeweils etwas unter der vollständig korrigierten Validität. Bei Verhältnissubstitution beträgt sie $\hat{\rho} = .587$.

Ein *kumulatives Baumdiagramm* veranschaulicht das Ergebnis der kumulativen Metaanalyse, bei der die einzelnen Studien *nacheinander* aufgenommen werden. Begonnen wird mit der ältesten Studie. Abgebildet werden jeweils die mittleren Effekte und 95% CIs. Dem metaanalytischen Ergebnis bei Inklusion aller, auch der jüngsten Studien, kann dann vertraut werden, wenn die mittleren Effekte bei enger werdenden Konfidenzinter-

Abbildung 4.2
Baumdiagramm der Lernleistungsstudien



Anmerkungen. X-Achse: beobachtete Korrelationen der einzelnen Studien sortiert nach Größe und Ergebnis der bare-bone Analyse (BB, letzte Zeile); Y-Achse: Studiennummern (vgl. Anhang C.1); Balkenbreite: Gewicht der Studie in der bare-bone Analyse, bei BB durchschnittliches Studiengewicht; Fehlerbalken: 95% Konfidenzintervalle.

Tabelle 4.4

Lernleistung: bare-bone Analyse

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
Null-Subst.	210	30 451	.369	.155	.072	.137	21.9	.35–.39	566
VW-Subst.	210	30 451	.371	.151	.072	.133	22.9	.35–.39	570
r_k -Subst.	210	30 451	.375	.147	.072	.128	24.1	.35–.39	577
Exkl. n.ber.	196	29 111	.382	.144	.071	.125	24.1	.36–.40	554
Exkl. MAg	169	25 518	.390	.142	.070	.124	24.1	.37–.41	491

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k .

vallen stabil bleiben. Das kumulative Baumdiagramm zu Lernleistung (Abbildung 4.3) enthält die vollständig korrigierten Ergebnisse bei Verhältnissubstitution. Es wird deutlich, dass die älteren Studien (bis ca. 1980) hohe mittlere Effekte berichten, die sogar *out of range* liegen. Bei der sukzessiven Inklusion der ca. 40 ältesten Studien kommt es auch zu verhältnismäßig großen Sprüngen des mittleren Effektes. Dies ist mit zunehmender Inklusion nicht mehr der Fall: der mittlere Effekt pendelt sich zwischen $\hat{\rho} = .60$ und $.65$ ein.

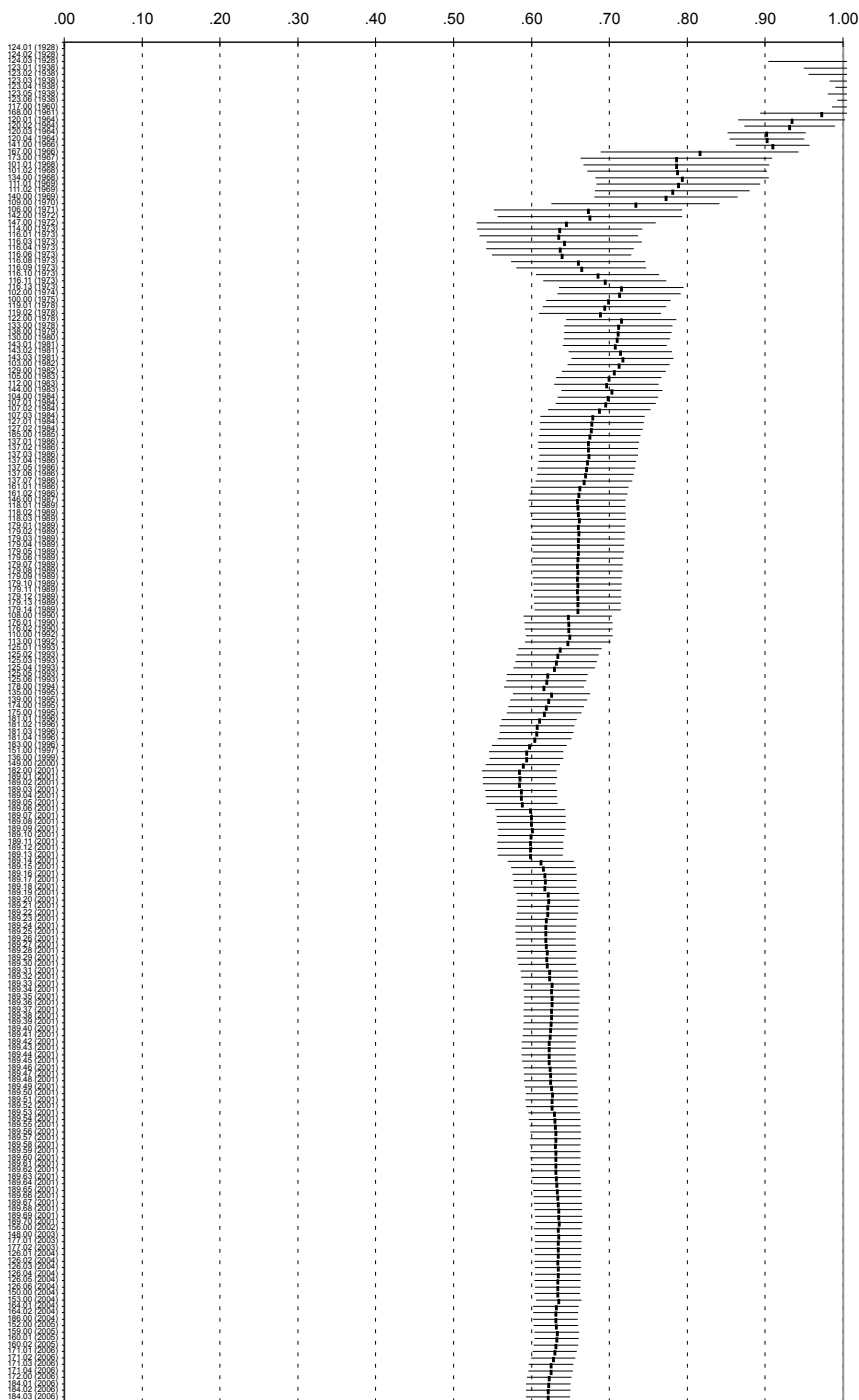
4.1.3. Sensitivitätsanalysen

Wie belastbar das metaanalytische Ergebnis ist, kann außer mit dem kumulativen Baumdiagramm auch mit weiteren Sensitivitätsanalysen eingeschätzt werden. Sensitivitätsanalysen werden durchgeführt zur Veröffentlichungs- bzw. Quellenverzerrung, dem Einfluss des BET-Manuals und der Studien, die vor 1945 veröffentlicht wurden sowie dem Einfluss der nicht-prädiktiven Validitäten.

Veröffentlichungsverzerrungen: Veröffentlichungsverzerrungen sollten sich in Funnelplots zeigen. Funnelplots sind Streudiagramme, bei denen die beobachteten Korrelationen im Verhältnis zur Stichprobengröße abgebildet werden (oder im Verhältnis zum Standardfehler, der hier nicht benutzt wird, da die Gewichtung der Studien auf Basis der Stichprobengröße vorgenommen wird). Die Streuung der beobachteten Korrelationen sollte mit zunehmendem n abnehmen und symmetrisch um \bar{r} sein. Ist dies nicht der Fall, kann dies als Hinweis auf Selektionsverzerrungen gewertet werden. Andere Ursachen für asymmetrische Funnelplots sind aber nicht auszuschließen, z.B. Qualitätsmängel von Studien mit kleinen Stichproben, wahre Heterogenität der Studien oder Zufall (Sterne et al., 2005; Sutton & Pigott, 2005).

Der Funnelplot zu berufsbezogener Lernleistung (Abbildung 4.4) liefert mehrere Hinweise auf einen möglichen Veröffentlichungsbias. Zum einen ist die Streuung der Studien

Abbildung 4.3
 Kumulatives Baumdiagramm der Lernleistungs-Metaanalyse



Anmerkungen. X-Achse: $\hat{\rho}$ (vollständig korrigiert, Verhältnissubstitution); Y-Achse: Studiennummern (vgl. Anhang C.1) sortiert nach Jahreszahl, die – von oben nach unten – *zusätzlich* in der Metaanalyse berücksichtigt werden; Fehlerbalken: 95% Konfidenzintervalle von $\hat{\rho}$.

Tabelle 4.5

Lernleistung: korrigierte Validität

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
Null-Subst.	210	30 451	.894	.639	.785	.585	.205	.101	.178	24.4	.56–.61	.29	TRA
VW-Subst.	210	30 451	.894	.639	.785	.587	.199	.101	.171	25.9	.56–.61	.31	TRA
r_k -Subst.	210	30 451	.894	.639	.785	.592	.192	.101	.163	27.6	.57–.62	.32	TRA
Exkl. n.ber.	196	29 111	.893	.639	.785	.602	.187	.099	.159	28.0	.58–.63	.34	TRA
Exkl. MAg	169	25 518	.892	.641	.787	.610	.185	.098	.157	27.9	.58–.64	.35	TRA
vollständig korrigierte Validität													
Null-Subst.	210	30 451	.894	.639	.785	.618	.216	.106	.188	24.1	.59–.65	.31	TRA
VW-Subst.	210	30 451	.894	.639	.785	.621	.210	.106	.181	25.5	.59–.65	.32	TRA
r_k -Subst.	210	30 451	.894	.639	.785	.626	.202	.105	.173	27.2	.60–.65	.34	TRA
Exkl. n.ber.	196	29 111	.893	.639	.785	.637	.197	.103	.168	27.2	.61–.66	.36	TRA
Exkl. MAg	169	25 518	.892	.641	.787	.646	.194	.101	.166	26.9	.62–.68	.37	TRA

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; TRA = transportability.

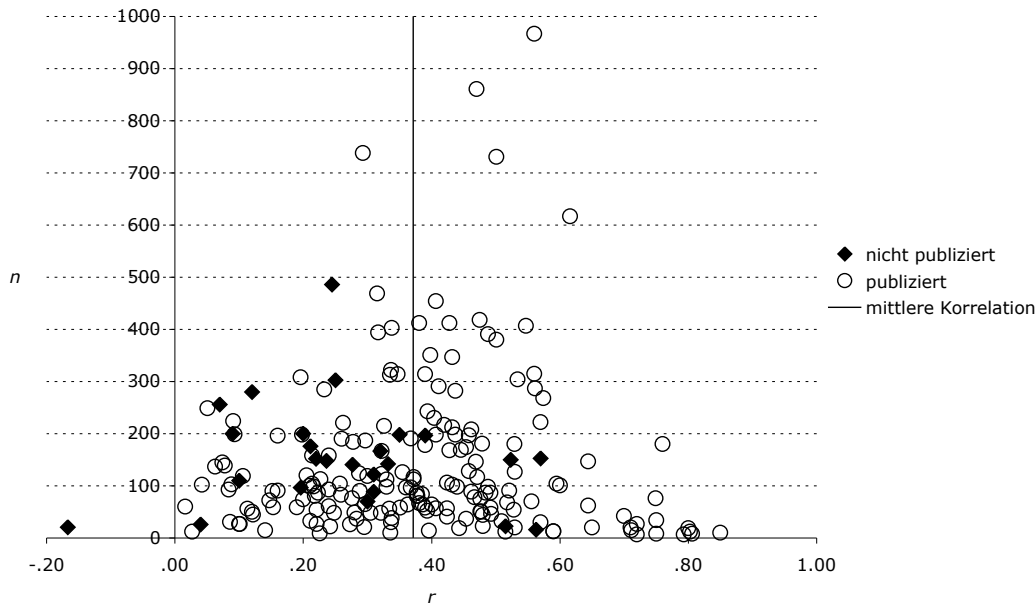
um \bar{r} nicht für jede Stichprobengröße symmetrisch. Es fehlen Studien mit kleinem n (< 200), die negative Korrelationen berichten, sowie Studien mit großem n (> 600), die Korrelationen berichten, welche kleiner als \bar{r} sind.

Zum anderen weisen die nicht-publizierten Studien überwiegend unterdurchschnittliche Korrelationen auf. Getrennte Metaanalysen für beide Gruppen zeigen dementsprechend deutliche Unterschiede (Tabellen 4.6 und 4.7): Die nicht-publizierten Studien (Variable PART = 1) weisen geringere mittlere Effekte auf ($\bar{r} = .254$, $\hat{\rho} = .424$ bei Verhältnissubstitution) als publizierte Studien ($1 < \text{PART} < 9$, $\bar{r} = .389$, $\hat{\rho} = .650$).

Die Höhe der Publikationsverzerrung kann mit der *trim and fill*-Methode von Duval (2005) geschätzt werden. Dazu werden zunächst die Studien exkludiert, die am weitesten über dem mittleren Effekt liegen. Dieser *trim*-Schritt wird so lange wiederholt, bis eine Symmetrie der Studien hergestellt ist. Der resultierende mittlere Effekt schätzt die Korrelation ohne Publikationsverzerrung, führt aber zu einer Unterschätzung der Varianz von $\hat{\rho}$. Deshalb werden im zweiten Schritt (*fill*) die zuvor exkludierten Studien und zusätzlich symmetrische Pendants von ihnen inkludiert. Der mittlere Effekt bleibt dabei unverändert. Duval (2005) empfiehlt zwei Kennwerte, mit denen die Anzahl der zu exkludierenden Studien bestimmt werden kann, bis Symmetrie erreicht ist: R_0 (Formel A.29) und L_0 (Formel A.30). Im vorliegenden Fall sind beide Kennwerte negativ (-2 bzw. -4), d.h. dass (weniger als) Null Studien zu exkludieren sind. Obwohl sich publizierte von

Abbildung 4.4

Funnelplot der Studien zu berufsbezogener Lernleistung



Anmerkung. Dargestellt sind beobachtete Korrelationen, nicht-berichtete sind verhältnissubstituiert.

Tabelle 4.6

Lernleistung: bare-bone Sensitivitätsanalyse publizierter und nicht-publizierter Studien

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
publizierte Studien									
Null-Subst.	185	26 528	.386	.151	.072	.133	22.6	.36–.41	530
VW-Subst.	185	26 528	.389	.147	.072	.128	23.8	.37–.41	534
r_k -Subst.	185	26 528	.392	.141	.071	.121	25.7	.37–.41	541
Exkl. n.ber.	171	25 188	.402	.136	.070	.116	26.4	.38–.42	518
Exkl. MAg	147	22 119	.410	.132	.068	.112	27.1	.39–.43	457
nicht-publizierte Studien									
ohne Subst.	25	3 923	.254	.129	.075	.105	34.0	.20–.30	39
Exkl. MAg	22	3 399	.260	.137	.076	.115	30.2	.20–.32	36

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k .

Tabelle 4.7

Lernleistung: korrigierte Validitäten publizierter und nicht-publizierter Studien

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
publizierte Studien													
Null-Subst.	185	26 528	.895	.637	.781	.612	.194	.100	.167	26.4	.58–.64	.34	TRA
VW-Subst.	185	26 528	.895	.637	.781	.615	.187	.100	.158	28.3	.59–.64	.35	TRA
r_k -Subst.	185	26 528	.895	.637	.781	.620	.177	.099	.147	31.4	.59–.65	.38	TRA
Exkl. n.ber.	171	25 188	.893	.636	.781	.633	.168	.097	.137	33.3	.61–.66	.41	TRA
Exkl. MAg	147	22 199	.892	.639	.781	.643	.162	.095	.131	34.6	.62–.67	.43	TRA
nicht-publizierte Studien													
ohne Subst.	25	3 923	.892	.657	.815	.400	.178	.108	.142	36.8	.33–.47	.17	TRA
Exkl. MAg	22	3 399	.891	.659	.822	.405	.188	.107	.154	32.6	.33–.48	.15	TRA
vollständig korrigierte Validität													
publizierte Studien													
Null-Subst.	185	26 528	.895	.637	.781	.647	.205	.105	.176	26.1	.62–.68	.36	TRA
VW-Subst.	185	26 528	.895	.637	.781	.650	.197	.104	.167	28.0	.62–.68	.38	TRA
r_k -Subst.	185	26 528	.895	.637	.781	.656	.186	.104	.155	30.9	.63–.68	.40	TRA
Exkl. n.ber.	171	25 188	.893	.636	.781	.670	.177	.100	.145	32.3	.64–.70	.43	TRA
Exkl. MAg	147	22 199	.892	.639	.781	.680	.170	.098	.138	33.4	.65–.71	.45	TRA
nicht-publizierte Studien													
ohne Subst.	25	3 923	.892	.657	.815	.424	.189	.116	.150	37.4	.35–.50	.18	TRA
Exkl. MAg	22	3 399	.891	.659	.822	.430	.200	.115	.163	33.1	.35–.51	.16	TRA

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; TRA = transportability.

nicht-publizierten Studien unterscheiden, kann dies als Indiz *gegen* das Vorliegen einer Publikationsverzerrung gewertet werden.

Quellenverzerrung: Zur Frage, ob sich eine *Quellenverzerrung* nachweisen lässt, werden die Metaanalysen getrennt für verschiedene Veröffentlichungsarten berechnet und untereinander sowie mit den nicht-veröffentlichten Studien verglichen. Die Veröffentlichungsarten werden zu den drei Gruppen a) *Testmanuale* (PART = 6), b) *Artikel in Zeitschriften oder Beiträge in Herausgeberwerken* (PART = 3 ∨ 5) und c) *Dissertationen bzw. Monographien ohne Testmanuale* (PART = 2 ∨ 4) zusammengefasst. Diese Einteilung in drei Gruppen wurde gewählt, da nicht alle kodierten Publikationsarten über ausreichend Studien für eine Metaanalyse verfügen (vgl. Tabelle 4.1). Ausschlaggebend für die Zusammenfassung ist die Hypothese, dass in kürzeren Artikeln auf Grund des begrenzten Raumes eher bedeutsame (d.h. signifikante) Ergebnisse dargestellt werden. In längeren Artikeln (Dissertationen und Monographien) ist diese Einschränkung nicht vonnöten. Es lässt sich zudem vermuten, dass in Testmanualen vor allem Ergebnisse berichtet werden, die die Validität des Testverfahrens stützen.

Die metaanalytischen Ergebnisse sprechen in der Tendenz für die genannten Unterschiede (Tabellen 4.8 und 4.9). Ergebnisse aus kürzeren Artikeln und längeren Monographien unterscheiden sich jedoch nur minimal voneinander. Die Validitäten bei Verhältnissubstitution betragen $\bar{r} = .336$ vs. $.316$ und $\hat{\rho} = .585$ vs. $.556$. Die Ergebnisse bei Exklusion von Studien, die insignifikante Korrelationen nicht berichten, übersteigen das theoretische Maximum. Die Validitäten, die in Testmanualen berichtet werden, sind deutlich höher als die der eben genannten beiden Gruppen ($\bar{r} = .434$, $\hat{\rho} = .702$, Substitution nicht erforderlich). Die geringsten Validitäten weisen nicht veröffentlichte Studien auf (s.o.), sie unterscheiden sich im unkorrigierten Fall deutlich von den Testmanualen und im vollständig korrigierten Fall deutlich von den Testmanualen und den Artikeln.

Die Exklusion von Studien mit Mittelwert-Aggregaten hat kaum Einfluss auf die Studien aus Testmanualen und die nicht-publizierten Studien. Bei Artikeln in Zeitschriften oder Herausgeberwerken sind die mittleren Effekte bei dieser Exklusion allerdings deutlich höher als bei ihrer Inklusion ($\bar{r} = .397$ vs. $.336$; $\hat{\rho} = .632$ vs. $.552$, vollständig korrigiert). Das Gegenteil ist der Fall bei Studien aus Monographien bzw. Dissertationen (Exklusion vs. Inklusion: $\bar{r} = .253$ vs. $.316$; $\hat{\rho} = .443$ vs. $.528$). Im Regelfall nimmt der mittlere Effekt bei Exklusion der Studien mit Mittelwert-Aggregaten ab. Dies ist hier nicht der Fall, da die Studie von Hossiep (1995) exkludiert wird, die vergleichsweise hohe Korrelationen berichtet ($\bar{r} = .470$) und einen starken Einfluss auf das Ergebnis dieser Sensitivitätsanalyse hat: ihr Gewicht bei der bare-bone Analyse beträgt 27.5%.

Der Anteil der durch den Stichprobenfehler erklärten Varianz liegt in allen Analysen zur Quellenverzerrung unter 75%, die vollständig korrigierten Ergebnisse können somit transportabel generalisiert werden.

Einfluss der Studien aus dem BET-Manual: Der Einfluss eines Testmanuals auf das Gesamtergebnis soll gesondert geprüft werden: es handelt sich um das BET-Manual, das mit 70 Studien ein großes Gewicht ausübt. Die Validität der Studien aus dem BET-

Tabelle 4.8

Lernleistung: bare-bone Analyse zu verschiedenen Quellen

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
Testmanuale									
ohne Subst.	81	14 915	.434	.100	.060	.079	36.4	.41–.46	271
Exkl. MAg	80	14 735	.433	.100	.060	.079	36.5	.41–.45	267
Artikel in Zeitschriften oder Herausgeberwerken									
Null-Subst.	69	8 482	.329	.190	.082	.172	18.4	.28–.37	159
VW-Subst.	69	8 482	.336	.181	.081	.162	20.1	.29–.38	163
r_k -Subst.	69	8 482	.347	.170	.080	.150	22.4	.31–.39	171
Exkl. n.ber.	58	7 383	.371	.166	.078	.147	21.9	.33–.41	158
Exkl. MAg	44	5 809	.397	.165	.074	.147	20.4	.35–.45	131
Monographien (ohne Testmanuale) und Dissertationen									
Null-Subst.	35	3 131	.315	.154	.097	.119	39.7	.26–.37	76
VW-Subst.	35	3 131	.316	.153	.097	.119	40.0	.27–.37	76
r_k -Subst.	35	3 131	.319	.152	.097	.118	40.2	.27–.37	77
Exkl. n.ber.	32	2 890	.321	.158	.096	.125	37.2	.27–.37	71
Exkl. MAg	23	1 575	.253	.142	.115	.082	66.4	.20–.31	36
nicht veröffentlicht									
ohne Subst.	25	3 923	.254	.129	.075	.105	34.0	.20–.30	39
Exkl. MAg	22	3 399	.260	.137	.076	.115	30.2	.20–.32	36

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k .

Tabelle 4.9

Lernleistung: korrigierte Validitäten für verschiedene Quellen

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
Testmanuale													
ohne Subst.	81	14 915	.894	.639	.785	.664	.119	.079	.089	44.0	.64–.69	.52	TRA
Exkl. MAg	80	14 735	.894	.639	.783	.664	.120	.079	.090	43.4	.64–.69	.51	TRA
Artikel in Zeitschriften oder Herausgeberwerken													
Null-Subst.	69	8 482	.892	.628	.773	.543	.259	.118	.230	20.9	.48–.60	.16	TRA
VW-Subst.	69	8 482	.892	.628	.773	.552	.243	.118	.213	23.5	.50–.61	.20	TRA
r_k -Subst.	69	8 482	.892	.628	.773	.567	.224	.116	.191	27.1	.51–.62	.25	TRA
Exkl. n.ber.	58	7 383	.889	.624	.771	.601	.212	.113	.180	28.1	.55–.66	.30	TRA
Exkl. MAg	44	5 809	.887	.631	.772	.632	.203	.109	.171	28.7	.57–.69	.35	TRA
Monographien (ohne Testmanuale) und Dissertationen													
Null-Subst.	35	3 131	.901	.647	.788	.526	.214	.144	.158	45.4	.46–.60	.27	TRA
VW-Subst.	35	3 131	.901	.647	.788	.528	.213	.144	.156	45.9	.46–.60	.27	TRA
r_k -Subst.	35	3 131	.901	.647	.788	.532	.211	.144	.154	46.7	.46–.60	.28	TRA
Exkl. n.ber.	32	2 890	.897	.648	.788	.538	.217	.143	.163	43.5	.46–.61	.27	TRA
Exkl. MAg	23	1 575	.898	.650	.794	.443	.221	.183	.124	68.6	.35–.53	.24	TRA
nicht veröffentlicht													
ohne Subst.	25	3 923	.892	.657	.815	.400	.178	.108	.142	36.8	.33–.47	.17	TRA
Exkl. MAg	22	3 399	.891	.659	.822	.405	.188	.107	.154	32.6	.33–.48	.15	TRA
vollständig korrigierte Validität													
Testmanuale													
ohne Subst.	81	14 915	.894	.639	.785	.702	.127	.085	.094	44.8	.67–.73	.55	TRA
Exkl. MAg	80	14 735	.894	.639	.783	.702	.128	.085	.096	44.3	.67–.73	.54	TRA
Artikel in Zeitschriften oder Herausgeberwerken													
Null-Subst.	69	8 482	.892	.628	.773	.575	.272	.121	.244	19.6	.51–.64	.17	TRA
VW-Subst.	69	8 482	.892	.628	.773	.585	.255	.120	.225	22.0	.52–.65	.21	TRA
r_k -Subst.	69	8 482	.892	.628	.773	.601	.234	.118	.202	25.3	.55–.66	.27	TRA
Exkl. n.ber.	58	7 383	.889	.624	.771	.637	.221	.111	.191	25.2	.58–.69	.32	TRA
Exkl. MAg	44	5 809	.887	.631	.772	.671	.209	.103	.182	24.4	.61–.73	.37	TRA
Monographien (ohne Testmanuale) und Dissertationen													
Null-Subst.	35	3 131	.901	.647	.788	.555	.226	.153	.167	45.7	.48–.63	.28	TRA
VW-Subst.	35	3 131	.901	.647	.788	.556	.225	.152	.165	46.1	.48–.63	.29	TRA
r_k -Subst.	35	3 131	.901	.647	.788	.561	.222	.152	.162	46.8	.49–.63	.29	TRA
Exkl. n.ber.	32	2 890	.897	.648	.788	.568	.229	.151	.172	43.5	.49–.65	.28	TRA
Exkl. MAg	23	1 575	.898	.650	.794	.468	.233	.193	.131	68.7	.37–.56	.25	TRA
nicht veröffentlicht													
ohne Subst.	25	3 923	.892	.657	.815	.424	.189	.116	.150	37.4	.35–.50	.18	TRA
Exkl. MAg	22	3 399	.891	.659	.822	.430	.200	.115	.163	33.1	.35–.51	.16	TRA

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; TRA = transportability.

Tabelle 4.10

Lernleistung: bare-bone Analyse zum Einfluss des BET-Manuals

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
alle Studien zu berufsbezogener Lernleistung ohne die Studien aus dem BET-Manual									
Null-Subst.	140	17 605	.330	.179	.080	.160	20.3	.30–.36	322
VW-Subst.	140	17 605	.333	.174	.080	.155	21.2	.30–.36	327
r_k -Subst.	140	17 605	.339	.169	.080	.148	22.5	.31–.37	335
Exkl. n.ber.	126	16 265	.350	.168	.078	.149	21.6	.32–.38	316
Exkl. MAg	99	12 672	.357	.174	.078	.155	20.2	.32–.39	255
Studien aus dem BET-Manual									
ohne Subst.	70	12 846	.423	.090	.061	.066	45.7	.40–.44	227

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k .

Manual liegt über derjenigen der anderen Studien (im vollständig korrigierten Fall: $\hat{\rho} = .688$ vs. $.570$; vgl. Tabellen 4.10 und 4.11). Dieser Unterschied ist bedeutsam: die Konfidenzintervalle überlappen sich nicht.

Die Ergebnisse der verschiedenen Substitutions- und Exklusionsverfahren entsprechen auch dann den Erwartungen, wenn die Studien aus dem BET-Manual exkludiert werden: die mittleren Effekte bei Verhältnissubstitution sind kleiner als bei Substitution mit r_k , diese kleiner als bei Exklusion der nicht-berichteten Werte. Studien mit Mittelwert-Aggregaten reduzieren den mittleren Effekt. Selbst bei ihrer Exklusion wird die Validität der BET-Studien nicht erreicht.

Tabelle 4.11

Lernleistung: korrigierte Validitäten zum Einfluss des BET-Manuals

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
alle Studien zu berufsbezogener Lernleistung ohne die Studien aus dem BET-Manual													
Null-Subst.	140	17 605	.894	.639	.786	.534	.242	.114	.214	22.1	.49–.57	.18	TRA
VW-Subst.	140	17 605	.894	.639	.786	.538	.236	.114	.206	23.3	.50–.58	.20	TRA
r_k -Subst.	140	17 605	.894	.639	.786	.547	.226	.113	.195	25.2	.51–.58	.23	TRA
Exkl. n.ber.	126	16 265	.892	.638	.786	.562	.224	.111	.195	24.5	.52–.60	.24	TRA
Exkl. MAg	99	12 672	.890	.643	.788	.568	.231	.110	.203	22.7	.52–.61	.23	TRA
Studien aus dem BET-Manual													
ohne Subst.	70	12 846	.895	.639	.784	.651	.111	.083	.074	55.4	.62–.68	.53	TRA
vollständig korrigierte Validität													
alle Studien zu berufsbezogener Lernleistung ohne die Studien aus dem BET-Manual													
Null-Subst.	140	17 605	.894	.639	.786	.565	.256	.120	.226	21.9	.52–.61	.19	TRA
VW-Subst.	140	17 605	.894	.639	.786	.569	.249	.120	.218	23.1	.53–.61	.21	TRA
r_k -Subst.	140	17 605	.894	.639	.786	.578	.238	.118	.206	24.8	.54–.62	.24	TRA
Exkl. n.ber.	126	16 265	.892	.638	.786	.595	.236	.115	.206	23.7	.55–.64	.26	TRA
Exkl. MAg	99	12 672	.890	.643	.788	.602	.244	.114	.216	21.8	.55–.65	.25	TRA
Studien aus dem BET-Manual													
ohne Subst.	70	12 846	.895	.639	.784	.688	.117	.087	.078	55.4	.66–.72	.56	TRA

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; TRA = transportability.

Tabelle 4.12

Lernleistung: bare-bone Analyse zum Einfluss der Veröffentlichungen vor 1945

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
Studien vor 1945									
ohne Subst.	9	121	.710	.101	.142	.000 ^a	100.0 ^a	.62–.80	55
Exkl. MAg	6	89	.715	.083	.132	.000 ^a	100.0 ^a	.61–.82	37
Studien nach 1945									
Null-Subst.	201	30 330	.368	.154	.071	.136	21.4	.35–.39	539
VW-Subst.	201	30 330	.370	.150	.071	.132	22.3	.35–.39	543
r_k -Subst.	201	30 330	.373	.146	.071	.127	23.6	.35–.39	550
Exkl. n.ber.	187	28 990	.381	.143	.069	.125	23.6	.36–.40	526
Exkl. MAg	163	25 429	.389	.141	.068	.123	23.7	.37–.41	472

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k . ^a *out of range*.

Studien vor 1945: Die Studien vor 1945 zeichnen sich durch hohe beobachtete Validitäten aus ($\bar{r} = .710$), die im vollständig korrigierten Fall einen Wert größer 1.00 (*out of range*) annehmen (vgl. Tabellen 4.12 und 4.13). Auf Grund der geringen Stichprobengröße der Studien vor 1945 von $\bar{n} = 13.4$ ist ihr Einfluss auf die verschiedenen Kennzahlen der Gesamtmetaanalyse sehr gering. Als Ursachen für die hohen berichteten Validitäten der Studien vor 1945 kommen andere Designansätze in Frage. Roloff (1928) hat z.B. nicht die gesamte anfallende Stichprobe ausgewertet, sondern auf eine hohe Variabilität der Leistungsbewertung geachtet. Es können auch Zufallsbefunde sein (vgl. Funnelplot, Abbildung 4.4).

Prädiktive vs. konkurrente Validität: Im Rahmen der Personalauswahl ist die Frage nach der prädiktiven Validität von Intelligenz von besonderem Interesse. Im Folgenden werden deshalb die Studien zur konkurrenten (VTYP = 4 & TDIFF = 0) und prädiktiven Validität (VTYP < 4 & TDIFF > 0) getrennten Metaanalysen unterzogen. Die retrograde Studie von Kersting (1999a) sowie die drei Studien mit unklarer Abfolge der Intelligenz- und Kriterienerhebung von Roloff (1928) bleiben unberücksichtigt.

Die konkurrenten Validitäten sind geringer als die prädiktiven ($\bar{r} = .294$ vs. $.372$, $\hat{\rho} = .546$ vs. $.621$, vollst. korr. bei Verhältnissubstitution; vgl. Tabellen 4.14 und 4.15). Die verschiedenen Substitutions- und Exklusionsverfahren haben auf das Ergebnis der prädiktiven Studien nur einen geringen Einfluss, bei den konkurrenten Studien ist keine Substitution bzw. Exklusion erforderlich.

Bei vollständiger Korrektur der konkurrenten Studien liegt die Varianz, die auf Artefakte zurückzuführen ist ($\hat{\rho} = .031$), über der beobachteten Varianz ($\hat{\rho} = .026$). Dies hat zur Folge, dass die Varianz von $\hat{\rho}$ negativ geschätzt wird und die Varianzaufklärung

Tabelle 4.13

Lernleistung: korrigierte Validitäten zum Einfluss der Veröffentlichungen vor 1945

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
Studien vor 1945													
ohne Subst.	9	121	.895	.520	.791	.990	.078	.078	.00 ^a	100 ^a	.94–1.0 ^a	.99	CSC
Exkl. MAg	6	89	.895	.520	.791	.993	.062	.062	.00 ^a	100 ^a	.94–1.0 ^a	.99	CSC
Studien nach 1945													
Null-Subst.	201	30 330	.894	.644	.785	.583	.203	.100	.177	24.2	.55–.61	.29	TRA
VW-Subst.	201	30 330	.894	.644	.785	.585	.197	.100	.170	25.7	.56–.61	.31	TRA
r_k -Subst.	201	30 330	.894	.644	.785	.590	.190	.100	.162	27.4	.56–.62	.32	TRA
Exkl. n.ber.	187	28 990	.893	.644	.785	.600	.186	.098	.158	27.8	.57–.63	.34	TRA
Exkl. MAg	163	25 429	.892	.646	.786	.609	.183	.097	.156	27.8	.58–.64	.35	TRA
vollständig korrigierte Validität													
Studien vor 1945													
ohne Subst.	9	121	.895	.520	.791	1.0 ^a	.082	.152	.00 ^a	100 ^a	.99–1.0 ^a	1.0 ^a	CSC
Exkl. MAg	6	89	.895	.520	.791	1.0 ^a	.065	.141	.00 ^a	100 ^a	1.0–1.1 ^a	1.1 ^a	CSC
Studien nach 1945													
Null-Subst.	201	30 330	.894	.644	.785	.616	.215	.105	.187	24.0	.59–.65	.31	TRA
VW-Subst.	201	30 330	.894	.644	.785	.619	.208	.105	.180	25.3	.59–.65	.32	TRA
r_k -Subst.	201	30 330	.894	.644	.785	.624	.201	.104	.171	27.0	.60–.65	.34	TRA
Exkl. n.ber.	187	28 990	.893	.644	.785	.635	.195	.101	.167	26.9	.61–.66	.36	TRA
Exkl. MAg	163	25 429	.892	.646	.786	.645	.192	.100	.165	26.8	.62–.67	.37	TRA

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; CSC = cross-situational consistency, TRA = transportability. ^a *out of range*.

Tabelle 4.14

Lernleistung: bare-bone Sensitivitätsanalyse zum Einfluss konkurrender Validierungen

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
Prädiktive Validität									
Null-Subst.	197	29 751	.370	.155	.071	.138	20.9	.35–.39	533
VW-Subst.	197	29 751	.372	.151	.071	.134	21.8	.35–.39	537
r_k -Subst.	197	29 751	.376	.147	.071	.129	23.1	.36–.40	544
Exkl. n.ber.	183	28 411	.384	.144	.069	.126	23.1	.36–.40	520
Exkl. MAg	166	25 105	.393	.141	.069	.123	24.2	.37–.41	486
Konkurrenente Validität									
ohne Subst.	8	563	.294	.118	.111	.041	88.0	.22–.37	16

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k .

über 100% liegt. Damit liegen beide Werte *out of range*. In diesem Fall gilt: „Because the estimated variance is negative, the estimated standard deviation is 0“ (Hunter & Schmidt, 2004, S. 89).

Da die Varianzaufklärung durch die Artefakte bei konkurrenten Studien größer ist als 75%, kann der mittlere Effekt generalisiert werden. Allerdings basiert das Ergebnis lediglich auf acht Studien: der VE-Wert ist liberal und das 95% Konfidenzintervall weit (.43 bis .66). Da \widehat{SD}_{ρ} mit .000 geschätzt wird, werden die Studien als homogen erachtet; es gilt das Modell fester Effekte (FM). Bei den prädiktiven Studien ist $\widehat{SD}_{\rho} > .000$, es gilt das Modell zufallsvariabler Effekte (RM). Die Ergebnisse zum Erfolgsbereich *berufsbezogene Lernleistung* sind noch einmal zusammengefasst in Kapitel 5.2 dargestellt.

4.1.4. Moderatoranalysen

Die durchgeführten Sensitivitätsanalysen dienen dazu, die Stabilität des metaanalytischen Gesamtergebnisses zu prüfen. Im Folgenden werden Moderatoranalysen mit dem Ziel durchgeführt, eine deskriptive Beschreibung der Validität von Intelligenz in Bezug auf Lernleistung zu erreichen – und zwar getrennt für verschieden komplexe Berufe (Skill Level), verschiedene Berufsgruppen, Intelligenztests und Kriterien (Fragestellungen 2 bis 4, vgl. Kapitel 2.3.4).

Berufskomplexität und Berufsgruppen

Zur Prüfung der Komplexitätsfragestellung werden die einzelnen Skill Level getrennt analysiert (Variable ISCOSL). Genügend Daten liegen vor für die Skill Level 2 ($k = 138$, $N = 21\,551$) und 3 ($k = 50$, $N = 5\,198$). In beiden Fällen übersteigt die mittlere

Tabelle 4.15

Lernleistung: korrigierte Validitäten zum Einfluss konkurrenter Validität

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
prädiktiv													
Null-Subst.	197	29 751	.895	.641	.786	.585	.205	.099	.180	23.5	.56–.61	.29	TRA
VW-Subst.	197	29 751	.895	.641	.786	.588	.200	.099	.173	24.8	.56–.62	.30	TRA
r_k -Subst.	197	29 751	.895	.641	.786	.592	.192	.099	.165	26.6	.57–.62	.32	TRA
Exkl. n.ber.	183	28 411	.893	.641	.786	.602	.187	.097	.160	26.8	.58–.63	.34	TRA
Exkl. MAg	166	25 105	.892	.642	.787	.612	.184	.097	.156	27.8	.58–.64	.36	TRA
konkurrent													
ohne Subst.	8	563	.895	.627	.775	.517	.153	.153	.00 ^a	100 ^a	.41–.62	.52	CSC
vollständig korrigierte Validität													
prädiktiv													
Null-Subst.	197	29 751	.895	.641	.786	.619	.217	.104	.190	23.2	.59–.65	.31	TRA
VW-Subst.	197	29 751	.895	.641	.786	.621	.211	.104	.183	24.4	.59–.65	.32	TRA
r_k -Subst.	197	29 751	.895	.641	.786	.626	.203	.103	.174	26.1	.60–.65	.34	TRA
Exkl. n.ber.	183	28 411	.893	.641	.786	.637	.197	.101	.170	26.0	.61–.67	.36	TRA
Exkl. MAg	166	25 105	.892	.642	.787	.648	.194	.100	.166	26.8	.62–.68	.38	TRA
konkurrent													
ohne Subst.	8	563	.895	.627	.775	.546	.161	.176	.00 ^a	100 ^a	.43–.66	.55	CSC

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; CSC = cross-situational consistency, TRA = transportability. ^a *out of range*.

Tabelle 4.16

Lernleistung: bare-bone-Analysen getrennt für Skill Level 2 und 3

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
Skill Level 2									
Null-Subst.	138	21 933	.397	.132	.068	.113	26.4	.37–.42	410
VW-Subst.	138	21 933	.397	.131	.068	.112	26.8	.38–.42	411
r_k -Subst.	138	21 933	.398	.129	.067	.110	27.2	.38–.42	412
Exkl. n.ber.	133	21 551	.400	.129	.067	.110	26.8	.38–.42	400
Exkl. MAg	117	19 026	.407	.127	.066	.109	26.9	.38–.43	360
Skill Level 3									
Null-Subst.	50	5 198	.297	.194	.090	.172	21.7	.24–.35	99
VW-Subst.	50	5 198	.305	.183	.090	.160	24.1	.25–.36	103
r_k -Subst.	50	5 198	.319	.170	.089	.144	27.6	.27–.37	110
Exkl. n.ber.	42	4 316	.351	.166	.087	.142	27.6	.30–.40	106
Exkl. MAg	32	3 442	.385	.155	.083	.131	28.5	.33–.44	92

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k .

Validität bei Exklusion der Studien, die insignifikante Korrelationen nicht berichten, die maximal mögliche (Tabellen 4.16 für bare-bone und 4.17 für korrigierte Ergebnisse). Die Ergebnisse entsprechen nicht den Annahmen: die mittlere beobachtete Korrelation bei Skill Level 2 ($\bar{r} = .397$, Verhältnissubstitution) ist höher als bei Skill Level 3 ($\bar{r} = .305$). Der Unterschied ist bei Betrachtung der Konfidenzintervalle als bedeutsam anzusehen.

Der Unterschied zwischen beiden Skill Leveln bleibt auch nach der Korrektur der Artefakte bestehen (Skill Level 2 vs. 3 bei Verhältnissubstitution: $\bar{\rho} = .649$ vs. $.542$). Die Varianzaufklärung in der bare-bone Analyse liegt bei 26.8% bzw. 24.1%, im korrigierten Fall liegt jeweils Transportabilität vor.

Die Studien des Skill Level 3 sind stärker von nicht-berichteten Korrelationen und Mittelwert-Aggregationen betroffen als die Studien des Skill Level 2 (vgl. Tabelle D.1). Die Exklusion der betroffenen Studien führt besonders beim Skill Level 3 zu einem höheren mittleren Effekt und breiteren Konfidenzintervallen. Die Unterschiede zwischen beiden Gruppen sind bei diesen Analysen nicht mehr bedeutsam, die Rangfolge der Validitäten beider Gruppen bleiben unverändert.

Den Skill Leveln liegt der allgemeinbildende Schulabschluss zu Grunde, der zur Ausübung der Berufe vorausgesetzt wird (vgl. Kapitel 3.4). Da einige Studien mit bekannten Schulabschlüssen vorliegen und im Hinblick darauf homogen sind, können die prädiktiven Validitäten getrennt für Stichproben errechnet werden, die höchstens über den Hauptschulabschluss verfügen (Hauptschul-Gruppe, Variable SBNIV1 = 100, $k = 7$, $N = 847$) bzw. die höhere Schulabschlüsse als die mittlere Reife haben (Hochschulreife-Gruppe,

Tabelle 4.17

Lernleistung: korrigierte Validitäten getrennt für Skill Level 2 und 3

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
Skill Level 2													
Null-Subst.	138	21 933	.895	.635	.794	.613	.172	.095	.143	30.5	.58–.64	.38	TRA
VW-Subst.	138	21 933	.895	.635	.794	.614	.170	.095	.141	31.1	.59–.64	.38	TRA
r_k -Subst.	138	21 933	.895	.635	.794	.615	.169	.095	.140	31.6	.59–.64	.39	TRA
Exkl. n.ber.	133	21 551	.895	.636	.794	.618	.167	.094	.139	31.4	.59–.65	.39	TRA
Exkl. MAg	117	19 026	.894	.638	.795	.626	.165	.094	.136	32.2	.60–.66	.40	TRA
Skill Level 3													
Null-Subst.	50	5 198	.894	.642	.768	.501	.278	.126	.247	20.8	.42–.58	.09	TRA
VW-Subst.	50	5 198	.894	.642	.768	.513	.260	.126	.227	23.5	.44–.58	.14	TRA
r_k -Subst.	50	5 198	.894	.642	.768	.531	.237	.124	.202	27.4	.46–.60	.20	TRA
Exkl. n.ber.	42	4 316	.891	.642	.771	.573	.228	.119	.194	27.4	.50–.64	.25	TRA
Exkl. MAg	32	3 442	.888	.642	.771	.618	.204	.111	.172	29.5	.55–.69	.34	TRA
vollständig korrigierte Validität													
Skill Level 2													
Null-Subst.	138	21 933	.895	.635	.794	.648	.180	.097	.151	29.2	.62–.68	.40	TRA
VW-Subst.	138	21 933	.895	.635	.794	.649	.178	.097	.149	29.7	.62–.68	.40	TRA
r_k -Subst.	138	21 933	.895	.635	.794	.650	.176	.097	.147	30.2	.62–.68	.41	TRA
Exkl. n.ber.	133	21 551	.895	.636	.794	.653	.175	.095	.147	29.7	.62–.68	.41	TRA
Exkl. MAg	117	19 026	.894	.638	.795	.662	.172	.094	.144	30.1	.63–.69	.43	TRA
Skill Level 3													
Null-Subst.	50	5 198	.894	.642	.768	.530	.296	.140	.261	22.3	.45–.61	.10	TRA
VW-Subst.	50	5 198	.894	.642	.768	.542	.277	.139	.240	25.0	.47–.62	.15	TRA
r_k -Subst.	50	5 198	.894	.642	.768	.561	.254	.136	.214	28.9	.49–.63	.21	TRA
Exkl. n.ber.	42	4 316	.891	.642	.771	.607	.243	.130	.206	28.7	.53–.68	.27	TRA
Exkl. MAg	32	3 442	.888	.642	.771	.656	.218	.121	.182	30.6	.58–.73	.36	TRA

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; TRA = transportability.

Tabelle 4.18

Lernleistung: bare-bone-Analyse zum Vergleich geringer und hoher Bildungsniveaus

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
maximal Hauptschul-Niveau									
ohne Subst.	7	847	.191	.131	.088	.096	45.7	.09–.29	7
Exkl. MAg	6	623	.228	.135	.094	.097	48.2	.12–.34	8
(Fach-)Hochschulreife-Niveau									
Null-Subst.	17	2018	.136	.136	.091	.102	44.2	.07–.20	7
VW-Subst.	17	2018	.158	.124	.090	.086	52.5	.10–.22	10
r_k -Subst.	17	2018	.191	.116	.089	.075	58.7	.14–.25	16
Exkl. n.ber.	9	1180	.205	.135	.084	.106	38.6	.12–.29	10
Exkl. MAg	7	791	.245	.151	.089	.122	34.8	.13–.36	11

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k .

Variable SBNIV3 = 100, k = 17, N = 2018).

Die Validität der Hauptschul-Gruppe beträgt \bar{r} = .191 bzw. $\hat{\rho}$ = .355 und die der Hochschulreife-Gruppe \bar{r} = .158 bzw. $\hat{\rho}$ = .288 (Tabellen 4.18 und 4.19). Acht der 17 Studien mit (Fach-)Hochschülern enthalten insignifikante Korrelationen. Werden diese Studien exkludiert, resultieren mittlere Validitäten von \bar{r} = .205 bzw. $\hat{\rho}$ = .368, die diejenigen der Hauptschul-Gruppe übersteigen. Sie liegen auch über dem maximal möglichen Wert bei Substitution mit r_k . Die Varianzaufklärung in der bare-bone Analyse liegt bei 45.7% bzw. 52.5%; bei vollständiger Artefaktkorrektur liegt Transportabilität vor.

Die Hochschulreife-Gruppe ist zudem stark von Studien beeinflusst, die Mittelwert-Aggregate beinhalten. Werden sie exkludiert, übersteigt ihr \bar{r} (.245) das der Hauptschul-Gruppe (.228) leicht, bei den korrigierten Werten ist die Rangfolge umgekehrt ($\hat{\rho}$ vollständig korrigiert: bei Hochschulreife .397 vs. bei Hauptschul-Niveau .429). Die Befunde bei Exklusion basieren auf wenigen Studien (k = 6 bzw. 7).

Für fünf Berufsgruppen (Variable ISCO88) können getrennte Metaanalysen berechnet werden: Militär (k = 7, N = 1 297), Techniker und gleichrangige nichttechnische Berufe (k = 50, N = 5 198), Bürokräfte (k = 35, N = 5 757), Handwerker (k = 86, N = 12 302) sowie Maschinenbediener (k = 8, N = 2 112). Im Falle der Techniker und Bürokräfte sind Substitutionen insignifikanter Korrelationen erforderlich. Ihre Exklusion führt zu mittleren Validitäten, die die maximal möglichen übersteigen (Tabellen 4.20 und 4.21).

Am höchsten fällt die beobachtete Validität bei den Handwerkern aus (\bar{r} = .407), gefolgt von Bürokräften (\bar{r} = .378) und Maschinenbedienern (\bar{r} = .355). Die drei Berufsgruppen unterscheiden sich nicht bedeutsam (vgl. Konfidenzintervalle in Tabelle 4.21). Die Gruppe der Techniker weist ein \bar{r} von .305 auf, das sich bedeutsam von dem Ergebnis der Handwerker unterscheidet. Die geringste Validität zeigt die militärische Berufsgruppe

Tabelle 4.19

Lernleistung: korrigierte Validitäten getrennt für geringe und hohe Bildungsniveaus

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
maximal Hauptschul-Niveau													
ohne Subst.	7	847	.908	.642	.765	.338	.218	.145	.164	43.8	.18–.50	.07	TRA
Exkl. MAg	6	623	.910	.642	.761	.410	.219	.154	.156	49.1	.23–.59	.15	TRA
(Fach-)Hochschulreife-Niveau													
Null-Subst.	17	2018	.897	.654	.800	.238	.207	.143	.150	47.6	.14–.34	-.01	SIS
VW-Subst.	17	2018	.897	.654	.800	.273	.183	.141	.116	59.7	.19–.36	.08	TRA
r_k -Subst.	17	2018	.897	.654	.800	.325	.163	.138	.087	71.7	.25–.40	.18	TRA
Exkl. n.ber.	9	1180	.887	.643	.805	.347	.197	.128	.150	42.3	.22–.48	.10	TRA
Exkl. MAg	7	791	.882	.667	.856	.373	.204	.123	.163	36.1	.22–.52	.10	TRA
vollständig korrigierte Validität													
maximal Hauptschul-Niveau													
ohne Subst.	7	847	.908	.642	.765	.355	.230	.153	.172	44.1	.18–.53	.07	TRA
Exkl. MAg	6	623	.910	.642	.761	.429	.231	.163	.164	49.6	.24–.61	.16	TRA
(Fach-)Hochschulreife-Niveau													
Null-Subst.	17	2018	.897	.654	.800	.251	.220	.153	.158	48.2	.15–.36	-.01	SIS
VW-Subst.	17	2018	.897	.654	.800	.288	.195	.151	.123	60.3	.20–.38	.09	TRA
r_k -Subst.	17	2018	.897	.654	.800	.343	.174	.148	.092	72.1	.26–.43	.19	TRA
Exkl. n.ber.	9	1180	.887	.643	.805	.368	.210	.137	.159	42.8	.23–.51	.11	TRA
Exkl. MAg	7	791	.882	.667	.856	.397	.218	.132	.174	36.6	.24–.56	.11	TRA

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; TRA = transportability, SIS = situational specificity.

Tabelle 4.20

Lernleistung: bare-bone-Analyse getrennt für ISCO-88 Berufs-Hauptgruppen

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
Hauptgruppe 0: Militär									
ohne Subst.	7	1 297	.206	.095	.071	.063	55.6	.14–.28	8
Hauptgruppe 3: Techniker und gleichrangige nichttechnische Berufe									
Null-Subst.	50	5 198	.297	.194	.090	.172	21.7	.24–.35	99
VW-Subst.	50	5 198	.305	.183	.090	.160	24.1	.25–.36	103
r_k -Subst.	50	5 198	.319	.170	.089	.144	27.6	.27–.37	110
Exkl. n.ber.	42	4 316	.351	.166	.087	.142	27.6	.30–.40	106
Exkl. MAg	32	3 442	.385	.155	.083	.131	28.5	.33–.44	92
Hauptgruppe 4: Bürokräfte									
Null-Subst.	35	5 757	.378	.120	.068	.099	31.8	.34–.42	98
VW-Subst.	35	5 757	.378	.119	.068	.097	32.4	.34–.42	98
r_k -Subst.	35	5 757	.380	.118	.067	.097	32.8	.34–.42	98
Exkl. n.ber.	33	5 708	.380	.117	.066	.097	31.5	.34–.42	93
Exkl. MAg	27	4 040	.376	.124	.071	.102	32.8	.33–.42	75
Hauptgruppe 7: Handwerker									
ohne Subst.	86	12 302	.407	.132	.071	.112	28.5	.38–.44	265
Exkl. MAg	77	11 593	.415	.126	.068	.106	29.1	.39–.44	243
Hauptgruppe 8: Maschinenbediener									
ohne Subst.	8	2 112	.355	.106	.054	.091	26.0	.28–.43	21
Exkl. MAg	7	1 964	.364	.104	.052	.091	24.8	.29–.44	19

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k .

Tabelle 4.21

Lernleistung: korrigierte Validitäten getrennt für ISCO-88 Berufs-Hauptgruppen

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD beo	SD art	\widehat{SD}_ρ	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
Hauptgruppe 0: Militär													
ohne Subst.	7	1 297	.880	.642	.776	.354	.150	.116	.096	59.4	.24–.47	.20	TRA
Hauptgruppe 3: Techniker und gleichrangige nichttechnische Berufe													
Null-Subst.	50	5 198	.894	.642	.768	.501	.278	.126	.247	20.8	.42–.58	.09	TRA
VW-Subst.	50	5 198	.894	.642	.768	.513	.260	.126	.227	23.5	.44–.58	.14	TRA
r_k -Subst.	50	5 198	.894	.642	.768	.531	.237	.124	.202	27.4	.46–.60	.20	TRA
Exkl. n.ber.	42	4 316	.891	.642	.771	.573	.228	.119	.194	27.4	.50–.64	.25	TRA
Exkl. MAg	32	3 442	.888	.642	.771	.618	.204	.111	.172	29.5	.55–.69	.34	TRA
Hauptgruppe 4: Bürokräfte													
Null-Subst.	35	5 757	.899	.638	.791	.599	.153	.095	.120	38.8	.55–.65	.40	TRA
VW-Subst.	35	5 757	.899	.638	.791	.600	.151	.095	.117	39.7	.55–.65	.41	TRA
r_k -Subst.	35	5 757	.899	.638	.791	.602	.150	.095	.117	39.8	.55–.65	.41	TRA
Exkl. n.ber.	33	5 708	.898	.642	.791	.602	.149	.092	.116	38.7	.55–.65	.41	TRA
Exkl. MAg	27	4 040	.897	.641	.791	.603	.156	.102	.118	43.0	.54–.66	.41	TRA
Hauptgruppe 7: Handwerker													
ohne Subst.	86	12 302	.894	.632	.798	.618	.177	.096	.149	29.0	.58–.66	.37	TRA
Exkl. MAg	77	11 593	.894	.635	.799	.626	.170	.092	.143	29.6	.59–.66	.39	TRA
Hauptgruppe 8: Maschinenbediener													
ohne Subst.	8	2 112	.878	.646	.776	.573	.140	.082	.114	34.4	.48–.67	.39	TRA
Exkl. MAg	7	1 964	.880	.647	.779	.582	.140	.078	.116	31.4	.48–.69	.39	TRA
vollständig korrigierte Validität													
Hauptgruppe 0: Militär													
ohne Subst.	7	1 297	.880	.642	.776	.378	.159	.123	.102	59.1	.26–.50	.21	TRA
Hauptgruppe 3: Techniker und gleichrangige nichttechnische Berufe													
Null-Subst.	50	5 198	.894	.642	.768	.530	.296	.140	.261	22.3	.45–.61	.10	TRA
VW-Subst.	50	5 198	.894	.642	.768	.542	.277	.139	.240	25.0	.47–.62	.15	TRA
r_k -Subst.	50	5 198	.894	.642	.768	.561	.254	.136	.214	28.9	.49–.63	.21	TRA
Exkl. n.ber.	42	4 316	.891	.642	.771	.607	.243	.130	.206	28.7	.53–.68	.27	TRA
Exkl. MAg	32	3 442	.888	.642	.771	.656	.218	.121	.182	30.6	.58–.73	.36	TRA
Hauptgruppe 4: Bürokräfte													
Null-Subst.	35	5 757	.899	.638	.791	.632	.162	.101	.126	39.1	.58–.69	.42	TRA
VW-Subst.	35	5 757	.899	.638	.791	.633	.160	.101	.124	40.0	.58–.69	.43	TRA
r_k -Subst.	35	5 757	.899	.638	.791	.635	.159	.101	.123	40.1	.58–.69	.43	TRA
Exkl. n.ber.	33	5 708	.898	.642	.791	.636	.157	.098	.123	38.8	.58–.69	.43	TRA
Exkl. MAg	27	4 040	.897	.641	.791	.636	.164	.106	.124	42.3	.57–.70	.43	TRA
Hauptgruppe 7: Handwerker													
ohne Subst.	86	12 302	.894	.632	.798	.653	.186	.099	.158	28.1	.61–.69	.39	TRA
Exkl. MAg	77	11 593	.894	.635	.799	.662	.178	.095	.151	28.5	.62–.7	.41	TRA

Tabelle 4.21
(Fortsetzung)

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
Hauptgruppe 8: Maschinenbediener													
ohne Subst.	8	2 112	.878	.646	.776	.611	.147	.084	.121	32.3	.51–.71	.41	TRA
Exkl. MAg	7	1 964	.880	.647	.779	.620	.147	.079	.123	29.3	.51–.73	.42	TRA

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; TRA = transportability.

($\bar{r} = .206$), sie unterscheidet sich bedeutsam von den Gruppen der Handwerker, Bürokräfte und Maschinenbediener. Die Varianzaufklärung liegt zwischen 24.1 und 55.6%.

Die in der vollständig korrigierten Analyse beobachtete Rangfolge entspricht der in der bare-bone Analyse, das gilt auch für die Bedeutsamkeit der Unterschiede (d.h. die Überlappungen der Konfidenzintervalle). Die $\hat{\rho}$ s betragen für die Handwerker .653, Bürokräfte .633, Maschinenbediener .611, Techniker .542 und die militärischen Stichproben .378 (Abbildung 4.5). In allen Fällen liegt Transportabilität vor.

Das Gewicht der unpublizierten Studien liegt in allen Gruppen zwischen 5.2% und 11.4%, bei den militärischen Studien beträgt es hingegen 87.1% (Tabelle D.1). Möglicherweise ist deshalb das Ergebnis der militärischen Gruppe weniger liberal durch einen Publikationsbias verzerrt. Der Einfluss der Studien mit Mittelwert-Aggregaten ist unterschiedlich: während die militärischen Studien keine Mittelwert-Aggregationen aufweisen, trifft dies für 33.8% der Techniker-Gruppe zu (vgl. Tabelle D.1). Ihre Exklusion resultiert in einer höheren Validität von Intelligenz in der Techniker-Gruppe ($\bar{r} = .385$, $\hat{\rho} = .656$). Sie unterscheidet sich dann nicht mehr bedeutsam von den Validitäten in den Berufsgruppen der Bürokräfte, Handwerker und Maschinenbediener. Zusammenfassungen der Ergebnisse für verschieden komplexe Berufe und Berufsgruppen enthalten die entsprechenden Kapitel (5.5 und 5.6) im Diskussionsteil.

Intelligenztest-Gruppen

Den Ergebnissen der bare-bone Metaanalysen getrennt für verschiedene Testverfahren (Tabelle 4.22) ist zu entnehmen, dass die Exklusion der Studien, die insignifikante Ergebnisse nicht berichten, zu \bar{r} s führen, die die maximal möglichen Werte übersteigen. Abgesehen von den Studien zum WIT, bei denen dieses Gewicht vergleichsweise hoch ist (30.3%), sind die Auswirkungen der Substitutionsverfahren auf die mittleren r s minimal: der Unterschied beträgt zwischen .000 bei GATB-basierten Verfahren und .015 beim LPS/PSB. Beim WIT hingegen unterscheiden sich beide Werte um .061. Die \bar{r} s auf Basis der Verhältnissubstitutionen liegen unter den \bar{r} s auf Basis der Exklusion.

Tabelle 4.22

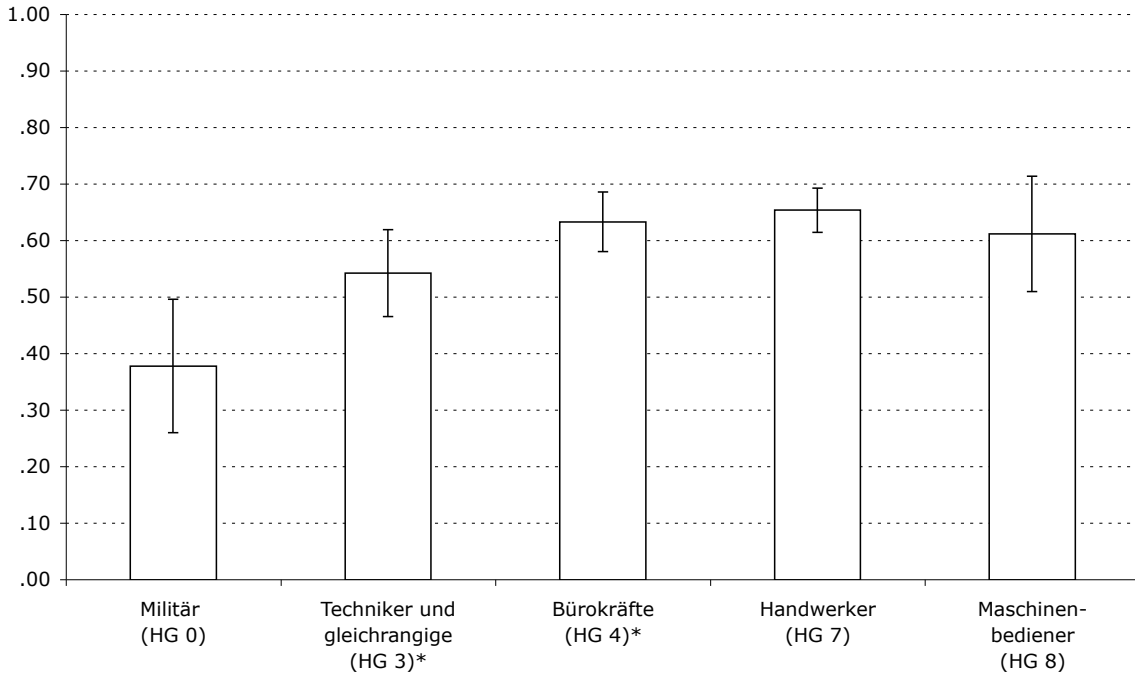
Lernleistung: bare-bone Moderatoranalyse verschiedener Intelligenztests

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
Intelligenz-Struktur-Test									
ohne Subst.	56	7 261	.359	.178	.077	.161	18.7	.31–.41	146
Wilde-Intelligenz-Test									
Null-Subst.	34	3 259	.238	.165	.098	.133	35.1	.18–.29	48
VW-Subst.	34	3 259	.253	.150	.097	.115	41.9	.20–.30	52
r_k -Subst.	34	3 259	.275	.133	.096	.092	52.2	.23–.32	60
Exkl. n.ber.	26	2 346	.314	.128	.097	.083	57.4	.27–.36	56
Exkl. MAg	23	1 929	.327	.128	.100	.080	60.8	.27–.38	53
Berliner Intelligenz-Struktur-Test									
Null-Subst.	9	883	.316	.134	.092	.097	46.9	.23–.40	20
VW-Subst.	9	883	.318	.130	.091	.092	49.5	.23–.40	20
r_k -Subst.	9	883	.325	.120	.091	.078	57.4	.25–.40	21
Exkl. n.ber.	8	838	.330	.121	.088	.084	52.2	.25–.41	19
Exkl. MAg	8	838	.330	.121	.088	.084	52.2	.25–.41	19
LPS / PSB									
Null-Subst.	15	2 542	.404	.177	.065	.164	13.5	.31–.49	46
VW-Subst.	15	2 542	.406	.174	.065	.162	13.8	.32–.49	46
r_k -Subst.	15	2 542	.409	.173	.065	.160	14.0	.32–.50	47
Exkl. n.ber.	12	2 301	.421	.176	.060	.166	11.5	.32–.52	39
Exkl. MAg	11	2 101	.442	.170	.059	.160	11.9	.34–.54	38
Matrizentests									
ohne Subst.	22	1 442	.285	.113	.116	.00 ^a	100 ^a	.24–.33	41
Exkl. MAg	15	1 187	.268	.115	.107	.041	87.4	.21–.32	26
Intelligenztests auf Basis der GATB									
Null-Subst.	74	13 155	.417	.100	.062	.078	38.9	.39–.44	235
VW-Subst.	74	13 155	.417	.099	.062	.077	39.5	.39–.44	235
r_k -Subst.	74	13 155	.417	.099	.062	.077	39.6	.39–.44	235
Exkl. n.ber.	73	13 133	.417	.099	.062	.077	39.0	.39–.44	232
Exkl. MAg	70	12 846	.423	.090	.061	.066	45.7	.40–.44	227
andere Verfahren									
Null-Subst.	24	3 263	.323	.190	.078	.173	17.1	.25–.40	54
VW-Subst.	24	3 263	.324	.188	.078	.171	17.4	.25–.40	54
r_k -Subst.	24	3 263	.327	.185	.078	.168	17.8	.25–.40	55
Exkl. n.ber.	23	3 144	.333	.186	.078	.169	17.4	.26–.41	54
Exkl. MAg	12	1 171	.260	.216	.097	.193	20.0	.14–.38	20

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k . ^a *out of range*.

Abbildung 4.5

Lernleistung: vollständig korrigierte Validitäten für verschiedene Berufsgruppen



Anmerkungen. X-Achse: ISCO-88-Berufshauptgruppen (HG); Y-Achse: $\hat{\rho}$ vollständig korrigiert bei Inklusion aller Studien (*ggf. mit Verhältnissubstitution); Fehlerbalken: 95% Konfidenzintervalle.

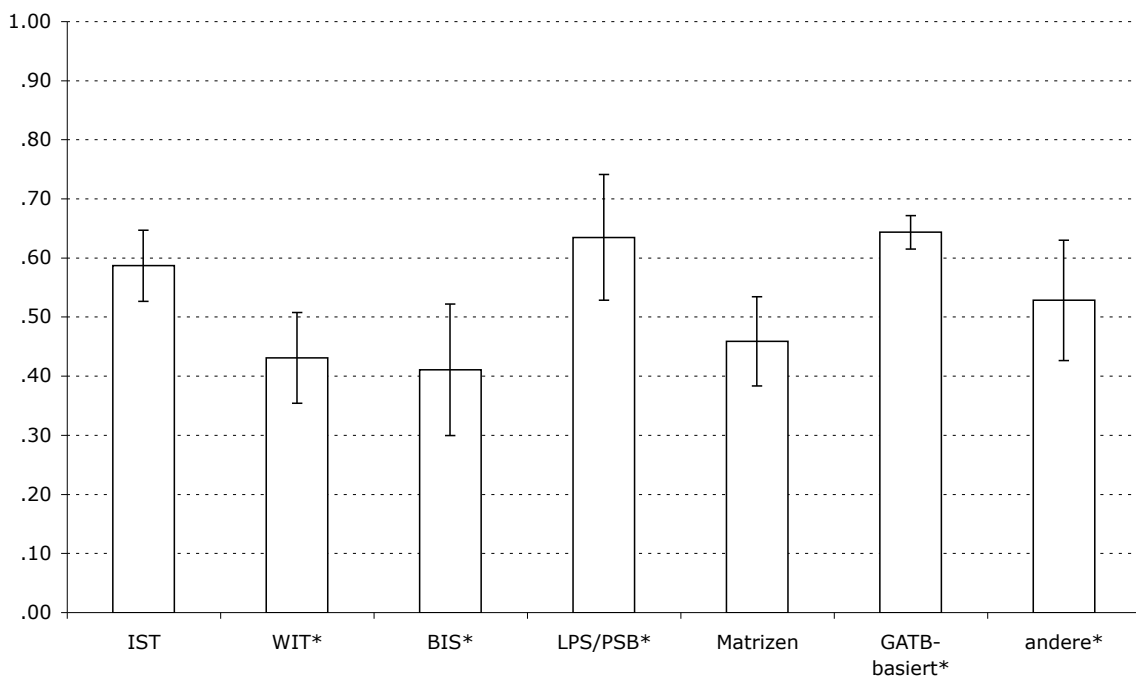
Die Auswirkungen der Substitutionsverfahren auf die Varianzen von \bar{r} und die Anteile der durch die Artefakte erklärten Varianzen sind uneinheitlich. Zum Beispiel ist die Standardabweichung von \bar{r} in den Studien, die den WIT als Prädiktor verwenden, bei Exklusion (.083) kleiner als bei der Verhältnissubstitution (.115) und die Varianzaufklärung durch die Artefakte größer (57.4% vs. 42.0%). Bei Studien, die den LPS bzw. PSB einsetzen, verhält es sich umgekehrt (vgl. Tabelle 4.22).

Die Validitäten der verschiedenen Intelligenztestverfahren sind klein ($\bar{r} = .253$ für WIT-Studien und .285 für Studien mit Matrizentests) bis mittelgroß ($\bar{r} = .318$ für BIS-Studien bis .417 für GATB-basierte Studien). Die Anzahl der inkludierten Studien reicht von $k = 9$ bei BIS-Studien bis $k = 74$ bei GATB-basierten Studien. Die unterschiedlichen k s machen sich in der Breite der Konfidenzintervalle und der Anzahl der Studien bemerkbar, die erforderlich wären, um \bar{r} unter .10 zu drücken.

Da in dieser Moderatoranalyse die Validitäten verschiedener Intelligenztests miteinander verglichen werden, werden den folgenden Ausführungen die *operationalen* Ergebnisse zu Grunde gelegt (Abbildung 4.6 und Tabelle 4.23). Die operationalen Validitäten reichen bei Verhältnissubstitution von $\hat{\rho} = .410$ (BIS) bis $\hat{\rho} = .643$ (GATB-basierte Verfahren). Die operationale Validität des WIT beträgt $\hat{\rho} = .431$, der Matrizentests $\hat{\rho} = .459$, des IST $\hat{\rho} = .587$ und von LPS/PSB $\hat{\rho} = .635$. Die Gruppe der anderen Verfahren weisen eine operationale Validität von $\hat{\rho} = .528$ auf.

Abbildung 4.6

Lernleistung: operationale Validitäten verschiedener Intelligenztests



Anmerkungen. X-Achse: Intelligenztest-Gruppen; Y-Achse: $\hat{\rho}$ operationale Validität bei Inklusion aller Studien (*ggf. mit Verhältnissubstitution); Fehlerbalken: 95% Konfidenzintervalle.

Tabelle 4.23

Lernleistung: operationale Validitäten verschiedener Intelligenztests

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
Intelligenz-Struktur-Test													
ohne Subst.	56	7 261	.867	.647	.771	.587	.230	.107	.204	21.7	.53–.65	.25	TRA
Wilde-Intelligenz-Test													
Null-Subst.	34	3 259	.917	.644	.789	.407	.256	.149	.208	33.8	.32–.49	.07	TRA
VW-Subst.	34	3 259	.917	.644	.789	.431	.228	.147	.174	41.7	.35–.51	.14	TRA
r_k -Subst.	34	3 259	.917	.644	.789	.465	.195	.144	.131	54.9	.40–.53	.25	TRA
Exkl. n.ber.	26	2 346	.917	.638	.793	.529	.183	.143	.115	60.7	.46–.60	.34	TRA
Exkl. MAg	23	1 929	.917	.645	.807	.532	.189	.143	.124	57.2	.46–.61	.33	TRA
Berliner Intelligenz-Struktur-Test													
Null-Subst.	9	883	.869	.681	.874	.409	.173	.114	.131	42.9	.30–.52	.19	TRA
VW-Subst.	9	883	.869	.681	.874	.410	.170	.113	.127	44.3	.30–.52	.20	TRA
r_k -Subst.	9	883	.869	.681	.874	.417	.164	.112	.120	46.7	.31–.52	.22	TRA
Exkl. n.ber.	8	838	.869	.686	.885	.418	.166	.107	.128	41.2	.30–.53	.21	TRA
Exkl. MAg	8	838	.869	.686	.885	.418	.166	.107	.128	41.2	.30–.53	.21	TRA
LPS / PSB													
Null-Subst.	15	2 542	.935	.634	.781	.632	.213	.091	.193	18.2	.52–.74	.31	TRA
VW-Subst.	15	2 542	.935	.634	.781	.635	.210	.091	.189	18.7	.53–.74	.32	TRA
r_k -Subst.	15	2 542	.935	.634	.781	.637	.207	.090	.186	19.1	.53–.74	.33	TRA
Exkl. n.ber.	12	2 301	.935	.632	.778	.654	.208	.083	.190	15.9	.54–.77	.34	TRA
Exkl. MAg	11	2 101	.935	.631	.777	.682	.191	.080	.174	17.6	.57–.79	.40	TRA
Matrizentests													
ohne Subst.	22	1 442	.893	.642	.820	.459	.181	.172	.057	90.1	.38–.53	.37	CSC
Exkl. MAg	15	1 187	.893	.642	.842	.431	.185	.161	.091	75.8	.34–.52	.28	CSC
Intelligenztests auf Basis der GATB													
Null-Subst.	74	13 155	.895	.638	.784	.643	.126	.085	.093	45.7	.61–.67	.49	TRA
VW-Subst.	74	13 155	.895	.638	.784	.643	.125	.085	.092	46.2	.61–.67	.49	TRA
r_k -Subst.	74	13 155	.895	.638	.784	.644	.125	.085	.092	46.1	.62–.67	.49	TRA
Exkl. n.ber.	73	13 133	.895	.639	.784	.643	.125	.084	.092	45.6	.61–.67	.49	TRA
Exkl. MAg	70	12 846	.895	.639	.784	.651	.111	.083	.074	55.4	.62–.68	.53	TRA
andere Verfahren													
Null-Subst.	24	3 263	.896	.596	.791	.526	.258	.107	.234	17.2	.42–.63	.14	TRA
VW-Subst.	24	3 263	.896	.596	.791	.528	.255	.107	.232	17.5	.43–.63	.15	TRA
r_k -Subst.	24	3 263	.896	.596	.791	.532	.250	.106	.226	18.1	.43–.63	.16	TRA
Exkl. n.ber.	23	3 144	.896	.594	.791	.540	.250	.105	.227	17.6	.44–.64	.17	TRA
Exkl. MAg	12	1 171	.893	.581	.780	.453	.317	.128	.290	16.3	.27–.63	-.02	SIS

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; CSC = cross-situational consistency, TRA = transportability, SIS = situational specificity.

Tabelle 4.24

Lernleistung: vollständig korrigierte Validitäten verschiedener Intelligenztests

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
Intelligenz-Struktur-Test													
ohne Subst.	56	7 261	.867	.647	.771	.630	.246	.113	.219	21.2	.57–.69	.27	TRA
Wilde-Intelligenz-Test													
Null-Subst.	34	3 259	.917	.644	.789	.425	.268	.158	.217	34.5	.34–.52	.07	TRA
VW-Subst.	34	3 259	.917	.644	.789	.450	.240	.156	.182	42.3	.37–.53	.15	TRA
r_k -Subst.	34	3 259	.917	.644	.789	.486	.205	.152	.137	55.5	.42–.55	.26	TRA
Exkl. n.ber.	26	2 346	.917	.638	.793	.553	.193	.152	.120	61.4	.48–.63	.36	TRA
Exkl. MAg	23	1 929	.917	.645	.807	.556	.199	.152	.129	58.2	.47–.64	.34	TRA
Berliner Intelligenz-Struktur-Test													
Null-Subst.	9	883	.869	.681	.874	.438	.186	.121	.141	42.7	.32–.56	.21	TRA
VW-Subst.	9	883	.869	.681	.874	.440	.182	.121	.136	44.2	.32–.56	.22	TRA
r_k -Subst.	9	883	.869	.681	.874	.448	.176	.120	.128	46.6	.33–.56	.24	TRA
Exkl. n.ber.	8	838	.869	.686	.885	.448	.178	.114	.137	41.0	.32–.57	.22	TRA
Exkl. MAg	8	838	.869	.686	.885	.448	.178	.114	.137	41.0	.32–.57	.22	TRA
LPS / PSB													
Null-Subst.	15	2 542	.935	.634	.781	.653	.221	.094	.200	18.2	.54–.76	.32	TRA
VW-Subst.	15	2 542	.935	.634	.781	.656	.217	.094	.196	18.7	.55–.77	.33	TRA
r_k -Subst.	15	2 542	.935	.634	.781	.659	.214	.093	.192	19.1	.55–.77	.34	TRA
Exkl. n.ber.	12	2 301	.935	.632	.778	.676	.215	.086	.197	15.9	.55–.80	.35	TRA
Exkl. MAg	11	2 101	.935	.631	.777	.705	.198	.083	.180	17.6	.59–.82	.41	TRA
Matrizentests													
ohne Subst.	22	1 442	.893	.642	.820	.486	.194	.184	.060	90.4	.40–.57	.39	CSC
Exkl. MAg	15	1 187	.893	.642	.842	.456	.199	.174	.097	76.5	.36–.56	.30	CSC
Intelligenztests auf Basis der GATB													
Null-Subst.	74	13 155	.895	.638	.784	.679	.133	.090	.098	45.7	.65–.71	.52	TRA
VW-Subst.	74	13 155	.895	.638	.784	.680	.132	.090	.097	46.2	.65–.71	.52	TRA
r_k -Subst.	74	13 155	.895	.638	.784	.680	.132	.090	.097	46.1	.65–.71	.52	TRA
Exkl. n.ber.	73	13 133	.895	.639	.784	.680	.132	.089	.097	45.6	.65–.71	.52	TRA
Exkl. MAg	70	12 846	.895	.639	.784	.688	.117	.087	.078	55.4	.66–.72	.56	TRA
andere Verfahren													
Null-Subst.	24	3 263	.896	.596	.791	.556	.272	.112	.248	17.1	.45–.67	.15	TRA
VW-Subst.	24	3 263	.896	.596	.791	.558	.269	.112	.245	17.4	.45–.67	.16	TRA
r_k -Subst.	24	3 263	.896	.596	.791	.562	.264	.112	.239	18.0	.46–.67	.17	TRA
Exkl. n.ber.	23	3 144	.896	.594	.791	.571	.264	.110	.240	17.4	.46–.68	.18	TRA
Exkl. MAg	12	1 171	.893	.581	.780	.480	.337	.140	.307	17.1	.29–.67	-.03	SIS

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; CSC = cross-situational consistency, TRA = transportability, SIS = situational specificity.

Von einer Exklusion der Studien mit substituierten Korrelationen profitiert vor allem der WIT, dessen operationale Validität in den durchgeführten Analysen auf $\hat{\rho} = .532$ ansteigt. BIS, LPS/PSB und GATB-basierte Verfahren profitieren ebenfalls etwas von der Exklusion der Mittelwert-Aggregate, bei Matrizen-tests und den *anderen Verfahren* verringert sich hingegen die operationale Validität geringfügig. Die letztgenannten weisen auch besondere Generalisierungsmöglichkeiten auf: bei Matrizen-tests liegt *cross-situational consistency* (CSC) vor, die Ergebnisse der anderen Verfahren können bei Exklusion der Studien mit Mittelwert-Aggregaten nicht mehr generalisiert werden (*situational specificity*, SIS).

Die vollständig korrigierten Ergebnisse unterscheiden sich von den operationalen nur geringfügig. Die Intelligenztests mit der höchsten vollständig korrigierten Validität sind LPS/PSB ($\hat{\rho} = .705$, vgl. Tabelle 4.24). Bei der Interpretation der Ergebnisse ist auch zu berücksichtigen, dass die Intelligenztest-Gruppen verschiedene Testverfahren bzw. Versionen von Testverfahren zusammenfassen. Sie können deshalb *nicht* als Validitäten der aktuellen Versionen angesehen werden. Eine Zusammenfassung dieser Ergebnisse zu den verschiedenen Intelligenztests enthält Kapitel 5.7.

Kriterien

Beim Vergleich der Studien, die Berufsschulnoten und betriebliche Beurteilungen getrennt berichten, übersteigt die aggregierte Validität bei Studienexklusion die maximal mögliche Validität (Tabelle 4.25). Den Ausführungen werden die Ergebnisse bei Verhältnissubstitution zu Grunde gelegt. Berufsschulnoten werden in 129 Studien berichtet (Variable KMM = 0, $N = 16\,695$). Das mittlere r liegt bei .339, die Varianzaufklärung durch die Artefakte in der bare-bone Analyse beträgt 20.4%. 28 Studien liegen vor, die betriebliche Beurteilungen als Kriterien verwenden (KMM = 1, $N = 2\,760$). Die beobachtete Validität unterschreitet mit $\bar{r} = .228$ die der Berufsschulnoten bedeutsam. Die Varianzaufklärung liegt bei 30.7%.

Die vollständig korrigierte Validität beim Kriterium Berufsschulnoten liegt ebenfalls mit $\hat{\rho} = .581$ deutlich über der der betrieblichen Beurteilungen mit $\hat{\rho} = .416$ (Tabelle 4.26). In beiden Fällen liegt Transportabilität vor, wobei diese bei den betrieblichen Beurteilungen und Verhältnissubstitution gerade noch erreicht wird (90% $CV_{\text{low}} = .01$). Die Studien, die betriebliche Beurteilungen berichten, sind etwas stärker von insignifikanten Korrelationen und Mittelwert-Aggregaten beeinflusst (Tabelle D.1). Ihre Exklusion führt in beiden Kriteriengruppen jedoch nur zu geringfügig höheren Werten als bei Verhältnissubstitution. Bei Exklusion der Mittelwert-Aggregate kann die korrigierte Validität nicht generalisiert werden (*situational specificity*).

Keinen bedeutsamen Unterschied gibt es zwischen Aus- (Variable KINHALT = 2) und Weiterbildungsleistung als Kriterium (KINHALT = 3; Tabellen 4.27 für bare-bone und 4.28 für korrigierte Analysen). Die beobachtete Validität liegt bei Verhältnissubstitution bei $\bar{r} = .369$ ($k = 190$, $N = 28\,724$) für Aus- und bei $\bar{r} = .407$ für Weiterbildungsleistung ($k = 20$, $N = 1\,727$). Die entsprechenden Werte bei Exklusion übersteigen die maximal möglichen Werte. Die Varianzaufklärung beträgt bei beiden Kriterien etwas mehr als 20%. Vollständig korrigiert beträgt die Validität von Ausbildungsleistung $\hat{\rho} = .618$ und

Tabelle 4.25

Lernleistung: bare-bone Metaanalysen von Berufsschulnoten und betrieblichen Beurteilungen

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	\widehat{SD}_{f}	VE (%)	CI (95%)	fs_k
Berufsschulnoten									
Null-Subst.	129	16 695	.335	.180	.079	.162	19.1	.30–.37	303
VW-Subst.	129	16 695	.339	.174	.079	.156	20.4	.31–.37	309
r_k -Subst.	129	16 695	.343	.170	.078	.151	21.2	.31–.37	314
Exkl. n.ber.	117	15 564	.355	.169	.077	.150	20.6	.32–.39	299
Exkl. MAg	95	12 561	.365	.169	.076	.151	20.3	.33–.40	252
betriebliche Beurteilungen									
Null-Subst.	28	2 760	.223	.179	.098	.151	29.6	.16–.29	35
VW-Subst.	28	2 760	.228	.176	.097	.146	30.7	.16–.29	36
r_k -Subst.	28	2 760	.237	.173	.097	.143	31.5	.17–.30	39
Exkl. n.ber.	25	2 489	.241	.179	.096	.151	28.9	.17–.31	36
Exkl. MAg	19	1 789	.244	.200	.099	.174	24.4	.15–.33	28

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; \widehat{SD}_{f} = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = fail-safe k .

von Weiterbildungsleistung $\hat{\rho} = .681$. In beiden Fällen liegt Transportabilität vor.

Die Studien mit Weiterbildungs-Leistungskriterien sind stärker von insignifikanten Korrelationen (18.8%) und Mittelwert-Aggregaten (36.9%) beeinflusst als die Studien mit Ausbildungs-Leistungskriterien (3.5% bzw. 15.0%). Von der Exklusion der betroffenen Studien profitiert deshalb insbesondere die Validität bei Weiterbildungsleistung: vollständig korrigiert beträgt sie $\hat{\rho} = .790$. Die Validitäten der Ausbildungsleistungsstudien liegen bei diesen Analysen deutlich darunter ($\hat{\rho} = .629$ bei Exklusion insignifikanter Korrelationen und $\hat{\rho} = .640$ bei Exklusion der Studien mit Mittelwert-Aggregaten). Abbildung 4.7 gibt eine Übersicht über die vollständig korrigierten Korrelationen der verschiedenen Lernleistungskriterien. Die wichtigsten Ergebnisse zur Validität von Intelligenz für verschiedene Lernleistungskriterien sind in Kapitel 5.2.1 zusammengefasst.

Tabelle 4.26

Lernleistung: korrigierte Validitäten getrennt für Berufsschulnoten und betriebliche Beurteilungen

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
Berufsschulnoten													
Null-Subst.	129	16 695	.894	.647	.785	.543	.243	.113	.216	21.5	.50–.58	.19	TRA
VW-Subst.	129	16 695	.894	.647	.785	.549	.233	.112	.204	23.3	.51–.59	.21	TRA
r_k -Subst.	129	16 695	.894	.647	.785	.555	.226	.112	.196	24.6	.52–.59	.23	TRA
Exkl. n.ber.	117	15 564	.892	.646	.786	.571	.222	.109	.193	24.3	.53–.61	.25	TRA
Exkl. MAg	95	12 561	.890	.649	.790	.582	.220	.109	.191	24.4	.54–.63	.27	TRA
betriebliche Beurteilungen													
Null-Subst.	28	2 760	.890	.520	.808	.386	.279	.145	.238	27.2	.28–.49	-.01	SIS
VW-Subst.	28	2 760	.890	.520	.808	.393	.273	.145	.231	28.3	.29–.49	.01	TRA
r_k -Subst.	28	2 760	.890	.520	.808	.406	.269	.144	.227	28.7	.31–.51	.03	TRA
Exkl. n.ber.	25	2 489	.887	.520	.812	.409	.277	.141	.238	26.0	.30–.52	.02	TRA
Exkl. MAg	19	1 789	.885	.520	.823	.400	.300	.142	.264	22.5	.27–.53	-.03	SIS
vollständig korrigierte Validität													
Berufsschulnoten													
Null-Subst.	129	16 695	.894	.647	.785	.574	.257	.119	.228	21.3	.53–.62	.20	TRA
VW-Subst.	129	16 695	.894	.647	.785	.581	.246	.118	.215	23.1	.54–.62	.23	TRA
r_k -Subst.	129	16 695	.894	.647	.785	.587	.238	.117	.207	24.3	.55–.63	.25	TRA
Exkl. n.ber.	117	15 564	.892	.646	.786	.604	.234	.114	.205	23.5	.56–.65	.27	TRA
Exkl. MAg	95	12 561	.890	.649	.790	.616	.232	.112	.203	23.3	.57–.66	.28	TRA
betriebliche Beurteilungen													
Null-Subst.	28	2 760	.890	.520	.808	.409	.295	.153	.252	27.0	.30–.52	-.01	SIS
VW-Subst.	28	2 760	.890	.520	.808	.416	.289	.153	.245	28.0	.31–.52	.01	TRA
r_k -Subst.	28	2 760	.890	.520	.808	.430	.284	.151	.241	28.2	.33–.54	.03	TRA
Exkl. n.ber.	25	2 489	.887	.520	.812	.435	.293	.148	.253	25.3	.32–.55	.02	TRA
Exkl. MAg	19	1 789	.885	.520	.823	.425	.317	.148	.281	21.7	.28–.57	-.04	SIS

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; TRA = transportability, SIS = situational specificity.

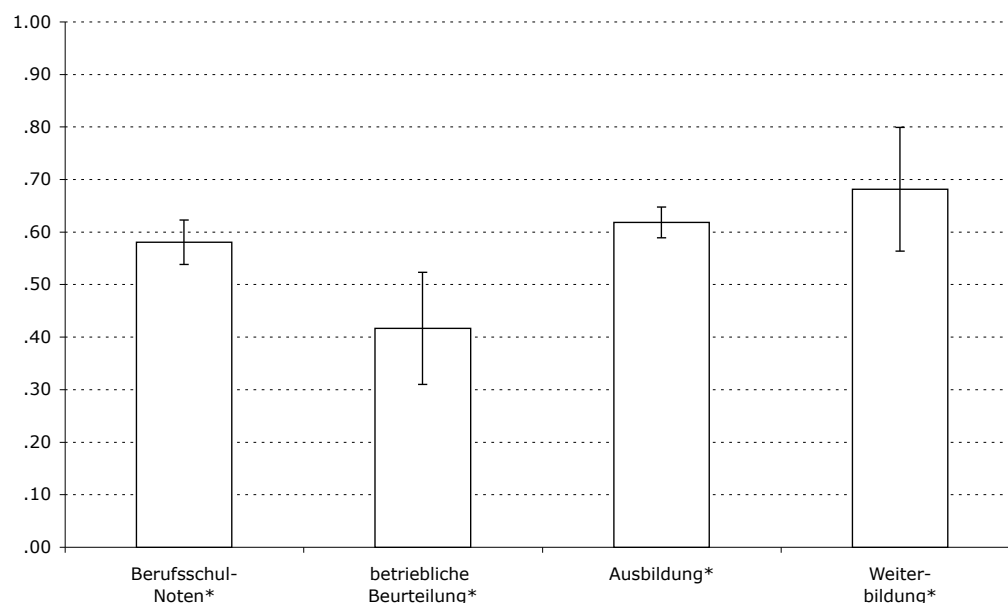
Tabelle 4.27

Lernleistung: bare-bone Metaanalysen für Aus- und Weiterbildungsleistung

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
Ausbildungsleistung									
Null-Subst.	190	28 724	.368	.151	.071	.133	22.2	.35–.39	509
VW-Subst.	190	28 724	.369	.148	.071	.130	23.0	.35–.39	513
r_k -Subst.	190	28 724	.372	.144	.071	.125	24.1	.35–.39	518
Exkl. n.ber.	178	27 709	.378	.142	.069	.124	23.7	.36–.40	495
Exkl. MAg	152	24 428	.386	.139	.068	.122	23.6	.36–.41	435
Weiterbildungsleistung									
Null-Subst.	20	1 727	.392	.213	.093	.192	19.0	.30–.49	59
VW-Subst.	20	1 727	.407	.194	.092	.171	22.4	.32–.49	62
r_k -Subst.	20	1 727	.414	.185	.091	.161	24.2	.33–.49	63
Exkl. n.ber.	18	1 402	.470	.150	.090	.120	36.3	.40–.54	67
Exkl. MAg	17	1 090	.498	.156	.096	.123	38.0	.42–.57	68

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k .

Abbildung 4.7

Vollständig korrigierte Validität bei unterschiedlichen Lernleistungskriterien

Anmerkungen. X-Achse: Kriterien; Y-Achse: $\hat{\rho}$ vollständig korrigierte Validität bei Inklusion aller Studien *mit Verhältnissubstitution; Fehlerbalken: 95% Konfidenzintervalle.

Tabelle 4.28

Lernleistung: korrigierte Validitäten getrennt für Aus- und Weiterbildungsleistung

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
Ausbildungsleistung													
Null-Subst.	190	28 724	.895	.642	.786	.583	.199	.101	.171	26.0	.55–.61	.30	TRA
VW-Subst.	190	28 724	.895	.642	.786	.585	.194	.101	.166	27.2	.56–.61	.31	TRA
r_k -Subst.	190	28 724	.895	.642	.786	.588	.189	.101	.160	28.6	.56–.61	.33	TRA
Exkl. n.ber.	178	27 709	.894	.642	.787	.595	.186	.099	.158	28.3	.57–.62	.34	TRA
Exkl. MAg	152	24 428	.894	.645	.787	.605	.181	.097	.153	28.8	.58–.63	.35	TRA
Weiterbildungsleistung													
Null-Subst.	20	1 727	.883	.614	.781	.617	.286	.100	.268	12.2	.49–.74	.18	TRA
VW-Subst.	20	1 727	.883	.614	.781	.640	.248	.100	.227	16.2	.53–.75	.27	TRA
r_k -Subst.	20	1 727	.883	.614	.781	.651	.230	.099	.207	18.7	.55–.75	.31	TRA
Exkl. n.ber.	18	1 402	.880	.601	.770	.739	.156	.101	.119	42.1	.67–.81	.54	TRA
Exkl. MAg	17	1 090	.878	.609	.786	.737	.195	.101	.167	27.0	.64–.83	.46	TRA
vollständig korrigierte Validität													
Ausbildungsleistung													
Null-Subst.	190	28 724	.895	.642	.786	.616	.209	.105	.181	25.4	.59–.65	.32	TRA
VW-Subst.	190	28 724	.895	.642	.786	.618	.204	.105	.175	26.5	.59–.65	.33	TRA
r_k -Subst.	190	28 724	.895	.642	.786	.622	.198	.105	.169	27.8	.59–.65	.34	TRA
Exkl. n.ber.	178	27 709	.894	.642	.787	.629	.195	.102	.167	27.2	.60–.66	.35	TRA
Exkl. MAg	152	24 428	.894	.645	.787	.640	.190	.099	.162	27.4	.61–.67	.37	TRA
Weiterbildungsleistung													
Null-Subst.	20	1 727	.883	.614	.781	.657	.310	.120	.285	15.1	.52–.79	.19	TRA
VW-Subst.	20	1 727	.883	.614	.781	.681	.269	.119	.241	19.4	.56–.80	.28	TRA
r_k -Subst.	20	1 727	.883	.614	.781	.693	.250	.118	.220	22.1	.58–.80	.33	TRA
Exkl. n.ber.	18	1 402	.880	.601	.770	.788	.170	.114	.126	44.7	.71–.87	.58	TRA
Exkl. MAg	17	1 090	.878	.609	.786	.787	.213	.116	.178	29.8	.69–.89	.49	TRA

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; TRA = transportability.

