

# Freie Universität Berlin

**Fachbereich Erziehungswissenschaft und Psychologie  
Freie Universität Berlin**

**Intelligenz und Wissen  
Schlüsselkompetenzen im Kontext personalpolitischer  
Eignungsdiagnostik  
Validierungsstudie des Wissenstests START-W für  
Berufseinsteiger**

Dissertation  
zur Erlangung des akademischen Grades  
Doktor der Philosophie (Dr. phil.)  
Doctor of Philosophy (Ph.D.)

vorgelegt von  
Diplom-Kaufmann  
**Christian Speckemeier**

Berlin, 2011

Erstgutachter:  
Herr Prof. Dr. Detlev Liepmann  
Freie Universität Berlin

Zweitgutachter:  
PD Dr. Hans-Uwe Hohner  
Freie Universität Berlin

Tag der Disputation: 24.05.2011

## **Danksagung**

Eine Vielzahl von Menschen haben mich auf dem Weg zur Doktorwürde begleitet, unterstützt und haben in unterschiedlicher Weise zum Gelingen der Arbeit beigetragen.

Mein ganz besonderer Dank gilt meinem Doktorvater (im wahrsten Sinne des Wortes) Herrn Prof. Dr. Detlev Liepmann, der mir nicht nur in fachspezifischen, sondern auch in privaten Gesprächen immer konstruktiv und offen zur Seite stand. Ohne die jederzeit angenehme, ermutigende, inspirierende, finanzielle und doch freiheitsgewährende Betreuung wären die wissenschaftlichen Arbeiten nicht möglich gewesen.

Im Hinblick auf die empirischen Datenerhebungen danke ich -in Erinnerung- Doris Stallmann. Zusammen mit meiner Mutter Ellen Speckemeier ebneten sie mir die Wege zur Untersuchung von Auszubildenden in den Schulen für Gesundheitsberufe.

In diesem Zusammenhang danke ich auch meinen Kooperationspartnern Herrn Söding, Frau Brinkmann-Dagci und Frau Hamann (Vivantes Instituts für berufliche Bildung), Frau Rösch (Charité- Gesundheitsakademie), Frau Vogler (Wannseeschule e.V.), Frau Schürmann sowie Herrn Schröder (Marie-Curie Gymnasium) für die außerordentlich sympathische, kompetente, tatkräftige und unkomplizierte Zusammenarbeit. Mit Bezug auf die Forschungsprojekte –die meine Liquidität sichergestellt haben– möchte ich mich zusätzlich bei Frau Dr. Brandt (Berufsförderungswerk- Brandenburg), Herrn Böschenbröker (Berufsförderungswerk-Berlin) und allen anderen Helfern bedanken.

Ich danke Dr. Kathrin Heinitz, Dr. Lars Michael, Dr. Jens Eisermann, Frau Liepmann, Stefanie Giljohann, Nicole Torjus, Stefanie Rieser für die angenehme Arbeitsatmosphäre, die zahlreichen guten Ratschläge, Ideen, Anregungen und ihre konstruktive Kritik.

Für die fachpraktische, interessierte Begleitung, die aufmerksame Beobachtung und auch die Redigierarbeit bin ich meinem Vater Heribert Speckemeier sehr dankbar. Meinem ehemaligen Deutschlehrer Herrn Lübbe danke ich für das Korrekturlesen und meinem Bruder Arne Speckemeier für den „letzten Schliff“ und die Unterstützung bei den Forschungsprojekten.

Für die kurzfristige, spontane Hilfsbereitschaft und Unterstützung danke ich PD Dr. Hans-Uwe Hohner für die Übernahme des Zweitgutachtens sowie Prof. Dr. Ernst-H. Hoff, Prof. Dr. Felfe und Dr. Kathrin Heinitz, die sich bereit erklärt haben, Mitglied der Promotionskommission zu sein.

Ich möchte mich auch bei meinen Freunden und meiner Freundin bedanken, die ein wichtiger Teil meines Lebens sind und stets für die erforderliche Abwechslung und Entspannung sorgten.

Ein großer, liebevoller und bedeutungsvoller Dank geht natürlich an meine Familie (Heribert, Ellen und Arne Speckemeier), die mich nicht nur in finanzieller Hinsicht unterstützt haben, sondern mir in jeglicher Hinsicht den Rücken frei gehalten haben. Ihr seid das Fundament für alles, was ich bisher erreicht habe und noch erreichen werde.

Schließlich sei all denen ein Dankeschön ausgesprochen, die nicht namentlich Erwähnung fanden, aber zum Gelingen dieser Arbeit beigetragen haben.

---

# Inhaltsverzeichnis

	Seite
1 Einleitung .....	1
2 Schlüsselqualifikationen im Kontext berufsbezogener Eignungsdiagnostik .....	8
2.1 Zum Konzept der Schlüsselqualifikationen- Herkunft des Begriffes .....	9
2.2 Konzepte zur Bestimmung der Schlüsselqualifikationen .....	12
2.2.1 Theoriebasierte Konzepte zur Identifikation und Relevanz von Schlüsselkompetenzen .....	12
2.2.2 Empirisch orientierte Konzepte zur Identifikation von Schlüsselkompetenzen .....	15
2.2.3 Bewertung der theoretischen und empirischen Konzepte .....	20
2.3 Ansätze zur Kategorisierung der Schlüsselqualifikationen .....	22
2.3.1 Ausgewählte Konzepte zur Klassifizierung der Schlüsselqualifikationen .....	22
2.3.2 Bewertung der Ansätze zur Kategorisierung der Schlüsselqualifikationen .....	26
2.4 Der Kompetenzbegriff - Weiterentwicklung des Konzepts der Schlüsselqualifikationen .....	28
2.4.1 Kompetenz- eine Begriffsbestimmung .....	29
2.4.2 Bewertung des Kompetenzkonzepts im Hinblick auf das Konzept der Schlüsselqualifikationen .....	34
2.4.3 Ausblick auf die Bedeutung der Schlüsselkompetenzen bzw. Schlüsselqualifikationen .....	35
3 Psychologische Diagnostik .....	37
3.1 Abgrenzung und Zielsetzung des Gegenstandsbereiches.....	38
3.2 Testverfahren im Rahmen der personalpolitischen Eignungsdiagnostik.....	42
3.2.1 Klassifizierung eignungsdiagnostischer Testverfahren .....	43
3.2.2 Der Testbegriff und die klassische Testtheorie.....	46
3.2.3 Leistungstests in der Eignungsdiagnostik.....	49
3.2.4 Gütekriterien der formalen Diagnostik .....	55

---

3.3	Methodische Konzepte zur Diagnostik von Schlüsselkompetenzen .....	59
3.4	Die Bedeutung von Leistungstests im Human Resource Management.....	62
3.4.1	Einsatzhäufigkeiten von Leistungstests.....	64
3.4.2	Internationale Forschungsbeiträge und Befunde zur Relevanz von Intelligenz- und Wissenstests.....	68
3.4.3	Anwendungsfelder kognitiver Fähigkeitstests.....	81
3.4.4	Exkurs: Betriebswirtschaftliche Bedeutung eignungsdiagnostischer Testverfahren.....	83
4	Intelligenz- Theorien, Befunde und Bedeutung.....	89
4.1	Intelligenzforschung im Überblick.....	89
4.1.1	Eingliederung der Intelligenzmessung in den psychologischen Kontext.....	89
4.1.2	Geschichtlicher Verlauf der Intelligenzmessung.....	90
4.1.3	Abschlussbetrachtung und Ausblick der Intelligenzforschung .....	96
4.2	Begriffsbestimmung der Intelligenz .....	97
4.2.1	Definition von Intelligenz .....	98
4.2.2	Intelligenz als Konstrukt .....	100
4.2.3	Zusammenhang zwischen dem Konstrukt-Verständnis und der Intelligenzmessung .....	103
4.3	Bedeutungsbereiche der Intelligenz.....	103
4.3.1	Methodische Ansätze der Intelligenzforschung .....	104
4.3.2	Die akademische Intelligenz als theorieübergreifender Bedeutungsbereich.....	106
4.4	Ausgewählte Strukturtheorien der Intelligenz .....	110
4.4.1	Die Zwei-Faktoren-Theorie von Spearman .....	111
4.4.2	Thurstones Primärfaktoren-Modell.....	112
4.4.3	Das Berliner Intelligenzstrukturmodell (BIS).....	115
4.4.4	Das Modell der fluiden und kristallisierten Intelligenz von Cattell.....	118
4.4.5	Abschließende Betrachtung und Bewertung der Strukturtheorien .....	122

---

5	Wissen- Konzepte, Befunde und Bedeutung .....	123
5.1	Die Taxonomie der Wissensarten .....	124
5.2	Wissenspsychologische Paradigmen und Aspekte .....	128
5.3	Wissenserwerb als Lernprozess .....	132
5.4	Ausblick- Wissen und seine strategische Relevanz .....	136
6	Wissen und Intelligenz – Eine konzeptionelle Zusammenführung .....	138
6.1	Grundlegende Zusammenhänge zwischen Intelligenz und Wissen .....	138
6.2	Das Beziehungsgeflecht zwischen fluider Intelligenz, kristallisierter Intelligenz und Wissen .....	145
6.3	Die Interaktion zwischen fluider Intelligenz und kristallisierter Intelligenz – Der Einfluss von Anlage und Umwelt .....	148
7	Überleitung zum empirischen Teil .....	160
8	Empirische Forschungsfragen und Forschungsannahmen .....	162
9	Darstellung der Forschungsmethode .....	170
9.1	Beschreibung der eingesetzten Testinstrumente .....	170
9.1.1	Intelligenz-Struktur-Test 2000 R (I-S-T- 2000 R) .....	170
9.1.2	START-W Wissenstest für Berufseinsteiger .....	175
9.2	Darstellung des Untersuchungsablaufs .....	177
9.3	Darstellung der Stichprobe .....	180
9.3.1	Datenvorbereitung .....	181
9.3.2	Deskriptive Statistik .....	182
9.4	Methodische Grundvoraussetzungen für die Datenanalyse .....	192
9.5	Methodische Vorgehensweise bei der Datenanalyse .....	194
10	Darstellung der Ergebnisse .....	196
10.1	Item- und Reliabilitätsanalysen .....	196
10.2	Darstellung der Ergebnisse zu den Forschungsfragen- Teil 1 .....	207
10.3	Darstellung der Ergebnisse zu den Forschungsfragen- Teil 2 .....	220
10.4	Darstellung der Ergebnisse zu den Forschungsfragen- Teil 3 .....	225

---

11	Diskussion der Ergebnisse .....	231
11.1	Item- und Reliabilitätsanalysen.....	231
11.2	Diskussion der Befunde zu den Forschungsfragen- Teil 1.....	235
11.3	Diskussion der Befunde zu den Forschungsfragen- Teil 2.....	247
11.4	Diskussion der Befunde zu den Forschungsfragen- Teil 3.....	252
12	Fazit und Ausblick .....	258
	Literaturverzeichnis .....	264
	Abbildungsverzeichnis .....	290
	Tabellenverzeichnis .....	291
	Verzeichnis des Anhangs .....	293



## 1 Einleitung

Wissen und Allgemeinbildung besitzen in unserer Kultur und Gesellschaft einen hohen Stellenwert. Sie erhalten eine besondere Bedeutung und Relevanz, weil sie als Garanten für den beruflichen, schulischen, akademischen und privaten Erfolg angesehen werden. Mit der allgemein bekannten Formulierung von Francis Bacon (1561- 1626) „Wissen ist Macht“ wird die Bedeutung des Wissens treffend umschrieben und die Wichtigkeit nachdrücklich unterstrichen. Ein weiterer Grundbaustein für den angesprochenen Erfolg bildet die Intelligenz. Obgleich die Begriffe Intelligenz und Wissen unterschiedliche Leistungsdimensionen bzw. Untersuchungsgegenstände darstellen, werden sie im Alltag wie auch in der Wissenschaft häufig in Verbindung gebracht. Eine Systematisierung des Zusammenhangs zwischen Intelligenz und Wissen stellt die Investmenttheorie von Cattell (1971a, 1987) dar. In seiner Strukturtheorie unterscheidet er diesbezüglich zwischen fluider Intelligenz und kristallisierter Intelligenz. In der aktuellen Praxis bietet der Intelligenz-Struktur-Test 2000 R (I-S-T 2000R, Liepmann, Beauducel, Brocke und Amthauer, 2007) somit eine Möglichkeit, die Konstrukte im Rahmen der Eignungsdiagnostik zu operationalisieren und hinsichtlich ihrer Schnittmenge zu analysieren. Denn insbesondere in der praxisorientierten Eignungsdiagnostik werden Intelligenztests und Wissenstests oft miteinander kombiniert, um differenzierte Aussagen über den Leistungsstand und das Leistungspotential bei Personen treffen zu können.

In diesem Kontext können die beiden Konstrukte auch als extrafunktionale Qualifikationen (i.w.S.), Schlüsselqualifikationen bzw. Basiskompetenzen verstanden werden (Stangel-Meseke, 1994). Im Gegensatz zu den fachlichen, funktionalen Qualifikationen sollen die Schlüsselqualifikationen dazu dienen, neue Wissensgebiete zu erschließen, um flexibel und selbständig auf zukünftige Veränderungen reagieren zu können (Mertens, 1974). Vor dem Hintergrund der kontinuierlichen Veränderungen in der global vernetzten Marktwirtschaft wird die Relevanz von extrafunktionalen Qualifikationen hervorgehoben. Die Zunahme von technologischen Innovationen hat zur Folge, dass das spezifische Fachwissen schneller veraltet. Die Konsequenz dieser Entwicklungen ist darin zu sehen, dass neben den Fachqualifikationen in verstärktem Maß auch extrafunktionale Qualifikationen gefordert werden, wie z.B. fachübergreifendes (Allgemein-) Wissen (z.B. Liepmann, 1993, Sonntag, 1996). Begriffe wie Wissensmanagement oder *intellectual capital* (Buck, 2005) unterstreichen die Tatsache, dass die Komponente Wissen mittlerweile als eine der wichtigsten Produktionsfaktoren der Zukunft gilt

(Picot & Scheuble, 2000). Mit dem Titel „Auf dem Weg zur Wissensgesellschaft“ thematisiert die Heinrich-Böll Stiftung ebenfalls den Wandel von der Industriegesellschaft zur Wissensgesellschaft und bezeichnet Bildung, Wissen und Kreativität als Schlüsselressourcen (Heinrich-Böll Stiftung, 2009).