4.2. Intelligenz und Arbeitsleistung

Die inkludierten Studien werden zunächst näher beschrieben (Kapitel 4.2.1), anschließend die Ergebnisse der metaanalytischen Integration (Kapitel 4.2.2) und der Sensitivitätsanalysen vorgestellt (Kapitel 4.2.3). Der eilige Leser findet Zusammenfassungen der wichtigsten Ergebnisse zum Erfolgsbereich Arbeitsleistung in der Diskussion (Kapitel 5.3).

4.2.1. Beschreibung der inkludierten Studien

In 18 Studien werden subjektive Fremdbeurteilungen der Arbeitsleistung berichtet (Variablen KINHALT = 1 & KMM = 1 \vee 2). Die Stichprobengröße beträgt $N = 2739$, im Durchschnitt $\bar{n} = 152.2$ (Range: 11 bis 913, $SD = 228.1$). Zwei Studien mit einem n von 913 (Hossiep, 1995) bzw. 597 (Nr. 116.12 aus Amthauer, 1973) dominieren die Ergebnisse: ihr Gewicht an der bare-bone Analyse beträgt zusammen 63.1% (vgl. Abbildungen 4.9 und 4.11). Die größte der restlichen Studien (Schuler, Funke et al., 1995) hat mit einem n von 144 lediglich einen Anteil von 5.3%. Die Studie 116.12 ist gleichzeitig die Studie, die das größte r berichtet (.71). Alle Studien wurden zwischen 1963 und 2006 veröffentlicht (15 Studien) bzw. durchgeführt (drei Studien). Nähere Angaben zum Gewicht dieser und weiterer kategorialer Variablen enthält Tabelle 4.29.

Von den 18 Stichproben sind zwei ausschließlich männlich, keine ausschließlich weiblich und neun gemischtgeschlechtlich. In sieben Studien werden keine Aussagen zum Geschlecht getroffen. Die Gesamtstichprobe besteht zu 40.2% aus Personen unbekanntes Geschlechts, die verbleibenden 59.8% sind zu einem Drittel weiblich und zu zwei Dritteln männlich. Sieben Studien berichten das durchschnittliche Alter der Beschäftigten, es liegt im Range von 17.1 bis 36.0 Jahren ($M = 25.3$, $SD = 8.4$). Der n -gewichtete Altersdurchschnitt beträgt 29.7 Jahre. Das Bildungsniveau ist in neun Studien angegeben. Darunter sind zwei Studien, die ausschließlich aus (Fach-)Hochschulabsolventen bestehen (Gewicht 7.8%). Das Bildungsniveau von 41.8% der Gesamtstichprobe ist unbekannt, 33.1% sind (Fach-)Hochschulabsolventen, 24.8% haben das mittlere Bildungsniveau und lediglich 0.4% verfügen höchstens über einen Hauptschulabschluss.

Circa ein Drittel der Gesamtstichprobe sind Techniker und gleichrangige nichttechnisch Ausgebildete (HG 3, 30.7%), etwas häufiger sind Bürokräfte (HG 4, 37.6%) vertreten; Wissenschaftler (HG 2) machen 9.0% der Stichprobe aus, leitend Tätige (HG 1) 4.9%, Handwerker (HG 7) 2.7% und Maschinenbediener (HG 8) 1.1%. Beschäftigte aus den Bereichen Militär, Dienstleistungen, Landwirtschaft/Fischerei sowie Hilfsarbeitende sind nicht mit eigenen Studien vertreten. 13.9% der Gesamtstichprobe können den ISCO-88 Berufs-Hauptgruppen nicht eindeutig zugeordnet werden.

Bis auf den Skill Level 1 sind alle Skill Level mit mindestens zwei Studien berücksichtigt. Die größte Gruppe kann dem Skill Level 2 zugeordnet werden (41.4%), es folgen in der Häufigkeit Skill Level 3 (30.7%) und 4 (9.0%). Den leitend Tätigen (4.9%) wird kein Skill Level zugewiesen. Es verbleiben 4 Studien, die nicht eindeutig einem Skill Level zugeordnet werden können. Ihr Gewicht an der bare-bone Analyse beträgt 13.9%.

Mit 43.9% hat der IST in den Versionen von 1953-1970 den größten Anteil an der Ge-

Tabelle 4.29

Gewicht kategorialer Moderatorvariablen an den Gesamtmetaanalysen zu Arbeitsleistung

Kategorie	<i>k</i>	Gewicht ^a (%)		
		bare-bone	operational	vollst. korr.
		Publikationsjahr		
1960-1969	4	6.1	1.8	1.8
1970-1979	7	39.8	51.9	51.1
1980-1989	0	0.0	0.0	0.0
1990-1999	3	42.2	36.1	36.8
2000-2006	4	11.9	10.2	10.4
		Publikationsart		
Zeitschriftenartikel	5	11.2	11.4	11.5
Dissertationen	4	41.9	33.0	33.5
Testmanuale	4	29.8	42.9	42.3
Monographien (ohne Testmanuale)	2	9.0	6.6	6.6
nicht veröffentlicht	3	8.1	6.2	6.1
		ISCO-88 Berufsgruppe		
1 Leitende Berufe	2	4.9	2.2	2.1
2 Wissenschaftler	2	9.0	6.6	6.6
3 Techniker und Gleichrangige	4	30.7	42.2	41.5
4 Bürokräfte	3	37.6	33.1	33.6
7 Handwerk	1	2.7	2.9	2.9
8 Maschinenbediener	2	1.1	1.0	1.0
gemischt / unklar	4	13.9	12.0	12.2
		ISCO-88 Skill Level		
ohne	2	4.9	2.2	2.1
1 (geringe Komplexität)	0	0.0	0.0	0.0
2	6	41.4	37.0	37.6
3	4	30.7	42.2	41.5
4 (hohe Komplexität)	2	9.0	6.6	6.6
gemischt / unklar	4	13.9	12.0	12.2
		Prädiktoren		
IST 1953-1970	9	43.8	52.1	51.4
WPT	3	10.4	9.0	9.1
Matrizentest FRT	2	1.1	1.0	1.0
BIS Form 4	1	3.6	3.1	3.1
WIT 1963	1	2.6	1.8	1.7
Eigenentwicklung	2	38.6	33.1	33.7
		Art der Validität		
prädiktiv	14	84.8	87.2	87.0
konkurrent	3	11.6	9.8	9.9
retrograd	1	3.6	3.1	3.1
		nicht-berichtete Korrelationen enthalten		
nein	17	97.4	98.2	98.3
ja	1	2.6	1.8	1.7
Gesamt	18	100.0	100.0	100.0

Tabelle 4.29
(Fortsetzung)

Kategorie	<i>k</i>	Gewicht ^a (%)		
		bare-bone	operational	vollst. korr.
Aggregationen enthalten einfache Mittelung				
nein	14	56.8	61.4	61.1
ja	4	43.2	38.6	38.9
Typ des Korrelationskoeffizienten				
Pearson-Korrelation ^b	17	96.2	95.4	95.5
andere Koeffizienten	1	3.8	4.6	4.5
Gesamt	18	100.0	100.0	100.0

Anmerkungen. ^aGewicht bei Inklusion der Studien, die die insignifikanten Werte nicht berichten; ^bPearson-Korrelationen für Intervalldaten (Produkt-Moment-Korrelationen) und ordinale Daten (Spearman-Rangkorrelationen).

samtstichprobe. Zwei Studien nutzen Eigenentwicklungen (38.6%), drei deutsche Übersetzungen des WPT (10.4%). Es folgen bezüglich des Gewichts an der bare-bone Analyse der BIS (Form 4, eine Studie, 3.6%), der WIT von 1963 (eine Studie, 2.6%) und der FRT (zwei Studien, 1.1%). Verarbeitungskapazität ist die BIS-Modell-Facette mit der mit Abstand größten Bedeutung (85.5% an der Gesamtstichprobe). Es folgen Merkfähigkeit (7.4%), Bearbeitungsgeschwindigkeit (6.6%) und Einfallsreichtum (0.5%). Das Gewicht der verbalen Aufgaben ist mit 46.8% deutlich größer als das Gewicht figuraler (27.5%) oder numerischer Aufgaben (25.7%; vgl. Tabelle 4.30).

Als Erfolgsindikatoren für Arbeitsleistung werden subjektive Fremdbeurteilungen inkludiert. Die Antwortformate sind Ratingskalen. Die Kriterien unterscheiden sich hinsichtlich der Anzahl und Inhalte der einzuschätzenden Leistungsdimensionen (Tabelle 3.3). In der Regel handelt es sich bei den Kriterien um Aggregate aus verschiedenen Urteilsdimensionen (Gewicht 61.5%), lediglich in zwei Studien von Althoff (1968) werden Zusammenhänge mit einem Globalurteil berichtet (Gewicht 5.0%). Bei Brandstätter (1970) und Amthauer (1973) ist nur bekannt, dass es sich beim Kriterium um Mittelwerte dienstlicher Beurteilungen handelt (Gewicht 33.5%).

Hinweise auf Varianzeinschränkungen im Kriterium gibt es in der Studie von Kersting (1999a): Die besten 19.2% wurden nicht erfasst, da sie eine Aufstiegszulassung erhalten und ein Aufbaustudium begonnen hatten.

In 15 der 18 Studien sind die Zeitabstände zwischen Intelligenz- und Kriterienerhebung mit hinreichender Sicherheit einschätzbar (Konfidenzrating 3 oder 4, vgl. Abbildung 4.8). Von einer retrograden Validität (-49 Monate, Kersting, 1999a) und einer konkurrenten Validität (Schuler, Funke et al., 1995) abgesehen, werden prädiktive Validitäten berichtet. Der Zeitabstand bei den prädiktiven Studien liegt zwischen 4.5 und 108 Monaten ($M = 65.1$, $SD = 27.2$ Monate). Eine Studie (Hossiep, 1995) enthält Ergebnisse zu mehreren Messzeitpunkten. Sie werden, da sie von einer Stichprobe stammen, zusammengefasst.

In elf Studien ist eine Einschätzung der Berufserfahrung möglich (Gewicht 62.9%):

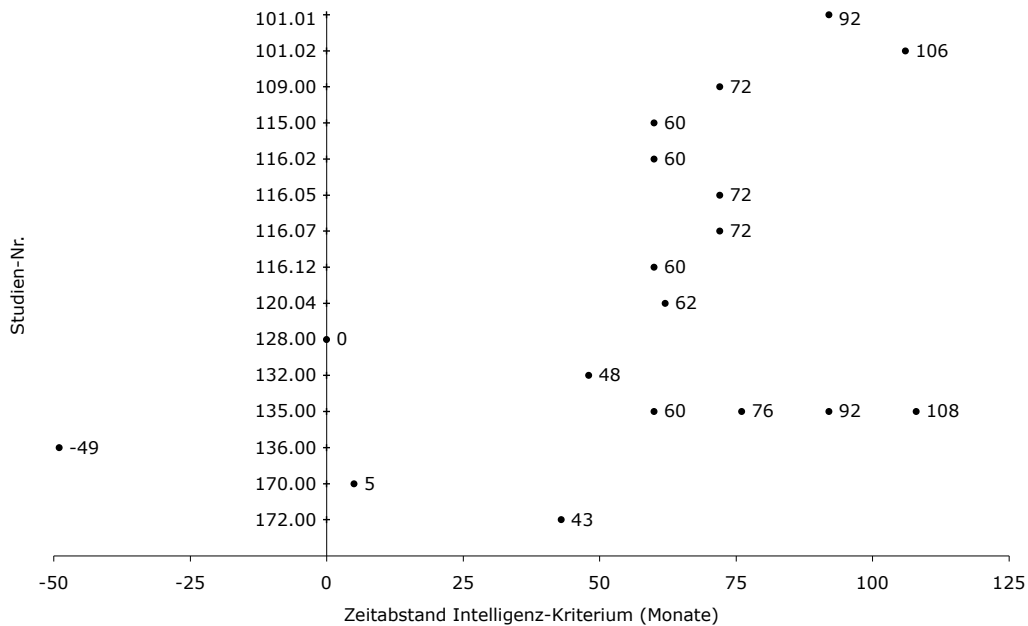
Tabelle 4.30

Geschätztes Gewicht der BIS-Zellen an der Metaanalyse zur Arbeitsleistung bei Inklusion aller Studien

Operationen	Inhalte (%)			Gesamt
	verbal	numerisch	figural	
Verarbeitungskapazität				
bare-bone	37.4	21.4	26.7	85.5
operationale Validität	38.4	21.5	26.1	86.0
vollständig korrigiert	38.3	21.5	26.2	86.0
Bearbeitungsgeschwindigkeit				
bare-bone	4.2	2.2	0.2	6.6
operationale Validität	3.6	1.9	0.1	5.6
vollständig korrigiert	3.6	1.9	0.1	5.7
Einfallsreichtum				
bare-bone	0.2	0.2	0.2	0.5
operationale Validität	0.1	0.1	0.1	0.4
vollständig korrigiert	0.1	0.1	0.1	0.4
Merkfähigkeit				
bare-bone	5.0	2.0	0.5	7.4
operationale Validität	5.9	1.7	0.4	8.0
vollständig korrigiert	5.8	1.7	0.4	7.9
Gesamt				
bare-bone	46.8	25.7	27.5	100.0
operationale Validität	48.1	25.2	26.7	100.0
vollständig korrigiert	48.0	25.3	26.8	100.0

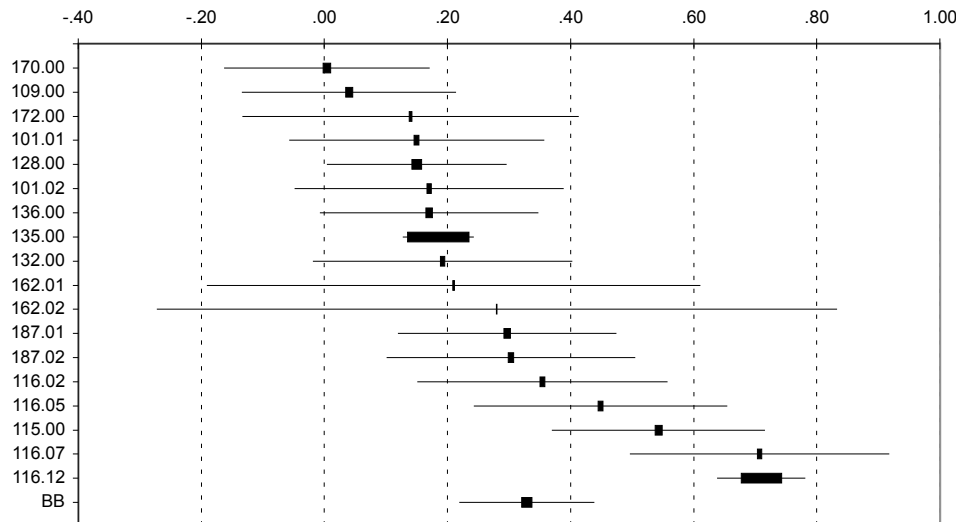
Abbildung 4.8

Zeitabstand zwischen Intelligenz- und Kriterienerhebung der Studien zur Arbeitsleistung



Anmerkungen. Abgebildet sind die Ergebnisse der 15 Studien (Studien-Nr. vgl. Anhang C.1), bei denen der Zeitabstand mit *hoher Sicherheit* eingeschätzt werden konnte (Konfidenzratings 3 und 4).

Abbildung 4.9
Baumdiagramm der Arbeitsleistungs-Studien



Anmerkungen. X-Achse: beobachtete Korrelationen der einzelnen Studien sortiert nach Größe und Ergebnis der bare-bone Analyse (BB, letzte Zeile); Y-Achse: Studiennummern (vgl. Anhang C.1); Balkenbreite: Gewicht der Studie in der bare-bone Analyse, bei BB durchschnittliches Studiengewicht; Fehlerbalken: 95% Konfidenzintervalle.

sie liegt im Range von zwölf und 104 Monaten und beträgt im Mittel 45.2 Monate ($SD = 26.7$). Insignifikante Korrelationen werden in einer Studie (Lometsch, 1975) nicht berichtet. Ihr Anteil an der bare-bone Metaanalyse beträgt 2.6%. Eine Unterschätzung der tatsächlichen Validitäten durch die Aggregatbildung über Mittelwerte ist in vier der 18 Studien möglich, die Mittelwertbildung betrifft ausschließlich die Kriterien, nicht die Prädiktoren (vgl. Tabelle E.35). Das Gewicht der Studien mit Mittelwert-Aggregation an der bare-bone Analyse beträgt 43.2%.

Dem Baumdiagramm (Abbildung 4.9) kann die Höhe der beobachteten Korrelationen, ihre 95% CIs sowie die Gewichte der Studien an der bare-bone Analyse entnommen werden. Eine tabellarische Übersicht der inkludierten Effekte pro Studie inklusive Angaben zu ihrer Berechnung und ihrer Gewichte an der Metaanalyse enthält Tabelle E.35.

4.2.2. Metaanalytische Ergebnisse bei Inklusion aller Arbeitsleistungsstudien

Auf Grund des geringen Gewichtes der insignifikanten nicht-berichteten Korrelationen (2.6%) sind die Ergebnisse für die bare-bone Analyse bei den verschiedenen Substitutionsverfahren nahezu identisch (vgl. Tabelle 4.31). \bar{r} bei Exklusion liegt im maximal möglichen Bereich. Die mittlere Validität beträgt $\bar{r} = .329$ bei Verhältnissubstitution. Sie liegt damit in der Tendenz unter der beobachteten Validität berufsbezogener Lernleistung, die $\bar{r} = .371$ beträgt. Der Anteil der durch den Stichprobenfehler aufgeklärten Varianz ist mit 9.4% sehr gering.

Die operationalen und vollständig korrigierten Validitäten betragen $\hat{\rho} = .615$ bzw. $.656$

Tabelle 4.31

Arbeitsleistung: bare-bone Metaanalyse

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
Null-Subst.	18	2 739	.327	.240	.073	.228	9.3	.22-.44	41
VW-Subst.	18	2 739	.329	.239	.073	.227	9.4	.22-.44	42
r_k -Subst.	18	2 739	.333	.238	.073	.226	9.4	.22-.44	42
Exkl. n.ber.	17	2 669	.333	.241	.072	.230	8.9	.22-.45	40
Exkl. MAg	14	1 555	.415	.274	.079	.262	8.4	.27-.56	45

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k sowie bei Exklusion dieser Studien. k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k .

(Tabelle 4.32). Im Gegensatz zur Exklusion von Studien mit insignifikanten Korrelationen führt die Exklusion von Studien mit Mittelwert-Aggregaten zu höheren mittleren Effekten ($\bar{r} = .415$, $\hat{\rho} = .804$), die auch über denen der Lernleistungsstudien liegen ($\bar{r} = .382$, $\hat{\rho} = .637$). Wegen der vergleichsweise geringen Anzahl von Arbeitsleistungsstudien (zwischen 14 und 18), sind die Konfidenzintervalle breit und die genannten Unterschiede nicht statistisch bedeutsam. Der Anteil der durch die Artefakte aufgeklärten Varianz beträgt im vollständig korrigierten und verhältnissubstituierten Fall 12.0%. Die Glaubwürdigkeitsintervalle schließen Null aus. Das Ergebnis kann somit transportabel generalisiert werden.

Die mittleren Effekte der kumulativen Metaanalyse (Abbildung 4.10) zeigen ebenfalls, dass weitere Studien das Vertrauen in die Ergebnisse steigern könnten. Zwar sind bei der Inklusion der jüngsten fünf Studien keine erheblichen Sprünge von $\hat{\rho}$ mehr zu erkennen, der mittlere Effekt nimmt jedoch kontinuierlich ab und das Konfidenzintervall hat bei Inklusion aller Studien einen Range von über .30.

Auf Grund der geringen Studienzahl ist es nicht möglich den Einfluss von Moderatoren über getrennte Metaanalysen zu bestimmen. Die Ergebnisse können aber mit einigen weiteren Sensitivitätsanalysen abgesichert werden.

4.2.3. Sensitivitätsanalysen

Sensitivitätsanalysen werden durchgeführt unter Ausschluss der unveröffentlichten Studien (Variable PART = 1), der Studien, die keine prädiktiven Validitäten berichten (VTYP = 4 \vee TDIFF \leq 0), der Studien, die keine Pearson-Korrelationen berichten (VKTYTYP > 2) und der Studie 116.12, die sowohl hinsichtlich r als auch hinsichtlich n im Funnelplot als Extremwert auffällt. Die metaanalytischen Ergebnisse können den Tabellen 4.33 und 4.34 entnommen werden. Die bare-bone Gewichte der Berufsgruppen, Skill Level, Veröffentlichungsarten, Prädiktoren und der Variablen, die den Sensitivitätsanalysen zu Grunde liegen, sind für die einzelnen Analysen in Tabelle D.2 wiedergegeben.

Im Funnelplot (Abbildung 4.11a) wird deutlich, dass die drei unveröffentlichten Stu-

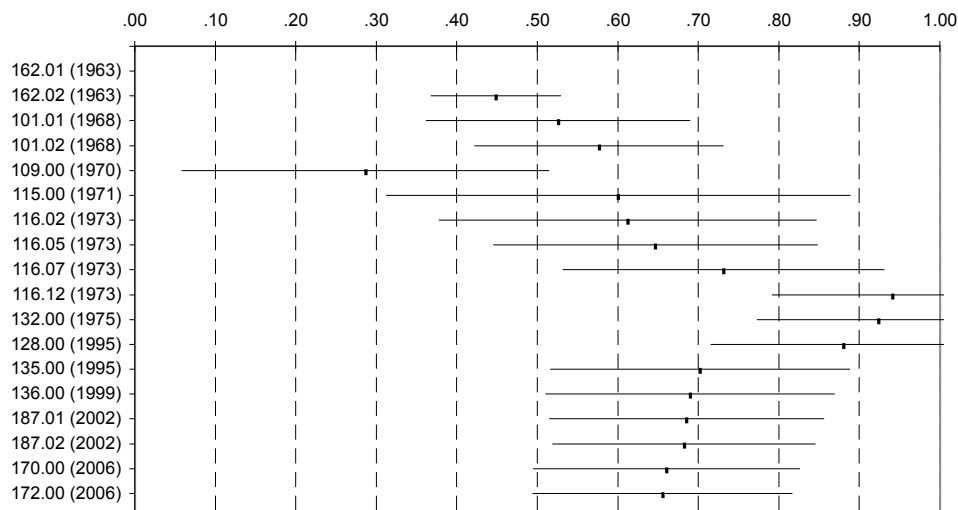
Tabelle 4.32

Arbeitsleistung: korrigierte Validitäten

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
Null-Subst.	18	2739	.880	.520	.743	.613	.326	.102	.309	9.7	.46–.76	.10	TRA
VW-Subst.	18	2739	.880	.520	.743	.615	.323	.101	.307	9.7	.47–.76	.11	TRA
r_k -Subst.	18	2739	.880	.520	.743	.619	.321	.099	.306	9.4	.47–.77	.12	TRA
Exkl. n.ber.	17	2669	.883	.520	.741	.620	.325	.095	.311	8.5	.47–.77	.11	TRA
Exkl. MAg	14	1555	.883	.520	.732	.755	.325	.100	.309	9.5	.59–.93	.25	TRA
vollständig korrigierte Validität													
Null-Subst.	18	2739	.880	.520	.743	.653	.352	.122	.330	12.0	.49–.82	.11	TRA
VW-Subst.	18	2739	.880	.520	.743	.656	.349	.121	.328	12.0	.49–.82	.12	TRA
r_k -Subst.	18	2739	.880	.520	.743	.660	.347	.120	.326	12.0	.50–.82	.12	TRA
Exkl. n.ber.	17	2669	.883	.520	.741	.660	.351	.118	.331	11.3	.49–.83	.12	TRA
Exkl. MAg	14	1555	.883	.520	.732	.804	.352	.126	.329	12.8	.62–.99	.26	TRA

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k sowie bei Exklusion dieser Studien. k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; TRA = transportability.

Abbildung 4.10

Kumulatives Baumdiagramm der Arbeitsleistungs-Metaanalyse

Anmerkungen. X-Achse: $\hat{\rho}$ (vollständig korrigiert); Y-Achse: Studiennummern (vgl. Anhang C.1) sortiert nach Jahreszahl, die – von oben nach unten – *zusätzlich* in der Metaanalyse berücksichtigt werden; Fehlerbalken: 95% Konfidenzintervalle von $\hat{\rho}$.

Tabelle 4.33

Arbeitsleistung: bare-bone Sensitivitätsanalysen

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
nur publizierte Studien									
ohne Subst.	15	2 517	.350	.236	.068	.226	8.4	.23–.47	38
Exkl. MAg	12	1 403	.456	.257	.074	.246	8.3	.31–.60	43
nur prädiktive Studien									
Null-Subst.	14	2 323	.347	.254	.069	.244	7.4	.21–.48	35
VW-Subst.	14	2 323	.349	.252	.069	.243	7.5	.22–.48	35
r_k -Subst.	14	2 323	.353	.251	.069	.241	7.5	.22–.48	36
Exkl. n.ber.	13	2 253	.354	.254	.067	.245	7.0	.22–.49	33
Exkl. MAg	11	1 237	.463	.287	.075	.277	6.9	.29–.63	40
nur Studien mit Pearson-Korrelationen									
Null-Subst.	17	2 636	.319	.241	.073	.229	9.2	.20–.43	38
VW-Subst.	17	2 636	.321	.239	.073	.228	9.2	.21–.43	38
r_k -Subst.	17	2 636	.324	.238	.073	.227	9.3	.21–.44	39
Exkl. n.ber.	16	2 566	.324	.242	.071	.231	8.7	.21–.44	36
Exkl. MAg	14	1 555	.415	.274	.079	.262	8.4	.27–.56	45
ohne Extremwert aus Studie 116.12									
Null-Subst.	17	2 142	.220	.146	.086	.118	34.4	.15–.29	21
VW-Subst.	17	2 142	.223	.145	.085	.117	34.8	.15–.29	21
r_k -Subst.	17	2 142	.227	.146	.085	.119	34.1	.16–.30	22
Exkl. n.ber.	16	2 072	.224	.147	.084	.121	32.8	.15–.30	20
Exkl. MAg	13	958	.232	.185	.111	.147	36.4	.13–.33	18

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAg; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k .

dien (Blickle et al., 2005; Lometsch, 1975; Schmitz, 2006) unterdurchschnittliche Validitäten berichten. Dies bedeutet, dass ein Veröffentlichungsbias vorliegen kann. Das \bar{r} der unveröffentlichten Studien beträgt bei Verhältnissubstitution .089, das der veröffentlichten Studien .350. Das Gewicht der unveröffentlichten Studien beträgt 8.1%. Ihre Exklusion führt nur zu geringfügig höheren mittleren Effekten: $\bar{r} = .35$, $\hat{\rho} = .69$.

Die Prüfung der Symmetrie des FunnelpLOTS mit dem *trim and fill*-Verfahren von Duval (2005, vgl. Kapitel 4.1.3) weist je nach verwendetem Kennwert entweder auf das Vorliegen einer positiven Verzerrung hin ($R_0 = 1$) oder impliziert, dass keine Verzerrung vorliegt ($L_0 = -3$). Wird von einer Verzerrung ausgegangen, sind drei Iterationen der R_0 -Berechnung erforderlich. Sie machen schließlich – um Symmetrie zu erreichen – die Exklusion von drei Studien nötig (116.12, 116.07 und 115.00). Die mittlere Korrelation der symmetrischen Verteilung (Abbildung 4.11b) ist mit $\bar{r} = .189$ deutlich vermindert, die Standardabweichung des mittleren Effektes deutlich erhöht ($SD_r = .109$). Der Unterschied zur ursprünglichen Verteilung muss jedoch nicht zwingend auf eine Publikationsverzerrung zurückzuführen sein (Kapitel 4.1.3).

Im vorliegenden Fall ist der Einfluss der Studie mitentscheidend, die die höchste Validität berichtet (Nr. 116.12): sie stellt sowohl hinsichtlich n als auch hinsichtlich r einen Extremwert dar. Berücksichtigt man die Studie nicht, sind die mittleren Effekte deutlich geringer: das beobachtete \bar{r} beträgt .223 und $\hat{\rho} = .459$. Die Exklusion dieser Studie hat einen deutlichen Anstieg der durch die Artefakte erklärten Varianz zur Folge: von 9.4% in der bare-bone Analyse auf 34.8% und von 12.0% in der vollständig korrigierten Analyse auf 37.9%.

Bei ausschließlicher Berücksichtigung prädiktiver Studien erhält man geringfügig höhere mittlere Effekte. Das \bar{r} beträgt .349, $\hat{\rho} = .688$, die anderen Kennwerte ändern sich nur unerheblich. Das gilt auch für die Sensitivitätsanalyse, bei der eine Studie exkludiert wird, deren Ergebnisse nicht auf Korrelationen, sondern auf transformierten Kontingenzkoeffizienten basieren (Dreyer & Grabitz, 1971). Die mittleren Effekte sind geringfügig kleiner als bei Berücksichtigung aller Studien ($\bar{r} = .321$ statt .329, $\hat{\rho} = .644$ statt .656 bei Verhältnissubstitution).

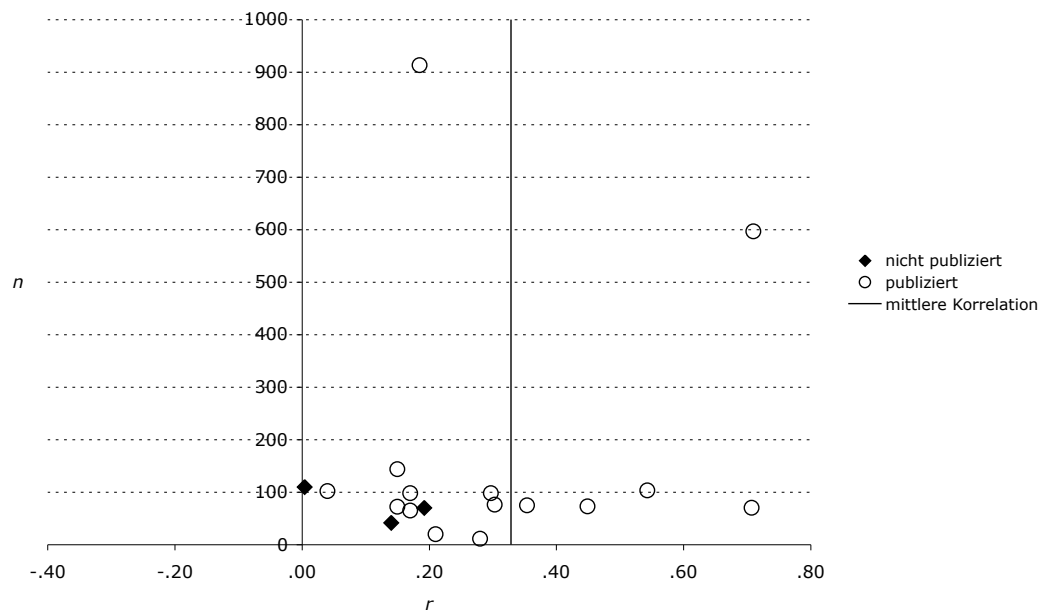
Insgesamt wird durch die Sensitivitätsanalysen deutlich, dass die Ergebnisse zur Arbeitsleistung weitaus weniger gesichert sind als die zu berufsbezogener Lernleistung. Dies liegt nicht zuletzt an der vergleichsweise geringen Anzahl an Studien. Die Ergebnisse zu diesem Erfolgsbereich sind in der Diskussion (Kapitel 5.3) noch einmal zusammengefasst dargestellt.

4.3. Intelligenz und Karriereerfolg

Zunächst werden die inkludierten Studien näher beschrieben (Kapitel 4.3.1), anschließend die metaanalytischen Ergebnisse berichtet (Kapitel 4.3.2). Dabei werden die Indikatoren *Einkommen* und *berufliches Vorankommen* jeweils getrennt betrachtet. Die Ergebnisse zum Erfolgsbereich *Karriere* sind im entsprechenden Kapitel der Diskussion (5.4) zusammengefasst.

Abbildung 4.11
Funnelplot der Studien zu Arbeitsleistung

a) bei Inklusion aller Studien



b) nach Import potentiell fehlender Studien (*trim and fill-Analyse*)

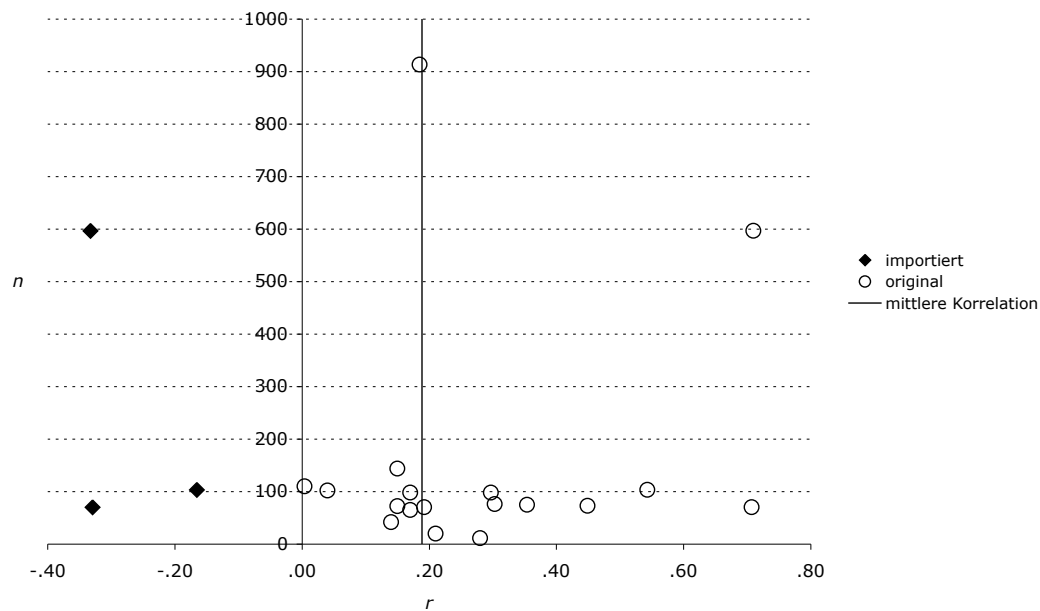


Tabelle 4.34

Arbeitsleistung: korrigierte Validitäten der Sensitivitätsanalysen

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD beo	SD art	\widehat{SD}_ρ	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
nur publizierte Studien													
ohne Subst.	15	2 517	.883	.520	.736	.646	.309	.086	.297	7.7	.49–.80	.16	TRA
Exkl. MAg	12	1 403	.883	.520	.723	.808	.274	.089	.259	10.6	.65–.96	.38	TRA
nur prädiktive Studien													
Null-Subst.	14	2 323	.875	.520	.736	.642	.336	.091	.323	7.4	.47–.82	.11	TRA
VW-Subst.	14	2 323	.875	.520	.736	.644	.333	.090	.321	7.4	.47–.82	.12	TRA
r_k -Subst.	14	2 323	.875	.520	.736	.649	.330	.088	.318	7.1	.48–.82	.13	TRA
Exkl. n.ber.	13	2 253	.879	.520	.733	.651	.334	.082	.324	6.1	.47–.83	.12	TRA
Exkl. MAg	11	1 237	.879	.520	.722	.815	.317	.089	.304	7.9	.63–1.00	.32	TRA
nur Studien mit Pearson-Korrelationen													
Null-Subst.	17	2 636	.881	.520	.741	.602	.329	.103	.313	9.7	.45–.76	.09	TRA
VW-Subst.	17	2 636	.881	.520	.741	.604	.327	.102	.311	9.7	.45–.76	.09	TRA
r_k -Subst.	17	2 636	.881	.520	.741	.609	.325	.100	.309	9.4	.45–.76	.10	TRA
Exkl. n.ber.	16	2 566	.884	.520	.738	.609	.329	.096	.315	8.5	.45–.77	.09	TRA
Exkl. MAg	14	1 555	.883	.520	.732	.755	.325	.100	.309	9.5	.59–.93	.25	TRA
ohne Extremwert aus Studie 116.12													
Null-Subst.	17	2 142	.881	.520	.743	.427	.238	.146	.188	37.5	.31–.54	.12	TRA
VW-Subst.	17	2 142	.881	.520	.743	.431	.236	.145	.187	37.6	.32–.54	.12	TRA
r_k -Subst.	17	2 142	.881	.520	.743	.438	.238	.143	.190	36.1	.33–.55	.13	TRA
Exkl. n.ber.	16	2 072	.884	.520	.741	.433	.240	.142	.193	35.0	.32–.55	.11	TRA
Exkl. MAg	13	958	.884	.520	.730	.478	.304	.195	.233	41.1	.31–.64	.09	TRA
vollständig korrigierte Validität													
nur publizierte Studien													
ohne Subst.	15	2 517	.883	.520	.736	.687	.335	.111	.316	10.9	.52–.86	.17	TRA
Exkl. MAg	12	1 403	.883	.520	.723	.860	.298	.114	.275	14.6	.69–1.03	.41	TRA
nur prädiktive Studien													
Null-Subst.	14	2 323	.875	.520	.736	.686	.363	.112	.345	9.5	.50–.88	.12	TRA
VW-Subst.	14	2 323	.875	.520	.736	.688	.360	.111	.343	9.5	.50–.88	.12	TRA
r_k -Subst.	14	2 323	.875	.520	.736	.694	.357	.110	.340	9.5	.51–.88	.13	TRA
Exkl. n.ber.	13	2 253	.879	.520	.733	.694	.362	.107	.346	8.8	.50–.89	.13	TRA
Exkl. MAg	11	1 237	.879	.520	.722	.870	.344	.114	.324	11.1	.67–1.07	.34	TRA
nur Studien mit Pearson-Korrelationen													
Null-Subst.	17	2 636	.881	.520	.741	.641	.355	.122	.334	11.9	.47–.81	.09	TRA
VW-Subst.	17	2 636	.881	.520	.741	.644	.353	.122	.331	11.9	.48–.81	.10	TRA
r_k -Subst.	17	2 636	.881	.520	.741	.649	.351	.121	.330	11.9	.48–.82	.11	TRA
Exkl. n.ber.	16	2 566	.884	.520	.738	.648	.355	.119	.335	11.2	.47–.82	.10	TRA
Exkl. MAg	14	1 555	.883	.520	.732	.804	.352	.126	.329	12.8	.62–.99	.26	TRA

Tabelle 4.34
(Fortsetzung)

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
ohne Extremwert aus Studie 116.12													
Null-Subst.	17	2142	.881	.520	.743	.456	.254	.156	.201	37.6	.33-.58	.13	TRA
VW-Subst.	17	2142	.881	.520	.743	.459	.252	.155	.199	37.9	.34-.58	.13	TRA
r_k -Subst.	17	2142	.881	.520	.743	.467	.254	.154	.203	36.6	.35-.59	.13	TRA
Exkl. n.ber.	16	2072	.884	.520	.741	.460	.256	.152	.205	35.4	.33-.59	.12	TRA
Exkl. MAG	13	958	.884	.520	.730	.508	.323	.208	.248	41.2	.33-.68	.10	TRA

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k , bei Exklusion dieser Studien (Exkl. n.ber) sowie bei Exklusion der Studien mit Mittelwert-Aggregaten (Exkl. MAG; fehlende Werte sind mit VW substituiert). k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; TRA = transportability.

4.3.1. Beschreibung der inkludierten Studien

Korrelationen zwischen Intelligenz und Einkommensmaßen berichten sieben Studien. Neun Studien berichten Korrelationen zwischen Intelligenz und Beförderung- bzw. Berufsstatus-Maßen (berufliches Vorankommen, vgl. Kapitel 3.2.2). Beide Karriereerfolgsbereiche werden getrennt untersucht, um möglichst homogene Kriterien zu erreichen.

Einkommen

Für den Erfolgsbereich *Einkommen* (KMM = 6 & KINHALT = 1) liegen neutrale Kennzahlen vor, die in drei Studien (Lometsch, 1975; Hossiep, 1995; Steinmann, 1997) Akten bzw. Dokumenten entnommen wurden und in vier Studien (Dahle & Erdmann, 2001; Nettelnstroth, 2004; Amelang & Steinmayr, 2006b; Blickle et al., 2006) Selbstberichten entstammen. In einigen Studien wird das Einkommen direkt erhoben, z.B. das mittlere Jahreseinkommen (Amelang & Steinmayr, 2006b) oder das Gehalt zu einem bestimmten Stichtag (Lometsch, 1975). Nettelnstroth (2004) fragt darüber hinaus nach der Einkommensentwicklung über die letzten zehn Jahre hinweg. Zwei Studien nutzen nicht das Einkommen als Kriterium, sondern die Tarifgruppe (Hossiep, 1995; Steinmann, 1997).

Insgesamt liegen Ergebnisse von sieben Studien vor ($N = 1980$, $\bar{n} = 282.9$, Range 70 bis 767). In einer Studie wird keine Angabe zum Geschlecht der Teilnehmer gemacht, fünf Studien sind gemischtgeschlechtlich, eine rein männlich. Für die Gesamtstichprobe bedeutet dies, dass das Geschlecht von 14.0% der untersuchten Personen unbekannt ist; die restlichen Personen sind zu 60.6% männlich und zu 39.4% weiblich. Fünf Studien berichten das durchschnittliche Alter der Beschäftigten: es liegt bei 30.6, n -gewichtet bei 31.1 Jahren ($SD = 8.4$ Jahre, Range 16.9 bis 39 Jahre). Das Bildungsniveau ist in sechs

Tabelle 4.35

Gewicht kategorialer Moderatorvariablen an den Metaanalysen zum Einkommen und beruflichen Vorankommen

Kategorie	Einkommen			Vorankommen				
	<i>k</i>	Gewicht ^a (%)			<i>k</i>	Gewicht ^a (%)		
		bare-b.	op.	korr.		bare-b.	op.	korr.
Publikationsjahr								
1970-1979	1	3.5	2.9	2.6	1	3.6	2.9	2.7
1980-1989	0	0.0	0.0	0.0	2	8.5	7.9	7.8
1990-1999	2	48.7	50.4	50.1	2	49.6	51.5	51.4
2000-2006	4	47.8	46.7	47.4	4	38.3	37.8	38.1
Publikationsart								
Dissertationen	1	38.7	40.1	39.6	1	35.4	38.3	38.1
Beiträge in Hrsg.-Werken	1	14.9	14.6	14.7	2	22.6	21.3	21.5
nicht veröffentlicht	5	46.4	45.4	45.7	6	42.0	40.3	40.4
ISCO-88 Berufsgruppe								
1 Leitende Berufe	1	3.5	2.9	2.6	1	3.9	3.1	2.9
3 Techniker und Gleichrangige	0	0.0	0.0	0.0	1	2.7	2.4	2.4
4 Bürokräfte	2	48.8	50.1	49.8	2	46.0	48.9	48.9
gemischt / unklar	4	47.6	47.0	47.6	5	47.5	45.5	45.8
ISCO-88 Skill Level								
ohne	1	3.5	2.9	2.6	1	3.9	3.1	2.9
1 (geringe Komplexität)	0	0.0	0.0	0.0	0	0.0	0.0	0.0
2	2	48.8	50.1	49.8	2	46.0	48.9	48.9
3	0	0.0	0.0	0.0	1	2.7	2.4	2.4
4 (hohe Komplexität)	0	0.0	0.0	0.0	0	0.0	0.0	0.0
gemischt/unklar	4	47.6	47.0	47.6	5	47.5	45.5	45.8
Prädiktoren								
IST 2000 R	2	25.4	25.9	26.1	2	27.6	26.6	27.0
LPS	2	25.8	25.1	25.7	1	10.6	10.6	10.8
WPT	1	6.6	6.1	6.0	2	13.5	12.7	12.7
WIT 1963	1	3.5	2.9	2.6	1	3.9	3.1	2.9
Eigenentwicklung	1	38.7	40.1	39.6	3	44.5	46.9	46.6
Art der Validität								
prädiktiv	4	62.8	64.3	63.8	6	73.6	75.4	75.2
konkurrent	1	6.6	6.1	6.0	2	23.7	22.1	22.3
retrograd	1	14.9	14.6	14.7	0	0.0	0.0	0.0
unklar	1	15.7	15.1	15.5	1	2.7	2.4	2.4
nicht-berichtete Korrelationen enthalten								
nein	6	96.5	97.1	97.4	7	79.9	81.7	81.7
ja	1	3.5	2.9	2.6	2	20.1	18.3	18.3
Aggregationen enthalten einfache Mittelung								
ja	6	85.1	85.4	85.3	8	96.1	96.9	97.1
nein	1	14.9	14.6	14.7	1	3.9	3.1	2.9
Gesamt	7	100.0	100.0	100.0	9	100.0	100.0	100.0

Tabelle 4.35
(Fortsetzung)

Kategorie	Einkommen			Vorankommen				
	<i>k</i>	Gewicht ^a (%)			<i>k</i>	Gewicht ^a (%)		
		bare-b.	op.	korr.		bare-b.	op.	korr.
	Typ des Korrelationskoeffizienten							
Pearson-Korrelation ^b	7	100.0	100.0	100.0	8	89.4	89.4	89.2
andere Koeffizienten	0	0.0	0.0	0.0	1	10.6	10.6	10.8
Gesamt	7	100.0	100.0	100.0	9	100.0	100.0	100.0

Anmerkungen. ^aGewicht bei Inklusion der Studien, die die insignifikanten Werte nicht berichten; ^bPearson-Korrelationen für Intervalldaten (Produkt-Moment-Korrelationen) und ordinale Daten (Spearman-Rangkorrelationen).

Studien berichtet; eine Studie besteht ausschließlich aus (Fach-)Hochschulabsolventen (Gewicht 3.5%; Lometsch, 1975). Die Gesamtstichprobe weist zu 14.9% ein unbekanntes Bildungsniveau auf, 38.2% können dem (Fach-)Hochschulniveau, 33.1% dem mittleren Schulniveau und 13.9% maximal dem Hauptschulniveau zugeordnet werden.

Zwei Studien mit einem Gewicht von insgesamt 48.8% untersuchen Bürokräfte mit dem Skill Level 2 (Hossiep, 1995; Steinmann, 1997), eine Studie leitende Berufstätige, die keinem Skill Level zugeordnet werden (Gewicht 3.5%; Lometsch, 1975). Die vier verbleibenden Studien lassen sich weder ISCO-88 Hauptgruppen noch Skill Leveln eindeutig zuordnen (Gewicht 47.6%). Eine Übersicht darüber und über die im Folgenden beschriebenen Variablen bietet Tabelle 4.35.

Je zwei Studien setzen den LPS (Gewicht 28.8%; Dahle & Erdmann, 2001; Steinmann, 1997) bzw. den IST 2000 R ein (25.4%; Amelang & Steinmayr, 2006b; Nettelstroth, 2004). Je eine Studie nutzt eine Eigenentwicklung (38.7%; Hossiep, 1995), eine deutsche Adaptation des WPT (6.6%; Blickle et al., 2006) oder den WIT von 1963 (3.5%; Lometsch, 1975). Die BIS-Facette *Verarbeitungskapazität* ist mit Abstand die bedeutendste (75.8%), gefolgt von Bearbeitungsgeschwindigkeit (21.2%). Merkfähigkeit und Einfallsreichtum spielen mit 1.8% und 1.3% kaum eine Rolle. Die drei Inhaltsfacetten des BIS haben ein Gewicht von 37.3% (verbal), 34.3% (figural) und 28.4% (numerisch; Tabelle 4.36).

Die Art der Validität ist in vier Studien prädiktiv (62.8%), in je einer konkurrent (6.6%; Blickle et al., 2006), retrograd (14.9%; Nettelstroth, 2004) oder nicht zuzuordnen (15.7%; Dahle & Erdmann, 2001). Eine Studie berichtet keine insignifikanten Korrelationen (3.5%; Lometsch, 1975), eine weitere enthält mehrere abhängige Ergebnisse, die nur durch einfache Mittelung aggregiert werden können (14.9%; Nettelstroth, 2004). Alle sieben Studien berichten Pearson-Produkt-Moment- oder Rangkorrelationen. Abbildung 4.12a veranschaulicht die Höhe der beobachteten Korrelationen und die Gewichte der einzelnen Studien an der Metaanalyse.

Tabelle 4.36

Geschätztes Gewicht der BIS-Zellen an der Metaanalyse zum Einkommen

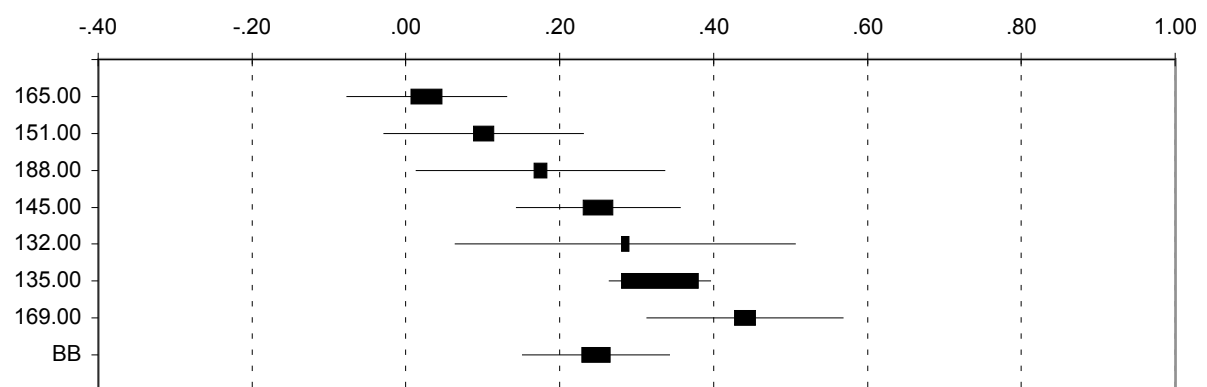
Operationen	Inhalte (%)			Gesamt
	verbal	numerisch	figural	
Verarbeitungskapazität				
bare-bone	23.3	22.1	30.4	75.8
operationale Validität	23.4	21.9	30.7	76.0
vollständig korrigiert	23.3	21.7	30.7	75.7
Bearbeitungsgeschwindigkeit				
bare-bone	12.6	4.6	3.9	21.2
operationale Validität	12.5	4.6	3.8	20.9
vollständig korrigiert	12.7	4.7	3.9	21.2
Einfallsreichtum				
bare-bone	1.3	0.0	0.0	1.3
operationale Validität	1.3	0.0	0.0	1.3
vollständig korrigiert	1.3	0.0	0.0	1.3
Merkfähigkeit				
bare-bone	0.0	1.8	0.0	1.8
operationale Validität	0.0	1.8	0.0	1.8
vollständig korrigiert	0.0	1.8	0.0	1.8
Gesamt				
bare-bone	37.3	28.4	34.3	100.0
operationale Validität	37.2	28.3	34.5	100.0
vollständig korrigiert	37.2	28.2	34.5	100.0

Anmerkung. Alle Studien inkludiert.

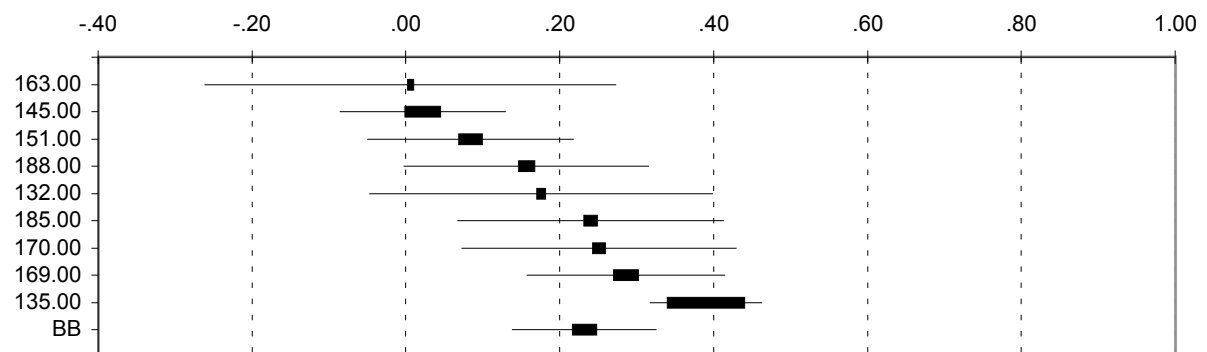
Abbildung 4.12

Baumdiagramme der Studien zum Einkommen und beruflichen Vorankommen

a) Einkommen



b) berufliches Vorankommen



Anmerkungen. X-Achse: beobachtete Korrelationen der einzelnen Studien sortiert nach Größe und Ergebnis der bare-bone Analyse (BB, letzte Zeile); Y-Achse: Studiennummern (vgl. Anhang C.1); Balkenbreite: Gewicht der Studie in der bare-bone Analyse, bei BB durchschnittliches Studiengewicht; Fehlerbalken: 95% Konfidenzintervalle.

Berufliches Vorankommen

Die Indikatoren des Erfolgsbereiches *berufliches Vorankommen* (KINHALT = 1 & KMM = 7 \vee 8) wurden in Kapitel 3.2.2 beschrieben. Insgesamt liegen Ergebnisse von neun Studien mit einem N von 1817 (\bar{n} = 201.9, Range von 70 bis 643) vor. In zwei Studien werden keine Angaben zum Geschlecht gemacht (Anteil an der Gesamtstichprobe 27.6%), die anderen sind gemischtgeschlechtlich. Die Personen, denen ein Geschlecht zugeordnet werden kann, sind zu 53.4% männlich und zu 46.6% weiblich. Das durchschnittliche Alter wird in fünf Studien berichtet, es liegt bei 30.7 Jahren, n -gewichtet beträgt das Mittel 31.4 Jahre (SD = 8.4 Jahre, Range: 16.9 bis 39 Jahre).

Weitere Informationen zur Bedeutung der im Folgenden beschriebenen Variablen enthält Tabelle 4.35. In zwei Studien ist das Bildungsniveau unbekannt, die Studie von Lometsch (1975) kann dem (Fach-)Hochschulniveau zugeordnet werden, die restlichen sechs Studien setzen sich aus Personen mit unterschiedlichen Bildungsniveaus zusammen. Das Bildungsniveau der Gesamtstichprobe ist zu 24.4% unbekannt, 42.4% haben die (Fach-)Hochschulreife, 31.8% das mittlere Bildungsniveau und 1.3% maximal den Hauptschulabschluss.

Ein großer Teil der Gesamtstichprobe (46.0%) entstammt zwei Studien, die Bürokräfte mit dem Skill Level 2 untersuchen (Hossiep, 1995; Steinmann, 1997); je eine Studie widmet sich leitenden Berufen (ohne Skill Level, 3.9%; Lometsch, 1975) und Technikern bzw. gleichrangigen Personen (Skill Level 3, 2.7%; Kreuscher, 1987). Die restlichen 47.5% lassen sich weder einer Hauptgruppe noch einem Skill Level zuordnen.

Die eingesetzten Intelligenzverfahren sind in drei Studien Eigenentwicklungen (44.5%; Hossiep, 1995; Kreuscher, 1987; Linz, 1985), in jeweils zwei Studien der IST 2000 R (27.6%; Amelang & Steinmayr, 2006b; Nettelnstroth, 2004) oder eine deutsche Adaptation des WPT (13.5%; Blickle et al., 2005, 2006). Je eine Studie setzt den LPS (10.6%; Steinmann, 1997) oder den WIT von 1963 ein (3.9%; Lometsch, 1975). Die Gewichte der BIS-Zellen enthält Tabelle 4.37. Auch bei den Studien zu diesem Erfolgsbereich ist insbesondere die Verarbeitungskapazitäts-Facette bedeutsam (79.8%), Bearbeitungsgeschwindigkeit hat ein geschätztes Gewicht von 12.2% und Merkfähigkeit von 1.6%. Einfallsreichtum wird nicht berücksichtigt. Ungefähr gleich stark vertreten sind verbale (35.9%), numerische (29.4%) und figurale (28.3%) Inhalte.

Sechs Studien berichten prädiktive (Gewicht 73.6%), zwei konkurrente (23.7%) Validitäten; in einer Studie ist es unklar, ob zuerst Intelligenz oder das Erfolgskriterium erhoben wurde (Kreuscher, 1987). Zwei Studien mit einem Gewicht von 20.1% berichten keine insignifikanten Korrelationen (Lometsch, 1975; Nettelnstroth, 2004). Bei einer Studie erfolgt die Aggregation der Kriterien durch einfache Mittelung (Lometsch, 1975) und eine Studie berichtet t -Werte statt Korrelationen als Effektmaß (Steinmann, 1997). Eine grafische Übersicht über das Gewicht der einzelnen Studien und die berichteten Validitäten gibt Abbildung 4.12b.

Tabelle 4.37

Geschätztes Gewicht der BIS-Zellen an den Metaanalysen zum beruflichen Vorankommen bei Inklusion aller Studien

Operationen	Inhalte (%)			
	verbal	numerisch	figural	Gesamt
Verarbeitungskapazität				
bare-bone	27.2	24.4	28.3	79.8
operationale Validität	27.0	23.9	28.6	79.6
vollständig korrigiert	27.0	23.9	28.7	79.6
Bearbeitungsgeschwindigkeit				
bare-bone	8.8	3.4	0.0	12.2
operationale Validität	9.0	3.5	0.0	12.5
vollständig korrigiert	9.1	3.5	0.0	12.6
Einfallsreichtum				
bare-bone	0.0	0.0	0.0	0.0
operationale Validität	0.0	0.0	0.0	0.0
vollständig korrigiert	0.0	0.0	0.0	0.0
Merkfähigkeit				
bare-bone	0.0	1.6	0.0	1.6
operationale Validität	0.0	1.7	0.0	1.7
vollständig korrigiert	0.0	1.7	0.0	1.7
Gesamt				
bare-bone	35.9	29.4	28.3	93.6 ^a
operationale Validität	36.1	29.2	28.6	93.9 ^a
vollständig korrigiert	36.1	29.1	28.7	93.9 ^a

Anmerkungen. ^aAn 100% fehlende Angaben lassen sich keinen Facetten zuordnen.

4.3.2. Metaanalytische Ergebnisse für die Indikatorgruppen Einkommen und berufliches Vorankommen

Einkommen

Die mittlere beobachtete Validität von Intelligenz in Bezug auf das Einkommen liegt bei Verhältnissubstitution bei .247, das vollständig korrigierte $\hat{\rho}$ beträgt .345 (Tabellen 4.38 und 4.39). Für den Erfolgsbereich Karriere erfolgt keine Korrektur der Unreliabilität im Kriterium (vgl. Kapitel 3.6.2). Die vollständig korrigierten $\hat{\rho}$ s der kumulativen Metaanalyse veranschaulicht Abbildung 4.13a. Die Validität von Intelligenz in Bezug auf das Einkommen ist somit geringer als in Bezug auf Lern- bzw. Arbeitsleistungskriterien; die Konfidenzintervalle überlappen sich mit Lernleistung nicht und mit Arbeitsleistung nur in der bare-bone Analyse (Abbildung 4.14).

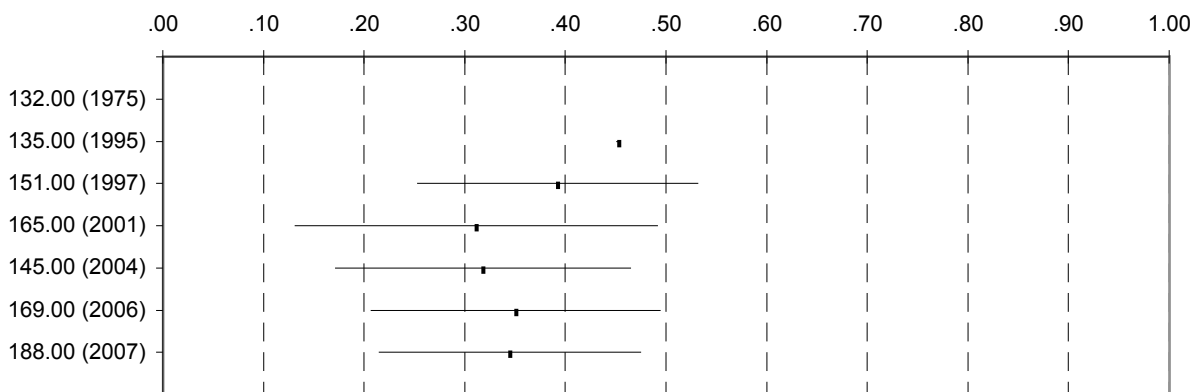
Das Gewicht der insignifikanten nicht-berichteten Korrelationen am Gesamtergebnis beträgt 2.9%. Auf Grund dieses geringen Gewichtes unterscheiden sich die metaanalytischen Ergebnisse zwischen den verschiedenen Substitutionsverfahren nur unbedeutend.

Eine Besonderheit ist darin zu sehen, dass sowohl \bar{r} als auch $\hat{\rho}$ bei Exklusion der

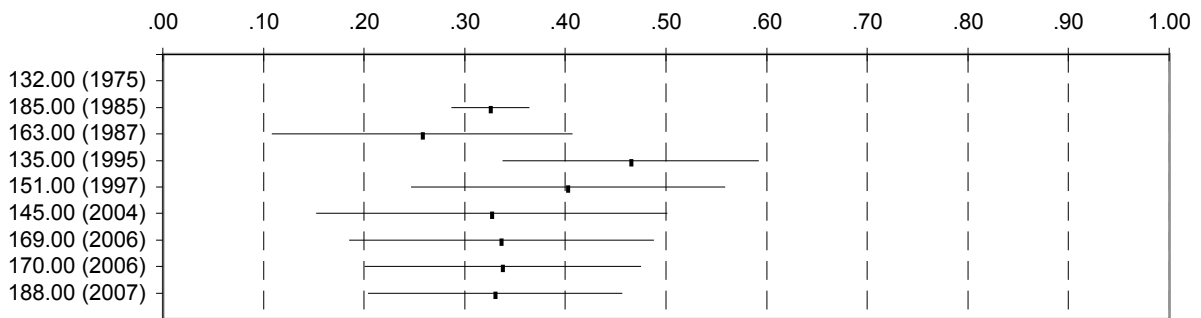
Abbildung 4.13

Kumulative Baumdiagramme der Metaanalysen zum Karriereerfolg

a) Einkommen



b) berufliches Vorankommen



Anmerkungen. X-Achse: $\hat{\rho}$ (verhältnissubstituiert, vollständig korrigiert); Y-Achse: Studiennummern (vgl. Anhang C.1) sortiert nach Jahreszahl, die – von oben nach unten – *zusätzlich* in der Metaanalyse berücksichtigt werden; Fehlerbalken: 95% Konfidenzintervalle von $\hat{\rho}$.

Tabelle 4.38

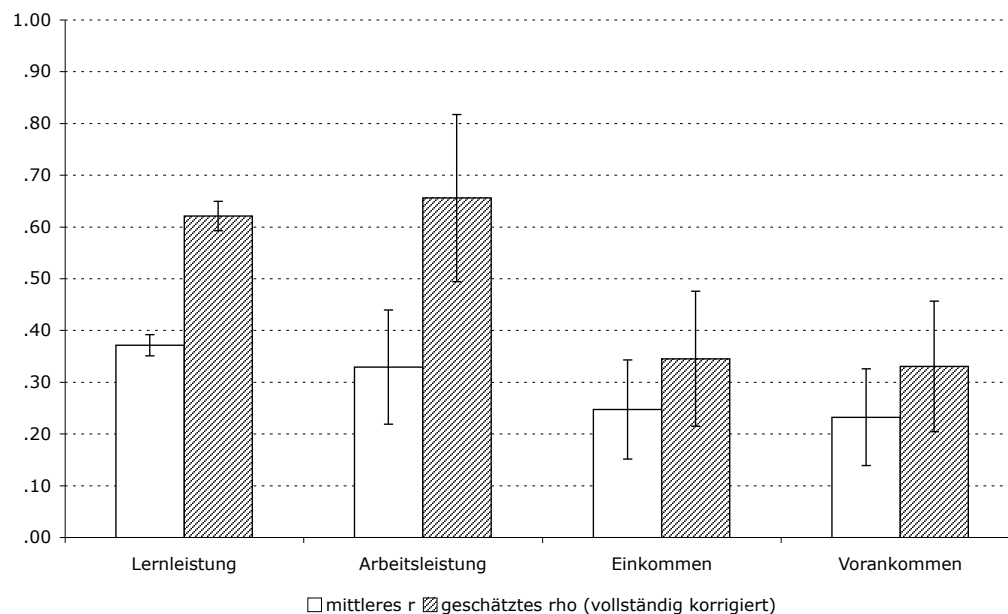
Karriereerfolg: bare-bone Metaanalysen zum Einkommen und beruflichen Vorankommen

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
Einkommen									
Null-Subst.	7	1980	.244	.129	.056	.117	18.8	.15–.34	11
VW-Subst.	7	1980	.247	.129	.056	.117	18.7	.15–.34	11
r_k -Subst.	7	1980	.249	.130	.056	.117	18.5	.15–.34	11
Exkl. n.ber.	6	1910	.246	.132	.053	.121	16.1	.14–.35	9
Exkl. MAg	6	1685	.247	.140	.056	.129	16.1	.13–.36	9
berufliches Vorankommen									
Null-Subst.	9	1817	.226	.150	.067	.134	20.0	.13–.32	12
VW-Subst.	9	1817	.232	.143	.067	.127	21.7	.14–.33	12
r_k -Subst.	9	1817	.250	.128	.066	.109	26.9	.17–.33	14
Exkl. n.ber.	7	1452	.278	.121	.064	.102	28.4	.19–.37	13
Exkl. MAg	8	1747	.234	.146	.064	.131	19.4	.13–.34	11

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k sowie bei Exklusion dieser Studien. k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k .

Abbildung 4.14

Validität von Intelligenz in Bezug auf vier berufliche Erfolgskriterien



Anmerkungen. X-Achse: Erfolgskriterien; Y-Achse: \bar{r} bzw. $\hat{\rho}$ vollständig korrigiert bei Inklusion aller Studien mit Verhältnissubstitution; Fehlerbalken: 95% Konfidenzintervalle.

Tabelle 4.39

Karriereerfolg: korrigierte Validitäten für Einkommen und berufliches Vorankommen

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
Einkommen													
Null-Subst.	7	1 980	.899	–	.780	.324	.165	.068	.150	16.8	.20–.45	.08	TRA
VW-Subst.	7	1 980	.899	–	.780	.327	.165	.066	.151	16.1	.20–.45	.08	TRA
r_k -Subst.	7	1 980	.899	–	.780	.329	.166	.066	.153	15.6	.21–.45	.08	TRA
Exkl. n.ber.	6	1 910	.912	–	.781	.327	.167	.061	.156	13.1	.19–.46	.07	TRA
Exkl. MAg	6	1 685	.897	–	.781	.326	.179	.066	.166	13.7	.18–.47	.05	TRA
berufliches Vorankommen													
Null-Subst.	9	1 817	.894	–	.779	.305	.191	.086	.170	20.2	.18–.43	.03	TRA
VW-Subst.	9	1 817	.894	–	.779	.312	.182	.085	.161	21.7	.19–.43	.05	TRA
r_k -Subst.	9	1 817	.894	–	.779	.334	.162	.082	.139	25.8	.23–.44	.11	TRA
Exkl. n.ber.	7	1 452	.901	–	.780	.369	.150	.079	.127	27.9	.26–.48	.16	TRA
Exkl. MAg	8	1 747	.903	–	.780	.315	.185	.080	.167	18.5	.19–.44	.04	TRA
vollständig korrigierte Validität													
Einkommen													
Null-Subst.	7	1 980	.899	–	.780	.342	.175	.074	.159	18.0	.21–.47	.08	TRA
VW-Subst.	7	1 980	.899	–	.780	.345	.176	.074	.160	17.6	.21–.48	.08	TRA
r_k -Subst.	7	1 980	.899	–	.780	.347	.177	.074	.161	17.3	.22–.48	.08	TRA
Exkl. n.ber.	6	1 910	.912	–	.781	.342	.178	.070	.163	15.4	.20–.48	.07	TRA
Exkl. MAg	6	1 685	.897	–	.781	.344	.190	.074	.175	15.1	.19–.50	.06	TRA
berufliches Vorankommen													
Null-Subst.	9	1 817	.894	–	.779	.323	.202	.091	.180	20.4	.19–.45	.03	TRA
VW-Subst.	9	1 817	.894	–	.779	.330	.193	.091	.171	22.0	.20–.46	.05	TRA
r_k -Subst.	9	1 817	.894	–	.779	.353	.172	.089	.147	26.8	.24–.47	.11	TRA
Exkl. n.ber.	7	1 452	.901	–	.780	.388	.159	.086	.134	29.0	.27–.51	.17	TRA
Exkl. MAg	8	1 747	.903	–	.780	.332	.196	.087	.176	19.5	.20–.47	.04	TRA

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k sowie bei Exklusion dieser Studien. k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; TRA = transportability.

Studien, die insignifikante Korrelationen nicht berichten, geringfügig unter denen der Maximalsubstitution und sogar geringfügig unter denen der Verhältnissubstitution bleiben. Dies liegt daran, dass die zu substituierende Studie (Lometsch, 1975) Korrelationen berichtet, die im Mittel über den beobachteten Werten der anderen Studien liegen ($r = .258$ vs. $\bar{r} = .246$) und im Vergleich zu den anderen Studien über ein vergleichsweise geringes n verfügt ($n = 70$ vs. $\bar{n} = 318.3$). Beides führt zu einem relativ hohen r_k von .198.

Die Exklusion von Studien mit Mittelwert-Aggregaten führt im Vergleich zu den Ergebnissen bei Verhältnissubstitution zu keinen bedeutsamen Veränderungen der mittleren Effekte. Der Anteil der aufgeklärten Varianz in der bare-bone Analyse liegt je nach Substitution bei 16.1% bis 18.8%, im korrigierten Fall liegt Transportabilität vor.

Berufliches Vorankommen

Wie in den meisten anderen Analysen liegt das Ergebnis bei Exklusion der Studien, die insignifikante Korrelationen nicht berichten, über dem maximal möglichen (Tabellen 4.38 und 4.39). Bei Verhältnissubstitution beträgt die beobachtete Validität $\bar{r} = .232$. 21.7% der Varianz können durch den Stichprobenfehler aufgeklärt werden. Vollständig (aber nicht für die Unreliabilität im Kriterium) korrigiert erhält man $\hat{\rho} = .330$. Die Exklusion der Studien mit Mittelwert-Aggregaten führt nur zu unbedeutend höheren mittleren Effekten ($\bar{r} = .234$, $\hat{\rho} = .332$). In allen Fällen liegt Transportabilität vor. Das kumulative Ergebnis der verhältnissubstituierten und vollständig korrigierten Analyse verdeutlicht Abbildung 4.13b.

Damit unterscheidet sich die Validität von Intelligenz in Bezug auf die beiden Karriereindikatoren *Einkommen* und *berufliches Vorankommen* kaum: weder in der Höhe noch in der Akkuratheit der Schätzung. Deutlich höher ist die Validität von Intelligenz in Bezug auf Lernleistung und – im korrigierten Fall – auch auf Arbeitsleistung (Abbildung 4.14).

4.3.3. Sensitivitätsanalysen

Ob ein Veröffentlichungsbias vorliegt, kann anhand der geringen Studienanzahl nicht entschieden werden. Zu beiden Karriereerfolgsbereichen liegen mehr nicht veröffentlichte als veröffentlichte Studien vor. Die FunnelpLOTS (Abbildung 4.15) machen deutlich, dass jeweils eine veröffentlichte Studie (sie stammt von Hossiep, 1995) mit überdurchschnittlich hoher Korrelation einen sehr großen Einfluss auf die Ergebnisse ausübt (Gewicht bei Einkommen 38.7% und bei beruflichem Vorankommen 35.4%).

Sensitivitätsanalysen, die ausschließlich auf prädiktiven Validitäten basieren, werden in Anbetracht der geringen Studienzahl für Einkommenskriterien nicht durchgeführt. Obwohl sechs Studien prädiktive Validitäten zum beruflichen Vorankommen berichten, ist eine Sensitivitätsanalyse auch für diesen Erfolgsindikator nicht aussagekräftig. Bei Exklusion einer Studie, die insignifikante Korrelationen nicht berichtet (Lometsch, 1975), sind nur fünf prädiktive Studien vorhanden. Eine Verhältnissubstitution ist ebenfalls nur

Tabelle 4.40

Berufliches Vorankommen: bare-bone Analyse bei Exklusion von Nicht-Pearson-Korrelationen

Analyse	k	N	\bar{r}	SD_{beo}	SD_{art}	$\widehat{SD}_{\bar{r}}$	VE (%)	CI (95%)	fs_k
Null-Subst.	8	1 625	.243	.150	.066	.134	19.6	.14–.35	12
VW-Subst.	8	1 625	.250	.142	.066	.126	21.7	.15–.35	12
r_k -Subst.	8	1 625	.270	.121	.065	.102	29.2	.19–.35	14
Exkl. n.ber.	6	1 260	.307	.101	.063	.079	38.5	.23–.39	13
Exkl. MAg	7	1 555	.253	.144	.063	.130	19.1	.15–.36	11

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k sowie bei Exklusion dieser Studien. k = Anzahl der Studien; SD_{beo} = beobachtete Standardabweichung; SD_{art} = durch den Stichprobenfehler erklärte Standardabweichung; $\widehat{SD}_{\bar{r}}$ = Standardabweichung von \bar{r} ; VE = Anteil der durch den Stichprobenfehler erklärten Varianz; CI = Konfidenzintervall; fs_k = faile-safe k .

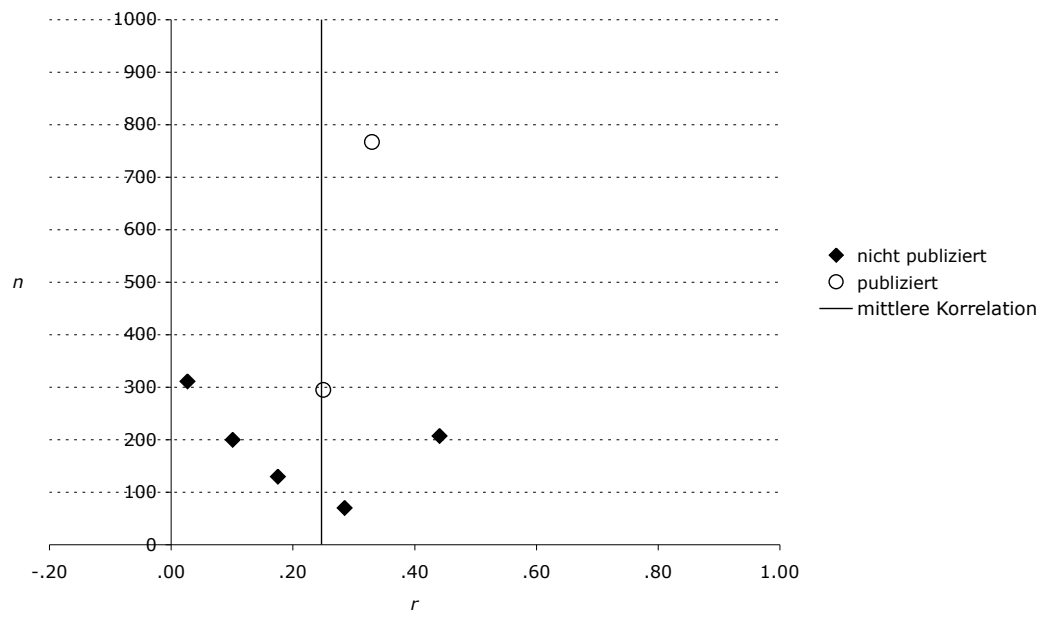
unzuverlässig möglich, da die Substitution nur auf Basis *eines* VW-Wertes (vgl. Kapitel 3.7.2) erfolgen würde.

Die Metaanalysen zum beruflichen Vorankommen enthalten eine Studie, die keine Korrelationen, sondern t -Werte berichtet (Steinmann, 1997). Wird sie exkludiert, erhält man geringfügig höhere Validitäten von Intelligenz mit beruflichem Vorankommen: sie betragen unkorrigiert $\bar{r} = .250$ statt $.232$ und korrigiert $\hat{\rho} = .356$ statt $.330$. Die anderen Kennwerte ändern sich ebenfalls nur geringfügig (Tabellen 4.40 und 4.41).

Die Gewichte ausgewählter Variablen (Berufskomplexität, Berufsgruppen, Prädiktoren) an den Sensitivitätsanalysen enthält Tabelle D.3 für den Erfolgsbereich *Einkommen* und Tabelle D.4 für den Erfolgsbereich *berufliches Vorankommen*. In den Tabellen wird auch deutlich, wie stark die Variablen, die den Sensitivitätsanalysen zu Grunde liegen, miteinander in Zusammenhang stehen. Eine Zusammenfassung der wichtigsten Ergebnisse zum Erfolgsbereich *Karriere* enthält Kapitel 5.4.

Abbildung 4.15
Funnelplots zum Karriereerfolg

a) Einkommen



b) berufliches Vorankommen

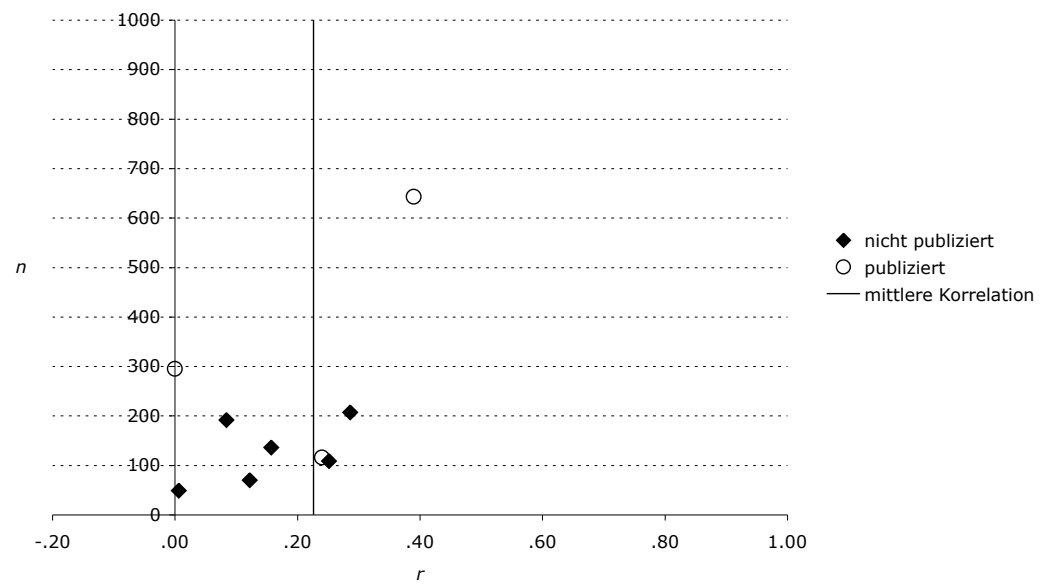


Tabelle 4.41

Berufliches Vorankommen: korrigierte Validitäten bei Exklusion von Nicht-Pearson-Korrelationen

Analyse	k	N	r_{xx}	r_{yy}	u_x	$\hat{\rho}$	SD_{beo}	SD_{art}	\widehat{SD}_{ρ}	VE (%)	CI 95%	CV 90%	Gen.
operationale Validität													
Null-Subst.	8	1 625	.891	–	.778	.329	.189	.085	.168	20.5	.20–.46	.05	TRA
VW-Subst.	8	1 625	.891	–	.778	.336	.178	.084	.157	22.5	.21–.46	.08	TRA
r_k -Subst.	8	1 625	.891	–	.778	.360	.150	.081	.126	29.2	.26–.46	.15	TRA
Exkl. n.ber.	6	1 260	.898	–	.779	.407	.121	.079	.092	42.4	.31–.50	.26	TRA
Exkl. MAg	7	1 555	.900	–	.778	.340	.181	.078	.163	18.8	.21–.47	.07	TRA
vollständig korrigierte Validität													
Null-Subst.	8	1 625	.891	–	.778	.348	.200	.090	.178	20.3	.21–.49	.05	TRA
VW-Subst.	8	1 625	.891	–	.778	.356	.189	.089	.166	22.4	.23–.49	.08	TRA
r_k -Subst.	8	1 625	.891	–	.778	.382	.160	.087	.134	29.9	.27–.49	.16	TRA
Exkl. n.ber.	6	1 260	.898	–	.779	.430	.128	.083	.097	42.6	.33–.53	.27	TRA
Exkl. MAg	7	1 555	.900	–	.778	.359	.191	.085	.172	19.6	.22–.50	.08	TRA

Anmerkungen. Analyse: bei Substitution insignifikanter nicht-berichteter Werte mit Null, VW oder r_k sowie bei Exklusion dieser Studien. k = Anzahl der Studien; r_{xx} = mittlere Reliabilität des Prädiktors; r_{yy} = mittlere Reliabilität des Kriteriums; u_x = mittlere beobachtete Varianzeinschränkung; SD_{beo} = beobachtete Standardabweichung der korrigierten Werte; SD_{art} = durch Artefakte erklärte Standardabweichung; \widehat{SD}_{ρ} = Standardabweichung von $\hat{\rho}$; VE = Anteil der durch Artefakte erklärten Varianz; CI = Konfidenzintervall; CV = untere Grenze des Glaubwürdigkeitsintervalls; Gen. = Generalisierung; TRA = transportability.

5. Diskussion

Ob Intelligenz ein valider Prädiktor für beruflichen Erfolg ist, wurde in dieser Arbeit im Hinblick auf drei Erfolgsbereiche untersucht: berufsbezogene Lernleistung, spezifische Arbeitsleistung und Karriereerfolg. Die Karriereerfolgs-Indikatoren Einkommen und berufliches Vorankommen wurden getrennt betrachtet. Berufliches Vorankommen umfasst Positions- und Beförderungskriterien (vgl. Kapitel 3.2.2). Die Ergebnisse sprechen dafür, dass Intelligenz in Bezug auf die genannten Erfolgsbereiche und Erfolgsindikatoren als valider Prädiktor gelten kann: Die Korrelationen zwischen Intelligenz und den Kriterien sind positiv und deutlich von Null verschieden. Dies gilt über unterschiedliche Berufe hinweg. Die Validität in Bezug auf die vier Indikatoren unterscheiden sich jedoch in ihrer Höhe, Akkuratheit, statistischen Belastbarkeit und Generalisierbarkeit.

Die Ergebnisse basieren zum Teil auf Studien, die insignifikante Korrelationen nicht berichten. Um diese Studien berücksichtigen zu können, wurde ein Substitutionsverfahren entwickelt, dessen Brauchbarkeit zunächst diskutiert wird. Im folgenden werden die wichtigsten Ergebnisse zunächst getrennt für die Erfolgsbereiche zusammengefasst und im Vergleich mit bereits vorliegenden metaanalytischen Ergebnissen bewertet. Damit wird die erste in Kapitel 2.3.4 formulierte Frage beantwortet: Sind Intelligenztests kriterienvalide in Bezug auf berufliche Lernleistung (Kapitel 5.2), spezifische Arbeitsleistung (Kapitel 5.3) und Karriereerfolg (Kapitel 5.4)? In Kapitel 5.2.1 wird die Moderatorwirkung verschiedener *Lernleistungskriterien* thematisiert (Forschungsfrage 2). Daran anschließend werden die Antworten auf die Forschungsfragen 3 und 4 gegeben, die sich mit den Moderatorwirkungen der Berufskomplexität (Kapitel 5.5) und der Berufsgruppe (Kapitel 5.6) sowie der verwendeten Intelligenztests (Kapitel 5.7) beschäftigen.

Schließlich wird betrachtet, welche allgemeinen Einschränkungen bei der Generalisierung der Befunde zu beachten sind (Kapitel 5.8), welcher weitere Forschungsbedarf sich daraus ergibt (Kapitel 5.9) und welche praktische Relevanz den Befunden zukommt (Kapitel 5.10).

5.1. Verhältnissubstitution

In der Regel werden Studien, die insignifikante Korrelationen nicht berichten, in Metaanalysen exkludiert. Alternative Vorgehensweisen sind die Nullsubstitution, die die wahren Effekte unterschätzt, und Regressionsverfahren, die jedoch nur zur Substitution zufällig fehlender Werten eingesetzt werden können (Rustenbach, 2003, vgl. Kapitel 3.7.2). Ziel dieser Arbeit ist die Berücksichtigung möglichst aller zur Verfügung stehender Informationen zur Validität von Intelligenz, auch wenn diese in ihrer genauen Höhe nicht bekannt sind und dafür in dichotomer Form vorliegen: nicht signifikant (vs. signifikant).

Dies ermöglicht die in Kapitel 3.7.2 vorgestellte Substitution der fehlenden Werte auf Basis der insignifikanten berichteten Korrelationen. Dazu wurden a) die Verhältnisse

der berichteten insignifikanten Korrelationen zu ihren kritischen r -Werten (r_k) berechnet und b) das mittlere Verhältnis der r_k -Werte den nicht berichteten insignifikanten Korrelationen zu Grunde gelegt (Verhältnissubstitution).

Mit Hilfe dieses Vorgehens konnten die Studien mit fehlenden Werten ebenfalls berücksichtigt werden. Die resultierenden mittleren Effekte sind wie erwartet höher als bei Substitution mit Null (die sie mit großer Wahrscheinlichkeit unterschätzen) und geringer als bei Substitution mit den maximal möglichen kritischen Werten (die sie überschätzen). Es konnte ferner gezeigt werden, dass die Exklusion der betroffenen Studien, die häufig angewendet wird, i.d.R. zu überschätzten mittleren Effekten führen: in 23 der 28 durchgeführten Metaanalysen mit substituierten Werten liegen die mittleren Effekte bei Exklusion über den maximal möglichen Werten bei Substitution mit r_k . Sie liegen nur dann innerhalb des möglichen Wertebereiches, wenn die Studien mit nicht-berichteten Korrelationen a) verhältnismäßig viele Korrelationen berichten und b) diese berichteten Korrelationen relativ groß sind. Beides führt zu vergleichsweise hohen Korrelationen, die exkludiert werden. Schließlich spielt c) die Stichprobengröße der Studien mit nicht-berichteten Korrelationen eine Rolle: sind sie relativ klein, resultiert dies in hohen r_k -Werten.

Die Substitutionswerte sind für jede nicht-berichtete Korrelation und jede Metaanalyse getrennt zu berechnen. Ihre Höhe hängt vom mittleren Verhältniswert (M_{VW}) und r_k ab. M_{VW} wird pro Metaanalyse berechnet. Die kritischen Werte r_k sind dagegen vom Signifikanzniveau, der Testseitigkeit und der Stichprobengröße der einzelnen Studie abhängig. Da sich die Signifikanzniveaus und Testseitigkeiten in den vorliegenden Studien nicht unterscheiden (5%, einseitig), sind die Substitutionswerte innerhalb einer Metaanalyse ausschließlich von der Stichprobengröße abhängig.

Die Stichprobengröße wird wiederum zur Korrektur der Stichprobenfehler herangezogen. Deshalb war eine Befürchtung, die Verhältnissubstitution könnte die Varianz überschätzen, die durch die Korrektur des Stichprobenfehlers aufgeklärt wird. Tatsächlich ist die Varianzaufklärung durch den Stichprobenfehler (VE% in den bare-bone Analysen) bei Verhältnissubstitution höher als bei Nullsubstitution, bei der ein konstanter Wert (Null) unabhängig von der Stichprobengröße zur Substitution verwendet wird. Sie ist aber i.d.R. geringer als bei Verwendung der kritischen Werte selbst, die ebenfalls von n abhängig sind. Außerdem ist sie i.d.R. geringer als bei Exklusion der Studien.

Die Unterschiede zwischen den Ergebnissen bei Verhältnissubstitution und Exklusion sind für die Metaanalysen auf Gesamt-Erfolgsbereich-Ebene minimal. Dies liegt an dem geringen Gewicht der Studien mit Korrelationen, die wegen Insignifikanz nicht berichtet sind, an den meisten Analysen (z.B. 4.4% bei Lernleistung, 2.6% bei Arbeitsleistung, 3.5% bei Einkommen als Kriterium). In einigen Analysen ist dieses Gewicht jedoch deutlich höher (z.B. 20.1% beim Kriterium berufliches Vorankommen). Den höchsten Anteil haben die Studien mit 41.5% bei der Moderatoranalyse zur Lernleistung von Abiturienten. In diesen Fällen ist auch der Unterschied zwischen den Substitutionsverfahren deutlicher. Die verschiedenen Substitutionsverfahren wurden in allen Studien durchgeführt, unabhängig vom Gewicht der fehlenden Werte, um die Auswirkungen der Verhältnissubstitution beobachten zu können.

Kritisch ist anzumerken, dass die verwendeten Substitutionswerte Schätzungen sind. Das zeigt sich z.B. darin, dass sie sich für *eine* Studie unterscheiden können, je nachdem, in welche Metaanalysen die Studie inkludiert wird (vgl. Tabelle 3.8). Sie kommen auf der Basis anderer Beobachtungen zu Stande, deren Generalisierung auf die betroffenen Studien fraglich erscheint. Insbesondere, wenn davon ausgegangen werden muss, dass jede Studie eine eigene Population darstellt (zufallsvariables Modell). Aus diesem Grund ist es empfehlenswert, Sensitivitätsanalysen durchzuführen, bei denen die betroffenen Studien exkludiert werden.

Theoretisch unabhängig von dem Problem der wegen Insignifikanz nicht berichteten Korrelationen ist ein zweites. Es betrifft Studien, die eine Aggregation mehrerer statistisch voneinander abhängiger Korrelationen erforderlich machen, weil mehrere (Sub-)Prädiktoren bzw. (Sub-)Kriterien verwendet werden. Die Berechnung der aggregierten Validität erfolgt im Idealfall mittels linearer Verbünde, die die Korrelationen innerhalb der Prädiktoren bzw. Kriterien berücksichtigen. Sind diese Intrakorrelationen nicht bekannt, kann die Aggregation nur durch einfache Mittelung erfolgen. Dabei wird der tatsächlich vorhandene Gesamteffekt unterschätzt (Hunter & Schmidt, 2004).

Tatsächlich sind die Stichproben zu berufsbezogener Lernleistung und Arbeitsleistung, die insignifikante Korrelationen nicht berichten, in hohem Maße auch von dem Aggregierungsproblem betroffen (zu 90.3% bei Lern- und zu 100.0% bei Arbeitsleistung). D.h. es handelt sich um Studien, die selbst bei angemessener Substitution der fehlenden Werte zur Unterschätzung des mittleren Effektes beitragen. Möglicherweise ist das der Grund dafür, dass bei Exklusion der Studien mit nicht berichteten Korrelationen mittlere Effekte resultieren, die über denen bei Substitution mit r_k liegen?

Um dies auszuschließen, wurden Sensitivitätsanalysen durchgeführt, bei denen die Studien, die vom *Aggregierungsproblem* betroffen sind exkludiert wurden. Dadurch wurde das Gewicht der Studien mit fehlenden Werten stark vermindert. Die Unterschiede zwischen den Substitutionsverfahren sind minimal, entsprechen jedoch den oben genannten Befunden.

Die Studien, die insignifikante Korrelationen nicht berichten, unterscheiden sich von denen ohne fehlende Werte auch in Bezug auf die anderen Variablen mehr oder weniger stark (Tabellen D.1, D.2, D.3 und D.4). Diese Studienmerkmale könnten mitverantwortlich sein für die unterschiedlichen Ergebnisse bei Substitution bzw. Exklusion nicht berichteter Werte.