Obwohl dem Wissen und der Intelligenz eine hohe Wichtigkeit beigemessen wird, belegen neue Studien in der schulischen und beruflichen Praxis, dass viele Personen erhebliche Defizite in der Beherrschung von Bildungstechniken und hinsichtlich des Allgemeinwissens aufweisen. Der Grund für diese Entwicklung wird in dem unzureichenden Bildungssystem gesehen. Es wird kritisiert, dass vor allem die Wissensvermittlung an den Bildungseinrichtungen nicht effektiv gestaltet und der Wissenserwerb nicht gefördert wird. Die PISA-Studie, die unter der Schirmherrschaft der OECD durchgeführt wurde (PISA, 2000, 2006), zeigt beispielsweise auf, dass ein Großteil der untersuchten 15-jährigen deutschen Schüler im Vergleich zu anderen OECD-Staaten hinsichtlich mathematischer und naturwissenschaftlicher Kenntnisse nicht einmal den Leistungsdurchschnitt erreichen. Eine Umfrage der Industrie- und Handelskammer zu Berlin (IHK, 2007) an 2165 Ausbildungsbetrieben bestätigt die Ergebnisse der PISA-Studie: 68,4% der befragten Unternehmen geben an, dass die Schulabgänger ein unzureichendes Leistungsvermögen und Leistungsverhalten haben. Das sind fünf Prozentpunkte mehr als in der 2003 gestarteten Umfrage. Aufgrund dessen konnten auch nicht alle Ausbildungsplätze besetzt werden. Als Grundvoraussetzung für die Aufnahme eines Auszubildenden erwarten 95,4% der befragten Unternehmen zumindest ausreichende Kenntnisse im Lesen, Schreiben und Rechnen. Des Weiteren werden wirtschaftliche Kenntnisse (79,9%) und ein solides Allgemeinwissen (77,8%) verlangt. Vergleichbare Ergebnisse liefert auch die bundesweit angelegte Ausbildungsumfrage 2010 (Deutsche Industrie- und Handelskammer, 2010). Von Seiten der IHK wird resümierend der Schluss gezogen, dass die Diagnose des Allgemeinwissens und der Intelligenz als extrafunktionale Qualifikationen eine wichtige Rolle für die Unternehmen spielen. Im Rahmen des Projekts „Passgenaue Vermittlung“, welches vom Bundeswirtschaftsministerium initiiert wurde, wird die Forderung nach geeigneten und systematischen Auswahlverfahren bereits aufgegriffen. Darüber hinaus stellen Bildungslücken für die Unternehmen und Ausbilder einen erheblichen Kostenfaktor dar. Das Erstellen einer allgemeinen Wissensgrundlage zu Lehrbeginn erfordert zusätzliche Zeitressourcen und erhöht somit die Ausbildungskosten. Eine fundierte Wissensgrundlage entwickelt sich somit zunehmend zu einer bedeutenden Eingangsvoraussetzung für die berufliche Laufbahn. Unternehmer und

Psychologen haben erkannt, wie bedeutsam eine gute Eignungsdiagnostik für die Marktwirtschaft ist und ergänzen, dass diese beiden Faktoren (Intelligenz und Wissen) Grundvoraussetzungen für die Bewältigung neuartiger bzw. sich stetig wandelnder beruflicher Anforderungen darstellen. Dabei wird insbesondere das Wissen als Indikator für zukünftiges Lernverhalten gesehen. In diesem Zusammenhang hebt die Bundesanstalt für Arbeit (2002) ebenfalls die Wichtigkeit einer geeigneten Kompetenzfeststellung (Eignungsdiagnostik) der Jugendlichen hervor, um die vorhandenen Kenntnisse, Ressourcen und Potentiale optimal einsetzen zu können. Es sollen dabei adäquate Stellenangebote mit dem Leistungsvermögen der Jugendlichen verknüpft werden. Wenn dieser Ansatz berücksichtigt wird, können womöglich Frustrationen und Enttäuschungen der Jugendlichen aufgrund von Fehlplatzierungen ausgeschlossen und die Motivation zum Erwerb von beruflichen Qualifikationen beim Einzelnen aktiviert werden.

Vor diesem Hintergrund sind in erster Linie Testverfahren von hoher Relevanz, die die Eignung der Berufseinsteiger für die bestimmten Anforderungsprofile der unterschiedlichen Tätigkeiten gut diagnostizieren und prognostizieren können. Im Bezug auf das Konzept der extrafunktionalen Qualifikationen lässt sich festhalten, dass insbesondere die eignungsdiagnostische Erfassung des Allgemeinwissens und der Intelligenz von hoher Relevanz sind. Ein geeignetes Mittel zur Erfassung der genannten, übergeordneten Qualifikationen ist das Konzept der fluiden und kristallisierten Intelligenz, das diesem Vorhaben zu Grunde liegt (vgl. u.a. Liepmann et al., 2007). Das Konzept bzw. das Testverfahren eröffnet der zukünftigen Praxis der Intelligenzmessung innovative Perspektiven, da es sowohl die funktionalen als auch die extrafunktionalen Qualifikationen differenziert erfasst. Das übergeordnete Ziel solcher Verfahren ist es, auf wissenschaftlicher Basis den Leistungsstand und das Leistungspotential einer Person zu bestimmen bzw. einen Zusammenhang zwischen den Persönlichkeitsmerkmalen einer Person (hier Intelligenz und Wissen) und dem beruflichen Erfolg herzustellen. In der einschlägigen Literatur wird die berufliche Eignungsdiagnostik als „Methodologie der Entwicklung, Prüfung und Anwendung psychologischer Verfahren zum Zwecke eignungsbezogener Erfolgsprognosen und Entscheidungshilfen im beruflichen Kontext“ (Schuler & Höft, 2006, S.102) verstanden. Die Voraussetzung hierfür ist, dass das verwendete Verfahren auch die geforderten Kompetenzen wie z.B. das Allgemeinwissen und die Intelligenz erfassen kann. Diesbezüglich wird ein valides Verfahren verlangt, welches eine zuverlässige Prognose über den beruflichen Erfolg erlaubt. Als weitere Gütekriterien werden Objektivität und Reliabilität gefordert (vgl. z.B. Lienert & Raatz, 1998). Außerdem soll

zusätzlich die DIN-Norm 33430 im Besonderen hervorgehoben werden, die die Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen reglementiert und die Gütequalität absichert (Kersting, 2006). Im Hinblick auf die ca. 100-jährige Erfolgsgeschichte der Intelligenzdiagnostik lässt sich bereits im Vorgriff sagen, dass ein sehr hoher Bedarf an innovativen Verfahren bei Intelligenz- und Wissenstests liegt (Schorr, 1991). Zum gegenwärtigen Zeitpunkt zeigt die Praxis der eignungsdiagnostischen Tätigkeit, dass sich ein Aufwärtstrend bezüglich der Einsatzhäufigkeit von Intelligenz- und Wissenstests abzeichnet (Schuler & Höft, 2006). Das liegt nicht zuletzt darin begründet, dass psychologische Intelligenz- und Wissenstests den wissenschaftlichen Gütekriterien genügen, sondern auch hinsichtlich ihrer prognostischen Validität zu den validesten Auswahlinstrumenten zählen (Schuler, Frier & Kaufmann, 1993, S.26). Vor diesem Hintergrund verlangen insbesondere Wirtschaftsunternehmen ein Instrument, das vor allem kostengünstig und weniger zeitaufwändig ist, als z.B. Assessment Center oder eine Arbeitsprobe. Vor diesem Hintergrund gewinnt die START- Testbatterie für Berufsanfänger (Liepmann & Beauducel, in Vorbereitung) an Bedeutung. Die START- Testbatterie bezeichnet ein umfangreiches Verfahren, mit der grundlegende Kompetenzen beim Berufseinstieg abgebildet werden können. Es werden Aspekte erfasst, die die Personalarbeit bei Auswahl- und Platzierungsverfahren ergänzen sollen. Die einzelnen Verfahren können hierbei entsprechend der eignungsdiagnostischen Fragestellungen miteinander kombiniert werden.

Mit der vorliegenden Arbeit wird das Ziel verfolgt, einen allgemeinen Wissenstest für Berufseinsteiger (START-W) durch entsprechende Validierungsstudien weiter zu entwickeln. Der wissenschaftliche Fortschritt psychologisch fundierter Eignungsdiagnostik spielt hierbei eine übergeordnete Rolle. Obgleich das Wissen sehr komplex ist und nicht endgültig erfasst werden kann, bietet der START-W-Wissenstest die Möglichkeit das Allgemeinwissen einer Person abzubilden und liefert zudem einen Indikator für zukünftiges Lernverhalten, wie es von der Praxis gefordert wird. Neben der Erfassung relevanter Themenbereiche des Allgemeinwissens werden verbale, numerische und figurale Inhalte berücksichtigt und zudem wird eine Interpretation des Faktorwertes „kristallisierte Intelligenz“ ermöglicht. Der zu validierende Allgemeinwissenstest wird Bestandteil der umfassenden START-Testbatterie für Berufseinsteiger sein, die sogenannte Basiskompetenzen (z.B. Fremdsprachenkenntnisse, Rechenfertigkeiten, Konzentrationsfähigkeit, Allgemeinwissen, etc.) erfassen soll.

Der Aufbau und die Systematik der Arbeit werden wie folgt dargestellt:

Im ersten Teil der vorliegenden Arbeit werden Schlüsselqualifikationen bzw. Schlüsselkompetenzen im Kontext der berufsbezogenen Diagnostik und der beruflichen Aus- und Weiterbildung bestimmt, erörtert und in ihrer Bedeutung hervorgehoben. Den Ausgangspunkt bildet das Konzept der Schlüsselqualifikationen, welches im Rahmen der Arbeitswelt zunehmend an Bedeutung und Relevanz gewinnt. Einleitend wird eine Begriffsbestimmung von Schlüsselqualifikationen vorgenommen. Daran anknüpfend werden einschlägige theoretische und empirische Klassifikationsansätze vorgestellt, die wiederum ausschlaggebend für die weiteren Kategorisierungskonzepte sind. Im Weiteren wird der Kompetenzbegriff anhand von zwei ausgewählten Ansätzen analysiert und eine gängige Systematisierung von Kompetenzen in Spezial- und Schlüsselkompetenzen vorgestellt. In Bezug auf den wirtschaftlichen, technologischen und gesellschaftlichen Wandel wird die Wichtigkeit zur Kompetenzentwicklung und diagnostischen Erfassung von Schlüsselkompetenzen im nationalen und internationalen Raum diskutiert und herausgestellt.

Darauf aufbauend wird der Bereich der psychologischen Eignungsdiagnostik in seinen Grundzügen dargestellt und die Relevanz von Testverfahren, genauer gesagt von Intelligenz- und Wissenstests herausgestellt. Der zweite und vor allem der dritte Gliederungspunkt sollen dazu beitragen, die wissenschaftliche und psychologisch fundierte Diagnostik mit den dazugehörigen Gütekriterien zu vertiefen. Die Notwendigkeit hierzu ergibt sich, weil die Bedeutung in der Arbeitswelt häufig unterschätzt wird. Demzufolge richten sich die Ausführungen insbesondere an berufsbezogene Praktiker, die grundlegende diagnostische Anforderungen und Zusammenhänge in kompakter Form erfassen möchten sowie anwendungsbezogen arbeiten. Im Kapitel 3.2.2 und 3.2.4 werden die Begriffe Test und Leistungstest definiert, klassifiziert und die zugrunde liegenden Gütekriterien vorgestellt. Anschließend werden die eignungsdiagnostischen Instrumente bzw. die Leistungstests im Kontext der Personalauswahl betrachtet und die Ergebnisse ausgewählter Studien wiedergegeben, die die Vorhersagbarkeit des beruflichen Erfolgs durch Intelligenz- und Wissenstests belegen. Danach wird die Bedeutung von Leistungstest für die Praxis herausgestellt.

Als Voraussetzung für die im empirischen Teil folgende Validierungsstudie des START-W, werden gemäß der klassischen Testtheorie die zu untersuchenden Konstrukte Intelligenz und Wissen und deren Beziehungsgeflecht detailliert analysiert.

Vorab wird ein historischer Abriss der Intelligenzmessung gegeben, in dem fundamentale Ansätze skizziert werden.

In Kapitel 4 werden das Konstrukt-Verständnis der Intelligenz erläutert und die Bedeutungsbereiche sowie Bedeutungskomponenten der Intelligenz herausgestellt.

Eine besondere Rolle kommt den Strukturtheorien zu, die sowohl den Bedeutungsbereich der Intelligenz spezifizieren als auch die Grundlage für verschiedene Intelligenztests bilden. Das Hauptaugenmerk liegt in diesem Kapitel auf dem hierarchischen Protomodell (HPI) und auf dem Modell der fluiden- und kristallisierten Intelligenz von Cattell (1971a, 1987). In diesem Zusammenhang wird insbesondere die kristallisierte Intelligenz hervorgehoben, die einen engen Bezug zum Wissen aufweist und daher einen Ansatzpunkt zur Überleitung auf das Kapitel 5 liefert.

In diesem Kapitel wird der Begriff Wissen expliziert und durch die beiden Wissensformen (deklaratives und prozedurales Wissen) spezifiziert. Danach werden grundlegende wissenspsychologische Aspekte zum Wissenserwerb und der Wissensanwendung dargestellt. Im Bezug auf die Stichprobenauswahl im empirischen Teil der Arbeit werden zusätzlich der Wissenserwerbsprozess (Lernen) und eine Möglichkeit zur Wissensvermittlung in schulischen und berufsorientierten Bildungseinrichtungen thematisiert.

Angesichts der engen Beziehung zwischen (kristallisierter) Intelligenz und Wissen werden in Kapitel 6 zunächst allgemeine Zusammenhänge erörtert. Bezüglich der angedeuteten Verknüpfung von kristallisierter Intelligenz und Wissen im Rahmen der Strukturtheorie von Cattell (1971a, 1987) und dem Intelligenz-Struktur-Test-2000 R (Liepmann et al., 2007) wird die komplexe Vernetzung zwischen fluider Intelligenz, kristallisierter Intelligenz und Wissen vor dem theoretischen Hintergrund gesondert betrachtet und diskutiert. Ausgehend von dieser Sichtweise werden die theoretisch fundierten Konstrukte der Intelligenz und des Wissens in die angewandte Differentielle Psychologie eingebettet und operationalisiert.

Auf der Basis von theoretischen Grundlagen wird eine umfassende Validierungsstudie durchgeführt. Der empirische Teil der Arbeit bezieht sich demnach auf die Weiterentwicklung des Wissenstest START-W. Das Ziel ist die Etablierung eines an den Richtlinien der DIN-Norm 33430 orientierten Tests, um den Leistungsstand und das Leistungspotential einer Person hinsichtlich der verschiedenen Anforderungen einer Tätigkeit realistisch abbilden zu können. Im Zuge dessen wurden Datenerhebungen an berufsbildenden Schulen, Berliner Oberstufenzentren, Oberschulen (Gymnasien) und

Universitäten realisiert, die einen deutlichen Bezug zur Zielpopulation der Berufseinsteiger erkennen lassen.

In einem ersten Schritt der Studie wird eine Binnenanalyse der Test-Items der Vorversion durchgeführt. Anhand der Item-Analysen werden mittels exploratorischer Methoden Skalen zur Erfassung des verbal, numerisch und figural kodierten Wissens gebildet und hinsichtlich ihrer Kennwerte evaluiert. Die Bildung von Skalen stellt dabei auch eine Grundvoraussetzung für die nachfolgenden Analysen und Interpretationen dar. Darauf aufbauend werden Reliabilitätsanalysen durchgeführt, die die Berechnung der *Retest-Reliabilität* sowie der *Split-Half-Reliabilität* umfassen.

Vor dem Hintergrund der theoretisch-empirischen Befundlage und Ausführungen, in dem die Interdependenz von Umweltfaktoren und Wissen thematisiert und der Einfluss verschiedener Moderatorvariablen diskutiert wird, werden zielführende Forschungsfragen abgeleitet. Aufgrund der Ansicht, dass das Wissen einer Person im Laufe der individuellen Entwicklung durch die Kultur und die Gesellschaft geprägt wird, werden innerhalb der aufgestellten Forschungsannahmen die Auswirkungen verschiedenartiger Umwelteinflüsse (Alter, Muttersprache, Geschlecht und Bildungsniveau) auf die Wissensakkumulation untersucht. In diesem Kontext werden auch erste Hinweise auf die Validität des Wissenstests generiert.

In einem weiteren Schritt wird das Verfahren auf seine kriterienbezogene Validität geprüft, bei der die Übereinstimmungsvalidität als auch die prognostische Validität analysiert werden. Anschließend wird die Konstruktvalidität analysiert. Hierbei wird die konvergente Validität untersucht und die faktorielle Validität zur Überprüfung der intendierten Struktur des Wissenstests begutachtet. In diesem Bezugsrahmen wird auch der Zusammenhang zwischen fluider Intelligenz und kristallisierter Intelligenz (Wissen) erforscht. Als Untersuchungsinstrument wird auch das Grundmodul des I-S-T 2000 R eingesetzt, das entsprechend den obigen Ausführungen geeignet ist, den Zusammenhang zu untersuchen und einen zusätzlichen Hinweis auf die konvergente Validität des Wissenstest START-W zu erhalten.

Abschließend werden die Ergebnisse dargestellt, die Befunde diskutiert und ein zusammenfassendes Fazit sowie ein Ausblick auf mögliche, zukünftige Forschungsarbeiten gegeben, in dem weitere Aspekte einer sich auch in Zukunft permanent weiter entwickelnden Bildungs- und Wissensfrage mit den involvierenden Konsequenzen für die berufsbezogene Eignungsdiagnostik aufgegriffen werden.

## 2 Schlüsselqualifikationen im Kontext berufsbezogener Eignungsdiagnostik

Dem Konzept der Schlüsselqualifikationen wird heute wie auch in der Zukunft eine besondere Bedeutung in der Wissenschaft und berufsbezogenen Praxis zuerkannt. Die Gründe für das verstärkte Interesse an extrafunktionalen Qualifikationen liegen in den kontinuierlichen Veränderungen der global vernetzten Marktwirtschaft. Vor dem Hintergrund der Globalisierung, der Zunahme von (informations-) technologischen Innovationen und der damit verbundenen Geschwindigkeit des Wissenszuwachses, wird deutlich, dass die Arbeitswelt einem stetigen Wandel unterliegt. Dadurch verändern sich nicht nur die einzelnen Berufsbilder, sondern auf einer höheren Aggregationsebene werden auch die Unternehmen und die gesamte Wirtschaft mit strukturellen und strategischen Umstellungen konfrontiert (Liepmann, 1993). Der Prozess der gesellschaftlichen Arbeitsteilung, der verstärkte Trend zur Kundenorientierung, die Tendenz zu flacheren Hierarchien und Projektarbeit, den immer kürzer werdenden Produktlebenszyklen und steigenden Innovationsraten als auch die Zunahme von neuen Managementkonzepten, wie z.B. *Diversity Management*, führen zu unterschiedlichen Qualitäten von Arbeit und zu einer Verlagerung von spezialisierten Arbeiten zu wissensorientierten Tätigkeiten (Sonntag & Schaper, 2006; Liepmann & v. Gilardi, 1993). Aufgrund des hohen Veränderungsdrucks auf die Unternehmen und der Notwendigkeit zur gleichzeitigen Anpassung ist es ersichtlich, dass das arbeitsplatzbezogene Fachwissen schneller veraltet, immer umfangreicher wird und letztendlich nicht mehr ausreicht, um den neuen Anforderungen gerecht zu werden. Die Konsequenz dieser Entwicklung ist darin zu sehen, dass neben den Fachqualifikationen in verstärktem Maß auch extrafunktionale Qualifikationen bzw. Schlüsselqualifikationen oder Schlüsselkompetenzen gefordert werden, um die Veränderungen erfolgreich zu bewältigen (Liepmann, 1993; Bürgel & Zeller, 1998; Eilles-Matthiessen, el Hage, Janssen & Osterholz, 2002). Die Forderung bezieht sich hierbei nicht ausschließlich auf die höheren Hierarchieebenen, sondern auf alle in der Organisation beteiligten Mitglieder. In diesem Zusammenhang wird schon von den Auszubildenden verlangt, dass sie die erforderlichen, fachübergreifenden Qualifikationen mitbringen, um flexibel und selbständig die vielfältigen Aufgaben zu bearbeiten (Renold, Frey & Balzer, 2002). Die erforderlichen Schlüsselkompetenzen, wie zum Beispiel Allgemeinwissen, Teamfähigkeit, Flexibilität, Lernbereitschaft, Denken in



Gesamtzusammenhängen und weiteren Fähigkeiten sollten demnach zum Standardrepertoire der Berufseinsteiger gehören. Damit für diesen Ansatz ein entsprechender Rahmen gegeben wird, gilt es gleichermaßen adäquate Organisationsstrukturen und Strategien zu entwickeln, die im Speziellen geeignete Personalauswahl, Personalplatzierungs- und Personalentwicklungsmaßnahmen umfassen (Liepmann, 1993; Stangel-Meseke, 1994).

Auf Basis dieser Ausführungen ist es ersichtlich und von erhöhter Relevanz, dass bereits in der berufsbezogenen Eignungsdiagnostik von Auszubildenden solche Kompetenzen berücksichtigt werden, wobei insbesondere die Erfassung des Allgemeinwissens und der Intelligenz von weitreichender Bedeutung ist. An dieser Stelle wird deutlich, dass der psychologischen Diagnostik auch im Allgemeinen eine außerordentliche Rolle zukommt und es wird die Notwendigkeit gesehen, „Wissen nicht nur indirekt (durch die Messung der Intelligenz), sondern verstärkt auch direkt bei der Prognose beruflicher Leistungen zu berücksichtigen“ (Kersting, 1999, S.256).

## **2.1 Zum Konzept der Schlüsselqualifikationen- Herkunft des Begriffes**

Der Begriff Schlüsselqualifikation ist in den letzten Jahren zum regelrechten Modebegriff avanciert. In der wissenschaftlichen psychologischen und betriebswirtschaftlichen Literatur werden mittlerweile ca. 800 verschiedene Schlüsselqualifikationen aufgeführt, von denen die Hälfte nur einmal genannt bzw. nicht detaillierter behandelt werden (Gnahn, 2007). Die Problematik liegt darin begründet, dass die Bedeutung der Begriffe unpräzise und uneinheitlich formuliert wird, die Bedeutungen bzw. Definitionen sich überschneiden oder es werden unterschiedliche Bedeutungen dem jeweiligen Konzept zugeordnet. Folglich sind im Zuge der wissenschaftlichen Forschungsarbeiten und in der beruflichen Praxis vielfältige Synonyme für extrafunktionale Qualifikationen entwickelt worden. Diesbezüglich wird von Schlüsselqualifikationen, überfachlichen Qualifikationen, dynamischen Qualifikationen, Basisqualifikationen, Zukunftsqualifikationen, fachübergreifenden Fähigkeiten, multifunktionalen Fähigkeiten, extrafunktionalem Wissen, etc. gesprochen. Im Vergleich zu klassischen psychologischen Dispositionen wirken einige Begriffe zwar fragmentarisch und unausgereift, aber im Endeffekt besteht „ihr Erklärungswert [...] jedoch darin, dass sie auf arbeits- und berufsrelevante

Anforderungen bezogen sind“ (Eilles-Mathiessen et al., 2002, S. 14). Aufgrund der Vielzahl inhaltlicher Überschneidungen und Verflechtungen werden die Begriffe Schlüsselqualifikationen, extrafunktionale Qualifikationen, Basiskompetenzen, außerfachliche Fähigkeiten oder Schlüsselkompetenzen daher im Folgenden synonym verwendet.

Der Ausgangspunkt des Begriffes liegt im wirtschafts- und arbeitsmarktpolitischen Bereich (vgl. Stangel-Meseke, 1994; Liepmann, 1993). Im deutschsprachigen Raum sind die zwei Autoren Dahrendorf (1956) und Offe (1970) zu nennen, die sich erstmals mit fachübergreifenden Qualifikationen auseinandergesetzt haben. Für Dahrendorf (1956) liegt das Erfordernis von Schlüsselqualifikationen im Ergebnis der wirtschaftlichen und industriellen Entwicklungsprozesse, genauer gesagt im Resultat zunehmender Arbeitsteilung begründet. Er differenziert im Weiteren zwischen funktionalen und extrafunktionalen Qualifikationen. Unter funktionalen Qualifikationen subsumiert Dahrendorf elementare bzw. grundlegende Kenntnisse, die dazu dienen, den reinen technischen und fachspezifischen Aufgaben der Arbeit gerecht zu werden. Die extrafunktionalen Qualifikationen beziehen sich dagegen weniger auf die technischen Anforderungen, sondern vielmehr auf die übergeordneten Arbeitsprozesse bzw. die organisatorischen und sozialen Zusammenhänge. Die folgenden Beispiele charakterisieren nach Dahrendorf die extrafunktionalen Fähigkeiten: (a) Anpassung an die gegebenen Arbeitsbedingungen (Fabrikatmosphäre, Arbeitsrhythmus, schwierige Arbeitsbedingungen), (b) Verantwortung für Arbeitsvollzüge und Arbeitsabläufe sowie für die Sicherheit von Personen, Maschinen oder Materialien (Entscheidungsfähigkeit), (c) Besitz latenter funktionaler Fertigkeiten (funktionale Kenntnisse, die über den Aufgabenbereich hinausgehen und den Arbeitsbereich bzw. Arbeitsprozesse betreffen). Im Rahmen seiner Forschungsarbeiten wird die Hypothese formuliert, dass im Zuge technologischer Innovationen funktionale Fertigkeiten zunehmend durch extrafunktionale Fertigkeiten substituiert werden (vgl. auch Stangel-Meseke, 1994). Offe (1970) begründet die Forderung nach Schlüsselqualifikation mit dem Wandel der Organisationsstrukturen bzw. dem Kontrollproblem innerhalb der Unternehmen. Aufgrund des rasanten Wachstums der Organisationen müssen äquivalente Formen der Kontrolle geschaffen werden, um deren Funktionsfähigkeit zu gewährleisten. Offe (1970) spricht hier unter anderem von *extra-funktionalen Orientierungen*, die die kooperativen Prozesse und die Konformität des Organisationsmitgliedes mit den Normen und Interessen sicherstellen sollen. Beiden Autoren ist gemein, dass sie die affektiv-motivationale Dimension der geforderten Qualifikationen exponiert

haben und so zum besseren Verständnis von Schlüsselqualifikationen beitragen. Es fehlt jedoch weiterhin an einer inhaltlichen Klassifikation der postulierten Schlüsselqualifikationen (Stangel-Meseke, 1994).