Um die Brauchbarkeit der Verhältnissubstitution abschließend zu klären, sind weitergehende Analysen (z.B. Simulationsstudien) erforderlich, die nicht Gegenstand dieser Arbeit sind. Die bisherigen Befunde sind vielversprechend und lassen die Verhältnissubstitution als eine brauchbare Alternative zu anderen Substitutionsverfahren und der Exklusion von Studien erscheinen. Dies gilt zumindest dann, wenn diese über ein bedeutendes Gewicht an der Metaanalyse verfügen und trotzdem genügend Studien vorhanden sind, die insignifikante Korrelationen berichten, um VW-Werte errechnen zu können.

5.2. Berufsbezogene Lernleistung

Zu berufsbezogener Lernleistung liegen 210 Studien ($N = 30\,451$) vor. Die mittlere beobachtete Korrelation beträgt bei Verhältnissubstitution $.37$ (95% CI = $.35-.39$), die vollständig korrigierte Validität $.62$ (CI = $.59-.65$). Sie liegt damit etwas über der operationalen Validität, die $.59$ beträgt (CI = $.56-.61$). Die Varianzaufklärung liegt mit 22.9–25.9% deutlich unter den 75%, die für eine Generalisierung der Werte erforderlich wären. Die Glaubwürdigkeitsintervalle liegen deutlich über Null, die Richtung des Effektes ist somit generalisierbar, der Zusammenhang zwischen Intelligenztest-Leistungen und Aus- bzw. Weiterbildungserfolg positiv. Nicht betrachtet wurde Studienerfolg als Kriterium, von dem ebenfalls angenommen werden kann, dass er positiv mit Intelligenz korreliert (Süß, 2001).

Der Einfluss der verschiedenen Substitutionsverfahren auf das Ergebnis zur Lernleistung ist marginal. Die Exklusion der neun Studien, die vor 1945 durchgeführt wurden, verfügen über eine deutlich höhere mittlere Korrelation ($\bar{r} = .71$), haben aber mit 0.4% ein geringes Gewicht an der Analyse. Die hohen Korrelationen der Studien vor 1945 können evtl. auf Besonderheiten im Design zurück zu führen sein. So basiert die Studie von Roloff (1928) auf wenigen Teilnehmern, die zudem auf Basis ihrer Leistungsbeurteilungen ausgewählt wurden, mit dem Ziel eine große Streuung im Kriterium zu erreichen. Die Beschränkung auf die 197 Studien, die prädiktive Validitäten berichten, hat ebenfalls keine nennenswerten Auswirkungen auf das Ergebnis, obwohl sich die acht konkurrenten Studien in der Höhe von $\bar{r} = .29$ (CI = $.22-.37$) und $\hat{\rho} = .55$ (CI = $.43-.66$) deutlich davon unterscheiden. Es ist nicht erwartungskonform, dass sie geringere Validitäten als prädiktive Studien aufweisen. Die Ursache dafür kann im großen Gewicht der Studien mit Mittelwert-Aggregaten an den konkurrenten Studien (45.5%) begründet sein. Die Exklusion der Studien, die Aggregationen über Mittelwerte erforderlich machen, führt insgesamt nur zu etwas höheren mittleren Effekten ($\bar{r} = .39$, $\hat{\rho} = .65$).

Die genannten Befunde sprechen für die Robustheit des Ergebnisses zur Lernleistung. Allerdings lassen sich Hinweise auf eine Quellenverzerrung finden: generell sind die berichteten Validitäten in Testmanualen ($.43$, CI = $.41-.46$) höher als in anderen Monographien und Dissertationen ($.32$, CI = $.27-.37$) bzw. Artikeln, die Zeitschriften oder Herausgeberwerken entstammen ($.34$, CI = $.29-.38$). Die geringsten Validitäten berichten nicht veröffentlichte Quellen ($.13$, CI = $.20-.30$). Der Funnelplot (Abbildung 4.4) ist augenscheinlich nicht symmetrisch, was ebenfalls den Einfluss einer Veröffentlichungsverzerrung nahe legt. Die Prüfung des Funnelplots mit der *trim and fill* Methode nach Duval (2005) konnte diesen Verdacht aber nicht erhärten (Kapitel 4.1.3).

Im Vergleich mit den bereits veröffentlichten Metaanalysen von Salgado und Anderson (2003) und Hülshager et al. (2007) konnten in dieser Arbeit deutlich mehr Studien inkludiert werden, was zu einer besseren Generalisierbarkeit und höheren Akkuratheit der Ergebnisse führt. Die beobachtete Validität von Intelligenz in Bezug auf berufliche Lernleistung auf Basis der hier durchgeführten Analyse ($.37$) liegt über der von Salgado et al. ($.34$) und Hülshager et al. ($.31$). Diese Werte sind unbeeinflusst von den methodischen Entscheidungen, die zur Berechnung der korrigierten Werte getroffen wurden. Die operationale Validität in dieser Arbeit ($.59$, 95% CI = $.56-.61$) liegt dementsprechend

deutlich über derjenigen, die Hülshager et al. berichten (.47, 95% CI = .43–.51) und unter der von Salgado et al. berichteten (.63, CI nicht angegeben).

Welche korrigierten Werte erzielt werden, hängt insbesondere von den gewählten Artefaktkorrekturen ab. Die Unreliabilität im Kriterium wurde in allen drei Arbeiten korrigiert. Die Korrektur ist bei Hülshager et al. am geringsten ($r_{yy} = .80$), höher in dieser Arbeit (.64) und am höchsten bei Salgado et al. (.56). Die Korrekturwerte basieren auf externen Daten, im Falle von Hülshager et al. und dieser Arbeit auf der gleichen Quelle (Baron-Boldt et al., 1989), die aber unterschiedlich interpretiert wird: Die Reliabilität ist bei Baron-Boldt et al. radiziert als Korrekturwert angegeben (.80). Er wird von Hülshager et al. als Reliabilitätskennwert zu Grunde gelegt wird. In dieser Arbeit wird $\sqrt{.80}$ verwendet.

Die Varianzeinschränkung im Prädiktor ist in dieser Arbeit geringer ($u_x = .79$) als bei Hülshager et al. (.69) bzw. Salgado et al. (.68). Diese Werte basieren auf Angaben in den inkludierten Studien, deren *SDs* zu denen in der Population in Beziehung gesetzt werden. Während bei Hülshager et al. und Salgado et al. die *SDs* der Gesamtbevölkerung als Referenzgröße dienen, wurden in dieser Arbeit die *SDs* der erwerbstätigen Bevölkerung zu Grunde gelegt. Außer in der Höhe der Varianzeinschränkungen unterscheiden sich die Arbeiten auch in der Korrektur dieses Artefakts. Hülshager et al. und Salgado et al. korrigieren *direkt*, in dieser Arbeit wurde *indirekt* korrigiert. Die indirekte Korrektur wird im vorliegenden Gegenstandsbereich von Hunter et al. (2006) für angemessen gehalten. Sie führt zu höheren korrigierten Validitäten als die direkte Korrektur (Kapitel 3.6.3).

Zur Berechnung der Varianzen von $\hat{\rho}$ wurde in allen drei Arbeiten zusätzlich die Unreliabilität im Prädiktor korrigiert. Die Korrektur dieses Artefaktes ist bei Hülshager et al. am geringsten ($r_{yy} = .90$), geringfügig höher in dieser Arbeit (.88) und am höchsten bei Salgado et al. (.83). Die Unterschiede zwischen der Analyse von Hülshager et al. und dieser Arbeit sind gering, obwohl unterschiedliche Vorgehensweisen zur Bestimmung der Korrekturwerte gewählt wurden: während bei Hülshager et al. die Retestreliabilität zu Grunde gelegt wird, ist es in dieser Arbeit die CES-Reliabilität nach Cronbach (1947, vgl. Kapitel 3.6.1).

Der Anteil der durch alle Artefakte aufgeklärten Varianz an der beobachteten Varianz der operationalen Validität ist bei Salgado et al. am höchsten (51%), geringer bei Hülshager et al. (34%) und am geringsten in der vorliegenden Arbeit (26%). Möglicherweise liegt dies daran, dass es leichter ist, eine hohe Varianzaufklärung bei Inklusion von wenigen Studien zu erreichen (N. Schmitt et al., 1984). Dies ist einer der Gründe, wieso von einer metaanalytischen Integration von weniger als sechs Studien abgeraten wird (vgl. Kapitel 3.8). Der Anteil der aufgeklärten Varianz ist bedeutsam für die Einschätzung der Generalisierungsmöglichkeiten. Da er in allen Arbeiten deutlich unter 75% liegt und die Glaubwürdigkeitsintervalle Null ausschließen, ist von Transportabilität auszugehen. Dies ist ein konsistenter Befund aller drei Metaanalysen, die Studien aus Deutschland betrachten.

Darin unterscheiden sich die Befunde aus Deutschland auch nicht von denen, die Salgado et al. (2003a) für europäische Länder finden. Die beobachteten und operationalen Validitäten den europäischen Ländern insgesamt sind etwas geringer ($\bar{r} = .28$, $\hat{\rho} = .54$).

Die amerikanischen Befunde werden von F. L. Schmidt und Hunter (1998a) mit $\hat{\rho} = .56$ angegeben, dabei handelt es sich allerdings um eine vollständig korrigierte Validität; der entsprechende hier ermittelte Wert für Deutschland liegt bei .62.

Die geringe Varianzaufklärung bei der Metaanalyse zur Lernleistung legt nahe, dass Moderatoren wirken. A priori wurde in dieser Arbeit nach dem Einfluss von vier potentiellen Moderatoren gefragt: dem Erfolgskriterium, der Berufskomplexität, der Berufsgruppe und dem Intelligenztest (Kapitel 2.3.4). Die Ergebnisse zu den Moderatoren, die über alle Erfolgsbereiche hinweg wirken, werden in den Kapiteln 5.5 (Berufskomplexität), Kapitel 5.6 (Berufsgruppe) und Kapitel 5.7 (Intelligenztest) diskutiert. Die Einflüsse der Lernleistungskriterien werden im Folgenden betrachtet.

5.2.1. Validität im Hinblick auf verschiedene Lernleistungskriterien

Es wurden zwei Unterscheidungen der Lernleistungskriterien vorgenommen: Berufsschulnoten vs. betriebliche Beurteilungen sowie Ausbildungs- vs. Weiterbildungserfolg. Die 129 Studien, die Berufsschulnoten berichten, weisen mit einem \bar{r} von .34 (95% CI = .31–.37) höhere beobachtete Validitäten auf als die 28 Studien, die betriebliche Beurteilungen berichten ($\bar{r} = .23$, CI = .16–.29). Bei Exklusion der Studien mit Mittelwert-Aggregation ist der Unterschied noch etwas deutlicher: .37 (CI = .33–.40) vs. .24 (CI = .15–.33).

Die Reliabilitäten der Kriterien, die der Korrektur zu Grunde gelegt wurden, sind .64 für Berufsschulnoten (Baron-Boldt et al., 1989) und .52 für betriebliche Beurteilungen (Rothstein, 1990; Viswesvaran et al., 2002). Die resultierenden korrigierten Werte liegen für Berufsschulnoten bei $\hat{\rho} = .58$ und für betriebliche Beurteilungen bei $\hat{\rho} = .42$ ($\hat{\rho}$ bei Exklusion .62 und .43). Diese Ergebnisse decken sich mit den Erwartungen, da Intelligenztests und Berufsschulnoten eine größere gemeinsame Methodenvarianz aufweisen als Intelligenztests und betriebliche Beurteilungen (vgl. Kapitel 2.3).

Hülshager et al. (2006) haben diese beiden Kriterien ebenfalls getrennt analysiert und für betriebliche Beurteilungen nahezu die selben Ergebnisse gefunden ($k = 11$, $\bar{r} = .23$, $\hat{\rho} = .43$). Als Maß für die Kriterienreliabilität wurde von ihnen ebenfalls .52 verwendet. Die Validitäten der Noten liegen auch in der Studie von Hülshager et al. über denen der betrieblichen Beurteilungen. Sie erreichen aber mit $\bar{r} = .32$ und $\hat{\rho} = .47$ nicht ganz die Höhe der hier gefundenen Werte. Dies erstaunt insofern nicht, da Hülshager et al. als Maß für die Kriterienreliabilität .80 verwenden und Varianzeinschränkungen direkt korrigieren, während in dieser Arbeit .64 verwendet wird und Varianzeinschränkungen indirekt korrigiert werden (s.o.).

Geringe Unterschiede sind zwischen 190 Aus- und 20 Weiterbildungsleistungs-Studien festzustellen. Erstere weisen beobachtete Validitäten von .37 auf (CI = .35–.39), letztere von .41 (CI = .32–.49). Die korrigierten Werte betragen .62 und .68. Die Exklusion der Studien mit Mittelwert-Aggregation wirkt sich zugunsten der Weiterbildungs-Studien aus: ihre Validität beträgt dann unkorrigiert .50 (CI = .42–.57), das $\hat{\rho}$ liegt bei .79. Die Werte der Ausbildungs-Studien ändern sich nur geringfügig auf $\bar{r} = .39$ (CI = .36–.41) und $\hat{\rho} = .64$.

Weiterbildungsleistung baut zu einem größeren Teil auf Vorwissen auf als Arbeitsleis-

tung; die Erfahrung spielt eine größere Rolle. Intelligenten Personen gelingt es besser, Fachwissen und Erfahrung anzusammeln und anzuwenden (Hunter, 1986). Dies macht die höhere Kriterienvalidität von Intelligenz in Bezug auf Weiterbildungsleistung plausibel.

5.3. Subjektiv bewertete Arbeitsleistung

Die fremdbeurteilte, subjektive Arbeitsleistung wurde in 18 Studien als Kriterium verwendet. Diese relativ geringe Studienzahl erlaubt es nicht, für die interessierenden Moderatorvariablen (Berufskomplexität, Berufsgruppen, Prädiktoren und Kriterien) getrennte Analysen zu berechnen.

Die Berufsgruppe *Techniker und Gleichrangige* ist mit vier, *Bürokräfte* mit drei, *leitende Berufe*, *Wissenschaftler* und *Maschinenbediener* sind mit je zwei Studien vertreten, eine Studie liegt zum *Handwerk* vor. Vier Studien können nicht eindeutig einer Berufsgruppe zugeordnet werden. *Dienstleistungs-*, *Landwirtschaft* und *Fischerei*-Berufe und *Hilfsarbeitskräfte* sind nicht vertreten. Eine Generalisierung der Befunde auf diese Berufsgruppen ist deshalb nicht möglich. Mit sechs Studien ist der Skill-Level 2 am einflussreichsten, vier Studien liegen zum Skill-Level 3 vor, zwei zum Skill-Level 4 und vier Studien lassen sich keinem Skill-Level zuordnen.

Die verwendeten Intelligenztest sind heterogen: der IST in den Versionen 1953–1970 (neun Studien), der WPT (drei Studien), der FRT (zwei Studien), Eigenentwicklungen (zwei Studien), sowie je einmal der BIS 4 und der WIT in der Version von 1963 kommen zum Einsatz, eine Generalisierung der Ergebnisse über verschiedene Testverfahren ist möglich.

Fremdurteiler waren überwiegend die Vorgesetzten der Beurteilten, in einer Studie zum Teil Kollegen. Die beurteilten Leistungsaspekte sind heterogen. Sie umfassen nicht nur Dimensionen der Kernleistung, sondern auch umfeldbezogene und adaptive Leistungsdimensionen (Tabelle 3.3). Nur in geringem Maße berücksichtigt ist z.B. Führungserfolg, für den ebenfalls positive Zusammenhänge mit Intelligenz zu erwarten sind. Dies zeigt zumindest eine Metaanalyse mit 151 internationalen Studien von Judge et al. (2004) mit subjektiv und objektiv gemessener Führungseffektivität als Kriterium. Sie finden einen Zusammenhang mit Intelligenz von $\bar{r} = .17$ bzw. nach Korrektur der Unreliabilitäten und Varianzeinschränkung im Prädiktor von $\hat{\rho} = .27$.

Validierungsstudien mit objektiven Arbeitsleistungs-Kriterien, z.B. Produktionsdaten liegen nicht in ausreichender Menge vor. Eine Studie, die diese berichten, wurde deshalb exkludiert (Wiegand, 1977). Eine Generalisierung der Befunde auf objektive Arbeitsleistungs-Kriterien ist deshalb nicht möglich. Hunter und Schmidt (1996) zeigen, dass diese Maße höher mit Intelligenz korrelieren als subjektiv bewertete Arbeitsleistung (vgl. Abbildung 2.3), da subjektive Einschätzungen stärker von nicht arbeitsbezogenen Faktoren beeinflusst sind (Kapitel 2.3.2). Die Kriterienvalidität von Intelligenz lässt sich somit nachweisen, *obwohl* ausschließlich subjektive Einschätzungen der Arbeitsleistung berücksichtigt wurden.

Die mittlere beobachtete Korrelation von Intelligenz mit subjektiv bewerteter Arbeits-

leistung beträgt .33 (95% CI = .22–.44), die vollständig korrigierte Validität liegt bei .66 (CI = .49–.82). Hinweise auf mögliche Veröffentlichungsverzerrungen liefert der Funnelplot (Abbildung 4.11). Die Prüfung durch zwei Kennwerte der *trim and fill* Methode (Duval, 2005) sprechen in einem Fall gegen und im anderen für die Wirkung einer Verzerrung. Folgt man dem Kennwert, der eine mögliche Verzerrung anzeigt und korrigiert diese, resultiert daraus ein deutlich geringeres \bar{r} von .19 mit höherer Varianz ($SD_{\bar{r}} = .11$).

Sensitivitätsanalysen zur Exklusion von Studien, die insignifikante Korrelationen nicht berichten, auf Nicht-Pearson-Korrelationen beruhen oder keine prädiktiven Validitäten berichten, wurden durchgeführt (Kapitel 4.2.3). Die Ergebnisse dieser Studien zeigen nur geringfügige Abweichungen: die \bar{r} s liegen zwischen .32 und .35, die $\hat{\rho}$ s zwischen .64 und .69. Die drei nicht publizierten Studien weisen unterdurchschnittliche Validitäten auf (vgl. Abbildung 4.11). Dies deutet neben der *trim and fill*-Analyse auf einen möglicherweise vorliegenden Veröffentlichungsbias hin.

Einen deutlichen Effekt hat auch die Exklusion der Studien, die Mittelwert-Aggregate enthalten: das \bar{r} der verbleibenden 14 Studien liegt bei .42 (CI = .27–.56), $\hat{\rho}$ bei .80. Die Mittelwert-Aggregate betreffen ausschließlich die Kriterien, eine Aggregation der Prädiktoren über Mittelwerte war nicht erforderlich. Das maximale berichtete r beträgt .71. Es stammt aus einer Studie, die ein Gewicht von 21.8% an der Gesamtanalyse hat (Studie 116.12 aus Amthauer, 1973). Wird dieser Extremwert exkludiert, resultieren deutlich verminderte mittlere Effekte ($\bar{r} = .22$, $\hat{\rho} = .46$). Der Anteil der durch die Artefakte aufgeklärten Varianz steigt um das über 3fache, erreicht aber trotzdem nicht annähernd die 75%-Marke.

Die Befunde zur Arbeitsleistung sind also mit einer größeren Unsicherheit behaftet als die Ergebnisse zur berufsbezogenen Lernleistung. Sie sind auch stärker von Artefakten beeinflusst: die korrigierte Validität bei den Arbeitsleistungsstudien übertrifft die der Lernleistungsstudien geringfügig, obwohl die beobachtete Validität der Arbeitsleistungsstudien geringer ist als die der Lernleistungsstudien. Der durch die Artefakte aufgeklärte Varianzanteil liegt bei Arbeitsleistung als Kriterium weit unter 75% (bei Inklusion aller Studien bei 12.0%), die unteren Grenzen der Glaubwürdigkeitsintervalle sind konsistent von Null verschieden. Damit liegt Transportabilität vor: die Richtung des Effektes kann verallgemeinert werden, nicht seine Höhe.

Die Validität von Intelligenz in Bezug auf subjektiv bewertete Arbeitsleistung in Deutschland wurde auch von Salgado und Anderson (2003) und Hülshager et al. (2007) untersucht. Die inkludierte Studienanzahl gegenüber den beiden Vorläuferstudien konnte ungefähr verdoppelt werden. Die unkorrigierten mittleren Korrelationen sind trotzdem nahezu die selben, sie betragen bei Salgado et al. .35 und bei Hülshager et al. .33. Die operationale Validität geben Salgado et al. mit .68 und Hülshager et al. mit .53 an. Der Unterschied in der operationalen Validität zwischen den Ergebnissen von Hülshager et al. und den hier gefundenen ist weder auf die beobachteten Validitäten zurückzuführen ($\bar{r} = .33$ in beiden Fällen) noch auf die Korrektur der Unreliabilität im Kriterium, da die Reliabilität in allen Arbeiten mit .52 angesetzt wurde. Die Unterschiede sind vermutlich in den Varianzeinschränkungen zu suchen. Sie betragen bei Hülshager et al. lediglich $u_x = .85$, während sie in dieser Arbeit .74 und bei Salgado et al. .68 betragen. Zudem wur-

den indirekte Korrekturen vorgenommen, während Hülshager et al. direkt korrigieren. Trotz dieser Unterschiede lässt sich konsistent feststellen, dass Intelligenzverfahren in Bezug auf Abwärtsbeurteilungen in Deutschland valide sind.

5.4. Karriereerfolg

Zum Erfolgsbereich *Karriere* liegen zehn Studien vor, die unterschiedliche Indikatoren für Karriererfolg verwenden: das aktuelle Einkommen, das Einkommen zu einem bestimmten Stichtag, die Einkommensentwicklung, die Tarifgruppe, die Anzahl der Beförderungen, die Hierarchiestufen, der erreichte berufliche Status und die Erlaubnis zur Teilnahme an Personalentwicklungs-Maßnahmen. Da für Metaanalysen möglichst homogene Kriterien benötigt werden, um interpretierbare Ergebnisse erhalten zu können, wurden aus diesen Indikatoren zwei Gruppen gebildet: die Einkommensmaße und Tarifgruppen wurden zur Indikatorgruppe *Einkommen* und die anderen genannten Maße zur Indikatorgruppe *berufliches Vorankommen* zusammengefasst.

5.4.1. Einkommen

Sieben Studien berichten Zusammenhänge zwischen Intelligenz- und Einkommensmaßen. Die Studien sind überwiegend nicht veröffentlicht, eine Veröffentlichungsverzerrung in positiver Richtung ist deshalb unwahrscheinlich. Die verwendeten Intelligenzmaße sind heterogen: es kommen veröffentlichte Strukturtests (IST 2000 R, WIT von 1963), Omnibusverfahren (LPS, WPT) und eine eigenentwickelte Testbatterie zum Einsatz. Die BIS-Inhaltsfacetten sind ungefähr gleich stark vertreten. Bei den Operationen gehen ca. 75% auf den Einfluss der Verarbeitungskapazität zurück, die restlichen 25% sind hauptsächlich der Bearbeitungsgeschwindigkeit zuzuschreiben. Einfallsreichtum und Merkfähigkeit spielen kaum eine Rolle. Ebenfalls heterogen sind die Berufe: vier Studien lassen sich nicht eindeutig Berufsgruppen oder Skill Levels zuordnen. In zwei Studien handelt es sich um Bürokräfte mit Skill-Level 2, in einer Studie werden leitende Berufstätige untersucht, denen kein Skill-Level zugeordnet wird. Das Gesamtergebnis wird zu 38.7% von einer Studie (Hossiep, 1995) bestimmt.

Der mittlere Effekt beträgt $\bar{r} = .25$ (95% CI = .15–.34), $\hat{\rho} = .35$ (CI = .21–.48). Diese Werte verändern sich nur unerheblich, wenn eine Studie, die Mittelwert-Aggregate enthält, exkludiert wird. Die Richtung des Zusammenhangs kann verallgemeinert werden (CV > Null), die Höhe nicht, da mit 17.6% deutlich weniger als die geforderten 75% der Varianz durch Artefakte aufgeklärt werden. Die geringe Anzahl der inkludierten Studien resultiert in einem breiten Konfidenzintervall und in einer geringen faile-safe- k -Anzahl von elf.

Weitergehende Sensitivitäts- und Moderatoranalysen sind auf Grund der geringen Studienzahl nicht möglich. Die Ergebnisse können deshalb nur als erster Hinweis dienen, in welcher Richtung Intelligenz mit Einkommensmaßen korreliert. Die Konfidenzintervalle der korrigierten Werte überlappen sich nicht mit denen bei Lern- und subjektiver Arbeitsleistung. Im Vergleich mit Lernleistung sind auch die Konfidenzintervalle der beob-

achteten Werte ohne Überschneidung. Es ist deshalb davon auszugehen, dass Intelligenz geringer mit Einkommensmaßen korreliert als mit Lern- und Arbeitsleistung.

Dieser Erfolgsbereich wurde noch nicht mit Studien aus Deutschland metaanalytisch untersucht. Es liegen aber Studien aus dem angloamerikanischen Kulturraum vor: N. Schmitt et al. (1984) beziffern die beobachtete Validität in Bezug auf Umsatzerlöse nach Auswertung von acht Studien auf .14 ($N = 12\,449$), Vinchur et al. (1998) aggregieren Korrelationen mit dem Einkommen von Verkäufern aus 18 Studien und erhalten quasi Nullkorrelationen ($N = 1\,876$). Ng et al. (2005) schließlich beziffern die für den Prädiktor attenuationskorrigierte Validität auf .27 ($k = 8$, $N = 9\,560$; vgl. Tabelle 2.7). Angesichts der geringen Studienzahl auch in diesen Analysen können die Ergebnisse nur als erste Hinweise auf die tatsächliche Validität von Intelligenz in Bezug auf Einkommen interpretiert werden.

5.4.2. Berufliches Vorankommen

Die Indikatoren, die in der Gruppe *berufliches Vorankommen* zusammengefasst wurden, sind nicht nur inhaltlich heterogen (Anzahl der Beförderungen, Hierarchiestufe, Berufsstatus, Erlaubnis zur Teilnahme an Personalentwicklungsmaßnahmen), sondern umfassen verschiedene Datenarten (neutrale Kennzahlen, Vergleiche mit Bezugsstandards) aus verschiedener Datenquellen (Akten, Fremdbeurteilungen und Selbstberichte).

Neun Studien wurden zum Erfolgsbereich berufliches Vorankommen gefunden und inkludiert. Drei davon sind veröffentlicht. Als Prädiktoren kommen ebenfalls Strukturtests (IST 2000 R, WIT von 1963), Omnibusverfahren (WPT und LPS) sowie selbst entwickelte Testbatterien zum Einsatz. Die BIS-Inhalte tragen zwischen 35.9% (verbal) und 28.3% (figural) zum Ergebnis bei, die Operation *Verarbeitungskapazität* ca. 80%, die verbleibenden 20% sind überwiegend der Bearbeitungskapazität zuzuschreiben und zu einem geringen Anteil der Merkfähigkeit. Die Operation *Einfallsreichtum* ist nicht vertreten.

Bei zwei Studien wurden Bürokräfte untersucht, mit dem Skill-Level 2, eine Studie konnte der Gruppe *Techniker und gleichrangige* zugeordnet werden (Skill-Level 3), eine weitere den *leitenden Berufen* (ohne Skill-Level). Die verbleibenden fünf Studien sind weder den Berufshauptgruppen noch den Skill-Levels eindeutig zuzuordnen. Die Studie von Hossiep (1995) trägt wiederum einen großen Teil zum Gesamtergebnis bei: ihr Gewicht in der *bare bone*-Analyse beträgt 35.4%.

Das \bar{r} beträgt .23 (95% CI = .14–.33), die vollständig korrigierte Validität .33 (CI = .20–.46). Die Exklusion einer Studie mit Mittelwert-Aggregaten führt zu unbedeutend höheren mittleren Effekten (Kapitel 4.3.3). Weitere Sensitivitäts- oder Moderatoranalysen sind auf Grund der geringen Studienzahl nicht möglich. Richtung aber nicht Höhe des Effektes können verallgemeinert werden. Die Anzahl der Studien mit Nullkorrelationen, die erforderlich wären, um den mittleren Effekt unter .10 zu drücken liegt bei zwölf. Die mittleren Validitäten im Bezug auf Einkommen und Vorankommen unterscheiden sich kaum voneinander, aber von denen bei Lern- und Arbeitsleistung als Kriterien. Im Vergleich mit den Befunden von N. Schmitt et al. (1984) zu Beförderungen als Kriterium ($\bar{r} = .28$, $k = 9$, $N = 21\,190$) sind die Validitäten etwas geringer.

5.5. Berufskomplexität

Die Berufskomplexität wurde in der Metaanalyse von Hunter und Hunter (1984) als wesentlicher Moderator der Validität identifiziert: je größer die Komplexität, desto höher die Validität von Intelligenz in Bezug auf Trainings- und Arbeitsleistung (vgl. Kapitel 2.3). In dieser Arbeit wurde die Berufskomplexität mit den Skill-Leveln der *International Standard Classification of Education* operationalisiert, die in den International Standard Classification of Occupations (ISCO-88; International Labour Office, 1990) enthalten sind (vgl. Kapitel 3.4.1).

Der Moderatoreinfluss konnte in Bezug auf Arbeitsleistung und Karriereerfolg nicht geprüft werden, da zu wenige Studien vorliegen. Bei Lernleistungsstudien war eine Prüfung eingeschränkt möglich: Berufe mit den mittleren Skill-Leveln 2 (Anlernberufe bis Meister) und 3 (Techniker und gleichrangige Berufe) liegen in ausreichender Anzahl vor. Zu den Skill-Leveln 1 (Hilfsarbeitskräfte) und 4 (Wissenschaftler) sind nicht genügend Studien für eine metaanalytische Integration vorhanden.

Die Validitäten der Skill-Level 2 und 3 unterscheiden sich *nicht* wie erwartet: das beobachtete \bar{r} der 138 Studien zu weniger komplexen Berufen (Skill-Level 2) beträgt .40, der entsprechende Wert für die 50 Studien zu den komplexeren Berufen (Skill-Level 3) beträgt .31. Wie in Kapitel 2.3.1 dargestellt, vermuten Schuler und Höft (2006) die Ursache in der höheren Varianzeinschränkung im Prädiktor bei komplexen Berufen. Der Unterschied dürfte sich dann jedoch nur im beobachteten \bar{r} und nicht im (varianzeinschränkungs-)korrigierten $\hat{\rho}$ zeigen. Dies ist nicht der Fall: Die $\hat{\rho}$ unterscheiden sich zwischen Skill-Level 2 und 3 ebenfalls mit .65 vs. .54 entsprechend den \bar{r} -Werten. Möglicherweise liegt die Ursache dafür in den Varianzeinschränkungen der vorliegenden Studien. Sie sind für beide Skill-Level nahezu gleich groß: $u = .79$ bei Skill-Level 2 vs. $u = .77$ bei Skill-Level 3. Die u -Werte basieren allerdings auf wenigen Studien und lassen die eigentlich interessanten extremeren Skill-Level 1 und 4 außer acht.

Ergänzend dazu ist zu beachten, dass die Unterschätzung durch die Mittelwert-Aggregationen bei Skill-Level 3 größer ist als bei Skill-Level 2. Berücksichtigt man dies, indem man nur Studien inkludiert, die ohne Mittelwert-Aggregation auskommen, verringert sich der Unterschied zwischen beiden Gruppen deutlich: $\bar{r} = .41$ (Skill-Level 2) vs. .39 (Skill-Level 3), die $\hat{\rho}$ s betragen .63 vs. .62. Insgesamt kann aus den Ergebnissen *kein* Anstieg der Validität mit zunehmender Berufskomplexität festgestellt werden. Eher scheint das Gegenteil der Fall zu sein.

Hülsheger et al. (2007) ermitteln höhere Werte für \bar{r} und $\hat{\rho}$ für beide Skill-Level-Gruppen (Level 2: $\bar{r} = .35$, $\hat{\rho} = .52$, Level 3: $\bar{r} = .29$, $\hat{\rho} = .45$) und gehen von höheren Varianzeinschränkungen im Prädiktor aus (Level 2: $u = .68$, Level 3: $u = .67$), die sie direkt korrigieren. Diese Unterschiede entsprechen denen, die für die Gesamtanalyse der Lernleistungsstudien gefunden wurden (Kapitel 5.2). Abgesehen von den Niveau-Unterschieden ähneln sich die Befunde stark: die Gruppe mit dem geringeren Komplexitätsniveau weist höhere Validitäten auf als die Gruppe mit dem höheren Komplexitätsniveau. Unterschiede in der Varianzeinschränkung können dafür nicht verantwortlich gemacht werden.

Hülsheger et al. (2007) berechnen zusätzlich eine getrennte Analyse für das *höchste*

Komplexitätsniveau und ermitteln mit $\bar{r} = .19$ und $\hat{\rho} = .30$ noch geringere Validitäten. Allerdings basieren diese Ergebnisse auf wenigen Studien ($k = 6$). Da keine getrennte Berechnung der Varianzeinschränkung für dieses Komplexitätslevel möglich war, verwendeten sie zur Korrektur den u -Wert von Skill-Level 3 (.67).

Die Ergebnisse widersprechen damit den europäischen Befunden von Salgado et al. (2003a), die zunehmende Validitäten mit zunehmender Berufskomplexität finden. Es unterscheiden sich bereits die unkorrigierten Validitäten: bei Berufen mit geringer Komplexität beträgt sie .23 vs. .29 bei mittlerer und hoher Komplexität. Die ermittelten Varianzeinschränkungen nehmen zu ($u = .83, .67$ und $.43$ bei geringer, mittlerer und hoher Komplexität). Dieser Unterschied ist für die deutlich ansteigende operationale Validität mit zunehmender Komplexität verantwortlich ($\hat{\rho} = .36, .53, .74$).

Die amerikanischen Befunde (Hunter, 1986) weisen ebenfalls auf eine Zunahme der Validität mit zunehmender Komplexität hin, allerdings verläuft diese Zunahme nicht kontinuierlich: Die Berufe mittlerer Komplexität weisen die höchste Validität auf ($\hat{\rho} = .57$, vollständig korrigiert), die wenig komplexen Berufen liegen bei $\hat{\rho} = .54$ und die Berufe mit hoher Komplexität bei $.50$. Die unkorrigierten Validitäten sind nicht berichtet, es ist jedoch anzunehmen, dass sie die gleiche Rangfolge aufweisen, da die Artefaktkorrekturen in allen drei Gruppen die gleichen waren ($u = .60$). Zu beachten ist ferner, dass die Ergebnisse von Hunter auf direkter Varianzeinschränkung basieren und ρ unterschätzen.

Die Unterschiede zwischen den Befunden aus Deutschland einerseits und denen zu Westeuropa bzw. Amerika andererseits können evtl. auf unterschiedliche Operationalisierungen der Berufskomplexität zurück geführt werden. Die Komplexitätsoperationalisierung der Berufe auf Basis der ISCO Skill-Level ist nicht so fein wie die von Hunter und Schmidt sowie Salgado et al. verwendete Komplexitätskodierung des DOT (vgl. Kapitel 2.3.1). Die Skill-Level-Kodierung stellt auch keine direkte Komplexitätseinschätzung dar, sondern eine Einschätzung des für die Berufsausübung erforderlichen *Bildungsniveaus*.

In einigen in dieser Arbeit inkludierten Studien wird das allgemeinbildende Schulniveau der Stichprobe mitgeteilt. Dies ermöglicht ergänzend zu den Skill-Level-Analysen einen Vergleich der 7 Studien mit dem Bildungsniveau *maximal Hauptschulabschluss* mit 17 Stichproben, die (*Fach-*)*Hochschul-Reife* besitzen. Bei den Hauptschulniveau-Stichproben handelt es sich überwiegend um Personen, die Skill-Level 2 Berufe ergriffen haben, insbesondere der Berufsgruppe der Handwerker. Die Abiturniveau-Stichproben sind überwiegend der Berufsgruppe *Techniker und Gleichrangige* mit Skill-Level 3 zuzuordnen.

Die mittleren Effekte beider Gruppen unterscheiden sich nur geringfügig voneinander: $\bar{r} = .19$ für Hauptschulniveau und $.16$ für Abiturniveau, bzw. $.23$ vs. $.25$ bei Exklusion der Studien mit Mittelwert-Aggregaten. Die vollständig korrigierten $\hat{\rho}$ s sind für das Hauptschulniveau höher als für das Abiturniveau ($.36$ vs. $.29$). Dies gilt auch dann, wenn Studien mit Mittelwert-Aggregaten exkludiert werden ($.43$ vs. $.40$). Die Unterschiede sind jedoch auf Grund der geringen Studienanzahl pro Analyse statistisch nicht bedeutsam. Insgesamt betrachtet sprechen auch diese Ergebnisse eher für eine abnehmende Validität mit zunehmendem Bildungsniveau.

Wie könnte die abnehmende Validität mit zunehmender Berufskomplexität erklärt

werden, wenn man Unterschiede in der Operationalisierung der Komplexität, unterschiedliche Messfehler, etc. bei Seite lässt und davon ausgeht, dass die Validität von Intelligenz tatsächlich abnimmt? Möglicherweise liegt die Ursache in den Varianzeinschränkungen im *Kriterium*, die weder in den Vorläufer-Studien noch in dieser Arbeit korrigiert wurden. Im Hinblick auf Arbeitsleistung diskutiert dies Jäger (1986): „Mit zunehmender Komplexität und Schwierigkeit der beruflichen Anforderungen, etwa bei Führungsaufgaben, wird der mit Intelligenztests vorhersagbare Varianzanteil geringer“ (S. 280) und dadurch die zu beobachtende Korrelation ebenfalls. Dem entgegen stehen allerdings Befunde von Hunter und Schmidt (1996, S. 451), die eine zunehmende Varianz in der Arbeitsleistung bei zunehmender Berufskomplexität finden. Die Rolle, die Varianzunterschiede im Kriterium spielen, sollte in weiterführenden Studien untersucht werden.

Möglicherweise lässt sich der Einfluss der Komplexität nicht zeigen, weil in Deutschland ein anderer Moderator mitentscheidend ist: die soziale Herkunft. Aus den PISA-Studien von 2000 und 2003 geht hervor, dass der Zusammenhang zwischen dem Kompetenzerwerb Jugendlicher und den sozioökonomischen und soziokulturellen Bedingungen des Elternhauses in Deutschland besonders eng ist (Ehmke, Hohensee, Heidemeier & Prenzel, 2004). Der große Einfluss der sozialen Herkunft wirkt sich auf die Wahl der weiterführenden Schulen – insbesondere das Gymnasium – aus und kanalisiert die Berufswahl (Baumert, Cortina & Leschinsky, 2003).

5.6. Berufsgruppen

Hunter und Schmidt weisen in verschiedenen Veröffentlichungen darauf hin, dass sich die Validitäten von Intelligenz über verschiedene Berufe hinweg generalisieren lassen. Um dies zu prüfen, wurden die Studien den Berufshauptgruppen der ISCO-88-Kodierung zugeordnet. Auch dieser Moderator kann auf Grund der geringen Studienanzahl der Erfolgsbereiche Arbeitsleistung und Karriere nur für Lernleistung untersucht werden. Für 19 Lernleistungsstudien war keine Berufsgruppen-Kodierung möglich, da sie unklar oder gemischt sind. Keine Studien sind zu den Berufsgruppen 6 *Landwirtschaft, Fischerei* und 9 *Hilfsarbeitskräfte* gefunden worden. Die Berufsgruppen 1 *Leitende Berufe*, 2 *Wissenschaftler* und 5 *Dienstleistungsberufe* sind mit zu wenigen Studien vertreten, um eigene Metaanalysen dafür durchzuführen.

Die inkludierten Tätigkeiten erfordern hauptsächlich den Umgang mit Dingen bzw. Daten, nicht mit Personen. Eine Generalisierung der Befunde auf personenorientierte Tätigkeiten und Berufsgruppen (z.B. Gesundheits- und Pflegeberufe) ist nicht möglich. Personenorientierte Berufe sind i.d.R. komplex, dinge- bzw. datenorientierte Berufe verteilen sich über das gesamte Komplexitätsspektrum (Gottfredson, 2003). Geht man von zunehmender Validität von Intelligenz bei zunehmender Berufskomplexität aus, ist zu vermuten, dass der Berufserfolg in personenorientierten eher stärker mit Intelligenz zusammenhängt als in dinge- bzw. datenorientierten Berufen.

Metaanalytisch verglichen werden können die Gruppen 0 *Militär* (7 Studien), 3 *Techniker und gleichrangige* (50 Studien), 4 *Bürokräfte* (35 Studien), 7 *Handwerker* (86 Stu-

dien) sowie 8 *Maschinenbediener* (8 Studien). Die Berufsgruppen unterscheiden sich entsprechend ihrer Skill-Level wie im vorherigen Kapitel dargestellt: die Berufe, die dem Skill-Level 2 angehören (Bürokräfte, Handwerker und Maschinenbediener) weisen tendenziell höhere Validitäten auf, als die Techniker-Gruppe, die der Skill-Level-3-Gruppe entspricht: das \bar{r} beträgt für Bürokräfte .38 (95% CI = .34-.42), für Handwerker .41 (CI = .38-.44), für Maschinenbediener .36 (CI = .28-.43) sowie für Techniker .31 (CI = .25-.36). Im Vergleich dazu geringer ist das mittlere r der militärischen Stichproben: .21 (CI = .14-.28). An dieser Rangfolge ändert sich auch bei Betrachtung der $\hat{\rho}$ s nichts: Bürokräfte: .63, Handwerker: .65, Maschinenbediener: .61, Techniker: .54 und militärische Stichproben: .39. In allen Fällen liegt Transportabilität vor, d.h. die Richtung des Effektes ist positiv, die Höhe hängt auch innerhalb der einzelnen Berufsgruppen von weiteren Variablen ab.

Vom Problem der Mittelwert-Aggregate sind vor allem die Gruppen der Techniker betroffen. Werden die Studien, die Mittelwert-Aggregate erforderlich machen, exkludiert, unterscheiden sich die Validitäten der Berufsgruppen kaum: Die Validitäten von Technikern, Bürokräften, Handwerkern und Maschinenbedienern betragen $.36 < \bar{r} < .42$ bzw. $.62 < \hat{\rho} < .66$. Die einzige Berufsgruppe, die sich davon deutlich abhebt, ist die militärische (s.o.). Sie enthält keine Studien mit Mittelwert-Aggregaten. Die militärischen Studien sind jedoch zu einem weitaus größeren Teil unpubliziert (Gewicht an der bare-bone-Analyse: 87.2%) als die Studien der restlichen Berufsgruppen (5.2-11.4%), möglicherweise also weniger stark positiv von einem *publication bias* verzerrt. Dass die Kriterienvalidität in militärischen Berufen geringer ist als in zivilen, lässt sich auch an internationalen Studien zeigen: Hunter (1986) schätzen sie in militärischen Berufen auf $\hat{\rho} = .53$ und in zivilen Berufen auf .75.

5.7. Intelligenztests

Eine getrennte Betrachtung der Validität für die verschiedenen Intelligenztest-Gruppen wurde für Lernleistungs-Studien durchgeführt, auf Grund der geringen Studienanzahl nicht für die anderen Erfolgsbereiche. Die berichteten Validitäten der Lernleistungs-Studien fallen unterschiedlich hoch aus, wenn man sie getrennt für verschiedene Intelligenztests betrachtet. WIT-Studien weisen ein \bar{r} von .25 (95% CI = .20-.30) auf, Matrizentest-Studien berichten .29 (CI = .24-.33), BIS-Studien .32 (CI = .23-.40), IST-Studien .36 (CI = .31-.41), LPS/PSB-Studien .41 (CI = .32-.49) und GATB-basierte Studien .42 (CI = .39-.44). Studien, die andere Testverfahren verwenden erreichen $\bar{r} = .32$ (CI = .25-.40). Von der Exklusion von Studien, die Mittelwert-Aggregate beinhalten, profitiert vor allem der WIT ($\bar{r} = .33$, CI = .27-.38). Die beobachteten Validitäten von LPS/PSB steigen geringfügig auf .44 (CI = .34-.54), die der *anderen Verfahren* sinkt auf .26 (CI = .14-.38).

Bei Inklusion aller Studien weisen die GATB-basierten Verfahren mit .64 die höchste *operationale* Validität auf (CI = .61-.67), es folgen LPS/PSB (.63, CI = .53-.74), IST (.59, CI = .53-.65), andere Verfahren (.53, CI = .43-.63), Matrizentests (.46, CI = .38-.53), WIT (.47, CI = .35-.51) und BIS (.41, CI = .30-.52). Bei Exklusion der Studien

mit Mittelwert-Aggregaten profitieren vor allem der WIT (.53, CI = .46–.61) und LPS/PSB (.68, CI = .57–.79). Die operationalen Validitäten von IST und BIS bleiben mit .59 und .42 nahezu unverändert. Die Exklusion der Studien mit Mittelwert-Aggregaten führt hingegen bei Matrizen tests (.43) und anderen Verfahren (.45) zu geringeren operationalen Validitäten als bei Inklusion aller Studien. Dies macht deutlich, dass die Exklusion der Studien mit Mittelwert-Aggregaten nicht zwangsläufig zu höheren Validitäten führen muss. Berichten die betroffenen Studien im Vergleich zu den anderen hohe Korrelationen, wirkt sich eine Exklusion trotz unterschätzter wahrer Effekte negativ aus.

Alle Intelligenztest-Gruppen weisen Glaubwürdigkeitsintervalle über Null auf und lassen deshalb die Generalisierung der Richtung der Korrelation zu: die Zusammenhänge sind positiv. Bei der Matrizen test-Gruppe klären die Artefakte zudem 90.4% der Varianz auf. In diesem Fall kann deshalb davon ausgegangen werden, dass sich auch die Höhe der Validität verallgemeinern lässt.

Hülshager et al. (2006) berichten ebenfalls die operationalen Validitäten getrennt für die Matrizen tests CFT-3/SPM ($k = 5$, $\bar{r} = .37$, $\hat{\rho} = .54$), den WIT ($k = 14$, $\bar{r} = .32$, $\hat{\rho} = .54$), den LPS/PSB ($k = 6$, $\bar{r} = .40$, $\hat{\rho} = .52$) sowie den IST ($k = 39$, $\bar{r} = .32$, $\hat{\rho} = .48$). LPS/PSB und IST weisen *transportability* auf, die restlichen Testverfahren *cross situational consistency*. Auf Grund der geringen Studienanzahl sind die Ergebnisse zu den Matrizen tests und LPS/PSB mit Vorsicht zu interpretieren. WIT und IST weisen bei Hülshager et al. die selben unkorrigierten Korrelationen mit Lernleistung auf. Die Artefaktkorrekturen wirken sich zu Gunsten des WIT aus, der über eine höhere operationale Validität verfügt als der IST. In dieser Arbeit werden bei Inklusion aller Studien für den IST deutlich höhere Korrelationen mit Lernleistung beobachtet (.36) als beim WIT (.25). Werden nur Studien ohne Mittelwert-Aggregate betrachtet, ist die Korrelation des IST (.36) nur geringfügig höher als die des WIT (.33). Entsprechende Ergebnisse werden für die operationalen Validitäten gefunden (s.o.).

Die Ergebnisse in beiden Arbeiten sprechen dafür, dass die verschiedenen Intelligenzverfahren in Bezug auf berufliche Lernleistung valide sind. Wie bei allen Moderator-Analysen dürfen die gefundenen Unterschiede zwischen den Intelligenztests nicht überbewertet werden, selbst wenn sich die Konfidenzintervalle nicht überlappen, weil die Unterschiede nicht unbedingt den Testverfahren allein zuzuschreiben sind. Hinzu kommt, dass die Ergebnisse des IST und WIT nicht auf den aktuellen Testversionen basieren. Der aktuelle IST 2000 R (Amthauer et al., 2001) ist nicht vertreten, der aktuelle WIT-2 (Kersting et al., 2008) nur in einer Studie (Kersting et al., 2005) in einer Vorversion. Der BIS schließlich kommt in seiner Kurzfassung zum Einsatz und das in vergleichsweise wenigen Studien ($k = 9$).

Der BET erweist sich im Hinblick auf Lernleistung als prädiktiv valide. Der Einsatz vergleichsweise zeitaufwändiger Strukturverfahren (IST, WIT, BIS) scheint nicht erforderlich zu sein, wenn man lediglich an einem Indikator für g interessiert ist, um z.B. ein schnelles Screening vieler Bewerber durchzuführen. In Berufen, in denen spezifische Fähigkeiten von besonderer Bedeutung sind (z.B. figurale bei Technischen Zeichnern), können anschließende Strukturtests sinnvoll sein, die die spezifischen Fähigkeiten mit erfassen. Allerdings erlauben nicht alle Strukturtests auch eine Auswertung auf spezifischer

Ebene. Dies gilt z.B. für den IST70 und den BIS 4 in der Kurzversion.

Das Gewicht der drei Inhaltsfacetten an den Metaanalysen ist insgesamt ungefähr gleich groß: Der Anteil der verbalen Facetten liegt etwas über den numerischen und figuralen Inhalten. Die Operationen unterscheiden sich deutlich: weitaus den größten Einfluss hat Verarbeitungskapazität. Dies entspricht den Anforderungen des BIS-Modells, in dem die Verarbeitungskapazität als die bedeutendste Operation angesehen wird (Jäger et al., 1997).

5.8. Generalisierbarkeit

Es wurde gezeigt, dass Intelligenz (gemessen mit verschiedenen Testverfahren) *generell positiv* mit Berufserfolg (gemessen mit verschiedenen Erfolgsindikatoren) korreliert. Dies kann nach F. L. Schmidt und Hunter (1998a) als Beleg für die Validität von Intelligenz gewertet werden. Diese Interpretation der Befunde wird von Borsboom, Mellenbergh und Heerden (2004) als zu weitgehend angesehen, da die vorgenommenen Korrekturen nicht die mangelnde Validität korrigieren. D.h. die korrigierten Werte entsprechen nicht den latenten Konstrukten (*construct scores*), sondern stellen lediglich wahre Werte (*true scores*) im Sinne der klassischen Testtheorie dar. Beides kann nicht gleichgesetzt werden (Borsboom & Mellenbergh, 2002). Borsboom et al. (2004) schlagen vor, die korrigierten Werte nicht als Nachweis der Validität, sondern der *Qualität* zu interpretieren. In diesem Sinne wurde gezeigt, dass Intelligenz eine hohe (prädiktive) Qualität in Bezug auf Berufserfolg aufweist.

Aber auch diese Feststellung gilt nicht uneingeschränkt. Es stellt sich die Frage, worauf sie generalisiert werden kann bzw. worauf nicht. In Anlehnung an Matt (2003) können fünf Generalisierungsbereiche unterschieden werden: die einzelnen Personen, die untersucht wurden (*units*), die Intelligenztests (bei Matt, der die Generalisierung von Experimenten betrachtet, *treatment*), die Kriterien (*outcomes*), die sozialen und physikalischen Umgebungen (*settings*) sowie die Zeit (*time*).

Zu *soziodemographischen Daten* und weiteren möglicherweise relevanten Variablen der Personen (z.B. die Berufserfahrung) sind in den Primärstudien häufig keine Angaben gemacht. Wie repräsentativ die verschiedenen Stichproben für die jeweiligen Populationen sind und wie repräsentativ die Gesamtstichprobe für die Population der erwerbstätigen Bevölkerung ist, kann deshalb nicht geklärt werden. Die vorliegenden Informationen zum Alter und Geschlecht der Studienteilnehmer deuten darauf hin, dass die Stichproben insgesamt im Vergleich mit der Population der Erwerbstätigen jünger sind und aus mehr Männern bestehen.

Das mittlere Alter der Personen, die ihre Ausbildung bereits abgeschlossen haben, liegt je nach Kriterienbereich zwischen 25.4 Jahren (Weiterbildungsleistung) und 31.4 Jahren (berufliches Vorankommen). In der Population der Erwerbstätigen beträgt es laut den Mikrozensus von 1998 bis 2006 ca. 40 Jahre (Statistisches Bundesamt, 2008). Dies ist relevant, da Berufserfolgs-Maße mit dem Alter der Beurteilten zusammenhängen können. Für die Einkommensentwicklung und die Anzahl der Beförderungen haben dies Ng et al. (2005) nachgewiesen. Auch auf Leistungsbeurteilungen wirkt sich das Alter

aus: laut Waldman und Avolio (1986) beträgt die Korrelation zwischen dem Alter der Beurteilten und Abwärtsbeurteilungen $\bar{r} = -.14$ ($k = 18$, $N = 3\,660$).

Der Anteil der Frauen in den Stichproben ist geringer als das bei Betrachtung der Population zu erwarten wäre. Der Frauenanteil an der erwerbstätigen Bevölkerung in Deutschland betrug 1957 36.6% und stieg seit dem kontinuierlich an auf 44.9% im Jahr 2006 (Statistisches Bundesamt, 2008). In den vorliegenden Daten liegt er zwischen 22.8% bei Lernleistungsstudien und 46.6% bei Studien zum beruflichen Vorankommen, wenn man die Studien unberücksichtigt lässt, die keine Angaben zum Geschlecht machen. Einen Einfluss des Geschlechts auf subjektive Leistungsbeurteilungen ist nachweisbar, wenn nur männliche Beurteiler die Einschätzung vornehmen sowie bei geschlechtsstereotypen Beurteilungsdimensionen (Bowen, Swim & Jacobs, 2000).

Die *Allgemeine Intelligenz* wird in der Metaanalyse mit unterschiedlichen Testverfahren operationalisiert. Die Operationalisierung der allgemeinen kognitiven Fähigkeiten ist mehr als die genetische Komponente von Intelligenz. Es ist die „developed ability at the time the test is administered“ (F. L. Schmidt, 2002, S. 188).

Im Vergleich mit den bereits durchgeführten Metaanalysen (Hülsheger et al., 2007; Salgado & Anderson, 2003) wurden zusätzlich die GATB basierten Verfahren berücksichtigt. Trotzdem bleibt die Einschränkung gültig, auf die Hülsheger (2006) hinweist: „eine Reihe von Intelligenztests (z.B. MIT, Büscher et al., 1986; BOMAT, Hossiep, Turck & Hassella, 2001; HAWIE, Tewes, 1991) ... [wurden] noch nicht hinsichtlich ihrer Kriteriumsvalidität überprüft.“ (S. 31). Die Kriteriumsvalidität dürfte allerdings auch bei nicht berücksichtigten Verfahren gegeben sein, wenn sie kunstgerecht konzipiert sind. Diese Vermutung liegt nahe, da der positive Zusammenhang mit Berufserfolg über alle inkludierten Testverfahren hinweg nachgewiesen werden konnte und die Ergebnisse verschiedener Intelligenztests hoch miteinander korrelieren (Jensen, 1984). Die eingesetzten Intelligenztests sind heterogen: es werden Matrizen-Tests, Strukturtests und Omnibusverfahren verwendet. Die Validität lässt sich selbst für Kriterienbereiche nachweisen, die bei der Testentwicklung nicht als relevant erachtet wurden. Kuncel et al. (2004) zeigt z.B., dass ein Test, der für schulische Settings entwickelt wurde (der Miller Analogies Test, MAT; Miller, 1960), auch in Bezug auf Arbeitsleistung und Berufserfolg prädiktiv valide ist.

Auf welche *Kriterien* generalisiert werden kann, ist in den Kapiteln 5.2.1 für Lernleistung, 5.3 für Arbeitsleistung und 5.4 für Karriereerfolg beschrieben. Die korrigierten mittleren Effekte pro Erfolgsbereich sind auch von der Validität der Kriterien beeinflusst. Zum Beispiel ist die Validität von Abwärtsbeurteilungen bei kurzer Dauer der Zusammenarbeit mit den Beurteilten vergleichsweise gering und steigt bis zu einer Zusammenarbeits-Dauer von zwei Jahren kontinuierlich an (Rothstein, 1990, vgl. auch Kapitel 3.6.2). Schuler (2004a) berichtet vergleichbare Ergebnisse und schlussfolgert: „Offensichtlich waren die Beurteiler erst bei langfristiger Kenntnis ihrer Mitarbeiter in der Lage, deren Leistungen richtig einzuschätzen. Daraus wird ersichtlich, dass unzulängliche Beurteilungsdaten nicht nur für sich genommen Schaden anrichten, sondern mitunter auch zu Fehlschlüssen über validierte Auswahlverfahren führen“ (S. 51). Eine größere Anzahl von Primärstudien mit unterschiedlichen Kriterien würde es erlauben,

ihre differentiellen Effekte zu quantifizieren.

Exkludiert wurde berufliche Zufriedenheit. Es kann davon ausgegangen werden, dass Zufriedenheitsmaße mit den inkludierten Erfolgsmaßen positiv korrelieren, da a) gemäß der Attributionstheorie Erfolg eher internal attribuiert wird, was zu einer Steigerung positiver Selbstwahrnehmungen führen sollte und b) gemäß der Theorie sozialer Vergleichsprozesse Leistungsindikatoren den sozialen Vergleichen zugrunde gelegt werden, was bei positivem Ergebnis ebenfalls zu einer Steigerung positiver Selbstwahrnehmung führt (Ng et al., 2005). In ihrer Metaanalyse konnten Ng et al. (2005) den Zusammenhang von Karrierezufriedenheit mit Einkommen ($\hat{\rho} = .30$, $k = 23$, $N = 10\,903$) als auch mit Beförderungen ($\hat{\rho} = .22$, $k = 12$, $N = 8\,701$) nachweisen. Die Korrelation zwischen Arbeitszufriedenheit und Arbeitsleistung beziffern Judge und Bono (2001) mit $\hat{\rho} = .30$ ($k = 312$, $N = 54\,471$). Bowling (2007) widerlegen allerdings die Annahme, dass Zufriedenheit und Leistung *kausal* miteinander zusammenhängen.

Der Bereich der *Settings* betrifft insbesondere die Berufsgruppen, die zum Teil mit sehr wenig Gewicht (soziale Berufe) oder gar nicht (künstlerische Berufe) vertreten sind. Wie dargestellt wurde, sind vor allem dinge- bzw. datenorientierte Tätigkeiten abgedeckt, personenorientierte fehlen.

Über welchen *Zeithorizont* hinweg können die Befunde generalisiert werden? Hülshager et al. (2007) betrachten das Veröffentlichungs- bzw. Durchführungsjahr der Studien als Moderator und stellen fest, dass die Kriterienvalidität von Intelligenz der älteren Studien (vor 1990, Lernleistung: $\bar{r} = .39$, $\hat{\rho} = .56$, Arbeitsleistung: $\bar{r} = .42$, $\hat{\rho} = .74$) höher ist als der jüngeren Studien (nach 1989, $\bar{r} = .24$, $\hat{\rho} = .38$, Arbeitsleistung: $\bar{r} = .25$, $\hat{\rho} = .34$). Dies wurde in der vorliegenden Studie nicht untersucht. Dennoch lässt sich am Ergebnis der kumulativen Metaanalyse zur Lernleistung (Abbildung 4.3) nachvollziehen, dass die Validität tendenziell abnimmt bei zusätzlicher Inklusion jüngerer Studien. Werden nur Studien vor 1960 inkludiert resultieren sogar deutlich höhere mittlere Effekte. Eine mögliche Ursache dafür ist eine geänderte Notengebung, die Ende der 60er Jahre eingeführt wurde: Der Bezugsstandard für Benotungen war bis dahin der Durchschnitt (sozial, relativ), dann wurde er an den Anforderungen (sachlich, absolut) ausgerichtet (Tent, 2001). Dennoch erstaunt die Abnahme der Kriterienvalidität von Intelligenz bei Inklusion jüngerer Studien. Die Berufsumwelten und Tätigkeitsanforderungen werden komplexer (Bertua et al., 2005), viele einfache Berufe fallen durch Automatisierung und Computerisierung weg (Hunter & Schmidt, 1996) und die Bedeutung berufsbezogenen Wissens nimmt zu (Campion in Morgeson et al., 2007). Die Bedeutung kognitiver Fähigkeiten am Arbeitsplatz sollte deshalb eher zu- statt abgenommen haben und in Zukunft weiter an Bedeutung gewinnen (Anderson, Lievens, Dam & Ryan, 2004).

5.9. Weiterer Forschungsbedarf

Die vorliegende Analyse hat deutlich gemacht, zu welchen Intelligenztests, Berufsgruppen und Kriterien wenige oder keine Validierungsstudien vorhanden sind. Um die Ergebnisse besser generalisieren und relevante Moderatoren entdecken zu können, sind generell mehr Primärstudien erforderlich und mehr Information zu den Primärstudien. Die 70

Studien aus dem BET-Manual enthalten z.B. außer den Korrelationen, Stichprobengrößen und Berufsgruppen keine weiteren Informationen.

Die Qualität metaanalytischer Ergebnisse hängt maßgeblich von den inkludierten Primärstudien ab (F. L. Schmidt et al., 1980), insbesondere von der Vollständigkeit ihrer Berichterstattung. Die integrierten Studien enthalten häufig keine Angaben zur Höhe der Artefakte, z.B. zur Varianzeinschränkung in Prädiktor und Kriterium. In einigen sind die insignifikanten Korrelationen nicht berichtet. Die fehlenden Werte machen Schätzungen nötig und dürften zu einem Teil für die vergleichsweise geringe Varianzaufklärung der Metaanalysen ursächlich sein. Die eingesetzten Messverfahren sind häufig nicht detailliert genug beschrieben um in Moderatoranalysen getrennt berücksichtigt werden zu können. Dies gilt besonders für die Kriterien. Bei Intelligenztests fehlen häufig aussagekräftige bzw. aktuelle Normen (z.B. für den BIS), was sich negativ auf Reliabilität und Validität der Primärstudienresultate auswirken kann (vgl. Kersting, 2007).

Zur Korrektur der Artefakte wurde auf Artefaktverteilungen zurückgegriffen, die nicht auf deutschen Stichproben basieren (Reliabilität subjektiver Leistungsbeurteilungen; Rothstein, 1990), älteren Datums sind (Reliabilität von Berufsschulnoten; Baron-Boldt et al., 1988) oder auf wenigen Angaben basieren (Varianzeinschränkung im Prädiktor). Andere Artefakte, z.B. die Varianzeinschränkung im Kriterium, konnten auf Grund der wenigen Angaben in den Studien nicht korrigiert werden. Da Reliabilitäten von Intelligenztests und Leistungskriterien kulturellen Einflüssen unterliegen können (Kapitel 2.3), würden Metaanalysen, die die Höhe der Artefakte in Deutschland quantifizieren, eine genauere Korrektur ermöglichen.

Geringe Varianzaufklärungen von unter 25% wurden auch in anderen Metaanalysen gefunden (z.B. Hülshager et al., 2007; N. Schmitt et al., 1984). Schmitt et al. sehen eine mögliche Ursache dafür in der hohen Variabilität der inkludierten Studien im Hinblick auf die verwendeten Prädiktoren, Datenerhebungs- und Auswertungsmethoden sowie die organisationalen Settings. Die Prüfung dieser Moderatoreinflüsse ist nur möglich, wenn eine große Zahl heterogener Studien vorhanden ist.

Ferner sollte die Frage geklärt werden, welche Kriterienvalidität alternative Personalauswahl-Verfahren in Deutschland aufweisen und wie hoch ihre inkrementelle Validität über Allgemeine Intelligenz hinaus ist. Erst dadurch wird eine Entscheidung darüber möglich, welches oder welche Verfahren angewendet werden können, um die höchste Kriterienvalidität zu erreichen.

Schließlich ist die metaanalytische Methode weiterzuentwickeln. Zur Durchführung der Metaanalysen liegen verschiedene Modelle vor (vgl. Kapitel 3.8), deren Anwendung nur zu bedingt miteinander vergleichbaren Ergebnissen führt. Insbesondere die Durchführung von Artefaktkorrekturen ist nicht unumstritten. Auch darüber, welche Kennwerte den Artefaktkorrekturen zu Grunde gelegt werden sollen, herrschen verschiedene Meinungen vor (vgl. Kapitel 3.6). Zur Kummulierung abhängiger Ergebnisse ist eine Entscheidung zwischen der Mittelwert-Bildung und der Bildung linearer Verbünde zu treffen, wobei erstere unterschiedliche Varianzanteile der einzelnen Prädiktoren bzw. Kriterien als Fehlervarianz betrachten, letztere als unterschiedliche relevante Aspekte am „wahren“ Prädiktor bzw. Kriterium. Tatsächlich dürften i.d.R. Kombinationen aus beidem vorliegen.

5.10. Bedeutung für die Personalauswahl

Der Nutzen eines Personalauswahl-Verfahrens hängt maßgeblich von der Höhe der Validität ab (Görlich & Schuler, 2006). Diese zu kennen ist deshalb von hoher praktischer Relevanz: „For purposes of practical validation, it does not matter why cognitive ability correlates with performance. The only practical question for selection is just how high is the validity?“ (Hunter, 1986, S. 354). Diese Frage kann für Lern- und Arbeitsleistung nun mit größerer Genauigkeit beantwortet werden. Es konnten etwas mehr Studien pro 1 Million Einwohner inkludiert werden (228 Studien / ca. 82 Millionen Einwohner = 2.8) als in der Metaanalyse von Hunter und Hunter (1984) in den USA, die allerdings schon 1984 durchgeführt wurde (730 Studien / 292 Millionen Einwohner = 2.5). In Großbritannien hingegen wurden mit 4.8 deutlich mehr Studien pro 1 Million Einwohner durchgeführt, wenn man die Metaanalyse von Bertua et al. (2005) als Maßstab nimmt (293 Studien / 60 Millionen Einwohner; Die Einwohnerzahlen stammen aus Hülsheger, 2006). Erstmals lässt sich die Frage nach der Validität von Intelligenz für Karriereerfolg in Deutschland beantworten.

F. L. Schmidt und Hunter (1998a) konnten zeigen, dass die Validität von 16 anderen Auswahlinstrumenten im Hinblick auf Arbeitsleistung nicht an die von GMA-Tests heranreicht, mit Ausnahme von Arbeitsproben und strukturierten Interviews. Deren Einsatz ist allerdings aufwändiger und teurer.

Häufig werden in der Personalauswahl Assessment-Center (ACs) als Personalauswahl-Verfahren eingesetzt (Schuler, Hell, Trapmann, Schaar & Boramir, 2007). Die Validität dieser Verfahren wird von F. L. Schmidt und Hunter (1998a) mit .36 für Arbeitsleistung angegeben. Der Wert geht zurück auf eine Metaanalyse von Gaugler, Rosenthal, Thornton und Bentson (1987). Hardison und Sackett (2007) erweiterten diese Metaanalyse um neuere Studien und ermitteln als AC-Validität $\hat{\rho} = .28$ in Bezug auf Arbeitsleistung. Für Trainingserfolg ermitteln Gaugler et al. (1987) sowie Hardison und Sackett (2007) $\hat{\rho} = .35$, für Karriereerfolg $\hat{\rho} = .36$. (Gaugler et al.) bzw. $\bar{r} = .27$ (Hardison & Sackett).

Die deutlich höheren Kosten von ACs lassen sich dadurch nicht rechtfertigen. Ein weiteres Problem anderer Personalauswahl-Instrumente, wie ACs, der Überprüfung von Referenzen oder der Messung biographischer Daten, sind unbekannt Konstruktzusammensetzungen (F. L. Schmidt & Hunter, 1998a, S. 37). Generell gilt, dass andere Personalauswahl-Verfahren eine umso höhere Validität aufweisen, je stärker sie die Allgemeine Intelligenz miterfassen (F. L. Schmidt, 2002, S. 199).

Obwohl somit bereits 1998 eine – im Vergleich mit anderen Verfahren, z.B. ACs – hohe prädiktive Validität von Intelligenz zumindest in den USA nachgewiesen wurde, werden sie in Deutschland bis heute weniger häufig eingesetzt als ACs: Schuler et al. (2007) haben 125 Unternehmen aus Deutschland danach gefragt, welche Auswahlverfahren sie nutzen, und diese Daten mit früheren Angaben (Schuler, Frier & Kauffmann, 1993; Schulz, Schuler & Stehle, 1985) verglichen. Von den 125 Unternehmen setzen 30.4% Intelligenztests ein und 57.6% ACs. Dies ist im Vergleich mit dem Ergebnis der Umfrage von 1993 ein Verwendungsrückgang bei Intelligenztests um 3.6 Prozentpunkte und eine Verwendungszunahme bei ACs um 18.6 Prozentpunkte. Die Ursache für diese Entwicklung liegt möglicherweise darin, dass Intelligenztests im Vergleich mit ACs von den 125

befragten Organisationen a) als weniger valide angesehen werden (!) und b) die Akzeptanz, die ihnen von Seiten der Bewerber entgegengebracht wird, geringer eingeschätzt wird (Schuler et al., 2007).

Auch im internationalen Vergleich werden in Deutschland Intelligenztests selten eingesetzt. Dies zeigen Ryan et al. (1999) anhand einer Befragung von 955 Organisationen in 18 Staaten: Intelligenztests werden lediglich in Italien und Hongkong seltener eingesetzt als in Deutschland, dagegen häufiger in den USA, Frankreich, Australien, Griechenland, Kanada, Irland, Singapur, Schweden, Großbritannien, Südafrika, Portugal, Neuseeland, Spanien, den Niederlanden und Belgien (genannt in der Reihenfolge der relativen Einsatzhäufigkeit).

Die Akzeptanz, die einem Verfahren von Seiten der Bewerber und der Organisationsangehörigen entgegengebracht wird, ist ein wesentliches Kriterium, dem Auswahlverfahren – neben der Validität und den Kosten – genügen sollen (Schuler & Höft, 2006). Für die Akzeptanz ist ausschlaggebend, dass ein Verfahren als fair empfunden wird. Bei Intelligenztests ist den Bewerbern und Organisationsangehörigen häufig nicht klar, welche Beziehung zwischen den Testitems und den Aufgaben am Arbeitsplatz bestehen. Da Intelligenz ein stabiles Persönlichkeitsmerkmal darstellt, ist auch eine persönlichkeitsförderliche Rückmeldung schwieriger als beim Einsatz von simulationsorientierten Selektionsverfahren.

Um die Akzeptanz von Intelligenztests zu steigern sind deshalb Verfahren mit höherem Anforderungsbezug zu entwickeln. Der Anforderungsbezug ist ein zentrales Merkmal zur Beurteilung der Qualität von Eignungsuntersuchungen gemäß der DIN 33430 (Deutsches Institut für Normung e.V., 2002). Deutschsprachige Verfahren zur Erfassung kognitiver Fähigkeiten, die ausdrücklich mit diesem Ziel entwickelt worden sind, sind die AZUBI-Tests für Büro- und kaufmännische Tätigkeiten (AZUBI-BK; Schuler & Klingner, 2005), für technische und handwerkliche Tätigkeiten (AZUBI-TH; Görlich & Schuler, 2007) sowie die neue Version des WIT (WIT-2; Kersting et al., 2008). Auch die DGP-Version des BIS (BIS-r-DGP; Deutsche Gesellschaft für Personalwesen e.V., 1996) wurde in dieser Hinsicht optimiert, indem simulationsorientierte Aufgaben ergänzt wurden (Kersting & Beauducel, 2004). Validierungsstudien zu diesen Tests wurden auch in der vorliegenden Arbeit berücksichtigt (Görlich & Schuler, 2004; Kersting, 2003; Kersting et al., 2005; Schuler & Klingner, 2005).

Dass der Einsatz klassischer Intelligenztests auf Vorbehalte stößt, wird auch von F. L. Schmidt (2002) gesehen und ein Stück weit geteilt: „Would it be better if GCA [general cognitive ability] were less important than it is in determining job performance? In my opinion, it would. ... This outcome seems much fairer and more democratic to me“ (S. 201-202). Trotzdem warnen Hunter und Schmidt (1996) davor, politisch gewollt auf den Einsatz von Intelligenztests in der Personalauswahl zu verzichten. Dieser Verzicht wirke sich langfristig in nachlassender Effizienz aus, was besonders in exportorientierten Branchen, die stärkerem Wettbewerb auf dem Weltmarkt ausgesetzt sind, die Konkurrenzfähigkeit reduziere. Folgt man dieser Argumentation, sollten Intelligenztests insbesondere in Deutschland, das eine stark exportorientierte Wirtschaft aufweist, verstärkt zur Personalauswahl eingesetzt werden.

Literaturverzeichnis

In die Metaanalyse inkludierte Quellen sind mit einem Sternchen (*) markiert.

- Abele-Brehm, A. E. & Stief, M. (2004). Die Prognose des Berufserfolgs von Hochschulabsolventinnen und -absolventen. *Zeitschrift für Arbeits- und Organisationspsychologie*, 48, 1-13.
- Ackerman, P. L. (1994). Intelligence, attention, and learning: Maximal and typical performance. In D. K. Ditterman (Hrsg.), *Current topics in human intelligence: Theories of intelligence* (Bd. 4, S. 1-27). Norwood, NJ: Ablex.
- Ackerman, P. L. & Humphreys, L. G. (1991). Individual differences theory in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Hrsg.), *Handbook of industrial and organizational psychology* (Bd. 1, S. 223-282). Palo Alto, CA: Consulting Psychologists Press.
- *Althoff, K. (1968). *Untersuchungen zum Einfluss von Berufsausbildungen und Berufstätigkeiten auf Testergebnisse und Eignungsdispositionen*. Unveröffentlichte Dissertation, Universität Gießen.
- *Althoff, K. (1974). Kurzbericht über eine Bewährungskontrolle der Eignungsprognosen für Bewerber der allgemeinen Verwaltung in Nordrhein-Westfalen. *DGP-Informationen*, 13, 12-21.
- *Althoff, K. (1975). Die Vorhersage des Ausbildungserfolgs von Aufstiegsbeamten der Kriminalpolizei. *DGP-Informationen*, 14, 13-23.
- *Althoff, K. (1977). Zusammenhänge zwischen den Ergebnissen von Eignungstests und beruflicher Bewährung – dargestellt am Beispiel der Auswahl von Bewerbern für den gehobenen und höheren Polizeivollzugsdienst. *Schriftenreihe der Polizeiführungsakademie*, 1, 6-27.
- *Althoff, K. (1978). [Unbekannt. Zitiert nach Jäger & Althoff, 1994, S. 37].
- Amelang, M. & Bartussek, D. (1997). *Differentielle Psychologie und Persönlichkeitsforschung* (4. überarb. und erw. Aufl.). Stuttgart: Kohlhammer.
- *Amelang, M. & Steinmayr, R. (2006a). Is there a validity increment for tests of emotional intelligence in explaining the variance of performance criteria? *Intelligence*, 34, 459-468.
- *Amelang, M. & Steinmayr, R. (2006b). [Is there a validity increment for tests of emotional intelligence in explaining the variance of performance?]. Unveröffentlichte Daten. Universität Heidelberg.
- Amthauer, R. (1953a). *Intelligenz-Struktur-Test*. Göttingen: Hogrefe.
- Amthauer, R. (1953b). Intelligenz und Beruf. Ergebnisse eines neuen Verfahrens zur Bestimmung der Intelligenz. *Zeitschrift für experimentelle und angewandte Psychologie*, 1, 102-144.
- Amthauer, R. (1955). *Intelligenz-Struktur-Test* (2. Aufl.). Göttingen: Hogrefe.

- *Amthauer, R. (1973). *I-S-T 70: Intelligenz-Struktur-Test* (3. erw. Aufl.). Göttingen: Hogrefe.
- Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (1999). *I-S-T 2000. Intelligenz-Struktur-Test 2000*. Göttingen: Hogrefe.
- Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (2001). *I-S-T 2000 R. Intelligenz-Struktur-Test 2000 R*. Göttingen: Hogrefe.
- Anderson, N., Lievens, F., Dam, K. van & Ryan, A. M. (2004). Future perspectives on employee selection: Key directions for future research and practice. *Applied Psychology: An International Review*, 53, 487-501.
- *Awiszus, M. (1967). Über den Vorhersagewert des Intelligenzstrukturtests (I-S-T) für die Leistung in einer Berufsaufbauschule. *Der Schulpsychologe*, 14, 50-61.
- *Balck, F., Leins, B. & Schröder, U. (1979). Gruppenunterschiede, Reliabilitäten und Validitäten des Berufseignungstests (BET) von Schmale und Schmidtke. *Psychologie und Praxis*, 23, 153-164.
- Baron-Boldt, J., Funke, U. & Schuler, H. (1989). Prognostische Validität von Schulnoten. Eine Metaanalyse der Prognose des Studien- und Ausbildungserfolgs. In R. Jäger, R. Horn & K. Ingenkamp (Hrsg.), *Tests und Trends 7* (S. 11-39). Weinheim: Beltz.
- Baron-Boldt, J., Schuler, H. & Funke, U. (1988). Prädiktive Validität von Schulabschlussnoten: Eine Metaanalyse. *Zeitschrift für Pädagogische Psychologie*, 2, 79-90.
- Barrett, G. V., Polomsky, M. D. & McDaniel, M. A. (1999). Selection tests for fire-fighters: A comprehensive review and meta-analysis. *Journal of Business and Psychology*, 13, 507-513.
- *Barthel, E., Fünfgelt, V. & Schuler, H. (1982). *Modellversuch zur Erschließung gewerblich-technischer Ausbildungsberufe für Mädchen. 2. Zwischenbericht der wissenschaftlichen Begleitung in Augsburg*. Unveröffentlichtes Manuskript, Universität Erlangen-Nürnberg.
- Baumert, J., Cortina, K. S. & Leschinsky, A. (2003). Grundlegende Entwicklungen und Strukturprobleme im allgemein bildenden Schulwesen. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland* (S. 52-147). Hamburg: Rowohlt.
- Beauducel, A. & Kersting, M. (2002). Fluid and crystallized intelligence and the Berlin model of intelligence structure (BIS). *European Journal of Psychological Assessment*, 18, 97-112.
- Bennett, G. K. & Fry. (1969). *Test of mechanical comprehension*. New York: Psychological Corporation.
- Bennett, G. K., Seashore, H. G. & Wesman, A. G. (1952). *Differential aptitude tests (DAT)*. New York: Psychological Corporation.
- Bennett, G. K., Seashore, H. G. & Wesman, A. G. (1956). The differential aptitude tests: An overview. *Personnel and Guidance Journal*, 35, 81-91.
- Bertua, C., Anderson, N. & Salgado, J. F. (2005). The predictive validity of cognitive ability tests: A UK meta-analysis. *Journal of Occupational and Organizational*

- Psychology*, 78, 387-409.
- *Biernat, J. & Scholz, B. (1995). Eine Bewährungskontrolle des DGP-Auswahlverfahrens für den mittleren feuerwehrtechnischen Dienst. *DGP-Informationen*, 54, 21-30.
- Binet, A. & Simon, T. (1905). Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux [Neue Methoden zur Diagnose des intellektuellen Niveaus von Anormalen]. *Année Psychologique*, 11, 191-336.
- *Blickle, G., Kramer, J. & Mierke, J. (2005). [Politische Fertigkeiten, Intelligenz und Arbeitsleistung]. Unveröffentlichte Daten.
- *Blickle, G., Witzki, A. & Schneider, P. B. (2006). [Längsschnittstudie zum selbstinitiierten Mentoring in Organisationen]. Unveröffentlichte Daten aus dem DFG-Projekt AZBL62.
- Bobko, P., Roth, P. L. & Potosky, D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology*, 52, 561-589.
- Bommer, W. H., Johnson, J. L., Rich, G. A., Podsakoff, P. M. & MacKenzie, S. B. (1995). On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. *Personnel Psychology*, 48, 587-605.
- Borenstein, M. (2005). Software for publication bias. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Hrsg.), *Publication bias in meta analysis: prevention, assessment and adjustments* (S. 193-220). Chichester: Wiley.
- Borman, W., Penner, L. A., Allen, T. D. & Motowidlo, S. J. (2001). Personality predictors of citizenship performance. *International Journal of Selection and Assessment*, 9, 52-69.
- Borsboom, D. & Mellenbergh, G. J. (2002). True scores, latent variables and constructs: A comment on Schmidt and Hunter. *Intelligence*, 30, 505-514.
- Borsboom, D., Mellenbergh, G. J. & Heerden, J. van. (2004). The concept of validity. *Psychological Review*, 111, 1061-1071.
- Bortz, J. (2005). *Statistik für Sozialwissenschaftler* (6. vollst. überarb. und aktual. Aufl.). Berlin: Springer.
- Bowen, C.-C., Swim, J. K. & Jacobs, R. R. (2000). Evaluating gender biases on actual job performance of real people: A meta-analysis. *Journal of Applied Social Psychology*, 30, 2194-2215.
- Bowling, N. A. (2007). Is the job satisfaction-job performance relationship spurious? A meta-analytic examination. *Journal of Vocational Behavior*, 71, 167-185.
- Brambring, M. (1983). Spezielle Eignungsdiagnostik. In K. J. Groffmann & L. Michel (Hrsg.), *Grundlagen psychologische Diagnostik* (Bd. 2, S. 414-481). Göttingen: Hogrefe.
- Brand, C. (1987). The importance of general intelligence. In S. Modgil & C. Modgil (Hrsg.), *Arthur Jensen: Consensus and controversy* (S. 251-265). New York: Falmer.
- *Brandstätter, H. (1970). *Leistungsprognose und Erfolgskontrolle*. Bern: Huber.
- Bredenkamp, J. (1998). *Lernen, Erinnern, Vergessen*. München: C. H. Beck.
- *Bretz, E. & Oldendörp, H. (1992). Bewährungskontrolle: Vorhersage des Ausbildungs-

- erfolges im Angestelltenlehrgang I. *DGP-Informationen*, 51, 75-84.
- Brickenkamp, R. (2002a). *Aufmerksamkeits-Belastungs-Test d2* (9. vollst. überarb. und erw. Aufl.). Göttingen: Hogrefe.
- Brickenkamp, R. (Hrsg.). (2002b). *Handbuch psychologischer und pädagogischer Tests* (3. Aufl., Bd. 1). Göttingen: Hogrefe.
- Brocke, B. & Beauducel, A. (2001). Intelligenz als Konstrukt. In E. Stern & J. Guthke (Hrsg.), *Perspektiven der Intelligenzforschung* (S. 13-42). Lengerich: Pabst.
- Brocke, B., Beauducel, A. & Tasche, K. (1998). Der Intelligenz-Struktur-Test: Analysen zur theoretischen Grundlage und technischen Güte. *Diagnostica*, 44, 84-99.
- Brody, N. (2000). Theories and measurements of intelligence. In R. J. Sternberg (Hrsg.), *Handbook of intelligence* (S. 16-33). Cambridge University Press.
- Bruggemann, A., Großkurth, P. & Ulrich, E. (1975). *Arbeitszufriedenheit*. Bern: Huber.
- Bundesagentur für Arbeit. (2008a). *Ausbildungswege*. Zugriff am 30.10.2008 unter http://www.arbeitsagentur.de/nm_26014/Navigation/zentral/Buerger/Ausbildung/Ausbildungswege/Wege-Nav.html
- Bundesagentur für Arbeit. (2008b). *BERUFENET*. Zugriff am 17.10.2008 unter <http://www.berufenet.de>
- Bundesanstalt für Arbeit. (1982). *EUB, Handanweisung für die Beratungsfachkräfte der Berufsberatung*. Nürnberg: Bundesanstalt für Arbeit.
- Bundesministerium für Bildung und Forschung. (2005). *Berufsbildungsgesetz (BBiG)*. Berlin: Bundesministerium für Bildung und Forschung.
- Burchard, U. (2000). *Managerkarrieren*. Frankfurt a. M.: Peter Lang.
- Carroll, J. B. (1982). The measurement of intelligence. In R. J. Sternberg (Hrsg.), *Handbook of human intelligence* (S. 29-120). New York: Academic Press.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.
- Cattell, R. B. (1957). *Personality and motivation structure and measurement*. New York: Yonkers-on-Hudson.
- Cattell, R. B., Weiß, R. H. & Osterland, J. (1997). *Grundintelligenztest Skala 1 (CFT 1)* (5. rev. Aufl.). Göttingen: Hogrefe.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Colquitt, J. A., LePine, J. A. & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology*, 85, 678-707.
- Conway, J. M. & Huffcutt, A. I. (1997). Psychometric properties of multisource performance ratings: A meta-analysis of subordinate, supervisor, peer, and self-ratings. *Human Performance*, 10, 331-360.
- Cooper, H. M. (1998). *Synthesizing research: A guide for literature reviews* (3. Aufl.). Thousand Oaks: Sage.
- Cronbach, L. J. (1947). Test „reliability“: Its meaning and determination. *Psychometrika*, 12, 1-16.
- *Dahle, K.-P. & Erdmann, K. (2001). [Aus der Berliner CRIME-Studie]. Unveröffentlichte Daten.

- *Damitz, M., Eißfeldt, H., Grasshoff, D., Lorenz, B., Pecena, Y. & Schwert, T. (2000). *Validierung des DLR-Auswahlverfahrens für Nachwuchsfluglotsen der DFS Deutsche Flugsicherung GmbH: Ergebnisse des Projektes Qualitätssicherung*. Hamburg: Deutsches Zentrum für Luft- und Raumfahrt e.V.
- Daniels, J. C. (1962a). *Figure reasoning test - Instruction manual*. London: Lockwood.
- Daniels, J. C. (1962b). *FRT - Figure reasoning test*. London: Lockwood.
- Daniels, J. C. (2004). *FRT/FRT-J - Figure reasoning test*. Göttingen: Hogrefe.
- *Derner, N. (1983). Der Vorhersagewert verschiedener Eingangsvariablen für den Ausbildungserfolg von Metallfacharbeitern. *Lernzielorientierter Unterricht*, 4, 16-26.
- DerSimonian, R. & Laird, N. M. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7, 177-188.
- *Deter, B. & Herbig, V. (1992). *Zusammenhänge zwischen standardisierten Ausbildungstests und den betrieblichen Beurteilungen angehender Chemikanten im zweiten Lehrjahr*. Unveröffentlichtes Manuskript.
- Detle, D. E., Abele, A. E. & Renner, O. (2004). Zur Definition und Messung von Berufserfolg. Theoretische Überlegungen und metaanalytische Befunde zum Zusammenhang von externen und internen Laufbahnerfolgsmäßen. *Zeitschrift für Personalpsychologie*, 3, 170-183.
- Deutsche Gesellschaft für Personalwesen e.V. (1996). *Berliner Intelligenzstruktur-Test der Deutschen Gesellschaft für Personalwesen e.V. (BIS-r-DGP)*. Hannover: Deutsche Gesellschaft für Personalwesen e.V.
- Deutsche Lufthansa AG. (1980). *Validierung der Auswahl von auszubildenden Flugzeugmechanikern*. Frankfurt a. M.: Deutsche Lufthansa AG.
- Deutsches Institut für Normung e.V. (2002). *DIN 33430: Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen*. Berlin: Beuth.
- Dickens, W. & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review*, 108, 346-369.
- Diehl, J. M. & Arbinger, R. (2001). *Einführung in die Inferenzstatistik* (3. korr. Aufl.). Eschborn: Klotz.
- Diehl, J. M. & Kohr, H. U. (2004). *Deskriptive Statistik* (13. überarb. Aufl.). Eschborn: Klotz.
- *Diepenbrock, W. & Thielepape, M. (1973). Untersuchung über die Vorhersage des Ausbildungserfolgs von Meisteranwärtern in der Automobilproduktion. *DGP-Informationen*, 12, 1-14.
- *Dreyer, H. & Grabitz, H. (1971). Über den Wert psychologischer Eignungsuntersuchungen bei der Auslese von potentiellen Führungskräften. *Psychologische Rundschau*, 22, 187-196.
- Duval, S. (2005). The trim and fill method. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Hrsg.), *Publication bias in meta analysis: prevention, assessment and adjustments* (S. 127-144). Chichester: Wiley.
- *Ehlers, T. (1961). Über den Vorhersagewert einiger psychologischer Untersuchungsverfahren. *Psychologie und Praxis*, 5, 157-163.
- Ehmke, T., Hohensee, F., Heidemeier, H. & Prenzel, M. (2004). Familiäre Lebensverhält-

- nisse, Bildungsbeteiligung und Kompetenzerwerb. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2003 Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (S. 225-254). Münster: Waxmann.
- Elias, P. & Birch, M. (1994). *Establishment of community-wide occupational statistics. ISCO 88 (COM). A guide for users*. UK: University of Warwick.
- Engelbrecht, W. (1975). Validierung einer Berufseignungs-Testbatterie und Verwendung der Ergebnisse für eine computerunterstützte berufsbezogene Testbefundinterpretation. *Diagnostica*, 21, 3-24 und 97-106.
- Engelbrecht, W. (1978). Weiterentwicklung der maschinellen Testbefundinterpretation zur EUB Testbatterie. *Diagnostica*, 24, 39-49.
- Engelbrecht, W., Schröder, R. & Elgert, W. (1991). *BWT, Entwicklungsarbeiten* (Bd. 46). Nürnberg: Bundesanstalt für Arbeit.
- Eysenck, H. J. & Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach*. New York: Plenum.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ-tests really measure. *Psychological Bulletin*, 101, 171-191.
- French, J. W., Ekstrom, R. B. & Price, L. A. (1963). *Manual for kit of reference tests for cognitive factors*. Princeton, NJ: Education Testing System.
- *Fünfgelt, V., Barthel, E. & Schuler, H. (1984). *Modellversuch zur Erschließung gewerblich-technischer Ausbildungsberufe für Mädchen. Abschlußbericht der wissenschaftlichen Begleitung in Augsburg*. Unveröffentlichtes Manuskript.
- Funke, J. & Vaterrodt-Plünnecke, B. (1998). *Was ist Intelligenz?* München: Beck.
- *Funke, U. (1986). Die Validität verschiedener eignungsdiagnostischer Verfahren bei Lehrstellenbewerbern. *Zeitschrift für Arbeits- und Organisationspsychologie*, 30, 92-97.
- Funke, U., Krauß, J., Schuler, H. & Stapf, K. H. (1987). Zur Prognostizierbarkeit wissenschaftlich-technischer Leistungen mittels Personvariablen: Eine Metaanalyse der Validität diagnostischer Verfahren im Bereich Forschung und Entwicklung. *Gruppendynamik*, 18, 407-428.
- Gaschok, J. (2002). Figure Reasoning Test (FRT). In R. Brickenkamp (Hrsg.), *Handbuch psychologischer und pädagogischer Tests* (3. Aufl., Bd. 1, S. 126-127). Göttingen: Hogrefe.
- Gattiker, U. E. & Larwood, L. (1986). Subjective career success: A study of managers and support personnel. *Journal of Business and Psychology*, 1, 78-94.
- Gaugler, B. B., Rosenthal, D. B., Thornton, G. C. & Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology*, 72, 493-511.
- Geis, A. (2006). *Handbuch für die Berufsvercodung*. Mannheim: Zentrum für Umfragen, Methoden und Analysen (ZUMA).
- Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology*, 26, 461-477.
- Glass, G. V., McGaw, B. & Smith, M. L. (1984). *Meta-analysis in social research* (3. Aufl.). Beverly Hills, CA: Sage.