Mertens (1974), ein weiterer Pionier in diesem Kontext, stellte in seinen Forschungen ebenfalls heraus, dass durch die hohen Innovationsraten, den rasanten (informations-) technologischen Entwicklungen und den damit einhergehenden Wandel in den verschiedenen Berufsfeldern und Organisationsstrukturen, die fachspezifischen Arbeitsplatzkenntnisse bzw. das berufliche Wissen eine schnellere Verfallzeit haben als fachübergreifende Fähigkeiten. Damit die neuartigen Anforderungen und Veränderungen dennoch erfolgreich bewältigt werden können, bedarf es nach Mertens sogenannter Schlüsselqualifikationen. Denn erst die Schlüsselqualifikationen ermöglichen es, die ständig wechselnden, zukünftigen beruflichen Anforderungen auf einer höheren Abstraktionsebene effektiv und flexibel zu bewerkstelligen. Mertens (1974) definiert Schlüsselqualifikationen als

„Kenntnisse, Fähigkeiten und Fertigkeiten, welche nicht unmittelbaren und begrenzten Bezug zu bestimmten, disparaten praktischen Tätigkeiten erbringen, sondern vielmehr die Eignung für eine große Zahl von Positionen und Funktionen als alternative Optionen zum gleichen Zeitpunkt [sicherstellen], und die Eignung für die Bewältigung einer Sequenz von (meist unvorhersehbaren) Änderungen von Anforderungen im Laufe des Lebens [ermöglichen].“ (S.40)

Des Weiteren unterteilt er die Schlüsselqualifikationen in (a) Basisqualifikationen, (b) Breiterelemente, (c) Horizontalqualifikationen und (d) Vintage-Faktoren. Zu den Basisqualifikationen zählt Mertens die logisch-analytischen strukturierenden und konzeptionellen kognitiven Fähigkeiten als auch die Fähigkeiten zur Kooperation und der Kreativität. Die Breiterelemente versteht er als allgemeine notwendige Kenntnisse und Fertigkeiten, die die Beherrschung von Grundrechenarten, Lesefähigkeiten und Kenntnisse von Arbeitsplatzsicherheit und Umweltschutz mit einschließen. Unter den Horizontalqualifikationen subsumiert er die Gewinnung, Verarbeitung, Verständnis und die effiziente Anwendung/Nutzung von Informationen. Die Vintage-Faktoren umfassen vor allem die grundlegende Bereitschaft zum lebenslangen Lernen sowie Sprachkenntnisse, Informatik, Rechtsgrundlagen, Geographie, Kunst/Kultur und naturwissenschaftliche Kenntnisse. Grundsätzlich fokussiert Mertens in seinem Konzept die kognitiven Fähigkeiten und Wissenskomponenten, die den Individuen eine flexible und schnelle

Aufnahme von neuem Wissen und die Anpassungen an veränderte Anforderungen ermöglichen. Die Kritik betrifft die unpräzise Abgrenzung der vier Dimensionen, die vage inhaltliche Differenzierung des Begriffes, die ungenügende theoretische Einbettung, die empirische Überprüfbarkeit und die Operationalisierung (Stangel-Meseke, 1994). Obwohl dem Konzept eher ein Vorschlagscharakter zugewiesen wird, gibt Mertens` Ansatz viele nützliche, praxisorientierte Hinweise, die eine Sensibilisierung für das Erfordernis von extrafunktionalen Qualifikation in der beruflichen Praxis hervorgerufen haben. Mertens kann somit als Vorreiter in der wissenschaftlichen und praxisbezogenen Erforschung von Schlüsselqualifikationen gesehen werden.

## **2.2 Konzepte zur Bestimmung der Schlüsselqualifikationen**

Im folgenden Kapitel werden ausgewählte theoretische und empirische Konzepte zur Identifikation extrafunktionaler Qualifikationen vorgestellt. Die Darstellung der Ansätze ist fundamental für die anschließende Kategorisierung der Schlüsselqualifikationen. Das erklärte Ziel ist, einen theoretischen Rahmen zu schaffen, der die Anforderungen an die Kategorisierung gewährleisten soll.

### **2.2.1 Theoriebasierte Konzepte zur Identifikation und Relevanz von Schlüsselkompetenzen**

Der erste theoriebasierte Ansatz von Fricke und Fricke (1976) umfasst eine grundlegende Einordnung der Schlüsselqualifikationen. Die Autoren ordnen das Konzept der extrafunktionalen Qualifikationen hierbei in den Bereich der Industriesoziologie bzw. in die autonomieorientierte Qualifikationsforschung ein. Die autonomieorientierte Qualifikationsforschung fokussiert im Allgemeinen „die berufliche Autonomie der Arbeitenden als Zielgröße bildungspolitischer Anstrengungen und komplementärer arbeitsorganisatorischer Innovationen, die den Handlungsspielraum der Arbeitenden und die Möglichkeiten der Anwendung von Qualifikationen im Arbeitsprozess erweitern sollen“ (Stangel-Meseke, 1994, S.38). Qualifikationen werden dabei als Handlungsspielraum der Arbeitenden verstanden, die den reibungslosen Arbeitsablauf mit seinen technisch organisatorischen Determinanten sicherstellen sollen. Die erforderlichen Qualifikationen werden wiederum in funktionale bzw. fachliche Qualifikationen und innovatorische Qualifikationen unterteilt. Die fachlichen Qualifikationen werden in diesem Kontext als

arbeitsplatzspezifische, funktionale Fähigkeiten aufgefasst, wohingegen die innovatorischen Qualifikationen eher übergeordneten Charakter haben, die dem Arbeiter mehr Autonomie im Industrieunternehmen erlauben. Der Forschungsschwerpunkt in diesem Bereich erstreckt sich auf die Untersuchung funktionaler Qualifikationen, die den rein technischen Arbeitsprozess betreffen und die Identifikation von erforderlichen fachübergreifenden Fähigkeiten (Fricke & Fricke, 1976; vgl. Stangel-Meseke, 1994).

Einen anderen Ansatz verfolgt Schmiel (1988). Sein Ausgangspunkt ist der dem globalisierten, marktwirtschaftlichen Prozess inhärenten Wandel von Anforderungen, die die Herausbildung von übergeordneten Qualifikationen erfordern. Schmiel (1988) spricht in diesem Zusammenhang von wichtigen Lernzielen, aus denen er die Ausweisung von extrafunktionalen Qualifikationen ableitet. Hierbei stützt er sich auf Grundlagen der Soziologie, anthropologischen Forschung und der Psychologie. Außerdem bezieht er sich auf Handlungsanforderungen, die die „moderne Zeit“ mit sich bringt. Der erste Bezugspunkt betrifft die Handlungsfähigkeit des Menschen, die auf folgenden Lernzielen (Qualifikationen) basiert: Reflexionsfähigkeit (Vorbereitung von Entscheidungen, Verantwortungsbewusstsein, Sachkenntnis), Selbständigkeit und Verantwortungsbewusstsein als Voraussetzungen für die Reflexionsfähigkeit und geistige Fähigkeit (Fähigkeiten zum abstrakten, analytischen, problemlösenden, kreativen und transferfähigen Denken). Auf der anderen Seite beinhalten die Handlungsanforderungen der „modernen Zeit“ sogenannte berufswichtige, berufsübergreifende und verantwortungsbezogene Fähigkeiten, soziale Kompetenzen, gutes Arbeitsverhalten, wie zum Beispiel Lern- und Leistungsbereitschaft, Zielstrebigkeit, Pünktlichkeit und die allgemeine Grundbildung (Sprachkenntnisse, Allgemeinwissen). Für die Identifikation und Beurteilung berufsbezogener Schlüsselqualifikationen sind nach Schmiel (1988, S.72) folgende Kriterien von essentieller Bedeutung:

- Die Qualifikationen müssen berufs- und wirkungsbereichsübergreifend wichtig sein.
- Sie müssen den besonderen Anforderungen der modernen Zeit mit ihren Technologien und technologischen Weiterentwicklungen entsprechen.
- Die Qualifikationen müssen mit denen in Einklang stehen, die den Besonderheiten des Menschseins entsprechen, da diejenigen Qualifikationen zu fördern sind, die der Sicherung des Menschlichen in der modernen Zeit dienen.
- Diejenigen Fähigkeiten sind den Schlüsselqualifikationen zuzuordnen, die für die Meisterung der Anforderungsänderungen benötigt werden.

Der individuumorientierte Ansatz von Reetz (1989, 1990) ist eine pädagogisch anthropologische Theorie, die wiederum auf der Persönlichkeitstheorie von Roth (1966, 1971) aufbaut. Im Kern des Konzepts geht es um ein *best fit* zwischen Person und Situation, die auf einer zweidimensionalen, grafischen Achse veranschaulicht werden können. Dementsprechend wird die Kongruenz zwischen situationalen Anforderungen und den Fähigkeiten einer Person in den Mittelpunkt der Betrachtung gestellt. Das heißt, sofern eine Person eine neuartige Situation erfolgreich bewältigt, gewinnt sie Handlungsfähigkeit. Von Qualifikation wird in diesem Bezugsrahmen gesprochen, wenn die Handlungsfähigkeit eines Individuums einen übergeordneten oder abgrenzbaren situativen Konnex besitzt. Mit anderen Worten müssen die Qualifikationen flexibel genug sein, um den wechselnden beruflichen Anforderungen gerecht zu werden. Demnach werden weniger die spezifischen Berufsanforderungen und funktionalen Fähigkeiten thematisiert, sondern vielmehr die menschliche Handlungsfähigkeit bzw. extrafunktionale Qualifikationen ins Zentrum der Forschung gerückt (vgl. Stangel-Meseke, 1994). Im Bezug auf die Persönlichkeitstheorie von Roth (1966, 1971) gliedert sich die Qualifikation eines Individuums in drei Dimensionen, die auch als Handlungssystem bezeichnet werden:

- Sacheinsichtiges Verhalten und Handeln (Sachkompetenz und intellektuelle Mündigkeit)
- Sozialeinsichtiges Verhalten (Sozialkompetenz und soziale Mündigkeit)
- Werteinsichtiges Verhalten (Selbstkompetenz und moralische Mündigkeit)

Diesem Handlungssystem werden wiederum fünf Systeme bzw. persönlich-charakterliche Basisfähigkeiten vorangestellt: das Wertungssystem (z.B. Verantwortungsbewusstsein), Orientierungssystem (z.B. Abstraktionsfähigkeit), Antriebssystem (z.B. Leistungsmotivation, Initiative), Lernsystem (Beherrschung von Bildungstechniken) und das Steuerungssystem (z.B. Kontinuität von Interesse). Die Schlüsselqualifikationen werden nach Reetz (1989) insoweit abgegrenzt, als dass sie sich stärker auf den beruflichen Kontext beziehen. Das gilt nicht nur für die tätigkeits- und leistungsorientierten kognitiven Fähigkeiten (wie z.B. Entwicklung von Innovationen und Problemlösungen, berufsspezifisches und allgemeines Wissen), sondern gleichermaßen für die sozialen und kommunikativen Fähigkeiten, die im Rahmen seines Konzepts zu den klassischen Schlüsselqualifikationen gerechnet werden können. Die Kriterien zur Bestimmung des berufsbezogenen Charakters der Schlüsselqualifikationen sind an

Schmiels Ansatz (1988) angelehnt. Ein besonderer Aspekt in Reetz' individuumsorientiertem Konzept ist, dass die einzelnen Dimensionen der Schlüsselqualifikationen in einem interdependenten Verhältnis zu Systemen der Persönlichkeit stehen. Durch die Berücksichtigung von emotionalen und motivationalen Grundfähigkeiten wird die „kognitive Kopflastigkeit“ im Vergleich zu anderen Schlüsselqualifikationsansätzen reduziert (Stangel-Meseke, 1994).

### **2.2.2 Empirisch orientierte Konzepte zur Identifikation von Schlüsselkompetenzen**

Im Rahmen der empirischen Analysen wird der Fokus auf die Identifikation relevanter extrafunktionaler Qualifikationen gelenkt, die vor allem in der beruflichen Praxis von Bedeutung sind. Insbesondere die personalpolitischen Teilbereiche, wie Berufseinstieg, Personalauswahl, Personalbeurteilung und -entwicklung erhalten hier eine herausragende Stellung. Hierzu werden im Folgenden einige ausgewählte Forschungsbeiträge vorgestellt, die sich auf die unterschiedlichen Arbeitsfelder der Personalpolitik beziehen.

In der Untersuchung von Krüger (1988) wurden zunächst die organisatorischen Einflussfaktoren auf die Schlüsselqualifikationen analysiert. Der Ausgangspunkt seiner Überlegungen und die zugrunde liegende Forschungshypothese war, dass Schlüsselqualifikationen einen dynamischen Charakter haben, die zur erfolgreichen Bewältigung organisatorischer und sozialer Probleme bzw. Entwicklungen im Spannungsfeld zwischen der Organisation und der Umwelt beitragen. In dem Maß wie die extrafunktionale Qualifikationen die Unsicherheiten der Organisation, bezogen auf die sich ständig wandelnde Umwelt, begrenzen können, entfalten sie ihre außerordentliche Bedeutung. Das Ziel seiner empirischen Untersuchung war es, die Anforderungen an extrafunktionale Qualifikationen zu eruieren, die von der Organisationsstruktur und der Umwelt determiniert werden.

Folglich formuliert Krüger (1988) vier unterschiedliche Dimensionen von Schlüsselqualifikationen, die Ausfluss generalisierter Erwartungen der Organisation sind:

- Zweckbezogene Orientierungen/Qualifikationen  
(formale Erwartungen an die Mitgliedschaft)
- Kontextbezogene Qualifikationen  
(Fähigkeiten, Fertigkeiten und Kenntnisse, die auf die Merkmale, Strukturen, Prozesse und das informale Regelwerk der Organisation bezogen sind.)

- Sozio-motorische Qualifikationen  
(Fähigkeiten, die es dem Organisationsmitglied auf der Grundlage kontextbezogener Kenntnisse möglich machen, Informationswege zu entschlüsseln, zu gestalten, zu nutzen und die Informationsversorgung effektiv und effizient zu betreiben. Die Qualifikationen sind transferierbar.)
- Soziale Qualifikation  
(Fähigkeiten im Umgang mit anderen Organisationsmitgliedern: formell und informell.)

Bezüglich seiner methodischen Herangehensweise sollte die Korrelation zwischen den von der Organisation vorgegebenen Anforderungen und den Schlüsselqualifikationen analysiert werden. Die Stichprobe von 34 Ökonomen aller Hierarchieebenen wurde an einem Versicherungsunternehmen erhoben. Das Untersuchungsdesign bestand aus einem teilstandardisiertem Fragebogen und einem offenem Interview. Die grundlegenden Ergebnisse waren folgende (vgl. auch Stangel-Meseke, 1994):

- Schlüsselqualifikationen sind in ihrer Funktion multidimensional, multivariat und multifunktional.
- Schlüsselqualifikationen inkludieren persönlich-charakterliche Eigenschaften sowie Wissen über arbeitsplatzbezogene und allgemeine Zusammenhänge, als auch kognitive Fähigkeiten.
- Zur Bestimmung der Anforderungen an Schlüsselqualifikationen spielen die unterschiedlichen Hierarchieebenen, Funktionen und Abteilungen als Erklärungsvariablen keine Rolle.
- Die kontextbezogenen und die sozio-motorischen Qualifikationen sind bei fast allen organisatorischen Merkmalen sowie für die Erfüllung der Aufgaben erforderlich.
- Die Anforderungen an einzelne Schlüsselqualifikationen sind abhängig von dem Ausmaß der gegebenen Handlungsspielräume.
- Koordinationsaufgaben und Kontrollaufgaben sowie das vorgegebene Entscheidungsniveau und die Entscheidungsvorbereitung erfordern eine große Bandbreite an Schlüsselqualifikationen. Insbesondere soziale Qualifikationen sind hier relevant.



- Die zweckbezogenen Qualifikationen bleiben in den Unterstellungsverhältnissen latent. Persönliche Kontakte erfordern hingegen sozio-motorische als auch soziale Qualifikationen.

Teichler, Butgereit und Holtkamp (1984) konzentrieren sich in ihren empirischen Untersuchungen auf die Identifikation, Analyse und Bedeutung von Schlüsselqualifikationen in der Personalauswahl. Die zugrunde liegende Stichprobe von 47 Personalexperthen wurde im Zeitraum von 1981-1982 in 40 der umsatzstärksten Organisationen erhoben. Die Instrumente innerhalb der Untersuchung setzten sich aus einem Fragebogen und strukturierten Interviews zusammen. Die ersten empirischen Ergebnisse von Teichler et al. (1984) dokumentieren, dass vor allem die außerfachlichen Kompetenzen einen hohen Stellenwert in der Bewerberauswahl besitzen. Es bleibt jedoch anzumerken, dass das Erfordernis bzw. die Schwerpunktsetzung auf bestimmte Schlüsselqualifikationen von Unternehmen zu Unternehmen variieren kann. Außerdem sind die inhaltlichen Ausgestaltungen der verschiedenen Dimensionen von Schlüsselqualifikationen teilweise uneindeutig und interdependent. Trotzdem konnten in diesem Bezugsrahmen fünf allgemeine Kategorien von extrafunktionalen Qualifikationen herausgestellt werden:

1. Generelle Arbeitstugenden  
(Effektivität, Effizienz, Leistungsbereitschaft, Belastbarkeit, Pünktlichkeit, Zielstrebigkeit)
2. Allgemeine Persönlichkeitsmerkmale  
(Allgemeinbildung, fundierte Wissensbasis, Persönlichkeit, Interessen und politische Bildung)
3. Kognitive Fähigkeiten  
(Problemlösungsfähigkeit, Fähigkeiten zum abstrakten, analytischen Denken)
4. Kooperative und kommunikative Fähigkeiten  
(Führungsverhalten, Kontaktfreudigkeit, Teamfähigkeit, Durchsetzungsvermögen)
5. Arbeitsübergreifende Aspekte  
(Allgemeines Interesse, Engagement, Flexibilität, Initiative)

In der Untersuchung von Kemmet, Linke und Wolf (1982) wurden Schlüsselqualifikationen herauskristallisiert, die beim Berufseinstieg von Wirtschaftswissenschaftlern gefordert werden.

Die empirische Datenerhebung wurde in ca. 350 Wirtschaftsunternehmen durchgeführt. Die Stichprobe wurde aus Personalleitern, Abteilungsleitern und Arbeitsplatzinhabern rekrutiert. Die verwendeten methodischen Instrumente umfassten ausführliche Interviews und eine von Infratest initiierte Repräsentativbefragung. Die Bedeutung der zu untersuchenden Schlüsselqualifikationen wurde im Weiteren nach drei Aufgabenfeldern differenziert:

1. Planung-Steuerung-Verwaltung
2. Beschaffung-Absatz-Management
3. Personal-und Sozialwesen

Die damit zusammenhängenden Schlüsselqualifikationen wurden wiederum in drei Fähigkeitsklassen unterteilt:

1. Kognitive Fähigkeiten  
(Fähigkeit zum abstrakten, analytischen, kritischen und schöpferischen Denken, Problemlösungsfähigkeit, Fähigkeit zu lernen bzw. Wissen zu akkumulieren und anzuwenden)
2. Soziale Fähigkeiten  
(Teamfähigkeit, Führungsqualitäten, Überzeugungskraft, Diskussionsfähigkeit, Sozialkompetenz im Umgang mit Anderen)
3. Motivationale Fähigkeiten  
(Verantwortungsbewusstsein, Entscheidungsfähigkeit, Interesse, Bereitschaft neue und übergeordnete Aufgaben zu übernehmen)

Die Ergebnisse zeigten, dass in allen drei Aufgabenbereichen die kognitiven Fähigkeiten hinsichtlich der unterschiedlichen Arbeitsplatzanforderungen für Wirtschaftswissenschaftler am wichtigsten sind. Insbesondere die Fähigkeit zum Problemlösen sowie die Fähigkeit zur schnellen Wissensakkumulation als auch das analytische, schöpferische Denken besitzen eine übergeordnete Relevanz. Hierzu gehören auch vorgelagerte Basiskompetenzen, wie Allgemeinbildung und die Beherrschung von Bildungstechniken (vgl. auch Stangel-Meseke, 1994).

Im Folgenden wird auf weitere empirische Forschungsstudien verwiesen, die einen wichtigen Beitrag zur Entwicklung des Konzepts der Schlüsselqualifikationen geleistet haben: Spreter-Müller (1988) analysierte die Bedeutung von extrafunktionalen Qualifikationen beim Berufseinstieg von Diplom-Kaufleuten und Diplom-Ingenieuren.

Der Autor stellte hierbei einen Kriterienkatalog zusammen, der vier grundsätzliche Fähigkeiten und Fertigkeiten aufwies:

1. Fachübergreifende Kenntnisse  
(Sprachkenntnisse, EDV-Kenntnisse, Allgemeinwissen, grundlegende juristische, betriebswirtschaftliche und technische Kenntnisse)
2. Kognitiv-intellektuelle Fähigkeiten  
(Fähigkeit zum abstrakten, analytischen, methodischen, kreativen und praxisnahen Denken, Denken in Gesamtzusammenhängen, Fähigkeit zu lernen und Wissen zu akkumulieren)
3. Soziale und kommunikative Fähigkeiten  
(Führungsqualitäten, Motivation, Durchsetzungsvermögen, Konfliktmanagement, Diskussionsbereitschaft)
4. Leistungsorientierte Fähigkeiten  
(Verantwortungsbewusstsein, Entscheidungsfähigkeit, Stressbewältigung)

Grabowski (1990) untersuchte die Bedeutung von außerfachlichen Qualifikationen in der beruflichen Erstausbildung in verschiedenen Metallberufen, wobei zwei bedeutende Begriffscluster aufgezeigt werden konnten:

1. Kognitive Fähigkeiten  
(analytisches, abstraktes, schlussfolgerndes Denken, Lernfähigkeiten bzw. Wissenserwerb)
2. Arbeitstugenden  
(Fleiß, Pünktlichkeit, Zielstrebigkeit, Interesse und Motivation)

In einer Studie des Instituts für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit wurde vor allem die soziale Qualifikation innerhalb des Berufsverlaufes eingehend betrachtet und hervorgehoben (Blaschke, 1987).

Auf der Basis der Studien von Teichler et al. (1984) führte Landsberg (1986) eine empirische Untersuchung bei Personalexperten an 107 Unternehmen durch. Das angestrebte Ziel seiner Arbeit war die Herausstellung eines beruflichen Erfolgsprofils technischer Führungskräfte.

Hierbei konnten vier Dimensionen faktoranalytisch nachgewiesen werden:

1. Führungspotential  
(u.a. Selbstsicherheit, Risikobereitschaft, Führungskompetenz)
2. Leistungswille  
(Anpassungs- und Denkfähigkeit, Lern- und Leistungsbereitschaft)
3. Sozialverhalten  
(sozial erwünschtes Verhalten, Kooperations- und Teamfähigkeit)
4. Kreatives Problemlösen  
(Breites Interessenspektrum, Problemlösungsfähigkeit, Allgemeinbildung und Wissen, Kreativität)

### **2.2.3 Bewertung der theoretischen und empirischen Konzepte**

Im Bezug auf die theoretischen Konzepte wird deutlich, dass Fricke und Fricke (1976), Reetz (1989) und Schmiel (1988) einen handlungstheoretischen Zugang zur Bestimmung der Schlüsselqualifikationen gewählt haben. Den Autoren ist ebenso gemein, dass sich die Qualifikation an den Anforderungen der aktuellen Technologie und wirtschaftlichen Strukturen orientieren müssen. Die extrahierten Schlüsselqualifikationen haben dabei einen übergeordneten, universellen Charakter, die die erfolgreiche Bewältigung neuer Situationen und Entwicklungen gewährleisten sollen. Durch den Rekurs auf die Persönlichkeitstheorie (Roth, 1971) hebt sich der Ansatz von Reetz (1989) insoweit ab, als dass er neben den vorrangig kognitiv-orientierten Ansätzen, die biografisch vorgelegten affektiv-motivationalen Komponenten der Schlüsselqualifikationen (persönlich-charakterliche Fähigkeiten) zusätzlich mit einbezieht. Hier stellt sich jedoch die Frage, ob die dargestellten affektiven extrafunktionalen Qualifikationen überhaupt Rückschlüsse auf eine präzisere Bestimmung jener zulassen oder ob sie einen sinnvollen Mehrwert für die Bestimmung von Schlüsselqualifikationen leisten. Ebenso ist es nicht schlüssig, ob in der beruflichen Ausbildung und Praxis die persönlich-charakterlichen Grundfähigkeiten in der geforderten Intensität berücksichtigt werden können (vgl. auch Stangel-Meseke, 1994).

Die Gemeinsamkeit der empirischen Ansätze ist darin zu sehen, dass den Schlüsselqualifikationen eine zunehmend wichtigere Rolle in der Berufspraxis zuerkannt wird. Es wird jedoch gleichzeitig darauf hingewiesen, dass einzelne Schlüsselqualifikationen immer im organisations- und tätigkeitsspezifischen Kontext betrachtet werden müssen

und in ihrer Bedeutung dementsprechend variieren können. Außerdem treten wie bei den theoretischen Ansätzen Uneinigheiten bei der inhaltlichen Abgrenzung auf. Die Vermischung der verschiedenen Dimensionen von Schlüsselqualifikationen, wie z.B. kognitive, soziale und persönliche Qualifikationsmerkmale, legen vielmehr den Schluss nahe, dass sie interdependent aufeinander bezogen sind. Es ist im Weiteren problematisch, dass den empirischen Untersuchungen unterschiedliche Stichproben und Forschungsfragen zugrunde liegen. Eine einheitliche Identifikation und Interpretation der Schlüsselqualifikationen wird dadurch erschwert und erlaubt eher tendenzielle Aussagen. Es herrscht jedoch Einigkeit darüber, dass insbesondere die kognitiven Fähigkeiten (u.a. analytisches, schlussfolgerndes, schöpferisches Denken, Denken in Gesamtzusammenhängen, Problemlösungsfähigkeiten, etc.) und die damit verbundenen fachspezifischen und tätigkeitsübergreifenden Wissenskomponenten einen bedeutenden Stellenwert einnehmen. Das Erfordernis von kognitiven Fähigkeiten und Wissen trifft hierbei auf fast alle organisatorischen Merkmale und deren Organisationsmitglieder zu. Im Vorgriff auf die nächsten Kapitel ist zu konstatieren, dass die aufgeführten kognitiven Fähigkeiten einen engen Bezug zum Konstrukt der Intelligenz aufweisen. Die eignungsdiagnostische Erfassung der Intelligenz und des Wissens ist damit von höchster Relevanz für die berufliche Praxis.

Auf der Grundlage der dargestellten Studien lässt sich festhalten, dass die theoretischen und vor allem die empirischen Ansätze einen wertvollen Beitrag zur Identifikation und Bedeutung der geforderten Schlüsselqualifikationen in der Berufspraxis leisten. Resümierend ergeben sich daher folgende Implikationen zur Bestimmung der Schlüsselqualifikationen im Rahmen der Klassifikationsansätze (Stangel-Meseke, 1994):

- Eine ganzheitliche Sichtweise verlangt sowohl die Berücksichtigung der theoretischen Ansätze als auch der empirischen Konzepte.
- Bezüglich der theoretischen Ansätze ist es sinnvoll, die extrafunktionalen Qualifikationen in einen handlungstheoretischen und praxisorientierten Kontext einzubetten.
- Beide Herangehensweisen müssen die Person (charakterliche Eigenschaften), allgemeine, umweltorientierte Anforderungen sowie die Unternehmenssituation (z.B. gesetzliche Vorgaben, internationale Verflechtungen, Anforderungen der Branche und des Marktes, etc.) als auch die spezifischen Unternehmensanforderungen (bedingt durch die verschiedenen organisatorischen Merkmale) berücksichtigen.

## **2.3 Ansätze zur Kategorisierung der Schlüsselqualifikationen**

Die dargestellten theoretischen und empirischen Ansätze zur Bestimmung der extrafunktionalen Qualifikationen stellen einen Grundbaustein für die folgende Klassifizierung der erforderlichen Kompetenzen dar. In der Literatur lassen sich vielfältige Kategorisierungsansätze finden, die einen wertvollen Beitrag zur Ordnung der Schlüsselqualifikationen geleistet haben.