- Goldstein, H. W., Zedeck, S. & Goldstein, I. L. (2002). g: Is this your final answer? *Human Performance*, 15, 123-142.
- *Görlich, Y. & Schuler, H. (2004). [Arbeitsprobe zur berufsbezogenen Intelligenz – Technische und handwerkliche Tätigkeiten (AZUBI-TH) und berufliche Ausbildungsleistung] Unveröffentlichte Daten.
- Görlich, Y. & Schuler, H. (2006). Personalentscheidungen, Nutzen und Fairness. In H. Schuler (Hrsg.), *Lehrbuch der Personalpsychologie* (S. 797-840). Göttingen: Hogrefe.
- Görlich, Y. & Schuler, H. (2007). *Arbeitsprobe zur berufsbezogenen Intelligenz – Technische und handwerkliche Tätigkeiten (AZUBI-TH)*. Göttingen: Hogrefe.
- Gottfredson, L. S. (2003). g, jobs and life. In H. Nyborg (Hrsg.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (S. 293-342). Oxford: Pergamon.
- *Graudenz, H. (1982). Vorhersage des Ausbildungserfolges von Beamtenanwärtern des gehobenen Dienstes beim RP Darmstadt. *DGP-Informationen*, 42, 32-50.
- *Graudenz, H. (1984). Personalauswahl für den mittleren Justizvollzugsdienst in Hessen. Eine Bewährungskontrolle. *DGP-Informationen*, 44, 1-17.
- Greguras, G. J. & Robie, C. (1998). A new look at within-source interrater reliability of 360-degree feedback ratings. *Journal of Applied Psychology*, 83, 960-968.
- *Greif, S. (1972). *Gruppenintelligenztests – Untersuchungen am WIT, IST, LPS und AIT*. Frankfurt: Lang.
- Greif, S. & Kluge, A. (2004). Lernen in Organisationen. In H. Schuler (Hrsg.), *Enzyklopädie der Psychologie – Organisationspsychologie* (S. 752-825). Göttingen: Hogrefe.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Guilford, J. P. (1977). *Way beyond the IQ: Guide to improving intelligence and creativity*. New York: McGraw-Hill.
- Guilford, J. P. (1981). Higher-order structure-of-intellect abilities. *Multivariate Behavioral Research*, 16, 411-435.
- Guttenberg, R. L., Arvey, R. D., Osborn, H. G. & Jeanneret, P. R. (1983). Moderating effects of decision-making/information processing job dimensions on test validities. *Journal of Applied Psychology*, 68, 602-608.
- Guttman, L. (1965). A faceted definition of intelligence. In R. Eiferman (Hrsg.), *Studies in psychology* (Bd. 14, S. 166-181). Jerusalem: The Hebrew University.
- Guttman, L. & Levy, S. (1991). Two structural laws for intelligence tests. *Intelligence*, 15, 79-103.
- Hacker, W. & Skell, W. (1993). *Lernen in der Arbeit*. Berlin: Bundesinstitut für Berufsbildung.
- Hardison, C. M. & Sackett, P. R. (2007). Kriteriumsbezogene Validität des Assessment Centers: lebendig und wohlauf? In H. Schuler (Hrsg.), *Assessment Center zur Potenzialanalyse* (S. 192-202). Göttingen: Hogrefe.
- Harris, M. M. & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41, 43-62.
- Hartigan, J. & Wigdor, A. K. (Hrsg.). (1989). *Fairness in employment testing. Validity*

- generalization, minority issues, and the General Aptitude Test Battery*. Washington, DC: National Academy Press.
- Hasselmann, D. (1993). *Computersimulierte komplexe Problemstellungen in der Management-Diagnostik*. Hamburg: Windmühle.
- Hedges, L. V. & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- *Herzog. (2004). *Eignungsfeststellung beim Zentrum für Nachwuchsgewinnung der Marine. Eine Prüfung der soldatischen Eignungsprognosen auf ihre Vorhersagevalidität*. Unveröffentlichte Diplomarbeit, [Ort und Institution unbekannt].
- Heyde, G. (1996). Weitere Daten zum IST-70. *Report Psychologie*, 21, 138.
- Hirsh, H. R., Northrop, L. C. & Schmidt, F. L. (1986). Validity generalization results for law enforcement occupations. *Personnel Psychology*, 39, 399-420.
- Holling, H. & Liepmann, D. (2004). Personalentwicklung. In H. Schuler (Hrsg.), *Lehrbuch Organisationspsychologie* (3. vollst. überarb. und erg. Aufl., S. 345-383). Bern: Huber.
- Holling, H., Preckel, F. & Vock, M. (2004). *Intelligenzdiagnostik*. Göttingen: Hogrefe.
- Hollingshead, A. B. (1975). *Four factor index of social status*. Unveröffentlichtes Manuskript.
- Hollstegge, S. (1994). *Berwerberauswahl und Ausbildungserfolg im öffentlichen Dienst*. Unveröffentlichte Diplomarbeit, Universität Bielefeld.
- Hopewell, S., Clarke, M. & Mallett, S. (2005). Grey literature and systematic reviews. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Hrsg.), *Publication bias in meta analysis: prevention, assessment and adjustments* (S. 49-72). Chichester: Wiley.
- Horn, W. (1956). *Begabungstestsystem (B-T-S)*. Göttingen: Verlag für Psychologie.
- Horn, W. (1962). *Leistungsprüfsystem (LPS)*. Göttingen: Verlag für Psychologie.
- Horn, W. (1969). *Prüfsystem für Schul- und Bildungsberatung PSB. Handanweisung*. Göttingen: Hogrefe.
- Horn, W. (1983). *Leistungsprüfsystem LPS. Handanweisung* (2. erw. und verb. Aufl.). Göttingen: Hogrefe.
- *Hossiep, R. (1995). *Berufseignungsdiagnostische Entscheidungen. Zur Bewährung eignungsdiagnostischer Ansätze*. Göttingen: Hogrefe.
- Hülshager, U. R. (2006). *Persönlichkeit im Arbeitsleben: Zur Bedeutung von Persönlichkeitseigenschaften und Intelligenz für berufliche Leistung und beruflichen Erfolg*. Unveröffentlichte Dissertation, Universität Bielefeld.
- Hülshager, U. R., Maier, G. W. & Stumpp, T. (2007). Validity of general mental ability for the prediction of job performance and training success in Germany: A meta-analysis. *International Journal of Selection and Assessment*, 15, 3-18.
- Hülshager, U. R., Maier, G. W., Stumpp, T. & Muck, P. M. (2006). Vergleich kriteriumsbezogener Validitäten verschiedener Intelligenztests zur Vorhersage von Ausbildungserfolg in Deutschland. *Zeitschrift für Personalpsychologie*, 5, 145-162.
- Hunt, M. (1997). *How science takes stock: The story of meta-analysis*. New York: Russel Sage Foundation.
- Hunter, J. E. (1980). *Validity generalization for 12,000 jobs: An application of synthetic*

- validity and validity generalization to the General Aptitude Test battery (GATB)*. Washington, DC: U.S. Department of Labor, Employment Service.
- Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, 29, 340-362.
- Hunter, J. E. & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72-98.
- Hunter, J. E. & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Beverly Hills, CA: Sage.
- Hunter, J. E. & Schmidt, F. L. (1996). Intelligence and job performance: Economic and social implications. *Psychology, Public Policy, and Law*, 2, 447-472.
- Hunter, J. E. & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment*, 8, 275-292.
- Hunter, J. E. & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2. Aufl.). Beverly Hills, CA: Sage.
- Hunter, J. E., Schmidt, F. L. & Le, H. (2006). Implications of direct and indirect range restriction for meta-analysis methods and findings. *Journal of Applied Psychology*, 91, 594-612.
- International Labour Office (Hrsg.). (1990). *International Standard Classification of Occupations: ISCO-88*. Geneva: International Labour Office.
- *Jäger, A. O. (1960). Zum prognostischen Wert psychologischer Eignungsuntersuchungen, eine Bewährungsstudie. *Psychologische Rundschau*, 6, 160-178.
- Jäger, A. O. (1963). Der Wilde-Test. Ein neues Intelligenzdiagnostikum. *Zeitschrift für experimentelle und angewandte Psychologie*, 10, 260-278.
- Jäger, A. O. (1970). Personalauslese. In A. Mayer & B. Herwig (Hrsg.), *Handbuch der Psychologie* (Bd. 9, S. 613-667). Göttingen: Hogrefe.
- Jäger, A. O. (1982). Mehrmodale Klassifikation von Intelligenzleistungen. Experimentell kontrollierte Weiterentwicklung eines deskriptiven Intelligenzstrukturmodells. *Diagnostica*, 28, 195-226.
- Jäger, A. O. (1986). Validität von Intelligenztests. *Diagnostica*, 32, 272-289.
- Jäger, A. O. & Althoff, K. (1983). *Der Wilde-Intelligenz-Test (WIT) – Ein Strukturdiagnostikum*. Göttingen: Hogrefe.
- Jäger, A. O. & Althoff, K. (1994). *Der Wilde-Intelligenz-Test (WIT) – Ein Strukturdiagnostikum* (2. rev. Aufl.). Göttingen: Hogrefe.
- Jäger, A. O., Süß, H. M. & Beauducel, A. (1997). *BIS-TEST Berliner Intelligenzstruktur-Test – Form 4*. Göttingen: Hogrefe.
- Jäger, A. O., Süß, H.-M. & Brühl, S. (1988). *BIS-1 Test für das Berliner Intelligenzstrukturmodell*. Unveröffentlichtes Manuskript, Freie Universität Berlin.
- Jäger, A. O. & Todt, E. (1964). Zur Faktorenstruktur des Wilde-Intelligenz-Tests (WIT) bei Siebzehnjährigen; Faktorenanalyse der WIT Langformen. *Diagnostica*, 10, 3-14.
- Jensen, A. R. (1984). Test bias: Concepts and criticisms. In C. R. Reynolds & R. T. Brown (Hrsg.), *Perspectives on bias in mental testing* (S. 507-586). New

- York: Plenum Press.
- Jensen, A. R. (1987). Individual differences in mental ability. In J. A. Glover & R. R. Ronnings (Hrsg.), *Historical foundations of educational psychology* (S. 61-88). New York: Plenum Press.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport: Praeger Publishers.
- Johnson, B. T., Mullen, B. & Salas, E. (1995). Comparison of three major meta-analytic approaches. *Journal of Applied Psychology*, *80*, 94-106.
- Johnson, W. & Bouchard, T. J. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 393-416.
- Judge, T. A. & Bono, J. E. (2001). Relationship of core self-evaluations traits – self-esteem, generalized self-efficacy, locus of control, and emotional stability – with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology*, *86*, 80-92.
- Judge, T. A., Cable, D. M., Boudreau, J. W. & Bretz, R. D. (1995). An empirical investigation of the predictors of executive career success. *Personnel Psychology*, *48*, 485-519.
- Judge, T. A., Colbert, A. E. & Ilies, R. (2004). Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology*, *89*, 542-552.
- Judge, T. A., Higgins, C. A., Thoresen, C. J. & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, *52*, 621-652.
- *Jungkunz, D. (1995). *Berufsausbildungserfolg in ausgewählten Ausbildungsberufen des Handwerks. Theoretische Klärung und empirische Analyse*. Weinheim: Deutscher Studien Verlag.
- *Jungkunz, D. & Bodinet, K. (1989). Korrelative Bedeutung von Testergebnissen, schulischer Vorbildung, Berufsschulnoten und Fähigkeitseinschätzungen für Berufsabschlussnoten. *Zeitschrift für Pädagogik*, *35*, 711-730.
- *Jungkunz, D. & John, W. (1991). Bedingungsfaktoren bei der Vergabe von Zeugnisnoten an Auszubildende handwerklicher Ausbildungsberufe. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, *87*, 481-494.
- Kanfer, R. & Ackerman, P. L. (2004). Aging, adult development, and work motivation. *Academy of Management Journal*, *29*, 440-458.
- Kanning, U. P. (Hrsg.). (2002). *Handbuch personaldiagnostischer Instrumente*. Göttingen: Hogrefe.
- Kemery, E. R., Mosholder, K. W. & Roth, L. (1987). The power of the Schmidt and Hunter additive model of validity generalization. *Journal of Applied Psychology*, *72*, 30-37.
- Kersting, M. (1998). Differentielle Aspekte der sozialen Akzeptanz von Intelligenztests und Problemlöseszenarien als Personalauswahlverfahren. *Zeitschrift für Arbeits- und Organisationspsychologie*, *42*, 61-75.

- *Kersting, M. (1999a). *Diagnostik und Personalauswahl mit computergestützten Problemlöseszenarien? Zur Kriteriumsvalidität von Problemlöseszenarien und Intelligenztests*. Göttingen: Hogrefe.
- Kersting, M. (1999b). Intelligenz-Struktur-Test 2000 (I-S-T 2000). In E. Fay (Hrsg.), *Tests unter der Lupe II* (S. 88-115). Lengerich: Pabst.
- Kersting, M. (1999c). Personalauswahl mit dem BIS-r-DGP Test: Evaluation der prognostischen Validität. *DGP-Informationen*, 56, 43-60.
- Kersting, M. (2000). Instrumente der Arbeits- und Organisationspsychologie: Rezension des „Intelligenz-Struktur-Test 2000“ von R. Amthauer, B. Brocke, D. Liepmann und A. Beauducel. *Zeitschrift für Arbeits- und Organisationspsychologie*, 44, 96-101.
- Kersting, M. (2001). Zur Konstrukt- und Kriteriumsvalidität von Problemlöseszenarien anhand der Vorhersage von Vorgesetztenurteilen über die berufliche Bewährung. *Diagnostica*, 47, 67-76.
- *Kersting, M. (2003). Assessment Center: Erfolgsmessung und Qualitätskontrolle. In S. Höft & B. Wolf (Hrsg.), *Qualitätsstandards für Personalentwicklung in Wirtschaft und Verwaltung* (S. 72-93). Hamburg: Windmühle.
- Kersting, M. (2004). Qualitätssicherung und -verbesserung: Zur Überprüfung der Gültigkeit berufsbezogener Eignungsbeurteilungen. In L. Hornke & U. Winterfeld (Hrsg.), *Eignungsbeurteilung auf dem Prüfstand: DIN 33430 zur Qualitätssicherung* (S. 103-128). Heidelberg: Spektrum Akademischer Verlag.
- Kersting, M. (2006). *Kurzinformation zum neuen WILDE Intelligenztest 2 (WIT-2)*. Zugriff am 6.11.2007 unter http://www.bopsych.rwth-aachen.de/contento/cms/upload/Kersting/wilde/Wilde_Kurzbeschreibung.pdf
- Kersting, M. (2007). Wenn Tests in die Jahre kommen. Probleme des Einsatzes überalteter Testverfahren. In C. Lorei (Hrsg.), *Polizei und Psychologie* (S. 565-577). Frankfurt: Verlag für Polizeiwissenschaft.
- *Kersting, M., Althoff, K. & Jäger, A. O. (2005, Juni). [Kriteriumsvalidität neuer Wilde Test]. Unveröffentlichte Daten.
- Kersting, M., Althoff, K. & Jäger, A. O. (2008). *Der WILDE Intelligenztest 2 (WIT-2)*. Göttingen: Hogrefe.
- Kersting, M. & Beauducel, A. (1997). Der neue DGP-Leistungstest auf der Basis des Berliner Intelligenzstrukturmodells: Informationen zu ausgewählten Testgütekriterien und zur Normierung. *DGP-Informationen*, 55, 93-103.
- Kersting, M. & Beauducel, A. (2004). BIS-r-DGP (sowie die Kurzformen A-1 und K-1). Berliner Intelligenzstruktur-Test der Deutschen Gesellschaft für Personalwesen e.V. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren* (2. überarb. und erw. Aufl., S. 149-157). Lengerich: Pabst Science Publishers.
- *Kettel, J. & Simmat, W. (1969). Beitrag zu einem psychologischen Berufsprofil für Fernmeldehandwerker. *Psychologie und Praxis*, 13, 169-174.
- Kirsch, H. (1961). Ergebnisse einer Bewährungskontrolle von quantitativen Verfahren für die Vorhersage der Fliegereignung. *Diagnostica*, 7, 17-123.

- *Kleinevoss, R. (1983). Untersuchung zur Vorhersage des Ausbildungserfolges von Anwärtern des gehobenen Dienstes einer Bundesbehörde. *DGP-Informationen*, 43, 41-72.
- *Knebelau, M. (2002). *Evaluation des Wilde-Intelligenz-Tests (WIT) und seiner Funktion als eignungsdiagnostisches Instrument*. Aachen: RWTH Aachen.
- *Kober, U. (1978). Eignungsuntersuchungs-Ergebnisse bei mehreren Ausbildungsberufen in einer Bewährungskontrolle durch verschiedene Methoden. *Psychologie und Praxis*, 22, 168-177.
- *Krause, D. E., Kersting, M. & Heggstad, E. D. (2006). Incremental validity of assessment center ratings over cognitive ability tests: A study at the executive management level. *International Journal of Selection and Assessment*, 14, 360-371.
- *Krause, D. E., Kersting, M., Heggstad, E. D. & Thornton, G. C. (2005, April). *Criterion validity of cognitive ability tests and assessment centers*. Paper presented at the 20th Annual Society of Industrial and Organizational Psychology (SIOP) Conference, Los Angeles.
- *Kreuscher, R. (1987). *Der Zusammenhang von Berufserfolg mit Handlungsstil und fünf weiteren, psychologischen Variablen bei selbständig im Außendienst tätigen*. Unveröffentlichte Diplomarbeit, Maximilians-Universität München.
- *Kuhl, J. & Greve, H. G. (1971). Validierung zweier Testsysteme zur Selektion und Beurteilung von Auszubildenden in Metall- und Elektroberufen. *Psychologie und Praxis*, 15, 81-88.
- Kuncel, N. R., Hezlett, S. A. & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, 86, 148-161.
- *Lang, K. (1990). *Entwicklung und längsschnittliche Validierung einer eignungsdiagnostischen Testbatterie für Auszubildende gewerblich-technischer Berufe*. Unveröffentlichte Diplomarbeit, Justus-Liebig-Universität Gießen.
- Legg, S. & Hutter, M. (2007). A collection of definitions of intelligence. *Frontiers in Artificial Intelligence and Applications*, 157, 17-24.
- Levine, E. L., Spector, P. E., Menon, P. E., Narayanan, L. & Cannon-Bowers, J. (1996). Validity generalisation for cognitive, psychomotor, and perceptual tests for craft jobs in the utility industry. *Human Performance*, 9, 1-22.
- Levy-Leboyer, C. (1994). Selection and assessment in Europe. In H. C. Triandis, M. D. Dunnette & L. M. Hough (Hrsg.), *Handbook of industrial and organizational psychology* (Bd. 4, S. 173-190). Palo Alto, CA: Consulting Psychologists Press.
- Lienert, G. A. (1958). Ein Test zur Beurteilung des mechanisch-technischen Verständnisses. *Zeitschrift für experimentelle und angewandte Psychologie*, 6, 605-620.
- Lienert, G. A. (1967). *Drahtbiegeprobe* (2. verb. Aufl.). Göttingen: Hogrefe.
- *Linz, P. (1985). Die kognitive Leistungsfähigkeit und die berufliche Integration hörgeschädigter Jugendlicher. In W. F. Kugemann, S. Preiser & K. A. Schneewind (Hrsg.), *Psychologie und Lebenswirklichkeit. Festschrift zum 65. Geburtstag von Walter Toman* (S. 309-320). Göttingen: Hogrefe.
- Lipsey, M. W. & Wilson, D. B. (2001). The way in which intervention studies have

- „personality” and why it is important to meta-analysis. *Evaluation & The Health Professions*, 24, 236-254.
- Lohman, D. F. (2000). Complex information processing and intelligence. In R. J. Sternberg (Hrsg.), *Handbook of intelligence* (S. 285-340). Cambridge University Press.
- *Lometsch, D. (1975). *Untersuchungen zur Vorhersage des Berufserfolgs von Führungsnachwuchskräften*. Unveröffentlichtes Manuskript, Hamburg: Deutsche Unilever GmbH.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *Psychologische Rundschau*, 58, 103-117.
- Lüken, K. H. (2002a). Grundintelligenztest Skala 2 mit Wortschatztest (WS) und Zahlenfolgentest (ZF) (CFT 20). In R. Brickenkamp (Hrsg.), *Handbuch psychologischer und pädagogischer Tests* (3. Aufl., Bd. 1, S. 135-138). Göttingen: Hogrefe.
- Lüken, K. H. (2002b). Grundintelligenztest Skala 3 (CFT 3). In R. Brickenkamp (Hrsg.), *Handbuch psychologischer und pädagogischer Tests* (3. Aufl., Bd. 1, S. 139-140). Göttingen: Hogrefe.
- Marcus, B. & Schuler, H. (2006). Leistungsbeurteilung. In H. Schuler (Hrsg.), *Lehrbuch der Personalpsychologie* (2. überarb. und erw. Aufl., S. 433-469). Göttingen: Hogrefe.
- *Marcus, B., Schuler, H., Quell, P. & Hümpfer, G. (2002). Measuring counter productivity: development and initial validation of a German self-report questionnaire. *International Journal of Selection and Assessment*, 10, 18-35.
- *Marschner, G. (1964). Auswahl des Steigernachwuchses. *Psychologie und Praxis*, 8, 49-65.
- *Marschner, G. (1966). Betriebspsychologische Erfahrungen mit dem I-S-T-Amthauer bei Auswahluntersuchungen. *Psychologie und Praxis*, 10, 145-153.
- *Marschner, G. (1972). Auswahluntersuchungen bei Führungsnachwuchskräften. *Psychologie und Praxis*, 16, 116-127.
- *Marschner, G. (1978). Differential-psychologische Untersuchungen an industriellen Facharbeitern. *Zeitschrift für Arbeitswissenschaft*, 32, 222-224.
- *Marschner, G. (1981). Untersuchungen mit dem Lern- und Gedächtnis-Test (LGT-3) bei Facharbeitern. *Diagnostica*, 27, 261-265.
- *Marschner, G. (1983a). Die geistige Leistungsfähigkeit im Verlauf der beruflichen Ausbildung. *Zeitschrift für Arbeitswissenschaft*, 37, 50-53.
- Marschner, G. (1983b). Schulprüfungängste bei männlichen Jugendlichen. *Psychologie in Erziehung und Unterricht*, 30, 150-156.
- Matt, G. E. (2003). Will it work in Münster? Meta-analysis and the empirical generalization of causal relationships. In R. Schulze, H. Holling & D. Böhning (Hrsg.), *Meta-analysis. New developments and applications in medical and social sciences* (S. 113-139). Göttingen: Hogrefe & Huber.
- McCormick, E. J., DeNisi, A. S. & Shaw, J. B. (1979). Use of the Position Analysis Questionnaire for establishing the job component validity of tests. *Journal of Applied Psychology*, 64, 51-56.

- Meili, R. (1966). *Analytischer Intelligenztest (AIT)*. Bern: Huber.
- Miller, W. S. (1960). *Technical manual for the Miller Analogies Test*. New York: The Psychological Corporation.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K. & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, *60*, 683-729.
- Moser, K. (2004a). Organisationale Sozialisation und berufliche Entwicklung. In H. Schuler (Hrsg.), *Enzyklopädie der Psychologie – Organisationspsychologie* (S. 533-595). Göttingen: Hogrefe.
- Moser, K. (2004b). Selbstbeurteilung. In H. Schuler (Hrsg.), *Beurteilung und Förderung beruflicher Leistung* (2. vollst. überarb. und erw. Aufl., S. 83-99). Göttingen: Hogrefe.
- Mosier, C. I. (1943). On the reliability of a weighted composite. *Psychometrika*, *8*, 161-168.
- Nettelstroth, W. (2003). *Intelligenz im Rahmen der beruflichen Tätigkeit. Zum Einfluss von Intelligenzfacetten, Personenmerkmalen und Organisationsstrukturen*. Unveröffentlichte Dissertation, Freie Universität Berlin.
- *Nettelstroth, W. (2004). Intelligenz und Berufserfolg. Eine Validierung der fluiden und kristallinen Intelligenz an subjektiven und objektiven Berufserfolgskriterien. In W. Bungard (Hrsg.), *Psychologie und Wirtschaft leben* (S. 415-420). München: Hampp.
- *Neumann, E. (1938). Psychotechnische Eignungsprüfung und Anlernung im Flugmotorenbau. *Industrielle Psychotechnik*, *15*, 111-162.
- Ng, T. W. H., Eby, L. T., Sorensen, K. L. & Feldman, D. C. (2005). Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology*, *58*, 367-408.
- Olkin, I. & Pratt, J. W. (1958). Unbiased estimation of certain correlation coefficients. *Annals of Mathematical Statistics*, *29*, 201-211.
- O'Reilly, C. A. & Chatman, J. A. (1994). Working smarter and harder: A longitudinal study of managerial success. *Administrative Science Quarterly*, *39*, 603-627.
- Orr, J. M., Sackett, P. R. & Mercer, M. (1989). The role of prescribed and nonprescribed behaviors in estimating the dollar value of performance. *Journal of Applied Psychology*, *74*, 34-40.
- Otis, A. S. (1918). An absolute point scale for the group measure of intelligence. *Journal of Educational Psychology*, *9*, 238-261, 333-348.
- *Paff, G. (1963). Ergebnisse einer Bewährungskontrolle bei Apparatewärtern und Maschinisten. *FORFA-Briefe*, *2*, 89-92.
- *Paff, G. (1966). *Schulleistung, Berufseignung und Bewährung*. Bern: Huber.
- Pawlowsky, P. & Bäumer, J. (1996). *Betriebliche Weiterbildung. Management von Qualifikation und Wissen*. München: Beck'sche Verlagsbuchhandlung.
- Pearlman, K., Schmidt, F. L. & Hunter, J. E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. *Journal of Applied Psychology*, *65*, 373-406.

- Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P. R., Fleishman, E. A., Levin, K. Y. et al. (2001). Understanding work using the Occupational Information Network (O*NET): Implications for practice and research. *Personnel Psychology*, *54*, 451-492.
- Pulakos, E. D., Arad, S., Donovan, M. A. & Plamondon, K. E. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology*, *85*, 612-624.
- Raju, N. S., Burke, M. J., Normand, J. & Langlois, G. M. (1991). A new meta-analytic approach. *Journal of Applied Psychology*, *76*, 432-446.
- Raven, J., Raven, J. C. & Court, J. H. (2000). *Raven manual: Section 3. SPM manual (including the parallel and plus versions). 2000 Edition*. Oxford: Oxford Psychologists Press.
- Raven, J., Raven, J. C. & Court, J. H. (2003). *Raven's standard progressive matrices plus - SPMPLS. Manual*. Wien: Schuhfried.
- Ree, M. J. & Carretta, T. R. (2002). g2K. *Human Performance*, *15*, 3-24.
- Reeve, C. L. & Hakel, M. D. (2002). Asking the right questions about g. *Human Performance*, *15*, 47-74.
- *Rolloff, H. P. (1928). Über Eignung und Bewährung. Forschungen zur industriellen Psychotechnik. *Beihefte zur Zeitschrift für Angewandte Psychologie*, *148*.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Rev. Aufl.). Newbury Park (CA): Sage.
- Rösing, I. (2004). *Intelligenz und Dummheit. Wissenschaftliche Konzepte, Alltagskonzepte, fremdkulturelle Konzepte. Ein Werk- und Denk-Buch*. Heidelberg: Asanger.
- Rothstein, H. R. (1990). Interrater reliability of job performance ratings: Growth to asymptote level with increasing opportunity to observe. *Journal of Applied Psychology*, *75*, 322-327.
- Rowe, D. C. & Rodgers, J. L. (2002). Expanding variance and the case of historical changes in IQ means: A critique of Dickens and Flynn (2001). *Psychological Review*, *109*, 759-763.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*, 581-592.
- Rustenbach, S. J. (2003). *Metaanalyse. Eine anwendungsorientierte Einführung*. Bern: Huber.
- Rüttinger, B. & Klein-Moddenborg, V. (1989). Aus-, Fort- und Weiterbildung. In E. Roth (Hrsg.), *Enzyklopädie der Psychologie – Organisationspsychologie* (S. 685-711). Göttingen: Hogrefe.
- Ryan, A. M., MacFarland, L., Baron, H. & Page, R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology*, *52*, 359-391.
- Sackett, P. R. (2007). Revisiting the origins of the typical-maximum performance distinction. *Human Performance*, *20*, 179-185.
- Sackett, P. R. & Ostgaard, D. J. (1994). Job-specific applicant pools and national norms for cognitive ability tests: Implications for range restriction corrections in validation research. *Journal of Applied Psychology*, *79*, 680-684.

- Salgado, J. F. & Anderson, N. (2002). Cognitive and GMA testing in the European Community: Issues and evidence. *Human Performance*, 15, 75-96.
- Salgado, J. F. & Anderson, N. (2003). Validity generalization of GMA tests across countries in the European Community. *European Journal of Work and Organizational Psychology*, 12, 1-17.
- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C. & de Fruyt, F. (2003a). International validity generalization of GMA and cognitive abilities: A European community meta-analysis. *Personnel Psychology*, 56, 573-605.
- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F. & Rolland, J.-P. (2003b). A meta-analytic study of general mental ability validity for different occupations in the European community. *Journal of Applied Psychology*, 88, 1068-1081.
- *Salzgitter Service und Technik GmbH. (2004). [Zusammenhang Intelligenzmaße und Leistungsmaße bei Auszubildenden der SZST (früher PPS) Salzgitter]. Unveröffentlichte Daten.
- Samson, G. E., Graue, E. M., Weinstein, T. & Wallberg, H. J. (1984). Academic and occupational performance: Quantitative synthesis. *American Educational Research Journal*, 21, 311-321.
- Sarges, W. & Wottawa, H. (Hrsg.). (2004). *Handbuch wirtschaftspsychologischer Testverfahren. Band I: Personalpsychologische Instrumente* (2. überarb. und erw. Aufl.). Lengerich: Pabst.
- Schaarschmidt, U. (1997). Intelligenz-Struktur-Test (IST 70). *Zeitschrift für Differentielle und Diagnostische Psychologie*, 18, 106-108.
- *Schäfer, N. (1986). *Eignung und Arbeitszufriedenheit. Beiträge zur Eignungsdiagnostik*. Frankfurt: Lang.
- Scherm, M. (2004). 360°-Beurteilung. In H. Schuler (Hrsg.), *Beurteilung und Förderung beruflicher Leistung* (2. vollst. überarb. und erw. Aufl., S. 61-81). Göttingen: Hogrefe.
- Schmale, H. (2004). BET. Berufseignungstest. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren* (2. überarb. und erw. Aufl., S. 115-118). Lengerich: Pabst Science Publishers.
- Schmale, H. & Schmidtke, H. (1966a). *BET Handanweisung für den Berufseignungstest*. Stuttgart: Huber.
- Schmale, H. & Schmidtke, H. (1966b). *BET Tabellenband*. Stuttgart: Huber.
- Schmale, H. & Schmidtke, H. (1969). *Eignungsprognose und Ausbildungserfolg*. Forschungsbericht des Landes Nordrhein-Westfalen Nr. 2044.
- Schmale, H. & Schmidtke, H. (1984). *BET Tabellenband* (2. vollst. überarb. Aufl.). Stuttgart: Huber.
- Schmale, H. & Schmidtke, H. (1995a). *Berufseignungstest BET Tabellenband* (3. Aufl.). Bern: Huber.
- Schmale, H. & Schmidtke, H. (1995b). *Handanweisung für den Berufseignungstest* (3. korr. Aufl.). Bern: Huber.
- *Schmale, H. & Schmidtke, H. (2001). *Berufseignungstest BET Tabellenband* (4. über-

- arb. und erg. Aufl.). Bern: Huber.
- Schmidt, F. L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance*, *15*, 187-211.
- Schmidt, F. L., Gast-Rosenberg, I. & Hunter, J. E. (1980). Validity generalization results for computer programmers. *Journal of Applied Psychology*, *65*, 643-661.
- Schmidt, F. L. & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Applied Psychology*, *86*, 162-173.
- Schmidt, F. L. & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology*, *62*, 529-540.
- Schmidt, F. L. & Hunter, J. E. (1998a). Meßbare Personmerkmale: Stabilität, Variabilität und Validität zur Vorhersage zukünftiger Berufsleistung und berufsbezogenen Lernens. In M. Kleinmann & B. Strauss (Hrsg.), *Potentialfeststellung und Personalentwicklung* (S. 15-43). Göttingen: Verlag für angewandte Psychologie.
- Schmidt, F. L. & Hunter, J. E. (1998b). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262-274.
- Schmidt, F. L., Hunter, J. E. & Caplan, J. R. (1981). Validity generalization results for two job groups in petroleum industry. *Journal of Applied Psychology*, *66*, 261-273.
- Schmidt, F. L., Hunter, J. E. & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology*, *71*, 432-439.
- Schmidt, F. L., Hunter, J. E. & Pearlman, K. (1981). Task difference as moderators of aptitude test validity in selection: A red herring. *Journal of Applied Psychology*, *66*, 166-185.
- Schmidt, F. L., Hunter, J. E., Pearlman, K. & Shane, G. S. (1979). Further tests of the Schmidt-Hunter bayesian validity generalization procedure. *Personnel Psychology*, *32*, 257-281.
- Schmidt, F. L. & Le, H. (2005). Software for the Hunter-Schmidt meta-analysis methods (Version 1.1) [Computer Software]. University of Iowa.
- Schmidt, F. L., Le, H. & Ilies, R. (2003). Beyond alpha: an empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual differences constructs. *Psychological Methods*, *8*, 206-224.
- Schmidt, F. L., Ones, D. S. & Hunter, J. E. (1992). Personnel selection. *Annual Review of Psychology*, *43*, 627-670.
- Schmidt, F. L. & Raju, N. S. (2007). Updating meta-analytic research findings: Bayesian approaches versus the medical model. *Journal of Applied Psychology*, *92*, 297-308.
- *Schmidt, J. U. (1987). Zur Vorhersagbarkeit des Ausbildungserfolgs im gehobenen Verwaltungsdienst durch einen auf der Grundlage des Berliner Intelligenzstrukturmodells entwickelten Test. *DGP-Informationen*, *38*, 29-45.
- Schmidt, K.-H. & Kleinbeck, U. (2004). Leistung und Leistungsförderung. In H. Schuler (Hrsg.), *Enzyklopädie der Psychologie – Organisationspsychologie* (S. 891-945). Göttingen: Hogrefe.

- Schmidt-Atzert, L. (1997). Replik zur Rezension des IST 70. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 18, 109-112.
- Schmidt-Atzert, L. (2002). Tests und Tools: Intelligenz-Struktur-Test 2000 R. *Zeitschrift für Personalpsychologie*, 1, 50-56.
- *Schmidt-Atzert, L. & Deter, B. (1993). Intelligenz und Ausbildungserfolg: Eine Untersuchung zur prognostischen Validität des I-S-T 70. *Zeitschrift für Arbeits- und Organisationspsychologie*, 37, 52-63.
- *Schmidt-Atzert, L., Deter, B. & Jaeckel, S. (2004). Prädiktion von Ausbildungserfolg: Allgemeine Intelligenz (g) oder spezifische kognitive Fähigkeiten? *Zeitschrift für Personalpsychologie*, 3, 147-158.
- *Schmidt-Atzert, L. & Eser, K.-H. (2006, März). *Intelligenz, Leistungsmotivation und Prognose des Ausbildungserfolgs bei lern- und mehrfachbehinderten Teilnehmern in Berufsbildungswerken*. Vortrag präsentiert auf der Konferenz „teilhaben2006.de – Ein Beruf gehört zum Leben“, Berlin.
- Schmidt-Atzert, L., Hommers, W. & Hess, M. (1995). Der IST-70: Eine Analyse und Neubewertung. *Diagnostica*, 41, 108-130.
- *Schmitt, A. A. (2006). [Faktoren für Ausbildungserfolg in den neugeordneten technischen Ausbildungsberufen unter dem besonderen Fokus sozialer Kompetenzen]. Unpublizierte Daten.
- Schmitt, N., Cortina, J. M., Ingerick, M. J. & Wiechmann, D. (2003). Personnel selection and employee performance. In W. C. Borman, D. R. Ilgen & R. J. Klimoski (Hrsg.), *Handbook of Psychology, Vol. 12: Industrial and organizational psychology* (S. 77-105). Hoboken, NJ: Wiley.
- Schmitt, N., Gooding, R. Z., Noe, R. A. & Kirsch, M. (1984). Metaanalyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407-422.
- *Schmitz, I. (2006). *Prognostische Validität des Auswahlverfahrens für Auszubildende bei zwei Sparkassen in Nordrhein-Westfalen*. Unveröffentlichte Diplomarbeit, Universität Bonn.
- Schneer, J. A. & Reitman, F. (1995). The impact of gender as managerial careers unfold. *Journal of Vocational Behavior*, 47, 290-315.
- Schuhfried, G. (2004). SPMPLS. Raven's Standard Progressive Matrices Plus. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren* (2. überarb. und erw. Aufl., S. 745-748). Lengerich: Pabst Science Publishers.
- Schuler, H. (2002a). *Das Einstellungsinterview*. Göttingen: Hogrefe.
- Schuler, H. (2002b). Emotionale Intelligenz - ein irreführender und unnötiger Begriff. *Zeitschrift für Personalpsychologie*, 1, 138-140.
- Schuler, H. (2004a). Der Prozess der Urteilsbildung und die Qualität der Beurteilungen. In H. Schuler (Hrsg.), *Beurteilung und Förderung beruflicher Leistung* (2. vollst. überarb. und erw. Aufl., S. 33-60). Göttingen: Hogrefe.
- Schuler, H. (2004b). Leistungsbeurteilung - Gegenstand, Funktionen und Formen. In H. Schuler (Hrsg.), *Beurteilung und Förderung beruflicher Leistung* (2. vollst. überarb. und erw. Aufl., S. 1-23). Göttingen: Hogrefe.

- *Schuler, H., Barthel, E. & Fünfgelt, V. (1980). *Modellversuch zur Erschließung gewerblich-technischer Ausbildungsberufe für Mädchen. 1. Zwischenbericht der wissenschaftlichen Begleitung in Augsburg*. Unveröffentlichtes Manuskript, Universität Erlangen-Nürnberg.
- *Schuler, H., Barthel, E. & Fünfgelt, V. (1984). Erfolg von Mädchen in gewerblich-technischen Ausbildungsberufen: Ein Modellversuch. *Zeitschrift für Arbeits- und Organisationspsychologie*, 2, 67-78.
- Schuler, H., Frier, D. & Kauffmann, M. (1993). *Personalauswahl im europäischen Vergleich*. Göttingen: Hogrefe.
- *Schuler, H., Funke, U., Moser, K. & Donat, M. (1995). *Personalauswahl in Forschung und Entwicklung. Eignung und Leistung von Wissenschaftlern und Ingenieuren*. Göttingen: Hogrefe.
- *Schuler, H., Funke, U., Moser, K., Donat, M., Barthelme, D., Bretz, E. et al. (1991). *Abschlussbericht zum Forschungsprojekt: „Entwicklung eines eignungsdiagnostischen Personalauswahlsystems für Wissenschaftler und Ingenieure im Bereich industrieller Forschung und Entwicklung“*. Berichtszeitraum 1.4.1986 - 30.9.1990. Unveröffentlichtes Manuskript, Universität Hohenheim.
- Schuler, H., Hell, B., Trapmann, S., Schaar, H. & Boramir, I. (2007). Die Nutzung psychologischer Verfahren der externen Personalauswahl in deutschen Unternehmen: Ein Vergleich über 20 Jahre. *Zeitschrift für Personalpsychologie*, 6, 60-70.
- Schuler, H. & Höft, S. (2004). Diagnose beruflicher Eignung und Leistung. In H. Schuler (Hrsg.), *Lehrbuch Organisationspsychologie* (3. vollst. überarb. und erg. Aufl., S. 289-343). Bern: Huber.
- Schuler, H. & Höft, S. (2006). Konstruktorientierte Verfahren der Personalauswahl. In H. Schuler (Hrsg.), *Lehrbuch der Personalpsychologie* (2., überarb. und erw. Aufl., S. 101-144). Göttingen: Hogrefe.
- *Schuler, H. & Klingner, Y. (2002). [Arbeitsprobe zur berufsbezogenen Intelligenz – Büro- und kaufmännische Tätigkeiten (AZUBI-BK) und Intelligenz-Struktur-Test 70 (IST-70) und berufliche Ausbildungsleistung]. Unveröffentlichte Daten.
- Schuler, H. & Klingner, Y. (2004). AZUBI-BK. Arbeitsprobe zur Berufsbezogenen Intelligenz - Büro- und kaufmännische Tätigkeiten. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren* (2. überarb. und erw. Aufl., S. 99-103). Lengerich: Pabst Science Publishers.
- *Schuler, H. & Klingner, Y. (2005). *Arbeitsprobe zur berufsbezogenen Intelligenz - Büro- und kaufmännische Tätigkeiten (AZUBI-BK)*. Göttingen: Hogrefe.
- *Schuler, H., Moser, K., Diemand, A. & Funke, U. (1995). Validität eines Einstellungsinterviews zur Prognose des Ausbildungserfolgs. *Zeitschrift für Pädagogische Psychologie*, 9, 45-54.
- Schulz, C., Schuler, H. & Stehle, W. (1985). Die Verwendung eignungsdiagnostischer Methoden in deutschen Unternehmen. In H. Schuler & W. Stehle (Hrsg.), *Organisationspsychologie und Unternehmenspraxis: Perspektiven der Kooperation* (S. 126-132). Göttingen: Hogrefe.
- Schulze, R. (2004). *Meta-analysis: A comparison of approaches*. Cambridge: Hogrefe &

- Huber.
- Schulze, R., Holling, H. & Böhning, D. (Hrsg.). (2003). *Meta-analysis. New developments and applications in medical and social sciences*. Göttingen: Hogrefe & Huber.
- *Schwadorf, H. (2003). *Berufliche Handlungskompetenz*. Hohenheim: ibw.
- Scotter, J. R. v. & Motowidlo, S. J. (1996). Interpersonal facilitation and job dedication as separate facets of contextual performance. *Journal of Applied Psychology*, 81, 525-531.
- *Seggebruch, G. (1982). Zur Vorhersage des Berufserfolgs von Bewerbern für den mittleren Verwaltungsdienst in einer Kommunalverwaltung. *DGP-Informationen*, 42, 6-31.
- *Seggebruch, G. (1984). Vergleich der Ergebnisse aus Eignungsuntersuchungen und Zwischenprüfungen im gehobenen allgemeinen Verwaltungsdienst – eine Bewährungskontrolle. *DGP-Informationen*, 44, 56-74.
- Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review*, 17, 881-901.
- Spearman, C. (1904). „General Intelligence“ objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Statistisches Bundesamt (Hrsg.). (1988). *Klassifizierung der Berufe, systematisches und alphabetisches Verzeichnis der Berufsbenennungen, Ausgabe 1988*. Stuttgart: Kohlhammer.
- Statistisches Bundesamt. (2008). [Alter und Geschlecht der Erwerbstätigen Bevölkerung über die Zeit]. (erhältlich vom Statistischen Bundesamt, Gustav-Stresemann-Ring 11, 65189 Wiesbaden)
- *Steinmann, V. (1997). *Die Prognosekraft des Eignungstests Bankkaufmann für den späteren Berufserfolg: eine Längsschnittstudie*. Unveröffentlichte Diplomarbeit, Universität Köln.
- Sternberg, R. J. (1987). Intelligence. In R. L. Gregory (Hrsg.), *The Oxford companion to the mind* (S. 375-383). Oxford University Press.
- Sternberg, R. J. & Detterman, D. K. (Hrsg.). (1986). *What is intelligence? Contemporary viewpoints on its nature and definition*. Norwood, NJ: Ablex.
- Sternberg, R. J. & Powell, J. S. (1982). Theories of intelligence. In R. J. Sternberg (Hrsg.), *Handbook of human intelligence* (S. 975-1005). Cambridge University Press.
- Sterne, J. A. C., Becker, B. J. & Egger, M. (2005). The funnel plot. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Hrsg.), *Publication bias in meta analysis: prevention, assessment and adjustments* (S. 75-98). Chichester: Wiley.
- *Streitkräfteamt der Bundeswehr. (1996a). [Arbeitsbericht]. Unveröffentlichte Daten.
- *Streitkräfteamt der Bundeswehr. (1996b). [Arbeitsbericht]. Unveröffentlichte Daten.
- *Streitkräfteamt der Bundeswehr. (2001). [Arbeitsbericht]. Unveröffentlichte Daten.
- *Streitkräfteamt der Bundeswehr. (2004). [Arbeitsbericht des Psychologischen Dienstes]. Unveröffentlichte Daten.
- Süß, H.-M. (2001). Prädiktive Validität der Intelligenz im schulischen und außerschulischen Bereich. In E. Stern & J. Guthke (Hrsg.), *Perspektiven der Intelligenzfor-*

- schung (S. 109-135). Lengerich: Pabst.
- Süß, H.-M. (2004). BIS-4. Berliner Intelligenzstruktur-Test – Form 4. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren* (2. überarb. und erw. Aufl., S. 139-147). Lengerich: Pabst Science Publishers.
- Süß, H.-M., Brühl, S. & Jäger, A. O. (1989). *Berliner Untersuchung zum Zusammenhang von Intelligenz, Wissen und komplexem Problemlösen. Materialband*. Unveröffentlichtes Manuskript, Freie Universität Berlin.
- Süß, H.-M. & Jäger, A. O. (1990). *Berliner Wiederholungsuntersuchung zum Zusammenhang von Intelligenz, Wissen und komplexem Problemlösen. Materialband*. Unveröffentlichtes Manuskript, Freie Universität Berlin.
- Süß, H.-M., Oberauer, K. & Jäger, A. O. (1988). Geltungsbereiche des Berliner Intelligenzstrukturmodells. In W. Schönplugg (Hrsg.), *Bericht über den 36. Kongreß der Deutschen Gesellschaft für Psychologie in Berlin 1988* (S. 300-301). Göttingen: Hogrefe.
- Süß, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O. & Schulze, R. (2002). Working-memory capacity explains reasoning ability – and a little bit more. *Intelligence*, 30, 261-288.
- Sutton, A. J. & Pigott, T. D. (2005). Bias in meta-analysis induced by incompletely reported studies. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Hrsg.), *Publication bias in meta analysis: prevention, assessment and adjustments* (S. 223-239). Chichester: Wiley.
- Tent, L. (1969). *Die Auslese von Schülern für weiterführende Schulen - Möglichkeiten und Grenzen*. Göttingen: Verlag für Psychologie.
- Tent, L. (2001). Zensuren. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (2. überarb. und erw. Aufl., S. 805-811). Weinheim: Psychologie Verlags Union.
- Terman, L. M. & Merrill, M. A. (1960). *Stanford-Binet intelligence scale: Manual for the third revision form L-M*. Boston: Houghton Mifflin.
- Tharenou, P., Latimer, S. & Conroy, D. (1994). How do you make it to the top? An examination of influences on women's and men's managerial advancement. *Academy of Management Journal*, 37, 899-931.
- *Thielepape, M. (1980). Untersuchung über die Vorhersage des Ausbildungserfolges von Krankenschwestern und -pflegern – eine Bewährungskontrolle. *DGP-Informationen*, 17, 21-43.
- *Thielepape, M. & Kersting, M. (2005a). Erfolgskontrolle und Qualitätsoptimierung in der Personalauswahl: 2 Bewährungskontrollen. *DGP-Informationen*, 49, 2-20.
- *Thielepape, M. & Kersting, M. (2005b). [Evaluation und Qualitätsoptimierung in der Personalauswahl: Zwei Bewährungskontrollen]. Unveröffentlichte Daten.
- U. S. Department of Labor. (1991). *Dictionary of occupational titles* (4. Aufl.). Washington, DC: U. S. Government Printing Office.
- U. S. Employment Service (Hrsg.). (1970). *General aptitude test battery (GATB)*. Washington, DC: U. S. Employment Service.
- van Maanen, J. & Schein, E. H. (1977). Career development. In J. R. Hackman &

- J. L. Suttle (Hrsg.), *Improving life at work: Behavioral science approaches to organizational change* (S. 30-95). Santa Monica, CA: Goodyear Publishing Company.
- Vernon, P. E. (1950). *The structure of human abilities*. London: Methuen.
- Vinchur, A. J., Schippmann, J. S., Switzer, F. S. & Roth, P. L. (1998). A meta-analytic review of predictors of job performance for salespersons. *Journal of Applied Psychology*, *83*, 586-597.
- Viswesvaran, C., Ones, D. S. & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, *81*, 557-574.
- Viswesvaran, C., Schmidt, F. L. & Ones, D. S. (2002). The moderating influence of job performance dimensions on convergence of supervisory and peer ratings of job performance: Unconfounding construct-level convergence and rating difficulty. *Journal of Applied Psychology*, *87*, 345-354.
- Waldman, D. A. & Avolio, B. J. (1986). A meta-analysis of age differences in job performance. *Journal of Applied Psychology*, *71*, 33-38.
- Waldman, D. A. & Avolio, B. J. (1991). Race effects in performance evaluations: Controlling for ability, education, and experience. *Journal of Applied Psychology*, *76*, 897-901.
- Wayne, S. J., Liden, R. C., Kraimer, M. L. & Graf, I. K. (1999). The role of human capital, motivation, and supervisor sponsorship in predicting career success. *Journal of Organizational Behavior*, *20*, 577-595.
- Wechsler, D. (1956). *Die Messung der Intelligenz Erwachsener*. Bern: Huber.
- Wegener, B. (1988). *Kritik des Prestiges*. Opladen: Westdeutscher Verlag.
- *Weiß, R. H. (1969). *Die Brauchbarkeit des Culture Free Intelligence Tests Skala 3 (CFT 3) bei begabungspsychologischen Untersuchungen*. Unveröffentlichte Dissertation, Universität Würzburg.
- *Weiß, R. H. (1980). *Grundintelligenztest Skala 3 - CFT 3*. Braunschweig: Georg Westermann.
- Weiß, R. H. (1987). *Grundintelligenztest Skala 2 (CFT 20). Handanweisung für die Durchführung, Auswertung und Interpretation*. Göttingen: Hogrefe.
- Whitener, E. M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology*, *75*, 315-321.
- Wiegand, D. (1977). Untersuchung über die Aussagekraft des Ausleseverfahrens der Bundeswehr für Militärkraftfahrer. *Wehrpsychologische Untersuchungen*, *12*, 1-113.
- Wilk, S. L., Desmarais, L. B. & Sackett, P. R. (1995). Gravitation to jobs commensurate with ability: Longitudinal and cross-sectional tests. *Journal of Applied Psychology*, *80*, 79-85.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe.
- *Wolf, B. (1990a). Untersuchung zur Vorhersage des Ausbildungserfolges von Aufstiegsbewerbern für den gehobenen Polizeivollzugsdienst – Kurzbericht über eine

- Bewährungskontrolle. *DGP-Informationen*, 50, 51-66.
- Wolf, B. (1990b). *Zusammenhänge zwischen Ergebnissen von eignungsdiagnostischen Verfahren und Ausbildungserfolg – dargestellt am Beispiel der Auswahl von Bewerbern für den gehobenen Polizeivollzugsdienst*. Unveröffentlichte Diplomarbeit, Universität Göttingen.
- *Wolff, P. & Voullaire, C. (1968). Eignungsbegutachtung von Körperbehinderten für einen Verwaltungsberuf – eine Bewährungskontrolle. *Diagnostica*, 14, 70-87.
- Wonderlic, E. F. & Hovland, C. I. (1939). The personnel test: A restandardized abridgment of the Otis S-A test for business and industrial use. *Journal of Applied Psychology*, 685-702.
- Wonderlic Inc. (1996). *Wonderlic personnel test (WPT-German version, forms A and B)*. Libertyville, IL: Wonderlic.
- Wonderlic Inc. (2002). *Wonderlic personnel test & scholastic level exam user's manual*. Libertyville, IL: Wonderlic.
- Wonderlic Inc. (2007). *Wonderlic personnel test - revised (WPT-R)*. Zugriff am 20.1.2007 unter <http://www.wonderlic.com/products/selection/wpnr/>
- *Zeißig, H. J. (1989). *Berufsfindung und Arbeitserprobung der Berufsförderungswerke, Möglichkeiten ihrer Ergänzung hin zum Assessment-Center*. Unveröffentlichte Dissertation, Universität Bielefeld.

A. Formeln

A.1. Korrektur punktbiserialer Korrelationen

Die Korrektur der punktbiserialen Korrelation r_{pb} wird nach [Hunter und Schmidt \(2004\)](#) vorgenommen:

$$r_{\text{korr}} = \frac{ar_{pb}}{\sqrt{(a^2 - 1)r_{pb}^2 + 1}} \quad (\text{A.1})$$

Der Korrekturwert a wird aus den Anteilen (p und q) der beiden Gruppen am Gesamt- N der Studie berechnet:

$$a = \sqrt{\frac{.25}{pq}}$$

Zur Korrektur der Standardfehler bei Anwendung von Formel [A.1](#) wird ein reduziertes n nach [Hunter und Schmidt \(2004\)](#) verwendet:

$$n_{\text{red}} = \frac{r_{pb}^2(n - 1)}{r_{\text{korr}}^2} + 1 \quad (\text{A.2})$$

A.2. Kummulierung abhängiger Werte

Die Bildung linearer Verbände wird wie folgt vorgenommen:

$$r_{xy} = \frac{\mathbf{1}'\mathbf{R}_{xy}\mathbf{1}}{\sqrt{\mathbf{1}'\mathbf{R}_{xx}\mathbf{1}}\sqrt{\mathbf{1}'\mathbf{R}_{yy}\mathbf{1}}} \quad (\text{A.3})$$

\mathbf{R}_{xy} bezeichnet die Korrelationsmatrix zwischen den Prädiktoren und Kriterien, \mathbf{R}_{xx} die Matrix der Korrelationen zwischen den einzelnen Prädiktoren und \mathbf{R}_{yy} die Matrix der Korrelationen zwischen den einzelnen Kriterien ([Hunter & Schmidt, 2004](#)).

Unterschiedliche Gewichte der einzelnen Prädiktoren bzw. Kriterien werden berücksichtigt nach

$$r_{xy} = \frac{\mathbf{v}'\mathbf{R}_{xy}\mathbf{w}}{\sqrt{\mathbf{v}'\mathbf{R}_{xx}\mathbf{v}}\sqrt{\mathbf{w}'\mathbf{R}_{yy}\mathbf{w}}} \quad (\text{A.4})$$

wobei \mathbf{v} der Vektor der Prädiktorgewichte und \mathbf{w} der Vektor der Kriteriengewichte darstellt ([Hunter & Schmidt, 2004](#)).

Die Reliabilität des Prädiktorverbundes wird nach Mosier (1943) berechnet:

$$r_{xx} = \frac{\mathbf{1}'(\mathbf{R}_{xx} - \mathbf{D} + \mathbf{D}_{rel})\mathbf{1}}{\mathbf{1}'\mathbf{R}_{xx}\mathbf{1}} \quad (\text{A.5})$$

wobei $\mathbf{R}_{xx} - \mathbf{D} + \mathbf{D}_{rel}$ die Korrelationsmatrix der Subprädiktoren darstellt (\mathbf{R}_{xx}), bei der die Inhalte der Diagonale gelöscht ($-\mathbf{D}$) und durch die Reliabilitäten der Subprädiktoren ersetzt werden ($+\mathbf{D}_{rel}$; Hunter & Schmidt, 2004).

Unterschiedliche Gewichte können dabei ebenfalls berücksichtigt werden:

$$r_{xx} = \frac{\mathbf{v}'(\mathbf{R}_{xx} - \mathbf{D} + \mathbf{D}_{rel})\mathbf{w}}{\mathbf{v}'\mathbf{R}_{xx}\mathbf{w}} \quad (\text{A.6})$$

A.3. Berechnungen der bare-bone Analysen

Die mittlere beobachtete Korrelation (\bar{r}) und die beobachtete Standardabweichung der Korrelationen (SD_{beo}) werden wie folgt berechnet (Hunter & Schmidt, 2004):

$$\bar{r} = \frac{\Sigma(n_i r_i)}{N} \quad (\text{A.7})$$

$$SD_{beo} = \sqrt{\frac{\Sigma(n_i [r_i - \bar{r}]^2)}{N}} \quad (\text{A.8})$$

Die Berechnung der Standardabweichung, die der Stichprobenfehler erklärt (SD_{art}), erfolgt ebenfalls nach Hunter und Schmidt (2004):

$$SD_{art} = \sqrt{\frac{\Sigma(n_i SD_{art_i}^2)}{N}} \quad (\text{A.9})$$

Dafür sind die Varianzen erforderlich, die sich pro Studie durch den Stichprobenfehler erklären lassen ($SD_{art_i}^2$, Hunter & Schmidt, 2004):

$$SD_{art_i}^2 = \frac{(1 - \bar{r}_i^2)^2}{n_i - 1} \quad (\text{A.10})$$

Die verbleibende Standardabweichung von \bar{r} ($SD_{\bar{r}}$) wird wie folgt geschätzt (Hunter & Schmidt, 2004):

$$\widehat{SD}_{\bar{r}} = \sqrt{SD_{beo}^2 - SD_{art}^2} \quad (\text{A.11})$$

Der Anteil der durch den Stichprobenfehler aufgeklärten Varianz an der beobachteten Varianz (VE) beträgt dementsprechend

$$VE = SD_{art}^2 / SD_{beo}^2 \quad (\text{A.12})$$

Die Berechnung der Standardfehler, die zur Bestimmung der Konfidenzintervalle benötigt werden, unterscheidet sich im homogenen (VE > 75%) und heterogenen Fall (VE

< 75%). Im homogenen Fall werden \bar{r} , N und k benötigt (Whitener, 1990):

$$SE_{\bar{r}} = \sqrt{\frac{(1 - \bar{r}^2)^2}{N - k}} \quad (\text{A.13})$$

Im heterogenen Fall wird zusätzlich die geschätzte Heterogenitätsvarianz $\widehat{SD}_{\bar{r}}^2$ berücksichtigt (Whitener, 1990):

$$SE_{\bar{r}} = \sqrt{\frac{(1 - \bar{r}^2)^2}{N - k} + \frac{\widehat{SD}_{\bar{r}}^2}{k}} \quad (\text{A.14})$$

Die Anzahl der Studien, die erforderlich wären, um \bar{r} unter .10 zu drücken werden nach Hunter und Schmidt (2004) berechnet und aufgerundet:

$$fs_k = \left\lceil k \left(\frac{\bar{r}}{.10} - 1 \right) \right\rceil \quad (\text{A.15})$$

A.4. Operationale Validität und vollständig korrigierte Analysen

A.4.1. Berechnungen auf Studienebene

Zur Berechnung von $\hat{\rho}$ werden die berichteten Korrelationen zunächst individuell auf Studienebene mit dem Verfahren der *indirekten* Varianzkorrektur von Hunter und Schmidt (2004) korrigiert. Dies erfolgt in mehreren Schritten. Zunächst werden die Reliabilitäten der varianzeingeschränkten Stichproben (r_{xx_i}) berechnet:

$$r_{xx_i} = 1 - U_x^2(1 - r_{xx}) \quad (\text{A.16})$$

dabei sind $U_x = 1/u_x$ und r_{xx} die Reliabilitäten der Intelligenztestverfahren auf Basis der Testmanuale (nicht varianzeingeschränkt). Anschließend erfolgt die Korrektur der Unreliabilitäten von Prädiktor und Kriterium:

$$\hat{\rho}_{\text{TP}} = \frac{r}{\sqrt{r_{xx_i} r_{yy}}} \quad (\text{A.17})$$

Schließlich wird die *wahre* Varianzeinschränkung korrigiert, die zunächst auf Basis der beobachteten (u_x) berechnet wird (Hunter & Schmidt, 2004):

$$u_T = \sqrt{\frac{u_x^2 - (1 - r_{xx})}{r_{xx}}} \quad (\text{A.18})$$

Zur Korrektur der Varianzeinschränkung wird u_T invertiert zu U_T und gemäß Hunter und Schmidt (2004) wie folgt verfahren:

$$\hat{\rho}_i = \frac{U_T \hat{\rho}_{\text{TP}}}{\sqrt{1 + U_T^2 \hat{\rho}_{\text{TP}}^2 - \hat{\rho}_{\text{TP}}^2}} \quad (\text{A.19})$$

Zur Bestimmung der operationalen Validität wird die Korrektur der Unreliabilität des Prädiktors wieder rückgängig gemacht:

$$\hat{\rho}_{\text{opi}} = \hat{\rho}_i \sqrt{r_{\text{xx}}} \quad (\text{A.20})$$

Das Gewicht der Studien in der Metaanalyse wird wie folgt bestimmt:

$$w = n(r/\hat{\rho}_i)^2 \quad (\text{A.21})$$

w_i wird in den Tabellen E.1 bis E.42, die die Ergebnisse auf Studienebene beinhalten, im Verhältnis zum Gesamtgewicht aller Studien angegeben. Zur Berechnung der Metaanalysen wird ferner die Varianz des Stichprobenfehlers der korrigierten Korrelation benötigt, die nach Hunter und Schmidt (2004) aus der Varianz der unkorrigierten Korrelation errechnet werden kann ($\text{Var}_{e_0} = SD_{\text{arti}}^2$ der bare-bone Analyse; vgl. Formel A.10):

$$SD_{\text{arti}}^2 = \frac{\text{Var}_{e_0}}{(r/\hat{\rho}_i)^2} \left(\frac{1}{(U_{\text{T}}^2 - 1)r^2 + 1} \right)^2 \quad (\text{A.22})$$

A.4.2. Berechnungen über alle Studien hinweg

Die Schätzung von $\hat{\rho}$ erfolgt nach Hunter und Schmidt (2004):

$$\hat{\rho} = \frac{\Sigma(w_i \hat{\rho}_i)}{\Sigma w_i} \quad (\text{A.23})$$

Die beobachtete Standardabweichung wird entsprechend folgender Formel bestimmt:

$$SD_{\text{beo}} = \sqrt{\frac{\Sigma(w_i [\hat{\rho}_i - \hat{\rho}]^2)}{\Sigma w_i}} \quad (\text{A.24})$$

und die Standardabweichung, die durch die Artefaktkorrekturen aufgeklärt wird, berechnet sich wie folgt:

$$SD_{\text{art}} = \sqrt{\frac{\Sigma(w_i SD_{\text{arti}}^2)}{\Sigma w_i}} \quad (\text{A.25})$$

Die verbleibende Standardabweichung von $\hat{\rho}$ wird mit Hilfe von SD_{beo} und SD_{art} geschätzt:

$$\widehat{SD}_{\rho} = \sqrt{SD_{\text{beo}}^2 - SD_{\text{art}}^2} \quad (\text{A.26})$$

Zur Berechnung der Konfidenzintervalle wird der Standardfehler von ρ nach Hunter und Schmidt (2004) wie folgt bestimmt:

$$SE_{\rho} = \frac{SD_{\text{beo}}}{\sqrt{k}} \quad (\text{A.27})$$

Die untere Grenze des 90% Glaubwürdigkeitsintervalls wird mittels \widehat{SD}_ρ berechnet (Hunter & Schmidt, 2004):

$$CV_{\text{low}} = \hat{\rho} - (1.645\widehat{SD}_\rho) \quad (\text{A.28})$$

A.5. Symmetriekennwerte für Funnelplots

Ob Funnelplots symmetrisch sind, wird mit zwei Kennwerten von Duval (2005) geprüft. Dazu werden die $|r_i|$ -Werte in eine Rangfolge gebracht. Die Anzahl der höchsten Ränge, die zu positiven Korrelationen gehören und über dem höchsten Rang der negativen Korrelationen liegen, wird mit y^* bezeichnet. Auf diesem Kennwert basiert das erste Symmetriemaß:

$$R_0 = y^* - 1 \quad (\text{A.29})$$

Das zweite Symmetriemaß basiert auf der Summe aller Ränge, die zu positiven Korrelationen gehören (S_{rank}), es wird zur Anzahl aller Korrelationen (k) in Beziehung gesetzt:

$$L_0 = \frac{4S_{\text{rank}} - k(k+1)}{2k-1} \quad (\text{A.30})$$

R_0 und L_0 werden auf die nächste *positive* Zahl aufgerundet (R_0^+ bzw. L_0^+) und geben dann an, wieviele Studien im 1. Schritt zu exkludieren sind, um einen symmetrischeren Funnelplot zu erhalten. Die Berechnungen werden so lange wiederholt, bis sich R_0^+ bzw. L_0^+ nicht mehr verändern.

B. Kodieranweisung

B.1. Studien-Kodierbogen

Grundsätzliches

- Das Interesse gilt dem Zusammenhang zwischen GMA, Subskalen von GMA und verschiedenen Maßen beruflichen Erfolgs (Validitätskoeffizienten, VK). Was dazu alles gehört wird näher spezifiziert in folgenden Variablen:
 1. Zusammenhangsmaße: VN, V, VTYP
 2. Prädiktoren: IQ
 3. Kriterien: KINHALT, KMM
- Um die Kodierung möglichst ökonomisch zu gestalten, ist wie folgt vorzugehen:
 1. Zunächst wird in den Quellen nach den berichteten VK zwischen GMA und Leistung gesucht (Variablen vom vorherigen Punkt in der Reihenfolge 1-3).
 2. Sind diese gefunden, werden davon ausgehend die restlichen Variablen kodiert.
- Die Kodierung wird in einer Exceltabelle vorgenommen entsprechend den Anweisungen in dieser Kodieranweisung. Spalten stehen für Variablen, Zeilen für Fälle. Ein Fall ist eine berichtete Korrelation zum Zusammenhang zwischen GMA und Leistung.
- Das Kürzel *ps* kennzeichnet Variablen, für die pro *Primärstudie* (PS) nur *eine* Kodierung vorgenommen werden kann. Eine Kodierung pro berichteter Korrelation ist *nicht* erforderlich.
- Es ist darauf zu achten, dass in numerischen Spalten keine Texte eingegeben werden. Sollen zu numerischen Spalten Anmerkungen (z.B. auch Seitenzahlen, damit die Inhalte schnell wiedergefunden werden können) gemacht werden, werden diese in einer separaten Word-Datei aufgeführt. Für einige Variablen sind auch spezielle Bemerkungsspalten in der Exceltabelle vorgesehen.
- Bei Text-Variablen können die Seitenangaben mit eingegeben werden.
- Bei der Excel-Kodierung können verschiedene Funktionen von Excel genutzt werden, die die Arbeit sehr erleichtern, z.B. farbige Hervorhebungen, das Ausblenden und wieder Einblenden von Spalten, das Fixieren der Variablenbezeichnungen (Spalten) und Studiennummern (Zeilen).
- Wenn im folgenden von *Prädiktoren* die Rede ist, sind Intelligenzmaße gemeint. Auch wenn diese konkurrent oder retrograd erhoben wurden.

Allgemeine Variablen

NR (Text, ps)

- Jede unabhängige PS bekommt eine eindeutige ID-Nummer (diese wird zwischen den Kodierern abgestimmt). Alle unabhängigen PS werden als eigene Studien angesehen!
- unabhängige PS bedeutet: es liegt eine *statistisch* unabhängige Stichprobe vor. Anders formuliert: alle statistisch abhängigen Stichproben werden als *eine* PS kodiert. Die statistische Unabhängigkeit bezieht sich auf die berichteten VK zw. GMA und Leistung. Beispiele:
 1. In einer Studie werden VK getrennt für Männer und Frauen berichtet, aber nicht für beide Geschlechter zusammen. Die zwei Stichproben (Männer + Frauen) sind unabhängig voneinander.
 2. Werden darüber hinaus auch Gesamtvaliditäten berichtet, ist folgendermaßen vorzugehen: *Eine* Studie definieren mit den drei Stichproben: Gesamt, Männer + Frauen.
 3. Ergebnisse zu verschiedenen Berufen werden wenn möglich als eigene Studien behandelt, auch wenn Gesamtergebnisse über mehrere Berufe hinweg berichtet sind, da sie später auch getrennt analysiert werden sollen.
- Statistisch abhängige Stichproben kann es auch zwischen verschiedenen Quellen geben: dann nämlich, wenn die Ergebnisse der gleichen Stichprobe öfters veröffentlicht wurden!
- Die NR besteht aus einer dreistelligen Nummer für die Quelle(n), gefolgt von einer zweistelligen Nummer für die Studie, beide werden mit einem Punkt getrennt. Beispiel: Studie 100.01 stammt aus Quelle(n) 100 und ist dort die erste Studie die berichtet wird. Wird in einer Quelle nur *eine* Studie berichtet, lautet die zweistellige Nummer 00.

SUBNR (numerisch)

- Mehrere abhängige Werte *einer* PS können in mehreren Zeilen abgelegt werden. Die Zeilen werden in dieser Variable durchnummeriert.
- Die 1. Zeile enthält Daten, die für die gesamte Primärstudie gültig sind! Die Zeile wird farblich hervorgehoben und bekommt die SUBNR = 0. Alle folgenden Zeilen werden von 1-n fortlaufend durchnummeriert.
- Wenn eine Variable für eine PS insgesamt gilt, kann sie einmal in der Zeile SUBNR = 0 kodiert werden. Es ist nicht erforderlich, die Kodierung in allen SUBNR > 0 -Zeilen zu wiederholen.

- Die Reihenfolge der zu kodierenden Korrelationen ist möglichst sinnvoll zu wählen. I.d.R. ist das die Reihenfolge, in der die Werte in der Quelle berichtet werden. Ggf. ist hier eine Absprache zw. den Kodierern zu treffen um das spätere Zusammenführen der Daten zu erleichtern.

PLABEL (Text, ps)

- Label des Quelle aus der die PS stammt. Es wird den Kodierern mitgeteilt.
- Das Label besteht i.d.R. aus dem Nachnamen des Erstautors, dem Publikationsjahr und evtl. einem Buchstaben
- Stammen die Ergebnisse aus mehreren Quellen, werden auch mehrere PLABEL eingetragen.

AU1 (Text, ps)

- Erstautor in der Form: Nachname, Vorname

AU2 (Text, ps)

- Zweitautor in der Form: Nachname, Vorname

AU3 (Text, ps)

- weitere Autoren in der Form: Nachname, Vorname; Nachname, Vorname; ...

PJAHR (intervall, ps)

- Publikationsjahr vierstellig, z.B. 1990
- bei mehreren veröffentlichten Quellen wird das früheste Veröffentlichungsjahr genannt
- bei nicht veröffentlichten Quellen das Jahr ihrer Fertigstellung bzw. Zusammenstellung
- Das PJAHR geht aus dem PLABEL hervor.

PART (numerisch, ps)

- Art der Publikation:
 - 1 = nicht veröffentlicht, z.B. Diplomarbeiten, graue Literatur, Folien von Kongressvorträgen, Literatur, die zur Veröffentlichung vorbereitet wird
 - 2 = Dissertation

- 3 = Zeitschriftenartikel
- 4 = Monographie (kein Testmanual)
- 5 = Beitrag in Herausgeberwerk
- 6 = Testmanual
- 9 = unklar

- Zur Kodierung dieser Variablen die Literaturquelle heranziehen!

PSLABEL (Text, ps)

- Bei Quellen, die VKs für mehrere unabhängige Stichproben (also mehrere Studien) berichten, kann ein beliebiger Label für jede Studie vergeben werden.

SEX (numerisch, ps)

- Geschlecht der Stichprobe:
 - 1 = weiblich
 - 2 = männlich
 - 3 = gemischt
 - 9 = unklar
- Bei Bergbau-Berufen und bei älteren PS mit militärischen Pb, die Dienst an der Waffe tun, kann männlich kodiert werden, auch wenn keine Angaben dazu gemacht werden.

WEIBL (absolut, ps)

- Bei SEX = 3: Frauenanteil in der Studie in Prozent.

MAGEX (absolut, ps)

- Mittleres Alter der SP in Jahren zum Zeitpunkt der Prädiktorerhebung

SDAGEX (absolut, ps)

- Standardabweichung der SP in Jahren zum Zeitpunkt der Prädiktorerhebung

AGEXBEM (Text, ps)

- Bemerkungsfeld zu MAGEX und SDAGEX
- Hier vermerken, wenn MAGEX Median ist oder auf ungenauen Angaben beruht, etc.

SBNIV1 (absolut, ps)

- Anteil der Stichprobe mit einem allgemeinbildenden Schulabschluss bis einschließlich Hauptschulniveau.
- Umfasst auch Personen ohne Schulabschluss, Sonderschulabschluss und Realschulabgänger ohne Realschulabschluss

SBNIV2 (absolut, ps)

- Anteil der Stichprobe mit einem allgemeinbildenden Schulabschluss der der mittleren Reife entspricht.

SBNIV3 (absolut, ps)

- Anteil der Stichprobe mit einem allgemeinbildenden Schulabschluss der der Fachhochschul- oder allgemeinen Hochschulreife entspricht in Prozent
- Es gilt: $SBNIV1 + SBNIV2 + SBNIV3 = 100$, was an 100 fehlt sind fehlende Angaben.

BERUF (Text, ps)

- Berufsbezeichnung, die in der PS verwendet wird
- Bei mehreren Berufen: wenn möglich alle Bezeichnungen angeben und mit Prozentangaben versehen.
- Bei Aus- und Weiterbildungen den angestrebten Beruf benennen!

ISCO88 (numerisch, ps)

- vierstelliger Tätigkeitsschlüssel laut ISCO-88
- Die erste Stelle gibt die Berufshauptkategorie an.
- Die ISCO-88-Kodierung wird in folgenden Schritten vorgenommen:
 1. [Geis \(2006\)](#)
 2. [International Labour Office \(1990\)](#)
 3. BERUFENET-Datenbank der Bundesagentur für Arbeit (www.berufenet.de)
- Einige Hinweise aus [Geis \(2006\)](#) sind hier noch einmal aufgelistet:
 - Für Verwaltungsberufe im Öffentlichen Dienst sind die Ausführungen auf S. 10-14 zu Rate zu ziehen.

- Andere Berufe im Öffentlichen Dienst, z.B. Feuerwehr oder Polizei sind wie folgt zu kodieren: mittlerer Dienst mit 5161 Feuerwehr, 5162 Polizei (einen „einfachen Dienst“ gibt es hier nicht; nur z.B. nicht-beamtete Hilfspolizisten, die mit 9152 kodiert werden). Gehobener Dienst bei der Polizei 3450, bei der Feuerwehr 3490; höherer Dienst 2000-1000.
- Meister = 7900
- Die ISCO-Kodierung ist wenn möglich bis zu den Vierstellern durchzuführen. Für die Metaanalyse entscheidend sind aber nur die Einsteller. Bei Unklarheiten können die Stellen 2-4 deshalb mit „0“ aufgefüllt werden.

ISCOSL (ordinal, ps)

- Skill Level nach ISCO-88
- Wird ausgehend von ISCO88 kodiert. Die erste Stelle des ISCO-88-Kodes ergibt den Skill Level wie folgt:
 - 0 = ISCO-88-Hauptgruppe 1 (leitende Tätigkeiten) und ISCO-88-Kode 0110 (Militär)
 - 1 = ISCO-88-Hauptgruppe 9 (ungelernte Hilfsarbeiter)
 - 2 = ISCO-88-Hauptgruppe 4-8 (angelernte Arbeiter, Ausbildungsberufe bis zum Meister). D.h. wenn keine Klarheit über die ISCO-88-Hauptgruppen 4-8 besteht, kann der Skill Level trotzdem kodiert werden!
 - 3 = ISCO-88-Hauptgruppe 3 (Ebene der Techniker)
 - 4 = ISCO-88-Hauptgruppe 2 (Wissenschaftler und Ingenieure außer: graduierte Ingenieure, sie können nicht eindeutig einem Skill Level zugeordnet werden)

ERF (absolut, ps)

- Berufserfahrung in *Monaten*
- Diese Variable ist nur bei Berufserfolgsstudien (KINHALT = 1) zu kodieren, nicht bei Aus- bzw. Weiterbildungen, da sie sich auf den *angestrebten* Beruf bezieht.
- Die Berufserfahrung enthält *nicht* die Ausbildungs- oder Weiterbildungszeit, sondern wird erst dann gezählt, wenn der Beruf erlernt ist.
- ggf. (gewichteten) Mittelwert bilden, wenn ERF für einzelne Substichproben berichtet ist.

VTYP (numerisch)

- Typ der Validität:
 - 1 = Längsschnittstudie: Intelligenz trug *nicht* zur SP-Auswahl bei. D.h. es fand bei der Personalauswahl *kein* Intelligenztest statt oder es wird ausdrücklich vermerkt, dass Intelligenzwerte keinen Einfluss auf die Personalauswahl hatten.
 - 2 = Längsschnittstudie: Intelligenz trug zur SP-Auswahl bei. Dies ist auch dann zu kodieren, wenn nicht der gleiche Prädiktor verwendet wurde, aber einer, der hoch korreliert ist (z.B. also ein anderer Intelligenztests). Es geht darum, festzustellen, ob eine Varianzeinschränkung für den Prädiktor vorliegt.
 - 3 = Längsschnittstudie: Rolle der Intelligenz bei der Personalauswahl unklar.
 - 4 = Querschnittstudie
 - 9 = unklar ob Längs- oder Querschnittstudie

Variablen zum Prädiktor**IQ (numerisch)**

- Intelligenztest:
 - 1 = IST
 - 2 = WIT
 - 3 = BIS
 - 4 = BOMAT
 - 5 = HAWIE
 - 6 = LPS/PSB
 - 7 = Matrizentests (CFT, Raven)
 - 8 = GATB-basierte Verfahren (BET, EUB)
 - 9 = andere

PM (absolut)

- Mittelwert des Prädiktors (Intelligenz)

PSD (absolut)

- Standardabweichung des Prädiktors

PVE (Text)

- Informationen zur Varianzeinschränkung im Prädiktor, z.B. die Selektionsquote

PIK (Text, ps)

- Informationen zur Korrelation verschiedener Prädiktoren untereinander

PREL (Text)

- Informationen zur Reliabilität des Prädiktors:
 - Höhe der Reliabilität
 - Art des Reliabilitätskoeffizienten
 - nähere Informationen (z.B. Zeitabstand bei Retest)
 - Reliabilitätswerte auf Basis der Primärstudie angegeben oder aus dritter Quelle z.B. einem Testmanual wiedergegeben?

Variablen zum Kriterium

KINHALT (numerisch)

- Inhalt der Kriterien:
 - 1 = Arbeitsleistung oder Karriererfolgsmaße
 - 2 = Ausbildungsleistung (Erstausbildungsberuf)
 - 3 = Weiterbildungsleistung (z.B. Weiterbildung zum Meister)
 - 8 = anderes, Kombination
 - 9 = unklar

KMM (numerisch)

- Kriterien-Messmethode
- subjektive Maße:
 - 0 = schulische Abwärtsbeurteilung in Form von Noten oder Punkten bei Ausbildungs- oder Weiterbildungsleistung (Berufsschulnoten; kann Abwärtsbeurteilungen im Rahmen der Ausbildung enthalten)
 - 1 = betriebliche Abwärtsbeurteilung (Vorgesetzte beurteilen ihre Mitarbeiter) für betriebliche Beurteilungen bei Ausbildungs-, Weiterbildungs- oder Arbeitsleistung.
 - 2 = Seitwärtsbeurteilung (Kollegen beurteilen sich untereinander)
 - 3 = andere subjektive Leistungsbeurteilung (z.B. Selbstbeurteilung, Aufwärtsbeurteilung, Kombination aus den Punkten 1 und 2)
- objektive Maße:
 - 4 = Fehlzeiten

- 5 = Produktionsdaten (z.B. Akkord; nur bei KINHALT = 1 möglich)
- 6 = Einkommen (z.B. Lohn, Gehalt; nur bei KINHALT = 1 möglich)
- 7 = Beförderung (z.B. Anzahl Beförderungen, Hierarchiestufe; nur bei KINHALT = 1 möglich)
- 8 = anderes objektives Maß
- subjektive und objektive Maße:
 - 9 = Kombination aus subjektiven und objektiven Maßen oder unklar

KM (absolut)

- Mittelwert des Kriteriums

KSD (absolut)

- Standardabweichung des Kriteriums

KVE (Text)

- Informationen zur Varianzeinschränkung im Kriterium

KIK (Text, ps)

- Informationen zur Korrelation verschiedener Kriterien untereinander?

KREL (Text)

- Informationen zur Reliabilität des Kriteriums:
 - Höhe der Reliabilität
 - Art des Reliabilitätskoeffizienten
 - nähere Informationen (z.B. Anzahl der Rater bei Beurteilerübereinstimmung, Zeitabstand bei Retest)
 - Reliabilitätswerte auf Basis der Primärstudie angegeben oder aus dritter Quelle bzw. anderer Stichprobe wiedergegeben

Variablen zu den Validitätskoeffizienten**X (Text)**

- Prädiktor in Worten (angelehnt an die Wortwahl der Quelle)

Y (Text)

- Kriterium in Worten (angelehnt an die Wortwahl der Quelle)

VALLAB (Text)

- Beliebiger Label, der für den Validitätskoeffizienten vergeben werden kann, um leichter die Übersicht zu behalten.
- Dieses Label kann auch dazu genutzt werden, verschiedene Substichproben kenntlich zu machen, z.B. Ergebnisse für Frauen und Männer.

TDIFF (absolut)

- Zeitabstand Prädiktor-Kriterium-Erhebung in *Monaten*
- Generell gilt: es wird der bestmögliche Schätzer für TDIFF gesucht und im Zweifelsfalle konservativ vorgegangen (kleinerer der in Frage kommenden Zeitabstände wird kodiert). Folgende Regeln sind zu beachten:
 1. Bsp. dreijährige Ausbildung, Kriterium am Ausbildungsende erhoben. In der Quelle wird darauf hingewiesen, dass es die Möglichkeit gab, um 1/2 Jahr zu verkürzen: TDIFF = 2,5 Jahre = 18 Monate
 2. Mehrere Substichproben, die alle unterschiedliche TDIFFs aufweisen: Gesamt-TDIFF = (gewichtetes) Mittel aus den Substichproben.
 3. Wenn ein Kriterium über einen Zeitraum kumuliert (z.B. Berufsschulnoten sich aus den Ergebnissen mehrere Jahre zusammensetzen oder das Mittel betrieblicher Beurteilungen zu verschiedenen Zeitpunkten als Kriterium verwendet wird), wird der gesamte Zeitraum angegeben über den sich die Kriterien erstrecken, d.h. der maximale.

TDIFFKR (ordinal)

- Konfidenzrating, das angibt, mit welcher Sicherheit TDIFF kodiert wurde:
 - 1 = unsicher: bei Schätzungen anhand anderer Quellen, z.B. vergleichbarer Studien
 - 2 = wahrscheinlich: bei Schätzungen anhand anderer Quellen (z.B. bei Leistungsstudien den Berufenet-Angaben, vergleichbaren Studien aus der selben Organisation etc.), wobei es wahrscheinlich ist, dass der Zeitabstand auch für die vorliegende Studie gilt
 - 3 = ziemlich sicher: Zeitraum in der Quelle grob angegeben
 - 4 = sicher: Zeitraum konkret und unmissverständlich angegeben

V (absolut)

- der unkorrigierte Validitätskoeffizient, auf den sich alle anderen Angaben in der zugehörigen Excel-Zeile beziehen!

- Jede erhobene Validität, die für die Analyse relevant ist, bekommt eine Zeile! Dies gilt auch für Validitäten, die wgn. Insignifikanz nicht berichtet werden. Dann wird die entsprechende V-Zelle als Missing-Value behandelt und bleibt leer. Keine Null eingeben, da Null für eine berichtete Nullkorrelation steht.
- Für alle Validitäten gilt: Kodierung in Richtung der Hypothese, d.h. Intelligenz trägt zum Berufserfolg positiv bei! Gerade bei Schulnoten werden auch negative Korrelationen berichtet, da ein geringer Notenwert (1) hohem Erfolg (sehr gut) entspricht. Bei Unklarheiten in der Word-Datei darauf hinweisen.
- Wenn keine Zusammenhänge angegeben werden, können auch Unterschiedsmaße (t -Werte) hier angegeben werden (vgl. VKTYP).

VN (absolut)

- Die Stichprobengröße, auf der der kodierte Validitätskoeffizient basiert, d.h. ohne Drop Outs, nur komplette Daten kodieren!
- Bei kodierten t -Werten (vgl. VKTYP) sind hier die Freiheitsgrade (df) anzugeben.

VKTYP (numerisch)

- Typ des Validitätskoeffizienten:
 - 1 = Pearson Produkt-Moment
 - 2 = Rang
 - 3 = Biserial
 - 4 = Tetrachisch
 - 5 = Punktbiserial
 - 6 = Phi
 - 7 = Kontingenz
 - 8 = multiple
 - 9 = unklar
 - 10 = t -Wert (nur kodieren, wenn kein r angegeben!, t -Wert selbst dann in V kodieren und df in N!)
 - 11 = χ^2 -Wert
 - 12 = korrigierte Korrelationen (nähere Beschreibung der Art der Korrelation und der vorgenommenen Korrekturen bei KTINFO)
- I.d.R. werden Produkt-Moment-Korrelationen und Rangkorrelationen berichtet. Zur Unterscheidung auf Skalenniveaus der Prädiktoren und Kriterien achten (vgl. Diehl & Kohr, 2004, S. 261).

KTINFO

- Informationen zur Validitätskorrektur, z.B. welcher Korrelationstyp, welche Korrekturformeln und Korrekturmaße verwendet wurden

B.2. BIS-Kodierbogen

Jede Aufgabengruppe ist – wenn möglich – den operativen *und* inhaltsgebundenen BIS-Facetten zuzuordnen. Zur Kodierung werden die Beschreibungen der Fähigkeitskonstrukte aus Jäger et al. (1997) herangezogen:

Operative Fähigkeiten:

- K *Verarbeitungskapazität*: Verarbeitung komplexer Informationen bei Aufgaben, die nicht auf Antrieb zu lösen sind, sondern Heranziehen, vielfältiges Beziehungsstiften, formallogisches exaktes Denken und sachgerechtes Beurteilen von Informationen erfordern.
- E *Einfallreichum*: Flexible Ideenproduktion, die Verfügbarkeit vielfältiger Informationen, Reichtum an Vorstellungen und das sehen vieler verschiedener Seiten, Varianten, Gründe und Möglichkeiten von Gegenständen und Problemen voraussetzt, wobei es um problemorientierte Lösungen geht, nicht um ein ungesteuertes Luxurieren der Phantasie.
- M *Merkfähigkeit*: Aktives Einprägen und kurzfristiges Wiedererkennen oder Reproduzieren von verschiedenartigem Material. Die frühere Bezeichnung Gedächtnis wurde geändert, da im Verlauf der Modellentwicklung nur noch kurzfristige Behaltensleistungen einbezogen werden konnten.
- B *Bearbeitungsgeschwindigkeit*: Arbeitstempo, Auffassungsleichtigkeit und Konzentrationskraft beim Lösen einfach strukturierter Aufgaben von geringem Schwierigkeitsniveau.

Inhaltsgebundene Fähigkeiten:

- V *Sprachgebundenes Denken*: Grad der Aneignung und der Verfügbarkeit des Beziehungssystems Sprache.
- N *Zahlengebundenes Denken*: Grad der Aneignung und der Verfügbarkeit des Beziehungssystems Zahlen.
- F *Anschaunungsgebundenes, figural-bildhaftes Denken*: Einheitsstiftendes Merkmal scheint hier die Eigenart des Aufgabenmaterials zu sein, dessen Bearbeitung figural-bildhaftes und/oder räumliches Vorstellen erfordert. (S. 6-7)

Darüber hinaus können die Aufgaben des BIS-4 als Beispielaufgaben herangezogen werden und ggf. die Ergebnisse empirischer Untersuchungen zum Zusammenhang der zu

kodierenden Testaufgabe mit den BIS-Facetten. Es geht aber nicht primär um eine empirisch fundierte Einschätzung, sondern um eine theoretische: Welche Zuordnung sollte sich in empirischen Studien auf Grund der Aufgabenbeschreibungen zeigen?

Die Abgrenzung der K- von der B-Operation kann schwierig sein, besonders bei Aufgaben mit Zeitbegrenzung. K-Aufgaben sind komplexer als B-Aufgaben und dauern i.d.R. länger. Kenntnisse von Regeln bzw. Heuristiken sind erforderlich. Evtl. sind mehrere Lösungswege möglich bzw. mehrere Entscheidungen zu treffen, d.h. die Lösung erfolgt in mehreren Schritten. B-Aufgaben sind vergleichsweise einfach und schnell zu lösen und bestehen i.d.R. nicht aus mehreren Schritten.

Ist eine Zuordnung mit Hilfe dieser Anweisungen nicht eindeutig möglich, können auch mehrere operative oder mehrere inhaltsgebundene Facetten vergeben werden. Die Nennung mehrere Operationen oder Inhalte kann bedeuten, dass die Aufgabengruppe mehrere Bereiche abdeckt (z.B. K *und* B) oder dass unklar ist welche der genannten Aufgabengruppen abgedeckt ist (z.B. K *oder* B). Ist z.B. von einer Aufgabengruppe bekannt, dass sie aus numerischen Items besteht, die Operationen sind aber nicht zu erschließen, so können a) alle vier Operationen angegeben werden, wenn klar ist, dass es sich um eine klassische Intelligenztestaufgabe handelt oder b) es kann „–“ vergeben werden, wenn klar ist, dass sie außerhalb des BIS-Modells liegt (Kontamination). Das entsprechende gilt für nicht eindeutig zu kodierende Inhalte.

C. Kodierungen

C.1. Quellen und Studiennummern

Tabelle C.1

Zuordnung der Quellen und Stichproben zu den Studiennummern

Studien-Nr.	Quellen	Stichprobe (Beruf)
100.00	Althoff (1975, 1977)	Beamter höherer Dienst Kripo
101.01	Althoff (1968)	RegierungsinspektorIn
101.02	Althoff (1968)	Revierförster
102.00	Althoff (1974)	(Verwaltungs-)Inspektoren
103.00	Graudenz (1982)	Beamter gehobener Dienst
104.00	Graudenz (1984)	Beamter mittlerer Dienst allgemein Justizvollzug
105.00	Kleinevoss (1983)	Beamter gehobener Dienst
106.00	Kuhl und Greve (1971)	Metall- und Elektroberufe für Haupt- und Sonderschulabsolventen
107.01	Seggebruch (1984)	Gehobener allgemeiner Verwaltungsdienst (Inspektoren, $n = 56$)
107.02	Seggebruch (1984)	Gehobener allgemeiner Verwaltungsdienst (Inspektoren, $n = 137$)
107.03	Seggebruch (1984)	Gehobener allgemeiner Verwaltungsdienst (Inspektoren, $n = 145$)
108.00	Wolf (1990a)	Schutz- und Kriminalpolizei (gehobener Dienst)
109.00	Brandstätter (1970)	Rechtspfleger
110.00	Bretz und Oldendörp (1992)	Angestellte Stadt Hannover: Angestellten-Lehrgang I
111.01	Weiß (1969)	gewerblich, kaufmännische Ausbildung
111.02	Weiß (1969)	gewerbliche Ausbildung
112.00	Derner (1983)	Metall-Facharbeiter; Schwerpunkte Fertigungstechnik, Metallbau, Installationstechnik
113.00	Deter und Herbig (1992)	Chemikanten
114.00	Diepenbrock und Thielepape (1973)	Meister: Schmied, Tischler, Werkzeugmacher, Polsterer, usw.
115.00	Dreyer und Grabitz (1971)	Steiger
116.01	Amthauer (1973)	Ingenieure grad.
116.02	Amthauer (1973)	Industriekaufleute
116.03	Amthauer (1973)	mathematisch-technische Assistenten
116.04	Amthauer (1973)	technisch handwerkliche Beruf
116.05	Amthauer (1973)	Betriebsschlosser
116.06	Amthauer (1973)	Laborfacharbeiter
116.07	Amthauer (1973)	Chemielaboranten ($n = 70$)
116.08	Amthauer (1973)	Chemielaboranten ($n = 304$)
116.09	Amthauer (1973)	Chemielaboranten ($n = 169$)
116.10	Amthauer (1973)	Chemiefacharbeiter
116.11	Amthauer (1973)	Bürogehilfin, Kontoristin
116.12	Amthauer (1973)	Chemielaboranten ($n = 597$)
116.13	Amthauer (1973)	Ingenieur grad.
117.00	Jäger (1960)	REFA-Grundlehrgang
118.01	Jungkunz und Bodinet (1989)	Friseur
118.02	Jungkunz und Bodinet (1989)	Tischler
118.03	Jungkunz und Bodinet (1989)	KFZ-Mechaniker
119.01	Kober (1978)	Chemielaboranten
119.02	Kober (1978)	Chemiefacharbeiter
120.01	Marschner (1964)	Grubensteiger ($n = 127$)
120.02	Marschner (1964)	Grubensteiger ($n = 101$)
120.03	Marschner (1964)	Grubensteiger ($n = 222$)
120.04	Marschner (1964)	Grubensteiger ($n = 13$)
122.00	Marschner (1978)	Facharbeiter Elektro/Metall; gewerblich, handwerklich, technische Berufsausbildung
123.01	Neumann (1938)	Fräser
123.02	Neumann (1938)	Bohrer

Tabelle C.1
(Fortsetzung)

Studien-Nr.	Quellen	Stichprobe (Beruf)
123.03	Neumann (1938)	Dreher
123.04	Neumann (1938)	Bohrwerksdreher
123.05	Neumann (1938)	Schleifer
123.06	Neumann (1938)	Kontrolleure
124.01	Roloff (1928)	Maschinenbauer ($n = 12$)
124.02	Roloff (1928)	Maschinenbauer ($n = 8$)
124.03	Roloff (1928)	Dreher
125.01	Schmidt-Atzert und Deter (1993)	Chemikanten
125.02	Schmidt-Atzert und Deter (1993)	Chemielaboranten
125.03	Schmidt-Atzert und Deter (1993)	Industriekaufleute
125.04	Schmidt-Atzert und Deter (1993)	Büroassistentinnen
125.05	Schmidt-Atzert und Deter (1993)	Maschinenschlosser, Betriebsschlosser, Mechaniker, Hochdruck-Rohrschlosser, Kessel- und Behälterbauer, Kunststoffschlosser
125.06	Schmidt-Atzert und Deter (1993)	Energieanlagenelektroniker, Energiegeräteelektroniker, Mess- und Regelmechaniker
126.01	Schmidt-Atzert, Deter und Jaeckel (2004)	Chemikanten
126.02	Schmidt-Atzert et al. (2004)	Chemie- und Physikalabornanten
126.03	Schmidt-Atzert et al. (2004)	Energieelektroniker, Industrieelektroniker, Mess- und Regelmechaniker
126.04	Schmidt-Atzert et al. (2004)	Industrie- und Anlagenmechaniker
126.05	Schmidt-Atzert et al. (2004)	Kaufleute für Bürokommunikation
126.06	Schmidt-Atzert et al. (2004)	Industriekaufleute
127.01	Barthel et al. (1982); Fünfgelt, Barthel und Schuler (1984); Schuler, Barthel und Fünfgelt (1980); Schuler et al. (1984)	Naachrichtengerätetechniker
127.02	Barthel et al. (1982); Fünfgelt et al. (1984); Schuler et al. (1980, 1984)	Werkzeugmacher, Mechaniker
128.00	Schuler et al. (1991); Schuler, Funke et al. (1995)	Forschung und Entwicklung
129.00	Seggebruch (1982)	mittlerer Dienst Kommunalverwaltung
130.00	Thielepape (1980)	Krankenschwestern und -pfleger
132.00	Lometsch (1975)	Management-Nachwuchskräfte
133.00	Althoff (1978)	Vermessungstechniker
134.00	Wolf und Voullaire (1968)	Reha-Lehrgang, Verwaltungsberuf
135.00	Hossiep (1995)	Bankkaufleute
136.00	Kersting (1999a, 2001)	gehobener und höherer Dienst Kripo und Schupo
137.01	Schäfer (1986)	Chemielaborant
137.02	Schäfer (1986)	Chemikant
137.03	Schäfer (1986)	Industriekaufleute
137.04	Schäfer (1986)	Büroassistentinnen
137.05	Schäfer (1986)	Mess- und Regelmechaniker
137.06	Schäfer (1986)	Betriebsschlosser
137.07	Schäfer (1986)	Industrieinformatiker
138.00	Balck et al. (1979)	Bankkaufleute
139.00	Biernat und Scholz (1995)	Beamter des mittleren feuerwehrtechnischen Dienstes
140.00	Kettel und Simmat (1969)	Fernmeldehandwerker
141.00	Marschner (1966)	Gruben-Fahrhauer-Meister
142.00	Marschner (1972)	Nachwuchs für mittleres Management (Steiger)
143.01	Marschner (1981)	Elektroanlageninstallateur
143.02	Marschner (1981)	Betriebsschlosser
143.03	Marschner (1981)	Bergmechaniker
144.00	Marschner (1983a)	Elektroanlageninstallateur, Betriebsschlosser, Bergmechaniker und Berg- und Maschinenmänner
145.00	Nettelstroth (2004)	verschiedene Berufe: Dienstleistung, Industrie, öffentliche Verwaltung
146.00	J. U. Schmidt (1987)	gehobener Verwaltungsdienst
147.00	Greif (1972)	Bereitschaftspolizei

Tabelle C.1
(Fortsetzung)

Studien-Nr.	Quellen	Stichprobe (Beruf)
148.00	Kersting (2003); Krause, Kersting und Heggestad (2006); Krause, Kersting, Heggestad und Thornton (2005)	Schutz-, Wasser- und Kriminalpolizisten; höherer Dienst
149.00	Damitz et al. (2000)	Fluglotsen
150.00	Herzog (2004)	Marinesoldaten
151.00	Steinmann (1997)	Bankkaufleute
152.00	Schuler und Klingner (2002, 2005)	Büro-, Bank-, Informatikkaufleute, Fachinformatiker, Dipl.-Betriebswirt (BA), dto. Volkswirt
153.00	Görlisch und Schuler (2004)	Industriemechaniker, -elektroniker, Mechatroniker
156.00	Knebelau (2002)	Verwaltungsinspektoren
159.00	Kersting et al. (2005)	Reha Vorbereitungslehrgang
160.01	Thielepape und Kersting (2005a, 2005b)	Kommunalverwaltung gehobener Dienst
160.02	Thielepape und Kersting (2005a, 2005b)	Landes- und Kommunalverwaltung gehobener Dienst
161.01	U. Funke (1986)	biologisch-technischer Assistent und chemisch-technischer Assistent ($n = 60$)
161.02	U. Funke (1986)	biologisch-technischer Assistent und chemisch-technischer Assistent ($n = 31$)
162.01	Paff (1963)	Apparatewärter
162.02	Paff (1963)	Maschinisten
163.00	Kreuscher (1987)	selbständige Vermögensberater im Außendienst
164.01	Salzgitter Service und Technik GmbH (2004)	Industriemechaniker, Zerspanungsmechaniker
164.02	Salzgitter Service und Technik GmbH (2004)	Industrieelektroniker, Produktionselektroniker versch. Fachrichtungen
165.00	Dahle und Erdmann (2001)	Strafgefangene in Berliner Forensik
167.00	Paff (1966)	Schlosser
168.00	Ehlers (1961)	metallgewerbl. Lehrlinge Hüttenwerk Salzgitter AG
169.00	Amelang und Steinmayr (2006a, 2006b)	verschiedene
170.00	Blickle et al. (2005)	verschiedene
171.01	A. A. Schmitt (2006)	Mechaniker
171.02	A. A. Schmitt (2006)	Elektriker/Elektroniker
171.03	A. A. Schmitt (2006)	Fertigungsmechaniker
171.04	A. A. Schmitt (2006)	Mechatroniker
172.00	Schmitz (2006)	Bankkaufleute
173.00	Awiszus (1967)	gewerbliche Berufe: vor allem Metall-, Bau-, Elektroberufe, wenige Nahrungs- und Kunstberufe
174.00	Schuler, Moser et al. (1995)	Bankkaufleute
175.00	Jungkunz (1995); Jungkunz und John (1991)	KFZ-Mechaniker (39.3%), Tischler (30.1%), Friseur (29.6%)
176.01	Lang (1990)	Energieelektroniker
176.02	Lang (1990)	Industriemechaniker
177.01	Schwadorf (2003)	Bankkaufleute (29.2%), Verkäufer/Einzelhandelskaufleute (28.2%), Industriekaufleute (27.9%), Steuerfachangestellte (14.7%)
177.02	Schwadorf (2003)	BA Bank (26.3%), Handel (10.1%), Industrie (28.3%), Steuern (35.4%)
178.00	Hollstegge (1994)	gehobener Dienst des Landes NRW
179.01	Zeißig (1989)	Bürokaufleute ($n = 12$)
179.02	Zeißig (1989)	Industriekaufleute ($n = 15$)
179.03	Zeißig (1989)	Bankkaufleute ($n = 9$)
179.04	Zeißig (1989)	Nachrichtengerätetechniker
179.05	Zeißig (1989)	Elektroniker ($n = 7$)
179.06	Zeißig (1989)	Hörgeräteakustiker ($n = 15$)
179.07	Zeißig (1989)	Augenoptiker ($n = 14$)
179.08	Zeißig (1989)	Bürokaufleute ($n = 51$)
179.09	Zeißig (1989)	Industriekaufleute ($n = 64$)
179.10	Zeißig (1989)	Bankkaufleute ($n = 23$)

Tabelle C.1
(Fortsetzung)

Studien-Nr.	Quellen	Stichprobe (Beruf)
179.11	Zeißig (1989)	Elektroniker ($n = 26$)
179.12	Zeißig (1989)	Hörgeräteakustiker ($n = 20$)
179.13	Zeißig (1989)	Augenoptiker ($n = 52$)
179.14	Zeißig (1989)	Masseure und med. Bademeister
181.01	Streitkräfteamt (1996a)	Offizierbewerber bei der Bundeswehr: Heer
181.02	Streitkräfteamt (1996a)	Offizierbewerber bei der Bundeswehr: Luftwaffe
181.03	Streitkräfteamt (1996a)	Offizierbewerber bei der Bundeswehr: Marine
181.04	Streitkräfteamt (1996a)	Offizierbewerber bei der Bundeswehr: Sanität
182.00	Streitkräfteamt (2001)	Offizierbewerber bei der Bundeswehr
183.00	Streitkräfteamt (1996b)	Gehobener Dienst in der Bundeswehrverwaltung (Zivilpersonal)
184.01	Schmidt-Atzert und Eser (2006)	verschiedene ($n = 152$)
184.02	Schmidt-Atzert und Eser (2006)	Mechaniker
184.03	Schmidt-Atzert und Eser (2006)	verschiedene ($n = 89$)
185.00	Linz (1985)	19 verschiedene Berufe in neun Berufsfeldern (BBW Nürnberg)
186.00	Streitkräfteamt der Bundeswehr (2004)	Unteroffiziere Marine
187.01	Marcus et al. (2002)	hauptsächlich Verkäufer, auch Back-Office
187.02	Marcus et al. (2002)	verschiedene
188.00	Blickle et al. (2006)	verschiedene
189.01	Schmale und Schmidtke (2001)	Betriebswirt (BA)
189.02	Schmale und Schmidtke (2001)	Bankkaufleute
189.03	Schmale und Schmidtke (2001)	Versicherungskaufleute
189.04	Schmale und Schmidtke (2001)	Sozialversicherungsfachangestellte (SOFA)
189.05	Schmale und Schmidtke (2001)	DV-Kaufleute
189.06	Schmale und Schmidtke (2001)	Industriekaufleute
189.07	Schmale und Schmidtke (2001)	Groß- und Außenhandelskaufleute
189.08	Schmale und Schmidtke (2001)	Einzelhandelskaufleute
189.09	Schmale und Schmidtke (2001)	Bürokaufmann Verwaltungsfachangestellter
189.10	Schmale und Schmidtke (2001)	Bürogehilfin
189.11	Schmale und Schmidtke (2001)	Phonotypistin
189.12	Schmale und Schmidtke (2001)	Kaufleute für Bürokommunikation und Büroorganisation: Fachrichtung Computergesteuerte Textverarbeitung
189.13	Schmale und Schmidtke (2001)	Kaufleute für Bürokommunikation und Büroorganisation: Fachrichtung Personalverwaltung
189.14	Schmale und Schmidtke (2001)	Maschinenschlosser
189.15	Schmale und Schmidtke (2001)	Betriebsschlosser
189.16	Schmale und Schmidtke (2001)	Bauschlosser/Stahlbaus Schlosser
189.17	Schmale und Schmidtke (2001)	KFZ-Schlosser
189.18	Schmale und Schmidtke (2001)	Rohrleitungsbauer
189.19	Schmale und Schmidtke (2001)	Werkzeugmacher/Feinmechaniker
189.20	Schmale und Schmidtke (2001)	Dreher
189.21	Schmale und Schmidtke (2001)	Fräser
189.22	Schmale und Schmidtke (2001)	Schweißer
189.23	Schmale und Schmidtke (2001)	Teilezurichter
189.24	Schmale und Schmidtke (2001)	Automateneinrichter
189.25	Schmale und Schmidtke (2001)	Feinblechner/Kupferschmied
189.26	Schmale und Schmidtke (2001)	Möbeltischler/Holzmechaniker
189.27	Schmale und Schmidtke (2001)	Kunstfaserschlosser
189.28	Schmale und Schmidtke (2001)	Mess- und Regelmechaniker
189.29	Schmale und Schmidtke (2001)	Starkstromelektriker
189.30	Schmale und Schmidtke (2001)	Nachrichtengerätemechaniker
189.31	Schmale und Schmidtke (2001)	Informationselektroniker
189.32	Schmale und Schmidtke (2001)	Energieanlagen-Installateur
189.33	Schmale und Schmidtke (2001)	Energieanlagen-Elektroniker
189.34	Schmale und Schmidtke (2001)	Fachkraft für Lagerwirtschaft
189.35	Schmale und Schmidtke (2001)	Schauwerbegestalter
189.36	Schmale und Schmidtke (2001)	Augenoptiker (Handwerk)
189.37	Schmale und Schmidtke (2001)	Reproduktionsfotograf
189.38	Schmale und Schmidtke (2001)	Tiefdruckretuscheur

Tabelle C.1
(Fortsetzung)

Studien- Nr.	Quellen	Stichprobe (Beruf)
189.39	Schmale und Schmidtke (2001)	Tiefdrucker
189.40	Schmale und Schmidtke (2001)	Hochdrucker (Buchdruck)
189.41	Schmale und Schmidtke (2001)	Flachdrucker (Offset, Farblitho)
189.42	Schmale und Schmidtke (2001)	Maschinensetzer
189.43	Schmale und Schmidtke (2001)	Handsetzer
189.44	Schmale und Schmidtke (2001)	Korrektor
189.45	Schmale und Schmidtke (2001)	Bauzeichner
189.46	Schmale und Schmidtke (2001)	Technischer Zeichner
189.47	Schmale und Schmidtke (2001)	Vermessungstechniker
189.48	Schmale und Schmidtke (2001)	Biologielaborant
189.49	Schmale und Schmidtke (2001)	Physiklaborant
189.50	Schmale und Schmidtke (2001)	Chemielaborant
189.51	Schmale und Schmidtke (2001)	Chemiekant (Chemiefacharbeiter)
189.52	Schmale und Schmidtke (2001)	Textillaborant
189.53	Schmale und Schmidtke (2001)	Industriemechaniker: Fachrichtung Produktionstechnik
189.54	Schmale und Schmidtke (2001)	Industriemechaniker: Fachrichtung Betriebstechnik
189.55	Schmale und Schmidtke (2001)	Industriemechaniker: Fachrichtung Maschinen- und Systemtechnik
189.56	Schmale und Schmidtke (2001)	Industriemechaniker: Fachrichtung Geräte- und Feinwerktechnik
189.57	Schmale und Schmidtke (2001)	Werkzeugmechaniker
189.58	Schmale und Schmidtke (2001)	Zerspanungsmechaniker
189.59	Schmale und Schmidtke (2001)	Konstruktionsmechaniker
189.60	Schmale und Schmidtke (2001)	Anlagenmechaniker: Fachrichtung Apparatetechnik
189.61	Schmale und Schmidtke (2001)	Anlagenmechaniker: Fachrichtung Versorgungstechnik
189.62	Schmale und Schmidtke (2001)	Automobilmechaniker
189.63	Schmale und Schmidtke (2001)	Elektromaschinenmonteur
189.64	Schmale und Schmidtke (2001)	Energieelektroniker: Fachrichtung Anlagentechnik
189.65	Schmale und Schmidtke (2001)	Energieelektroniker: Fachrichtung Betriebstechnik
189.66	Schmale und Schmidtke (2001)	Industrieelektroniker: Fachrichtung Produktionstechnik
189.67	Schmale und Schmidtke (2001)	Industrieelektroniker: Fachrichtung Gerätetechnik
189.68	Schmale und Schmidtke (2001)	Kommunikationselektroniker
189.69	Schmale und Schmidtke (2001)	IT-Systemkaufmann
189.70	Schmale und Schmidtke (2001)	Fachinformatiker

Tabelle C.2

Zuordnung der Studiennummern zu den inkludierten Quellen und Stichproben

Quellen	Stichprobe (Beruf)	Studien-Nr.
Althoff (1968)	RegierungsinspektorIn	101.01
Althoff (1968)	Revierförster	101.02
Althoff (1974)	(Verwaltungs-)Inspektoren	102.00
Althoff (1975, 1977)	Beamter höherer Dienst Kripo	100.00
Althoff (1978)	Vermessungstechniker	133.00
Amelang und Steinmayr (2006a, 2006b)	verschiedene	169.00
Amthauer (1973)	Ingenieure grad.	116.01
Amthauer (1973)	Industriekaufleute	116.02
Amthauer (1973)	mathematisch-technische Assistenten	116.03
Amthauer (1973)	technisch handwerkliche Beruf	116.04
Amthauer (1973)	Betriebsschlosser	116.05
Amthauer (1973)	Laborfacharbeiter	116.06
Amthauer (1973)	Chemielaboranten ($n = 70$)	116.07
Amthauer (1973)	Chemielaboranten ($n = 304$)	116.08
Amthauer (1973)	Chemielaboranten ($n = 169$)	116.09
Amthauer (1973)	Chemiefacharbeiter	116.10
Amthauer (1973)	Bürogehilfin, Kontoristin	116.11
Amthauer (1973)	Chemielaboranten ($n = 597$)	116.12
Amthauer (1973)	Ingenieur grad.	116.13
Awiszus (1967)	gewerbliche Berufe: vor allem Metall-, Bau-, Elektroberufe, wenige Nahrungs- und Kunstberufe	173.00
Balck et al. (1979)	Bankkaufleute	138.00
Barthel et al. (1982)	Naachrichtengerätetechniker	127.01
Barthel et al. (1982)	Werkzeugmacher, Mechaniker	127.02
Biernat und Scholz (1995)	Beamter des mittleren feuerwehrtechnischen Dienstes	139.00
Blickle et al. (2005)	verschiedene	170.00
Blickle et al. (2006)	verschiedene	188.00
Brandstätter (1970)	Rechtspfleger	109.00
Bretz und Oldendorp (1992)	Angestellte Stadt Hannover: Angestellten-Lehrgang I	110.00
Dahle und Erdmann (2001)	Strafgefangene in Berliner Forensik	165.00
Damitz et al. (2000)	Fluglotsen	149.00
Derner (1983)	Metall-Facharbeiter; Schwerpunkte Fertigungstechnik, Metallbau, Installationstechnik	112.00
Deter und Herbig (1992)	Chemikanten	113.00
Diepenbrock und Thielepape (1973)	Meister: Schmied, Tischler, Werkzeugmacher, Polsterer, usw.	114.00
Dreyer und Grabitz (1971)	Steiger	115.00
Ehlers (1961)	metallgewerbl. Lehrlinge Hüttenwerk Salzgitter AG	168.00
Fünfgelt et al. (1984)	Naachrichtengerätetechniker	127.01
Fünfgelt et al. (1984)	Werkzeugmacher, Mechaniker	127.02
U. Funke (1986)	biologisch-technischer Assistent und chemisch-technischer Assistent ($n = 60$)	161.01
U. Funke (1986)	biologisch-technischer Assistent und chemisch-technischer Assistent ($n = 31$)	161.02
Görlich und Schuler (2004)	Industriemechaniker, -elektroniker, Mechatroniker	153.00
Graudenz (1982)	Beamter gehobener Dienst	103.00
Graudenz (1984)	Beamter mittlerer Dienst allgemein Justizvollzug	104.00
Greif (1972)	Bereitschaftspolizei	147.00
Herzog (2004)	Marinesoldaten	150.00
Hollstegge (1994)	gehobener Dienst des Landes NRW	178.00
Hossiep (1995)	Bankkaufleute	135.00
Jäger (1960)	REFA-Grundlehrgang	117.00
Jungkunz (1995)	KFZ-Mechaniker (39.3%), Tischler (30.1%), Friseur (29.6%)	175.00
Jungkunz und Bodinet (1989)	Friseur	118.01
Jungkunz und Bodinet (1989)	Tischler	118.02
Jungkunz und Bodinet (1989)	KFZ-Mechaniker	118.03
Jungkunz und John (1991)	KFZ-Mechaniker (39.3%), Tischler (30.1%), Friseur (29.6%)	175.00
Kersting (1999a, 2001)	gehobener und höherer Dienst Kripo und Schupo	136.00
Kersting (2003)	Schutz-, Wasser- und Kriminalpolizisten; höherer Dienst	148.00

Tabelle C.2
(Fortsetzung)

Quellen	Stichprobe (Beruf)	Studien-Nr.
Kersting et al. (2005)	Reha Vorbereitungslehrgang	159.00
Kettel und Simmat (1969)	Fernmeldehandwerker	140.00
Kleinevoss (1983)	Beamter gehobener Dienst	105.00
Knebelau (2002)	Verwaltungsinspektoren	156.00
Kober (1978)	Chemielaboranten	119.01
Kober (1978)	Chemiefacharbeiter	119.02
Krause et al. (2005, 2006)	Schutz-, Wasser- und Kriminalpolizisten; höherer Dienst	148.00
Kreuscher (1987)	selbständige Vermögensberater im Außendienst	163.00
Kuhl und Greve (1971)	Metall- und Elektroberufe für Hauptüind Sonderschulabsolventen	106.00
Lang (1990)	Energieelektroniker	176.01
Lang (1990)	Industriemechaniker	176.02
Linz (1985)	19 verschiedene Berufe in neun Berufsfeldern (BBW Nürnberg)	185.00
Lometsch (1975)	Management-Nachwuchskräfte	132.00
Marcus et al. (2002)	hauptsächlich Verkäufer, auch Back-Office	187.01
Marcus et al. (2002)	verschiedene	187.02
Marschner (1964)	Grubensteiger ($n = 127$)	120.01
Marschner (1964)	Grubensteiger ($n = 101$)	120.02
Marschner (1964)	Grubensteiger ($n = 222$)	120.03
Marschner (1964)	Grubensteiger ($n = 13$)	120.04
Marschner (1966)	Gruben-Fahrhauer-Meister	141.00
Marschner (1972)	Nachwuchs für mittleres Management (Steiger)	142.00
Marschner (1978)	Facharbeiter Elektro/Metall; gewerblich, handwerklich, technische Berufsausbildung	122.00
Marschner (1981)	Elektroanlageninstallateur	143.01
Marschner (1981)	Betriebsschlosser	143.02
Marschner (1981)	Bergmechaniker	143.03
Marschner (1983a)	Elektroanlageninstallateur, Betriebsschlosser, Bergmechaniker und Berg- und Maschinenmänner	144.00
Nettelstroth (2004)	verschiedene Berufe: Dienstleistung, Industrie, öffentliche Verwaltung	145.00
Neumann (1938)	Fräser	123.01
Neumann (1938)	Bohrer	123.02
Neumann (1938)	Dreher	123.03
Neumann (1938)	Bohrwerksdreher	123.04
Neumann (1938)	Schleifer	123.05
Neumann (1938)	Kontrolleure	123.06
Paff (1963)	Apparatewärter	162.01
Paff (1963)	Maschinisten	162.02
Paff (1966)	Schlosser	167.00
Roloff (1928)	Maschinenbauer ($n = 12$)	124.01
Roloff (1928)	Maschinenbauer ($n = 8$)	124.02
Roloff (1928)	Dreher	124.03
Salzgitter Service und Technik GmbH (2004)	Industriemechaniker, Zerspanungsmechaniker	164.01
Salzgitter Service und Technik GmbH (2004)	Industrieelektroniker, Produktionselektroniker versch. Fachrichtungen	164.02
Schäfer (1986)	Chemielaborant	137.01
Schäfer (1986)	Chemikant	137.02
Schäfer (1986)	Industriekaufleute	137.03
Schäfer (1986)	Büroassistentinnen	137.04
Schäfer (1986)	Mess- und Regelmechaniker	137.05
Schäfer (1986)	Betriebsschlosser	137.06
Schäfer (1986)	Industrieinformatiker	137.07
Schmale und Schmidtke (2001)	Betriebswirt (BA)	189.01
Schmale und Schmidtke (2001)	Bankkaufleute	189.02
Schmale und Schmidtke (2001)	Versicherungskaufleute	189.03
Schmale und Schmidtke (2001)	Sozialversicherungsfachangestellte (SOFA)	189.04
Schmale und Schmidtke (2001)	DV-Kaufleute	189.05
Schmale und Schmidtke (2001)	Industriekaufleute	189.06
Schmale und Schmidtke (2001)	Groß- und Außenhandelskaufleute	189.07

Tabelle C.2
(Fortsetzung)

Quellen	Stichprobe (Beruf)	Studien-Nr.
Schmale und Schmidtke (2001)	Einzelhandelskaufleute	189.08
Schmale und Schmidtke (2001)	Bürokaufmann Verwaltungsfachangestellter	189.09
Schmale und Schmidtke (2001)	Bürogehilfin	189.10
Schmale und Schmidtke (2001)	Phonotypistin	189.11
Schmale und Schmidtke (2001)	Kaufleute für Bürokommunikation und Büroorganisation: Fachrichtung Computergesteuerte Textverarbeitung	189.12
Schmale und Schmidtke (2001)	Kaufleute für Bürokommunikation und Büroorganisation: Fachrichtung Personalverwaltung	189.13
Schmale und Schmidtke (2001)	Maschinenschlosser	189.14
Schmale und Schmidtke (2001)	Betriebsschlosser	189.15
Schmale und Schmidtke (2001)	Bauschlosser/Stahlbauschlosser	189.16
Schmale und Schmidtke (2001)	KFZ-Schlosser	189.17
Schmale und Schmidtke (2001)	Rohrleitungsbauer	189.18
Schmale und Schmidtke (2001)	Werkzeugmacher/Feinmechaniker	189.19
Schmale und Schmidtke (2001)	Dreher	189.20
Schmale und Schmidtke (2001)	Fräser	189.21
Schmale und Schmidtke (2001)	Schweißer	189.22
Schmale und Schmidtke (2001)	Teilezurichter	189.23
Schmale und Schmidtke (2001)	Automateneinrichter	189.24
Schmale und Schmidtke (2001)	Feinblechner/Kupferschmied	189.25
Schmale und Schmidtke (2001)	Möbeltischler/Holzmechaniker	189.26
Schmale und Schmidtke (2001)	Kunstfaserschlosser	189.27
Schmale und Schmidtke (2001)	Mess- und Regelmechaniker	189.28
Schmale und Schmidtke (2001)	Starkstromelektriker	189.29
Schmale und Schmidtke (2001)	Nachrichtengerätetechniker	189.30
Schmale und Schmidtke (2001)	Informationselektroniker	189.31
Schmale und Schmidtke (2001)	Energieanlagen-Installateur	189.32
Schmale und Schmidtke (2001)	Energieanlagen-Elektroniker	189.33
Schmale und Schmidtke (2001)	Fachkraft für Lagerwirtschaft	189.34
Schmale und Schmidtke (2001)	Schauerbegehalter	189.35
Schmale und Schmidtke (2001)	Augenoptiker (Handwerk)	189.36
Schmale und Schmidtke (2001)	Reproduktionsfotograf	189.37
Schmale und Schmidtke (2001)	Tiefdruckretuscheur	189.38
Schmale und Schmidtke (2001)	Tiefdrucker	189.39
Schmale und Schmidtke (2001)	Hochdrucker (Buchdruck)	189.40
Schmale und Schmidtke (2001)	Flachdrucker (Offset, Farblitho)	189.41
Schmale und Schmidtke (2001)	Maschinensetzer	189.42
Schmale und Schmidtke (2001)	Handsetzer	189.43
Schmale und Schmidtke (2001)	Korrektor	189.44
Schmale und Schmidtke (2001)	Bauzeichner	189.45
Schmale und Schmidtke (2001)	Technischer Zeichner	189.46
Schmale und Schmidtke (2001)	Vermessungstechniker	189.47
Schmale und Schmidtke (2001)	Biologielaborant	189.48
Schmale und Schmidtke (2001)	Physiklaborant	189.49
Schmale und Schmidtke (2001)	Chemielaborant	189.50
Schmale und Schmidtke (2001)	Chemikant (Chemiefacharbeiter)	189.51
Schmale und Schmidtke (2001)	Textillaborant	189.52
Schmale und Schmidtke (2001)	Industriemechaniker: Fachrichtung Produktionstechnik	189.53
Schmale und Schmidtke (2001)	Industriemechaniker: Fachrichtung Betriebstechnik	189.54
Schmale und Schmidtke (2001)	Industriemechaniker: Fachrichtung Maschinen- und Systemtechnik	189.55
Schmale und Schmidtke (2001)	Industriemechaniker: Fachrichtung Geräte- und Feinwerktechnik	189.56
Schmale und Schmidtke (2001)	Werkzeugmechaniker	189.57
Schmale und Schmidtke (2001)	Zerspanungsmechaniker	189.58
Schmale und Schmidtke (2001)	Konstruktionsmechaniker	189.59
Schmale und Schmidtke (2001)	Anlagenmechaniker: Fachrichtung Apparatechnik	189.60
Schmale und Schmidtke (2001)	Anlagenmechaniker: Fachrichtung Versorgungstechnik	189.61
Schmale und Schmidtke (2001)	Automobilmechaniker	189.62
Schmale und Schmidtke (2001)	Elektromaschinenmonteur	189.63
Schmale und Schmidtke (2001)	Energieelektroniker: Fachrichtung Anlagentechnik	189.64

Tabelle C.2
(Fortsetzung)

Quellen	Stichprobe (Beruf)	Studien-Nr.
Schmale und Schmidtke (2001)	Energieelektroniker: Fachrichtung Betriebstechnik	189.65
Schmale und Schmidtke (2001)	Industrieelektroniker: Fachrichtung Produktionstechnik	189.66
Schmale und Schmidtke (2001)	Industrieelektroniker: Fachrichtung Gerätetechnik	189.67
Schmale und Schmidtke (2001)	Kommunikationselektroniker	189.68
Schmale und Schmidtke (2001)	IT-Systemkaufmann	189.69
Schmale und Schmidtke (2001)	Fachinformatiker	189.70
J. U. Schmidt (1987)	gehobener Verwaltungsdienst	146.00
Schmidt-Atzert und Deter (1993)	Chemikanten	125.01
Schmidt-Atzert und Deter (1993)	Chemielaboranten	125.02
Schmidt-Atzert und Deter (1993)	Industriekaufleute	125.03
Schmidt-Atzert und Deter (1993)	Büroassistentinnen	125.04
Schmidt-Atzert und Deter (1993)	Maschinenschlosser, Betriebsschlosser, Mechaniker, Hochdruck-Rohr-schlosser, Kessel- und Behälterbauer, Kunststoffschlosser	125.05
Schmidt-Atzert und Deter (1993)	Energieanlagenelektroniker, Energiegeräteelektroniker, Mess- und Regelmechaniker	125.06
Schmidt-Atzert et al. (2004)	Chemikanten	126.01
Schmidt-Atzert et al. (2004)	Chemie- und Physiklaboranten	126.02
Schmidt-Atzert et al. (2004)	Energieelektroniker, Industrieelektroniker, Mess- und Regelmechaniker	126.03
Schmidt-Atzert et al. (2004)	Industrie- und Anlagenmechaniker	126.04
Schmidt-Atzert et al. (2004)	Kaufleute für Bürokommunikation	126.05
Schmidt-Atzert et al. (2004)	Industriekaufleute	126.06
Schmidt-Atzert und Eser (2006)	verschiedene ($n = 152$)	184.01
Schmidt-Atzert und Eser (2006)	Mechaniker	184.02
Schmidt-Atzert und Eser (2006)	verschiedene ($n = 89$)	184.03
A. A. Schmitt (2006)	Mechaniker	171.01
A. A. Schmitt (2006)	Elektriker/Elektroniker	171.02
A. A. Schmitt (2006)	Fertigungsmechaniker	171.03
A. A. Schmitt (2006)	Mechatroniker	171.04
Schmitz (2006)	Bankkaufleute	172.00
Schuler et al. (1980, 1984)	Naachrichtengerätetechniker	127.01
Schuler et al. (1980, 1984)	Werkzeugmacher, Mechaniker	127.02
Schuler et al. (1991); Schuler, Funke et al. (1995)	Forschung und Entwicklung	128.00
Schuler und Klingner (2002, 2005)	Büro-, Bank-, Informatikkaufleute, Fachinformatiker, Dipl.-Betriebs-wirt (BA), dto. Volkswirt	152.00
Schuler, Moser et al. (1995)	Bankkaufleute	174.00
Schwadorf (2003)	Bankkaufleute (29.2%), Verkäufer/Einzelhandelskaufleute (28.2%), In-dustriekaufleute (27.9%), Steuerfachangestellte (14.7%)	177.01
Schwadorf (2003)	BA Bank (26.3%), Handel (10.1%), Industrie (28.3%), Steuern (35.4%)	177.02
Seggebruch (1982)	mittlerer Dienst Kommunalverwaltung	129.00
Seggebruch (1984)	Gehobener allgemeiner Verwaltungsdienst (Inspektoren, $n = 56$)	107.01
Seggebruch (1984)	Gehobener allgemeiner Verwaltungsdienst (Inspektoren, $n = 137$)	107.02
Seggebruch (1984)	Gehobener allgemeiner Verwaltungsdienst (Inspektoren, $n = 145$)	107.03
Steinmann (1997)	Bankkaufleute	151.00
Streitkräfteamt (1996a)	Offizierbewerber bei der Bundeswehr: Heer	181.01
Streitkräfteamt (1996a)	Offizierbewerber bei der Bundeswehr: Luftwaffe	181.02
Streitkräfteamt (1996a)	Offizierbewerber bei der Bundeswehr: Marine	181.03
Streitkräfteamt (1996a)	Offizierbewerber bei der Bundeswehr: Sanität	181.04
Streitkräfteamt (1996b)	Gehobener Dienst in der Bundeswehrverwaltung (Zivilpersonal)	183.00
Streitkräfteamt (2001)	Offizierbewerber bei der Bundeswehr	182.00
Streitkräfteamt (2004)	Unteroffiziere Marine	186.00
Thielepape (1980)	Krankenschwestern und -pfleger	130.00
Thielepape und Kersting (2005a, 2005b)	Kommunalverwaltung gehobener Dienst	160.01
Thielepape und Kersting (2005a, 2005b)	Landes- und Kommunalverwaltung gehobener Dienst	160.02
Weiß (1969)	gewerblich, kaufmännische Ausbildung	111.01

Tabelle C.2
(Fortsetzung)

Quellen	Stichprobe (Beruf)	Studien-Nr.
Weiß (1969)	gewerbliche Ausbildung	111.02
Wolf (1990a)	Schutz- und Kriminalpolizei (gehobener Dienst)	108.00
Wolff und Voullaire (1968)	Reha-Lehrgang, Verwaltungsberuf	134.00
Zeifig (1989)	Bürokaufleute ($n = 12$)	179.01
Zeifig (1989)	Industriekaufleute ($n = 15$)	179.02
Zeifig (1989)	Bankkaufleute ($n = 9$)	179.03
Zeifig (1989)	Nachrichtengerätetechniker	179.04
Zeifig (1989)	Elektroniker ($n = 7$)	179.05
Zeifig (1989)	Hörgeräteakustiker ($n = 15$)	179.06
Zeifig (1989)	Augenoptiker ($n = 14$)	179.07
Zeifig (1989)	Bürokaufleute ($n = 51$)	179.08
Zeifig (1989)	Industriekaufleute ($n = 64$)	179.09
Zeifig (1989)	Bankkaufleute ($n = 23$)	179.10
Zeifig (1989)	Elektroniker ($n = 26$)	179.11
Zeifig (1989)	Hörgeräteakustiker ($n = 20$)	179.12
Zeifig (1989)	Augenoptiker ($n = 52$)	179.13
Zeifig (1989)	Masseure und med. Bademeister	179.14

C.2. Übersicht der Metaanalysen pro Studie

Tabelle C.3

Übersicht über die Metaanalysen pro Studie

Studien-Nr.	Metaanalysen zur berufsbezogenen Lernleistung ...																											Arbeit		Karriere														
	Gesamt publiziert	nicht publiziert	Testmanuale	Artikel	Monographien	ohne BET-Manual	BET-Manual	vor 1945	nach 1945	prädiktiv	konkurrent	Skill Level 2	Skill Level 3	max. Hauptschule	Hochschulreife	Berufs-HG 0	Berufs-HG 3	Berufs-HG 4	Berufs-HG 7	Berufs-HG 8	IST	WIT	BIS	LPS/PSB	Matrizen	GATB-basiert	anderer IQ-Test	Schule	Betrieb	Ausbildung	Weiterbildung	Gesamt publiziert	prädiktiv	Pearson-r	Einkommen	Vorankommen	Vorank. Pearson-r							
100.00	1	1	0	0	1	0	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0		
101.01	1	1	0	0	0	1	1	0	0	1	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	1	1	0	1	1	1	1	1	0	0	0	0			
101.02	1	1	0	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	1	1	1	1	1	0	0	0	0			
102.00	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0		
103.00	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	1	0	1	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0		
104.00	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0		
105.00	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	1	0	1	0	0	0	0	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0		
106.00	1	1	0	0	1	0	1	0	0	1	1	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0		
107.01	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	1	0	1	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0		
107.02	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	1	0	1	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0		
107.03	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	1	0	1	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
108.00	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	1	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
109.00	1	1	0	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	1	1	1	1	1	0	0	0	0	0		
110.00	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
111.01	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	
111.02	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	
112.00	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	
113.00	1	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	
114.00	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
115.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
116.01	1	1	0	1	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
116.02	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
116.03	1	1	0	1	0	0	1	0	0	1	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
116.04	1	1	0	1	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
116.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
116.06	1	1	0	1	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
116.07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
116.08	1	1	0	1	0	0	1	0	0	1	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
116.09	1	1	0	1	0	0	1	0	0	1	1	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
116.10	1	1	0	1	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
116.11	1	1	0	1	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
116.12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
116.13	1	1	0	1	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
117.00	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
118.01	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
118.02	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
118.03	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
119.01	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
119.02	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
120.01	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
120.02	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
120.03	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
120.04	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0	1																										

Tabelle C.3
(Fortsetzung)

Studien-Nr.	Metaanalysen zur berufsbezogenen Lernleistung ...																												Arbeit		Kar-																		
	Gesamt publiziert	nicht publiziert	Testmanuale	Artikel	Monographien	ohne BET-Manual	BET-Manual vor 1945	nach 1945	prädiaktiv	konkurrent	Skill Level 2	Skill Level 3	max. Hauptschule	Hochschulreife	Berufs-HG 0	Berufs-HG 3	Berufs-HG 4	Berufs-HG 7	Berufs-HG 8	IST	WIT	BIS	LPS/PSB	Matrizen	GATB-basiert	anderer IQ-Test	Schule	Betrieb	Ausbildung	Weiterbildung	Gesamt publiziert	prädiaktiv	Pearson-r	Einkommen	Vorankommen	Vorank. Pearson-r													
123.05	1	1	0	0	1	0	1	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0								
123.06	1	1	0	0	1	0	1	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0							
124.01	1	1	0	0	1	0	1	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
124.02	1	1	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
124.03	1	1	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
125.01	1	1	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
125.02	1	1	0	0	1	0	1	0	1	1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
125.03	1	1	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0					
125.04	1	1	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0					
125.05	1	1	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
125.06	1	1	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
126.01	1	1	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	0	1	0	1	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
126.02	1	1	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
126.03	1	1	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
126.04	1	1	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	0	0	1	0	0	1	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
126.05	1	1	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
126.06	1	1	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
127.01	1	1	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
127.02	1	1	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
128.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0			
129.00	1	1	0	0	1	0	1	0	1	1	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
130.00	1	1	0	0	1	0	1	0	1	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
132.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1			
133.00	1	1	0	1	0	0	1	0	0	1	1	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
134.00	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
135.00	1	1	0	0	1	1	0	0	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
136.00	1	1	0	0	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	1	0	1	0	1	0	1	0	1	0	1	0	0	0	0		
137.01	1	1	0	0	0	1	1	0	0	1	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
137.02	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
137.03	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
137.04	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
137.05	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
137.06	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
137.07	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
138.00	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
139.00	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
140.00	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
141.00	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
142.00	1	1	0	0	1	0	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
143.01	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
143.02	1	1	0	0	1	0	1	0	0																																								

Tabelle C.3
(Fortsetzung)

Studien-Nr.	Metaanalysen zur berufsbezogenen Lernleistung ...																																			Arbeit		Karriere								
	Gesamt publiziert	nicht publiziert	Testmanuale	Artikel	Monographien	ohne BET-Manual	BET-Manual vor 1945	nach 1945	prädiaktiv	konkurrent	Skill Level 2	Skill Level 3	max. Hauptschule	Hochschulreife	Berufs-HG 0	Berufs-HG 3	Berufs-HG 4	Berufs-HG 7	Berufs-HG 8	IST	WIT	BIS	LPS/PSB	Matrizen	GATB-basiert	anderer IQ-Test	Schule	Betrieb	Ausbildung	Weiterbildung	Gesamt publiziert	prädiaktiv	Pearson-r	Einkommen	Vorankommen	Vorank. Pearson-r										
153.00	1	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0						
156.00	1	0	1	0	0	0	1	0	0	1	1	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0				
159.00	1	0	1	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0			
160.01	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0			
160.02	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	1	0	1	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0			
161.01	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0			
161.02	1	1	0	0	1	0	1	0	0	1	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0		
162.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0				
162.02	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0				
163.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1			
164.01	1	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0		
164.02	1	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0		
165.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0		
167.00	1	1	0	0	0	1	1	0	0	1	1	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
168.00	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
169.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1			
170.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	0	1	1	1			
171.01	1	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
171.02	1	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
171.03	1	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
171.04	1	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
172.00	1	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	1	1	0	1	0	1	0	1	1	0	1	1	0	0	0	0		
173.00	1	1	0	0	1	0	1	0	0	1	1	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
174.00	1	1	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
175.00	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
176.01	1	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
176.02	1	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
177.01	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
177.02	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
178.00	1	0	1	0	0	0	1	0	0	1	1	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
179.01	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
179.02	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
179.03	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
179.04	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
179.05	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
179.06	1	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
179.07	1	1	0	0	0	1	1	0	0	1	1	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
179.08	1	1	0	0	0	1	1	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
179.09	1	1	0	0	0	1	1	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
179.10	1	1	0	0	0	1	1	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
179.11	1	1	0	0	0	1	1	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
179.12	1	1	0	0	0	1	1	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0																	

C.3. BIS-Kodierung

Tabelle C.4

Einordnung der Intelligenzaufgaben in die BIS-Klassifikation

Aufabengruppe	Operationen				Inhalte		
	K	E	M	B	F	V	N
AZUBI-BK (Schuler & Klingner, 2004, 2005)							
Gedächtnis			M			V	N
Postmodul				B		V	N
Rechnen	K						N
Sprache	K					V	
Begabungstest (Paff, 1966)							
Figurenalogien	K				F		
logisches Denken	K					V	N
Rechenaufgaben	K						N
BET (Schmale & Schmidtke, 1995b)							
1 Werkzeugvergleich				B	F		
2 Körperabwicklungen	K				F		
3 Adressenvergleich				B		V	
4 Grundrechnen	K						N
5 Figurenlesen				B	F		
6 Rechenaufgabe	K						N
7 Begriffsähnlichkeit, -gegensatz	K					V	
Biernat und Scholz (1995)							
Abwicklungen (AW) / Drehfiguren (DF)	K				F		
Figurales Gedächtnis (GTF)			M		F		
Numerisches Gedächtnis (GTN)			M				N
Rechts-Links-Wendungen (RLW)	K				F		
Unterschiede Erkennen (UE)				B	F		
Verbales Gedächtnis (GTV)			M			V	
BIS (Jäger et al., 1997)							
Abwicklungen (AW)	K				F		
Analogien (AN)	K				F		
Anwendungs-Möglichkeiten (AM)		E				V	
Bongard (BG)	K				F		
Buchstaben-Durchstreichen (BD)				B	F		
Buchstabenreihen (BR)	K						N
Charkow (CH)	K				F		
Divergentes Rechnen (DR)		E					N
Eigenschaften-Fähigkeiten (EF)		E				V	
Figuren-Auswahl (FA)	K				F		
Firmen-Zeichen (FM)			M		F		
Insight-Test (IT)		E				V	
Klassifizieren von Wörtern (KW)				B		V	
Layout (LO)		E			F		
Masselon (MA)		E				V	
Objekt-Gestaltung (OJ)		E			F		

Tabelle C.4
(Fortsetzung)

Aufgabengruppe	Operationen				Inhalte		
	K	E	M	B	F	V	N
Old English (OE)				B	F		
Orientierungsgedächtnis (OG)			M		F		
Phantasiesprache (PS)			M			V	
Rechen-Zeichen (RZ)				B			N
Rechnerisches Denken (RD)	K						N
Schätzen (SC)	K						N
Schlüsse (SL)	K					V	
Schlüsse-Vergleichen (SV)	K					V	
Sieben-Teilbar (SI)				B			N
Sinnvoller Text (ST)			M			V	
Tabellen-Lesen (TL)	K						N
Tatsache-Meinung (TM)	K					V	
Teil-Ganzes (TG)				B		V	
Telefon-Nummern (TN)		E					N
Unvollständige Wörter (UW)				B		V	
Wege Erinnern (WE)			M		F		
Wortanalogien (WA)	K					V	
Worte Merken (WM)			M			V	
Wortschatz (WS)	K					V	
X-Größer (XG)				B			N
Zahlen-Gleichungen (ZG)		E					N
Zahlen-Paare (ZP)			M				N
Zahlenrätsel (ZR)		E					N
Zahlenreihen (ZN)	K						N
Zahlen-Symbol-Test (ZS)				B	F		
Zahlen-Wiedererkennen (ZW)			M				N
Zeichen-Fortsetzen (ZF)		E			F		
Zeichen-Kombinieren (ZK)		E			F		
Zweistellige Zahlen (ZZ)			M				N
BIS-r-DGP (Kersting & Beauducel, 1997; Beauducel & Kersting, 2002)							
Abwicklungen (AW)	K				F		
Analogien, figural (AN)	K				F		
Analogien, verbal (AG)	K					V	
Anwendungs-Möglichkeiten (AM)		E				V	
Bongard (BG)	K				F		
Buchstaben-Durchstreichen (BD)				B	F		
Buchstabenreihen (BR)	K						N
Drehfiguren (DF)	K				F		
Eigenschaften-Fähigkeiten (EF)		E				V	
Ergebnisse Schätzen (ES)	K						N
Figuren Vergleichen (FI)				B	F		
Figuren-Auswahl (FA)	K				F		
Figuren-Paare (FP)			M		F		
Firmen-Zeichen (FM)			M		F		
Marigold (MG)				B	F		
Masselon (MA)		E				V	

Tabelle C.4
(Fortsetzung)

Aufgabengruppe	Operationen				Inhalte		
	K	E	M	B	F	V	N
Objekt-Gestaltung (OJ)		E			F		
Orientierungs-Gedächtnis (OG)			M		F		
Phantasiesprache (PS)			M			V	
Rechen-Zeichen (RZ)				B			N
Schlüsse (SL)	K					V	
Schlüsse-Vergleichen (SV)	K					V	
Sechs-Teilbar (SE)				B			N
Sinnvoller Text (ST)			M			V	
Tabellen/Statistiken (TS)	K						N
Teil-Ganzes (TG)				B		V	
Textanalyse (TA)	K					V	
Textrechnen (TX)	K						N
Unvollständige Wörter (UW)				B		V	
Worte Merken (WM)			M			V	
Wörter Klassifizieren (KW)				B		V	
Wortklassifikation (KL)	K					V	
X-Größer (XG)				B			N
Zahlenmatrizen (ZM)	K						N
Zahlen-Paare (ZP)			M				N
Zahlen-Wiedererkennen (ZW)			M				N
Zeichen-Fortsetzen (ZF)		E			F		
Zweistellige Zahlen (ZZ)			M				N
CFT3 (Weiß, 1980)	K				F		
Damitz et al. (2000)							
Akustische Richtungsangaben (ARA)				B	F		
Bourdon-Buchstabendurchstreichtest (BOU)				B		V	
Clearance-Test (CLE)			M				N
Flugzeug-Positionstest (FPT)				B	F		
Konzentrations-Belastungs-Test (KBT)				B	F		
Kopfrechnen (KRN)	K						N
Mathematik (MAT)	K						N
Merkfähigkeitstest (MEK)			M		F		N
Rechenaufgaben (RAG)	K						N
Versteckte Figuren finden (VFF)			M	B	F		
Vigilanz-Test (VIG)				B	F	V	N
Würfel-Summen berechnen (WSB)	K		M		F		N
EUB (Engelbrecht, 1975)							
Begriffe vergleichen	K					V	
Diktat	K					V	
Formbeziehungen	K				F		
Grundrechnen	K						N
Netze falten	K				F		
Richtung prüfen	K				F		
Spiegelbildzeichnen	K				F		
Textrechnen	K						N

Tabelle C.4
(Fortsetzung)

Aufgabengruppe	Operationen				Inhalte		
	K	E	M	B	F	V	N
Textrechnen II	K						N
Wortbedeutungen	K					V	
Wortbeziehungen	K					V	
FRT	K				F		
Hollstegge (1994)							
Ähnliche Wortbedeutungen	K					V	
Analogien	K					V	
Beobachtung				B	F		
Ergebnisse Schätzen	K						N
Figurenmatrizen	K				F		
Gedächtnis, numerisch			M				N
Gedächtnis, verbal			M			V	
Kontrolltest-Sorgfalt				B		V	
Kontrolltest-Tempo				B		V	
Schlussfolgerungen	K					V	
Tabellen und Statistiken	K						N
Textanalyse	K					V	
Textrechenaufgaben	K						N
Verschiedene Beziehungen	K						N
Wortschatz	K					V	
Zahlenmatrizen	K						N
Zahlensuchfeld				B			N
Hossiep (1995)							
Analogien (AN)	K					V	
Begriffliche Gemeinsamkeiten (BG)	K					V	
Beobachtungsfähigkeit (BO)	K				F		
Entdecken von Regelwidrigkeiten (ER)	K				F		
Erkennen von Gemeinsamkeiten (EG)	K				F		
Rechenaufgaben (RE)	K						N
Satzergänzen (SE)	K					V	
Schlussfolgerndes Denken (SD)	K				F		
Zahlenreihen (ZR)	K						N
IST 2000 R (Amthauer et al., 2001)							
Analogien (AN)	K					V	
Figurenauswahl (FA)	K				F		
Gemeinsamkeiten (GE)	K					V	
Matrizen (MA)	K				F		
Merkfähigkeit, figural			M			V	
Merkfähigkeit, verbal			M		F		
Rechenaufgaben (RE)	K						N
Rechenzeichen (RZ)	K			B			N
Satzergänzung (SE)	K					V	
Würfelaufgaben (WÜ)	K				F		
Zahlenreihen (ZR)	K						N

Tabelle C.4
(Fortsetzung)

Aufgabengruppe	Operationen				Inhalte		
	K	E	M	B	F	V	N
IST 70 (Amthauer, 1973)							
Analogien (AN)	K					V	
Figurenauswahl (FA)	K				F		
Gemeinsamkeiten (GE)	K					V	
Merkaufgaben (ME)			M			V	
Rechenaufgaben (RA/RE)	K						N
Satzergänzung (SE)	K					V	
Wortauswahl (WA)	K					V	
Würfelaufgaben (WÜ)	K				F		
Zahlenreihen (ZR)	K						N
Kersting (1999a)							
Abwicklungen (AW)	K				F		
Ähnliche Wortbedeutungen (ÄW)	K					V	
Analogien (AG)	K					V	
Buchstaben Durchstreichen (BD)				B	F		
Charkow (CH)	K				F		
Computerausdruck (CA)				B			N
Eigenschaften-Fähigkeiten (EF)		E				V	
Ergebnisse Schätzen (ES)	K						N
Objekt-Gestaltung (OJ)		E			F		
Orientierungs-Gedächtnis (OG)			M		F		
Phantasiesprache (PS)			M			V	
Postaufgabe (PA)				B			N
Schlüsse (SL)	K					V	
Sieben-Teilbar (SI)				B			N
Tabellen und Statistiken (TS)	K						N
Tatsache-Meinung (TM)	K						N
Telefon-Nummern (TN)		E					N
Textanalyse (TA)	K					V	
Textrechenaufgaben (TX)	K						N
Verschiedene Beziehungen (VB)	K				F		
Wörter Klassifizieren (KW)				B		V	
Wortschatz (WS)	K					V	
Zahlen-Paare (ZP)			M				N
Kleinevoss (1983); Seggebruch (1982, 1984); Wolf (1990b)							
Absurde Schlüsse (AS)	K				F	V	N
Figurenmatrizen (FZ)	K				F		
Verschiedene Beziehungen (VB)	K				F	V	N
Zahlenmatrizen (ZZ)	K				F		N
Kreuscher (1987)							
Analogien finden	K				F		N
Intelligenz-Organisation	K				F		
Sätze ergänzen	K					V	
Text-Rechenaufgaben lösen	K						N

Tabelle C.4
(Fortsetzung)

Aufgabengruppe	Operationen				Inhalte		
	K	E	M	B	F	V	N
Zahlenreihen fortführen	K						N
Kuhl und Greve (1971)							
Analogien	K				F		N
Denkrechnen	K						N
Figurenreihen	K	E	M	B	F		
Listenvergleich				B	F	V	N
Werkzeuge vergleichen				B	F		
Zahlenrechnen	K	E	M	B			N
LPS (Horn, 1983)							
3	K				F		
4	K						N
5				B		V	
6		E				V	
7				B	F		
8	K				F		
9	K				F		
10				B	F		
11				B	F		
12				B		V	
1+2				B		V	
13+14				B			N
Neumann (1938)							
Aufmerksamkeit			M			V	
Gedächtnis			M			V	
Kopfrechen-Aufgaben	K						N
Beurteilung technischer Skizzen	-	-	-	-	-	-	-
PSB (Horn, 1969)							
3	K				F		
4	K						N
5		E				V	
6				B		V	
7	K				F		
8				B	F		
9	K		M				N
10				B			N
1+2				B		V	
Roloff (1928)							
Flächenelement	K				F		
Gedächtnis für verknüpfte Wörter			M			V	
Körpervergleich	K				F		
Lückentest	K					V	
räumliche Konstruktionen	K				F		
Rybakowfiguren	K				F		
Unterschiede in Zeichnungen	K	E	M	B	F		

Tabelle C.4
(Fortsetzung)

Aufgabengruppe	Operationen				Inhalte		
	K	E	M	B	F	V	N
Beurteilung technischer Skizzen	–	–	–	–	–	–	–
Türschlösser	–	–	–	–	–	–	–
Apparateskizze und -beschreibung	–	–	–	–	–	–	–
Raven Matrizen	K				F		
RVT (Paff, 1966)				B	F		
Salzgitter Service und Technik GmbH (2004)	K					V	
Schuler und Höft (2004)							
räumlich	K			B	F		
rechnerisch	K			B			N
Rechtschreibung							
technisch	K						
Schuler, Funke et al. (1995)							
Assoziationen	K					V	
Bild-Zahlen-Paar			M		F		N
Fehler im Text				B		V	
Funktionsumwandlung		E			F	V	N
Hypothesen		E				V	
Mängel erkennen	K				F		
Matrizen	K				F		
Reasoning	K				F	V	
Versteckte Figuren				B	F		
Wortauswahl	K					V	
WIT (Knebelau, 2002; Jäger, 1963; Greif, 1972)							
Abwicklungen (AW)	K				F		
Analogien (AL)	K					V	
Beobachtung (BO)				B	F		
Buchstabenreihen (BR)	K						N
Eingekleidete Rechenaufgaben (ER)	K						N
Gedächtnis (GD/GDT)			M		F	V	N
Gleiche Wortbedeutung (GW)	K					V	
Grundrechnen (GR)	K						N
Klassifikation (KL; nur Version von 1963)	K					V	
Schätzen (Sch)	K						N
Spiegelbilder (SP)	K				F		
Sprichwörter (SW)	K					V	
Wortgewandtheit (WG)		E				V	
Zahlen-Merken (ZM)			M				N
Zahlenreihen (ZN)	K						N
WPT (Wonderlic Inc., 1996)	K			B	F	V	N

Anmerkungen. Operationen sind die BIS-Facetten K = Verarbeitungskapazität, E = Einfallsreichtum, M = Merkfähigkeit und B = Bearbeitungsgeschwindigkeit; Inhalte sind die BIS-Facetten F = figuralbildhaft, V = verbal und N = numerisch (Jäger et al., 1997, S. 21).

C.4. ISCO-88-Kodierung

Tabelle C.5

ISCO-88-kodierte Berufe der Studien zur berufsbezogenen Lernleistung

ISCO-88 Kode	Skill Level	Berufsbezeichnung	Studie Nr.
0110	–	Marinesoldaten	150.00
0110	–	Offizierbewerber bei der Bundeswehr: Heer	181.01
0110	–	Offizierbewerber bei der Bundeswehr: Luftwaffe	181.02
0110	–	Offizierbewerber bei der Bundeswehr: Marine	181.03
0110	–	Offizierbewerber bei der Bundeswehr: Sanitäter	181.04
0110	–	Offizierbewerber bei der Bundeswehr	182.00
0110	–	Unteroffiziere Marine	186.00
1000	–	Revierförster	101.02
2429	4	Rechtspfleger	109.00
2441	4	Betriebswirt/in (BA)	189.01
3000	3	biologisch-technischer Assistent und chemisch-technischer Assistent	161.01
3000	3	biologisch-technischer Assistent und chemisch-technischer Assistent	161.02
3111	3	Mathem.Techn. Assistenten	116.03
3111	3	Chemielaborant	116.08
3111	3	Chemielaborant	116.09
3111	3	Chemielaboranten	119.01
3111	3	Chemielaboranten	125.02
3111	3	Chemie- und Physiklaboranten	126.02
3111	3	Chemielaborant	137.01
3111	3	Physiklaborant	189.49
3111	3	Chemielaborant	189.50
3111	3	Textillaborant	189.52
3112	3	Vermessungstechniker	133.00
3112	3	Vermessungstechniker	189.47
3117	3	Grubensteiger	120.01
3117	3	Grubensteiger	120.02
3117	3	Grubensteiger	120.03
3117	3	Grubensteiger	120.04
3117	3	Nachwuchs für mittleres Management (Steiger)	142.00
3118	3	Bauzeichner	189.45
3118	3	Technischer Zeichner	189.46
3120	3	Fachinformatiker	189.70
3131	3	Reproduktionsfotograf	189.37
3144	3	Fluglotsen	149.00
3211	3	Biologielaborant	189.48
3224	3	Augenoptiker	179.07
3224	3	Augenoptiker	179.13
3224	3	Augenoptiker (Handwerk)	189.36
3226	3	Masseure und med. Bademeister	179.14
3229	3	Hörgeräteakustiker	179.06
3229	3	Hörgeräteakustiker	179.12
3231	3	Krankenschwestern und -pfleger	130.00
3400	3	Verwaltungsinspektoren	156.00
3400	3	gehobener Dienst des Landes NRW	178.00

Tabelle C.5
(Fortsetzung)

ISCO-88 Kode	Skill Level	Berufsbezeichnung	Studie Nr.
3400	3	Gehobener Dienst in der Bundeswehrverwaltung (Zivilpersonal)	183.00
3430	3	RegierungsinspektorIn	101.01
3430	3	Inspektoren(Verwaltungs-)	102.00
3430	3	Beamter gehobener Dienst	103.00
3430	3	Beamter gehobener Dienst	105.00
3430	3	Gehobener allgemeiner Verwaltungsdienst (Inspektoren)	107.01
3430	3	Gehobener allgemeiner Verwaltungsdienst (Inspektoren)	107.02
3430	3	Gehobener allgemeiner Verwaltungsdienst (Inspektoren)	107.03
3430	3	Angestellte Stadt Hannover: Angestellten-Lehrgang I	110.00
3430	3	mittlerer Dienst Kommunalverwaltung	129.00
3430	3	gehobener Verwaltungsdienst	146.00
3430	3	Kommunalverwaltung gehobener Dienst	160.01
3430	3	Landes- und Kommunalverwaltung gehobener Dienst	160.02
3432	3	Beamter mittlerer Dienst allgemein Justizvollzug	104.00
3450	3	Schutz- und Kriminalpolizei (gehobener Dienst)	108.00
3471	3	Schauwerbegestalter	189.35
4000	2	Büroassistentinnen	125.04
4000	2	Reha-Lehrgang, Verwaltungsberuf	134.00
4000	2	BA Bank (26.3%), Handel (10.1%), Industrie (28.3%), Steuern (35.4%)	177.02
4000	2	Groß- und Außenhandelskaufmann/frau	189.07
4000	2	Einzelhandelskaufmann/frau	189.08
4000	2	Kaufmann/frau für Bürokommunikation und Büroorganisation: Fachrichtung Computergesteuerte Textverarbeitung	189.12
4000	2	Kaufmann/frau für Bürokommunikation und Büroorganisation: Fachrichtung Personalverwaltung	189.13
4100	2	Industriekaufleute	125.03
4100	2	Industriekaufleute	126.06
4100	2	Industriekaufleute	137.03
4100	2	Bürokaufleute	179.01
4100	2	Industriekaufleute	179.02
4100	2	Bürokaufleute	179.08
4100	2	Industriekaufleute	179.09
4100	2	Bürokaufmann Verwaltungsfachangestellter	189.09
4100	2	IT-Systemkaufmann	189.69
4110	2	Bürogehilfin, Kontoristin	116.11
4112	2	Phonotypistin	189.11
4120	2	DV-Kaufmann/frau	189.05
4120	2	Industriekaufmann/frau	189.06
4122	2	Bankkaufmann/frau	189.02
4122	2	Versicherungskaufmann/frau	189.03
4122	2	Sozialversicherungsfachangestellte/r (SOFA)	189.04
4131	2	Fachkraft für Lagerwirtschaft	189.34
4140	2	Bürogehilfin	189.10
4200	2	Energieelektroniker, Industrieelektroniker, Mess- und Regelmechaniker	126.03
4200	2	Kaufleute für Bürokommunikation	126.05
4200	2	Büroassistentinnen	137.04

Tabelle C.5
(Fortsetzung)

ISCO-88 Kode	Skill Level	Berufsbezeichnung	Studie Nr.
4212	2	Bankkaufmann / -frau	135.00
4212	2	Bankkaufleute	138.00
4212	2	Bankkaufleute	151.00
4212	2	Bankkaufleute	172.00
4212	2	Bankkaufleute	174.00
4212	2	Bankkaufleute	179.03
4212	2	Bankkaufleute	179.10
5141	2	Friseur	118.01
5161	2	Beamter des mittleren feuerwehrtechnischen Dienstes	139.00
7000	2	REFA-Grundlehrgang	117.00
7000	2	Elektroanlageninstallateur, Betriebsschlosser, Bergmechaniker und Berg- und Maschinenmänner	144.00
7000	2	Rohrleitungsbauer	189.18
7000	2	Teilezurichter	189.23
7000	2	Kunstfaserschlosser	189.27
7111	2	Bergmechaniker	143.03
7200	2	Metall- und Elektroberufe für Haupt- und Sonderschulabsolventen	106.00
7200	2	Metall-Facharbeiter; Schwerpunkte Fertigungstechnik, Metallbau, Installationstechnik	112.00
7200	2	Maschinenbauer	124.01
7200	2	Maschinenbauer	124.02
7200	2	Dreher	124.03
7200	2	Maschinenschlosser, Betriebsschlosser, Mechaniker, Hochdruck-Rohrschlosser, Kessel- und Behälterbauer, Kunststoffschlosser	125.05
7200	2	Werkzeugmacher, Mechaniker	127.02
7200	2	metallgewerbl. Lehrlinge Hüttenwerk Salzgitter AG	168.00
7200	2	Mechatroniker	171.04
7200	2	Bauschlosser/Stahlbaus Schlosser	189.16
7200	2	Konstruktionsmechaniker	189.59
7200	2	Anlagenmechaniker: Fachrichtung Apparatechnik	189.60
7200	2	Anlagenmechaniker: Fachrichtung Versorgungstechnik	189.61
7212	2	Schweißer	189.22
7213	2	Feinblechner/Kupferschmied	189.25
7220	2	Werkzeugmechaniker	189.57
7222	2	Werkzeugmacher/Feinmechaniker	189.19
7223	2	Fräser	123.01
7223	2	Bohrer	123.02
7223	2	Dreher	123.03
7223	2	Bohrwdreher	123.04
7223	2	Dreher	189.20
7223	2	Fräser	189.21
7223	2	Zerspanungsmechaniker	189.58
7224	2	Schleifer	123.05
7230	2	Betriebsschlosser	137.06
7230	2	Betriebsschlosser	143.02

Tabelle C.5
(Fortsetzung)

ISCO-88 Kode	Skill Level	Berufsbezeichnung	Studie Nr.
7230	2	Industriemechaniker, Zerspanungsmechaniker	164.01
7230	2	Schlosser	167.00
7230	2	Mechaniker	171.01
7230	2	Fertigungsmechaniker	171.03
7230	2	Industriemechaniker	176.02
7230	2	Mechaniker	184.02
7231	2	KFZ-Mechaniker	118.03
7231	2	KFZ-Schlosser	189.17
7231	2	Automobilmechaniker	189.62
7233	2	Industrie- und Anlagenmechaniker	126.04
7233	2	Maschinenschlosser	189.14
7233	2	Betriebsschlosser	189.15
7233	2	Automateneinrichter	189.24
7233	2	Industriemechaniker: Fachrichtung Produktionstechnik	189.53
7233	2	Industriemechaniker: Fachrichtung Betriebstechnik	189.54
7233	2	Industriemechaniker: Fachrichtung Maschinen- und Systemtechnik	189.55
7233	2	Industriemechaniker: Fachrichtung Geräte- und Feinwerktechnik	189.56
7240	2	Energieanlagenelektroniker, Energiegeräteelektroniker, Mess- und Regelmechaniker	125.06
7240	2	Naachrichtengerätetechnik	127.01
7240	2	Mess- und Regelmechaniker	137.05
7240	2	Industriemechaniker, -elektroniker, Mechatroniker	153.00
7240	2	Industrieelektroniker, Produktionselektroniker versch. Fachrichtungen	164.02
7240	2	Elektriker/Elektroniker	171.02
7240	2	Energieelektroniker	176.01
7240	2	Nachrichtengerätetechnik	179.04
7240	2	Elektroniker	179.05
7240	2	Elektroniker	179.11
7240	2	Nachrichtengerätetechnik	189.30
7240	2	Informationselektroniker	189.31
7240	2	Energieanlagen-Elektroniker	189.33
7240	2	Kommunikationselektroniker	189.68
7241	2	Elektroanlageninstallateur	143.01
7241	2	Mess- und Regelmechaniker	189.28
7241	2	Starkstromelektriker	189.29
7241	2	Energieanlagen-Installateur	189.32
7241	2	Elektromaschinenmonteur	189.63
7241	2	Energieelektroniker: Fachrichtung Anlagentechnik	189.64
7241	2	Energieelektroniker: Fachrichtung Betriebstechnik	189.65
7241	2	Industrieelektroniker: Fachrichtung Produktionstechnik	189.66
7241	2	Industrieelektroniker: Fachrichtung Gerätetechnik	189.67
7244	2	Fernmeldehandwerker	140.00
7340	2	Tiefdruckretuscheur	189.38
7341	2	Tiefdrucker	189.39
7341	2	Flachdrucker (Offset, Farblitho)	189.41
7341	2	Maschinensetzer	189.42

Tabelle C.5
(Fortsetzung)

ISCO-88 Kode	Skill Level	Berufsbezeichnung	Studie Nr.
7341	2	Handsetzer	189.43
7341	2	Korrektor	189.44
7343	2	Hochdrucker (Buchdruck)	189.40
7422	2	Tischler	118.02
7422	2	Möbeltischler/Holzmechaniker	189.26
7900	2	Meister: Schmied, Tischler, Werkzeugmacher, Polsterer, usw.	114.00
7900	2	Kontrolleure	123.06
7900	2	Gruben-Fahrhauer-Meister	141.00
8159	2	Chemikanten	113.00
8159	2	Chemiefacharb.	116.10
8159	2	Chemiefacharbeiter	119.02
8159	2	Chemikanten	125.01
8159	2	Chemikanten	126.01
8159	2	Chemikant	137.02
8159	2	Chemiekant (Chemiefacharbeiter)	189.51
8229	2	Laborfacharbeiter	116.06
	2	gewerblich, kaufmännische Ausbildung	111.01
	2	gewerbliche Ausbildung	111.02
	2	Facharbeiter Elektro/Metall; gewerbl. Handwerkl. Techn. Berufsausbildung	122.00
	2	Industrieinformatiker	137.07
	2	gewerbliche Berufe: vorallem Metall-, Bau-, Elektroberufe, wenige Nahrungs- und Kunstberufe	173.00
	2	KFZ-Mechaniker (39.3%), Tischler (30.1%), Friseur (29.6%)	175.00
	2	Bankaufleute (29.2%), Verkäufer/Einzelhandelsaufleute (28.2%), Industriekaufleute (27.9%), Steuerfachangestellte (14.7%)	177.01
		Beamter höherer Dienst Kripo	100.00
		Ingenieure grad.	116.01
		techn.handwerkl. Beruf	116.04
		Ingenieur grad.	116.13
		gehobener und höherer Dienst Kripo und Schupo	136.00
		Bereitschaftspolizei (West-Berlin)	147.00
		Schutz-, Wasser- und Kriminalpolizisten; höherer Dienst	148.00
		Büro-, Bank-, Informatikkaufleute, Fachinformatiker, Dipl-Betriebswirt-(BA), dto. Volkswirt	152.00
		Reha Vorbereitungslehrgang	159.00
		verschiedene	184.01
		verschiedene	184.03
		19 verschiedene Berufe in 9 Berufsfeldern (BBW Nürnberg)	185.00

Anmerkungen. keine Angaben: Codes bzw. Skill Level können nicht zugeordnet werden

Tabelle C.6

ISCO-88-kodierte Berufe der Studien zur Arbeitsleistung

ISCO-88 Kode	Skill Level	Berufsbezeichnung	Studie Nr.
1000	–	Revierförster	101.02
1000	–	Management-Nachwuchskräfte	132.00
2000	4	Forschung und Entwicklung	128.00
2429	4	Rechtspfleger	109.00
3111	3	Chemielaboranten	116.07
3111	3	Chemielaborant	116.12
3117	3	Steiger	115.00
3430	3	RegierungsinspektorIn	101.01
4100	2	Industriekaufleute	116.02
4212	2	Bankkaufmannleute	135.00
4212	2	Bankkaufleute	172.00
7233	2	Betriebsschlosser	116.05
8000	2	Apparatewärter	162.01
8000	2	Maschinisten	162.02
		gehobener und höherer Dienst Kripo und Schupo	136.00
		verschiedene	170.00
		hauptsächlich Verkäufer, auch Back-Office	187.01
		verschiedene, hauptsächlich blue-color worker	187.02

Anmerkungen. keine Angaben: Kodes bzw. Skill Level können nicht zugeordnet werden

Tabelle C.7

ISCO-88-kodierte Berufe der Studien zum Karriereerfolg

ISCO-88 Kode	Skill Level	Berufsbezeichnung	Studie Nr.
1000	–	Management-Nachwuchskräfte	132.00
3410	3	Vermögensberater	163.00
4212	2	Bankkaufmannleute	135.00
4212	2	Bankkaufleute	151.00
		verschiedene Berufe: Dienstleistung, Industrie, öffentliche Verwaltung	145.00
		verschiedene	165.00
		verschiedene	169.00
		verschiedene	170.00
		19 verschiedene Berufe in 9 Berufsfeldern (BBW Nürnberg)	185.00
		verschiedene	188.00

Anmerkungen. keine Angaben: Kodes bzw. Skill Level können nicht zugeordnet werden

D. Gewichte ausgewählter Variablen an den Sensitivitäts- und Moderatoranalysen

Tabelle D.1

Lernleistung: bare-bone Gewichte ausgewählter Variablen an den Sensitivitäts- bzw. Moderatoranalysen

Gewicht der Variablen in Zeile an den Analysen in Spalte (%)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
insignifikante Korrelationen ...																		
1 ... berichtet	100.0	0.0	99.5	75.5	94.9	100.0	100.0	92.3	87.0	100.0	92.4	100.0	100.0	95.6	95.5	100.0	98.3	83.0
2 ... nicht berichtet	0.0	100.0	0.5	24.5	5.1	0.0	0.0	7.7	13.0	0.0	7.6	0.0	0.0	4.4	4.5	0.0	1.7	17.0
Studien mit Mittelwert-Aggregaten																		
3 nein	87.2	9.7	100.0	0.0	83.4	86.6	98.8	50.3	68.5	86.6	72.0	100.0	73.6	83.8	84.4	54.7	86.7	66.2
4 ja	12.8	90.3	0.0	100.0	16.6	13.4	1.2	49.7	31.5	13.4	28.0	0.0	26.4	16.2	15.6	45.3	13.3	33.8
Studien publiziert																		
5 ja	86.5	100.0	86.7	89.4	100.0	0.0	100.0	100.0	100.0	0.0	77.7	100.0	100.0	87.1	86.8	100.0	91.6	89.3
6 nein	13.5	0.0	13.3	10.6	0.0	100.0	0.0	0.0	0.0	100.0	22.3	0.0	0.0	12.9	13.2	0.0	8.4	10.7
Art der Publikation																		
7 Testmanuale	51.2	0.0	57.7	3.6	56.2	0.0	100.0	0.0	0.0	0.0	11.8	100.0	0.0	49.2	50.1	0.0	55.9	35.7
8 Artikel in Zeitschriften / Hrsg.-Werken	9.9	18.0	6.2	31.5	11.8	0.0	0.0	100.0	0.0	0.0	17.8	0.0	0.0	10.3	8.3	100.0	8.8	8.9
9 Monographien / Dissertationen	25.4	82.0	22.8	54.2	32.0	0.0	0.0	0.0	100.0	0.0	48.2	0.0	100.0	27.6	28.4	0.0	26.9	44.7
10 nicht publiziert	13.5	0.0	13.3	10.6	0.0	100.0	0.0	0.0	0.0	100.0	22.3	0.0	0.0	12.9	13.2	0.0	8.4	10.7
Studien aus dem BET-Manual																		
11 nein	55.9	100.0	49.7	100.0	51.6	100.0	13.9	100.0	100.0	100.0	100.0	0.0	100.0	57.6	56.8	100.0	47.4	76.9
12 ja	44.1	0.0	50.3	0.0	48.4	0.0	86.1	0.0	0.0	0.0	0.0	100.0	0.0	42.4	43.2	0.0	52.6	23.1
Veröffentlichungs- bzw. Durchführungsjahr																		
13 vor 1945	0.4	0.0	0.3	0.6	0.5	0.0	0.0	0.0	1.4	0.0	0.7	0.0	100.0	0.0	0.3	0.0	0.6	0.0
14 nach 1945	99.6	100.0	99.7	99.4	99.5	100.0	100.0	100.0	98.6	100.0	99.3	100.0	0.0	100.0	99.7	100.0	99.4	100.0
Art der Validität																		
15 prädiktive Studien	97.6	100.0	98.4	94.2	97.4	100.0	100.0	78.9	99.5	100.0	96.0	100.0	73.6	97.8	100.0	0.0	99.1	98.1
16 konkurrente Studien	1.9	0.0	1.2	5.2	2.1	0.0	0.0	18.0	0.0	0.0	3.2	0.0	0.0	1.9	0.0	100.0	0.7	1.8
Berufskomplexität																		
17 Skill Level 2	74.0	28.5	74.6	58.9	75.7	47.1	82.2	61.5	69.6	47.1	59.0	89.8	100.0	71.9	73.1	29.1	100.0	0.0
18 Skill Level 3	14.8	65.8	13.5	35.6	17.5	14.1	12.5	14.9	27.4	14.1	22.7	9.4	0.0	17.1	17.1	16.2	0.0	100.0
Schulbildung																		
19 maximal Hauptschul-Abschluss	2.9	0.0	2.4	4.5	2.2	6.8	0.0	4.4	5.2	6.8	4.8	0.0	0.0	2.8	2.8	0.0	2.8	0.0
20 (Fach-)Hochschulreife	4.1	62.5	3.1	24.9	5.5	14.1	0.0	10.3	13.5	14.1	11.5	0.0	0.0	6.7	6.5	0.0	0.1	33.2

Tabelle D.1
(Fortsetzung)

Gewicht der Variablen in Zeile an den Analysen in Spalte (%)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
ISCO-88-Berufs-Hauptgruppe (HG)																		
21 Militär (HG 0)	4.5	0.0	5.1	0.0	0.6	28.8	0.0	5.3	0.0	28.8	7.4	0.0	0.0	4.3	4.4	0.0	0.0	0.0
22 Techniker und gleichrangige (HG 3)	14.8	65.8	13.5	35.6	17.5	14.1	12.5	14.9	27.4	14.1	22.7	9.4	0.0	17.1	17.1	16.2	0.0	100.0
23 Bürokräfte (HG 4)	19.6	3.7	15.8	34.8	20.6	7.6	21.3	37.2	13.2	7.6	15.1	24.1	0.0	19.0	18.9	24.5	26.2	0.0
24 Handwerker (HG 7)	42.3	0.0	45.4	14.4	41.1	35.8	55.9	8.2	27.2	35.8	22.5	64.9	100.0	40.2	41.2	4.6	56.1	0.0
25 Maschinenbediener (HG 8)	7.3	0.0	7.7	3.0	7.4	3.8	5.0	2.0	13.7	3.8	11.4	0.8	0.0	7.0	7.1	0.0	9.6	0.0
Intelligenztest ^a																		
26 IST	23.7	0.0	25.6	7.4	22.2	25.7	12.5	29.5	36.5	25.7	39.2	0.0	0.0	22.7	22.4	40.9	17.0	29.4
27 WIT	7.2	68.1	5.4	33.3	9.5	12.4	0.8	4.6	26.8	12.4	17.1	0.0	0.0	9.9	9.8	18.2	4.6	32.1
28 BIS	2.9	3.4	3.3	0.9	1.4	12.8	0.0	3.1	3.4	12.8	5.0	0.0	0.0	2.9	2.6	0.0	2.3	3.7
29 LPS / PSB	6.8	18.0	6.9	9.4	7.0	9.6	0.0	11.0	17.8	9.6	12.7	0.0	0.0	7.3	7.1	18.2	9.0	0.0
30 Matrizentests	3.9	0.0	3.9	3.0	1.4	19.7	0.0	8.2	1.3	19.7	6.5	0.0	0.0	3.8	3.4	22.6	4.5	1.2
31 GATB basierte Verfahren	45.0	1.6	50.3	5.4	49.4	0.0	86.1	0.0	3.1	0.0	1.5	100.0	0.0	43.2	44.1	0.0	53.8	23.1
32 andere Intelligenztests	10.5	8.9	4.6	40.6	9.0	19.8	0.6	43.6	11.1	19.8	18.0	0.0	100.0	10.1	10.6	0.0	8.7	10.5
Kriterien ^a																		
33 Berufsschulnoten	50.9	87.5	45.8	86.9	47.1	88.5	13.3	83.8	93.2	88.5	90.8	0.0	0.0	52.7	52.5	45.3	42.5	73.8
34 betriebliche Beurteilungen	5.0	12.5	3.8	13.1	4.4	11.5	0.6	16.2	6.8	11.5	9.2	0.0	100.0	5.0	4.3	54.7	4.9	3.1
35 Ausbildungsleistung	95.2	75.7	95.7	87.1	93.5	100.0	100.0	96.9	80.8	100.0	90.2	100.0	26.4	94.6	94.5	100.0	97.0	84.7
36 Weiterbildungsleistung	4.8	24.3	4.3	12.9	6.5	0.0	0.0	3.1	19.2	0.0	9.8	0.0	73.6	5.4	5.5	0.0	3.0	15.3

Tabelle D.1
(Fortsetzung)

Gewicht der Variablen in Zeile an den Analysen in Spalte (%)	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
insignifikante Korrelationen ...																		
1 ... berichtet	100.0	58.5	100.0	83.0	99.1	100.0	100.0	100.0	69.7	94.9	89.2	100.0	99.8	96.2	92.7	89.7	96.5	81.2
2 ... nicht berichtet	0.0	41.5	0.0	17.0	0.9	0.0	0.0	0.0	30.3	5.1	10.8	0.0	0.2	3.8	7.3	10.3	3.5	18.8
Studien mit Mittelwert-Aggregaten																		
3 nein	73.6	39.2	100.0	66.2	70.2	94.2	93.0	94.7	45.5	94.9	79.3	87.0	98.0	36.9	73.2	60.2	85.0	63.1
4 ja	26.4	60.8	0.0	33.8	29.8	5.8	7.0	5.3	54.5	5.1	20.7	13.0	2.0	63.1	26.8	39.8	15.0	36.9
Studien publiziert																		
5 ja	68.5	72.5	12.9	89.3	94.8	88.6	93.0	85.4	83.8	43.3	83.0	32.4	100.0	75.5	78.3	72.3	86.3	100.0
6 nein	31.5	27.5	87.1	10.7	5.2	11.4	7.0	14.6	16.2	56.7	17.0	67.6	0.0	24.5	21.7	27.7	13.7	0.0
Art der Publikation																		
7 Testmanuale	0.0	0.0	0.0	35.7	55.2	67.8	35.1	27.1	3.7	0.0	0.0	0.0	98.0	2.8	12.4	5.5	51.9	0.0
8 Artikel in Zeitschriften / Hrsg.-Werken	16.4	16.0	12.9	8.9	20.2	2.1	3.0	13.4	4.8	11.0	15.4	22.4	0.0	43.1	16.4	31.2	10.6	5.6
9 Monographien / Dissertationen	52.1	56.5	0.0	44.7	19.4	18.7	54.9	44.9	75.3	32.3	67.6	10.0	2.0	29.7	49.5	35.6	23.9	94.4
10 nicht publiziert	31.5	27.5	87.1	10.7	5.2	11.4	7.0	14.6	16.2	56.7	17.0	67.6	0.0	24.5	21.7	27.7	13.7	0.0
Studien aus dem BET-Manual																		
11 nein	100.0	100.0	100.0	76.9	46.1	32.2	95.1	100.0	100.0	100.0	100.0	100.0	2.0	100.0	100.0	100.0	55.3	100.0
12 ja	0.0	0.0	0.0	23.1	53.9	67.8	4.9	0.0	0.0	0.0	0.0	0.0	98.0	0.0	0.0	0.0	44.7	0.0
Veröffentlichungs- bzw. Durchführungsjahr																		
13 vor 1945	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.8	0.0	7.4	0.1	5.2
14 nach 1945	100.0	100.0	100.0	100.0	100.0	99.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	96.2	100.0	92.6	99.9	94.8
Design																		
15 prädiktive Studien	100.0	95.2	100.0	98.1	97.6	99.5	100.0	96.5	96.6	89.0	95.4	88.9	100.0	99.0	97.7	79.1	97.9	93.9
16 konkurrente Studien	0.0	0.0	0.0	1.8	2.4	0.2	0.0	3.3	3.4	0.0	4.6	11.1	0.0	0.0	1.6	19.0	2.0	0.0
Berufskomplexität																		
17 Skill Level 2	71.5	1.3	0.0	0.0	100.0	100.0	100.0	54.0	33.7	56.7	88.6	87.0	90.1	60.3	58.4	65.6	74.0	38.7
18 Skill Level 3	0.0	85.6	0.0	100.0	0.0	0.0	0.0	22.1	55.4	22.0	0.0	5.2	9.2	17.2	24.0	9.9	15.3	46.0
Schulbildung																		
19 maximal Hauptschul-Abschluss	100.0	0.0	0.0	0.0	0.5	3.2	0.0	2.8	0.9	0.0	8.0	7.8	1.7	4.4	4.9	4.3	2.3	11.0
20 (Fach-)Hochschulreife	0.0	100.0	0.0	33.2	0.5	0.0	0.0	0.0	38.3	43.3	1.2	0.0	0.0	14.3	12.3	3.2	5.2	29.7

Tabelle D.1
(Fortsetzung)

Gewicht der Variablen in Zeile an den Analysen in Spalte (%)	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
ISCO-88-Berufs-Hauptgruppe (HG)																		
21 Militär (HG 0)	0.0	0.0	100.0	0.0	0.0	0.0	0.0	11.1	0.0	0.0	0.0	0.0	0.0	16.8	8.1	0.0	4.5	0.0
22 Techniker und gleichrangige (HG 3)	0.0	85.6	0.0	100.0	0.0	0.0	0.0	22.1	55.4	22.0	0.0	5.2	9.2	17.2	24.0	9.9	15.3	46.0
23 Bürokräfte (HG 4)	3.2	1.3	0.0	0.0	100.0	0.0	0.0	10.7	8.3	0.0	10.2	7.6	23.8	41.9	14.7	18.8	20.0	0.0
24 Handwerker (HG 7)	45.9	0.0	0.0	0.0	0.0	100.0	0.0	14.1	14.7	56.7	23.3	71.7	65.4	14.6	21.4	33.4	41.2	27.7
25 Maschinenbediener (HG 8)	0.0	0.0	0.0	0.0	0.0	0.0	100.0	24.4	10.7	0.0	0.0	0.0	0.8	0.0	11.8	7.6	7.4	0.0
Intelligenztest ^a																		
26 IST	22.4	0.0	58.8	29.4	12.8	7.9	79.8	100.0	0.0	0.0	0.0	0.0	0.0	0.0	40.6	25.0	21.4	42.7
27 WIT	3.2	57.2	0.0	32.1	4.4	3.6	15.2	0.0	100.0	0.0	0.0	0.0	0.0	0.0	17.7	11.0	8.0	41.3
28 BIS	0.0	18.9	0.0	3.7	0.0	4.1	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	4.0	15.4	2.4	10.9
29 LPS / PSB	21.0	1.3	0.0	0.0	3.9	4.2	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	12.7	12.1	7.8	0.0
30 Matrizentests	10.5	0.0	0.0	1.2	1.5	6.7	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	7.2	0.0	4.0	0.0
31 GATB basierte Verfahren	26.4	0.0	0.0	23.1	54.2	69.8	4.9	0.0	0.0	0.0	0.0	0.0	100.0	0.0	1.5	1.4	45.7	0.0
32 andere Intelligenztests	16.4	22.5	41.2	10.5	23.1	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	16.3	35.2	10.7	5.2
Kriterien ^a																		
33 Berufsschulnoten	91.8	97.4	100.0	73.8	40.8	27.8	89.2	94.1	94.1	71.6	91.2	100.0	1.9	82.0	100.0	0.0	49.9	94.8
34 betriebliche Beurteilungen	8.2	2.6	0.0	3.1	5.3	4.4	5.8	5.9	5.9	28.4	8.8	0.0	0.2	18.0	0.0	100.0	5.3	5.2
35 Ausbildungsleistung	77.6	74.6	100.0	84.7	100.0	96.1	100.0	89.3	76.3	78.7	100.0	100.0	100.0	97.2	89.8	94.5	100.0	0.0
36 Weiterbildungsleistung	22.4	25.4	0.0	15.3	0.0	3.9	0.0	10.7	23.7	21.3	0.0	0.0	0.0	2.8	10.2	5.5	0.0	100.0

Anmerkungen. Die Werte geben den Anteil in % an, den die Variablenausprägungen (Zeilen) an den entsprechenden Metaanalysen (Spalten) haben. D.h. die Spalten ergeben pro Bereich 100%. An 100% fehlende Werte können den Variablen nicht zugeordnet werden. ^aMehrfachnennungen möglich, Aufteilung entsprechend der Gewichtung in den Metaanalysen.

Tabelle D.2

Arbeitsleistung: bare-bone Gewichte ausgewählter Variablen an den Sensitivitäts- bzw. Moderatoranalysen

Gewicht der Variablen in Zeile an den Analysen in Spalte (%)	1	2	3	4	5	6	7	8	9	10	11	12
insignifikante Korrelationen ...												
1 ... berichtet	100.0	0.0	100.0	94.1	100.0	68.5	97.0	100.0	97.3	100.0	96.7	100.0
2 ... nicht berichtet	0.0	100.0	0.0	5.9	0.0	31.5	3.0	0.0	2.7	0.0	3.3	0.0
Studien mit Mittelwert-Aggregaten												
3 nein	58.3	0.0	100.0	0.0	55.7	68.5	53.3	76.4	59.0	0.0	44.7	100.0
4 ja	41.7	100.0	0.0	100.0	44.3	31.5	46.7	23.6	41.0	100.0	55.3	0.0
Studien publiziert												
5 ja	94.3	0.0	90.2	94.1	100.0	0.0	90.4	100.0	91.6	100.0	89.6	100.0
6 nein	5.7	100.0	9.8	5.9	0.0	100.0	9.6	0.0	8.4	0.0	10.4	0.0
Art der Publikation												
Testmanuale	30.5	0.0	52.4	0.0	32.4	0.0	35.1	0.0	30.9	0.0	10.2	100.0
Monographien / Dissertationen	52.2	0.0	24.6	85.4	55.4	0.0	49.6	58.2	52.9	0.0	65.1	0.0
Artikel in Zeitschr. / Hrsg.-Werken	11.5	0.0	13.2	8.7	12.2	0.0	5.8	41.8	7.8	100.0	14.4	0.0
nicht publiziert	5.7	100.0	9.8	5.9	0.0	100.0	9.6	0.0	8.4	0.0	10.4	0.0
Art der Validität												
7 prädiktive Studien	84.4	100.0	79.5	91.7	83.5	100.0	100.0	0.0	84.2	100.0	80.6	100.0
8 konkurrenz, retrograd, unklar	15.6	0.0	20.5	8.3	16.5	0.0	0.0	100.0	15.8	0.0	19.4	0.0
Effektmaß												
9 Pearson-Korrelationen	96.1	100.0	100.0	91.3	95.9	100.0	95.6	100.0	100.0	0.0	95.2	100.0
10 andere	3.9	0.0	0.0	8.7	4.1	0.0	4.4	0.0	0.0	100.0	4.8	0.0
Outlier-Analyse												
11 ohne Studie 116.12	77.6	100.0	61.6	100.0	76.3	100.0	74.3	100.0	77.4	100.0	100.0	0.0
12 Studie 116.12	22.4	0.0	38.4	0.0	23.7	0.0	25.7	0.0	22.6	0.0	0.0	100.0
Berufskomplexität												
ohne Skill Level	2.4	100.0	4.2	5.9	2.6	31.5	5.8	0.0	5.1	0.0	6.3	0.0
Skill Level 2	42.5	0.0	14.2	77.1	43.4	18.9	48.8	0.0	43.0	0.0	52.9	0.0
Skill Level 3	31.5	0.0	47.5	8.7	33.5	0.0	36.2	0.0	28.0	100.0	11.4	100.0
Skill Level 4	9.2	0.0	15.8	0.0	9.8	0.0	4.4	34.6	9.3	0.0	11.5	0.0
gemischt / unklar	14.4	0.0	18.3	8.3	10.7	49.6	4.8	65.4	14.5	0.0	17.9	0.0
ISCO-88-Berufs-Hauptgruppe (HG)												
Leitende Berufe (HG 1)	2.4	100.0	4.2	5.9	2.6	31.5	5.8	0.0	5.1	0.0	6.3	0.0
Wissenschaftler (HG 2)	9.2	0.0	15.8	0.0	9.8	0.0	4.4	34.6	9.3	0.0	11.5	0.0
Techniker und gleichrangige (HG 3)	31.5	0.0	47.5	8.7	33.5	0.0	36.2	0.0	28.0	100.0	11.4	100.0
Bürokräfte (HG 4)	38.6	0.0	7.5	77.1	39.3	18.9	44.3	0.0	39.1	0.0	48.1	0.0
Handwerker (HG 7)	2.7	0.0	4.7	0.0	2.9	0.0	3.1	0.0	2.8	0.0	3.4	0.0
Maschinenbediener (HG 8)	1.2	0.0	2.0	0.0	1.2	0.0	1.3	0.0	1.2	0.0	1.4	0.0
gemischt / unklar	14.4	0.0	18.3	8.3	10.7	49.6	4.8	65.4	14.5	0.0	17.9	0.0
Intelligenztest												
IST	44.9	0.0	70.5	8.7	46.0	18.9	51.6	0.0	41.6	100.0	28.1	100.0
WIT	0.0	100.0	0.0	5.9	0.0	31.5	3.0	0.0	2.7	0.0	3.3	0.0
BIS	3.7	0.0	0.0	8.3	3.9	0.0	0.0	23.6	3.7	0.0	4.6	0.0
WPT	10.6	0.0	18.3	0.0	6.9	49.5	4.7	41.8	10.8	0.0	13.3	0.0
FRT	1.2	0.0	2.0	0.0	1.2	0.0	1.3	0.0	1.2	0.0	1.4	0.0
andere	39.6	0.0	9.3	77.1	42.0	0.0	39.3	34.6	40.1	0.0	49.3	0.0
Gesamt	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Anmerkungen. Die Werte geben den Anteil in % an, den die Variablenausprägungen (Zeilen) an den entsprechenden Metaanalysen (Spalten) haben. D.h. die Spalten ergeben pro Bereich 100%.

Tabelle D.3

Einkommen: bare-bone Gewichte ausgewählter Variablen an den Sensitivitätsanalysen

Gewicht der Variablen in Zeile an den Analysen in Spalte (%)	1	2	3	4	5	6	7	8
insignifikante Korrelationen ...								
1 ... berichtet	100.0	0.0	95.8	100.0	100.0	92.4	94.4	100.0
2 ... nicht berichtet	0.0	100.0	4.2	0.0	0.0	7.6	5.6	0.0
Studien mit Mittelwert-Aggregaten								
3 nein	84.6	100.0	100.0	0.0	72.2	100.0	100.0	59.9
4 ja	15.4	0.0	0.0	100.0	27.8	0.0	0.0	40.1
Studien publiziert								
5 ja	55.6	0.0	45.5	100.0	100.0	0.0	61.7	40.1
6 nein	44.4	100.0	54.5	0.0	0.0	100.0	38.3	59.9
Art der Publikation								
Monographien / Dissertationen	40.2	0.0	45.5	0.0	72.2	0.0	61.7	0.0
Artikel in Zeitschriften / Hrsg.-Werken	15.4	0.0	0.0	100.0	27.8	0.0	0.0	40.1
nicht publiziert	44.4	100.0	54.5	0.0	0.0	100.0	38.3	59.9
Art der Validität								
7 prädiktiv	61.5	100.0	73.8	0.0	72.2	52.0	100.0	0.0
8 konkurrent, retrograd, unklar	38.5	0.0	26.2	100.0	27.8	48.0	0.0	100.0
Berufskomplexität								
ohne Skill Level	0.0	100.0	4.2	0.0	0.0	7.6	5.6	0.0
Skill Level 2	50.6	0.0	57.4	0.0	72.2	21.8	77.7	0.0
gemischt / unklar	49.4	0.0	38.4	100.0	27.8	70.6	16.7	100.0
ISCO-88-Berufs-Hauptgruppe (HG)								
Leitende Berufe (HG 1)	0.0	100.0	4.2	0.0	0.0	7.6	5.6	0.0
Bürokräfte (HG 4)	50.6	0.0	57.4	0.0	72.2	21.8	77.7	0.0
gemischt / unklar	49.4	0.0	38.4	100.0	27.8	70.6	16.7	100.0
Intelligenztest								
IST 2000 R	26.3	0.0	12.3	100.0	27.8	22.5	16.6	40.1
WIT (1963)	0.0	100.0	4.2	0.0	0.0	7.6	5.6	0.0
LPS	26.8	0.0	30.3	0.0	0.0	55.7	16.1	42.3
WPT (1996)	6.8	0.0	7.7	0.0	0.0	14.2	0.0	17.7
andere	40.2	0.0	45.5	0.0	72.2	0.0	61.7	0.0
Gesamt	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Anmerkungen. Die Werte geben den Anteil in % an, den die Variablenausprägungen (Zeilen) an den entsprechenden Metaanalysen (Spalten) haben. D.h. die Spalten ergeben pro Bereich 100%.