### **2.3.1 Ausgewählte Konzepte zur Klassifizierung der Schlüsselqualifikationen**

Der Ordnungsvorschlag von Laur-Ernst (1990) differenziert analytisch drei Klassen von Schlüsselqualifikationen. Ein besonderer Aspekt bei diesem Ansatz ist, dass er sowohl Aussagen über eine grundlegende Systematisierung trifft als auch Lehr-Lernbedingungen zum Erwerb und zur Vermittlung von Schlüsselqualifikationen berücksichtigt. Die drei Klassen werden unterteilt in (a) interdisziplinäres, überlappendes Wissen, (b) methodisches, verfahrens- und verhaltenstechnisches Können und (c) persönlichkeitsbezogene Fähigkeiten und Bereitschaften. Unter interdisziplinärem Wissen versteht Laur-Ernst zwischenberufliches bzw. berufsübergreifendes Wissen, das dem Berufseinsteiger oder Arbeitnehmer die Möglichkeit geben soll, mit Kollegen aus anderen Berufsbereichen oder Abteilungen effektiv und effizient zu kooperieren. Mit anderen Worten bezeichnet diese Kategorie Sachwissen, das sowohl berufsnah als auch berufsübergreifend ist, wie z.B. psychologische Kenntnisse für Personalberater, produktionstechnische Grundkenntnisse für Kaufleute, etc. sowie elementares Wissen zur Arbeitssicherheit, Umweltschutz, Arbeitsorganisation (z.B. Konzepte der Arbeitsteilung), betrieblichen Mitbestimmung und allgemeinberufliche Kenntnisse über Gesetzmäßigkeiten, Verfahrensweisen, Prinzipien und Begriffe anderer Berufsbereiche. Mit der zweiten Kategorie sind Vorgehensweisen, Techniken und Strategien gemeint, die auch auf den erfolgreichen Erwerb methodischer berufsübergreifender Kompetenzen abzielen. Das verfahrenstechnische Können inkludiert hierbei die Anwendung von Informations- und Kommunikationstechniken, wie z.B. grundlegende PC-Kenntnisse über Office, Outlook, Internet, Intranet, Zugriff auf Datenbanken, etc. Außerdem zählen zu dieser Kategorie beispielsweise der Erwerb von Präsentationstechniken, der Umgang mit Konflikten sowie die geeignete Strukturierung und Moderation von Diskussionen und anderen Gesprächstypen. Die persönlichkeitsbezogenen Fähigkeiten hingegen lassen keine einheitliche bzw. präzise Bestimmung zu. Sie sind vielmehr ein Konglomerat

verschiedener Eigenschaften, zu denen Kommunikationsfähigkeit, Kreativität, Lernmotivation, soziale Sensibilität, Selbstvertrauen und kontextuelles Denken gehören. Die komplexen Fähigkeiten werden als Orientierung und Leitideen verstanden, die vom Lernenden in Auseinandersetzung mit der Umwelt entwickelt werden (vgl. Stangel-Meseke, 1994).

Der Ansatz von Bunk (1981) bezieht sich in seinen Grundzügen auf Mertens (1974) Konzept der Schlüsselqualifikationen, wobei die personale Komponente zusätzliche Berücksichtigung findet. Der Ausgangspunkt seiner Überlegungen ist der ständige technische, wirtschaftliche und soziale (gesellschaftliche) Wandel, der die Herausbildung von antizipativen Kompetenzen nötig macht. Der angesprochene wirtschaftliche Wandel umfasst dabei die Veränderungen im Primären, Sekundären und Tertiären Sektor. Der technische Wandel betrifft vor allem die innovativen informationstechnologischen Entwicklungen als auch Verfahren zur rationellen Güterproduktion und Güterverteilung, wohingegen der soziale Wandel auf die sich verändernden Strukturen und Formen des menschlichen Zusammenlebens abzielt (wie z.B. politisches Gemeinwesen, Familienpolitik, Bildung, Kultur, soziale Schichtung, Freizeitgestaltungen, etc.). Unter Bezugnahme auf die unterschiedlichen Bereiche gesellschaftlichen Wandels nimmt Bunk (1981) wiederum eine Dreiteilung der notwendigen Qualifikationen vor, die entweder berufsbezogen, berufsübergreifend oder einen „nicht-fachlichen“ Charakter haben. Die berufsbezogenen Fähigkeiten beinhalten berufsspezifische, beruflich weiterführende und polyvalente Einstellungen, Kenntnisse und Fertigkeiten. Berufsübergreifende Qualifikationen schließen die Bereitschaft und Fähigkeit zur tätigkeitsbezogenen Flexibilität, Mobilität, Partizipation und Akzeptanz mit ein. Mit „nicht-beruflichen“ Qualifikationen sind zum Einen elementare, außerfachliche Kenntnisse und eine solide Allgemeinbildung (-wissen) gemeint und zum Anderen werden grundlegende Einstellungen und Einsichten gegenüber Politik, Kultur, Ethik, etc. als erforderliche Kompetenzen hervorgehoben. Darauf aufbauend entwickelte Bunk (1981, S.264-265, 1982, S.193) eine Matrix für Arbeitsqualifikationen, die den Grundbaustein für die weitere Differenzierung in materielle, formale und personale Qualifikationselemente bildet (Bunk, Kaiser & Zedler 1991).

In Anlehnung an Bunk (1981, 1982, 1991) und auf der Grundlage weitreichender Analysen und Studien verschiedener Klassifizierungsansätze aus Verwaltung, Wirtschaft und Wissenschaft übernimmt Stangel-Meseke (1994) die oben genannte dreiteilige Gliederung. Zusammenfassend enthält die Klassifikation vorrangig Aspekte und

Elemente der Intelligenz, des Wissens und der Persönlichkeit. Unter materiellen Schlüsselqualifikationen subsumiert die Autorin berufsbezogene und interdisziplinäre Fertigkeiten und Kenntnisse. Die Fähigkeiten sollen zudem eine große Breitenwirkung und Beständigkeit aufweisen als auch zukunftsorientiert angelegt sein. Als Beispiele werden ein fundiertes Allgemeinwissen, grundlegendes informationstechnologisches „Know-How“, elementare EDV-Kenntnisse oder die Beherrschung von Kulturtechniken, Fremdsprachen sowie wirtschaftliche und technische Grundkenntnisse angeführt. Die formalen Qualifikationen umfassen vor allem psychomotorische und kognitive Fähigkeiten. Ersteres beinhaltet beispielsweise Reaktions- und Umstellungsleistungen oder manuelle Fertigkeiten. Die kognitiven Fähigkeiten bezeichnen problemlösende, transferfähige und antizipative Denk- und Lernleistungen als Voraussetzung zum selbständigen, anwendungsbezogenen Handeln oder zur Erschließung neuer Wissensgebiete. Die personalen Aspekte beziehen sich auf Werthaltungen und affektive Persönlichkeitsmerkmale. Dazu gehören nach Stangel-Meseke (1994) individuelle Verhaltensweisen (z.B. Motivation, Anpassungsbereitschaft), soziale Verhaltensweisen (z.B. Kooperations-, Kommunikations- und Teamfähigkeit, Fairness), Grundüberzeugungen und Tugenden (z.B. Ehrlichkeit, Pünktlichkeit, Zuverlässigkeit) sowie sittliche und ethische Verhaltensweisen.

Ein weiteres nützliches und eher praxisorientiertes Klassifikationskonzept liefern Eilles-Mathiessen, el Hage, Janssen und Osterholz (2002). Im Vergleich zu den anderen Ansätzen weist dieses Konzept eine stärkere Verknüpfung zum Kompetenzkonzept auf, dass im nächsten Unterkapitel detaillierter betrachtet wird. Demzufolge werden insbesondere solche Schlüsselqualifikationen berücksichtigt, die im Rahmen der berufsbezogenen Eignungsdiagnostik anhand von psychologischen Messinstrumenten ermittelt werden können. Die Klassifikation der Schlüsselqualifikationen erfolgt hier nach verschiedenen Anforderungsebenen. Dazu zählen (a) Basisanforderungen, (b) Anforderungen im Umgang mit Anderen und (c) Anforderungen an spezielle Situationen. Unter Basisanforderungen verstehen die Autoren generelle Fähigkeiten, die für unterschiedliche berufliche Tätigkeiten und Positionen eine allgemeingültige Relevanz aufweisen.



Es werden folgende Qualifikationsmerkmale genannt (Eilles-Mathiessen et al., 2002, S.81ff):

- Intellektuelle Fähigkeiten  
(Konzentrationsfähigkeit, schnelle Auffassungsgabe, kreatives Denken, Problemlösefähigkeit, räumliches Vorstellungsvermögen, systematisch-analytisches Denken)
- Motivation und Engagement  
(Zielstrebigkeit, Durchhaltevermögen, Eigeninitiative, Leistungs- und Lernbereitschaft)
- Handlungskompetenzen  
(Belastbarkeit, Entscheidungsfähigkeit, realisierungsorientiertes Denken, Selbstmanagement, Gewissenhaftigkeit)

Durch den konzeptuellen Einbezug intellektueller Kompetenzen, wie z.B. systematisch-analytisches, abstraktes und schlussfolgerndes Denken, Problemlösefähigkeiten oder räumliches Vorstellungsvermögen, wird deutlich, dass insbesondere klassische Facetten der Intelligenz für berufsorientierte Tätigkeiten hervorgehoben werden.

Unter der zweiten Anforderungsebene werden diejenigen Schlüsselqualifikationen subsumiert, die zur erfolgreichen Bewältigung sozialer Interaktion Kooperation und Kommunikation beitragen. Dazu zählen:

- Soziale Kompetenzen  
(Durchsetzungsfähigkeit, Empathie, Konfliktmanagement, Teamfähigkeit, *Networking*, soziales Wissen über Normen und Standards,
- Kommunikative Kompetenzen  
(Fremdsprachenkenntnisse, Artikulation)
- Führungskompetenzen  
(Delegationsfähigkeit, Feedbackfähigkeit, Animationsfähigkeiten, Verantwortung, Zielorientierung)

Die dritte Ebene umfasst Fertigkeiten und Kenntnisse, die einerseits für spezielle Situationen und Positionen von Bedeutung sind und andererseits tätigkeitsübergreifenden Charakter haben:

- Interkulturelle Kompetenzen  
(Kulturelles Wissen über Geschichte, Politik, Geografie, Kunst, etc., kulturelle Aufgeschlossenheit, Respekt und Toleranz, Ambiguitätstoleranz)
- Unternehmerische Kompetenzen  
(Kundenorientierung, kaufmännisches Wissen und Denken, Verhandlungsgeschick, Risikobereitschaft, visionäres/strategisches Denken und Handeln)
- Arbeitstechniken  
(Informationsmanagement, Moderations- und Präsentationstechniken, Projektmanagement, Zeitmanagement)

Resümierend werden auch in diesem Konzept vor allem die kognitiven Fähigkeiten bzw. Aspekte der Intelligenz als auch Bereiche der Allgemeinbildung und des Wissens als notwendige, entscheidende und übergeordnete Schlüsselqualifikation in den Vordergrund gestellt.

### **2.3.2 Bewertung der Ansätze zur Kategorisierung der Schlüsselqualifikationen**

Zuerst werden die theoretischen Ansätze hinsichtlich ihrer Stringenz von Bezugspunkt und der gewählten Kategorisierung beurteilt. Danach werden die Unterschiede und Gemeinsamkeiten der dargestellten Ansätze verdeutlicht. Der Ordnungsvorschlag von Laur-Ernst (1990) bezieht sich in erster Linie auf die Flexibilität und Offenheit der Bildungsziele. In diesem Ansatz wird das Kategorienschema aus der Formulierung von Bildungszielen und der möglichen Gestaltung beruflicher Lehr- und Lernprozesse abgeleitet, so dass die Kategorisierung beim aktuellen Stand der Forschung seine Bedeutung entfaltet. Es ist jedoch fraglich und nicht weiter begründet, warum dieser eher undeutliche Bezugspunkt eine Klassifikation von Schlüsselqualifikationen zulässt (vgl. auch Stangel-Meseke, 1994). Laur-Ernst (1990) räumt hingegen ein, dass sie sich aufgrund der uneinheitlichen Begriffsverwendung nicht auf einen Qualifikationskatalog oder methodisch-didaktisches Konzept für Schlüsselqualifikationen festlegt, sondern den Ordnungsvorschlag offen konzipiert. Bunk (1981, 1982) rechtfertigt das Erfordernis von Schlüsselqualifikationen mit dem gesellschaftlichen Wandel und entwickelte daraufhin

ein Schema mit drei Kategorien: berufsbezogene, berufsübergreifende und nicht-fachliche Qualifikationen. Einerseits werden die Kategorien hinsichtlich des erläuterten Wandels inhaltlich schlüssig dargestellt und fundiert. Andererseits geht diese Stringenz bei der aufgestellten Matrix für Arbeitsqualifikationen verloren, indem die nicht-fachlichen Qualifikationen keine Berücksichtigung findet. Eine adäquate Weiterentwicklung erfährt die Matrix durch die Dreigliederung in eine materiale, formale und personale Kategorie (Bunk et al., 1991). Der Klassifikationsvorschlag von Eilles-Mathiessen et al. (2002) stützt sich vorrangig auf praxisorientierte, empirische Studien an verschiedenen Institutionen. Die Systematisierung lässt eine Stringenz von Bezugspunkt und der Kategorisierung erkennen. Außerdem enthält die Klassifikation wertvolle Hinweise für die Zuordnung von geforderten Schlüsselqualifikationen oder überfachlichen Anforderungen zu psychologischen Messverfahren, indem Definitionen und Indikatoren (Verhaltensbeispiele) derselben aufgelistet werden.