Tabelle D.4

Berufliches Vorankommen: bare-bone Gewichte ausgewählter Variablen an den Sensitivitätsanalysen

Gewicht der Variablen in Zeile an den Analysen in Spalte (%)	1	2	3	4	5	6	7	8	9	10
insignifikante Korrelationen ...										
1 ... berichtet	100.0	0.0	83.1	0.0	72.0	90.8	94.8	38.5	77.5	100.0
2 ... nicht berichtet	0.0	100.0	16.9	100.0	28.0	9.2	5.2	61.5	22.5	0.0
Studien mit Mittelwert-Aggregaten										
3 nein	100.0	80.8	100.0	0.0	100.0	90.8	94.8	100.0	95.7	100.0
4 ja	0.0	19.2	0.0	100.0	0.0	9.2	5.2	0.0	4.3	0.0
Studien publiziert										
5 ja	52.3	80.8	60.3	0.0	100.0	0.0	56.8	61.5	64.9	0.0
6 nein	47.7	19.2	39.7	100.0	0.0	100.0	43.2	38.5	35.1	100.0
Art der Publikation										
Monographien / Dissertationen	44.3	0.0	36.8	0.0	61.0	0.0	48.1	0.0	39.6	0.0
Artikel in Zeitschriften / Hrsg.-Werken	8.0	80.8	23.5	0.0	39.0	0.0	8.7	61.5	25.3	0.0
nicht publiziert	47.7	19.2	39.7	100.0	0.0	100.0	43.2	38.5	35.1	100.0
Art der Validität										
7 prädiktive Studien	87.3	19.2	72.5	100.0	72.0	75.8	100.0	0.0	70.5	100.0
8 konkurrent, retrograd, unklar	12.7	80.8	27.5	0.0	28.0	24.2	0.0	100.0	29.5	0.0
Effektmaß										
9 Pearson-Korrelationen	86.8	100.0	89.0	100.0	100.0	74.8	85.6	100.0	100.0	0.0
10 andere	13.2	0.0	11.0	0.0	0.0	25.2	14.4	0.0	0.0	100.0
Berufskomplexität										
ohne Skill Level	0.0	19.2	0.0	100.0	0.0	9.2	5.2	0.0	4.3	0.0
Skill Level 2	57.5	0.0	47.8	0.0	61.0	25.2	62.5	0.0	39.6	100.0
Skill Level 3	3.4	0.0	2.8	0.0	0.0	6.4	0.0	10.2	3.0	0.0
gemischt / unklar	31.1	80.8	42.8	0.0	28.0	59.2	23.6	89.8	46.0	0.0
ISCO-88-Berufs-Hauptgruppe (HG)										
Leitende Berufe (HG 1)	0.0	19.2	0.0	100.0	0.0	9.2	5.2	0.0	4.3	0.0
Techniker und gleichrangige (HG 3)	3.4	0.0	2.8	0.0	0.0	6.4	0.0	10.2	3.0	0.0
Bürokräfte (HG 4)	57.5	0.0	47.8	0.0	61.0	25.2	62.5	0.0	39.6	100.0
gemischt / unklar	31.1	80.8	42.8	0.0	28.0	59.2	23.6	89.8	46.0	0.0
Intelligenztest										
IST 2000 R	14.3	80.8	28.7	0.0	28.0	27.1	15.5	61.5	30.9	0.0
WIT (1963)	0.0	19.2	0.0	100.0	0.0	9.2	5.2	0.0	4.3	0.0
LPS	13.2	0.0	11.0	0.0	0.0	25.2	14.4	0.0	0.0	100.0
WPT (1996)	16.9	0.0	14.0	0.0	0.0	32.1	8.2	28.3	15.1	0.0
andere	55.6	0.0	46.3	0.0	72.0	6.4	56.8	10.2	49.7	0.0
Gesamt	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Anmerkungen. Die Werte geben den Anteil in % an, den die Variablenausprägungen (Zeilen) an den entsprechenden Metaanalysen (Spalten) haben. D.h. die Spalten ergeben pro Bereich 100%.

E. Übersicht über inkludierte Studien und Ergebnisse auf Studienebene pro Metaanalyse

Anmerkungen zu den Tabellen: n.ber. = Anteil der insignifikanten nicht-berichteten Korrelationen in % (nur angegeben, wenn Studien mit nicht-berichteten Korrelationen vorhanden); Aggregat X = Aggregationen mehrerer Intelligenzmaße vorgenommen; Aggregat Y = Aggregationen mehrerer Erfolgsmaße vorgenommen; Art: C = lineare Verbände, M = Mittelwerte, - = keine Aggregation; Gew.: X = Gewichtung der Subprädiktoren bzw. -kriterien, - = keine Gewichtung; r = beobachtete Korrelation, bei n.ber. > 0 auf Basis der Verhältnissubstitution; Artefakt-Werte: angegeben, wenn korrigiert; r_{xx} = Reliabilität des Prädiktors; r_{yy} = Reliabilität des Kriteriums; u_x = beobachtete Varianzeinschränkung (wenn < 1) bzw. Varianzvergrößerung (wenn > 1) im Prädiktor; op. Val. = operationale Validität; vollst. korr. = vollständig korrigierte Validität; $\hat{\rho}$ = geschätzter „wahrer“ Effekt (wenn > 1 : *out of range*); w = Gewicht der Studie in %.

Tabelle E.1
Lernleistung

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
100.00	60.0	-	-	C	-	.273	76	.911	.820	1.003	.300	0.49	.315	0.50
101.01	0.0	-	-	C	-	.147	72	.863	.639	.467	.667	0.03	.718	0.03
101.02	0.0	-	-	C	-	.296	65	.863	.784	.489	.804	0.07	.866	0.07
102.00	54.5	C	-	-	-	.218	54	.918	.642	.705	.403	0.12	.421	0.13
103.00	0.0	C	X	M	-	.152	90	.914	.647	.752	.266	0.23	.278	0.24
104.00	0.0	M	X	-	-	.240	158	.911	.642	.752	.411	0.42	.431	0.43
105.00	88.9	M	X	M	X	.086	103	.911	.639	.752	.154	0.25	.161	0.26
106.00	0.0	M	-	M	X	.091	224	.895	.647	.791	.153	0.62	.162	0.62
107.01	86.4	M	-	M	X	.110	56	.911	.647	.752	.195	0.14	.204	0.14
107.02	94.4	M	-	M	X	.061	137	.911	.647	.752	.109	0.34	.114	0.35
107.03	88.9	M	-	M	X	.072	145	.911	.647	.752	.128	0.36	.134	0.37
108.00	90.9	M	-	M	-	.050	249	.911	.642	.752	.090	0.61	.094	0.62
109.00	0.0	C	-	C	-	.042	102	.882	.811	.752	.069	0.30	.074	0.29
110.00	0.0	M	X	-	-	.476	75	.911	.642	.752	.727	0.25	.761	0.26
111.01	0.0	C	-	-	-	.467	78	.913	.642	.791	.692	0.28	.724	0.28
111.02	0.0	C	-	-	-	.440	98	.913	.642	.791	.660	0.34	.691	0.35
112.00	0.0	-	-	-	-	.240	93	.843	.642	.791	.407	0.25	.443	0.24
113.00	0.0	-	-	M	X	.236	148	.863	.639	.750	.425	0.36	.457	0.35
114.00	0.0	-	-	M	-	.335	313	.911	.647	.752	.549	0.91	.576	0.93
116.01	0.0	C	-	-	-	.355	126	.882	.642	.776	.572	0.38	.609	0.38
116.03	0.0	C	-	-	-	.556	70	.882	.642	.752	.815	0.26	.868	0.25
116.04	0.0	C	-	-	-	.326	215	.882	.642	.776	.532	0.63	.566	0.63
116.06	0.0	C	-	-	-	.404	230	.882	.642	.791	.625	0.76	.665	0.75
116.08	0.0	C	-	-	-	.534	304	.882	.642	.752	.795	1.08	.846	1.06
116.09	0.0	C	-	-	-	.446	169	.882	.642	.752	.702	0.53	.748	0.53
116.10	0.0	C	-	-	-	.547	407	.882	.642	.791	.786	1.55	.837	1.53
116.11	0.0	C	-	-	-	.749	76	.882	.642	.791	.959	0.36	1.021	0.36
116.13	0.0	-	-	-	-	.760	180	.863	.642	.776	.967	0.87	1.040	0.84
117.00	0.0	-	-	-	-	.700	42	.863	.642	.791	.923	0.19	.993	0.18

Tabelle E.1
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
118.01	0.0	-	-	C	-	.493	46	.935	.642	.791	.717	0.17	.741	0.18
118.02	0.0	-	-	C	-	.570	30	.935	.642	.791	.800	0.12	.827	0.13
118.03	0.0	-	-	C	-	.528	54	.935	.642	.791	.756	0.21	.781	0.22
119.01	0.0	-	-	C	-	.210	100	.863	.680	.752	.370	0.25	.398	0.24
119.02	0.0	-	-	C	-	.216	99	.863	.680	.791	.354	0.29	.381	0.28
120.01	0.0	-	-	-	-	.530	127	.863	.642	.752	.796	0.44	.856	0.43
120.02	0.0	-	-	-	-	.600	101	.863	.642	.752	.857	0.39	.923	0.38
120.03	0.0	-	-	-	-	.570	222	.863	.642	.844	.786	0.92	.846	0.89
120.04	0.0	-	-	-	-	.590	13	.863	.639	.752	.850	0.05	.915	0.05
122.00	0.0	-	-	-	-	.560	967	.935	.642	.791	.789	3.82	.816	4.00
123.01	0.0	-	-	-	-	.710	20	.895	.520	.791	.988	0.08	1.044	0.08
123.02	0.0	-	-	-	-	.650	20	.895	.520	.791	.939	0.08	.993	0.08
123.03	0.0	-	-	-	-	.800	19	.895	.520	.791	1.050	0.09	1.110	0.09
123.04	0.0	-	-	-	-	.720	7	.895	.520	.791	.995	0.03	1.052	0.03
123.05	0.0	-	-	-	-	.590	13	.895	.520	.791	.885	0.05	.936	0.05
123.06	0.0	-	-	-	-	.850	10	.895	.520	.791	1.080	0.05	1.142	0.05
124.01	0.0	M	-	-	-	.803	12	.895	.520	.791	1.052	0.05	1.112	0.05
124.02	0.0	M	-	-	-	.806	8	.895	.520	.791	1.054	0.04	1.114	0.04
124.03	0.0	M	-	-	-	.515	12	.895	.520	.791	.809	0.04	.855	0.04
125.01	0.0	-	-	C	-	.293	738	.863	.642	.811	.466	2.29	.502	2.21
125.02	0.0	-	-	C	-	.262	221	.863	.642	.778	.444	0.60	.478	0.58
125.03	0.0	-	-	-	-	.160	196	.863	.642	.644	.377	0.28	.405	0.27
125.04	0.0	-	-	C	-	.198	198	.863	.642	.711	.390	0.40	.420	0.39
125.05	0.0	-	-	C	-	.233	285	.863	.642	.900	.331	1.11	.356	1.07
125.06	0.0	-	-	C	-	.205	120	.863	.642	.722	.394	0.26	.424	0.25
126.01	0.0	-	-	C	-	.337	322	.911	.642	.689	.601	0.79	.630	0.81
126.02	0.0	-	-	C	-	.283	37	.911	.642	.737	.487	0.10	.510	0.10
126.03	0.0	-	-	C	-	.386	64	.911	.642	.625	.717	0.15	.752	0.15
126.04	0.0	-	-	C	-	.378	81	.911	.642	.641	.694	0.19	.727	0.19
126.05	0.0	-	-	C	-	.153	59	.911	.642	.673	.311	0.11	.326	0.11
126.06	0.0	-	-	C	-	.259	83	.911	.642	.609	.557	0.14	.583	0.14
127.01	0.0	M	X	C	-	.211	33	.895	.642	.791	.348	0.10	.368	0.10
127.02	0.0	M	X	C	-	.337	30	.895	.642	.791	.533	0.09	.563	0.09
129.00	50.0	M	X	-	-	.082	93	.911	.642	.752	.146	0.23	.153	0.23
130.00	0.0	M	X	-	-	.305	48	.911	.639	.752	.510	0.13	.534	0.14
133.00	0.0	-	-	-	-	.330	112	.925	.642	.785	.518	0.36	.539	0.37
134.00	0.0	-	-	-	-	.720	27	.911	.642	.890	.920	0.13	.964	0.13
135.00	0.0	-	-	M	-	.470	861	.895	.647	.791	.698	3.06	.738	3.07
136.00	0.0	-	-	-	-	.360	97	.869	.642	.776	.583	0.29	.625	0.28
137.01	0.0	-	-	-	-	.200	74	.863	.642	.752	.363	0.18	.391	0.17
137.02	0.0	-	-	-	-	.400	64	.863	.642	.791	.625	0.21	.673	0.20
137.03	0.0	-	-	-	-	.480	44	.863	.642	.791	.720	0.15	.775	0.15
137.04	0.0	-	-	-	-	.190	59	.863	.642	.791	.322	0.16	.347	0.16
137.05	0.0	-	-	-	-	.100	27	.863	.642	.791	.173	0.07	.186	0.07
137.06	0.0	-	-	-	-	.280	49	.863	.642	.791	.461	0.14	.497	0.14
137.07	0.0	-	-	-	-	.120	50	.863	.642	.791	.207	0.13	.223	0.13
138.00	86.1	C	-	M	-	.236	22	.895	.520	.791	.426	0.05	.450	0.05
139.00	72.2	M	-	M	-	.104	119	.895	.642	.791	.175	0.33	.185	0.33
140.00	0.0	C	-	-	-	.469	146	.940	.642	.791	.688	0.53	.710	0.56
141.00	0.0	-	-	-	-	.750	34	.863	.642	.791	.959	0.16	1.032	0.16
142.00	0.0	-	-	-	-	.750	8	.863	.642	.752	.962	0.04	1.035	0.04
143.01	0.0	-	-	-	-	.322	123	.863	.642	.791	.522	0.37	.561	0.36
143.02	0.0	-	-	-	-	.644	147	.863	.642	.791	.878	0.62	.946	0.60
143.03	0.0	-	-	-	-	.644	62	.863	.642	.791	.878	0.26	.946	0.25
144.00	0.0	-	-	-	-	.560	315	.935	.642	.791	.789	1.24	.816	1.30
146.00	75.9	M	-	-	-	.120	45	.869	.642	.791	.206	0.12	.221	0.12
147.00	0.0	C	X	-	-	.196	308	.895	.520	.776	.367	0.69	.388	0.69
148.00	0.0	C	-	-	-	.521	91	.869	.642	.700	.817	0.29	.876	0.28
149.00	0.0	M	-	C	-	.093	199	.895	.642	.752	.168	0.48	.178	0.48

Tabelle E.1
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
150.00	0.0	-	-	-	-	.322	167	.895	.642	.776	.522	0.50	.552	0.50
151.00	0.0	C	-	M	-	.200	200	.935	.647	.791	.321	0.61	.332	0.64
152.00	0.0	C	-	M	-	.529	180	.934	.639	.939	.688	0.83	.712	0.87
153.00	0.0	-	-	-	-	.570	153	.870	.639	.737	.839	0.55	.900	0.54
156.00	0.0	-	-	-	-	.310	122	.911	.642	.970	.399	0.58	.418	0.59
159.00	0.0	C	X	C	-	.523	150	.877	.642	.776	.771	0.54	.824	0.53
160.01	0.0	C	-	-	-	.351	58	.861	.642	.926	.475	0.25	.512	0.24
160.02	0.0	C	-	-	-	.160	91	.855	.642	.867	.242	0.31	.261	0.30
161.01	0.0	M	-	-	-	.016	60	.895	.642	.873	.024	0.21	.025	0.21
161.02	0.0	M	-	-	-	.086	31	.895	.642	.752	.156	0.07	.165	0.07
164.01	0.0	-	-	-	-	.245	486	.895	.642	.791	.400	1.43	.423	1.43
164.02	0.0	-	-	-	-	.350	198	.895	.642	.791	.550	0.63	.582	0.63
167.00	0.0	C	X	C	-	.078	139	.895	.639	.791	.132	0.38	.140	0.38
168.00	0.0	-	-	M	-	.476	51	.895	.639	.791	.708	0.18	.749	0.18
171.01	0.0	-	-	C	-	.390	197	.875	.740	.950	.478	1.03	.511	1.01
171.02	0.0	-	-	C	-	.332	142	.875	.720	1.018	.384	0.83	.410	0.82
171.03	0.0	-	-	C	-	.277	141	.875	.711	1.084	.300	0.94	.320	0.93
171.04	0.0	-	-	C	-	-.167	21	.875	.748	.757	-.279	0.06	-.298	0.06
172.00	0.0	-	-	C	-	.196	97	.869	.665	1.034	.231	0.55	.248	0.53
173.00	0.0	-	-	-	-	.260	190	.863	.642	.700	.505	0.40	.543	0.38
174.00	0.0	-	-	M	X	.315	469	.895	.639	.791	.503	1.44	.532	1.45
175.00	45.5	-	-	M	X	.296	187	.935	.639	.791	.465	0.59	.481	0.62
176.01	0.0	C	-	-	-	.515	23	.915	.642	.791	.746	0.09	.780	0.09
176.02	0.0	C	-	-	-	.563	16	.915	.642	.791	.796	0.06	.832	0.06
177.01	33.3	M	-	-	-	.221	27	.935	.642	.791	.354	0.08	.366	0.09
177.02	66.7	M	-	-	-	.100	27	.935	.642	.791	.164	0.08	.169	0.08
178.00	0.0	C	-	M	X	.211	176	.911	.647	.752	.364	0.46	.381	0.47
179.01	0.0	C	-	-	-	.027	12	.947	.642	.907	.038	0.05	.039	0.05
179.02	0.0	C	-	-	-	.711	15	.947	.642	.907	.906	0.07	.931	0.08
179.03	0.0	C	-	-	-	.226	9	.947	.642	.907	.312	0.04	.320	0.04
179.04	0.0	C	-	-	-	.336	10	.947	.642	.907	.458	0.04	.471	0.04
179.05	0.0	C	-	-	-	.793	7	.947	.642	.907	.991	0.04	1.019	0.04
179.06	0.0	C	-	-	-	.141	15	.947	.642	.907	.196	0.06	.201	0.06
179.07	0.0	C	-	-	-	.396	14	.947	.642	.907	.536	0.06	.550	0.06
179.08	0.0	M	-	-	-	.477	51	.895	.642	.791	.708	0.18	.749	0.18
179.09	0.0	M	-	-	-	.363	64	.895	.642	.791	.568	0.21	.600	0.21
179.10	0.0	M	-	-	-	.480	23	.895	.642	.791	.712	0.08	.752	0.08
179.11	0.0	M	-	-	-	.273	26	.895	.642	.791	.442	0.08	.467	0.08
179.12	0.0	M	-	-	-	.530	20	.895	.642	.752	.788	0.07	.832	0.07
179.13	0.0	M	-	-	-	.393	52	.895	.642	.752	.634	0.16	.670	0.16
179.14	0.0	M	-	-	-	.443	19	.895	.642	.752	.694	0.06	.734	0.06
181.01	0.0	-	-	-	-	.120	280	.869	.642	.776	.212	0.71	.227	0.69
181.02	0.0	-	-	-	-	.250	303	.869	.642	.776	.425	0.82	.456	0.80
181.03	0.0	-	-	-	-	.300	70	.869	.642	.776	.499	0.20	.536	0.19
181.04	0.0	-	-	-	-	.100	110	.869	.642	.776	.177	0.28	.190	0.27
182.00	0.0	-	-	-	-	.090	200	.895	.642	.776	.156	0.52	.165	0.52
183.00	0.0	-	-	-	-	.070	256	.895	.642	.752	.127	0.61	.134	0.61
184.01	0.0	-	-	-	-	.220	152	.935	.642	.721	.390	0.38	.403	0.40
184.02	0.0	-	-	-	-	.040	26	.935	.642	.721	.074	0.06	.076	0.06
184.03	0.0	-	-	-	-	.310	89	.920	.642	.741	.520	0.25	.542	0.26
185.00	0.0	-	-	M	-	.288	90	.895	.642	.776	.473	0.26	.500	0.26
186.00	0.0	-	-	-	-	.320	167	.895	.642	.776	.519	0.50	.549	0.50
189.01	0.0	C	-	-	-	.432	102	.895	.639	.752	.683	0.32	.722	0.32
189.02	0.0	C	-	-	-	.338	403	.895	.639	.791	.535	1.26	.566	1.26
189.03	0.0	C	-	-	-	.437	282	.895	.639	.791	.663	0.96	.700	0.96
189.04	0.0	C	-	-	-	.377	87	.895	.639	.791	.587	0.28	.621	0.28
189.05	0.0	C	-	-	-	.518	68	.895	.639	.791	.755	0.25	.798	0.25
189.06	0.0	C	-	-	-	.501	731	.895	.639	.791	.736	2.66	.778	2.66
189.07	0.0	C	-	-	-	.437	198	.895	.639	.791	.663	0.68	.700	0.68

Tabelle E.1
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.08	0.0	C	-	-	-	.372	112	.895	.639	.791	.581	0.36	.614	0.36
189.09	0.0	C	-	-	-	.398	351	.895	.639	.791	.614	1.16	.649	1.16
189.10	0.0	C	-	-	-	.317	394	.895	.639	.791	.506	1.21	.535	1.22
189.11	0.0	C	-	-	-	.218	81	.895	.639	.791	.359	0.23	.380	0.23
189.12	0.0	C	-	-	-	.368	191	.895	.639	.791	.575	0.61	.608	0.61
189.13	0.0	C	-	-	-	.368	97	.895	.639	.791	.575	0.31	.608	0.31
189.14	0.0	C	-	-	-	.616	617	.895	.639	.791	.852	2.53	.900	2.54
189.15	0.0	C	-	-	-	.475	418	.895	.639	.791	.707	1.48	.748	1.48
189.16	0.0	C	-	-	-	.428	412	.895	.639	.791	.652	1.39	.689	1.40
189.17	0.0	C	-	-	-	.394	243	.895	.639	.791	.609	0.80	.644	0.80
189.18	0.0	C	-	-	-	.330	98	.895	.639	.791	.524	0.30	.554	0.31
189.19	0.0	C	-	-	-	.561	287	.895	.639	.791	.799	1.11	.845	1.11
189.20	0.0	C	-	-	-	.411	291	.895	.639	.791	.631	0.97	.667	0.97
189.21	0.0	C	-	-	-	.278	184	.895	.639	.791	.450	0.55	.476	0.55
189.22	0.0	C	-	-	-	.385	84	.895	.639	.791	.598	0.27	.632	0.27
189.23	0.0	C	-	-	-	.214	158	.895	.639	.791	.353	0.46	.373	0.46
189.24	0.0	C	-	-	-	.300	119	.895	.639	.791	.482	0.36	.509	0.36
189.25	0.0	C	-	-	-	.381	67	.895	.639	.791	.593	0.22	.626	0.22
189.26	0.0	C	-	-	-	.407	57	.895	.639	.791	.626	0.19	.661	0.19
189.27	0.0	C	-	-	-	.377	81	.895	.639	.776	.598	0.25	.632	0.25
189.28	0.0	C	-	-	-	.479	181	.895	.639	.791	.712	0.64	.752	0.64
189.29	0.0	C	-	-	-	.347	314	.895	.639	.791	.547	0.99	.579	0.99
189.30	0.0	C	-	-	-	.407	454	.895	.639	.791	.626	1.51	.661	1.51
189.31	0.0	C	-	-	-	.501	380	.895	.639	.791	.736	1.38	.778	1.38
189.32	0.0	C	-	-	-	.381	412	.895	.639	.791	.593	1.34	.626	1.34
189.33	0.0	C	-	-	-	.488	391	.895	.639	.791	.722	1.40	.763	1.40
189.34	0.0	C	-	-	-	.321	48	.895	.639	.791	.512	0.15	.541	0.15
189.35	0.0	C	-	-	-	.390	58	.895	.639	.752	.631	0.17	.667	0.17
189.36	0.0	C	-	-	-	.424	57	.895	.639	.752	.673	0.18	.712	0.18
189.37	0.0	C	-	-	-	.223	87	.895	.639	.752	.390	0.22	.413	0.22
189.38	0.0	C	-	-	-	.334	56	.895	.639	.791	.530	0.17	.560	0.18
189.39	0.0	C	-	-	-	.248	48	.895	.639	.791	.405	0.14	.428	0.14
189.40	0.0	C	-	-	-	.240	61	.895	.639	.791	.393	0.18	.416	0.18
189.41	0.0	C	-	-	-	.214	104	.895	.639	.791	.353	0.30	.373	0.30
189.42	0.0	C	-	-	-	.227	113	.895	.639	.791	.373	0.33	.395	0.33
189.43	0.0	C	-	-	-	.338	38	.895	.639	.791	.535	0.12	.566	0.12
189.44	0.0	C	-	-	-	.295	21	.895	.639	.791	.475	0.06	.502	0.06
189.45	0.0	C	-	-	-	.390	178	.895	.639	.752	.631	0.53	.667	0.53
189.46	0.0	C	-	-	-	.432	347	.895	.639	.752	.683	1.09	.722	1.09
189.47	0.0	C	-	-	-	.454	37	.895	.639	.752	.708	0.12	.749	0.12
189.48	0.0	C	-	-	-	.424	41	.895	.639	.752	.673	0.13	.712	0.13
189.49	0.0	C	-	-	-	.595	104	.895	.639	.752	.849	0.40	.897	0.40
189.50	0.0	C	-	-	-	.454	174	.895	.639	.752	.708	0.56	.749	0.56
189.51	0.0	C	-	-	-	.257	104	.895	.639	.791	.419	0.31	.443	0.31
189.52	0.0	C	-	-	-	.509	32	.895	.639	.752	.768	0.11	.811	0.11
189.53	0.0	C	-	-	-	.574	268	.895	.639	.791	.812	1.05	.858	1.05
189.54	0.0	C	-	-	-	.432	212	.895	.639	.791	.657	0.72	.694	0.72
189.55	0.0	C	-	-	-	.488	98	.895	.639	.791	.722	0.35	.763	0.35
189.56	0.0	C	-	-	-	.484	87	.895	.639	.791	.717	0.31	.758	0.31
189.57	0.0	C	-	-	-	.424	106	.895	.639	.791	.647	0.36	.684	0.36
189.58	0.0	C	-	-	-	.287	124	.895	.639	.791	.463	0.37	.489	0.37
189.59	0.0	C	-	-	-	.428	168	.895	.639	.791	.652	0.57	.689	0.57
189.60	0.0	C	-	-	-	.372	117	.895	.639	.791	.581	0.38	.614	0.38
189.61	0.0	C	-	-	-	.407	198	.895	.639	.791	.626	0.66	.661	0.66
189.62	0.0	C	-	-	-	.390	314	.895	.639	.791	.604	1.03	.639	1.03
189.63	0.0	C	-	-	-	.420	217	.895	.639	.791	.642	0.73	.679	0.73
189.64	0.0	C	-	-	-	.462	208	.895	.639	.791	.692	0.73	.732	0.73
189.65	0.0	C	-	-	-	.458	197	.895	.639	.791	.688	0.69	.727	0.69
189.66	0.0	C	-	-	-	.458	128	.895	.639	.791	.688	0.45	.727	0.45

Tabelle E.1
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.67	0.0	C	-	-	-	.471	117	.895	.639	.791	.703	0.41	.743	0.41
189.68	0.0	C	-	-	-	.462	89	.895	.639	.791	.692	0.31	.732	0.31
189.69	0.0	C	-	-	-	.492	58	.895	.639	.791	.726	0.21	.768	0.21
189.70	0.0	C	-	-	-	.492	87	.895	.639	.752	.750	0.29	.793	0.29

Anmerkungen. siehe Seite 275.

Tabelle E.2
Lernleistung, Studien ohne Mittelwert-Aggregate

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
100.00	60.0	-	-	C	-	.268	76	.911	.820	1.003	.295	0.58	.309	0.59
101.01	0.0	-	-	C	-	.147	72	.863	.639	.467	.667	0.03	.718	0.03
101.02	0.0	-	-	C	-	.296	65	.863	.784	.489	.804	0.08	.866	0.08
102.00	54.5	C	-	-	-	.212	54	.918	.642	.705	.393	0.14	.410	0.15
109.00	0.0	C	-	C	-	.042	102	.882	.811	.752	.069	0.35	.074	0.34
111.01	0.0	C	-	-	-	.467	78	.913	.642	.791	.692	0.33	.724	0.34
111.02	0.0	C	-	-	-	.440	98	.913	.642	.791	.660	0.40	.691	0.41
112.00	0.0	-	-	-	-	.240	93	.843	.642	.791	.407	0.30	.443	0.28
116.01	0.0	C	-	-	-	.355	126	.882	.642	.776	.572	0.45	.609	0.44
116.03	0.0	C	-	-	-	.556	70	.882	.642	.752	.815	0.30	.868	0.30
116.04	0.0	C	-	-	-	.326	215	.882	.642	.776	.532	0.74	.566	0.74
116.06	0.0	C	-	-	-	.404	230	.882	.642	.791	.625	0.89	.665	0.88
116.08	0.0	C	-	-	-	.534	304	.882	.642	.752	.795	1.27	.846	1.25
116.09	0.0	C	-	-	-	.446	169	.882	.642	.752	.702	0.63	.748	0.62
116.10	0.0	C	-	-	-	.547	407	.882	.642	.791	.786	1.82	.837	1.80
116.11	0.0	C	-	-	-	.749	76	.882	.642	.791	.959	0.43	1.021	0.42
116.13	0.0	-	-	-	-	.760	180	.863	.642	.776	.967	1.03	1.040	0.99
117.00	0.0	-	-	-	-	.700	42	.863	.642	.791	.923	0.22	.993	0.22
118.01	0.0	-	-	C	-	.493	46	.935	.642	.791	.717	0.20	.741	0.21
118.02	0.0	-	-	C	-	.570	30	.935	.642	.791	.800	0.14	.827	0.15
118.03	0.0	-	-	C	-	.528	54	.935	.642	.791	.756	0.24	.781	0.25
119.01	0.0	-	-	C	-	.210	100	.863	.680	.752	.370	0.30	.398	0.29
119.02	0.0	-	-	C	-	.216	99	.863	.680	.791	.354	0.34	.381	0.33
120.01	0.0	-	-	-	-	.530	127	.863	.642	.752	.796	0.52	.856	0.50
120.02	0.0	-	-	-	-	.600	101	.863	.642	.752	.857	0.46	.923	0.44
120.03	0.0	-	-	-	-	.570	222	.863	.642	.844	.786	1.08	.846	1.04
120.04	0.0	-	-	-	-	.590	13	.863	.639	.752	.850	0.06	.915	0.06
122.00	0.0	-	-	-	-	.560	967	.935	.642	.791	.789	4.49	.816	4.71
123.01	0.0	-	-	-	-	.710	20	.895	.520	.791	.988	0.10	1.044	0.10
123.02	0.0	-	-	-	-	.650	20	.895	.520	.791	.939	0.09	.993	0.09
123.03	0.0	-	-	-	-	.800	19	.895	.520	.791	1.050	0.10	1.110	0.10
123.04	0.0	-	-	-	-	.720	7	.895	.520	.791	.995	0.03	1.052	0.03
123.05	0.0	-	-	-	-	.590	13	.895	.520	.791	.885	0.05	.936	0.05
123.06	0.0	-	-	-	-	.850	10	.895	.520	.791	1.080	0.06	1.142	0.06
125.01	0.0	-	-	C	-	.293	738	.863	.642	.811	.466	2.69	.502	2.60
125.02	0.0	-	-	C	-	.262	221	.863	.642	.778	.444	0.71	.478	0.69
125.03	0.0	-	-	-	-	.160	196	.863	.642	.644	.377	0.33	.405	0.32
125.04	0.0	-	-	C	-	.198	198	.863	.642	.711	.390	0.47	.420	0.46
125.05	0.0	-	-	C	-	.233	285	.863	.642	.900	.331	1.30	.356	1.26
125.06	0.0	-	-	C	-	.205	120	.863	.642	.722	.394	0.30	.424	0.29
126.01	0.0	-	-	C	-	.337	322	.911	.642	.689	.601	0.93	.630	0.95
126.02	0.0	-	-	C	-	.283	37	.911	.642	.737	.487	0.12	.510	0.12
126.03	0.0	-	-	C	-	.386	64	.911	.642	.625	.717	0.17	.752	0.17

Tabelle E.2
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
126.04	0.0	-	-	C	-	.378	81	.911	.642	.641	.694	0.22	.727	0.23
126.05	0.0	-	-	C	-	.153	59	.911	.642	.673	.311	0.13	.326	0.13
126.06	0.0	-	-	C	-	.259	83	.911	.642	.609	.557	0.17	.583	0.17
133.00	0.0	-	-	-	-	.330	112	.925	.642	.785	.518	0.42	.539	0.43
134.00	0.0	-	-	-	-	.720	27	.911	.642	.890	.920	0.15	.964	0.16
136.00	0.0	-	-	-	-	.360	97	.869	.642	.776	.583	0.34	.625	0.33
137.01	0.0	-	-	-	-	.200	74	.863	.642	.752	.363	0.21	.391	0.20
137.02	0.0	-	-	-	-	.400	64	.863	.642	.791	.625	0.24	.673	0.23
137.03	0.0	-	-	-	-	.480	44	.863	.642	.791	.720	0.18	.775	0.17
137.04	0.0	-	-	-	-	.190	59	.863	.642	.791	.322	0.19	.347	0.18
137.05	0.0	-	-	-	-	.100	27	.863	.642	.791	.173	0.08	.186	0.08
137.06	0.0	-	-	-	-	.280	49	.863	.642	.791	.461	0.17	.497	0.16
137.07	0.0	-	-	-	-	.120	50	.863	.642	.791	.207	0.16	.223	0.15
140.00	0.0	C	-	-	-	.469	146	.940	.642	.791	.688	0.63	.710	0.66
141.00	0.0	-	-	-	-	.750	34	.863	.642	.791	.959	0.19	1.032	0.19
142.00	0.0	-	-	-	-	.750	8	.863	.642	.752	.962	0.04	1.035	0.04
143.01	0.0	-	-	-	-	.322	123	.863	.642	.791	.522	0.43	.561	0.42
143.02	0.0	-	-	-	-	.644	147	.863	.642	.791	.878	0.73	.946	0.71
143.03	0.0	-	-	-	-	.644	62	.863	.642	.791	.878	0.31	.946	0.30
144.00	0.0	-	-	-	-	.560	315	.935	.642	.791	.789	1.46	.816	1.53
147.00	0.0	C	X	-	-	.196	308	.895	.520	.776	.367	0.81	.388	0.81
148.00	0.0	C	-	-	-	.521	91	.869	.642	.700	.817	0.34	.876	0.33
150.00	0.0	-	-	-	-	.322	167	.895	.642	.776	.522	0.59	.552	0.59
153.00	0.0	-	-	-	-	.570	153	.870	.639	.737	.839	0.65	.900	0.64
156.00	0.0	-	-	-	-	.310	122	.911	.642	.970	.399	0.68	.418	0.69
159.00	0.0	C	X	C	-	.523	150	.877	.642	.776	.771	0.64	.824	0.63
160.01	0.0	C	-	-	-	.351	58	.861	.642	.926	.475	0.29	.512	0.28
160.02	0.0	C	-	-	-	.160	91	.855	.642	.867	.242	0.37	.261	0.35
164.01	0.0	-	-	-	-	.245	486	.895	.642	.791	.400	1.68	.423	1.69
164.02	0.0	-	-	-	-	.350	198	.895	.642	.791	.550	0.74	.582	0.74
167.00	0.0	C	X	C	-	.078	139	.895	.639	.791	.132	0.45	.140	0.45
171.01	0.0	-	-	C	-	.390	197	.875	.740	.950	.478	1.21	.511	1.19
171.02	0.0	-	-	C	-	.332	142	.875	.720	1.018	.384	0.98	.410	0.96
171.03	0.0	-	-	C	-	.277	141	.875	.711	1.084	.300	1.11	.320	1.09
171.04	0.0	-	-	C	-	-.167	21	.875	.748	.757	-.279	0.07	-.298	0.07
172.00	0.0	-	-	C	-	.196	97	.869	.665	1.034	.231	0.65	.248	0.63
173.00	0.0	-	-	-	-	.260	190	.863	.642	.700	.505	0.47	.543	0.45
176.01	0.0	C	-	-	-	.515	23	.915	.642	.791	.746	0.10	.780	0.10
176.02	0.0	C	-	-	-	.563	16	.915	.642	.791	.796	0.07	.832	0.08
179.01	0.0	C	-	-	-	.027	12	.947	.642	.907	.038	0.06	.039	0.06
179.02	0.0	C	-	-	-	.711	15	.947	.642	.907	.906	0.09	.931	0.09
179.03	0.0	C	-	-	-	.226	9	.947	.642	.907	.312	0.04	.320	0.05
179.04	0.0	C	-	-	-	.336	10	.947	.642	.907	.458	0.05	.471	0.05
179.05	0.0	C	-	-	-	.793	7	.947	.642	.907	.991	0.04	1.019	0.04
179.06	0.0	C	-	-	-	.141	15	.947	.642	.907	.196	0.07	.201	0.08
179.07	0.0	C	-	-	-	.396	14	.947	.642	.907	.536	0.07	.550	0.07
181.01	0.0	-	-	-	-	.120	280	.869	.642	.776	.212	0.83	.227	0.81
181.02	0.0	-	-	-	-	.250	303	.869	.642	.776	.425	0.97	.456	0.94
181.03	0.0	-	-	-	-	.300	70	.869	.642	.776	.499	0.23	.536	0.23
181.04	0.0	-	-	-	-	.100	110	.869	.642	.776	.177	0.32	.190	0.32
182.00	0.0	-	-	-	-	.090	200	.895	.642	.776	.156	0.61	.165	0.62
183.00	0.0	-	-	-	-	.070	256	.895	.642	.752	.127	0.72	.134	0.72
184.01	0.0	-	-	-	-	.220	152	.935	.642	.721	.390	0.45	.403	0.47
184.02	0.0	-	-	-	-	.040	26	.935	.642	.721	.074	0.07	.076	0.07
184.03	0.0	-	-	-	-	.310	89	.920	.642	.741	.520	0.29	.542	0.30
186.00	0.0	-	-	-	-	.320	167	.895	.642	.776	.519	0.59	.549	0.59
189.01	0.0	C	-	-	-	.432	102	.895	.639	.752	.683	0.38	.722	0.38
189.02	0.0	C	-	-	-	.338	403	.895	.639	.791	.535	1.48	.566	1.49
189.03	0.0	C	-	-	-	.437	282	.895	.639	.791	.663	1.13	.700	1.14

Tabelle E.2
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.04	0.0	C	-	-	-	.377	87	.895	.639	.791	.587	0.33	.621	0.33
189.05	0.0	C	-	-	-	.518	68	.895	.639	.791	.755	0.30	.798	0.30
189.06	0.0	C	-	-	-	.501	731	.895	.639	.791	.736	3.12	.778	3.13
189.07	0.0	C	-	-	-	.437	198	.895	.639	.791	.663	0.79	.700	0.80
189.08	0.0	C	-	-	-	.372	112	.895	.639	.791	.581	0.42	.614	0.43
189.09	0.0	C	-	-	-	.398	351	.895	.639	.791	.614	1.36	.649	1.36
189.10	0.0	C	-	-	-	.317	394	.895	.639	.791	.506	1.43	.535	1.43
189.11	0.0	C	-	-	-	.218	81	.895	.639	.791	.359	0.27	.380	0.28
189.12	0.0	C	-	-	-	.368	191	.895	.639	.791	.575	0.72	.608	0.72
189.13	0.0	C	-	-	-	.368	97	.895	.639	.791	.575	0.37	.608	0.37
189.14	0.0	C	-	-	-	.616	617	.895	.639	.791	.852	2.98	.900	2.99
189.15	0.0	C	-	-	-	.475	418	.895	.639	.791	.707	1.74	.748	1.75
189.16	0.0	C	-	-	-	.428	412	.895	.639	.791	.652	1.64	.689	1.64
189.17	0.0	C	-	-	-	.394	243	.895	.639	.791	.609	0.94	.644	0.94
189.18	0.0	C	-	-	-	.330	98	.895	.639	.791	.524	0.36	.554	0.36
189.19	0.0	C	-	-	-	.561	287	.895	.639	.791	.799	1.30	.845	1.31
189.20	0.0	C	-	-	-	.411	291	.895	.639	.791	.631	1.14	.667	1.14
189.21	0.0	C	-	-	-	.278	184	.895	.639	.791	.450	0.65	.476	0.65
189.22	0.0	C	-	-	-	.385	84	.895	.639	.791	.598	0.32	.632	0.32
189.23	0.0	C	-	-	-	.214	158	.895	.639	.791	.353	0.53	.373	0.54
189.24	0.0	C	-	-	-	.300	119	.895	.639	.791	.482	0.43	.509	0.43
189.25	0.0	C	-	-	-	.381	67	.895	.639	.791	.593	0.26	.626	0.26
189.26	0.0	C	-	-	-	.407	57	.895	.639	.791	.626	0.22	.661	0.22
189.27	0.0	C	-	-	-	.377	81	.895	.639	.776	.598	0.30	.632	0.30
189.28	0.0	C	-	-	-	.479	181	.895	.639	.791	.712	0.76	.752	0.76
189.29	0.0	C	-	-	-	.347	314	.895	.639	.791	.547	1.16	.579	1.17
189.30	0.0	C	-	-	-	.407	454	.895	.639	.791	.626	1.77	.661	1.78
189.31	0.0	C	-	-	-	.501	380	.895	.639	.791	.736	1.62	.778	1.63
189.32	0.0	C	-	-	-	.381	412	.895	.639	.791	.593	1.57	.626	1.58
189.33	0.0	C	-	-	-	.488	391	.895	.639	.791	.722	1.65	.763	1.65
189.34	0.0	C	-	-	-	.321	48	.895	.639	.791	.512	0.17	.541	0.17
189.35	0.0	C	-	-	-	.390	58	.895	.639	.752	.631	0.20	.667	0.20
189.36	0.0	C	-	-	-	.424	57	.895	.639	.752	.673	0.21	.712	0.21
189.37	0.0	C	-	-	-	.223	87	.895	.639	.752	.390	0.26	.413	0.26
189.38	0.0	C	-	-	-	.334	56	.895	.639	.791	.530	0.21	.560	0.21
189.39	0.0	C	-	-	-	.248	48	.895	.639	.791	.405	0.17	.428	0.17
189.40	0.0	C	-	-	-	.240	61	.895	.639	.791	.393	0.21	.416	0.21
189.41	0.0	C	-	-	-	.214	104	.895	.639	.791	.353	0.35	.373	0.35
189.42	0.0	C	-	-	-	.227	113	.895	.639	.791	.373	0.39	.395	0.39
189.43	0.0	C	-	-	-	.338	38	.895	.639	.791	.535	0.14	.566	0.14
189.44	0.0	C	-	-	-	.295	21	.895	.639	.791	.475	0.07	.502	0.08
189.45	0.0	C	-	-	-	.390	178	.895	.639	.752	.631	0.63	.667	0.63
189.46	0.0	C	-	-	-	.432	347	.895	.639	.752	.683	1.28	.722	1.29
189.47	0.0	C	-	-	-	.454	37	.895	.639	.752	.708	0.14	.749	0.14
189.48	0.0	C	-	-	-	.424	41	.895	.639	.752	.673	0.15	.712	0.15
189.49	0.0	C	-	-	-	.595	104	.895	.639	.752	.849	0.47	.897	0.47
189.50	0.0	C	-	-	-	.454	174	.895	.639	.752	.708	0.66	.749	0.66
189.51	0.0	C	-	-	-	.257	104	.895	.639	.791	.419	0.36	.443	0.36
189.52	0.0	C	-	-	-	.509	32	.895	.639	.752	.768	0.13	.811	0.13
189.53	0.0	C	-	-	-	.574	268	.895	.639	.791	.812	1.23	.858	1.24
189.54	0.0	C	-	-	-	.432	212	.895	.639	.791	.657	0.85	.694	0.85
189.55	0.0	C	-	-	-	.488	98	.895	.639	.791	.722	0.41	.763	0.41
189.56	0.0	C	-	-	-	.484	87	.895	.639	.791	.717	0.37	.758	0.37
189.57	0.0	C	-	-	-	.424	106	.895	.639	.791	.647	0.42	.684	0.42
189.58	0.0	C	-	-	-	.287	124	.895	.639	.791	.463	0.44	.489	0.44
189.59	0.0	C	-	-	-	.428	168	.895	.639	.791	.652	0.67	.689	0.67
189.60	0.0	C	-	-	-	.372	117	.895	.639	.791	.581	0.44	.614	0.44
189.61	0.0	C	-	-	-	.407	198	.895	.639	.791	.626	0.77	.661	0.78
189.62	0.0	C	-	-	-	.390	314	.895	.639	.791	.604	1.21	.639	1.21

Tabelle E.2
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.63	0.0	C	-	-	-	.420	217	.895	.639	.791	.642	0.86	.679	0.86
189.64	0.0	C	-	-	-	.462	208	.895	.639	.791	.692	0.85	.732	0.86
189.65	0.0	C	-	-	-	.458	197	.895	.639	.791	.688	0.81	.727	0.81
189.66	0.0	C	-	-	-	.458	128	.895	.639	.791	.688	0.52	.727	0.53
189.67	0.0	C	-	-	-	.471	117	.895	.639	.791	.703	0.48	.743	0.49
189.68	0.0	C	-	-	-	.462	89	.895	.639	.791	.692	0.37	.732	0.37
189.69	0.0	C	-	-	-	.492	58	.895	.639	.791	.726	0.25	.768	0.25
189.70	0.0	C	-	-	-	.492	87	.895	.639	.752	.750	0.35	.793	0.35

Anmerkungen. siehe Seite 275.

Tabelle E.3
Lernleistung, Studien mit Mittelwert-Aggregaten

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
103.00	0.0	C	X	M	-	.152	90	.914	.647	.752	.266	1.6	.278	1.6
104.00	0.0	M	X	-	-	.240	158	.911	.642	.752	.411	2.8	.431	2.9
105.00	88.9	M	X	M	X	.090	103	.911	.639	.752	.161	1.7	.168	1.7
106.00	0.0	M	-	M	X	.091	224	.895	.647	.791	.153	4.2	.162	4.1
107.01	86.4	M	-	M	X	.116	56	.911	.647	.752	.205	0.9	.215	1.0
107.02	94.4	M	-	M	X	.065	137	.911	.647	.752	.116	2.3	.121	2.3
107.03	88.9	M	-	M	X	.076	145	.911	.647	.752	.135	2.4	.142	2.4
108.00	90.9	M	-	M	-	.052	249	.911	.642	.752	.093	4.1	.098	4.1
110.00	0.0	M	X	-	-	.476	75	.911	.642	.752	.727	1.7	.761	1.7
113.00	0.0	-	-	M	X	.236	148	.863	.639	.750	.425	2.4	.457	2.3
114.00	0.0	-	-	M	-	.335	313	.911	.647	.752	.549	6.1	.576	6.2
124.01	0.0	M	-	-	-	.803	12	.895	.520	.791	1.052	0.4	1.112	0.4
124.02	0.0	M	-	-	-	.806	8	.895	.520	.791	1.054	0.3	1.114	0.2
124.03	0.0	M	-	-	-	.515	12	.895	.520	.791	.809	0.3	.855	0.3
127.01	0.0	M	X	C	-	.211	33	.895	.642	.791	.348	0.6	.368	0.6
127.02	0.0	M	X	C	-	.337	30	.895	.642	.791	.533	0.6	.563	0.6
129.00	50.0	M	X	-	-	.085	93	.911	.642	.752	.152	1.5	.159	1.6
130.00	0.0	M	X	-	-	.305	48	.911	.639	.752	.510	0.9	.534	0.9
135.00	0.0	-	-	M	-	.470	861	.895	.647	.791	.698	20.6	.738	20.3
138.00	86.1	C	-	M	-	.248	22	.895	.520	.791	.445	0.4	.471	0.4
139.00	72.2	M	-	M	-	.107	119	.895	.642	.791	.180	2.2	.190	2.2
146.00	75.9	M	-	-	-	.125	45	.869	.642	.791	.214	0.8	.230	0.8
149.00	0.0	M	-	C	-	.093	199	.895	.642	.752	.168	3.2	.178	3.2
151.00	0.0	C	-	M	-	.200	200	.935	.647	.791	.321	4.1	.332	4.2
152.00	0.0	C	-	M	-	.529	180	.934	.639	.939	.688	5.6	.712	5.8
161.01	0.0	M	-	-	-	.016	60	.895	.642	.873	.024	1.4	.025	1.4
161.02	0.0	M	-	-	-	.086	31	.895	.642	.752	.156	0.5	.165	0.5
168.00	0.0	-	-	M	-	.476	51	.895	.639	.791	.708	1.2	.749	1.2
174.00	0.0	-	-	M	X	.315	469	.895	.639	.791	.503	9.7	.532	9.6
175.00	45.5	-	-	M	X	.297	187	.935	.639	.791	.467	4.0	.483	4.1
177.01	33.3	M	-	-	-	.222	27	.935	.642	.791	.355	0.6	.368	0.6
177.02	66.7	M	-	-	-	.102	27	.935	.642	.791	.167	0.5	.173	0.6
178.00	0.0	C	-	M	X	.211	176	.911	.647	.752	.364	3.1	.381	3.1
179.08	0.0	M	-	-	-	.477	51	.895	.642	.791	.708	1.2	.749	1.2
179.09	0.0	M	-	-	-	.363	64	.895	.642	.791	.568	1.4	.600	1.4
179.10	0.0	M	-	-	-	.480	23	.895	.642	.791	.712	0.6	.752	0.6
179.11	0.0	M	-	-	-	.273	26	.895	.642	.791	.442	0.5	.467	0.5
179.12	0.0	M	-	-	-	.530	20	.895	.642	.752	.788	0.5	.832	0.5
179.13	0.0	M	-	-	-	.393	52	.895	.642	.752	.634	1.1	.670	1.0

Tabelle E.3
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
179.14	0.0	M	–	–	–	.443	19	.895	.642	.752	.694	0.4	.734	0.4
185.00	0.0	–	–	M	–	.288	90	.895	.642	.776	.473	1.8	.500	1.7

Anmerkungen. siehe Seite 275.

Tabelle E.4
Lernleistung, publizierte Studien

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
100.00	60.0	–	–	C	–	.268	76	.911	.820	1.003	.295	0.57	.309	0.58
101.01	0.0	–	–	C	–	.147	72	.863	.639	.467	.667	0.03	.718	0.03
101.02	0.0	–	–	C	–	.296	65	.863	.784	.489	.804	0.08	.866	0.08
102.00	54.5	C	–	–	–	.212	54	.918	.642	.705	.393	0.14	.410	0.15
103.00	0.0	C	X	M	–	.152	90	.914	.647	.752	.266	0.27	.278	0.27
104.00	0.0	M	X	–	–	.240	158	.911	.642	.752	.411	0.49	.431	0.49
105.00	88.9	M	X	M	X	.082	103	.911	.639	.752	.147	0.29	.154	0.30
106.00	0.0	M	–	M	X	.091	224	.895	.647	.791	.153	0.72	.162	0.72
107.01	86.4	M	–	M	X	.105	56	.911	.647	.752	.186	0.16	.195	0.16
107.02	94.4	M	–	M	X	.057	137	.911	.647	.752	.102	0.39	.106	0.40
107.03	88.9	M	–	M	X	.069	145	.911	.647	.752	.123	0.41	.129	0.42
108.00	90.9	M	–	M	–	.047	249	.911	.642	.752	.084	0.70	.088	0.71
109.00	0.0	C	–	C	–	.042	102	.882	.811	.752	.069	0.34	.074	0.34
110.00	0.0	M	X	–	–	.476	75	.911	.642	.752	.727	0.29	.761	0.30
111.01	0.0	C	–	–	–	.467	78	.913	.642	.791	.692	0.32	.724	0.33
111.02	0.0	C	–	–	–	.440	98	.913	.642	.791	.660	0.39	.691	0.40
112.00	0.0	–	–	–	–	.240	93	.843	.642	.791	.407	0.29	.443	0.27
114.00	0.0	–	–	M	–	.335	313	.911	.647	.752	.549	1.05	.576	1.07
116.01	0.0	C	–	–	–	.355	126	.882	.642	.776	.572	0.44	.609	0.43
116.03	0.0	C	–	–	–	.556	70	.882	.642	.752	.815	0.29	.868	0.29
116.04	0.0	C	–	–	–	.326	215	.882	.642	.776	.532	0.73	.566	0.72
116.06	0.0	C	–	–	–	.404	230	.882	.642	.791	.625	0.87	.665	0.86
116.08	0.0	C	–	–	–	.534	304	.882	.642	.752	.795	1.24	.846	1.22
116.09	0.0	C	–	–	–	.446	169	.882	.642	.752	.702	0.62	.748	0.61
116.10	0.0	C	–	–	–	.547	407	.882	.642	.791	.786	1.78	.837	1.75
116.11	0.0	C	–	–	–	.749	76	.882	.642	.791	.959	0.42	1.021	0.41
116.13	0.0	–	–	–	–	.760	180	.863	.642	.776	.967	1.00	1.040	0.97
117.00	0.0	–	–	–	–	.700	42	.863	.642	.791	.923	0.22	.993	0.21
118.01	0.0	–	–	C	–	.493	46	.935	.642	.791	.717	0.20	.741	0.21
118.02	0.0	–	–	C	–	.570	30	.935	.642	.791	.800	0.14	.827	0.14
118.03	0.0	–	–	C	–	.528	54	.935	.642	.791	.756	0.24	.781	0.25
119.01	0.0	–	–	C	–	.210	100	.863	.680	.752	.370	0.29	.398	0.28
119.02	0.0	–	–	C	–	.216	99	.863	.680	.791	.354	0.33	.381	0.32
120.01	0.0	–	–	–	–	.530	127	.863	.642	.752	.796	0.51	.856	0.49
120.02	0.0	–	–	–	–	.600	101	.863	.642	.752	.857	0.45	.923	0.43
120.03	0.0	–	–	–	–	.570	222	.863	.642	.844	.786	1.05	.846	1.02
120.04	0.0	–	–	–	–	.590	13	.863	.639	.752	.850	0.06	.915	0.05
122.00	0.0	–	–	–	–	.560	967	.935	.642	.791	.789	4.39	.816	4.59
123.01	0.0	–	–	–	–	.710	20	.895	.520	.791	.988	0.09	1.044	0.09
123.02	0.0	–	–	–	–	.650	20	.895	.520	.791	.939	0.09	.993	0.09
123.03	0.0	–	–	–	–	.800	19	.895	.520	.791	1.050	0.10	1.110	0.10
123.04	0.0	–	–	–	–	.720	7	.895	.520	.791	.995	0.03	1.052	0.03
123.05	0.0	–	–	–	–	.590	13	.895	.520	.791	.885	0.05	.936	0.05
123.06	0.0	–	–	–	–	.850	10	.895	.520	.791	1.080	0.06	1.142	0.06
124.01	0.0	M	–	–	–	.803	12	.895	.520	.791	1.052	0.06	1.112	0.06

Tabelle E.4
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
124.02	0.0	M	-	-	-	.806	8	.895	.520	.791	1.054	0.04	1.114	0.04
124.03	0.0	M	-	-	-	.515	12	.895	.520	.791	.809	0.04	.855	0.04
125.01	0.0	-	-	C	-	.293	738	.863	.642	.811	.466	2.63	.502	2.54
125.02	0.0	-	-	C	-	.262	221	.863	.642	.778	.444	0.70	.478	0.67
125.03	0.0	-	-	-	-	.160	196	.863	.642	.644	.377	0.32	.405	0.31
125.04	0.0	-	-	C	-	.198	198	.863	.642	.711	.390	0.46	.420	0.44
125.05	0.0	-	-	C	-	.233	285	.863	.642	.900	.331	1.28	.356	1.23
125.06	0.0	-	-	C	-	.205	120	.863	.642	.722	.394	0.29	.424	0.28
126.01	0.0	-	-	C	-	.337	322	.911	.642	.689	.601	0.91	.630	0.93
126.02	0.0	-	-	C	-	.283	37	.911	.642	.737	.487	0.11	.510	0.11
126.03	0.0	-	-	C	-	.386	64	.911	.642	.625	.717	0.17	.752	0.17
126.04	0.0	-	-	C	-	.378	81	.911	.642	.641	.694	0.22	.727	0.22
126.05	0.0	-	-	C	-	.153	59	.911	.642	.673	.311	0.13	.326	0.13
126.06	0.0	-	-	C	-	.259	83	.911	.642	.609	.557	0.16	.583	0.17
127.01	0.0	M	X	C	-	.211	33	.895	.642	.791	.348	0.11	.368	0.11
127.02	0.0	M	X	C	-	.337	30	.895	.642	.791	.533	0.11	.563	0.11
129.00	50.0	M	X	-	-	.079	93	.911	.642	.752	.141	0.26	.148	0.27
130.00	0.0	M	X	-	-	.305	48	.911	.639	.752	.510	0.16	.534	0.16
133.00	0.0	-	-	-	-	.330	112	.925	.642	.785	.518	0.41	.539	0.42
134.00	0.0	-	-	-	-	.720	27	.911	.642	.890	.920	0.15	.964	0.15
135.00	0.0	-	-	M	-	.470	861	.895	.647	.791	.698	3.52	.738	3.52
136.00	0.0	-	-	-	-	.360	97	.869	.642	.776	.583	0.33	.625	0.32
137.01	0.0	-	-	-	-	.200	74	.863	.642	.752	.363	0.20	.391	0.20
137.02	0.0	-	-	-	-	.400	64	.863	.642	.791	.625	0.24	.673	0.23
137.03	0.0	-	-	-	-	.480	44	.863	.642	.791	.720	0.18	.775	0.17
137.04	0.0	-	-	-	-	.190	59	.863	.642	.791	.322	0.18	.347	0.18
137.05	0.0	-	-	-	-	.100	27	.863	.642	.791	.173	0.08	.186	0.08
137.06	0.0	-	-	-	-	.280	49	.863	.642	.791	.461	0.16	.497	0.16
137.07	0.0	-	-	-	-	.120	50	.863	.642	.791	.207	0.15	.223	0.15
138.00	86.1	C	-	M	-	.225	22	.895	.520	.791	.407	0.06	.431	0.06
139.00	72.2	M	-	M	-	.101	119	.895	.642	.791	.170	0.38	.180	0.38
140.00	0.0	C	-	-	-	.469	146	.940	.642	.791	.688	0.61	.710	0.64
141.00	0.0	-	-	-	-	.750	34	.863	.642	.791	.959	0.19	1.032	0.18
142.00	0.0	-	-	-	-	.750	8	.863	.642	.752	.962	0.04	1.035	0.04
143.01	0.0	-	-	-	-	.322	123	.863	.642	.791	.522	0.42	.561	0.41
143.02	0.0	-	-	-	-	.644	147	.863	.642	.791	.878	0.71	.946	0.69
143.03	0.0	-	-	-	-	.644	62	.863	.642	.791	.878	0.30	.946	0.29
144.00	0.0	-	-	-	-	.560	315	.935	.642	.791	.789	1.43	.816	1.50
146.00	75.9	M	-	-	-	.114	45	.869	.642	.791	.196	0.14	.210	0.13
147.00	0.0	C	X	-	-	.196	308	.895	.520	.776	.367	0.79	.388	0.79
148.00	0.0	C	-	-	-	.521	91	.869	.642	.700	.817	0.33	.876	0.32
149.00	0.0	M	-	C	-	.093	199	.895	.642	.752	.168	0.55	.178	0.55
150.00	0.0	-	-	-	-	.322	167	.895	.642	.776	.522	0.57	.552	0.57
152.00	0.0	C	-	M	-	.529	180	.934	.639	.939	.688	0.96	.712	1.00
160.01	0.0	C	-	-	-	.351	58	.861	.642	.926	.475	0.29	.512	0.27
160.02	0.0	C	-	-	-	.160	91	.855	.642	.867	.242	0.36	.261	0.34
161.01	0.0	M	-	-	-	.016	60	.895	.642	.873	.024	0.25	.025	0.25
161.02	0.0	M	-	-	-	.086	31	.895	.642	.752	.156	0.09	.165	0.09
167.00	0.0	C	X	C	-	.078	139	.895	.639	.791	.132	0.44	.140	0.44
168.00	0.0	-	-	M	-	.476	51	.895	.639	.791	.708	0.21	.749	0.21
173.00	0.0	-	-	-	-	.260	190	.863	.642	.700	.505	0.46	.543	0.44
174.00	0.0	-	-	M	X	.315	469	.895	.639	.791	.503	1.66	.532	1.66
175.00	45.5	-	-	M	X	.295	187	.935	.639	.791	.464	0.68	.480	0.71
177.01	33.3	M	-	-	-	.220	27	.935	.642	.791	.352	0.09	.364	0.10
177.02	66.7	M	-	-	-	.098	27	.935	.642	.791	.160	0.09	.166	0.10
179.01	0.0	C	-	-	-	.027	12	.947	.642	.907	.038	0.06	.039	0.06
179.02	0.0	C	-	-	-	.711	15	.947	.642	.907	.906	0.08	.931	0.09
179.03	0.0	C	-	-	-	.226	9	.947	.642	.907	.312	0.04	.320	0.05
179.04	0.0	C	-	-	-	.336	10	.947	.642	.907	.458	0.05	.471	0.05

Tabelle E.4
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
179.05	0.0	C	-	-	-	.793	7	.947	.642	.907	.991	0.04	1.019	0.04
179.06	0.0	C	-	-	-	.141	15	.947	.642	.907	.196	0.07	.201	0.07
179.07	0.0	C	-	-	-	.396	14	.947	.642	.907	.536	0.07	.550	0.07
179.08	0.0	M	-	-	-	.477	51	.895	.642	.791	.708	0.21	.749	0.21
179.09	0.0	M	-	-	-	.363	64	.895	.642	.791	.568	0.24	.600	0.24
179.10	0.0	M	-	-	-	.480	23	.895	.642	.791	.712	0.09	.752	0.09
179.11	0.0	M	-	-	-	.273	26	.895	.642	.791	.442	0.09	.467	0.09
179.12	0.0	M	-	-	-	.530	20	.895	.642	.752	.788	0.08	.832	0.08
179.13	0.0	M	-	-	-	.393	52	.895	.642	.752	.634	0.18	.670	0.18
179.14	0.0	M	-	-	-	.443	19	.895	.642	.752	.694	0.07	.734	0.07
185.00	0.0	-	-	M	-	.288	90	.895	.642	.776	.473	0.30	.500	0.30
189.01	0.0	C	-	-	-	.432	102	.895	.639	.752	.683	0.37	.722	0.37
189.02	0.0	C	-	-	-	.338	403	.895	.639	.791	.535	1.45	.566	1.45
189.03	0.0	C	-	-	-	.437	282	.895	.639	.791	.663	1.11	.700	1.11
189.04	0.0	C	-	-	-	.377	87	.895	.639	.791	.587	0.32	.621	0.32
189.05	0.0	C	-	-	-	.518	68	.895	.639	.791	.755	0.29	.798	0.29
189.06	0.0	C	-	-	-	.501	731	.895	.639	.791	.736	3.05	.778	3.06
189.07	0.0	C	-	-	-	.437	198	.895	.639	.791	.663	0.78	.700	0.78
189.08	0.0	C	-	-	-	.372	112	.895	.639	.791	.581	0.41	.614	0.42
189.09	0.0	C	-	-	-	.398	351	.895	.639	.791	.614	1.33	.649	1.33
189.10	0.0	C	-	-	-	.317	394	.895	.639	.791	.506	1.40	.535	1.40
189.11	0.0	C	-	-	-	.218	81	.895	.639	.791	.359	0.27	.380	0.27
189.12	0.0	C	-	-	-	.368	191	.895	.639	.791	.575	0.71	.608	0.71
189.13	0.0	C	-	-	-	.368	97	.895	.639	.791	.575	0.36	.608	0.36
189.14	0.0	C	-	-	-	.616	617	.895	.639	.791	.852	2.91	.900	2.91
189.15	0.0	C	-	-	-	.475	418	.895	.639	.791	.707	1.70	.748	1.70
189.16	0.0	C	-	-	-	.428	412	.895	.639	.791	.652	1.60	.689	1.60
189.17	0.0	C	-	-	-	.394	243	.895	.639	.791	.609	0.92	.644	0.92
189.18	0.0	C	-	-	-	.330	98	.895	.639	.791	.524	0.35	.554	0.35
189.19	0.0	C	-	-	-	.561	287	.895	.639	.791	.799	1.28	.845	1.28
189.20	0.0	C	-	-	-	.411	291	.895	.639	.791	.631	1.12	.667	1.12
189.21	0.0	C	-	-	-	.278	184	.895	.639	.791	.450	0.63	.476	0.63
189.22	0.0	C	-	-	-	.385	84	.895	.639	.791	.598	0.31	.632	0.31
189.23	0.0	C	-	-	-	.214	158	.895	.639	.791	.353	0.52	.373	0.52
189.24	0.0	C	-	-	-	.300	119	.895	.639	.791	.482	0.42	.509	0.42
189.25	0.0	C	-	-	-	.381	67	.895	.639	.791	.593	0.25	.626	0.25
189.26	0.0	C	-	-	-	.407	57	.895	.639	.791	.626	0.22	.661	0.22
189.27	0.0	C	-	-	-	.377	81	.895	.639	.776	.598	0.29	.632	0.29
189.28	0.0	C	-	-	-	.479	181	.895	.639	.791	.712	0.74	.752	0.74
189.29	0.0	C	-	-	-	.347	314	.895	.639	.791	.547	1.14	.579	1.14
189.30	0.0	C	-	-	-	.407	454	.895	.639	.791	.626	1.73	.661	1.73
189.31	0.0	C	-	-	-	.501	380	.895	.639	.791	.736	1.59	.778	1.59
189.32	0.0	C	-	-	-	.381	412	.895	.639	.791	.593	1.54	.626	1.54
189.33	0.0	C	-	-	-	.488	391	.895	.639	.791	.722	1.61	.763	1.61
189.34	0.0	C	-	-	-	.321	48	.895	.639	.791	.512	0.17	.541	0.17
189.35	0.0	C	-	-	-	.390	58	.895	.639	.752	.631	0.20	.667	0.20
189.36	0.0	C	-	-	-	.424	57	.895	.639	.752	.673	0.20	.712	0.20
189.37	0.0	C	-	-	-	.223	87	.895	.639	.752	.390	0.26	.413	0.26
189.38	0.0	C	-	-	-	.334	56	.895	.639	.791	.530	0.20	.560	0.20
189.39	0.0	C	-	-	-	.248	48	.895	.639	.791	.405	0.16	.428	0.16
189.40	0.0	C	-	-	-	.240	61	.895	.639	.791	.393	0.21	.416	0.21
189.41	0.0	C	-	-	-	.214	104	.895	.639	.791	.353	0.34	.373	0.34
189.42	0.0	C	-	-	-	.227	113	.895	.639	.791	.373	0.38	.395	0.38
189.43	0.0	C	-	-	-	.338	38	.895	.639	.791	.535	0.14	.566	0.14
189.44	0.0	C	-	-	-	.295	21	.895	.639	.791	.475	0.07	.502	0.07
189.45	0.0	C	-	-	-	.390	178	.895	.639	.752	.631	0.61	.667	0.61
189.46	0.0	C	-	-	-	.432	347	.895	.639	.752	.683	1.25	.722	1.25
189.47	0.0	C	-	-	-	.454	37	.895	.639	.752	.708	0.14	.749	0.14
189.48	0.0	C	-	-	-	.424	41	.895	.639	.752	.673	0.15	.712	0.15

Tabelle E.4
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.49	0.0	C	-	-	-	.595	104	.895	.639	.752	.849	0.46	.897	0.46
189.50	0.0	C	-	-	-	.454	174	.895	.639	.752	.708	0.65	.749	0.65
189.51	0.0	C	-	-	-	.257	104	.895	.639	.791	.419	0.35	.443	0.35
189.52	0.0	C	-	-	-	.509	32	.895	.639	.752	.768	0.13	.811	0.13
189.53	0.0	C	-	-	-	.574	268	.895	.639	.791	.812	1.21	.858	1.21
189.54	0.0	C	-	-	-	.432	212	.895	.639	.791	.657	0.83	.694	0.83
189.55	0.0	C	-	-	-	.488	98	.895	.639	.791	.722	0.40	.763	0.40
189.56	0.0	C	-	-	-	.484	87	.895	.639	.791	.717	0.36	.758	0.36
189.57	0.0	C	-	-	-	.424	106	.895	.639	.791	.647	0.41	.684	0.41
189.58	0.0	C	-	-	-	.287	124	.895	.639	.791	.463	0.43	.489	0.43
189.59	0.0	C	-	-	-	.428	168	.895	.639	.791	.652	0.65	.689	0.65
189.60	0.0	C	-	-	-	.372	117	.895	.639	.791	.581	0.43	.614	0.43
189.61	0.0	C	-	-	-	.407	198	.895	.639	.791	.626	0.76	.661	0.76
189.62	0.0	C	-	-	-	.390	314	.895	.639	.791	.604	1.18	.639	1.18
189.63	0.0	C	-	-	-	.420	217	.895	.639	.791	.642	0.84	.679	0.84
189.64	0.0	C	-	-	-	.462	208	.895	.639	.791	.692	0.84	.732	0.84
189.65	0.0	C	-	-	-	.458	197	.895	.639	.791	.688	0.79	.727	0.79
189.66	0.0	C	-	-	-	.458	128	.895	.639	.791	.688	0.51	.727	0.51
189.67	0.0	C	-	-	-	.471	117	.895	.639	.791	.703	0.47	.743	0.47
189.68	0.0	C	-	-	-	.462	89	.895	.639	.791	.692	0.36	.732	0.36
189.69	0.0	C	-	-	-	.492	58	.895	.639	.791	.726	0.24	.768	0.24
189.70	0.0	C	-	-	-	.492	87	.895	.639	.752	.750	0.34	.793	0.34

Tabelle E.5
Lernleistung unpubliziert

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
113.00	-	-	M	X	.236	148	.863	.639	.750	.425	2.8	.457	2.7
151.00	C	-	M	-	.200	200	.935	.647	.791	.321	4.7	.332	4.9
153.00	-	-	-	-	.570	153	.870	.639	.737	.839	4.2	.900	4.2
156.00	-	-	-	-	.310	122	.911	.642	.970	.399	4.4	.418	4.5
159.00	C	X	C	-	.523	150	.877	.642	.776	.771	4.2	.824	4.1
164.01	M	-	-	-	.245	486	.895	.642	.791	.400	11.0	.423	11.1
164.02	M	-	-	-	.350	198	.895	.642	.791	.550	4.8	.582	4.9
171.01	-	-	C	-	.390	197	.875	.740	.950	.478	7.9	.511	7.8
171.02	-	-	C	-	.332	142	.875	.720	1.018	.384	6.4	.410	6.3
171.03	-	-	C	-	.277	141	.875	.711	1.084	.300	7.2	.320	7.1
171.04	-	-	C	-	-.167	21	.875	.748	.757	-.279	0.5	-.298	0.5
172.00	-	-	C	-	.196	97	.869	.665	1.034	.231	4.2	.248	4.1
176.01	C	-	-	-	.515	23	.915	.642	.791	.746	0.7	.780	0.7
176.02	C	-	-	-	.563	16	.915	.642	.791	.796	0.5	.832	0.5
178.00	C	-	M	X	.211	176	.911	.647	.752	.364	3.6	.381	3.7
181.01	-	-	-	-	.120	280	.869	.642	.776	.212	5.4	.227	5.3
181.02	-	-	-	-	.250	303	.869	.642	.776	.425	6.3	.456	6.2
181.03	-	-	-	-	.300	70	.869	.642	.776	.499	1.5	.536	1.5
181.04	-	-	-	-	.100	110	.869	.642	.776	.177	2.1	.190	2.1
182.00	-	-	-	-	.090	200	.895	.642	.776	.156	4.0	.165	4.0
183.00	-	-	-	-	.070	256	.895	.642	.752	.127	4.7	.134	4.7
184.01	-	-	-	-	.220	152	.935	.642	.721	.390	2.9	.403	3.1
184.02	-	-	-	-	.040	26	.935	.642	.721	.074	0.5	.076	0.5
184.03	-	-	-	-	.310	89	.920	.642	.741	.520	1.9	.542	2.0
186.00	-	-	-	-	.320	167	.895	.642	.776	.519	3.8	.549	3.8

Anmerkungen. siehe Seite 275.

Tabelle E.6
Lernleistung, Testmanuale

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
116.01	C	–	–	–	.355	126	.882	.642	.776	.572	0.75	.609	0.74
116.03	C	–	–	–	.556	70	.882	.642	.752	.815	0.50	.868	0.49
116.04	C	–	–	–	.326	215	.882	.642	.776	.532	1.24	.566	1.23
116.06	C	–	–	–	.404	230	.882	.642	.791	.625	1.48	.665	1.46
116.08	C	–	–	–	.534	304	.882	.642	.752	.795	2.11	.846	2.09
116.09	C	–	–	–	.446	169	.882	.642	.752	.702	1.05	.748	1.04
116.10	C	–	–	–	.547	407	.882	.642	.791	.786	3.03	.837	2.99
116.11	C	–	–	–	.749	76	.882	.642	.791	.959	0.71	1.021	0.70
116.13	–	–	–	–	.760	180	.863	.642	.776	.967	1.71	1.040	1.66
133.00	–	–	–	–	.330	112	.925	.642	.785	.518	0.70	.539	0.72
152.00	C	–	M	–	.529	180	.934	.639	.939	.688	1.64	.712	1.71
189.01	C	–	–	–	.432	102	.895	.639	.752	.683	0.63	.722	0.63
189.02	C	–	–	–	.338	403	.895	.639	.791	.535	2.48	.566	2.48
189.03	C	–	–	–	.437	282	.895	.639	.791	.663	1.89	.700	1.89
189.04	C	–	–	–	.377	87	.895	.639	.791	.587	0.55	.621	0.55
189.05	C	–	–	–	.518	68	.895	.639	.791	.755	0.49	.798	0.49
189.06	C	–	–	–	.501	731	.895	.639	.791	.736	5.21	.778	5.22
189.07	C	–	–	–	.437	198	.895	.639	.791	.663	1.33	.700	1.33
189.08	C	–	–	–	.372	112	.895	.639	.791	.581	0.71	.614	0.71
189.09	C	–	–	–	.398	351	.895	.639	.791	.614	2.27	.649	2.27
189.10	C	–	–	–	.317	394	.895	.639	.791	.506	2.38	.535	2.39
189.11	C	–	–	–	.218	81	.895	.639	.791	.359	0.46	.380	0.46
189.12	C	–	–	–	.368	191	.895	.639	.791	.575	1.20	.608	1.20
189.13	C	–	–	–	.368	97	.895	.639	.791	.575	0.61	.608	0.61
189.14	C	–	–	–	.616	617	.895	.639	.791	.852	4.97	.900	4.98
189.15	C	–	–	–	.475	418	.895	.639	.791	.707	2.90	.748	2.91
189.16	C	–	–	–	.428	412	.895	.639	.791	.652	2.74	.689	2.74
189.17	C	–	–	–	.394	243	.895	.639	.791	.609	1.57	.644	1.57
189.18	C	–	–	–	.330	98	.895	.639	.791	.524	0.60	.554	0.60
189.19	C	–	–	–	.561	287	.895	.639	.791	.799	2.18	.845	2.18
189.20	C	–	–	–	.411	291	.895	.639	.791	.631	1.90	.667	1.91
189.21	C	–	–	–	.278	184	.895	.639	.791	.450	1.08	.476	1.08
189.22	C	–	–	–	.385	84	.895	.639	.791	.598	0.54	.632	0.54
189.23	C	–	–	–	.214	158	.895	.639	.791	.353	0.89	.373	0.89
189.24	C	–	–	–	.300	119	.895	.639	.791	.482	0.71	.509	0.71
189.25	C	–	–	–	.381	67	.895	.639	.791	.593	0.43	.626	0.43
189.26	C	–	–	–	.407	57	.895	.639	.791	.626	0.37	.661	0.37
189.27	C	–	–	–	.377	81	.895	.639	.776	.598	0.50	.632	0.50
189.28	C	–	–	–	.479	181	.895	.639	.791	.712	1.26	.752	1.26
189.29	C	–	–	–	.347	314	.895	.639	.791	.547	1.94	.579	1.95
189.30	C	–	–	–	.407	454	.895	.639	.791	.626	2.96	.661	2.96
189.31	C	–	–	–	.501	380	.895	.639	.791	.736	2.71	.778	2.71
189.32	C	–	–	–	.381	412	.895	.639	.791	.593	2.62	.626	2.63
189.33	C	–	–	–	.488	391	.895	.639	.791	.722	2.75	.763	2.76
189.34	C	–	–	–	.321	48	.895	.639	.791	.512	0.29	.541	0.29
189.35	C	–	–	–	.390	58	.895	.639	.752	.631	0.34	.667	0.34
189.36	C	–	–	–	.424	57	.895	.639	.752	.673	0.35	.712	0.35
189.37	C	–	–	–	.223	87	.895	.639	.752	.390	0.44	.413	0.44
189.38	C	–	–	–	.334	56	.895	.639	.791	.530	0.34	.560	0.34
189.39	C	–	–	–	.248	48	.895	.639	.791	.405	0.28	.428	0.28
189.40	C	–	–	–	.240	61	.895	.639	.791	.393	0.35	.416	0.35
189.41	C	–	–	–	.214	104	.895	.639	.791	.353	0.59	.373	0.59
189.42	C	–	–	–	.227	113	.895	.639	.791	.373	0.64	.395	0.64
189.43	C	–	–	–	.338	38	.895	.639	.791	.535	0.23	.566	0.23
189.44	C	–	–	–	.295	21	.895	.639	.791	.475	0.12	.502	0.13
189.45	C	–	–	–	.390	178	.895	.639	.752	.631	1.05	.667	1.05
189.46	C	–	–	–	.432	347	.895	.639	.752	.683	2.14	.722	2.14
189.47	C	–	–	–	.454	37	.895	.639	.752	.708	0.23	.749	0.23
189.48	C	–	–	–	.424	41	.895	.639	.752	.673	0.25	.712	0.25

Tabelle E.6
(Fortsetzung)

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.49	C	–	–	–	.595	104	.895	.639	.752	.849	0.79	.897	0.79
189.50	C	–	–	–	.454	174	.895	.639	.752	.708	1.10	.749	1.10
189.51	C	–	–	–	.257	104	.895	.639	.791	.419	0.60	.443	0.60
189.52	C	–	–	–	.509	32	.895	.639	.752	.768	0.22	.811	0.22
189.53	C	–	–	–	.574	268	.895	.639	.791	.812	2.06	.858	2.07
189.54	C	–	–	–	.432	212	.895	.639	.791	.657	1.41	.694	1.42
189.55	C	–	–	–	.488	98	.895	.639	.791	.722	0.69	.763	0.69
189.56	C	–	–	–	.484	87	.895	.639	.791	.717	0.61	.758	0.61
189.57	C	–	–	–	.424	106	.895	.639	.791	.647	0.70	.684	0.70
189.58	C	–	–	–	.287	124	.895	.639	.791	.463	0.73	.489	0.73
189.59	C	–	–	–	.428	168	.895	.639	.791	.652	1.12	.689	1.12
189.60	C	–	–	–	.372	117	.895	.639	.791	.581	0.74	.614	0.74
189.61	C	–	–	–	.407	198	.895	.639	.791	.626	1.29	.661	1.29
189.62	C	–	–	–	.390	314	.895	.639	.791	.604	2.02	.639	2.02
189.63	C	–	–	–	.420	217	.895	.639	.791	.642	1.43	.679	1.43
189.64	C	–	–	–	.462	208	.895	.639	.791	.692	1.43	.732	1.43
189.65	C	–	–	–	.458	197	.895	.639	.791	.688	1.35	.727	1.35
189.66	C	–	–	–	.458	128	.895	.639	.791	.688	0.87	.727	0.88
189.67	C	–	–	–	.471	117	.895	.639	.791	.703	0.81	.743	0.81
189.68	C	–	–	–	.462	89	.895	.639	.791	.692	0.61	.732	0.61
189.69	C	–	–	–	.492	58	.895	.639	.791	.726	0.41	.768	0.41
189.70	C	–	–	–	.492	87	.895	.639	.752	.750	0.58	.793	0.58

Anmerkungen. siehe Seite 275.

Tabelle E.7
Lernleistung, Artikel in Zeitschriften oder Herausgeberwerken

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
100.00	60.0	–	–	C	–	.274	76	.911	.820	1.003	.302	1.9	.316	1.9
102.00	54.5	C	–	–	–	.219	54	.918	.642	.705	.405	0.5	.423	0.5
103.00	0.0	C	X	M	–	.152	90	.914	.647	.752	.266	0.9	.278	0.9
104.00	0.0	M	X	–	–	.240	158	.911	.642	.752	.411	1.6	.431	1.6
105.00	88.9	M	X	M	X	.086	103	.911	.639	.752	.154	1.0	.161	1.0
106.00	0.0	M	–	M	X	.091	224	.895	.647	.791	.153	2.3	.162	2.3
107.01	86.4	M	–	M	X	.111	56	.911	.647	.752	.196	0.5	.206	0.5
107.02	94.4	M	–	M	X	.062	137	.911	.647	.752	.110	1.3	.116	1.3
107.03	88.9	M	–	M	X	.073	145	.911	.647	.752	.130	1.4	.136	1.4
108.00	90.9	M	–	M	–	.050	249	.911	.642	.752	.090	2.3	.094	2.3
110.00	0.0	M	X	–	–	.476	75	.911	.642	.752	.727	1.0	.761	1.0
112.00	0.0	–	–	–	–	.240	93	.843	.642	.791	.407	1.0	.443	0.9
114.00	0.0	–	–	M	–	.335	313	.911	.647	.752	.549	3.4	.576	3.5
117.00	0.0	–	–	–	–	.700	42	.863	.642	.791	.923	0.7	.993	0.7
118.01	0.0	–	–	C	–	.493	46	.935	.642	.791	.717	0.6	.741	0.7
118.02	0.0	–	–	C	–	.570	30	.935	.642	.791	.800	0.5	.827	0.5
118.03	0.0	–	–	C	–	.528	54	.935	.642	.791	.756	0.8	.781	0.8
119.01	0.0	–	–	C	–	.210	100	.863	.680	.752	.370	1.0	.398	0.9
119.02	0.0	–	–	C	–	.216	99	.863	.680	.791	.354	1.1	.381	1.1
120.01	0.0	–	–	–	–	.530	127	.863	.642	.752	.796	1.7	.856	1.6
120.02	0.0	–	–	–	–	.600	101	.863	.642	.752	.857	1.5	.923	1.4
120.03	0.0	–	–	–	–	.570	222	.863	.642	.844	.786	3.5	.846	3.3
120.04	0.0	–	–	–	–	.590	13	.863	.639	.752	.850	0.2	.915	0.2
122.00	0.0	–	–	–	–	.560	967	.935	.642	.791	.789	14.4	.816	15.0
123.01	0.0	–	–	–	–	.710	20	.895	.520	.791	.988	0.3	1.044	0.3

Tabelle E.7
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
123.02	0.0	-	-	-	-	.650	20	.895	.520	.791	.939	0.3	.993	0.3
123.03	0.0	-	-	-	-	.800	19	.895	.520	.791	1.050	0.3	1.110	0.3
123.04	0.0	-	-	-	-	.720	7	.895	.520	.791	.995	0.1	1.052	0.1
123.05	0.0	-	-	-	-	.590	13	.895	.520	.791	.885	0.2	.936	0.2
123.06	0.0	-	-	-	-	.850	10	.895	.520	.791	1.080	0.2	1.142	0.2
124.01	0.0	M	-	-	-	.803	12	.895	.520	.791	1.052	0.2	1.112	0.2
124.02	0.0	M	-	-	-	.806	8	.895	.520	.791	1.054	0.1	1.114	0.1
124.03	0.0	M	-	-	-	.515	12	.895	.520	.791	.809	0.1	.855	0.1
125.01	0.0	-	-	C	-	.293	738	.863	.642	.811	.466	8.6	.502	8.3
125.02	0.0	-	-	C	-	.262	221	.863	.642	.778	.444	2.3	.478	2.2
125.03	0.0	-	-	-	-	.160	196	.863	.642	.644	.377	1.0	.405	1.0
125.04	0.0	-	-	C	-	.198	198	.863	.642	.711	.390	1.5	.420	1.5
125.05	0.0	-	-	C	-	.233	285	.863	.642	.900	.331	4.2	.356	4.0
125.06	0.0	-	-	C	-	.205	120	.863	.642	.722	.394	1.0	.424	0.9
126.01	0.0	-	-	C	-	.337	322	.911	.642	.689	.601	3.0	.630	3.0
126.02	0.0	-	-	C	-	.283	37	.911	.642	.737	.487	0.4	.510	0.4
126.03	0.0	-	-	C	-	.386	64	.911	.642	.625	.717	0.6	.752	0.6
126.04	0.0	-	-	C	-	.378	81	.911	.642	.641	.694	0.7	.727	0.7
126.05	0.0	-	-	C	-	.153	59	.911	.642	.673	.311	0.4	.326	0.4
126.06	0.0	-	-	C	-	.259	83	.911	.642	.609	.557	0.5	.583	0.5
127.01	0.0	M	X	C	-	.211	33	.895	.642	.791	.348	0.4	.368	0.4
127.02	0.0	M	X	C	-	.337	30	.895	.642	.791	.533	0.4	.563	0.4
129.00	50.0	M	X	-	-	.083	93	.911	.642	.752	.148	0.9	.155	0.9
130.00	0.0	M	X	-	-	.305	48	.911	.639	.752	.510	0.5	.534	0.5
134.00	0.0	-	-	-	-	.720	27	.911	.642	.890	.920	0.5	.964	0.5
138.00	86.1	C	-	M	-	.238	22	.895	.520	.791	.429	0.2	.453	0.2
139.00	72.2	M	-	M	-	.104	119	.895	.642	.791	.175	1.2	.185	1.2
140.00	0.0	C	-	-	-	.469	146	.940	.642	.791	.688	2.0	.710	2.1
141.00	0.0	-	-	-	-	.750	34	.863	.642	.791	.959	0.6	1.032	0.6
142.00	0.0	-	-	-	-	.750	8	.863	.642	.752	.962	0.1	1.035	0.1
143.01	0.0	-	-	-	-	.322	123	.863	.642	.791	.522	1.4	.561	1.3
143.02	0.0	-	-	-	-	.644	147	.863	.642	.791	.878	2.3	.946	2.3
143.03	0.0	-	-	-	-	.644	62	.863	.642	.791	.878	1.0	.946	1.0
144.00	0.0	-	-	-	-	.560	315	.935	.642	.791	.789	4.7	.816	4.9
146.00	75.9	M	-	-	-	.120	45	.869	.642	.791	.206	0.5	.221	0.4
148.00	0.0	C	-	-	-	.521	91	.869	.642	.700	.817	1.1	.876	1.1
160.01	0.0	C	-	-	-	.351	58	.861	.642	.926	.475	0.9	.512	0.9
160.02	0.0	C	-	-	-	.160	91	.855	.642	.867	.242	1.2	.261	1.1
161.01	0.0	M	-	-	-	.016	60	.895	.642	.873	.024	0.8	.025	0.8
161.02	0.0	M	-	-	-	.086	31	.895	.642	.752	.156	0.3	.165	0.3
168.00	0.0	-	-	M	-	.476	51	.895	.639	.791	.708	0.7	.749	0.7
173.00	0.0	-	-	-	-	.260	190	.863	.642	.700	.505	1.5	.543	1.4
174.00	0.0	-	-	M	X	.315	469	.895	.639	.791	.503	5.4	.532	5.4
185.00	0.0	-	-	M	-	.288	90	.895	.642	.776	.473	1.0	.500	1.0

Anmerkungen. siehe Seite 275.

Tabelle E.8
Lernleistung, Monographien (ohne Testmanuale) und Dissertationen

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
101.01	0.0	-	-	C	-	.147	72	.863	.639	.467	.667	0.3	.718	0.3
101.02	0.0	-	-	C	-	.296	65	.863	.784	.489	.804	0.7	.866	0.7
109.00	0.0	C	-	C	-	.042	102	.882	.811	.752	.069	3.2	.074	3.1

Tabelle E.8
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
111.01	0.0	C	-	-	-	.467	78	.913	.642	.791	.692	3.0	.724	3.0
111.02	0.0	C	-	-	-	.440	98	.913	.642	.791	.660	3.6	.691	3.7
135.00	0.0	-	-	M	-	.470	861	.895	.647	.791	.698	32.6	.738	32.5
136.00	0.0	-	-	-	-	.360	97	.869	.642	.776	.583	3.1	.625	3.0
137.01	0.0	-	-	-	-	.200	74	.863	.642	.752	.363	1.9	.391	1.8
137.02	0.0	-	-	-	-	.400	64	.863	.642	.791	.625	2.2	.673	2.1
137.03	0.0	-	-	-	-	.480	44	.863	.642	.791	.720	1.6	.775	1.6
137.04	0.0	-	-	-	-	.190	59	.863	.642	.791	.322	1.7	.347	1.7
137.05	0.0	-	-	-	-	.100	27	.863	.642	.791	.173	0.8	.186	0.7
137.06	0.0	-	-	-	-	.280	49	.863	.642	.791	.461	1.5	.497	1.5
137.07	0.0	-	-	-	-	.120	50	.863	.642	.791	.207	1.4	.223	1.4
147.00	0.0	C	X	-	-	.196	308	.895	.520	.776	.367	7.3	.388	7.3
149.00	0.0	M	-	C	-	.093	199	.895	.642	.752	.168	5.1	.178	5.1
150.00	0.0	-	-	-	-	.322	167	.895	.642	.776	.522	5.3	.552	5.3
167.00	0.0	C	X	C	-	.078	139	.895	.639	.791	.132	4.0	.140	4.0
175.00	45.5	-	-	M	X	.292	187	.935	.639	.791	.460	6.3	.475	6.6
177.01	33.3	M	-	-	-	.218	27	.935	.642	.791	.349	0.9	.361	0.9
177.02	66.7	M	-	-	-	.093	27	.935	.642	.791	.152	0.8	.158	0.9
179.01	0.0	C	-	-	-	.027	12	.947	.642	.907	.038	0.5	.039	0.6
179.02	0.0	C	-	-	-	.711	15	.947	.642	.907	.906	0.8	.931	0.8
179.03	0.0	C	-	-	-	.226	9	.947	.642	.907	.312	0.4	.320	0.4
179.04	0.0	C	-	-	-	.336	10	.947	.642	.907	.458	0.5	.471	0.5
179.05	0.0	C	-	-	-	.793	7	.947	.642	.907	.991	0.4	1.019	0.4
179.06	0.0	C	-	-	-	.141	15	.947	.642	.907	.196	0.7	.201	0.7
179.07	0.0	C	-	-	-	.396	14	.947	.642	.907	.536	0.6	.550	0.7
179.08	0.0	M	-	-	-	.477	51	.895	.642	.791	.708	1.9	.749	1.9
179.09	0.0	M	-	-	-	.363	64	.895	.642	.791	.568	2.2	.600	2.2
179.10	0.0	M	-	-	-	.480	23	.895	.642	.791	.712	0.9	.752	0.9
179.11	0.0	M	-	-	-	.273	26	.895	.642	.791	.442	0.8	.467	0.8
179.12	0.0	M	-	-	-	.530	20	.895	.642	.752	.788	0.8	.832	0.8
179.13	0.0	M	-	-	-	.393	52	.895	.642	.752	.634	1.7	.670	1.7
179.14	0.0	M	-	-	-	.443	19	.895	.642	.752	.694	0.7	.734	0.6

Anmerkungen. siehe Seite 275.

Tabelle E.9
Lernleistung ohne BET-Manual

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
100.00	60.0	-	-	C	-	.213	76	.911	.820	1.003	.234	0.87	.246	0.89
101.01	0.0	-	-	C	-	.147	72	.863	.639	.467	.667	0.05	.718	0.05
101.02	0.0	-	-	C	-	.296	65	.863	.784	.489	.804	0.12	.866	0.12
102.00	54.5	C	-	-	-	.213	54	.918	.642	.705	.395	0.22	.412	0.22
103.00	0.0	C	X	M	-	.152	90	.914	.647	.752	.266	0.41	.278	0.42
104.00	0.0	M	X	-	-	.240	158	.911	.642	.752	.411	0.75	.431	0.76
105.00	88.9	M	X	M	X	.086	103	.911	.639	.752	.154	0.45	.161	0.46
106.00	0.0	M	-	M	X	.091	224	.895	.647	.791	.153	1.10	.162	1.10
107.01	86.4	M	-	M	X	.106	56	.911	.647	.752	.188	0.25	.197	0.25
107.02	94.4	M	-	M	X	.058	137	.911	.647	.752	.103	0.60	.108	0.61
107.03	88.9	M	-	M	X	.070	145	.911	.647	.752	.125	0.63	.131	0.65
108.00	90.9	M	-	M	-	.048	249	.911	.642	.752	.086	1.08	.090	1.10
109.00	0.0	C	-	C	-	.042	102	.882	.811	.752	.069	0.52	.074	0.52
110.00	0.0	M	X	-	-	.476	75	.911	.642	.752	.727	0.45	.761	0.46
111.01	0.0	C	-	-	-	.467	78	.913	.642	.791	.692	0.49	.724	0.50

Tabelle E.9
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
111.02	0.0	C	-	-	-	.440	98	.913	.642	.791	.660	0.60	.691	0.62
112.00	0.0	-	-	-	-	.240	93	.843	.642	.791	.407	0.45	.443	0.42
113.00	0.0	-	-	M	X	.236	148	.863	.639	.750	.425	0.63	.457	0.61
114.00	0.0	-	-	M	-	.335	313	.911	.647	.752	.549	1.61	.576	1.64
116.01	0.0	C	-	-	-	.355	126	.882	.642	.776	.572	0.67	.609	0.67
116.03	0.0	C	-	-	-	.556	70	.882	.642	.752	.815	0.45	.868	0.45
116.04	0.0	C	-	-	-	.326	215	.882	.642	.776	.532	1.12	.566	1.11
116.06	0.0	C	-	-	-	.404	230	.882	.642	.791	.625	1.33	.665	1.32
116.08	0.0	C	-	-	-	.534	304	.882	.642	.752	.795	1.90	.846	1.88
116.09	0.0	C	-	-	-	.446	169	.882	.642	.752	.702	0.94	.748	0.93
116.10	0.0	C	-	-	-	.547	407	.882	.642	.791	.786	2.73	.837	2.70
116.11	0.0	C	-	-	-	.749	76	.882	.642	.791	.959	0.64	1.021	0.63
116.13	0.0	-	-	-	-	.760	180	.863	.642	.776	.967	1.54	1.040	1.49
117.00	0.0	-	-	-	-	.700	42	.863	.642	.791	.923	0.33	.993	0.32
118.01	0.0	-	-	C	-	.493	46	.935	.642	.791	.717	0.30	.741	0.32
118.02	0.0	-	-	C	-	.570	30	.935	.642	.791	.800	0.21	.827	0.22
118.03	0.0	-	-	C	-	.528	54	.935	.642	.791	.756	0.37	.781	0.38
119.01	0.0	-	-	C	-	.210	100	.863	.680	.752	.370	0.45	.398	0.43
119.02	0.0	-	-	C	-	.216	99	.863	.680	.791	.354	0.51	.381	0.49
120.01	0.0	-	-	-	-	.530	127	.863	.642	.752	.796	0.78	.856	0.75
120.02	0.0	-	-	-	-	.600	101	.863	.642	.752	.857	0.69	.923	0.66
120.03	0.0	-	-	-	-	.570	222	.863	.642	.844	.786	1.62	.846	1.56
120.04	0.0	-	-	-	-	.590	13	.863	.639	.752	.850	0.09	.915	0.08
122.00	0.0	-	-	-	-	.560	967	.935	.642	.791	.789	6.74	.816	7.06
123.01	0.0	-	-	-	-	.710	20	.895	.520	.791	.988	0.14	1.044	0.14
123.02	0.0	-	-	-	-	.650	20	.895	.520	.791	.939	0.13	.993	0.13
123.03	0.0	-	-	-	-	.800	19	.895	.520	.791	1.050	0.15	1.110	0.15
123.04	0.0	-	-	-	-	.720	7	.895	.520	.791	.995	0.05	1.052	0.05
123.05	0.0	-	-	-	-	.590	13	.895	.520	.791	.885	0.08	.936	0.08
123.06	0.0	-	-	-	-	.850	10	.895	.520	.791	1.080	0.09	1.142	0.09
124.01	0.0	M	-	-	-	.803	12	.895	.520	.791	1.052	0.10	1.112	0.10
124.02	0.0	M	-	-	-	.806	8	.895	.520	.791	1.054	0.06	1.114	0.06
124.03	0.0	M	-	-	-	.515	12	.895	.520	.791	.809	0.07	.855	0.07
125.01	0.0	-	-	C	-	.293	738	.863	.642	.811	.466	4.04	.502	3.90
125.02	0.0	-	-	C	-	.262	221	.863	.642	.778	.444	1.07	.478	1.03
125.03	0.0	-	-	-	-	.160	196	.863	.642	.644	.377	0.49	.405	0.47
125.04	0.0	-	-	C	-	.198	198	.863	.642	.711	.390	0.71	.420	0.68
125.05	0.0	-	-	C	-	.233	285	.863	.642	.900	.331	1.96	.356	1.89
125.06	0.0	-	-	C	-	.205	120	.863	.642	.722	.394	0.45	.424	0.44
126.01	0.0	-	-	C	-	.337	322	.911	.642	.689	.601	1.40	.630	1.43
126.02	0.0	-	-	C	-	.283	37	.911	.642	.737	.487	0.17	.510	0.18
126.03	0.0	-	-	C	-	.386	64	.911	.642	.625	.717	0.26	.752	0.26
126.04	0.0	-	-	C	-	.378	81	.911	.642	.641	.694	0.33	.727	0.34
126.05	0.0	-	-	C	-	.153	59	.911	.642	.673	.311	0.20	.326	0.20
126.06	0.0	-	-	C	-	.259	83	.911	.642	.609	.557	0.25	.583	0.25
127.01	0.0	M	X	C	-	.211	33	.895	.642	.791	.348	0.17	.368	0.17
127.02	0.0	M	X	C	-	.337	30	.895	.642	.791	.533	0.17	.563	0.17
129.00	50.0	M	X	-	-	.080	93	.911	.642	.752	.143	0.40	.150	0.41
130.00	0.0	M	X	-	-	.305	48	.911	.639	.752	.510	0.24	.534	0.24
133.00	0.0	-	-	-	-	.330	112	.925	.642	.785	.518	0.63	.539	0.65
134.00	0.0	-	-	-	-	.720	27	.911	.642	.890	.920	0.23	.964	0.23
135.00	0.0	-	-	M	-	.470	861	.895	.647	.791	.698	5.40	.738	5.42
136.00	0.0	-	-	-	-	.360	97	.869	.642	.776	.583	0.51	.625	0.50
137.01	0.0	-	-	-	-	.200	74	.863	.642	.752	.363	0.31	.391	0.30
137.02	0.0	-	-	-	-	.400	64	.863	.642	.791	.625	0.36	.673	0.35
137.03	0.0	-	-	-	-	.480	44	.863	.642	.791	.720	0.27	.775	0.26
137.04	0.0	-	-	-	-	.190	59	.863	.642	.791	.322	0.28	.347	0.27
137.05	0.0	-	-	-	-	.100	27	.863	.642	.791	.173	0.12	.186	0.12
137.06	0.0	-	-	-	-	.280	49	.863	.642	.791	.461	0.25	.497	0.24

Tabelle E.9
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
137.07	0.0	-	-	-	-	.120	50	.863	.642	.791	.207	0.23	.223	0.23
138.00	86.1	C	-	M	-	.228	22	.895	.520	.791	.412	0.09	.436	0.09
139.00	72.2	M	-	M	-	.102	119	.895	.642	.791	.172	0.58	.182	0.58
140.00	0.0	C	-	-	-	.469	146	.940	.642	.791	.688	0.94	.710	0.99
141.00	0.0	-	-	-	-	.750	34	.863	.642	.791	.959	0.29	1.032	0.28
142.00	0.0	-	-	-	-	.750	8	.863	.642	.752	.962	0.07	1.035	0.07
143.01	0.0	-	-	-	-	.322	123	.863	.642	.791	.522	0.65	.561	0.63
143.02	0.0	-	-	-	-	.644	147	.863	.642	.791	.878	1.09	.946	1.06
143.03	0.0	-	-	-	-	.644	62	.863	.642	.791	.878	0.46	.946	0.45
144.00	0.0	-	-	-	-	.560	315	.935	.642	.791	.789	2.20	.816	2.30
146.00	75.9	M	-	-	-	.116	45	.869	.642	.791	.199	0.21	.214	0.21
147.00	0.0	C	X	-	-	.196	308	.895	.520	.776	.367	1.22	.388	1.22
148.00	0.0	C	-	-	-	.521	91	.869	.642	.700	.817	0.51	.876	0.50
149.00	0.0	M	-	C	-	.093	199	.895	.642	.752	.168	0.84	.178	0.84
150.00	0.0	-	-	-	-	.322	167	.895	.642	.776	.522	0.88	.552	0.88
151.00	0.0	C	-	M	-	.200	200	.935	.647	.791	.321	1.08	.332	1.13
152.00	0.0	C	-	M	-	.529	180	.934	.639	.939	.688	1.47	.712	1.54
153.00	0.0	-	-	-	-	.570	153	.870	.639	.737	.839	0.98	.900	0.95
156.00	0.0	-	-	-	-	.310	122	.911	.642	.970	.399	1.02	.418	1.04
159.00	0.0	C	X	C	-	.523	150	.877	.642	.776	.771	0.96	.824	0.94
160.01	0.0	C	-	-	-	.351	58	.861	.642	.926	.475	0.44	.512	0.42
160.02	0.0	C	-	-	-	.160	91	.855	.642	.867	.242	0.55	.261	0.53
161.01	0.0	M	-	-	-	.016	60	.895	.642	.873	.024	0.38	.025	0.38
161.02	0.0	M	-	-	-	.086	31	.895	.642	.752	.156	0.13	.165	0.13
164.01	0.0	-	-	-	-	.245	486	.895	.642	.791	.400	2.53	.423	2.53
164.02	0.0	-	-	-	-	.350	198	.895	.642	.791	.550	1.11	.582	1.11
167.00	0.0	C	X	C	-	.078	139	.895	.639	.791	.132	0.67	.140	0.67
168.00	0.0	-	-	M	-	.476	51	.895	.639	.791	.708	0.32	.749	0.32
171.01	0.0	-	-	C	-	.390	197	.875	.740	.950	.478	1.82	.511	1.78
171.02	0.0	-	-	C	-	.332	142	.875	.720	1.018	.384	1.47	.410	1.44
171.03	0.0	-	-	C	-	.277	141	.875	.711	1.084	.300	1.67	.320	1.63
171.04	0.0	-	-	C	-	-.167	21	.875	.748	.757	-.279	0.10	-.298	0.10
172.00	0.0	-	-	C	-	.196	97	.869	.665	1.034	.231	0.97	.248	0.94
173.00	0.0	-	-	-	-	.260	190	.863	.642	.700	.505	0.70	.543	0.68
174.00	0.0	-	-	M	X	.315	469	.895	.639	.791	.503	2.55	.532	2.55
175.00	45.5	-	-	M	X	.295	187	.935	.639	.791	.464	1.05	.480	1.10
176.01	0.0	C	-	-	-	.515	23	.915	.642	.791	.746	0.15	.780	0.16
176.02	0.0	C	-	-	-	.563	16	.915	.642	.791	.796	0.11	.832	0.11
177.01	33.3	M	-	-	-	.220	27	.935	.642	.791	.352	0.15	.364	0.15
177.02	66.7	M	-	-	-	.098	27	.935	.642	.791	.160	0.14	.166	0.15
178.00	0.0	C	-	M	X	.211	176	.911	.647	.752	.364	0.82	.381	0.84
179.01	0.0	C	-	-	-	.027	12	.947	.642	.907	.038	0.09	.039	0.09
179.02	0.0	C	-	-	-	.711	15	.947	.642	.907	.906	0.13	.931	0.14
179.03	0.0	C	-	-	-	.226	9	.947	.642	.907	.312	0.07	.320	0.07
179.04	0.0	C	-	-	-	.336	10	.947	.642	.907	.458	0.07	.471	0.08
179.05	0.0	C	-	-	-	.793	7	.947	.642	.907	.991	0.06	1.019	0.07
179.06	0.0	C	-	-	-	.141	15	.947	.642	.907	.196	0.11	.201	0.11
179.07	0.0	C	-	-	-	.396	14	.947	.642	.907	.536	0.11	.550	0.11
179.08	0.0	M	-	-	-	.477	51	.895	.642	.791	.708	0.32	.749	0.32
179.09	0.0	M	-	-	-	.363	64	.895	.642	.791	.568	0.36	.600	0.36
179.10	0.0	M	-	-	-	.480	23	.895	.642	.791	.712	0.14	.752	0.15
179.11	0.0	M	-	-	-	.273	26	.895	.642	.791	.442	0.14	.467	0.14
179.12	0.0	M	-	-	-	.530	20	.895	.642	.752	.788	0.13	.832	0.13
179.13	0.0	M	-	-	-	.393	52	.895	.642	.752	.634	0.28	.670	0.28
179.14	0.0	M	-	-	-	.443	19	.895	.642	.752	.694	0.11	.734	0.11
181.01	0.0	-	-	-	-	.120	280	.869	.642	.776	.212	1.24	.227	1.21
181.02	0.0	-	-	-	-	.250	303	.869	.642	.776	.425	1.45	.456	1.42
181.03	0.0	-	-	-	-	.300	70	.869	.642	.776	.499	0.35	.536	0.34
181.04	0.0	-	-	-	-	.100	110	.869	.642	.776	.177	0.49	.190	0.47

Tabelle E.9
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
182.00	0.0	–	–	–	–	.090	200	.895	.642	.776	.156	0.92	.165	0.92
183.00	0.0	–	–	–	–	.070	256	.895	.642	.752	.127	1.08	.134	1.08
184.01	0.0	–	–	–	–	.220	152	.935	.642	.721	.390	0.67	.403	0.70
184.02	0.0	–	–	–	–	.040	26	.935	.642	.721	.074	0.11	.076	0.11
184.03	0.0	–	–	–	–	.310	89	.920	.642	.741	.520	0.44	.542	0.45
185.00	0.0	–	–	M	–	.288	90	.895	.642	.776	.473	0.46	.500	0.46
186.00	0.0	–	–	–	–	.320	167	.895	.642	.776	.519	0.88	.549	0.88

Anmerkungen. siehe Seite 275.

Tabelle E.10
Lernleistung, nur BET-Manual

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.01	C	–	–	–	.432	102	.895	.639	.752	.683	0.74	.722	0.74
189.02	C	–	–	–	.338	403	.895	.639	.791	.535	2.91	.566	2.91
189.03	C	–	–	–	.437	282	.895	.639	.791	.663	2.22	.700	2.22
189.04	C	–	–	–	.377	87	.895	.639	.791	.587	0.65	.621	0.65
189.05	C	–	–	–	.518	68	.895	.639	.791	.755	0.58	.798	0.58
189.06	C	–	–	–	.501	731	.895	.639	.791	.736	6.13	.778	6.13
189.07	C	–	–	–	.437	198	.895	.639	.791	.663	1.56	.700	1.56
189.08	C	–	–	–	.372	112	.895	.639	.791	.581	0.83	.614	0.83
189.09	C	–	–	–	.398	351	.895	.639	.791	.614	2.67	.649	2.67
189.10	C	–	–	–	.317	394	.895	.639	.791	.506	2.80	.535	2.80
189.11	C	–	–	–	.218	81	.895	.639	.791	.359	0.54	.380	0.54
189.12	C	–	–	–	.368	191	.895	.639	.791	.575	1.41	.608	1.41
189.13	C	–	–	–	.368	97	.895	.639	.791	.575	0.72	.608	0.72
189.14	C	–	–	–	.616	617	.895	.639	.791	.852	5.84	.900	5.84
189.15	C	–	–	–	.475	418	.895	.639	.791	.707	3.41	.748	3.41
189.16	C	–	–	–	.428	412	.895	.639	.791	.652	3.22	.689	3.22
189.17	C	–	–	–	.394	243	.895	.639	.791	.609	1.84	.644	1.84
189.18	C	–	–	–	.330	98	.895	.639	.791	.524	0.70	.554	0.70
189.19	C	–	–	–	.561	287	.895	.639	.791	.799	2.56	.845	2.56
189.20	C	–	–	–	.411	291	.895	.639	.791	.631	2.24	.667	2.24
189.21	C	–	–	–	.278	184	.895	.639	.791	.450	1.27	.476	1.27
189.22	C	–	–	–	.385	84	.895	.639	.791	.598	0.63	.632	0.63
189.23	C	–	–	–	.214	158	.895	.639	.791	.353	1.05	.373	1.05
189.24	C	–	–	–	.300	119	.895	.639	.791	.482	0.84	.509	0.84
189.25	C	–	–	–	.381	67	.895	.639	.791	.593	0.50	.626	0.50
189.26	C	–	–	–	.407	57	.895	.639	.791	.626	0.44	.661	0.44
189.27	C	–	–	–	.377	81	.895	.639	.776	.598	0.58	.632	0.58
189.28	C	–	–	–	.479	181	.895	.639	.791	.712	1.48	.752	1.48
189.29	C	–	–	–	.347	314	.895	.639	.791	.547	2.29	.579	2.29
189.30	C	–	–	–	.407	454	.895	.639	.791	.626	3.48	.661	3.48
189.31	C	–	–	–	.501	380	.895	.639	.791	.736	3.19	.778	3.19
189.32	C	–	–	–	.381	412	.895	.639	.791	.593	3.09	.626	3.09
189.33	C	–	–	–	.488	391	.895	.639	.791	.722	3.24	.763	3.24
189.34	C	–	–	–	.321	48	.895	.639	.791	.512	0.34	.541	0.34
189.35	C	–	–	–	.390	58	.895	.639	.752	.631	0.40	.667	0.40
189.36	C	–	–	–	.424	57	.895	.639	.752	.673	0.41	.712	0.41
189.37	C	–	–	–	.223	87	.895	.639	.752	.390	0.51	.413	0.51
189.38	C	–	–	–	.334	56	.895	.639	.791	.530	0.40	.560	0.40
189.39	C	–	–	–	.248	48	.895	.639	.791	.405	0.33	.428	0.33
189.40	C	–	–	–	.240	61	.895	.639	.791	.393	0.41	.416	0.41

Tabelle E.10
(Fortsetzung)

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.41	C	–	–	–	.214	104	.895	.639	.791	.353	0.69	.373	0.69
189.42	C	–	–	–	.227	113	.895	.639	.791	.373	0.76	.395	0.76
189.43	C	–	–	–	.338	38	.895	.639	.791	.535	0.27	.566	0.27
189.44	C	–	–	–	.295	21	.895	.639	.791	.475	0.15	.502	0.15
189.45	C	–	–	–	.390	178	.895	.639	.752	.631	1.23	.667	1.23
189.46	C	–	–	–	.432	347	.895	.639	.752	.683	2.52	.722	2.52
189.47	C	–	–	–	.454	37	.895	.639	.752	.708	0.28	.749	0.28
189.48	C	–	–	–	.424	41	.895	.639	.752	.673	0.29	.712	0.29
189.49	C	–	–	–	.595	104	.895	.639	.752	.849	0.93	.897	0.93
189.50	C	–	–	–	.454	174	.895	.639	.752	.708	1.29	.749	1.29
189.51	C	–	–	–	.257	104	.895	.639	.791	.419	0.71	.443	0.71
189.52	C	–	–	–	.509	32	.895	.639	.752	.768	0.25	.811	0.25
189.53	C	–	–	–	.574	268	.895	.639	.791	.812	2.42	.858	2.42
189.54	C	–	–	–	.432	212	.895	.639	.791	.657	1.66	.694	1.66
189.55	C	–	–	–	.488	98	.895	.639	.791	.722	0.81	.763	0.81
189.56	C	–	–	–	.484	87	.895	.639	.791	.717	0.72	.758	0.72
189.57	C	–	–	–	.424	106	.895	.639	.791	.647	0.82	.684	0.82
189.58	C	–	–	–	.287	124	.895	.639	.791	.463	0.86	.489	0.86
189.59	C	–	–	–	.428	168	.895	.639	.791	.652	1.31	.689	1.31
189.60	C	–	–	–	.372	117	.895	.639	.791	.581	0.87	.614	0.87
189.61	C	–	–	–	.407	198	.895	.639	.791	.626	1.52	.661	1.52
189.62	C	–	–	–	.390	314	.895	.639	.791	.604	2.37	.639	2.37
189.63	C	–	–	–	.420	217	.895	.639	.791	.642	1.68	.679	1.68
189.64	C	–	–	–	.462	208	.895	.639	.791	.692	1.68	.732	1.68
189.65	C	–	–	–	.458	197	.895	.639	.791	.688	1.58	.727	1.58
189.66	C	–	–	–	.458	128	.895	.639	.791	.688	1.03	.727	1.03
189.67	C	–	–	–	.471	117	.895	.639	.791	.703	0.95	.743	0.95
189.68	C	–	–	–	.462	89	.895	.639	.791	.692	0.72	.732	0.72
189.69	C	–	–	–	.492	58	.895	.639	.791	.726	0.48	.768	0.48
189.70	C	–	–	–	.492	87	.895	.639	.752	.750	0.68	.793	0.68

Anmerkungen. siehe Seite 275.

Tabelle E.11
Lernleistung vor 1945

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
123.01	–	–	–	–	.710	20	.895	.520	.791	.988	16.4	1.044	16.4
123.02	–	–	–	–	.650	20	.895	.520	.791	.939	15.2	.993	15.2
123.03	–	–	–	–	.800	19	.895	.520	.791	1.050	17.5	1.110	17.5
123.04	–	–	–	–	.720	7	.895	.520	.791	.995	5.8	1.052	5.8
123.05	–	–	–	–	.590	13	.895	.520	.791	.885	9.2	.936	9.2
123.06	–	–	–	–	.850	10	.895	.520	.791	1.080	9.8	1.142	9.8
124.01	M	–	–	–	.803	12	.895	.520	.791	1.052	11.1	1.112	11.1
124.02	M	–	–	–	.806	8	.895	.520	.791	1.054	7.4	1.114	7.4
124.03	M	–	–	–	.515	12	.895	.520	.791	.809	7.7	.855	7.7

Anmerkungen. siehe Seite 275.

Tabelle E.12
Lernleistung nach 1945

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
100.00	60.0	-	-	C	-	.273	76	.911	.820	1.003	.300	0.49	.315	0.50
101.01	0.0	-	-	C	-	.147	72	.863	.639	.467	.667	0.03	.718	0.03
101.02	0.0	-	-	C	-	.296	65	.863	.784	.489	.804	0.07	.866	0.07
102.00	54.5	C	-	-	-	.217	54	.918	.642	.705	.401	0.12	.419	0.13
103.00	0.0	C	X	M	-	.152	90	.914	.647	.752	.266	0.23	.278	0.24
104.00	0.0	M	X	-	-	.240	158	.911	.642	.752	.411	0.42	.431	0.43
105.00	88.9	M	X	M	X	.085	103	.911	.639	.752	.152	0.25	.159	0.26
106.00	0.0	M	-	M	X	.091	224	.895	.647	.791	.153	0.63	.162	0.63
107.01	86.4	M	-	M	X	.110	56	.911	.647	.752	.195	0.14	.204	0.14
107.02	94.4	M	-	M	X	.061	137	.911	.647	.752	.109	0.34	.114	0.35
107.03	88.9	M	-	M	X	.072	145	.911	.647	.752	.128	0.36	.134	0.37
108.00	90.9	M	-	M	-	.050	249	.911	.642	.752	.090	0.61	.094	0.62
109.00	0.0	C	-	C	-	.042	102	.882	.811	.752	.069	0.30	.074	0.29
110.00	0.0	M	X	-	-	.476	75	.911	.642	.752	.727	0.25	.761	0.26
111.01	0.0	C	-	-	-	.467	78	.913	.642	.791	.692	0.28	.724	0.29
111.02	0.0	C	-	-	-	.440	98	.913	.642	.791	.660	0.34	.691	0.35
112.00	0.0	-	-	-	-	.240	93	.843	.642	.791	.407	0.25	.443	0.24
113.00	0.0	-	-	M	X	.236	148	.863	.639	.750	.425	0.36	.457	0.35
114.00	0.0	-	-	M	-	.335	313	.911	.647	.752	.549	0.92	.576	0.94
116.01	0.0	C	-	-	-	.355	126	.882	.642	.776	.572	0.38	.609	0.38
116.03	0.0	C	-	-	-	.556	70	.882	.642	.752	.815	0.26	.868	0.25
116.04	0.0	C	-	-	-	.326	215	.882	.642	.776	.532	0.64	.566	0.63
116.06	0.0	C	-	-	-	.404	230	.882	.642	.791	.625	0.76	.665	0.75
116.08	0.0	C	-	-	-	.534	304	.882	.642	.752	.795	1.08	.846	1.07
116.09	0.0	C	-	-	-	.446	169	.882	.642	.752	.702	0.54	.748	0.53
116.10	0.0	C	-	-	-	.547	407	.882	.642	.791	.786	1.55	.837	1.53
116.11	0.0	C	-	-	-	.749	76	.882	.642	.791	.959	0.37	1.021	0.36
116.13	0.0	-	-	-	-	.760	180	.863	.642	.776	.967	0.88	1.040	0.85
117.00	0.0	-	-	-	-	.700	42	.863	.642	.791	.923	0.19	.993	0.18
118.01	0.0	-	-	C	-	.493	46	.935	.642	.791	.717	0.17	.741	0.18
118.02	0.0	-	-	C	-	.570	30	.935	.642	.791	.800	0.12	.827	0.13
118.03	0.0	-	-	C	-	.528	54	.935	.642	.791	.756	0.21	.781	0.22
119.01	0.0	-	-	C	-	.210	100	.863	.680	.752	.370	0.25	.398	0.25
119.02	0.0	-	-	C	-	.216	99	.863	.680	.791	.354	0.29	.381	0.28
120.01	0.0	-	-	-	-	.530	127	.863	.642	.752	.796	0.44	.856	0.43
120.02	0.0	-	-	-	-	.600	101	.863	.642	.752	.857	0.39	.923	0.38
120.03	0.0	-	-	-	-	.570	222	.863	.642	.844	.786	0.92	.846	0.89
120.04	0.0	-	-	-	-	.590	13	.863	.639	.752	.850	0.05	.915	0.05
122.00	0.0	-	-	-	-	.560	967	.935	.642	.791	.789	3.84	.816	4.02
125.01	0.0	-	-	C	-	.293	738	.863	.642	.811	.466	2.30	.502	2.22
125.02	0.0	-	-	C	-	.262	221	.863	.642	.778	.444	0.61	.478	0.59
125.03	0.0	-	-	-	-	.160	196	.863	.642	.644	.377	0.28	.405	0.27
125.04	0.0	-	-	C	-	.198	198	.863	.642	.711	.390	0.40	.420	0.39
125.05	0.0	-	-	C	-	.233	285	.863	.642	.900	.331	1.11	.356	1.08
125.06	0.0	-	-	C	-	.205	120	.863	.642	.722	.394	0.26	.424	0.25
126.01	0.0	-	-	C	-	.337	322	.911	.642	.689	.601	0.80	.630	0.81
126.02	0.0	-	-	C	-	.283	37	.911	.642	.737	.487	0.10	.510	0.10
126.03	0.0	-	-	C	-	.386	64	.911	.642	.625	.717	0.15	.752	0.15
126.04	0.0	-	-	C	-	.378	81	.911	.642	.641	.694	0.19	.727	0.19
126.05	0.0	-	-	C	-	.153	59	.911	.642	.673	.311	0.11	.326	0.11
126.06	0.0	-	-	C	-	.259	83	.911	.642	.609	.557	0.14	.583	0.14
127.01	0.0	M	X	C	-	.211	33	.895	.642	.791	.348	0.10	.368	0.10
127.02	0.0	M	X	C	-	.337	30	.895	.642	.791	.533	0.09	.563	0.09
129.00	50.0	M	X	-	-	.082	93	.911	.642	.752	.146	0.23	.153	0.23
130.00	0.0	M	X	-	-	.305	48	.911	.639	.752	.510	0.14	.534	0.14
133.00	0.0	-	-	-	-	.330	112	.925	.642	.785	.518	0.36	.539	0.37
134.00	0.0	-	-	-	-	.720	27	.911	.642	.890	.920	0.13	.964	0.13
135.00	0.0	-	-	M	-	.470	861	.895	.647	.791	.698	3.08	.738	3.08
136.00	0.0	-	-	-	-	.360	97	.869	.642	.776	.583	0.29	.625	0.28

Tabelle E.12
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
137.01	0.0	-	-	-	-	.200	74	.863	.642	.752	.363	0.18	.391	0.17
137.02	0.0	-	-	-	-	.400	64	.863	.642	.791	.625	0.21	.673	0.20
137.03	0.0	-	-	-	-	.480	44	.863	.642	.791	.720	0.15	.775	0.15
137.04	0.0	-	-	-	-	.190	59	.863	.642	.791	.322	0.16	.347	0.16
137.05	0.0	-	-	-	-	.100	27	.863	.642	.791	.173	0.07	.186	0.07
137.06	0.0	-	-	-	-	.280	49	.863	.642	.791	.461	0.14	.497	0.14
137.07	0.0	-	-	-	-	.120	50	.863	.642	.791	.207	0.13	.223	0.13
138.00	86.1	C	-	M	-	.235	22	.895	.520	.791	.424	0.05	.448	0.05
139.00	72.2	M	-	M	-	.104	119	.895	.642	.791	.175	0.33	.185	0.33
140.00	0.0	C	-	-	-	.469	146	.940	.642	.791	.688	0.54	.710	0.56
141.00	0.0	-	-	-	-	.750	34	.863	.642	.791	.959	0.16	1.032	0.16
142.00	0.0	-	-	-	-	.750	8	.863	.642	.752	.962	0.04	1.035	0.04
143.01	0.0	-	-	-	-	.322	123	.863	.642	.791	.522	0.37	.561	0.36
143.02	0.0	-	-	-	-	.644	147	.863	.642	.791	.878	0.62	.946	0.60
143.03	0.0	-	-	-	-	.644	62	.863	.642	.791	.878	0.26	.946	0.25
144.00	0.0	-	-	-	-	.560	315	.935	.642	.791	.789	1.25	.816	1.31
146.00	75.9	M	-	-	-	.120	45	.869	.642	.791	.206	0.12	.221	0.12
147.00	0.0	C	X	-	-	.196	308	.895	.520	.776	.367	0.69	.388	0.69
148.00	0.0	C	-	-	-	.521	91	.869	.642	.700	.817	0.29	.876	0.28
149.00	0.0	M	-	C	-	.093	199	.895	.642	.752	.168	0.48	.178	0.48
150.00	0.0	-	-	-	-	.322	167	.895	.642	.776	.522	0.50	.552	0.50
151.00	0.0	C	-	M	-	.200	200	.935	.647	.791	.321	0.61	.332	0.64
152.00	0.0	C	-	M	-	.529	180	.934	.639	.939	.688	0.84	.712	0.88
153.00	0.0	-	-	-	-	.570	153	.870	.639	.737	.839	0.56	.900	0.54
156.00	0.0	-	-	-	-	.310	122	.911	.642	.970	.399	0.58	.418	0.59
159.00	0.0	C	X	C	-	.523	150	.877	.642	.776	.771	0.54	.824	0.53
160.01	0.0	C	-	-	-	.351	58	.861	.642	.926	.475	0.25	.512	0.24
160.02	0.0	C	-	-	-	.160	91	.855	.642	.867	.242	0.31	.261	0.30
161.01	0.0	M	-	-	-	.016	60	.895	.642	.873	.024	0.21	.025	0.22
161.02	0.0	M	-	-	-	.086	31	.895	.642	.752	.156	0.07	.165	0.07
164.01	0.0	-	-	-	-	.245	486	.895	.642	.791	.400	1.44	.423	1.44
164.02	0.0	-	-	-	-	.350	198	.895	.642	.791	.550	0.63	.582	0.63
167.00	0.0	C	X	C	-	.078	139	.895	.639	.791	.132	0.38	.140	0.38
168.00	0.0	-	-	M	-	.476	51	.895	.639	.791	.708	0.18	.749	0.18
171.01	0.0	-	-	C	-	.390	197	.875	.740	.950	.478	1.04	.511	1.01
171.02	0.0	-	-	C	-	.332	142	.875	.720	1.018	.384	0.84	.410	0.82
171.03	0.0	-	-	C	-	.277	141	.875	.711	1.084	.300	0.95	.320	0.93
171.04	0.0	-	-	C	-	-.167	21	.875	.748	.757	-.279	0.06	-.298	0.06
172.00	0.0	-	-	C	-	.196	97	.869	.665	1.034	.231	0.55	.248	0.54
173.00	0.0	-	-	-	-	.260	190	.863	.642	.700	.505	0.40	.543	0.38
174.00	0.0	-	-	M	X	.315	469	.895	.639	.791	.503	1.45	.532	1.45
175.00	45.5	-	-	M	X	.296	187	.935	.639	.791	.465	0.60	.481	0.62
176.01	0.0	C	-	-	-	.515	23	.915	.642	.791	.746	0.09	.780	0.09
176.02	0.0	C	-	-	-	.563	16	.915	.642	.791	.796	0.06	.832	0.06
177.01	33.3	M	-	-	-	.221	27	.935	.642	.791	.354	0.08	.366	0.09
177.02	66.7	M	-	-	-	.100	27	.935	.642	.791	.164	0.08	.169	0.08
178.00	0.0	C	-	M	X	.211	176	.911	.647	.752	.364	0.47	.381	0.48
179.01	0.0	C	-	-	-	.027	12	.947	.642	.907	.038	0.05	.039	0.05
179.02	0.0	C	-	-	-	.711	15	.947	.642	.907	.906	0.07	.931	0.08
179.03	0.0	C	-	-	-	.226	9	.947	.642	.907	.312	0.04	.320	0.04
179.04	0.0	C	-	-	-	.336	10	.947	.642	.907	.458	0.04	.471	0.04
179.05	0.0	C	-	-	-	.793	7	.947	.642	.907	.991	0.04	1.019	0.04
179.06	0.0	C	-	-	-	.141	15	.947	.642	.907	.196	0.06	.201	0.07
179.07	0.0	C	-	-	-	.396	14	.947	.642	.907	.536	0.06	.550	0.06
179.08	0.0	M	-	-	-	.477	51	.895	.642	.791	.708	0.18	.749	0.18
179.09	0.0	M	-	-	-	.363	64	.895	.642	.791	.568	0.21	.600	0.21
179.10	0.0	M	-	-	-	.480	23	.895	.642	.791	.712	0.08	.752	0.08
179.11	0.0	M	-	-	-	.273	26	.895	.642	.791	.442	0.08	.467	0.08
179.12	0.0	M	-	-	-	.530	20	.895	.642	.752	.788	0.07	.832	0.07

Tabelle E.12
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
179.13	0.0	M	-	-	-	.393	52	.895	.642	.752	.634	0.16	.670	0.16
179.14	0.0	M	-	-	-	.443	19	.895	.642	.752	.694	0.06	.734	0.06
181.01	0.0	-	-	-	-	.120	280	.869	.642	.776	.212	0.71	.227	0.69
181.02	0.0	-	-	-	-	.250	303	.869	.642	.776	.425	0.83	.456	0.80
181.03	0.0	-	-	-	-	.300	70	.869	.642	.776	.499	0.20	.536	0.19
181.04	0.0	-	-	-	-	.100	110	.869	.642	.776	.177	0.28	.190	0.27
182.00	0.0	-	-	-	-	.090	200	.895	.642	.776	.156	0.53	.165	0.53
183.00	0.0	-	-	-	-	.070	256	.895	.642	.752	.127	0.61	.134	0.61
184.01	0.0	-	-	-	-	.220	152	.935	.642	.721	.390	0.38	.403	0.40
184.02	0.0	-	-	-	-	.040	26	.935	.642	.721	.074	0.06	.076	0.06
184.03	0.0	-	-	-	-	.310	89	.920	.642	.741	.520	0.25	.542	0.26
185.00	0.0	-	-	M	-	.288	90	.895	.642	.776	.473	0.26	.500	0.26
186.00	0.0	-	-	-	-	.320	167	.895	.642	.776	.519	0.50	.549	0.50
189.01	0.0	C	-	-	-	.432	102	.895	.639	.752	.683	0.32	.722	0.32
189.02	0.0	C	-	-	-	.338	403	.895	.639	.791	.535	1.27	.566	1.27
189.03	0.0	C	-	-	-	.437	282	.895	.639	.791	.663	0.97	.700	0.97
189.04	0.0	C	-	-	-	.377	87	.895	.639	.791	.587	0.28	.621	0.28
189.05	0.0	C	-	-	-	.518	68	.895	.639	.791	.755	0.25	.798	0.25
189.06	0.0	C	-	-	-	.501	731	.895	.639	.791	.736	2.67	.778	2.67
189.07	0.0	C	-	-	-	.437	198	.895	.639	.791	.663	0.68	.700	0.68
189.08	0.0	C	-	-	-	.372	112	.895	.639	.791	.581	0.36	.614	0.36
189.09	0.0	C	-	-	-	.398	351	.895	.639	.791	.614	1.16	.649	1.16
189.10	0.0	C	-	-	-	.317	394	.895	.639	.791	.506	1.22	.535	1.22
189.11	0.0	C	-	-	-	.218	81	.895	.639	.791	.359	0.24	.380	0.24
189.12	0.0	C	-	-	-	.368	191	.895	.639	.791	.575	0.62	.608	0.62
189.13	0.0	C	-	-	-	.368	97	.895	.639	.791	.575	0.31	.608	0.31
189.14	0.0	C	-	-	-	.616	617	.895	.639	.791	.852	2.54	.900	2.55
189.15	0.0	C	-	-	-	.475	418	.895	.639	.791	.707	1.49	.748	1.49
189.16	0.0	C	-	-	-	.428	412	.895	.639	.791	.652	1.40	.689	1.40
189.17	0.0	C	-	-	-	.394	243	.895	.639	.791	.609	0.80	.644	0.80
189.18	0.0	C	-	-	-	.330	98	.895	.639	.791	.524	0.31	.554	0.31
189.19	0.0	C	-	-	-	.561	287	.895	.639	.791	.799	1.12	.845	1.12
189.20	0.0	C	-	-	-	.411	291	.895	.639	.791	.631	0.97	.667	0.98
189.21	0.0	C	-	-	-	.278	184	.895	.639	.791	.450	0.55	.476	0.55
189.22	0.0	C	-	-	-	.385	84	.895	.639	.791	.598	0.27	.632	0.28
189.23	0.0	C	-	-	-	.214	158	.895	.639	.791	.353	0.46	.373	0.46
189.24	0.0	C	-	-	-	.300	119	.895	.639	.791	.482	0.36	.509	0.36
189.25	0.0	C	-	-	-	.381	67	.895	.639	.791	.593	0.22	.626	0.22
189.26	0.0	C	-	-	-	.407	57	.895	.639	.791	.626	0.19	.661	0.19
189.27	0.0	C	-	-	-	.377	81	.895	.639	.776	.598	0.25	.632	0.25
189.28	0.0	C	-	-	-	.479	181	.895	.639	.791	.712	0.65	.752	0.65
189.29	0.0	C	-	-	-	.347	314	.895	.639	.791	.547	1.00	.579	1.00
189.30	0.0	C	-	-	-	.407	454	.895	.639	.791	.626	1.51	.661	1.52
189.31	0.0	C	-	-	-	.501	380	.895	.639	.791	.736	1.39	.778	1.39
189.32	0.0	C	-	-	-	.381	412	.895	.639	.791	.593	1.34	.626	1.35
189.33	0.0	C	-	-	-	.488	391	.895	.639	.791	.722	1.41	.763	1.41
189.34	0.0	C	-	-	-	.321	48	.895	.639	.791	.512	0.15	.541	0.15
189.35	0.0	C	-	-	-	.390	58	.895	.639	.752	.631	0.17	.667	0.17
189.36	0.0	C	-	-	-	.424	57	.895	.639	.752	.673	0.18	.712	0.18
189.37	0.0	C	-	-	-	.223	87	.895	.639	.752	.390	0.22	.413	0.22
189.38	0.0	C	-	-	-	.334	56	.895	.639	.791	.530	0.18	.560	0.18
189.39	0.0	C	-	-	-	.248	48	.895	.639	.791	.405	0.14	.428	0.14
189.40	0.0	C	-	-	-	.240	61	.895	.639	.791	.393	0.18	.416	0.18
189.41	0.0	C	-	-	-	.214	104	.895	.639	.791	.353	0.30	.373	0.30
189.42	0.0	C	-	-	-	.227	113	.895	.639	.791	.373	0.33	.395	0.33
189.43	0.0	C	-	-	-	.338	38	.895	.639	.791	.535	0.12	.566	0.12
189.44	0.0	C	-	-	-	.295	21	.895	.639	.791	.475	0.06	.502	0.06
189.45	0.0	C	-	-	-	.390	178	.895	.639	.752	.631	0.54	.667	0.54
189.46	0.0	C	-	-	-	.432	347	.895	.639	.752	.683	1.10	.722	1.10

Tabelle E.12
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.47	0.0	C	-	-	-	.454	37	.895	.639	.752	.708	0.12	.749	0.12
189.48	0.0	C	-	-	-	.424	41	.895	.639	.752	.673	0.13	.712	0.13
189.49	0.0	C	-	-	-	.595	104	.895	.639	.752	.849	0.40	.897	0.40
189.50	0.0	C	-	-	-	.454	174	.895	.639	.752	.708	0.56	.749	0.56
189.51	0.0	C	-	-	-	.257	104	.895	.639	.791	.419	0.31	.443	0.31
189.52	0.0	C	-	-	-	.509	32	.895	.639	.752	.768	0.11	.811	0.11
189.53	0.0	C	-	-	-	.574	268	.895	.639	.791	.812	1.06	.858	1.06
189.54	0.0	C	-	-	-	.432	212	.895	.639	.791	.657	0.72	.694	0.72
189.55	0.0	C	-	-	-	.488	98	.895	.639	.791	.722	0.35	.763	0.35
189.56	0.0	C	-	-	-	.484	87	.895	.639	.791	.717	0.31	.758	0.31
189.57	0.0	C	-	-	-	.424	106	.895	.639	.791	.647	0.36	.684	0.36
189.58	0.0	C	-	-	-	.287	124	.895	.639	.791	.463	0.38	.489	0.38
189.59	0.0	C	-	-	-	.428	168	.895	.639	.791	.652	0.57	.689	0.57
189.60	0.0	C	-	-	-	.372	117	.895	.639	.791	.581	0.38	.614	0.38
189.61	0.0	C	-	-	-	.407	198	.895	.639	.791	.626	0.66	.661	0.66
189.62	0.0	C	-	-	-	.390	314	.895	.639	.791	.604	1.03	.639	1.03
189.63	0.0	C	-	-	-	.420	217	.895	.639	.791	.642	0.73	.679	0.73
189.64	0.0	C	-	-	-	.462	208	.895	.639	.791	.692	0.73	.732	0.73
189.65	0.0	C	-	-	-	.458	197	.895	.639	.791	.688	0.69	.727	0.69
189.66	0.0	C	-	-	-	.458	128	.895	.639	.791	.688	0.45	.727	0.45
189.67	0.0	C	-	-	-	.471	117	.895	.639	.791	.703	0.41	.743	0.42
189.68	0.0	C	-	-	-	.462	89	.895	.639	.791	.692	0.31	.732	0.31
189.69	0.0	C	-	-	-	.492	58	.895	.639	.791	.726	0.21	.768	0.21
189.70	0.0	C	-	-	-	.492	87	.895	.639	.752	.750	0.30	.793	0.30

Anmerkungen. siehe Seite 275.

Tabelle E.13
Lernleistung, prädiktive Validität

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
100.00	60.0	-	-	C	-	.268	76	.911	.820	1.003	.295	0.50	.309	0.51
101.01	0.0	-	-	C	-	.147	72	.863	.639	.467	.667	0.03	.718	0.03
101.02	0.0	-	-	C	-	.296	65	.863	.784	.489	.804	0.07	.866	0.07
102.00	54.5	C	-	-	-	.213	54	.918	.642	.705	.395	0.13	.412	0.13
103.00	0.0	C	X	M	-	.152	90	.914	.647	.752	.266	0.24	.278	0.24
104.00	0.0	M	X	-	-	.240	158	.911	.642	.752	.411	0.43	.431	0.44
105.00	88.9	M	X	M	X	.082	103	.911	.639	.752	.147	0.26	.154	0.26
106.00	0.0	M	-	M	X	.091	224	.895	.647	.791	.153	0.63	.162	0.64
107.01	86.4	M	-	M	X	.106	56	.911	.647	.752	.188	0.14	.197	0.15
107.02	94.4	M	-	M	X	.058	137	.911	.647	.752	.103	0.35	.108	0.35
107.03	88.9	M	-	M	X	.069	145	.911	.647	.752	.123	0.37	.129	0.37
108.00	90.9	M	-	M	-	.047	249	.911	.642	.752	.084	0.62	.088	0.63
109.00	0.0	C	-	C	-	.042	102	.882	.811	.752	.069	0.30	.074	0.30
110.00	0.0	M	X	-	-	.476	75	.911	.642	.752	.727	0.26	.761	0.26
111.01	0.0	C	-	-	-	.467	78	.913	.642	.791	.692	0.28	.724	0.29
111.02	0.0	C	-	-	-	.440	98	.913	.642	.791	.660	0.35	.691	0.36
112.00	0.0	-	-	-	-	.240	93	.843	.642	.791	.407	0.26	.443	0.24
113.00	0.0	-	-	M	X	.236	148	.863	.639	.750	.425	0.37	.457	0.35
114.00	0.0	-	-	M	-	.335	313	.911	.647	.752	.549	0.93	.576	0.95
116.01	0.0	C	-	-	-	.355	126	.882	.642	.776	.572	0.39	.609	0.38
116.03	0.0	C	-	-	-	.556	70	.882	.642	.752	.815	0.26	.868	0.26
116.04	0.0	C	-	-	-	.326	215	.882	.642	.776	.532	0.65	.566	0.64
116.06	0.0	C	-	-	-	.404	230	.882	.642	.791	.625	0.77	.665	0.76

Tabelle E.13
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
116.08	0.0	C	-	-	-	.534	304	.882	.642	.752	.795	1.10	.846	1.08
116.09	0.0	C	-	-	-	.446	169	.882	.642	.752	.702	0.55	.748	0.54
116.10	0.0	C	-	-	-	.547	407	.882	.642	.791	.786	1.58	.837	1.56
116.11	0.0	C	-	-	-	.749	76	.882	.642	.791	.959	0.37	1.021	0.37
116.13	0.0	-	-	-	-	.760	180	.863	.642	.776	.967	0.89	1.040	0.86
117.00	0.0	-	-	-	-	.700	42	.863	.642	.791	.923	0.19	.993	0.19
118.01	0.0	-	-	C	-	.493	46	.935	.642	.791	.717	0.17	.741	0.18
118.02	0.0	-	-	C	-	.570	30	.935	.642	.791	.800	0.12	.827	0.13
118.03	0.0	-	-	C	-	.528	54	.935	.642	.791	.756	0.21	.781	0.22
119.01	0.0	-	-	C	-	.210	100	.863	.680	.752	.370	0.26	.398	0.25
119.02	0.0	-	-	C	-	.216	99	.863	.680	.791	.354	0.29	.381	0.28
120.01	0.0	-	-	-	-	.530	127	.863	.642	.752	.796	0.45	.856	0.44
120.02	0.0	-	-	-	-	.600	101	.863	.642	.752	.857	0.40	.923	0.38
120.03	0.0	-	-	-	-	.570	222	.863	.642	.844	.786	0.94	.846	0.90
120.04	0.0	-	-	-	-	.590	13	.863	.639	.752	.850	0.05	.915	0.05
122.00	0.0	-	-	-	-	.560	967	.935	.642	.791	.789	3.90	.816	4.08
123.01	0.0	-	-	-	-	.710	20	.895	.520	.791	.988	0.08	1.044	0.08
123.02	0.0	-	-	-	-	.650	20	.895	.520	.791	.939	0.08	.993	0.08
123.03	0.0	-	-	-	-	.800	19	.895	.520	.791	1.050	0.09	1.110	0.09
123.04	0.0	-	-	-	-	.720	7	.895	.520	.791	.995	0.03	1.052	0.03
123.05	0.0	-	-	-	-	.590	13	.895	.520	.791	.885	0.05	.936	0.05
123.06	0.0	-	-	-	-	.850	10	.895	.520	.791	1.080	0.05	1.142	0.05
125.01	0.0	-	-	C	-	.293	738	.863	.642	.811	.466	2.33	.502	2.25
125.02	0.0	-	-	C	-	.262	221	.863	.642	.778	.444	0.62	.478	0.60
125.03	0.0	-	-	-	-	.160	196	.863	.642	.644	.377	0.28	.405	0.27
125.04	0.0	-	-	C	-	.198	198	.863	.642	.711	.390	0.41	.420	0.39
125.05	0.0	-	-	C	-	.233	285	.863	.642	.900	.331	1.13	.356	1.09
125.06	0.0	-	-	C	-	.205	120	.863	.642	.722	.394	0.26	.424	0.25
126.01	0.0	-	-	C	-	.337	322	.911	.642	.689	.601	0.81	.630	0.83
126.02	0.0	-	-	C	-	.283	37	.911	.642	.737	.487	0.10	.510	0.10
126.03	0.0	-	-	C	-	.386	64	.911	.642	.625	.717	0.15	.752	0.15
126.04	0.0	-	-	C	-	.378	81	.911	.642	.641	.694	0.19	.727	0.20
126.05	0.0	-	-	C	-	.153	59	.911	.642	.673	.311	0.11	.326	0.12
126.06	0.0	-	-	C	-	.259	83	.911	.642	.609	.557	0.14	.583	0.15
127.01	0.0	M	X	C	-	.211	33	.895	.642	.791	.348	0.10	.368	0.10
127.02	0.0	M	X	C	-	.337	30	.895	.642	.791	.533	0.10	.563	0.10
129.00	50.0	M	X	-	-	.080	93	.911	.642	.752	.143	0.23	.150	0.24
130.00	0.0	M	X	-	-	.305	48	.911	.639	.752	.510	0.14	.534	0.14
133.00	0.0	-	-	-	-	.330	112	.925	.642	.785	.518	0.36	.539	0.38
134.00	0.0	-	-	-	-	.720	27	.911	.642	.890	.920	0.13	.964	0.14
135.00	0.0	-	-	M	-	.470	861	.895	.647	.791	.698	3.12	.738	3.13
137.01	0.0	-	-	-	-	.200	74	.863	.642	.752	.363	0.18	.391	0.17
137.02	0.0	-	-	-	-	.400	64	.863	.642	.791	.625	0.21	.673	0.20
137.03	0.0	-	-	-	-	.480	44	.863	.642	.791	.720	0.16	.775	0.15
137.04	0.0	-	-	-	-	.190	59	.863	.642	.791	.322	0.16	.347	0.16
137.05	0.0	-	-	-	-	.100	27	.863	.642	.791	.173	0.07	.186	0.07
137.06	0.0	-	-	-	-	.280	49	.863	.642	.791	.461	0.14	.497	0.14
137.07	0.0	-	-	-	-	.120	50	.863	.642	.791	.207	0.13	.223	0.13
138.00	86.1	C	-	M	-	.226	22	.895	.520	.791	.409	0.05	.432	0.05
139.00	72.2	M	-	M	-	.101	119	.895	.642	.791	.170	0.34	.180	0.34
140.00	0.0	C	-	-	-	.469	146	.940	.642	.791	.688	0.54	.710	0.57
141.00	0.0	-	-	-	-	.750	34	.863	.642	.791	.959	0.17	1.032	0.16
143.01	0.0	-	-	-	-	.322	123	.863	.642	.791	.522	0.38	.561	0.36
143.02	0.0	-	-	-	-	.644	147	.863	.642	.791	.878	0.63	.946	0.61
143.03	0.0	-	-	-	-	.644	62	.863	.642	.791	.878	0.27	.946	0.26
144.00	0.0	-	-	-	-	.560	315	.935	.642	.791	.789	1.27	.816	1.33
146.00	75.9	M	-	-	-	.115	45	.869	.642	.791	.197	0.12	.212	0.12
148.00	0.0	C	-	-	-	.521	91	.869	.642	.700	.817	0.30	.876	0.29
149.00	0.0	M	-	C	-	.093	199	.895	.642	.752	.168	0.49	.178	0.49

Tabelle E.13
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
150.00	0.0	-	-	-	-	.322	167	.895	.642	.776	.522	0.51	.552	0.51
151.00	0.0	C	-	M	-	.200	200	.935	.647	.791	.321	0.62	.332	0.65
152.00	0.0	C	-	M	-	.529	180	.934	.639	.939	.688	0.85	.712	0.89
153.00	0.0	-	-	-	-	.570	153	.870	.639	.737	.839	0.57	.900	0.55
156.00	0.0	-	-	-	-	.310	122	.911	.642	.970	.399	0.59	.418	0.60
159.00	0.0	C	X	C	-	.523	150	.877	.642	.776	.771	0.55	.824	0.54
160.01	0.0	C	-	-	-	.351	58	.861	.642	.926	.475	0.25	.512	0.24
160.02	0.0	C	-	-	-	.160	91	.855	.642	.867	.242	0.32	.261	0.31
161.01	0.0	M	-	-	-	.016	60	.895	.642	.873	.024	0.22	.025	0.22
161.02	0.0	M	-	-	-	.086	31	.895	.642	.752	.156	0.08	.165	0.08
164.01	0.0	-	-	-	-	.245	486	.895	.642	.791	.400	1.46	.423	1.46
164.02	0.0	-	-	-	-	.350	198	.895	.642	.791	.550	0.64	.582	0.64
167.00	0.0	C	X	C	-	.078	139	.895	.639	.791	.132	0.39	.140	0.39
168.00	0.0	-	-	M	-	.476	51	.895	.639	.791	.708	0.18	.749	0.18
171.01	0.0	-	-	C	-	.390	197	.875	.740	.950	.478	1.05	.511	1.03
171.02	0.0	-	-	C	-	.332	142	.875	.720	1.018	.384	0.85	.410	0.83
171.03	0.0	-	-	C	-	.277	141	.875	.711	1.084	.300	0.96	.320	0.94
171.04	0.0	-	-	C	-	-.167	21	.875	.748	.757	-.279	0.06	-.298	0.06
172.00	0.0	-	-	C	-	.196	97	.869	.665	1.034	.231	0.56	.248	0.54
173.00	0.0	-	-	-	-	.260	190	.863	.642	.700	.505	0.40	.543	0.39
174.00	0.0	-	-	M	X	.315	469	.895	.639	.791	.503	1.47	.532	1.47
175.00	45.5	-	-	M	X	.295	187	.935	.639	.791	.464	0.61	.480	0.63
176.01	0.0	C	-	-	-	.515	23	.915	.642	.791	.746	0.09	.780	0.09
176.02	0.0	C	-	-	-	.563	16	.915	.642	.791	.796	0.06	.832	0.07
177.01	33.3	M	-	-	-	.220	27	.935	.642	.791	.352	0.08	.364	0.09
177.02	66.7	M	-	-	-	.117	27	.935	.642	.791	.191	0.08	.198	0.08
178.00	0.0	C	-	M	X	.211	176	.911	.647	.752	.364	0.47	.381	0.48
179.01	0.0	C	-	-	-	.027	12	.947	.642	.907	.038	0.05	.039	0.05
179.02	0.0	C	-	-	-	.711	15	.947	.642	.907	.906	0.07	.931	0.08
179.03	0.0	C	-	-	-	.226	9	.947	.642	.907	.312	0.04	.320	0.04
179.04	0.0	C	-	-	-	.336	10	.947	.642	.907	.458	0.04	.471	0.05
179.05	0.0	C	-	-	-	.793	7	.947	.642	.907	.991	0.04	1.019	0.04
179.06	0.0	C	-	-	-	.141	15	.947	.642	.907	.196	0.06	.201	0.07
179.07	0.0	C	-	-	-	.396	14	.947	.642	.907	.536	0.06	.550	0.06
181.01	0.0	-	-	-	-	.120	280	.869	.642	.776	.212	0.72	.227	0.70
181.02	0.0	-	-	-	-	.250	303	.869	.642	.776	.425	0.84	.456	0.82
181.03	0.0	-	-	-	-	.300	70	.869	.642	.776	.499	0.20	.536	0.20
181.04	0.0	-	-	-	-	.100	110	.869	.642	.776	.177	0.28	.190	0.27
182.00	0.0	-	-	-	-	.090	200	.895	.642	.776	.156	0.53	.165	0.53
183.00	0.0	-	-	-	-	.070	256	.895	.642	.752	.127	0.62	.134	0.62
184.01	0.0	-	-	-	-	.220	152	.935	.642	.721	.390	0.39	.403	0.41
184.02	0.0	-	-	-	-	.040	26	.935	.642	.721	.074	0.06	.076	0.06
184.03	0.0	-	-	-	-	.310	89	.920	.642	.741	.520	0.25	.542	0.26
185.00	0.0	-	-	M	-	.288	90	.895	.642	.776	.473	0.27	.500	0.27
186.00	0.0	-	-	-	-	.320	167	.895	.642	.776	.519	0.51	.549	0.51
189.01	0.0	C	-	-	-	.432	102	.895	.639	.752	.683	0.33	.722	0.33
189.02	0.0	C	-	-	-	.338	403	.895	.639	.791	.535	1.29	.566	1.29
189.03	0.0	C	-	-	-	.437	282	.895	.639	.791	.663	0.98	.700	0.98
189.04	0.0	C	-	-	-	.377	87	.895	.639	.791	.587	0.29	.621	0.29
189.05	0.0	C	-	-	-	.518	68	.895	.639	.791	.755	0.26	.798	0.26
189.06	0.0	C	-	-	-	.501	731	.895	.639	.791	.736	2.71	.778	2.71
189.07	0.0	C	-	-	-	.437	198	.895	.639	.791	.663	0.69	.700	0.69
189.08	0.0	C	-	-	-	.372	112	.895	.639	.791	.581	0.37	.614	0.37
189.09	0.0	C	-	-	-	.398	351	.895	.639	.791	.614	1.18	.649	1.18
189.10	0.0	C	-	-	-	.317	394	.895	.639	.791	.506	1.24	.535	1.24
189.11	0.0	C	-	-	-	.218	81	.895	.639	.791	.359	0.24	.380	0.24
189.12	0.0	C	-	-	-	.368	191	.895	.639	.791	.575	0.63	.608	0.63
189.13	0.0	C	-	-	-	.368	97	.895	.639	.791	.575	0.32	.608	0.32
189.14	0.0	C	-	-	-	.616	617	.895	.639	.791	.852	2.58	.900	2.59

Tabelle E.13
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.15	0.0	C	-	-	-	.475	418	.895	.639	.791	.707	1.51	.748	1.51
189.16	0.0	C	-	-	-	.428	412	.895	.639	.791	.652	1.42	.689	1.42
189.17	0.0	C	-	-	-	.394	243	.895	.639	.791	.609	0.81	.644	0.81
189.18	0.0	C	-	-	-	.330	98	.895	.639	.791	.524	0.31	.554	0.31
189.19	0.0	C	-	-	-	.561	287	.895	.639	.791	.799	1.13	.845	1.13
189.20	0.0	C	-	-	-	.411	291	.895	.639	.791	.631	0.99	.667	0.99
189.21	0.0	C	-	-	-	.278	184	.895	.639	.791	.450	0.56	.476	0.56
189.22	0.0	C	-	-	-	.385	84	.895	.639	.791	.598	0.28	.632	0.28
189.23	0.0	C	-	-	-	.214	158	.895	.639	.791	.353	0.46	.373	0.47
189.24	0.0	C	-	-	-	.300	119	.895	.639	.791	.482	0.37	.509	0.37
189.25	0.0	C	-	-	-	.381	67	.895	.639	.791	.593	0.22	.626	0.22
189.26	0.0	C	-	-	-	.407	57	.895	.639	.791	.626	0.19	.661	0.19
189.27	0.0	C	-	-	-	.377	81	.895	.639	.776	.598	0.26	.632	0.26
189.28	0.0	C	-	-	-	.479	181	.895	.639	.791	.712	0.66	.752	0.66
189.29	0.0	C	-	-	-	.347	314	.895	.639	.791	.547	1.01	.579	1.01
189.30	0.0	C	-	-	-	.407	454	.895	.639	.791	.626	1.54	.661	1.54
189.31	0.0	C	-	-	-	.501	380	.895	.639	.791	.736	1.41	.778	1.41
189.32	0.0	C	-	-	-	.381	412	.895	.639	.791	.593	1.36	.626	1.37
189.33	0.0	C	-	-	-	.488	391	.895	.639	.791	.722	1.43	.763	1.43
189.34	0.0	C	-	-	-	.321	48	.895	.639	.791	.512	0.15	.541	0.15
189.35	0.0	C	-	-	-	.390	58	.895	.639	.752	.631	0.18	.667	0.18
189.36	0.0	C	-	-	-	.424	57	.895	.639	.752	.673	0.18	.712	0.18
189.37	0.0	C	-	-	-	.223	87	.895	.639	.752	.390	0.23	.413	0.23
189.38	0.0	C	-	-	-	.334	56	.895	.639	.791	.530	0.18	.560	0.18
189.39	0.0	C	-	-	-	.248	48	.895	.639	.791	.405	0.14	.428	0.14
189.40	0.0	C	-	-	-	.240	61	.895	.639	.791	.393	0.18	.416	0.18
189.41	0.0	C	-	-	-	.214	104	.895	.639	.791	.353	0.31	.373	0.31
189.42	0.0	C	-	-	-	.227	113	.895	.639	.791	.373	0.33	.395	0.34
189.43	0.0	C	-	-	-	.338	38	.895	.639	.791	.535	0.12	.566	0.12
189.44	0.0	C	-	-	-	.295	21	.895	.639	.791	.475	0.06	.502	0.07
189.45	0.0	C	-	-	-	.390	178	.895	.639	.752	.631	0.54	.667	0.54
189.46	0.0	C	-	-	-	.432	347	.895	.639	.752	.683	1.11	.722	1.11
189.47	0.0	C	-	-	-	.454	37	.895	.639	.752	.708	0.12	.749	0.12
189.48	0.0	C	-	-	-	.424	41	.895	.639	.752	.673	0.13	.712	0.13
189.49	0.0	C	-	-	-	.595	104	.895	.639	.752	.849	0.41	.897	0.41
189.50	0.0	C	-	-	-	.454	174	.895	.639	.752	.708	0.57	.749	0.57
189.51	0.0	C	-	-	-	.257	104	.895	.639	.791	.419	0.31	.443	0.31
189.52	0.0	C	-	-	-	.509	32	.895	.639	.752	.768	0.11	.811	0.11
189.53	0.0	C	-	-	-	.574	268	.895	.639	.791	.812	1.07	.858	1.07
189.54	0.0	C	-	-	-	.432	212	.895	.639	.791	.657	0.73	.694	0.74
189.55	0.0	C	-	-	-	.488	98	.895	.639	.791	.722	0.36	.763	0.36
189.56	0.0	C	-	-	-	.484	87	.895	.639	.791	.717	0.32	.758	0.32
189.57	0.0	C	-	-	-	.424	106	.895	.639	.791	.647	0.36	.684	0.37
189.58	0.0	C	-	-	-	.287	124	.895	.639	.791	.463	0.38	.489	0.38
189.59	0.0	C	-	-	-	.428	168	.895	.639	.791	.652	0.58	.689	0.58
189.60	0.0	C	-	-	-	.372	117	.895	.639	.791	.581	0.38	.614	0.38
189.61	0.0	C	-	-	-	.407	198	.895	.639	.791	.626	0.67	.661	0.67
189.62	0.0	C	-	-	-	.390	314	.895	.639	.791	.604	1.05	.639	1.05
189.63	0.0	C	-	-	-	.420	217	.895	.639	.791	.642	0.74	.679	0.74
189.64	0.0	C	-	-	-	.462	208	.895	.639	.791	.692	0.74	.732	0.74
189.65	0.0	C	-	-	-	.458	197	.895	.639	.791	.688	0.70	.727	0.70
189.66	0.0	C	-	-	-	.458	128	.895	.639	.791	.688	0.45	.727	0.46
189.67	0.0	C	-	-	-	.471	117	.895	.639	.791	.703	0.42	.743	0.42
189.68	0.0	C	-	-	-	.462	89	.895	.639	.791	.692	0.32	.732	0.32
189.69	0.0	C	-	-	-	.492	58	.895	.639	.791	.726	0.21	.768	0.21
189.70	0.0	C	-	-	-	.492	87	.895	.639	.752	.750	0.30	.793	0.30

Anmerkungen. siehe Seite 275.

Tabelle E.14

Lernleistung, konkurrente Validität

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
147.00	C	X	–	–	.196	308	.895	.520	.776	.367	45.2	.388	45.2
179.08	M	–	–	–	.477	51	.895	.642	.791	.708	11.9	.749	11.9
179.09	M	–	–	–	.363	64	.895	.642	.791	.568	13.5	.600	13.5
179.10	M	–	–	–	.480	23	.895	.642	.791	.712	5.4	.752	5.4
179.11	M	–	–	–	.273	26	.895	.642	.791	.442	5.1	.467	5.1
179.12	M	–	–	–	.530	20	.895	.642	.752	.788	4.7	.832	4.7
179.13	M	–	–	–	.393	52	.895	.642	.752	.634	10.3	.670	10.3
179.14	M	–	–	–	.443	19	.895	.642	.752	.694	4.0	.734	4.0

Anmerkungen. siehe Seite 275.

Tabelle E.15

Lernleistung, Skill Level 2

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
106.00	0.0	M	–	M	X	.091	224	.895	.647	.791	.153	0.84	.162	0.84
111.01	0.0	C	–	–	–	.467	78	.913	.642	.791	.692	0.38	.724	0.38
111.02	0.0	C	–	–	–	.440	98	.913	.642	.791	.660	0.46	.691	0.47
112.00	0.0	–	–	–	–	.240	93	.843	.642	.791	.407	0.34	.443	0.32
113.00	0.0	–	–	M	X	.236	148	.863	.639	.750	.425	0.48	.457	0.47
114.00	0.0	–	–	M	–	.335	313	.911	.647	.752	.549	1.23	.576	1.25
116.06	0.0	C	–	–	–	.404	230	.882	.642	.791	.625	1.02	.665	1.00
116.10	0.0	C	–	–	–	.547	407	.882	.642	.791	.786	2.08	.837	2.05
116.11	0.0	C	–	–	–	.749	76	.882	.642	.791	.959	0.49	1.021	0.48
117.00	0.0	–	–	–	–	.700	42	.863	.642	.791	.923	0.26	.993	0.25
118.01	0.0	–	–	C	–	.493	46	.935	.642	.791	.717	0.23	.741	0.24
118.02	0.0	–	–	C	–	.570	30	.935	.642	.791	.800	0.16	.827	0.17
118.03	0.0	–	–	C	–	.528	54	.935	.642	.791	.756	0.28	.781	0.29
119.02	0.0	–	–	C	–	.216	99	.863	.680	.791	.354	0.39	.381	0.38
122.00	0.0	–	–	–	–	.560	967	.935	.642	.791	.789	5.14	.816	5.38
123.01	0.0	–	–	–	–	.710	20	.895	.520	.791	.988	0.11	1.044	0.11
123.02	0.0	–	–	–	–	.650	20	.895	.520	.791	.939	0.10	.993	0.10
123.03	0.0	–	–	–	–	.800	19	.895	.520	.791	1.050	0.12	1.110	0.12
123.04	0.0	–	–	–	–	.720	7	.895	.520	.791	.995	0.04	1.052	0.04
123.05	0.0	–	–	–	–	.590	13	.895	.520	.791	.885	0.06	.936	0.06
123.06	0.0	–	–	–	–	.850	10	.895	.520	.791	1.080	0.07	1.142	0.07
124.01	0.0	M	–	–	–	.803	12	.895	.520	.791	1.052	0.07	1.112	0.07
124.02	0.0	M	–	–	–	.806	8	.895	.520	.791	1.054	0.05	1.114	0.05
124.03	0.0	M	–	–	–	.515	12	.895	.520	.791	.809	0.05	.855	0.05
125.01	0.0	–	–	C	–	.293	738	.863	.642	.811	.466	3.08	.502	2.97
125.03	0.0	–	–	–	–	.160	196	.863	.642	.644	.377	0.37	.405	0.36
125.04	0.0	–	–	C	–	.198	198	.863	.642	.711	.390	0.54	.420	0.52
125.05	0.0	–	–	C	–	.233	285	.863	.642	.900	.331	1.49	.356	1.44
125.06	0.0	–	–	C	–	.205	120	.863	.642	.722	.394	0.34	.424	0.33
126.01	0.0	–	–	C	–	.337	322	.911	.642	.689	.601	1.07	.630	1.09
126.03	0.0	–	–	C	–	.386	64	.911	.642	.625	.717	0.20	.752	0.20
126.04	0.0	–	–	C	–	.378	81	.911	.642	.641	.694	0.25	.727	0.26
126.05	0.0	–	–	C	–	.153	59	.911	.642	.673	.311	0.15	.326	0.15
126.06	0.0	–	–	C	–	.259	83	.911	.642	.609	.557	0.19	.583	0.19
127.01	0.0	M	X	C	–	.211	33	.895	.642	.791	.348	0.13	.368	0.13
127.02	0.0	M	X	C	–	.337	30	.895	.642	.791	.533	0.13	.563	0.13
134.00	0.0	–	–	–	–	.720	27	.911	.642	.890	.920	0.17	.964	0.18
135.00	0.0	–	–	M	–	.470	861	.895	.647	.791	.698	4.12	.738	4.13

Tabelle E.15
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
137.02	0.0	-	-	-	-	.400	64	.863	.642	.791	.625	0.28	.673	0.27
137.03	0.0	-	-	-	-	.480	44	.863	.642	.791	.720	0.21	.775	0.20
137.04	0.0	-	-	-	-	.190	59	.863	.642	.791	.322	0.22	.347	0.21
137.05	0.0	-	-	-	-	.100	27	.863	.642	.791	.173	0.10	.186	0.09
137.06	0.0	-	-	-	-	.280	49	.863	.642	.791	.461	0.19	.497	0.18
137.07	0.0	-	-	-	-	.120	50	.863	.642	.791	.207	0.18	.223	0.17
138.00	86.1	C	-	M	-	.236	22	.895	.520	.791	.426	0.07	.450	0.07
139.00	72.2	M	-	M	-	.104	119	.895	.642	.791	.175	0.44	.185	0.44
140.00	0.0	C	-	-	-	.469	146	.940	.642	.791	.688	0.72	.710	0.75
141.00	0.0	-	-	-	-	.750	34	.863	.642	.791	.959	0.22	1.032	0.21
143.01	0.0	-	-	-	-	.322	123	.863	.642	.791	.522	0.50	.561	0.48
143.02	0.0	-	-	-	-	.644	147	.863	.642	.791	.878	0.84	.946	0.81
143.03	0.0	-	-	-	-	.644	62	.863	.642	.791	.878	0.35	.946	0.34
144.00	0.0	-	-	-	-	.560	315	.935	.642	.791	.789	1.68	.816	1.75
151.00	0.0	C	-	M	-	.200	200	.935	.647	.791	.321	0.82	.332	0.86
153.00	0.0	-	-	-	-	.570	153	.870	.639	.737	.839	0.75	.900	0.73
164.01	0.0	-	-	-	-	.245	486	.895	.642	.791	.400	1.93	.423	1.93
164.02	0.0	-	-	-	-	.350	198	.895	.642	.791	.550	0.85	.582	0.85
167.00	0.0	C	X	C	-	.078	139	.895	.639	.791	.132	0.51	.140	0.51
168.00	0.0	-	-	M	-	.476	51	.895	.639	.791	.708	0.24	.749	0.24
171.01	0.0	-	-	C	-	.390	197	.875	.740	.950	.478	1.39	.511	1.36
171.02	0.0	-	-	C	-	.332	142	.875	.720	1.018	.384	1.12	.410	1.10
171.03	0.0	-	-	C	-	.277	141	.875	.711	1.084	.300	1.27	.320	1.24
171.04	0.0	-	-	C	-	-.167	21	.875	.748	.757	-.279	0.08	-.298	0.08
172.00	0.0	-	-	C	-	.196	97	.869	.665	1.034	.231	0.74	.248	0.72
173.00	0.0	-	-	-	-	.260	190	.863	.642	.700	.505	0.53	.543	0.51
174.00	0.0	-	-	M	X	.315	469	.895	.639	.791	.503	1.94	.532	1.94
175.00	45.5	-	-	M	X	.296	187	.935	.639	.791	.465	0.80	.481	0.84
176.01	0.0	C	-	-	-	.515	23	.915	.642	.791	.746	0.12	.780	0.12
176.02	0.0	C	-	-	-	.563	16	.915	.642	.791	.796	0.08	.832	0.09
177.01	33.3	M	-	-	-	.221	27	.935	.642	.791	.354	0.11	.366	0.12
177.02	66.7	M	-	-	-	.100	27	.935	.642	.791	.164	0.11	.169	0.11
179.01	0.0	C	-	-	-	.027	12	.947	.642	.907	.038	0.07	.039	0.07
179.02	0.0	C	-	-	-	.711	15	.947	.642	.907	.906	0.10	.931	0.10
179.03	0.0	C	-	-	-	.226	9	.947	.642	.907	.312	0.05	.320	0.05
179.04	0.0	C	-	-	-	.336	10	.947	.642	.907	.458	0.06	.471	0.06
179.05	0.0	C	-	-	-	.793	7	.947	.642	.907	.991	0.05	1.019	0.05
179.08	0.0	M	-	-	-	.477	51	.895	.642	.791	.708	0.24	.749	0.24
179.09	0.0	M	-	-	-	.363	64	.895	.642	.791	.568	0.28	.600	0.28
179.10	0.0	M	-	-	-	.480	23	.895	.642	.791	.712	0.11	.752	0.11
179.11	0.0	M	-	-	-	.273	26	.895	.642	.791	.442	0.11	.467	0.11
184.02	0.0	-	-	-	-	.040	26	.935	.642	.721	.074	0.08	.076	0.08
189.02	0.0	C	-	-	-	.338	403	.895	.639	.791	.535	1.70	.566	1.70
189.03	0.0	C	-	-	-	.437	282	.895	.639	.791	.663	1.30	.700	1.30
189.04	0.0	C	-	-	-	.377	87	.895	.639	.791	.587	0.38	.621	0.38
189.05	0.0	C	-	-	-	.518	68	.895	.639	.791	.755	0.34	.798	0.34
189.06	0.0	C	-	-	-	.501	731	.895	.639	.791	.736	3.58	.778	3.58
189.07	0.0	C	-	-	-	.437	198	.895	.639	.791	.663	0.91	.700	0.91
189.08	0.0	C	-	-	-	.372	112	.895	.639	.791	.581	0.49	.614	0.49
189.09	0.0	C	-	-	-	.398	351	.895	.639	.791	.614	1.56	.649	1.56
189.10	0.0	C	-	-	-	.317	394	.895	.639	.791	.506	1.64	.535	1.64
189.11	0.0	C	-	-	-	.218	81	.895	.639	.791	.359	0.31	.380	0.32
189.12	0.0	C	-	-	-	.368	191	.895	.639	.791	.575	0.83	.608	0.83
189.13	0.0	C	-	-	-	.368	97	.895	.639	.791	.575	0.42	.608	0.42
189.14	0.0	C	-	-	-	.616	617	.895	.639	.791	.852	3.41	.900	3.41
189.15	0.0	C	-	-	-	.475	418	.895	.639	.791	.707	1.99	.748	1.99
189.16	0.0	C	-	-	-	.428	412	.895	.639	.791	.652	1.88	.689	1.88
189.17	0.0	C	-	-	-	.394	243	.895	.639	.791	.609	1.07	.644	1.07
189.18	0.0	C	-	-	-	.330	98	.895	.639	.791	.524	0.41	.554	0.41

Tabelle E.15
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.19	0.0	C	-	-	-	.561	287	.895	.639	.791	.799	1.49	.845	1.50
189.20	0.0	C	-	-	-	.411	291	.895	.639	.791	.631	1.31	.667	1.31
189.21	0.0	C	-	-	-	.278	184	.895	.639	.791	.450	0.74	.476	0.74
189.22	0.0	C	-	-	-	.385	84	.895	.639	.791	.598	0.37	.632	0.37
189.23	0.0	C	-	-	-	.214	158	.895	.639	.791	.353	0.61	.373	0.61
189.24	0.0	C	-	-	-	.300	119	.895	.639	.791	.482	0.49	.509	0.49
189.25	0.0	C	-	-	-	.381	67	.895	.639	.791	.593	0.29	.626	0.29
189.26	0.0	C	-	-	-	.407	57	.895	.639	.791	.626	0.25	.661	0.25
189.27	0.0	C	-	-	-	.377	81	.895	.639	.776	.598	0.34	.632	0.34
189.28	0.0	C	-	-	-	.479	181	.895	.639	.791	.712	0.87	.752	0.87
189.29	0.0	C	-	-	-	.347	314	.895	.639	.791	.547	1.33	.579	1.33
189.30	0.0	C	-	-	-	.407	454	.895	.639	.791	.626	2.03	.661	2.03
189.31	0.0	C	-	-	-	.501	380	.895	.639	.791	.736	1.86	.778	1.86
189.32	0.0	C	-	-	-	.381	412	.895	.639	.791	.593	1.80	.626	1.80
189.33	0.0	C	-	-	-	.488	391	.895	.639	.791	.722	1.89	.763	1.89
189.34	0.0	C	-	-	-	.321	48	.895	.639	.791	.512	0.20	.541	0.20
189.38	0.0	C	-	-	-	.334	56	.895	.639	.791	.530	0.24	.560	0.24
189.39	0.0	C	-	-	-	.248	48	.895	.639	.791	.405	0.19	.428	0.19
189.40	0.0	C	-	-	-	.240	61	.895	.639	.791	.393	0.24	.416	0.24
189.41	0.0	C	-	-	-	.214	104	.895	.639	.791	.353	0.40	.373	0.40
189.42	0.0	C	-	-	-	.227	113	.895	.639	.791	.373	0.44	.395	0.44
189.43	0.0	C	-	-	-	.338	38	.895	.639	.791	.535	0.16	.566	0.16
189.44	0.0	C	-	-	-	.295	21	.895	.639	.791	.475	0.09	.502	0.09
189.51	0.0	C	-	-	-	.257	104	.895	.639	.791	.419	0.41	.443	0.41
189.53	0.0	C	-	-	-	.574	268	.895	.639	.791	.812	1.42	.858	1.42
189.54	0.0	C	-	-	-	.432	212	.895	.639	.791	.657	0.97	.694	0.97
189.55	0.0	C	-	-	-	.488	98	.895	.639	.791	.722	0.47	.763	0.47
189.56	0.0	C	-	-	-	.484	87	.895	.639	.791	.717	0.42	.758	0.42
189.57	0.0	C	-	-	-	.424	106	.895	.639	.791	.647	0.48	.684	0.48
189.58	0.0	C	-	-	-	.287	124	.895	.639	.791	.463	0.50	.489	0.50
189.59	0.0	C	-	-	-	.428	168	.895	.639	.791	.652	0.77	.689	0.77
189.60	0.0	C	-	-	-	.372	117	.895	.639	.791	.581	0.51	.614	0.51
189.61	0.0	C	-	-	-	.407	198	.895	.639	.791	.626	0.89	.661	0.89
189.62	0.0	C	-	-	-	.390	314	.895	.639	.791	.604	1.38	.639	1.38
189.63	0.0	C	-	-	-	.420	217	.895	.639	.791	.642	0.98	.679	0.98
189.64	0.0	C	-	-	-	.462	208	.895	.639	.791	.692	0.98	.732	0.98
189.65	0.0	C	-	-	-	.458	197	.895	.639	.791	.688	0.92	.727	0.92
189.66	0.0	C	-	-	-	.458	128	.895	.639	.791	.688	0.60	.727	0.60
189.67	0.0	C	-	-	-	.471	117	.895	.639	.791	.703	0.56	.743	0.56
189.68	0.0	C	-	-	-	.462	89	.895	.639	.791	.692	0.42	.732	0.42
189.69	0.0	C	-	-	-	.492	58	.895	.639	.791	.726	0.28	.768	0.28

Anmerkungen. siehe Seite 275.

Tabelle E.16
Lernleistung, Skill Level 3

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
101.01	0.0	–	–	C	–	.147	72	.863	.639	.467	.667	0.18	.718	0.17
102.00	54.5	C	–	–	–	.219	54	.918	.642	.705	.405	0.79	.423	0.82
103.00	0.0	C	X	M	–	.152	90	.914	.647	.752	.266	1.48	.278	1.51
104.00	0.0	M	X	–	–	.240	158	.911	.642	.752	.411	2.70	.431	2.76
105.00	88.9	M	X	M	X	.086	103	.911	.639	.752	.154	1.62	.161	1.65
107.01	86.4	M	–	M	X	.111	56	.911	.647	.752	.196	0.90	.206	0.92
107.02	94.4	M	–	M	X	.062	137	.911	.647	.752	.110	2.17	.116	2.21
107.03	88.9	M	–	M	X	.073	145	.911	.647	.752	.130	2.30	.136	2.35
108.00	90.9	M	–	M	–	.050	249	.911	.642	.752	.090	3.90	.094	3.98
110.00	0.0	M	X	–	–	.476	75	.911	.642	.752	.727	1.62	.761	1.65
116.03	0.0	C	–	–	–	.556	70	.882	.642	.752	.815	1.63	.868	1.62
116.08	0.0	C	–	–	–	.534	304	.882	.642	.752	.795	6.89	.846	6.82
116.09	0.0	C	–	–	–	.446	169	.882	.642	.752	.702	3.42	.748	3.39
119.01	0.0	–	–	C	–	.210	100	.863	.680	.752	.370	1.62	.398	1.57
120.01	0.0	–	–	–	–	.530	127	.863	.642	.752	.796	2.83	.856	2.74
120.02	0.0	–	–	–	–	.600	101	.863	.642	.752	.857	2.48	.923	2.40
120.03	0.0	–	–	–	–	.570	222	.863	.642	.844	.786	5.86	.846	5.68
120.04	0.0	–	–	–	–	.590	13	.863	.639	.752	.850	0.31	.915	0.30
125.02	0.0	–	–	C	–	.262	221	.863	.642	.778	.444	3.87	.478	3.75
126.02	0.0	–	–	C	–	.283	37	.911	.642	.737	.487	0.63	.510	0.64
129.00	50	M	X	–	–	.083	93	.911	.642	.752	.148	1.47	.155	1.50
130.00	0.0	M	X	–	–	.305	48	.911	.639	.752	.510	0.86	.534	0.88
133.00	0.0	–	–	–	–	.330	112	.925	.642	.785	.518	2.28	.539	2.37
137.01	0.0	–	–	–	–	.200	74	.863	.642	.752	.363	1.13	.391	1.09
142.00	0.0	–	–	–	–	.750	8	.863	.642	.752	.962	0.24	1.035	0.24
146.00	75.9	M	–	–	–	.121	45	.869	.642	.791	.208	0.77	.223	0.75
149.00	0.0	M	–	C	–	.093	199	.895	.642	.752	.168	3.05	.178	3.06
156.00	0.0	–	–	–	–	.310	122	.911	.642	.970	.399	3.70	.418	3.78
160.01	0.0	C	–	–	–	.351	58	.861	.642	.926	.475	1.59	.512	1.53
160.02	0.0	C	–	–	–	.160	91	.855	.642	.867	.242	2.00	.261	1.92
161.01	0.0	M	–	–	–	.016	60	.895	.642	.873	.024	1.37	.025	1.37
161.02	0.0	M	–	–	–	.086	31	.895	.642	.752	.156	0.47	.165	0.48
178.00	0.0	C	–	M	X	.211	176	.911	.647	.752	.364	2.97	.381	3.04
179.06	0.0	C	–	–	–	.141	15	.947	.642	.907	.196	0.39	.201	0.42
179.07	0.0	C	–	–	–	.396	14	.947	.642	.907	.536	0.38	.550	0.41
179.12	0.0	M	–	–	–	.530	20	.895	.642	.752	.788	0.45	.832	0.46
179.13	0.0	M	–	–	–	.393	52	.895	.642	.752	.634	1.00	.670	1.01
179.14	0.0	M	–	–	–	.443	19	.895	.642	.752	.694	0.39	.734	0.39
183.00	0.0	–	–	–	–	.070	256	.895	.642	.752	.127	3.90	.134	3.91
189.35	0.0	C	–	–	–	.390	58	.895	.639	.752	.631	1.11	.667	1.12
189.36	0.0	C	–	–	–	.424	57	.895	.639	.752	.673	1.14	.712	1.14
189.37	0.0	C	–	–	–	.223	87	.895	.639	.752	.390	1.42	.413	1.43
189.45	0.0	C	–	–	–	.390	178	.895	.639	.752	.631	3.41	.667	3.42
189.46	0.0	C	–	–	–	.432	347	.895	.639	.752	.683	6.98	.722	7.00
189.47	0.0	C	–	–	–	.454	37	.895	.639	.752	.708	0.76	.749	0.77
189.48	0.0	C	–	–	–	.424	41	.895	.639	.752	.673	0.82	.712	0.82
189.49	0.0	C	–	–	–	.595	104	.895	.639	.752	.849	2.56	.897	2.57
189.50	0.0	C	–	–	–	.454	174	.895	.639	.752	.708	3.59	.749	3.60
189.52	0.0	C	–	–	–	.509	32	.895	.639	.752	.768	0.71	.811	0.71
189.70	0.0	C	–	–	–	.492	87	.895	.639	.752	.750	1.88	.793	1.89

Anmerkungen. siehe Seite 275.

Tabelle E.17

Lernleistung, maximal Hauptschul-Abschluss

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
106.00	M	–	M	X	.091	224	.895	.647	.791	.153	28.1	.162	27.9
134.00	–	–	–	–	.720	27	.911	.642	.890	.920	5.9	.964	5.9
167.00	C	X	C	–	.078	139	.895	.639	.791	.132	17.2	.140	17.0
173.00	–	–	–	–	.260	190	.863	.642	.700	.505	17.9	.543	17.1
184.01	–	–	–	–	.220	152	.935	.642	.721	.390	17.2	.403	17.8
184.02	–	–	–	–	.040	26	.935	.642	.721	.074	2.7	.076	2.8
184.03	–	–	–	–	.310	89	.920	.642	.741	.520	11.2	.542	11.4

Anmerkungen. siehe Seite 275.

Tabelle E.18

Lernleistung, (Fach-)Hochschulreife

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
100.00	60.0	–	–	C	–	.276	76	.911	.820	1.003	.304	8.4	.318	8.5
103.00	0.0	C	X	M	–	.152	90	.914	.647	.752	.266	3.9	.278	4.0
105.00	88.9	M	X	M	X	.088	103	.911	.639	.752	.157	4.3	.165	4.4
107.01	86.4	M	–	M	X	.113	56	.911	.647	.752	.200	2.4	.209	2.4
107.02	94.4	M	–	M	X	.063	137	.911	.647	.752	.112	5.8	.118	5.8
107.03	88.9	M	–	M	X	.074	145	.911	.647	.752	.132	6.1	.138	6.2
108.00	90.9	M	–	M	–	.051	249	.911	.642	.752	.091	10.3	.096	10.5
136.00	0.0	–	–	–	–	.360	97	.869	.642	.776	.583	4.9	.625	4.8
146.00	75.9	M	–	–	–	.122	45	.869	.642	.791	.209	2.0	.224	2.0
148.00	0.0	C	–	–	–	.521	91	.869	.642	.700	.817	4.9	.876	4.8
149.00	0.0	M	–	C	–	.093	199	.895	.642	.752	.168	8.1	.178	8.1
156.00	0.0	–	–	–	–	.310	122	.911	.642	.970	.399	9.8	.418	9.9
160.01	0.0	C	–	–	–	.351	58	.861	.642	.926	.475	4.2	.512	4.0
160.02	0.0	C	–	–	–	.160	91	.855	.642	.867	.242	5.3	.261	5.1
177.02	66.7	M	–	–	–	.101	27	.935	.642	.791	.165	1.3	.171	1.4
178.00	0.0	C	–	M	X	.211	176	.911	.647	.752	.364	7.9	.381	8.0
183.00	0.0	–	–	–	–	.070	256	.895	.642	.752	.127	10.3	.134	10.3

Anmerkungen. siehe Seite 275.

Tabelle E.19

Lernleistung, Militär (HG 0)

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
150.00	–	–	–	–	.322	167	.895	.642	.776	.522	14.2	.552	14.4
181.01	–	–	–	–	.120	280	.869	.642	.776	.212	20.0	.227	19.8
181.02	–	–	–	–	.250	303	.869	.642	.776	.425	23.4	.456	23.1
181.03	–	–	–	–	.300	70	.869	.642	.776	.499	5.6	.536	5.6
181.04	–	–	–	–	.100	110	.869	.642	.776	.177	7.8	.190	7.7
182.00	–	–	–	–	.090	200	.895	.642	.776	.156	14.8	.165	15.1
186.00	–	–	–	–	.320	167	.895	.642	.776	.519	14.1	.549	14.4

Anmerkungen. siehe Seite 275.