Im Hinblick auf die Gemeinsamkeiten und Unterschiede fällt auf, dass eine dreiteilige Kategorienstruktur anhand von Oberkategorien bei vielen Ansätzen verwendet wird (Stangel-Meseke, 1994). Weiterhin kann auch eine grundlegende inhaltliche Übereinstimmung konstatiert werden. Im Bezug auf Bunk et al. (1991) lässt sich darüber hinaus feststellen, dass die Systematisierung eindeutige Analogien zu anderen Kategorisierungsansätzen aufweist oder umgekehrt. Dasselbe gilt auch für die inhaltliche Charakterisierung der einzelnen Kategorien. Genauer gesagt, werden unter materialen Fähigkeiten und Kenntnissen Qualifikationselemente angegeben, die vornehmlich wissensbasierte, berufsübergreifende und langfristige Eigenschaften haben. Formale Kompetenzen beinhalten Lern- und Denkfähigkeiten (z.B. schlussfolgerndes, abstraktes, logisches Denken) und die personalen Qualifikationen inkludieren soziale und individuelle Verhaltenselemente. Parallel zu dieser Einteilung bezeichnet Laur-Ernst (1990) die erste Kategorie als überlappendes Wissen, die formalen Fähigkeiten entsprechen der Kategorie methodisches, verfahrens-technisches Können und die persönlichkeitsbezogenen Fähigkeiten können mit der Klasse Verhalten und Einstellung gleichgesetzt werden. Eine ähnliche Kategorienstruktur zeigt sich auch im Vergleich zum Ansatz von Eilles-Mathiessens et al. (2002). Die Kategorie „Anforderungen für spezielle Situationen“ gleicht den materialen Fähigkeiten aufgrund der eher wissensorientierten Kompetenzen. Die Basisanforderungen weisen eindeutige Analogien zu den formalen Fähigkeiten auf, da hier vorwiegend Denk- und Lernleistungen angesprochen werden.

Der Bereich „Anforderungen im Umgang mit Anderen“ kann den personalen Fähigkeiten zugeordnet werden.

Ein gravierender Unterschied zu dem Ansatz von Laur-Ernst (1990) besteht darin, dass Bunk et al. (1991) und Eilles-Mathiessen et al. (2002) die erforderlichen Schlüsselqualifikationen vor allem aus den Anforderungen der Berufspraxis ableiten, wohingegen Laur-Ernst (1990) sich stärker auf die Bildungsziele und bildungspolitischen Inhalte bezieht. Hinsichtlich der teils unterschiedlichen inhaltlichen Differenzierung von Schlüsselqualifikationen kommt es letztendlich auch auf die individuelle Zielsetzung der Unternehmung oder anderer Institutionen an. Demzufolge sollte die Bestimmung von Schlüsselqualifikationselementen in einem unternehmens- und ausbildungsspezifischen Kontext betrachtet werden, bei dem je ein anderer Grad der Operationalisierung des angewendeten Kategorisierungsschemas möglich ist. Dementsprechend erscheint eine praxisorientierte, empirische Eruiierung von Schlüsselqualifikationen sinnvoller als theoretische Herangehensweisen. Zusammenfassend lässt sich festhalten, dass die Dreigliederung in materiale, formale und personale Fähigkeiten eine geeignete Basis für die Kategorisierung von Schlüsselqualifikationen bietet, sofern die inhaltliche Ausgestaltung bzw. der Einbezug von anderen Qualifikationselementen offen angelegt ist. Ein wichtiger und durchgängiger Aspekt, der in diesem Bezugsrahmen hervorgehoben werden sollte ist, dass sowohl kognitive Fähigkeiten bzw. Intelligenzfacetten, wie schlussfolgerndes, abstraktes, logisches Denken und der Bereich des (Allgemein-) Wissens bei vielen Klassifikationsansätzen integriert und als übergeordnete Schlüsselqualifikationen gesehen werden können.

## **2.4 Der Kompetenzbegriff - Weiterentwicklung des Konzepts der Schlüsselqualifikationen**

Auf der Grundlage weitreichender Studien zum Konzept extrafunktionaler Qualifikationen (s.o.) ist heutzutage zunehmend der Begriff „Kompetenz“ Gegenstand nationaler und internationaler Forschungen geworden. Betrachtet man beispielsweise die Anzahl der Veröffentlichungen in der Deutschen Nationalbibliothek in Frankfurt, werden seit 2004 unter dem Schlagwort Kompetenz 740 Publikationen aufgelistet. Die Literatur umfasst dabei Studien zu allgemeinen Diagnose- und Messverfahren über die Erforschung bestimmter Kompetenzaspekte bis hin zu berufspolitischen Konzepten zur

Kompetenzerfassung und Kompetenzentwicklung. Die weitverbreitete Würdigung des Kompetenzkonzepts lässt sich damit begründen, dass es in weiten Teilen anschlussfähig ausgelegt ist, verschiedene Sichtweisen zusammenführt und Ansätze aus der Schlüsselqualifikationsforschung integriert. Das Konzept ermöglicht eine Verknüpfung von allgemeinen und berufsbildenden Bereichen, es verbindet individuelle und betriebliche Perspektiven und eröffnet der Kompetenzfeststellung im Rahmen formalisierter Bildungskontexte neue Möglichkeiten. Außerdem vertritt das Kompetenzkonzept eine stärkere outputorientierte Anschauungsweise, so dass die Idee des lebenslangen Lernens transparenter wird (vgl. Gnahs, 2007).

#### **2.4.1 Kompetenz- eine Begriffsbestimmung**

Wie bereits angedeutet zeigt das Kompetenzkonzept deutliche Analogien zu den Ansätzen der Schlüsselqualifikationen auf. Vor diesem Hintergrund gewinnen auch die bereits dargestellten Konzepte an Bedeutung und sollten bei der Begriffsbestimmung berücksichtigt werden. Anhand der Etymologie lässt sich der Begriff Kompetenz aus dem ursprünglichen lateinischen Nomen „competentia“ (Zusammentreffen) und dem Verb „competere“ (zu etwas fähig sein, ausreichen, zustehen, zusammentreffen) ableiten (Langenscheidt Wörterbuch, 2007). Aus der Übersetzung kann man bereits entnehmen, dass die Kompetenzen einer Person die Fähigkeit geben, in einer bestimmten Situation adäquat zu handeln. Mit anderen Worten können Kompetenzen als individuelles Potential an Fertigkeiten und Kenntnissen verstanden werden, die beim „Zusammentreffen“ mit situativen Erfordernissen eine geeignete Handlung des Individuums implizieren.

Als Vorreiter in der wissenschaftlichen Kompetenzforschung können White (1959) und Mc Clelland (1973) genannt werden, die den Begriff erstmals in die Psychologie einführten. White (1959) brachte den Kompetenzbegriff insbesondere mit der Motivationspsychologie in Verbindung. Für White ist die Kompetenz das Resultat einer vom Individuum selbstorganisierten Herausbildung fundamentaler Fähigkeiten. Demgemäß wird die Kompetenz als Voraussetzung für Leistung (hier Performanz) interpretiert, die die Person im Laufe der Zeit durch die Interaktion mit seiner Umwelt entwickelt. Daraus lässt sich entnehmen, dass die Bestimmung der Kompetenz auch motivationale Aspekte einschließt. Eine ähnliche Herangehensweise nimmt auch Chomsky (1973) vor, der Kompetenzen als Potential begreift, eine bestimmte Performanz zu erreichen. Weinert (2001) bezeichnet Kompetenz als „die bei Individuen verfügbaren oder durch sie

erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können.“(S.27f). Eine weitere Definition geben Süß, Weiß und Seidel (2005), die unter Kompetenz das Potential verstehen, „erforderliches Verhalten in einer gegebenen Situation zeigen zu können“ (S. 350). Nach Staudt, Kriegesmann und Muschik (2003) ist Kompetenz die „Grundlage von Handlungen bzw. Aufgabenerfüllungen und basiert auf individueller Ebene auf einem Zusammenspiel der

- Handlungsfähigkeit (explizites, implizites Wissen und Fertigkeiten) als kognitive Basis,
- Handlungsbereitschaft als motivationale Basis und
- Zuständigkeit als organisatorische Legitimation und Einbindung in den Unternehmenskontext“ (S.160).

Eine vergleichsweise weitverbreitete Definition, die bisweilen große nationale und internationale Akzeptanz findet, wurde im Rahmen des OECD-Projekts DeSeCo (Defining and Selecting Key Competencies) entwickelt: „A competence is defined as the ability to successfully meet complex demands in a particular context. Competent performance or effective action implies the mobilization of knowledge, cognitive and practical skills, as well as social and behavior components such as attitudes, emotions and values and motivations. A competence- a holistic notion-is therefore not reducible to its cognitive dimension, and thus the terms of competence and skill are not synonymus“ (OECD, 2003, S.2).

In Analogie zu dem Ansatz von White (1959) bezeichnen Erpenbeck und Heyse (2007) Kompetenzen als *Selbsorganisationsdisposition* einer Person. Kompetenzen befähigen das Individuum Handlungen selbst motiviert auszuführen, deren Resultate bedingt durch die Situation, den Handlungsprozess und der Komplexität der Person nicht gänzlich vorhergesagt werden können. Demnach haben die Handlungen reflexiven Charakter und beziehen sich beispielsweise auf aktivitätsbetonte Tätigkeiten und Selbsteinschätzungen. Unter den geistig-instrumentellen Handlungen werden kreative Denkprozesse oder Problemlösungsprozesse subsumiert, während kommunikative Handlungen Verkaufstätigkeiten, soziale Interaktionen oder Selbstprofilierung umfassen. Mithin sind Dispositionen nicht beobachtbare, innere Antezedenzen bzw. Voraussetzungen, die eine Tätigkeit regulieren als auch individuelle Eigenschaften sowie Entwicklungsergebnisse

beinhalten. Die Dispositionen, die letztendlich für das selbstorganisierte Handeln nötig sind, werden nunmehr als Kompetenzen bezeichnet. Diesbezüglich unterscheiden Erpenbeck und von Rosenstil (2007) vier verschiedene Kompetenzklassen bzw. Schlüsselkompetenzen:

#### 1. Personale Kompetenzen

Personale Kompetenzen sind Dispositionen, die von einem Individuum dazu gebraucht werden, reflexiv selbstorganisiert zu agieren und selbstkritisch mit sich umzugehen. Damit sind Persönlichkeitseigenschaften gemeint, wie zum Beispiel die Entwicklung von Werthaltungen, Selbstbildern und Motiven, Motivation, Selbstsicherheit, Zielstrebigkeit, Engagement, Kreativität, etc.

#### 2. Sozial-kommunikative Kompetenzen

Diese Kompetenzen werden als soziale Kommunikationsvoraussetzungen verstanden. Sie sollen die Person einerseits dazu befähigen, kooperativ und kommunikativ zu handeln und andererseits dienen die Eigenschaften dazu, teamorientiert und partnerschaftlich zu fungieren sowie die soziale Interaktion mit anderen effektiv zu gestalten. In diesem Sinne definiert auch Wenninger (2001) soziale Kompetenz als Sammelbegriff „für solche Wissensbestandteile, Fähigkeiten, und Fertigkeiten einer Person, die eine notwendige Voraussetzung für die Bewältigung sozialer Interaktion darstellen“ (S.197).

#### 3. Fachlich-methodische Kompetenzen

Die fachlich-methodischen Kompetenzen schließen geistige und physische Dispositionen ein, die zur Lösung von sachlich-gegenständlichen Problemen beitragen. Selbstorganisiertes Handeln wird nach den Autoren nur dann möglich sein, sobald die Person imstande ist, Wissen sinnorientiert zu erfassen und zu bewerten sowie funktionale und fachspezifische Kenntnisse, Fähigkeiten und Fertigkeiten zur kreativen Problemlösung einzubringen.

#### 4. Aktivitäts- und umsetzungsorientierte Kompetenzen

Unter dieser Klasse werden Dispositionen subsumiert, die auf das aktive und gesamtgesellschaftliche Handeln eines Individuums abzielen. Ferner geht es um die erfolgreiche Realisierung von Absichten, Plänen und Handlungen, wobei die Person sämtliche Persönlichkeitseigenschaften, wie Motivation, eigene Emotionen, Erfahrungen sowie die anderen dargestellten Kompetenzen in den intrinsisch veranlagten Willensantrieb mit einfließen lassen muss.

Letztendlich können die geforderten Dispositionen auch unter der Oberkategorie Handlungskompetenz zusammengefasst werden, die nach Holling und Liepmann (2007) „alle beruflichen Kenntnisse, Fertigkeiten beziehungsweise Kompetenzen und Fähigkeiten [beinhaltet], die als Voraussetzung für eine erfolgreiche Ausübung der beruflichen Aufgaben fungieren“ (S.345). Die Klasse der Handlungskompetenz integriert demnach die aufgeführten Kompetenzarten oder anders ausgedrückt bildet sie die Schnittmenge, in der alle Kompetenzen für die erfolgreiche Ausführung einer Aufgabe zusammenwirken. Gnahs (2007) macht in seinem Ansatz zudem deutlich, dass sich die beobachtbare Performanz bzw. die Handlungskompetenz einer Person erst aus dem Zusammenspiel von Mensch und Situation entfaltet. Anhand der Abbildung 1 sollen die grundlegenden Zusammenhänge transparenter gemacht werden. Jede Person verfügt über ein spezifisches Potential an Dispositionen, Fähigkeiten, Fertigkeiten, Motivationen, Werte und Wissen, die von den physischen und genetischen Grundlagen abhängig sind und je nach individuellen biografischen Erfahrungen variieren können. In einer bestimmten (Problem-) Situation werden die Kompetenzen entsprechend den Rahmenbedingungen und Anforderungen eingesetzt und das Individuum reagiert mit Handeln. An diesem Punkt zeigt sich, inwieweit das vorhandene Potential Anwendung findet und welche Handlungskompetenz die Person mitbringt. Es ist ersichtlich, dass alle Elemente interdependent aufeinander bezogen sind und in ihrer Gesamtheit zur Performanz beitragen. Im Weiteren sollen die aufgeführten Elemente näher dargestellt werden (vgl. Gnahs, 2007). Unter der Komponente „Dispositionen“ werden verschiedene Persönlichkeitseigenschaften subsumiert, deren Merkmale als relativ stabil im Lebenslauf gesehen werden. Eine breite Akzeptanz haben in diesem Zusammenhang die folgenden Persönlichkeitsdimensionen bzw. die so genannten *Big Five* gefunden: *Neurotizismus*, *Extraversion*, *Verträglichkeit*, *Gewissenhaftigkeit* und *Offenheit für neue Erfahrungen*. Das Element „Fertigkeiten“ bezieht sich vor allem auf die sensomotorischen und physischen Aspekte der individuellen Leistungsfähigkeiten. Dazu gehören beispielsweise Fingerfertigkeiten, handwerkliches Geschick, Beherrschung von Arbeitstechniken, Sprachvermögen und die Nutzung aller Sinnesorgane. Im Gegensatz zu den eher stabilen persönlichen können die Fertigkeiten durch Training optimiert bzw. verbessert werden. Der Faktor „Werte“ bezeichnet grundsätzliche Einstellungen des Individuums gegenüber ideologischen Anschauungen, Verhaltensweisen, Personen (-gruppen) und allgemeinen Dingen. Die Werte können dabei politischen, religiösen oder kulturellen Charakter haben, die im organisatorischen oder auch familiären Kontext zum Tragen kommen. Die Motivation zielt auf



die inneren Antriebskräfte und Interessen des Individuums ab. Handlungsauslösende Motive können dabei intrinsischer und extrinsischer Natur sein. Die Komponente „Wissen“ umfasst sowohl Faktenwissen sowie überfachliches Allgemeinwissen, das von einer Person in verschiedenen Situationen abgerufen werden kann. Die Wissensbestände sind, ähnlich den Fertigkeiten, grundsätzlich veränderbar, so dass der Ausspruch des lebenslangen Lernens an Bedeutung zunimmt. Der Fähigkeitsbegriff setzt sich im engeren Sinn aus den Komponenten Wissen und Fertigkeiten zusammen und zielt damit auf das Zusammenspiel jener ab. So schließt die Lesefähigkeit beispielsweise die Fertigkeit ein, Töne zu produzieren und das erworbene Wissen wird genutzt, um die Bedeutung von Wörtern und Sätzen zu erkennen. Im weiteren Sinn kann der Fähigkeitsbegriff auch mit dem Begriff der Kompetenz gleichgesetzt werden. Erpenbeck und v. Rosenstil (2003) definieren die Fähigkeit als „verfestigte Systeme verallgemeinerter psychophysischer Handlungsprozesse (...), einschließlich der zur Ausführung einer Tätigkeit oder Handlung erforderlichen inneren psychischen Bedingungen (...) und der lebensgeschichtlich erworbenen Eigenschaften (...)“ (S. XXIX).