Tabelle E.20
Lernleistung, Techniker und gleichrangige (HG 3)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
101.01	0.0	-	-	C	-	.147	72	.863	.639	.467	.667	0.18	.718	0.17
102.00	54.5	C	-	-	-	.219	54	.918	.642	.705	.405	0.79	.423	0.82
103.00	0.0	C	X	M	-	.152	90	.914	.647	.752	.266	1.48	.278	1.51
104.00	0.0	M	X	-	-	.240	158	.911	.642	.752	.411	2.70	.431	2.76
105.00	88.9	M	X	M	X	.086	103	.911	.639	.752	.154	1.62	.161	1.65
107.01	86.4	M	-	M	X	.111	56	.911	.647	.752	.196	0.90	.206	0.92
107.02	94.4	M	-	M	X	.062	137	.911	.647	.752	.110	2.17	.116	2.21
107.03	88.9	M	-	M	X	.073	145	.911	.647	.752	.130	2.30	.136	2.35
108.00	90.9	M	-	M	-	.050	249	.911	.642	.752	.090	3.90	.094	3.98
110.00	0.0	M	X	-	-	.476	75	.911	.642	.752	.727	1.62	.761	1.65
116.03	0.0	C	-	-	-	.556	70	.882	.642	.752	.815	1.63	.868	1.62
116.08	0.0	C	-	-	-	.534	304	.882	.642	.752	.795	6.89	.846	6.82
116.09	0.0	C	-	-	-	.446	169	.882	.642	.752	.702	3.42	.748	3.39
119.01	0.0	-	-	C	-	.210	100	.863	.680	.752	.370	1.62	.398	1.57
120.01	0.0	-	-	-	-	.530	127	.863	.642	.752	.796	2.83	.856	2.74
120.02	0.0	-	-	-	-	.600	101	.863	.642	.752	.857	2.48	.923	2.40
120.03	0.0	-	-	-	-	.570	222	.863	.642	.844	.786	5.86	.846	5.68
120.04	0.0	-	-	-	-	.590	13	.863	.639	.752	.850	0.31	.915	0.30
125.02	0.0	-	-	C	-	.262	221	.863	.642	.778	.444	3.87	.478	3.75
126.02	0.0	-	-	C	-	.283	37	.911	.642	.737	.487	0.63	.510	0.64
129.00	50.0	M	X	-	-	.083	93	.911	.642	.752	.148	1.47	.155	1.50
130.00	0.0	M	X	-	-	.305	48	.911	.639	.752	.510	0.86	.534	0.88
133.00	0.0	-	-	-	-	.330	112	.925	.642	.785	.518	2.28	.539	2.37
137.01	0.0	-	-	-	-	.200	74	.863	.642	.752	.363	1.13	.391	1.09
142.00	0.0	-	-	-	-	.750	8	.863	.642	.752	.962	0.24	1.035	0.24
146.00	75.9	M	-	-	-	.121	45	.869	.642	.791	.208	0.77	.223	0.75
149.00	0.0	M	-	C	-	.093	199	.895	.642	.752	.168	3.05	.178	3.06
156.00	0.0	-	-	-	-	.310	122	.911	.642	.970	.399	3.70	.418	3.78
160.01	0.0	C	-	-	-	.351	58	.861	.642	.926	.475	1.59	.512	1.53
160.02	0.0	C	-	-	-	.160	91	.855	.642	.867	.242	2.00	.261	1.92
161.01	0.0	M	-	-	-	.016	60	.895	.642	.873	.024	1.37	.025	1.37
161.02	0.0	M	-	-	-	.086	31	.895	.642	.752	.156	0.47	.165	0.48
178.00	0.0	C	-	M	X	.211	176	.911	.647	.752	.364	2.97	.381	3.04
179.06	0.0	C	-	-	-	.141	15	.947	.642	.907	.196	0.39	.201	0.42
179.07	0.0	C	-	-	-	.396	14	.947	.642	.907	.536	0.38	.550	0.41
179.12	0.0	M	-	-	-	.530	20	.895	.642	.752	.788	0.45	.832	0.46
179.13	0.0	M	-	-	-	.393	52	.895	.642	.752	.634	1.00	.670	1.01
179.14	0.0	M	-	-	-	.443	19	.895	.642	.752	.694	0.39	.734	0.39
183.00	0.0	-	-	-	-	.070	256	.895	.642	.752	.127	3.90	.134	3.91
189.35	0.0	C	-	-	-	.390	58	.895	.639	.752	.631	1.11	.667	1.12
189.36	0.0	C	-	-	-	.424	57	.895	.639	.752	.673	1.14	.712	1.14
189.37	0.0	C	-	-	-	.223	87	.895	.639	.752	.390	1.42	.413	1.43
189.45	0.0	C	-	-	-	.390	178	.895	.639	.752	.631	3.41	.667	3.42
189.46	0.0	C	-	-	-	.432	347	.895	.639	.752	.683	6.98	.722	7.00
189.47	0.0	C	-	-	-	.454	37	.895	.639	.752	.708	0.76	.749	0.77
189.48	0.0	C	-	-	-	.424	41	.895	.639	.752	.673	0.82	.712	0.82
189.49	0.0	C	-	-	-	.595	104	.895	.639	.752	.849	2.56	.897	2.57
189.50	0.0	C	-	-	-	.454	174	.895	.639	.752	.708	3.59	.749	3.60
189.52	0.0	C	-	-	-	.509	32	.895	.639	.752	.768	0.71	.811	0.71
189.70	0.0	C	-	-	-	.492	87	.895	.639	.752	.750	1.88	.793	1.89

Anmerkungen. siehe Seite 275.

Tabelle E.21

Lernleistung, Bürokräfte (HG 4)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
116.11	0.0	C	-	-	-	.749	76	.882	.642	.791	.959	2.0	1.021	1.9
125.03	0.0	-	-	-	-	.160	196	.863	.642	.644	.377	1.5	.405	1.4
125.04	0.0	-	-	C	-	.198	198	.863	.642	.711	.390	2.2	.420	2.1
126.03	0.0	-	-	C	-	.386	64	.911	.642	.625	.717	0.8	.752	0.8
126.05	0.0	-	-	C	-	.153	59	.911	.642	.673	.311	0.6	.326	0.6
126.06	0.0	-	-	C	-	.259	83	.911	.642	.609	.557	0.8	.583	0.8
134.00	0.0	-	-	-	-	.720	27	.911	.642	.890	.920	0.7	.964	0.7
135.00	0.0	-	-	M	-	.470	861	.895	.647	.791	.698	16.4	.738	16.4
137.03	0.0	-	-	-	-	.480	44	.863	.642	.791	.720	0.8	.775	0.8
137.04	0.0	-	-	-	-	.190	59	.863	.642	.791	.322	0.9	.347	0.8
138.00	86.1	C	-	M	-	.213	22	.895	.520	.791	.387	0.3	.409	0.3
151.00	0.0	C	-	M	-	.200	200	.935	.647	.791	.321	3.3	.332	3.4
172.00	0.0	-	-	C	-	.196	97	.869	.665	1.034	.231	3.0	.248	2.9
174.00	0.0	-	-	M	X	.315	469	.895	.639	.791	.503	7.7	.532	7.7
177.02	66.7	M	-	-	-	.101	27	.935	.642	.791	.165	0.4	.171	0.4
179.01	0.0	C	-	-	-	.027	12	.947	.642	.907	.038	0.3	.039	0.3
179.02	0.0	C	-	-	-	.711	15	.947	.642	.907	.906	0.4	.931	0.4
179.03	0.0	C	-	-	-	.226	9	.947	.642	.907	.312	0.2	.320	0.2
179.08	0.0	M	-	-	-	.477	51	.895	.642	.791	.708	1.0	.749	1.0
179.09	0.0	M	-	-	-	.363	64	.895	.642	.791	.568	1.1	.600	1.1
179.10	0.0	M	-	-	-	.480	23	.895	.642	.791	.712	0.4	.752	0.4
189.02	0.0	C	-	-	-	.338	403	.895	.639	.791	.535	6.8	.566	6.8
189.03	0.0	C	-	-	-	.437	282	.895	.639	.791	.663	5.2	.700	5.2
189.04	0.0	C	-	-	-	.377	87	.895	.639	.791	.587	1.5	.621	1.5
189.05	0.0	C	-	-	-	.518	68	.895	.639	.791	.755	1.4	.798	1.4
189.06	0.0	C	-	-	-	.501	731	.895	.639	.791	.736	14.3	.778	14.3
189.07	0.0	C	-	-	-	.437	198	.895	.639	.791	.663	3.6	.700	3.6
189.08	0.0	C	-	-	-	.372	112	.895	.639	.791	.581	1.9	.614	1.9
189.09	0.0	C	-	-	-	.398	351	.895	.639	.791	.614	6.2	.649	6.2
189.10	0.0	C	-	-	-	.317	394	.895	.639	.791	.506	6.5	.535	6.5
189.11	0.0	C	-	-	-	.218	81	.895	.639	.791	.359	1.3	.380	1.3
189.12	0.0	C	-	-	-	.368	191	.895	.639	.791	.575	3.3	.608	3.3
189.13	0.0	C	-	-	-	.368	97	.895	.639	.791	.575	1.7	.608	1.7
189.34	0.0	C	-	-	-	.321	48	.895	.639	.791	.512	0.8	.541	0.8
189.69	0.0	C	-	-	-	.492	58	.895	.639	.791	.726	1.1	.768	1.1

Anmerkungen. siehe Seite 275.

Tabelle E.22

Lernleistung, Handwerker (HG 7)

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
106.00	M	-	M	X	.091	224	.895	.647	.791	.153	1.45	.162	1.45
112.00	-	-	-	-	.240	93	.843	.642	.791	.407	0.59	.443	0.56
114.00	-	-	M	-	.335	313	.911	.647	.752	.549	2.13	.576	2.17
117.00	-	-	-	-	.700	42	.863	.642	.791	.923	0.44	.993	0.43
118.02	-	-	C	-	.570	30	.935	.642	.791	.800	0.28	.827	0.29
118.03	-	-	C	-	.528	54	.935	.642	.791	.756	0.48	.781	0.50
123.01	-	-	-	-	.710	20	.895	.520	.791	.988	0.19	1.044	0.19
123.02	-	-	-	-	.650	20	.895	.520	.791	.939	0.18	.993	0.18
123.03	-	-	-	-	.800	19	.895	.520	.791	1.050	0.20	1.110	0.20
123.04	-	-	-	-	.720	7	.895	.520	.791	.995	0.07	1.052	0.07
123.05	-	-	-	-	.590	13	.895	.520	.791	.885	0.11	.936	0.11
123.06	-	-	-	-	.850	10	.895	.520	.791	1.080	0.11	1.142	0.11

Tabelle E.22
(Fortsetzung)

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
124.01	M	-	-	-	.803	12	.895	.520	.791	1.052	0.13	1.112	0.13
124.02	M	-	-	-	.806	8	.895	.520	.791	1.054	0.09	1.114	0.09
124.03	M	-	-	-	.515	12	.895	.520	.791	.809	0.09	.855	0.09
125.05	-	-	C	-	.233	285	.863	.642	.900	.331	2.58	.356	2.50
125.06	-	-	C	-	.205	120	.863	.642	.722	.394	0.60	.424	0.57
126.04	-	-	C	-	.378	81	.911	.642	.641	.694	0.44	.727	0.45
127.01	M	X	C	-	.211	33	.895	.642	.791	.348	0.22	.368	0.22
127.02	M	X	C	-	.337	30	.895	.642	.791	.533	0.22	.563	0.22
137.05	-	-	-	-	.100	27	.863	.642	.791	.173	0.16	.186	0.16
137.06	-	-	-	-	.280	49	.863	.642	.791	.461	0.33	.497	0.32
140.00	C	-	-	-	.469	146	.940	.642	.791	.688	1.24	.710	1.31
141.00	-	-	-	-	.750	34	.863	.642	.791	.959	0.38	1.032	0.37
143.01	-	-	-	-	.322	123	.863	.642	.791	.522	0.86	.561	0.83
143.02	-	-	-	-	.644	147	.863	.642	.791	.878	1.44	.946	1.40
143.03	-	-	-	-	.644	62	.863	.642	.791	.878	0.61	.946	0.59
144.00	-	-	-	-	.560	315	.935	.642	.791	.789	2.90	.816	3.03
153.00	-	-	-	-	.570	153	.870	.639	.737	.839	1.29	.900	1.26
164.01	-	-	-	-	.245	486	.895	.642	.791	.400	3.34	.423	3.34
164.02	-	-	-	-	.350	198	.895	.642	.791	.550	1.46	.582	1.47
167.00	C	X	C	-	.078	139	.895	.639	.791	.132	0.89	.140	0.89
168.00	-	-	M	-	.476	51	.895	.639	.791	.708	0.42	.749	0.42
171.01	-	-	C	-	.390	197	.875	.740	.950	.478	2.40	.511	2.35
171.02	-	-	C	-	.332	142	.875	.720	1.018	.384	1.94	.410	1.90
171.03	-	-	C	-	.277	141	.875	.711	1.084	.300	2.20	.320	2.16
171.04	-	-	C	-	-.167	21	.875	.748	.757	-.279	0.14	-.298	0.13
176.01	C	-	-	-	.515	23	.915	.642	.791	.746	0.20	.780	0.21
176.02	C	-	-	-	.563	16	.915	.642	.791	.796	0.15	.832	0.15
179.04	C	-	-	-	.336	10	.947	.642	.907	.458	0.10	.471	0.10
179.05	C	-	-	-	.793	7	.947	.642	.907	.991	0.08	1.019	0.09
179.11	M	-	-	-	.273	26	.895	.642	.791	.442	0.18	.467	0.18
184.02	-	-	-	-	.040	26	.935	.642	.721	.074	0.14	.076	0.15
189.14	C	-	-	-	.616	617	.895	.639	.791	.852	5.90	.900	5.91
189.15	C	-	-	-	.475	418	.895	.639	.791	.707	3.45	.748	3.45
189.16	C	-	-	-	.428	412	.895	.639	.791	.652	3.25	.689	3.26
189.17	C	-	-	-	.394	243	.895	.639	.791	.609	1.86	.644	1.86
189.18	C	-	-	-	.330	98	.895	.639	.791	.524	0.71	.554	0.71
189.19	C	-	-	-	.561	287	.895	.639	.791	.799	2.59	.845	2.59
189.20	C	-	-	-	.411	291	.895	.639	.791	.631	2.26	.667	2.26
189.21	C	-	-	-	.278	184	.895	.639	.791	.450	1.28	.476	1.29
189.22	C	-	-	-	.385	84	.895	.639	.791	.598	0.64	.632	0.64
189.23	C	-	-	-	.214	158	.895	.639	.791	.353	1.06	.373	1.06
189.24	C	-	-	-	.300	119	.895	.639	.791	.482	0.84	.509	0.85
189.25	C	-	-	-	.381	67	.895	.639	.791	.593	0.51	.626	0.51
189.26	C	-	-	-	.407	57	.895	.639	.791	.626	0.44	.661	0.44
189.27	C	-	-	-	.377	81	.895	.639	.776	.598	0.59	.632	0.59
189.28	C	-	-	-	.479	181	.895	.639	.791	.712	1.50	.752	1.50
189.29	C	-	-	-	.347	314	.895	.639	.791	.547	2.31	.579	2.31
189.30	C	-	-	-	.407	454	.895	.639	.791	.626	3.51	.661	3.52
189.31	C	-	-	-	.501	380	.895	.639	.791	.736	3.22	.778	3.22
189.32	C	-	-	-	.381	412	.895	.639	.791	.593	3.12	.626	3.12
189.33	C	-	-	-	.488	391	.895	.639	.791	.722	3.27	.763	3.27
189.38	C	-	-	-	.334	56	.895	.639	.791	.530	0.41	.560	0.41
189.39	C	-	-	-	.248	48	.895	.639	.791	.405	0.33	.428	0.33
189.40	C	-	-	-	.240	61	.895	.639	.791	.393	0.42	.416	0.42
189.41	C	-	-	-	.214	104	.895	.639	.791	.353	0.70	.373	0.70
189.42	C	-	-	-	.227	113	.895	.639	.791	.373	0.76	.395	0.77
189.43	C	-	-	-	.338	38	.895	.639	.791	.535	0.28	.566	0.28
189.44	C	-	-	-	.295	21	.895	.639	.791	.475	0.15	.502	0.15
189.53	C	-	-	-	.574	268	.895	.639	.791	.812	2.45	.858	2.45

Tabelle E.22
(Fortsetzung)

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.54	C	–	–	–	.432	212	.895	.639	.791	.657	1.68	.694	1.68
189.55	C	–	–	–	.488	98	.895	.639	.791	.722	0.82	.763	0.82
189.56	C	–	–	–	.484	87	.895	.639	.791	.717	0.72	.758	0.73
189.57	C	–	–	–	.424	106	.895	.639	.791	.647	0.83	.684	0.83
189.58	C	–	–	–	.287	124	.895	.639	.791	.463	0.87	.489	0.87
189.59	C	–	–	–	.428	168	.895	.639	.791	.652	1.32	.689	1.33
189.60	C	–	–	–	.372	117	.895	.639	.791	.581	0.88	.614	0.88
189.61	C	–	–	–	.407	198	.895	.639	.791	.626	1.53	.661	1.53
189.62	C	–	–	–	.390	314	.895	.639	.791	.604	2.39	.639	2.40
189.63	C	–	–	–	.420	217	.895	.639	.791	.642	1.70	.679	1.70
189.64	C	–	–	–	.462	208	.895	.639	.791	.692	1.69	.732	1.70
189.65	C	–	–	–	.458	197	.895	.639	.791	.688	1.60	.727	1.60
189.66	C	–	–	–	.458	128	.895	.639	.791	.688	1.04	.727	1.04
189.67	C	–	–	–	.471	117	.895	.639	.791	.703	0.96	.743	0.96
189.68	C	–	–	–	.462	89	.895	.639	.791	.692	0.72	.732	0.73

Anmerkungen. siehe Seite 275.

Tabelle E.23
Lernleistung, Maschinenbediener (HG 8)

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
113.00	–	–	M	X	.236	148	.863	.639	.750	.425	5.5	.457	5.4
116.06	C	–	–	–	.404	230	.882	.642	.791	.625	11.6	.665	11.6
116.10	C	–	–	–	.547	407	.882	.642	.791	.786	23.6	.837	23.8
119.02	–	–	C	–	.216	99	.863	.680	.791	.354	4.4	.381	4.4
125.01	–	–	C	–	.293	738	.863	.642	.811	.466	35.0	.502	34.4
126.01	–	–	C	–	.337	322	.911	.642	.689	.601	12.1	.630	12.6
137.02	–	–	–	–	.400	64	.863	.642	.791	.625	3.1	.673	3.1
189.51	C	–	–	–	.257	104	.895	.639	.791	.419	4.7	.443	4.8

Anmerkungen. siehe Seite 275.

Tabelle E.24
Lernleistung, IST

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
101.01	–	–	C	–	.221	72	.863	.639	.467	.796	0.19	.857	0.19
101.02	–	–	C	–	.296	65	.863	.784	.489	.804	0.30	.866	0.30
109.00	C	–	C	–	.042	102	.882	.811	.752	.069	1.29	.074	1.31
111.01	–	–	–	–	.430	78	.863	.642	.791	.662	1.12	.713	1.11
111.02	–	–	–	–	.450	98	.863	.642	.791	.686	1.44	.738	1.43
113.00	–	–	M	X	.236	148	.863	.639	.750	.425	1.56	.457	1.55
116.01	C	–	–	–	.355	126	.882	.642	.776	.572	1.66	.609	1.68
116.03	C	–	–	–	.556	70	.882	.642	.752	.815	1.11	.868	1.13
116.04	C	–	–	–	.326	215	.882	.642	.776	.532	2.75	.566	2.80
116.06	C	–	–	–	.404	230	.882	.642	.791	.625	3.28	.665	3.33
116.08	C	–	–	–	.534	304	.882	.642	.752	.795	4.68	.846	4.75

Tabelle E.24
(Fortsetzung)

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
116.09	C	–	–	–	.446	169	.882	.642	.752	.702	2.32	.748	2.36
116.10	C	–	–	–	.547	407	.882	.642	.791	.786	6.71	.837	6.82
116.11	C	–	–	–	.749	76	.882	.642	.791	.959	1.58	1.021	1.60
116.13	–	–	–	–	.760	180	.863	.642	.776	.967	3.79	1.040	3.77
117.00	–	–	–	–	.700	42	.863	.642	.791	.923	0.82	.993	0.82
119.01	–	–	C	–	.210	100	.863	.680	.752	.370	1.10	.398	1.09
119.02	–	–	C	–	.216	99	.863	.680	.791	.354	1.26	.381	1.25
120.01	–	–	–	–	.530	127	.863	.642	.752	.796	1.92	.856	1.91
120.02	–	–	–	–	.600	101	.863	.642	.752	.857	1.69	.923	1.68
120.03	–	–	–	–	.570	222	.863	.642	.844	.786	3.98	.846	3.95
120.04	–	–	–	–	.590	13	.863	.639	.752	.850	0.21	.915	0.21
125.01	–	–	C	–	.293	738	.863	.642	.811	.466	9.93	.502	9.87
125.02	–	–	C	–	.262	221	.863	.642	.778	.444	2.63	.478	2.61
125.03	–	–	–	–	.160	196	.863	.642	.644	.377	1.21	.405	1.20
125.04	–	–	C	–	.198	198	.863	.642	.711	.390	1.74	.420	1.73
125.05	–	–	C	–	.233	285	.863	.642	.900	.331	4.82	.356	4.79
125.06	–	–	C	–	.205	120	.863	.642	.722	.394	1.11	.424	1.10
137.01	–	–	–	–	.200	74	.863	.642	.752	.363	0.76	.391	0.76
137.02	–	–	–	–	.400	64	.863	.642	.791	.625	0.89	.673	0.89
137.03	–	–	–	–	.480	44	.863	.642	.791	.720	0.67	.775	0.66
137.04	–	–	–	–	.190	59	.863	.642	.791	.322	0.70	.347	0.69
137.05	–	–	–	–	.100	27	.863	.642	.791	.173	0.31	.186	0.31
137.06	–	–	–	–	.280	49	.863	.642	.791	.461	0.62	.497	0.61
137.07	–	–	–	–	.120	50	.863	.642	.791	.207	0.57	.223	0.57
140.00	–	–	–	–	.550	146	.863	.642	.791	.793	2.39	.854	2.38
141.00	–	–	–	–	.750	34	.863	.642	.791	.959	0.71	1.032	0.70
142.00	–	–	–	–	.750	8	.863	.642	.752	.962	0.17	1.035	0.16
143.01	–	–	–	–	.322	123	.863	.642	.791	.522	1.60	.561	1.59
143.02	–	–	–	–	.644	147	.863	.642	.791	.878	2.69	.946	2.68
143.03	–	–	–	–	.644	62	.863	.642	.791	.878	1.14	.946	1.13
147.00	C	–	–	–	.160	308	.875	.520	.776	.308	2.84	.329	2.86
152.00	–	–	–	–	.465	138	.863	.639	.939	.614	2.70	.661	2.68
172.00	–	–	C	–	.196	97	.867	.665	1.034	.231	2.39	.248	2.38
173.00	–	–	–	–	.260	19	.863	.642	.700	.505	0.17	.543	0.17
179.08	–	–	–	–	.710	50	.863	.642	.791	.930	0.99	1.002	0.99
179.09	–	–	–	–	.440	64	.863	.642	.791	.674	0.93	.725	0.92
179.10	–	–	–	–	.470	19	.863	.642	.791	.709	0.28	.763	0.28
179.11	–	–	–	–	.240	25	.863	.642	.791	.401	0.31	.432	0.30
179.12	–	–	–	–	.700	19	.863	.642	.752	.931	0.37	1.002	0.36
179.13	–	–	–	–	.470	51	.863	.642	.752	.735	0.71	.791	0.71
179.14	–	–	–	–	.440	19	.863	.642	.752	.702	0.25	.756	0.25
181.01	–	–	–	–	.120	280	.867	.642	.776	.212	3.05	.228	3.04
181.02	–	–	–	–	.250	303	.867	.642	.776	.426	3.56	.457	3.56
181.03	–	–	–	–	.300	70	.867	.642	.776	.500	0.86	.537	0.86
181.04	–	–	–	–	.100	110	.867	.642	.776	.178	1.19	.191	1.19

Anmerkungen. siehe Seite 275.

Tabelle E.25
Lernleistung, WIT

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
100.00	60.0	–	–	C	–	.275	76	.917	.820	1.003	.303	5.28	.316	5.30
102.00	54.5	C	–	–	–	.220	54	.918	.642	.705	.407	1.33	.424	1.34

Tabelle E.25
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
103.00	0.0	C	X	M	-	.152	90	.914	.647	.752	.266	2.48	.278	2.48
104.00	0.0	M	X	-	-	.240	158	.917	.642	.752	.409	4.58	.427	4.59
105.00	88.9	M	X	M	X	.087	103	.917	.639	.752	.155	2.75	.161	2.76
107.01	86.4	M	-	M	X	.112	56	.917	.647	.752	.197	1.52	.206	1.53
107.02	94.4	M	-	M	X	.062	137	.917	.647	.752	.110	3.68	.115	3.69
107.03	88.9	M	-	M	X	.074	145	.917	.647	.752	.131	3.90	.137	3.91
108.00	90.9	M	-	M	-	.051	249	.917	.642	.752	.091	6.62	.095	6.64
110.00	0.0	M	X	-	-	.476	75	.917	.642	.752	.725	2.72	.757	2.73
114.00	0.0	-	-	C	-	.335	313	.917	.647	.752	.547	9.88	.572	9.91
126.01	0.0	-	-	C	-	.337	322	.917	.642	.689	.598	8.61	.624	8.64
126.02	0.0	-	-	C	-	.283	37	.917	.642	.737	.485	1.06	.506	1.07
126.03	0.0	-	-	C	-	.386	64	.917	.642	.625	.713	1.58	.745	1.58
126.04	0.0	-	-	C	-	.378	81	.917	.642	.641	.690	2.05	.721	2.05
126.05	0.0	-	-	C	-	.153	59	.917	.642	.673	.309	1.22	.322	1.23
126.06	0.0	-	-	C	-	.259	83	.917	.642	.609	.551	1.54	.576	1.55
129.00	50.0	M	X	-	-	.083	93	.917	.642	.752	.147	2.49	.154	2.50
130.00	0.0	M	X	-	-	.305	48	.917	.639	.752	.508	1.46	.530	1.46
133.00	0.0	-	-	-	-	.330	112	.925	.642	.785	.518	3.82	.539	3.87
134.00	0.0	-	-	-	-	.720	27	.917	.642	.890	.920	1.39	.961	1.40
147.00	0.0	C	-	-	-	.192	308	.915	.520	.776	.354	7.61	.371	7.62
156.00	0.0	-	-	-	-	.310	122	.917	.642	.970	.399	6.20	.417	6.22
159.00	0.0	C	X	C	-	.497	150	.860	.642	.776	.749	5.57	.807	5.24
176.01	0.0	C	-	-	-	.515	23	.915	.642	.791	.746	0.92	.780	0.92
176.02	0.0	C	-	-	-	.563	16	.915	.642	.791	.796	0.67	.832	0.67
178.00	0.0	C	-	M	X	.211	176	.917	.647	.752	.362	5.03	.378	5.05
179.01	0.0	-	-	-	-	-.232	12	.925	.642	.907	-.321	0.53	-.334	0.53
179.02	0.0	-	-	-	-	.881	15	.925	.642	.907	1.076	0.85	1.118	0.86
179.03	0.0	-	-	-	-	.612	9	.925	.642	.907	.798	0.45	.830	0.45
179.04	0.0	-	-	-	-	.265	10	.925	.642	.907	.366	0.44	.380	0.45
179.05	0.0	-	-	-	-	.805	7	.925	.642	.907	1.003	0.38	1.043	0.38
179.06	0.0	-	-	-	-	.630	15	.925	.642	.907	.819	0.75	.851	0.76
179.07	0.0	-	-	-	-	.360	14	.925	.642	.907	.491	0.63	.511	0.64

Anmerkungen. siehe Seite 275.

Tabelle E.26
Lernleistung, BIS

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
136.00	0.0	-	-	-	-	.360	97	.869	.642	.776	.583	7.0	.625	7.0
146.00	75.9	M	-	-	-	.085	45	.869	.642	.791	.147	2.9	.157	2.9
148.00	0.0	C	-	-	-	.521	91	.869	.642	.700	.817	7.0	.876	7.0
160.01	0.0	C	-	-	-	.351	58	.861	.642	.926	.475	6.0	.512	5.9
160.02	0.0	C	-	-	-	.160	91	.855	.642	.867	.242	7.6	.261	7.4
171.01	0.0	-	-	C	-	.390	197	.875	.740	.950	.478	25.0	.511	25.1
171.02	0.0	-	-	C	-	.332	142	.875	.720	1.018	.384	20.2	.410	20.3
171.03	0.0	-	-	C	-	.277	141	.875	.711	1.084	.300	22.9	.320	23.0
171.04	0.0	-	-	C	-	-.167	21	.875	.748	.757	-.279	1.4	-.298	1.4

Anmerkungen. siehe Seite 275.

Tabelle E.27
Lernleistung, LPS/PSB

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
118.01	0.0	–	–	C	–	.493	46	.935	.642	.791	.717	2.0	.741	2.0
118.02	0.0	–	–	C	–	.570	30	.935	.642	.791	.800	1.4	.827	1.4
118.03	0.0	–	–	C	–	.528	54	.935	.642	.791	.756	2.4	.781	2.4
122.00	0.0	–	–	–	–	.560	967	.935	.642	.791	.789	43.7	.816	43.7
127.01	0.0	–	–	C	–	.050	33	.935	.642	.791	.082	1.1	.085	1.1
127.02	0.0	–	–	C	–	.250	30	.935	.642	.791	.397	1.1	.411	1.1
140.00	0.0	–	–	–	–	.310	146	.935	.642	.791	.484	5.4	.501	5.4
144.00	0.0	–	–	–	–	.560	315	.935	.642	.791	.789	14.2	.816	14.2
147.00	0.0	–	–	–	–	.174	302	.935	.520	.776	.318	8.1	.329	8.1
151.00	0.0	C	–	M	–	.200	200	.935	.647	.791	.321	7.0	.332	7.0
175.00	45.5	–	–	M	X	.302	187	.935	.639	.791	.474	6.8	.490	6.8
177.01	33.3	M	–	–	–	.224	27	.935	.642	.791	.358	1.0	.371	1.0
177.02	66.7	M	–	–	–	.109	27	.935	.642	.791	.178	0.9	.184	0.9
184.01	0.0	–	–	–	–	.220	152	.935	.642	.721	.390	4.4	.403	4.4
184.02	0.0	–	–	–	–	.040	26	.935	.642	.721	.074	0.7	.076	0.7

Anmerkungen. siehe Seite 275.

Tabelle E.28
Lernleistung, Matrizentests

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
111.01	–	–	–	–	.430	78	.843	.642	.791	.668	5.8	.727	5.5
111.02	–	–	–	–	.360	98	.843	.642	.791	.580	6.7	.632	6.4
112.00	–	–	–	–	.240	93	.843	.642	.791	.407	5.8	.443	5.5
127.01	–	–	–	–	.210	33	.893	.642	.791	.347	2.2	.367	2.2
127.02	–	–	–	–	.250	30	.893	.642	.791	.408	2.0	.432	2.0
164.01	–	–	–	–	.210	486	.893	.642	.791	.347	31.8	.367	32.0
164.02	–	–	–	–	.350	198	.893	.642	.791	.551	14.2	.583	14.3
179.01	–	–	–	–	.279	12	.910	.642	.907	.386	1.1	.405	1.1
179.02	–	–	–	–	.369	15	.910	.642	.907	.504	1.4	.529	1.5
179.03	–	–	–	–	-.220	9	.910	.642	.907	-.306	0.8	-.321	0.9
179.04	–	–	–	–	.326	10	.910	.642	.907	.448	0.9	.470	1.0
179.05	–	–	–	–	.588	7	.910	.642	.907	.772	0.7	.809	0.7
179.06	–	–	–	–	-.380	15	.910	.642	.907	-.519	1.4	-.544	1.5
179.07	–	–	–	–	.336	14	.910	.642	.907	.462	1.3	.484	1.4
179.08	M	–	–	–	.360	51	.893	.642	.791	.564	3.7	.597	3.7
179.09	M	–	–	–	.325	64	.893	.642	.791	.517	4.5	.547	4.5
179.10	M	–	–	–	.485	23	.893	.642	.791	.718	1.9	.760	1.9
179.11	M	–	–	–	.290	26	.893	.642	.791	.467	1.8	.494	1.8
179.12	M	–	–	–	.445	20	.893	.642	.752	.697	1.5	.738	1.5
179.13	M	–	–	–	.355	52	.893	.642	.752	.585	3.4	.620	3.4
179.14	M	–	–	–	.445	19	.893	.642	.752	.697	1.4	.738	1.4
184.03	–	–	–	–	.310	89	.920	.642	.741	.520	5.6	.542	5.8

Anmerkungen. siehe Seite 275.

Tabelle E.29
Lernleistung, GATB-basierte Verfahren

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
106.00	0.0	M	–	M	X	.091	224	.895	.647	.791	.153	1.41	.162	1.41
127.01	0.0	M	X	C	–	.305	33	.895	.642	.791	.488	0.23	.516	0.23
127.02	0.0	M	X	C	–	.437	30	.895	.642	.791	.661	0.23	.699	0.23
138.00	86.1	C	–	M	–	.274	22	.895	.520	.791	.487	0.12	.515	0.12
189.01	0.0	C	–	–	–	.432	102	.895	.639	.752	.683	0.72	.722	0.72
189.02	0.0	C	–	–	–	.338	403	.895	.639	.791	.535	2.85	.566	2.85
189.03	0.0	C	–	–	–	.437	282	.895	.639	.791	.663	2.18	.700	2.18
189.04	0.0	C	–	–	–	.377	87	.895	.639	.791	.587	0.64	.621	0.64
189.05	0.0	C	–	–	–	.518	68	.895	.639	.791	.755	0.57	.798	0.57
189.06	0.0	C	–	–	–	.501	731	.895	.639	.791	.736	6.01	.778	6.01
189.07	0.0	C	–	–	–	.437	198	.895	.639	.791	.663	1.53	.700	1.53
189.08	0.0	C	–	–	–	.372	112	.895	.639	.791	.581	0.82	.614	0.82
189.09	0.0	C	–	–	–	.398	351	.895	.639	.791	.614	2.61	.649	2.61
189.10	0.0	C	–	–	–	.317	394	.895	.639	.791	.506	2.75	.535	2.75
189.11	0.0	C	–	–	–	.218	81	.895	.639	.791	.359	0.53	.380	0.53
189.12	0.0	C	–	–	–	.368	191	.895	.639	.791	.575	1.39	.608	1.39
189.13	0.0	C	–	–	–	.368	97	.895	.639	.791	.575	0.70	.608	0.70
189.14	0.0	C	–	–	–	.616	617	.895	.639	.791	.852	5.73	.900	5.73
189.15	0.0	C	–	–	–	.475	418	.895	.639	.791	.707	3.35	.748	3.35
189.16	0.0	C	–	–	–	.428	412	.895	.639	.791	.652	3.15	.689	3.15
189.17	0.0	C	–	–	–	.394	243	.895	.639	.791	.609	1.80	.644	1.80
189.18	0.0	C	–	–	–	.330	98	.895	.639	.791	.524	0.69	.554	0.69
189.19	0.0	C	–	–	–	.561	287	.895	.639	.791	.799	2.51	.845	2.51
189.20	0.0	C	–	–	–	.411	291	.895	.639	.791	.631	2.19	.667	2.19
189.21	0.0	C	–	–	–	.278	184	.895	.639	.791	.450	1.25	.476	1.25
189.22	0.0	C	–	–	–	.385	84	.895	.639	.791	.598	0.62	.632	0.62
189.23	0.0	C	–	–	–	.214	158	.895	.639	.791	.353	1.03	.373	1.03
189.24	0.0	C	–	–	–	.300	119	.895	.639	.791	.482	0.82	.509	0.82
189.25	0.0	C	–	–	–	.381	67	.895	.639	.791	.593	0.49	.626	0.49
189.26	0.0	C	–	–	–	.407	57	.895	.639	.791	.626	0.43	.661	0.43
189.27	0.0	C	–	–	–	.377	81	.895	.639	.776	.598	0.57	.632	0.57
189.28	0.0	C	–	–	–	.479	181	.895	.639	.791	.712	1.45	.752	1.45
189.29	0.0	C	–	–	–	.347	314	.895	.639	.791	.547	2.24	.579	2.24
189.30	0.0	C	–	–	–	.407	454	.895	.639	.791	.626	3.41	.661	3.41
189.31	0.0	C	–	–	–	.501	380	.895	.639	.791	.736	3.12	.778	3.12
189.32	0.0	C	–	–	–	.381	412	.895	.639	.791	.593	3.02	.626	3.02
189.33	0.0	C	–	–	–	.488	391	.895	.639	.791	.722	3.17	.763	3.17
189.34	0.0	C	–	–	–	.321	48	.895	.639	.791	.512	0.34	.541	0.34
189.35	0.0	C	–	–	–	.390	58	.895	.639	.752	.631	0.39	.667	0.39
189.36	0.0	C	–	–	–	.424	57	.895	.639	.752	.673	0.40	.712	0.40
189.37	0.0	C	–	–	–	.223	87	.895	.639	.752	.390	0.50	.413	0.50
189.38	0.0	C	–	–	–	.334	56	.895	.639	.791	.530	0.40	.560	0.40
189.39	0.0	C	–	–	–	.248	48	.895	.639	.791	.405	0.32	.428	0.32
189.40	0.0	C	–	–	–	.240	61	.895	.639	.791	.393	0.40	.416	0.40
189.41	0.0	C	–	–	–	.214	104	.895	.639	.791	.353	0.68	.373	0.68
189.42	0.0	C	–	–	–	.227	113	.895	.639	.791	.373	0.74	.395	0.74
189.43	0.0	C	–	–	–	.338	38	.895	.639	.791	.535	0.27	.566	0.27
189.44	0.0	C	–	–	–	.295	21	.895	.639	.791	.475	0.14	.502	0.14
189.45	0.0	C	–	–	–	.390	178	.895	.639	.752	.631	1.21	.667	1.21
189.46	0.0	C	–	–	–	.432	347	.895	.639	.752	.683	2.47	.722	2.47
189.47	0.0	C	–	–	–	.454	37	.895	.639	.752	.708	0.27	.749	0.27
189.48	0.0	C	–	–	–	.424	41	.895	.639	.752	.673	0.29	.712	0.29
189.49	0.0	C	–	–	–	.595	104	.895	.639	.752	.849	0.91	.897	0.91
189.50	0.0	C	–	–	–	.454	174	.895	.639	.752	.708	1.27	.749	1.27
189.51	0.0	C	–	–	–	.257	104	.895	.639	.791	.419	0.70	.443	0.70
189.52	0.0	C	–	–	–	.509	32	.895	.639	.752	.768	0.25	.811	0.25
189.53	0.0	C	–	–	–	.574	268	.895	.639	.791	.812	2.38	.858	2.38
189.54	0.0	C	–	–	–	.432	212	.895	.639	.791	.657	1.63	.694	1.63
189.55	0.0	C	–	–	–	.488	98	.895	.639	.791	.722	0.79	.763	0.79

Tabelle E.29
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.56	0.0	C	–	–	–	.484	87	.895	.639	.791	.717	0.70	.758	0.70
189.57	0.0	C	–	–	–	.424	106	.895	.639	.791	.647	0.81	.684	0.81
189.58	0.0	C	–	–	–	.287	124	.895	.639	.791	.463	0.85	.489	0.85
189.59	0.0	C	–	–	–	.428	168	.895	.639	.791	.652	1.29	.689	1.29
189.60	0.0	C	–	–	–	.372	117	.895	.639	.791	.581	0.85	.614	0.85
189.61	0.0	C	–	–	–	.407	198	.895	.639	.791	.626	1.49	.661	1.49
189.62	0.0	C	–	–	–	.390	314	.895	.639	.791	.604	2.32	.639	2.32
189.63	0.0	C	–	–	–	.420	217	.895	.639	.791	.642	1.65	.679	1.65
189.64	0.0	C	–	–	–	.462	208	.895	.639	.791	.692	1.64	.732	1.64
189.65	0.0	C	–	–	–	.458	197	.895	.639	.791	.688	1.55	.727	1.55
189.66	0.0	C	–	–	–	.458	128	.895	.639	.791	.688	1.01	.727	1.01
189.67	0.0	C	–	–	–	.471	117	.895	.639	.791	.703	0.93	.743	0.93
189.68	0.0	C	–	–	–	.462	89	.895	.639	.791	.692	0.70	.732	0.70
189.69	0.0	C	–	–	–	.492	58	.895	.639	.791	.726	0.47	.768	0.47
189.70	0.0	C	–	–	–	.492	87	.895	.639	.752	.750	0.66	.793	0.66

Anmerkungen. siehe Seite 275.

Tabelle E.30
Lernleistung, andere Intelligenztests

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
123.01	0.0	–	–	–	–	.710	20	.895	.520	.791	.988	0.8	1.044	0.8
123.02	0.0	–	–	–	–	.650	20	.895	.520	.791	.939	0.7	.993	0.7
123.03	0.0	–	–	–	–	.800	19	.895	.520	.791	1.050	0.8	1.110	0.8
123.04	0.0	–	–	–	–	.720	7	.895	.520	.791	.995	0.3	1.052	0.3
123.05	0.0	–	–	–	–	.590	13	.895	.520	.791	.885	0.4	.936	0.4
123.06	0.0	–	–	–	–	.850	10	.895	.520	.791	1.080	0.5	1.142	0.5
124.01	0.0	M	–	–	–	.803	12	.895	.520	.791	1.052	0.5	1.112	0.5
124.02	0.0	M	–	–	–	.806	8	.895	.520	.791	1.054	0.4	1.114	0.4
124.03	0.0	M	–	–	–	.515	12	.895	.520	.791	.809	0.4	.855	0.4
135.00	0.0	–	–	M	–	.470	861	.895	.647	.791	.698	29.4	.738	29.3
139.00	72.2	M	–	M	–	.094	119	.895	.642	.791	.159	3.2	.168	3.1
149.00	0.0	M	–	C	–	.093	199	.895	.642	.752	.168	4.6	.178	4.6
150.00	0.0	–	–	–	–	.322	167	.895	.642	.776	.522	4.8	.552	4.8
152.00	0.0	–	–	M	–	.520	180	.934	.639	.939	.677	8.0	.701	8.3
153.00	0.0	–	–	–	–	.570	153	.870	.639	.737	.839	5.3	.900	5.2
161.01	0.0	M	–	–	–	.016	60	.895	.642	.873	.024	2.1	.025	2.1
161.02	0.0	M	–	–	–	.086	31	.895	.642	.752	.156	0.7	.165	0.7
167.00	0.0	C	X	C	–	.078	139	.895	.639	.791	.132	3.6	.140	3.6
168.00	0.0	–	–	M	–	.476	51	.895	.639	.791	.708	1.7	.749	1.7
174.00	0.0	–	–	M	X	.315	469	.895	.639	.791	.503	13.8	.532	13.8
182.00	0.0	–	–	–	–	.090	200	.895	.642	.776	.156	5.0	.165	5.0
183.00	0.0	–	–	–	–	.070	256	.895	.642	.752	.127	5.8	.134	5.8
185.00	0.0	–	–	M	–	.288	90	.895	.642	.776	.473	2.5	.500	2.5
186.00	0.0	–	–	–	–	.320	167	.895	.642	.776	.519	4.8	.549	4.8

Anmerkungen. siehe Seite 275.

Tabelle E.31
Lernleistung, Berufsschulnoten

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
100.00	60.0	-	-	C	-	.269	76	.911	.820	1.003	.296	0.92	.310	0.94
101.01	0.0	-	-	C	-	.250	72	.863	.647	.467	.825	0.10	.888	0.09
101.02	0.0	-	-	C	-	.296	65	.863	.784	.489	.804	0.13	.866	0.12
102.00	54.5	C	-	-	-	.213	54	.918	.642	.705	.395	0.23	.412	0.24
103.00	0.0	C	X	M	-	.152	90	.914	.647	.752	.266	0.43	.278	0.44
104.00	0.0	M	X	-	-	.240	158	.911	.642	.752	.411	0.79	.431	0.80
105.00	83.3	M	X	-	-	.099	103	.911	.642	.752	.176	0.48	.185	0.49
106.00	0.0	M	-	M	X	.091	224	.895	.647	.791	.153	1.16	.162	1.16
107.01	86.4	M	-	M	X	.106	56	.911	.647	.752	.188	0.26	.197	0.27
107.02	94.4	M	-	M	X	.058	137	.911	.647	.752	.103	0.63	.108	0.64
107.03	88.9	M	-	M	X	.070	145	.911	.647	.752	.125	0.67	.131	0.68
108.00	90.9	M	-	M	-	.047	249	.911	.642	.752	.084	1.14	.088	1.16
109.00	0.0	C	-	C	-	.042	102	.882	.811	.752	.069	0.55	.074	0.55
110.00	0.0	M	X	-	-	.476	75	.911	.642	.752	.727	0.47	.761	0.48
111.01	0.0	C	-	-	-	.467	78	.913	.642	.791	.692	0.52	.724	0.53
111.02	0.0	C	-	-	-	.440	98	.913	.642	.791	.660	0.64	.691	0.65
112.00	0.0	-	-	-	-	.240	93	.843	.642	.791	.407	0.47	.443	0.45
113.00	0.0	-	-	M	-	.219	100	.863	.642	.750	.396	0.45	.426	0.43
114.00	0.0	-	-	M	-	.335	313	.911	.647	.752	.549	1.70	.575	1.74
116.01	0.0	C	-	-	-	.355	126	.882	.642	.776	.572	0.71	.609	0.70
116.03	0.0	C	-	-	-	.556	70	.882	.642	.752	.815	0.48	.868	0.47
116.04	0.0	C	-	-	-	.326	215	.882	.642	.776	.532	1.18	.566	1.17
116.06	0.0	C	-	-	-	.404	230	.882	.642	.791	.625	1.41	.665	1.39
116.08	0.0	C	-	-	-	.534	304	.882	.642	.752	.795	2.01	.846	1.98
116.09	0.0	C	-	-	-	.446	169	.882	.642	.752	.702	1.00	.748	0.99
116.10	0.0	C	-	-	-	.547	407	.882	.642	.791	.786	2.88	.837	2.85
116.11	0.0	C	-	-	-	.749	76	.882	.642	.791	.959	0.68	1.021	0.67
116.13	0.0	-	-	-	-	.760	180	.863	.642	.776	.967	1.63	1.040	1.57
117.00	0.0	-	-	-	-	.700	42	.863	.642	.791	.923	0.35	.993	0.34
118.01	0.0	-	-	C	-	.493	46	.935	.642	.791	.717	0.32	.741	0.33
118.02	0.0	-	-	C	-	.570	30	.935	.642	.791	.800	0.22	.827	0.23
118.03	0.0	-	-	C	-	.528	54	.935	.642	.791	.756	0.39	.781	0.40
119.01	0.0	-	-	-	-	.330	100	.863	.642	.752	.563	0.50	.606	0.49
119.02	0.0	-	-	-	-	.240	99	.863	.642	.791	.401	0.52	.432	0.50
120.01	0.0	-	-	-	-	.530	127	.863	.642	.752	.796	0.82	.856	0.80
120.02	0.0	-	-	-	-	.600	101	.863	.642	.752	.857	0.72	.923	0.70
120.03	0.0	-	-	-	-	.570	222	.863	.642	.844	.786	1.71	.846	1.65
120.04	0.0	-	-	-	-	.590	13	.863	.642	.752	.849	0.09	.914	0.09
122.00	0.0	-	-	-	-	.560	967	.935	.642	.791	.789	7.12	.816	7.46
125.01	0.0	-	-	C	-	.293	738	.863	.642	.811	.466	4.26	.502	4.12
125.02	0.0	-	-	C	-	.262	221	.863	.642	.778	.444	1.13	.478	1.09
125.03	0.0	-	-	-	-	.160	196	.863	.642	.644	.377	0.52	.405	0.50
125.04	0.0	-	-	C	-	.198	198	.863	.642	.711	.390	0.75	.420	0.72
125.05	0.0	-	-	C	-	.233	285	.863	.642	.900	.331	2.07	.356	2.00
125.06	0.0	-	-	C	-	.205	120	.863	.642	.722	.394	0.48	.424	0.46
126.01	0.0	-	-	C	-	.337	322	.911	.642	.689	.601	1.48	.630	1.51
126.02	0.0	-	-	C	-	.283	37	.911	.642	.737	.487	0.18	.510	0.19
126.03	0.0	-	-	C	-	.386	64	.911	.642	.625	.717	0.27	.752	0.28
126.04	0.0	-	-	C	-	.378	81	.911	.642	.641	.694	0.35	.727	0.36
126.05	0.0	-	-	C	-	.153	59	.911	.642	.673	.311	0.21	.326	0.21
126.06	0.0	-	-	C	-	.259	83	.911	.642	.609	.557	0.26	.583	0.27
127.01	0.0	M	X	-	-	.215	33	.895	.642	.791	.340	0.18	.359	0.18
127.02	0.0	M	X	-	-	.339	30	.895	.642	.791	.569	0.18	.601	0.18
129.00	50.0	M	X	-	-	.080	93	.911	.642	.752	.143	0.43	.150	0.44
130.00	0.0	M	X	-	-	.391	51	.911	.642	.752	.626	0.29	.655	0.30
133.00	0.0	-	-	-	-	.330	112	.925	.642	.785	.518	0.66	.539	0.69
134.00	0.0	-	-	-	-	.720	27	.911	.642	.890	.920	0.24	.964	0.25
135.00	0.0	-	-	M	-	.470	861	.895	.647	.791	.698	5.71	.738	5.73
136.00	0.0	-	-	-	-	.360	97	.869	.642	.776	.583	0.54	.625	0.53

Tabelle E.31
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
137.01	0.0	-	-	-	-	.200	74	.863	.642	.752	.363	0.33	.391	0.32
137.02	0.0	-	-	-	-	.400	64	.863	.642	.791	.625	0.38	.673	0.37
137.03	0.0	-	-	-	-	.480	44	.863	.642	.791	.720	0.29	.775	0.28
137.04	0.0	-	-	-	-	.190	59	.863	.642	.791	.322	0.30	.347	0.29
137.05	0.0	-	-	-	-	.100	27	.863	.642	.791	.173	0.13	.186	0.13
137.06	0.0	-	-	-	-	.280	49	.863	.642	.791	.461	0.26	.497	0.26
137.07	0.0	-	-	-	-	.120	50	.863	.642	.791	.207	0.25	.223	0.24
139.00	72.2	M	-	M	-	.101	119	.895	.642	.791	.170	0.61	.180	0.61
140.00	0.0	C	-	-	-	.469	146	.940	.642	.791	.688	0.99	.710	1.05
141.00	0.0	-	-	-	-	.750	34	.863	.642	.791	.959	0.30	1.032	0.29
142.00	0.0	-	-	-	-	.750	8	.863	.642	.752	.962	0.07	1.035	0.07
143.01	0.0	-	-	-	-	.322	123	.863	.642	.791	.522	0.69	.561	0.66
143.02	0.0	-	-	-	-	.644	147	.863	.642	.791	.878	1.16	.946	1.12
143.03	0.0	-	-	-	-	.644	62	.863	.642	.791	.878	0.49	.946	0.47
144.00	0.0	-	-	-	-	.560	315	.935	.642	.791	.789	2.32	.816	2.43
146.00	75.9	M	-	-	-	.115	45	.869	.642	.791	.197	0.22	.212	0.22
148.00	0.0	C	-	-	-	.521	91	.869	.642	.700	.817	0.54	.876	0.53
149.00	0.0	M	-	C	-	.093	199	.895	.642	.752	.168	0.89	.178	0.89
150.00	0.0	-	-	-	-	.322	167	.895	.642	.776	.522	0.93	.552	0.93
151.00	0.0	C	-	M	-	.200	200	.935	.647	.791	.321	1.14	.332	1.19
152.00	0.0	C	-	-	-	.688	53	.934	.642	.939	.874	0.48	.904	0.50
153.00	0.0	-	-	-	-	.445	153	.870	.647	.737	.714	0.87	.766	0.85
156.00	0.0	-	-	-	-	.310	122	.911	.642	.970	.399	1.08	.418	1.10
159.00	0.0	C	X	C	-	.523	150	.877	.642	.776	.771	1.01	.824	0.99
160.01	0.0	C	-	-	-	.351	58	.861	.642	.926	.475	0.46	.512	0.45
160.02	0.0	C	-	-	-	.160	91	.855	.642	.867	.242	0.58	.261	0.56
161.01	0.0	M	-	-	-	.016	60	.895	.642	.873	.024	0.40	.025	0.40
161.02	0.0	M	-	-	-	.086	31	.895	.642	.752	.156	0.14	.165	0.14
164.01	0.0	-	-	-	-	.245	486	.895	.642	.791	.400	2.67	.423	2.68
164.02	0.0	-	-	-	-	.350	198	.895	.642	.791	.550	1.17	.582	1.18
167.00	0.0	C	X	-	-	.028	139	.895	.642	.791	.047	0.71	.050	0.71
168.00	0.0	-	-	C	-	.455	51	.895	.784	.791	.630	0.39	.666	0.39
171.01	0.0	-	-	C	-	.466	197	.875	.642	.950	.607	1.70	.649	1.67
171.02	0.0	-	-	C	-	.384	142	.875	.642	1.018	.471	1.38	.503	1.35
171.03	0.0	-	-	C	-	.314	141	.875	.642	1.084	.359	1.58	.384	1.55
171.04	0.0	-	-	C	-	-.118	21	.875	.642	.757	-.215	0.09	-.230	0.09
172.00	0.0	-	-	-	-	.400	97	.869	.642	1.034	.483	0.97	.518	0.95
173.00	0.0	-	-	-	-	.260	190	.863	.642	.700	.505	0.74	.543	0.71
174.00	0.0	-	-	-	-	.440	182	.895	.642	.791	.665	1.17	.703	1.17
175.00	0.0	-	-	M	-	.470	187	.935	.647	.791	.688	1.28	.712	1.34
176.01	0.0	C	-	-	-	.515	23	.915	.642	.791	.746	0.16	.780	0.16
176.02	0.0	C	-	-	-	.563	16	.915	.642	.791	.796	0.12	.832	0.12
177.01	33.3	M	-	-	-	.220	27	.935	.642	.791	.352	0.15	.364	0.16
177.02	66.7	M	-	-	-	.098	27	.935	.642	.791	.160	0.15	.166	0.15
178.00	0.0	C	-	M	X	.211	176	.911	.647	.752	.364	0.87	.381	0.88
179.01	0.0	C	-	-	-	.027	12	.947	.642	.907	.038	0.09	.039	0.10
179.02	0.0	C	-	-	-	.711	15	.947	.642	.907	.906	0.14	.931	0.14
179.03	0.0	C	-	-	-	.226	9	.947	.642	.907	.312	0.07	.320	0.07
179.04	0.0	C	-	-	-	.336	10	.947	.642	.907	.458	0.08	.471	0.08
179.05	0.0	C	-	-	-	.793	7	.947	.642	.907	.991	0.07	1.019	0.07
179.06	0.0	C	-	-	-	.141	15	.947	.642	.907	.196	0.11	.201	0.12
179.07	0.0	C	-	-	-	.396	14	.947	.642	.907	.536	0.11	.550	0.12
179.08	0.0	M	-	-	-	.477	51	.895	.642	.791	.708	0.34	.749	0.34
179.09	0.0	M	-	-	-	.363	64	.895	.642	.791	.568	0.38	.600	0.38
179.10	0.0	M	-	-	-	.480	23	.895	.642	.791	.712	0.15	.752	0.15
179.11	0.0	M	-	-	-	.273	26	.895	.642	.791	.442	0.15	.467	0.15
179.12	0.0	M	-	-	-	.530	20	.895	.642	.752	.788	0.13	.832	0.13
179.13	0.0	M	-	-	-	.393	52	.895	.642	.752	.634	0.29	.670	0.29
179.14	0.0	M	-	-	-	.443	19	.895	.642	.752	.694	0.11	.734	0.11

Tabelle E.31
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
181.01	0.0	-	-	-	-	.120	280	.869	.642	.776	.212	1.31	.227	1.28
181.02	0.0	-	-	-	-	.250	303	.869	.642	.776	.425	1.54	.456	1.49
181.03	0.0	-	-	-	-	.300	70	.869	.642	.776	.499	0.37	.536	0.36
181.04	0.0	-	-	-	-	.100	110	.869	.642	.776	.177	0.51	.190	0.50
182.00	0.0	-	-	-	-	.090	200	.895	.642	.776	.156	0.97	.165	0.98
183.00	0.0	-	-	-	-	.070	256	.895	.642	.752	.127	1.14	.134	1.14
184.01	0.0	-	-	-	-	.220	152	.935	.642	.721	.390	0.71	.403	0.74
184.02	0.0	-	-	-	-	.040	26	.935	.642	.721	.074	0.11	.076	0.12
184.03	0.0	-	-	-	-	.310	89	.920	.642	.741	.520	0.46	.542	0.48
185.00	0.0	-	-	M	-	.288	90	.895	.642	.776	.473	0.49	.500	0.49
186.00	0.0	-	-	-	-	.320	167	.895	.642	.776	.519	0.93	.549	0.93

Anmerkungen. siehe Seite 275.

Tabelle E.32
Lernleistung, betriebliche Beurteilungen

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
101.01	0.0	-	-	-	-	.080	72	.863	.520	.467	.480	0.2	.517	0.2
105.00	91.7	M	X	M	-	.062	103	.911	.520	.752	.123	2.6	.129	2.7
113.00	0.0	-	-	M	-	.252	148	.863	.520	.750	.492	3.9	.530	3.8
119.01	0.0	-	-	-	-	.010	100	.863	.520	.752	.021	2.3	.023	2.2
119.02	0.0	-	-	-	-	.110	99	.863	.520	.791	.211	2.7	.227	2.6
123.01	0.0	-	-	-	-	.710	20	.895	.520	.791	.988	1.0	1.044	1.0
123.02	0.0	-	-	-	-	.650	20	.895	.520	.791	.939	1.0	.993	1.0
123.03	0.0	-	-	-	-	.800	19	.895	.520	.791	1.050	1.1	1.110	1.1
123.04	0.0	-	-	-	-	.720	7	.895	.520	.791	.995	0.4	1.052	0.4
123.05	0.0	-	-	-	-	.590	13	.895	.520	.791	.885	0.6	.936	0.6
123.06	0.0	-	-	-	-	.850	10	.895	.520	.791	1.080	0.6	1.142	0.6
124.01	0.0	M	-	-	-	.803	12	.895	.520	.791	1.052	0.7	1.112	0.7
124.02	0.0	M	-	-	-	.806	8	.895	.520	.791	1.054	0.5	1.114	0.5
124.03	0.0	M	-	-	-	.515	12	.895	.520	.791	.809	0.5	.855	0.5
130.00	0.0	M	X	-	-	.305	51	.911	.520	.752	.557	1.5	.583	1.6
138.00	86.1	C	-	M	-	.222	22	.895	.520	.791	.402	0.7	.425	0.7
147.00	0.0	C	X	-	-	.196	308	.895	.520	.776	.367	8.8	.388	8.8
152.00	0.0	C	-	-	-	.371	180	.934	.520	.939	.542	8.4	.561	8.9
153.00	0.0	-	-	-	-	.590	153	.870	.520	.737	.904	6.5	.969	6.4
167.00	0.0	C	X	C	-	.128	139	.895	.520	.791	.238	4.0	.252	4.0
168.00	0.0	-	-	C	-	.462	51	.895	.520	.791	.748	1.9	.791	2.0
171.01	0.0	-	-	C	-	.267	197	.875	.520	.950	.392	9.1	.419	9.0
171.02	0.0	-	-	C	-	.214	142	.875	.520	1.018	.291	7.7	.311	7.6
171.03	0.0	-	-	C	-	.184	141	.875	.520	1.084	.232	8.9	.248	8.7
171.04	0.0	-	-	C	-	-.193	21	.875	.520	.757	-.379	0.5	-.406	0.5
172.00	0.0	-	-	-	-	-.090	97	.869	.520	1.034	-.120	5.5	-.128	5.4
174.00	0.0	-	-	M	X	.190	469	.895	.520	.791	.348	14.0	.368	14.1
175.00	71.4	-	-	M	X	.118	146	.935	.520	.791	.214	4.5	.221	4.7

Anmerkungen. siehe Seite 275.

Tabelle E.33
Lernleistung, Ausbildung

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
101.01	0.0	-	-	C	-	.147	72	.863	.639	.467	.667	0.03	.718	0.03
101.02	0.0	-	-	C	-	.296	65	.863	.784	.489	.804	0.07	.866	0.07
102.00	54.5	C	-	-	-	.216	54	.918	.642	.705	.400	0.13	.417	0.13
103.00	0.0	C	X	M	-	.152	90	.914	.647	.752	.266	0.25	.278	0.25
104.00	0.0	M	X	-	-	.240	158	.911	.642	.752	.411	0.45	.431	0.46
105.00	88.9	M	X	M	X	.084	103	.911	.639	.752	.150	0.27	.157	0.27
106.00	0.0	M	-	M	X	.091	224	.895	.647	.791	.153	0.66	.162	0.66
107.01	86.4	M	-	M	X	.108	56	.911	.647	.752	.191	0.15	.200	0.15
107.02	94.4	M	-	M	X	.059	137	.911	.647	.752	.105	0.36	.110	0.37
107.03	88.9	M	-	M	X	.071	145	.911	.647	.752	.126	0.38	.132	0.39
109.00	0.0	C	-	C	-	.042	102	.882	.811	.752	.069	0.31	.074	0.31
111.01	0.0	C	-	-	-	.467	78	.913	.642	.791	.692	0.30	.724	0.30
111.02	0.0	C	-	-	-	.440	98	.913	.642	.791	.660	0.36	.691	0.37
112.00	0.0	-	-	-	-	.240	93	.843	.642	.791	.407	0.27	.443	0.25
113.00	0.0	-	-	M	X	.236	148	.863	.639	.750	.425	0.38	.457	0.37
116.01	0.0	C	-	-	-	.355	126	.882	.642	.776	.572	0.40	.609	0.40
116.03	0.0	C	-	-	-	.556	70	.882	.642	.752	.815	0.27	.868	0.27
116.04	0.0	C	-	-	-	.326	215	.882	.642	.776	.532	0.67	.566	0.66
116.06	0.0	C	-	-	-	.404	230	.882	.642	.791	.625	0.80	.665	0.79
116.08	0.0	C	-	-	-	.534	304	.882	.642	.752	.795	1.14	.846	1.13
116.09	0.0	C	-	-	-	.446	169	.882	.642	.752	.702	0.57	.748	0.56
116.10	0.0	C	-	-	-	.547	407	.882	.642	.791	.786	1.64	.837	1.62
116.11	0.0	C	-	-	-	.749	76	.882	.642	.791	.959	0.39	1.021	0.38
116.13	0.0	-	-	-	-	.760	180	.863	.642	.776	.967	0.93	1.040	0.89
118.01	0.0	-	-	C	-	.493	46	.935	.642	.791	.717	0.18	.741	0.19
118.02	0.0	-	-	C	-	.570	30	.935	.642	.791	.800	0.13	.827	0.13
118.03	0.0	-	-	C	-	.528	54	.935	.642	.791	.756	0.22	.781	0.23
119.01	0.0	-	-	C	-	.210	100	.863	.680	.752	.370	0.27	.398	0.26
119.02	0.0	-	-	C	-	.216	99	.863	.680	.791	.354	0.31	.381	0.30
122.00	0.0	-	-	-	-	.560	967	.935	.642	.791	.789	4.05	.816	4.24
124.01	0.0	M	-	-	-	.803	12	.895	.520	.791	1.052	0.06	1.112	0.06
124.02	0.0	M	-	-	-	.806	8	.895	.520	.791	1.054	0.04	1.114	0.04
124.03	0.0	M	-	-	-	.515	12	.895	.520	.791	.809	0.04	.855	0.04
125.01	0.0	-	-	C	-	.293	738	.863	.642	.811	.466	2.43	.502	2.34
125.02	0.0	-	-	C	-	.262	221	.863	.642	.778	.444	0.64	.478	0.62
125.03	0.0	-	-	-	-	.160	196	.863	.642	.644	.377	0.29	.405	0.28
125.04	0.0	-	-	C	-	.198	198	.863	.642	.711	.390	0.42	.420	0.41
125.05	0.0	-	-	C	-	.233	285	.863	.642	.900	.331	1.18	.356	1.14
125.06	0.0	-	-	C	-	.205	120	.863	.642	.722	.394	0.27	.424	0.26
126.01	0.0	-	-	C	-	.337	322	.911	.642	.689	.601	0.84	.630	0.86
126.02	0.0	-	-	C	-	.283	37	.911	.642	.737	.487	0.10	.510	0.11
126.03	0.0	-	-	C	-	.386	64	.911	.642	.625	.717	0.15	.752	0.16
126.04	0.0	-	-	C	-	.378	81	.911	.642	.641	.694	0.20	.727	0.20
126.05	0.0	-	-	C	-	.153	59	.911	.642	.673	.311	0.12	.326	0.12
126.06	0.0	-	-	C	-	.259	83	.911	.642	.609	.557	0.15	.583	0.15
127.01	0.0	M	X	C	-	.211	33	.895	.642	.791	.348	0.10	.368	0.10
127.02	0.0	M	X	C	-	.337	30	.895	.642	.791	.533	0.10	.563	0.10
129.00	50.0	M	X	-	-	.081	93	.911	.642	.752	.145	0.24	.151	0.25
130.00	0.0	M	X	-	-	.305	48	.911	.639	.752	.510	0.14	.534	0.15
133.00	0.0	-	-	-	-	.330	112	.925	.642	.785	.518	0.38	.539	0.39
134.00	0.0	-	-	-	-	.720	27	.911	.642	.890	.920	0.14	.964	0.14
135.00	0.0	-	-	M	-	.470	861	.895	.647	.791	.698	3.25	.738	3.25
137.01	0.0	-	-	-	-	.200	74	.863	.642	.752	.363	0.19	.391	0.18
137.02	0.0	-	-	-	-	.400	64	.863	.642	.791	.625	0.22	.673	0.21
137.03	0.0	-	-	-	-	.480	44	.863	.642	.791	.720	0.16	.775	0.16
137.04	0.0	-	-	-	-	.190	59	.863	.642	.791	.322	0.17	.347	0.16
137.05	0.0	-	-	-	-	.100	27	.863	.642	.791	.173	0.08	.186	0.07
137.06	0.0	-	-	-	-	.280	49	.863	.642	.791	.461	0.15	.497	0.15
137.07	0.0	-	-	-	-	.120	50	.863	.642	.791	.207	0.14	.223	0.14

Tabelle E.33
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
138.00	86.1	C	-	M	-	.232	22	.895	.520	.791	.419	0.06	.443	0.06
139.00	72.2	M	-	M	-	.103	119	.895	.642	.791	.174	0.35	.183	0.35
140.00	0.0	C	-	-	-	.469	146	.940	.642	.791	.688	0.57	.710	0.59
143.01	0.0	-	-	-	-	.322	123	.863	.642	.791	.522	0.39	.561	0.38
143.02	0.0	-	-	-	-	.644	147	.863	.642	.791	.878	0.66	.946	0.64
143.03	0.0	-	-	-	-	.644	62	.863	.642	.791	.878	0.28	.946	0.27
144.00	0.0	-	-	-	-	.560	315	.935	.642	.791	.789	1.32	.816	1.38
146.00	75.9	M	-	-	-	.118	45	.869	.642	.791	.203	0.13	.217	0.12
147.00	0.0	C	X	-	-	.196	308	.895	.520	.776	.367	0.73	.388	0.73
149.00	0.0	M	-	C	-	.093	199	.895	.642	.752	.168	0.51	.178	0.51
150.00	0.0	-	-	-	-	.322	167	.895	.642	.776	.522	0.53	.552	0.53
151.00	0.0	C	-	M	-	.200	200	.935	.647	.791	.321	0.65	.332	0.68
152.00	0.0	C	-	M	-	.529	180	.934	.639	.939	.688	0.89	.712	0.93
153.00	0.0	-	-	-	-	.570	153	.870	.639	.737	.839	0.59	.900	0.57
156.00	0.0	-	-	-	-	.310	122	.911	.642	.970	.399	0.61	.418	0.62
159.00	0.0	C	X	C	-	.523	150	.877	.642	.776	.771	0.57	.824	0.56
160.01	0.0	C	-	-	-	.351	58	.861	.642	.926	.475	0.26	.512	0.25
160.02	0.0	C	-	-	-	.160	91	.855	.642	.867	.242	0.33	.261	0.32
161.01	0.0	M	-	-	-	.016	60	.895	.642	.873	.024	0.23	.025	0.23
161.02	0.0	M	-	-	-	.086	31	.895	.642	.752	.156	0.08	.165	0.08
164.01	0.0	-	-	-	-	.245	486	.895	.642	.791	.400	1.52	.423	1.52
164.02	0.0	-	-	-	-	.350	198	.895	.642	.791	.550	0.67	.582	0.67
167.00	0.0	C	X	C	-	.078	139	.895	.639	.791	.132	0.40	.140	0.40
168.00	0.0	-	-	M	-	.476	51	.895	.639	.791	.708	0.19	.749	0.19
171.01	0.0	-	-	C	-	.390	197	.875	.740	.950	.478	1.09	.511	1.07
171.02	0.0	-	-	C	-	.332	142	.875	.720	1.018	.384	0.89	.410	0.87
171.03	0.0	-	-	C	-	.277	141	.875	.711	1.084	.300	1.00	.320	0.98
171.04	0.0	-	-	C	-	-.167	21	.875	.748	.757	-.279	0.06	-.298	0.06
172.00	0.0	-	-	C	-	.196	97	.869	.665	1.034	.231	0.58	.248	0.57
174.00	0.0	-	-	M	X	.315	469	.895	.639	.791	.503	1.53	.532	1.53
175.00	45.5	-	-	M	X	.295	187	.935	.639	.791	.464	0.63	.480	0.66
176.01	0.0	C	-	-	-	.515	23	.915	.642	.791	.746	0.09	.780	0.09
176.02	0.0	C	-	-	-	.563	16	.915	.642	.791	.796	0.07	.832	0.07
177.01	33.3	M	-	-	-	.221	27	.935	.642	.791	.354	0.09	.366	0.09
177.02	66.7	M	-	-	-	.099	27	.935	.642	.791	.162	0.08	.168	0.09
178.00	0.0	C	-	M	X	.211	176	.911	.647	.752	.364	0.49	.381	0.50
179.01	0.0	C	-	-	-	.027	12	.947	.642	.907	.038	0.05	.039	0.05
179.02	0.0	C	-	-	-	.711	15	.947	.642	.907	.906	0.08	.931	0.08
179.03	0.0	C	-	-	-	.226	9	.947	.642	.907	.312	0.04	.320	0.04
179.04	0.0	C	-	-	-	.336	10	.947	.642	.907	.458	0.04	.471	0.05
179.05	0.0	C	-	-	-	.793	7	.947	.642	.907	.991	0.04	1.019	0.04
179.06	0.0	C	-	-	-	.141	15	.947	.642	.907	.196	0.06	.201	0.07
179.07	0.0	C	-	-	-	.396	14	.947	.642	.907	.536	0.06	.550	0.07
179.08	0.0	M	-	-	-	.477	51	.895	.642	.791	.708	0.19	.749	0.19
179.09	0.0	M	-	-	-	.363	64	.895	.642	.791	.568	0.22	.600	0.22
179.10	0.0	M	-	-	-	.480	23	.895	.642	.791	.712	0.09	.752	0.09
179.11	0.0	M	-	-	-	.273	26	.895	.642	.791	.442	0.08	.467	0.08
179.12	0.0	M	-	-	-	.530	20	.895	.642	.752	.788	0.08	.832	0.08
179.13	0.0	M	-	-	-	.393	52	.895	.642	.752	.634	0.17	.670	0.17
179.14	0.0	M	-	-	-	.443	19	.895	.642	.752	.694	0.06	.734	0.06
181.01	0.0	-	-	-	-	.120	280	.869	.642	.776	.212	0.75	.227	0.73
181.02	0.0	-	-	-	-	.250	303	.869	.642	.776	.425	0.87	.456	0.85
181.03	0.0	-	-	-	-	.300	70	.869	.642	.776	.499	0.21	.536	0.20
181.04	0.0	-	-	-	-	.100	110	.869	.642	.776	.177	0.29	.190	0.28
182.00	0.0	-	-	-	-	.090	200	.895	.642	.776	.156	0.55	.165	0.56
183.00	0.0	-	-	-	-	.070	256	.895	.642	.752	.127	0.65	.134	0.65
184.01	0.0	-	-	-	-	.220	152	.935	.642	.721	.390	0.40	.403	0.42
184.02	0.0	-	-	-	-	.040	26	.935	.642	.721	.074	0.06	.076	0.07
184.03	0.0	-	-	-	-	.310	89	.920	.642	.741	.520	0.26	.542	0.27

Tabelle E.33
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
185.00	0.0	-	-	M	-	.288	90	.895	.642	.776	.473	0.28	.500	0.28
186.00	0.0	-	-	-	-	.320	167	.895	.642	.776	.519	0.53	.549	0.53
189.01	0.0	C	-	-	-	.432	102	.895	.639	.752	.683	0.34	.722	0.34
189.02	0.0	C	-	-	-	.338	403	.895	.639	.791	.535	1.34	.566	1.34
189.03	0.0	C	-	-	-	.437	282	.895	.639	.791	.663	1.02	.700	1.02
189.04	0.0	C	-	-	-	.377	87	.895	.639	.791	.587	0.30	.621	0.30
189.05	0.0	C	-	-	-	.518	68	.895	.639	.791	.755	0.27	.798	0.27
189.06	0.0	C	-	-	-	.501	731	.895	.639	.791	.736	2.82	.778	2.82
189.07	0.0	C	-	-	-	.437	198	.895	.639	.791	.663	0.72	.700	0.72
189.08	0.0	C	-	-	-	.372	112	.895	.639	.791	.581	0.38	.614	0.38
189.09	0.0	C	-	-	-	.398	351	.895	.639	.791	.614	1.23	.649	1.23
189.10	0.0	C	-	-	-	.317	394	.895	.639	.791	.506	1.29	.535	1.29
189.11	0.0	C	-	-	-	.218	81	.895	.639	.791	.359	0.25	.380	0.25
189.12	0.0	C	-	-	-	.368	191	.895	.639	.791	.575	0.65	.608	0.65
189.13	0.0	C	-	-	-	.368	97	.895	.639	.791	.575	0.33	.608	0.33
189.14	0.0	C	-	-	-	.616	617	.895	.639	.791	.852	2.69	.900	2.69
189.15	0.0	C	-	-	-	.475	418	.895	.639	.791	.707	1.57	.748	1.57
189.16	0.0	C	-	-	-	.428	412	.895	.639	.791	.652	1.48	.689	1.48
189.17	0.0	C	-	-	-	.394	243	.895	.639	.791	.609	0.85	.644	0.85
189.18	0.0	C	-	-	-	.330	98	.895	.639	.791	.524	0.32	.554	0.32
189.19	0.0	C	-	-	-	.561	287	.895	.639	.791	.799	1.18	.845	1.18
189.20	0.0	C	-	-	-	.411	291	.895	.639	.791	.631	1.03	.667	1.03
189.21	0.0	C	-	-	-	.278	184	.895	.639	.791	.450	0.59	.476	0.59
189.22	0.0	C	-	-	-	.385	84	.895	.639	.791	.598	0.29	.632	0.29
189.23	0.0	C	-	-	-	.214	158	.895	.639	.791	.353	0.48	.373	0.48
189.24	0.0	C	-	-	-	.300	119	.895	.639	.791	.482	0.38	.509	0.38
189.25	0.0	C	-	-	-	.381	67	.895	.639	.791	.593	0.23	.626	0.23
189.26	0.0	C	-	-	-	.407	57	.895	.639	.791	.626	0.20	.661	0.20
189.27	0.0	C	-	-	-	.377	81	.895	.639	.776	.598	0.27	.632	0.27
189.28	0.0	C	-	-	-	.479	181	.895	.639	.791	.712	0.68	.752	0.68
189.29	0.0	C	-	-	-	.347	314	.895	.639	.791	.547	1.05	.579	1.05
189.30	0.0	C	-	-	-	.407	454	.895	.639	.791	.626	1.60	.661	1.60
189.31	0.0	C	-	-	-	.501	380	.895	.639	.791	.736	1.47	.778	1.47
189.32	0.0	C	-	-	-	.381	412	.895	.639	.791	.593	1.42	.626	1.42
189.33	0.0	C	-	-	-	.488	391	.895	.639	.791	.722	1.49	.763	1.49
189.34	0.0	C	-	-	-	.321	48	.895	.639	.791	.512	0.16	.541	0.16
189.35	0.0	C	-	-	-	.390	58	.895	.639	.752	.631	0.18	.667	0.18
189.36	0.0	C	-	-	-	.424	57	.895	.639	.752	.673	0.19	.712	0.19
189.37	0.0	C	-	-	-	.223	87	.895	.639	.752	.390	0.24	.413	0.24
189.38	0.0	C	-	-	-	.334	56	.895	.639	.791	.530	0.19	.560	0.19
189.39	0.0	C	-	-	-	.248	48	.895	.639	.791	.405	0.15	.428	0.15
189.40	0.0	C	-	-	-	.240	61	.895	.639	.791	.393	0.19	.416	0.19
189.41	0.0	C	-	-	-	.214	104	.895	.639	.791	.353	0.32	.373	0.32
189.42	0.0	C	-	-	-	.227	113	.895	.639	.791	.373	0.35	.395	0.35
189.43	0.0	C	-	-	-	.338	38	.895	.639	.791	.535	0.13	.566	0.13
189.44	0.0	C	-	-	-	.295	21	.895	.639	.791	.475	0.07	.502	0.07
189.45	0.0	C	-	-	-	.390	178	.895	.639	.752	.631	0.57	.667	0.57
189.46	0.0	C	-	-	-	.432	347	.895	.639	.752	.683	1.16	.722	1.16
189.47	0.0	C	-	-	-	.454	37	.895	.639	.752	.708	0.13	.749	0.13
189.48	0.0	C	-	-	-	.424	41	.895	.639	.752	.673	0.14	.712	0.14
189.49	0.0	C	-	-	-	.595	104	.895	.639	.752	.849	0.43	.897	0.43
189.50	0.0	C	-	-	-	.454	174	.895	.639	.752	.708	0.60	.749	0.60
189.51	0.0	C	-	-	-	.257	104	.895	.639	.791	.419	0.33	.443	0.33
189.52	0.0	C	-	-	-	.509	32	.895	.639	.752	.768	0.12	.811	0.12
189.53	0.0	C	-	-	-	.574	268	.895	.639	.791	.812	1.12	.858	1.12
189.54	0.0	C	-	-	-	.432	212	.895	.639	.791	.657	0.76	.694	0.77
189.55	0.0	C	-	-	-	.488	98	.895	.639	.791	.722	0.37	.763	0.37
189.56	0.0	C	-	-	-	.484	87	.895	.639	.791	.717	0.33	.758	0.33
189.57	0.0	C	-	-	-	.424	106	.895	.639	.791	.647	0.38	.684	0.38

Tabelle E.33
(Fortsetzung)

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
189.58	0.0	C	-	-	-	.287	124	.895	.639	.791	.463	0.40	.489	0.40
189.59	0.0	C	-	-	-	.428	168	.895	.639	.791	.652	0.60	.689	0.60
189.60	0.0	C	-	-	-	.372	117	.895	.639	.791	.581	0.40	.614	0.40
189.61	0.0	C	-	-	-	.407	198	.895	.639	.791	.626	0.70	.661	0.70
189.62	0.0	C	-	-	-	.390	314	.895	.639	.791	.604	1.09	.639	1.09
189.63	0.0	C	-	-	-	.420	217	.895	.639	.791	.642	0.77	.679	0.77
189.64	0.0	C	-	-	-	.462	208	.895	.639	.791	.692	0.77	.732	0.77
189.65	0.0	C	-	-	-	.458	197	.895	.639	.791	.688	0.73	.727	0.73
189.66	0.0	C	-	-	-	.458	128	.895	.639	.791	.688	0.47	.727	0.47
189.67	0.0	C	-	-	-	.471	117	.895	.639	.791	.703	0.44	.743	0.44
189.68	0.0	C	-	-	-	.462	89	.895	.639	.791	.692	0.33	.732	0.33
189.69	0.0	C	-	-	-	.492	58	.895	.639	.791	.726	0.22	.768	0.22
189.70	0.0	C	-	-	-	.492	87	.895	.639	.752	.750	0.31	.793	0.31

Anmerkungen. siehe Seite 275.

Tabelle E.34
Lernleistung, Weiterbildung

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
100.00	60.0	-	-	C	-	.323	76	.911	.820	1.003	.356	8.5	.373	8.7
108.00	90.9	M	-	M	-	.076	249	.911	.642	.752	.136	10.6	.142	10.9
110.00	0.0	M	X	-	-	.476	75	.911	.642	.752	.727	4.4	.761	4.5
114.00	0.0	-	-	M	-	.335	313	.911	.647	.752	.549	15.7	.576	16.2
117.00	0.0	-	-	-	-	.700	42	.863	.642	.791	.923	3.3	.993	3.2
120.01	0.0	-	-	-	-	.530	127	.863	.642	.752	.796	7.6	.856	7.4
120.02	0.0	-	-	-	-	.600	101	.863	.642	.752	.857	6.7	.923	6.5
120.03	0.0	-	-	-	-	.570	222	.863	.642	.844	.786	15.8	.846	15.4
120.04	0.0	-	-	-	-	.590	13	.863	.639	.752	.850	0.9	.915	0.8
123.01	0.0	-	-	-	-	.710	20	.895	.520	.791	.988	1.4	1.044	1.4
123.02	0.0	-	-	-	-	.650	20	.895	.520	.791	.939	1.3	.993	1.3
123.03	0.0	-	-	-	-	.800	19	.895	.520	.791	1.050	1.5	1.110	1.5
123.04	0.0	-	-	-	-	.720	7	.895	.520	.791	.995	0.5	1.052	0.5
123.05	0.0	-	-	-	-	.590	13	.895	.520	.791	.885	0.8	.936	0.8
123.06	0.0	-	-	-	-	.850	10	.895	.520	.791	1.080	0.8	1.142	0.9
136.00	0.0	-	-	-	-	.360	97	.869	.642	.776	.583	5.0	.625	4.9
141.00	0.0	-	-	-	-	.750	34	.863	.642	.791	.959	2.8	1.032	2.8
142.00	0.0	-	-	-	-	.750	8	.863	.642	.752	.962	0.7	1.035	0.6
148.00	0.0	C	-	-	-	.521	91	.869	.642	.700	.817	5.0	.876	4.9
173.00	0.0	-	-	-	-	.260	190	.863	.642	.700	.505	6.8	.543	6.7

Anmerkungen. siehe Seite 275.

Tabelle E.35
Arbeitsleistung

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
101.01	0.0	–	–	–	–	.150	72	.863	.520	.467	.710	0.4	.764	0.3
101.02	0.0	–	–	–	–	.170	65	.863	.520	.489	.694	0.4	.747	0.4
109.00	0.0	C	–	–	–	.040	102	.882	.520	.752	.082	2.6	.087	2.6
115.00	0.0	–	–	M	–	.543	103	.863	.520	.791	.843	4.6	.908	4.5
116.02	0.0	C	–	–	–	.354	75	.882	.520	.791	.611	2.7	.651	2.7
116.05	0.0	C	–	–	–	.449	73	.882	.520	.791	.736	2.9	.783	2.9
116.07	0.0	C	–	–	–	.707	70	.882	.520	.752	.984	3.9	1.048	3.9
116.12	0.0	–	–	–	–	.710	597	.863	.520	.752	.984	33.4	1.059	32.8
128.00	0.0	–	–	–	–	.150	144	.895	.520	.752	.297	4.0	.314	4.0
132.00	80.4	C	X	M	–	.192	70	.825	.520	.776	.382	1.9	.420	1.8
135.00	0.0	–	–	M	–	.185	913	.895	.520	.791	.339	29.1	.359	29.7
136.00	0.0	–	–	M	X	.170	98	.895	.520	.776	.321	3.0	.339	3.0
162.01	0.0	–	–	–	–	.210	20	.895	.520	.791	.382	0.7	.404	0.7
162.02	0.0	–	–	–	–	.280	11	.895	.520	.791	.497	0.4	.525	0.4
170.00	0.0	–	–	–	–	.004	110	.895	.520	.776	.008	3.2	.008	3.2
172.00	0.0	–	–	–	–	.140	42	.869	.520	.791	.265	1.3	.284	1.2
187.01	0.0	–	–	C	–	.297	98	.895	.520	.776	.533	3.3	.564	3.3
187.02	0.0	–	–	C	–	.303	76	.895	.520	.776	.542	2.6	.573	2.6

Anmerkungen. siehe Seite 275.

Tabelle E.36
Arbeitsleistung, publizierte Studien

PS-Nr.	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
	Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
101.01	–	–	–	–	.150	72	.863	.520	.467	.710	0.4	.764	0.4
101.02	–	–	–	–	.170	65	.863	.520	.489	.694	0.5	.747	0.4
109.00	C	–	–	–	.040	102	.882	.520	.752	.082	2.8	.087	2.8
115.00	–	–	M	–	.543	103	.863	.520	.791	.843	4.9	.908	4.8
116.02	C	–	–	–	.354	75	.882	.520	.791	.611	2.9	.651	2.9
116.05	C	–	–	–	.449	73	.882	.520	.791	.736	3.1	.783	3.1
116.07	C	–	–	–	.707	70	.882	.520	.752	.984	4.1	1.048	4.2
116.12	–	–	–	–	.710	597	.863	.520	.752	.984	35.6	1.059	34.9
128.00	–	–	–	–	.150	144	.895	.520	.752	.297	4.2	.314	4.3
135.00	–	–	M	n	.185	913	.895	.520	.791	.339	31.1	.359	31.6
136.00	–	–	M	X	.170	98	.895	.520	.776	.321	3.2	.339	3.2
162.01	–	–	–	–	.210	20	.895	.520	.791	.382	0.7	.404	0.7
162.02	–	–	–	–	.280	11	.895	.520	.791	.497	0.4	.525	0.4
187.01	–	–	C	–	.297	98	.895	.520	.776	.533	3.5	.564	3.5
187.02	–	–	C	–	.303	76	.895	.520	.776	.542	2.7	.573	2.8

Anmerkungen. siehe Seite 275.

Tabelle E.37
Arbeitsleistung, prädiktive Validität

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
101.01	0.0	–	–	–	–	.150	72	.863	.520	.467	.710	0.4	.764	0.4
101.02	0.0	–	–	–	–	.170	65	.863	.520	.489	.694	0.5	.747	0.5
109.00	0.0	C	–	–	–	.040	102	.882	.520	.752	.082	3.0	.087	3.0
115.00	0.0	–	–	M	–	.543	103	.863	.520	.791	.843	5.3	.908	5.2
116.02	0.0	C	–	–	–	.354	75	.882	.520	.791	.611	3.1	.651	3.1
116.05	0.0	C	–	–	–	.449	73	.882	.520	.791	.736	3.3	.783	3.4
116.07	0.0	C	–	–	–	.707	70	.882	.520	.752	.984	4.4	1.048	4.5
116.12	0.0	–	–	–	–	.710	597	.863	.520	.752	.984	38.2	1.059	37.6
132.00	80.4	C	X	M	–	.192	70	.825	.520	.776	.382	2.2	.420	2.1
135.00	0.0	–	–	M	n	.185	913	.895	.520	.791	.339	33.4	.359	34.1
162.01	0.0	–	–	–	–	.210	20	.895	.520	.791	.382	0.7	.404	0.8
162.02	0.0	–	–	–	–	.280	11	.895	.520	.791	.497	0.4	.525	0.4
170.00	0.0	–	–	–	–	.004	110	.895	.520	.776	.008	3.6	.008	3.7
172.00	0.0	–	–	–	–	.140	42	.869	.520	.791	.265	1.4	.284	1.4

Anmerkungen. siehe Seite 275.

Tabelle E.38
Arbeitsleistung, nur Pearson-Korrelationen

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
101.01	0.0	–	–	–	–	.150	72	.863	.520	.467	.710	0.4	.764	0.4
101.02	0.0	–	–	–	–	.170	65	.863	.520	.489	.694	0.4	.747	0.4
109.00	0.0	C	–	–	–	.040	102	.882	.520	.752	.082	2.7	.087	2.7
116.02	0.0	C	–	–	–	.354	75	.882	.520	.791	.611	2.8	.651	2.8
116.05	0.0	C	–	–	–	.449	73	.882	.520	.791	.736	3.1	.783	3.1
116.07	0.0	C	–	–	–	.707	70	.882	.520	.752	.984	4.1	1.048	4.1
116.12	0.0	–	–	–	–	.710	597	.863	.520	.752	.984	35.0	1.059	34.3
128.00	0.0	–	–	–	–	.150	144	.895	.520	.752	.297	4.1	.314	4.2
132.00	80.4	C	X	M	–	.192	70	.825	.520	.776	.382	2.0	.420	1.9
135.00	0.0	–	–	M	n	.185	913	.895	.520	.791	.339	30.5	.359	31.1
136.00	0.0	–	–	M	X	.170	98	.895	.520	.776	.321	3.1	.339	3.2
162.01	0.0	–	–	–	–	.210	20	.895	.520	.791	.382	0.7	.404	0.7
162.02	0.0	–	–	–	–	.280	11	.895	.520	.791	.497	0.4	.525	0.4
170.00	0.0	–	–	–	–	.004	110	.895	.520	.776	.008	3.3	.008	3.4
172.00	0.0	–	–	–	–	.140	42	.869	.520	.791	.265	1.3	.284	1.3
187.01	0.0	–	–	C	–	.297	98	.895	.520	.776	.533	3.4	.564	3.5
187.02	0.0	–	–	C	–	.303	76	.895	.520	.776	.542	2.7	.573	2.7

Anmerkungen. siehe Seite 275.

Tabelle E.39
Arbeitsleistung ohne Studie 116.12

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
101.01	0.0	–	–	–	–	.150	72	.863	.520	.467	.710	0.5	.764	0.5
101.02	0.0	–	–	–	–	.170	65	.863	.520	.489	.694	0.6	.747	0.6
109.00	0.0	C	–	–	–	.040	102	.882	.520	.752	.082	3.9	.087	3.9
115.00	0.0	–	–	M	–	.543	103	.863	.520	.791	.843	6.9	.908	6.7
116.02	0.0	C	–	–	–	.354	75	.882	.520	.791	.611	4.1	.651	4.0
116.05	0.0	C	–	–	–	.449	73	.882	.520	.791	.736	4.4	.783	4.4
116.07	0.0	C	–	–	–	.707	70	.882	.520	.752	.984	5.8	1.048	5.8
128.00	0.0	–	–	–	–	.150	144	.895	.520	.752	.297	5.9	.314	6.0
132.00	80.4	C	X	M	–	.192	70	.825	.520	.776	.382	2.9	.420	2.7
135.00	0.0	–	–	M	n	.185	913	.895	.520	.791	.339	43.7	.359	44.1
136.00	0.0	–	–	M	X	.170	98	.895	.520	.776	.321	4.4	.339	4.5
162.01	0.0	–	–	–	–	.210	20	.895	.520	.791	.382	1.0	.404	1.0
162.02	0.0	–	–	–	–	.280	11	.895	.520	.791	.497	0.6	.525	0.6
170.00	0.0	–	–	–	–	.004	110	.895	.520	.776	.008	4.7	.008	4.8
172.00	0.0	–	–	–	–	.140	42	.869	.520	.791	.265	1.9	.284	1.9
187.01	0.0	–	–	C	–	.297	98	.895	.520	.776	.533	4.9	.564	4.9
187.02	0.0	–	–	C	–	.303	76	.895	.520	.776	.542	3.8	.573	3.9

Anmerkungen. siehe Seite 275.

Tabelle E.40
Einkommen

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
132.00	81.8	C	X	–	–	.285	70	.825	–	.776	.406	3.0	.447	2.8
135.00	0.0	–	–	–	–	.330	767	.895	–	.791	.429	40.0	.453	39.5
145.00	0.0	–	–	M	–	.250	295	.913	–	.776	.335	14.5	.350	14.6
151.00	0.0	–	–	–	–	.101	200	.920	–	.791	.134	10.0	.140	10.2
165.00	0.0	–	–	C	–	.027	311	.935	–	.776	.036	15.0	.038	15.5
169.00	0.0	–	–	–	–	.441	207	.913	–	.776	.559	11.3	.585	11.4
188.00	0.0	–	–	–	–	.175	130	.895	–	.776	.241	6.1	.255	6.0

Anmerkungen. siehe Seite 275.

Tabelle E.41
Berufliches Vorankommen

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
132.00	75.0	C	X	M	–	.176	70	.825	–	.776	.259	3.2	.285	2.9
135.00	0.0	–	–	–	–	.390	643	.895	–	.791	.499	38.3	.527	38.1
145.00	100.0	–	–	–	–	.022	295	.913	–	.776	.030	15.2	.032	15.4
151.00	0.0	C	–	–	–	.084	192	.920	–	.791	.112	10.6	.116	10.8
163.00	0.0	C	X	–	–	.006	49	.895	–	.776	.008	2.5	.009	2.4
169.00	0.0	–	–	–	–	.286	207	.913	–	.776	.380	11.5	.397	11.6
170.00	0.0	–	–	–	–	.251	109	.895	–	.776	.340	5.8	.360	5.7
185.00	0.0	–	–	–	–	.240	116	.895	–	.776	.326	6.1	.345	6.1
188.00	0.0	–	–	C	–	.157	136	.895	–	.776	.217	7.0	.229	6.9

Anmerkungen. siehe Seite 275.

Tabelle E.42

Berufliches Vorankommen, nur Pearson-Korrelationen

PS-Nr.	n.ber. (%)	Aggregat X		Aggregat Y		r	n	Artefakt-Werte			op. Val.		vollst. korr.	
		Art	Gew.	Art	Gew.			r_{xx}	r_{yy}	u_x	$\hat{\rho}$	w (%)	$\hat{\rho}$	w (%)
132.00	75.0	C	X	M	–	.176	70	.825	–	.776	.259	3.5	.285	3.3
135.00	0.0	–	–	–	–	.390	643	.895	–	.791	.499	42.9	.527	42.7
145.00	100.0	–	–	–	–	.022	295	.913	–	.776	.030	17.0	.032	17.3
163.00	0.0	C	X	–	–	.006	49	.895	–	.776	.008	2.7	.009	2.7
169.00	0.0	–	–	–	–	.286	207	.913	–	.776	.380	12.8	.397	13.0
170.00	0.0	–	–	–	–	.251	109	.895	–	.776	.340	6.5	.360	6.4
185.00	0.0	–	–	–	–	.240	116	.895	–	.776	.326	6.8	.345	6.8
188.00	0.0	–	–	C	–	.157	136	.895	–	.776	.217	7.8	.229	7.8

Anmerkungen. siehe Seite 275.