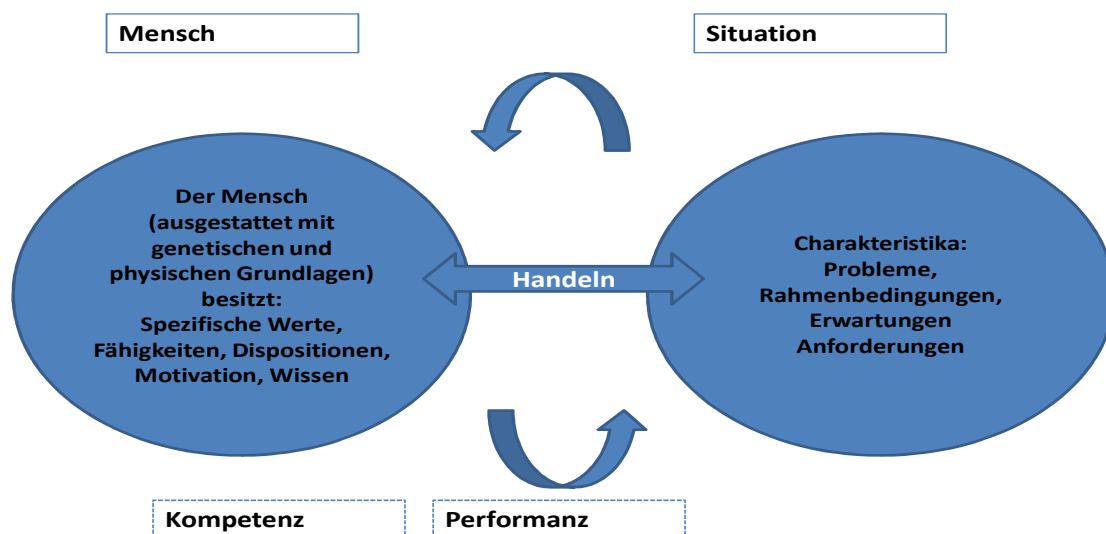


Abbildung 1: Kompetenz und Performanz, nach Gnahn, 2007

Mithin lassen sich in der Literatur weitere Differenzierungen des Kompetenzbegriffs ausmachen. Eine häufig verwendete Aufspaltung ist die in Fachkompetenzen und überfachliche Kompetenzen, zu denen auch die Sozialkompetenz (soziale Interaktion), Methodenkompetenz und die personale Kompetenzen (personenbezogene Eigenschaften).

Eine weitaus interessantere Aufteilung hinsichtlich der obigen Ausführungen ist die Unterscheidung in Spezial- und Schlüsselkompetenzen. Die Fachkompetenz inkludiert Kenntnisse und Fähigkeiten, die auf spezialisierte berufsorientierte oder private Bereiche bezogen sind. Die Schlüsselkompetenzen hingegen können als Synonym für das bereits dargestellte Konzept der Schlüsselqualifikation aufgefasst werden (Gnahs, 2007).

#### **2.4.2 Bewertung des Kompetenzkonzepts im Hinblick auf das Konzept der Schlüsselqualifikationen**

Obwohl das Kompetenzkonzept bisweilen eine breite Resonanz findet und dem Konzept eine hohe Wichtigkeit beigemessen wird, ist es gleichermaßen Kristallisationspunkt für Auseinandersetzungen. Insbesondere die Debatte über die vermeintlichen Unterschiede zu den Begriffen wie „Qualifikation“, „Bildung“ oder „Humankapital“ ist hier von Relevanz. Bezogen auf die Unterschiede zwischen Qualifikationen und Kompetenzen werden Qualifikationen im Allgemeinen als Wissens- und Fertigkeitsdispositionen verstanden, die in geplanten Bildungsprozessen unterrichtet werden, wohingegen Kompetenzen als Dispositionen selbstorganisierten Handelns interpretiert werden. Im Bereich der beruflichen Bildungsprozesse wird die konzeptionelle Vermittlung von Inhalten auf eine praktische Verwertbarkeit ausgerichtet, so dass eine Person zu kompetentem Handeln befähigt wird. Im Ergebnis kann der Erfolg anhand von Tests evaluiert werden. Die Qualifikation ist demnach eine prüfungsrelevante Bestandsaufnahme des aktuellen Wissens und der vorhandenen Fertigkeiten. Im Gegensatz dazu können die Kompetenzen bzw. Selbstorganisationsdispositionen nicht ohne weiteres durch Prüfungen und Zertifizierungen erfasst werden (Gnahs, 2007; Erpenbeck & Rosenstil, 2003, 2007).

Hinsichtlich der Analogien zwischen Schlüsselqualifikationen und Kompetenzen bestätigen Hartig und Klieme (2006), dass eine konzeptionelle Trennung nur bedingt möglich ist, weil die Abgrenzung beider Konstrukte vor allem im Bereich der Schlüsselkompetenzen verwischt. Genauer gesagt wird die begriffliche Trennung schwierig, sobald der entsprechende anforderungsspezifische- und situationsspezifische Kontext breiter wird. Die Autoren sind der Auffassung, dass vielfältige Verflechtungen mit dem Konzept der Schlüsselqualifikationen eminent sind und in die Begriffsbestimmung der Kompetenz mit einfließen. Gnahs (2007) unterstützt diese Sichtweise und verweist auf die engen Verbindungen zu dem von Mertens eingeführten Begriff der Schlüsselqualifikation.

Trotz der teils uneinheitlichen Definitionen des Kompetenzbegriffes und der Schlüsselqualifikationen lässt sich festhalten, dass Elemente wie Wissen und kognitive Fähigkeiten (Intelligenz) in nahezu allen Ansätzen einbezogen und im Besonderen hervorgehoben werden, so dass die psychologischen Konstrukte als übergeordnete Dimensionen unterschiedlicher extrafunktionaler Qualifikationen und (Schlüssel-) Kompetenzen verstanden bzw. interpretiert werden können.

### **2.4.3 Ausblick auf die Bedeutung der Schlüsselkompetenzen bzw. Schlüsselqualifikationen**

Das Konzept extrafunktionaler Qualifikationen findet nicht nur in wissenschaftlich-psychologischen und betriebswirtschaftlichen Studien besondere Beachtung, auch von den Medien, Politikern, Wirtschaftsvertretern und Bildungsinstitutionen erhält das Konzept der Schlüsselqualifikationen in der Berufspraxis und der beruflichen Ausbildung einen hohen Stellenwert und wird dort eingehend diskutiert.

Gnahs (2007) stellt in diesem Zusammenhang heraus, dass „das inhaltlich-konzeptionelle Substrat des Kompetenzansatzes zu einem politisch wichtigen strategischen Hebel [wird], der gleichermaßen bedeutsam ist für die persönliche, soziale und ökonomische Entwicklung“ (S.11). Auch vor dem Hintergrund der Globalisierung, dem stetigen technologischen Wandel und der fortschreitenden Entwicklung zur Wissensgesellschaft reagiert die EU (genauer der Europäische Rat) mit Setzungen strategischer Ziele, um den bevorstehenden Herausforderungen adäquat zu begegnen. Die übergeordneten strategischen Ziele im Zuge der zu forcierenden Kompetenzentwicklung umfassen dabei die persönliche Entfaltung, gesellschaftliche Teilhabe und Beschäftigungsfähigkeit (Europäischer Rat, 2000, Tz. 5). Hierzu unterbreiten der Europäische Rat und das Europäische Parlament die folgende Empfehlung vom 18.12.2006 „die Bemühungen der Mitgliedstaaten (...) [zu] unterstützen, [so] dass junge Menschen nach der Grundbildung und Ausbildung die Schlüsselkompetenzen erworben haben, die sie für das Erwachsenenleben rüsten und eine Grundlage für das weitere Lernen sowie das Arbeitsleben bilden, und dass Erwachsene ihre Schlüsselkompetenzen ein Leben lang weiterentwickeln und aktualisieren können“ (Europäisches Parlament, ABI. L 394 vom 30.12.2006, S.13). Die „Kompetenzkonjunktur“ zeigt sich bereits auf internationaler Ebene und wurde beispielsweise in Frankreich unter dem Oberbegriff „bilans de compétences“ (vgl. Drexel, 1997) sowie in Italien/Südtirol (Haller, 2003) in Form des Kompetenzpasses (KOMPASS) als bildungs- und marktpolitisches Instrument eingeführt.

Auch im nationalen Raum der Bundesrepublik Deutschland wurde das Konzept der Schlüsselqualifikationen bereits in Novellierungen der Berufsbildungsverordnungen verschiedener Ausbildungszweige integriert, wie z.B. in der Neuordnung der Metall- und Elektroberufe von 1987 (vgl. Kultusministerkonferenz [KMK], 1991) oder im Rahmen verschiedener Ansätze zur berufsqualifizierenden Ausbildung Jugendlicher von der Bundesanstalt für Arbeit (2003) nahe gelegt. Außerdem wurde das Konzept im Rahmen des Bologna-Prozesses als Grundlage für die aktuelle Studienreform mit einbezogen (Barz, Hopbach & Meyer-Guckel, 2004). Darauf aufbauend unterstreicht die Arbeitsgemeinschaft für betriebliche Weiterbildungsforschung (ABWF) folgendes: „Im internationalen Wettbewerb sind hohe Kompetenzen der Mitarbeiter ein nachhaltiger Wettbewerbsvorteil. Deutschland verfügt über solche Ressourcen. Es geht darum, diese weiterzuentwickeln und zu erschließen“ (Arbeitsgemeinschaft betriebliche Weiterbildungsforschung e.V. [ABWF], 2006, S.1).

Daran anknüpfend ist es ebenso ersichtlich, dass für eine zielorientierte Kompetenzentwicklung geeignete Instrumente der Kompetenzerfassung entwickelt und eingesetzt werden müssen. Nach der Kommission der Europäischen Gemeinschaften (2001) „ist [demnach] ein neues Konzept zur Lernbewertung notwendig, um Brücken zwischen verschiedenen Lernkontexten und Lernformen zu schlagen und den Zugang zu individuellen Lernwegen zu erleichtern“ (S.16). In einer Empfehlung des Europäischen Parlaments und des Rates vom 18.12.2006 wird ergänzend vorgeschlagen, „die Schlüsselkompetenzen zu ermitteln und zu definieren, die in einer Wissensgesellschaft für persönliche Entfaltung, aktive Bürgerschaft, sozialen Zusammenhalt und Beschäftigungsfähigkeit nötig sind“ (Europäisches Parlament und Rat, ABI. L 394 vom 30.12.2006, S.13). In diesem Kontext lässt sich ein Wandel von der „Input- zur Outcome-Orientierung“ feststellen, die vor allem die gewonnenen Lernergebnisse und deren Verwendungsmöglichkeiten in den Vordergrund stellt (*learning outcomes*) (vgl. Gnahn, 2007). Ordnungspolitische Anstrengungen, die in diese Richtung gehen, sind zum Beispiel der „European Qualifications Framework-EQF“ (vgl. Kommission der Europäischen Gemeinschaften, 2006), das European Credit System for Vocational Education and Training-ECVT (vgl. Bundesministerium für Bildung und Forschung [BMBF] 2006, S.321f) oder der von der OECD (2005) initiierte PIAAC-Prozess (Kompetenzerfassung bei Erwachsenen). Die aufgeführten Ansätze dienen in erster Linie dazu, Kompetenzen bzw. Qualifikationen sichtbar und international vergleichbar zu machen. Entsprechende Aktionsprogramme wurden auch von der Bundesregierung angestoßen.

Laut dem Bundesministerium für Forschung und Bildung (2001) soll demzufolge ein Netzwerk sämtlicher Bildungsbereiche geschaffen werden, in denen „gemeinsam die Zertifizierung selbstgesteuerter Lernerfolge für externe Bewerber/innen“ entwickelt und erprobt werden soll bzw. „Verfahren zur Messung und Bewertung individueller Kompetenzentwicklung (...) erarbeitet werden sollen“(S. 6).

Im Hinblick auf die obigen Ausführungen wird deutlich, dass insbesondere die Erfassung des (Allgemein-) Wissens und der kognitiven Fähigkeiten eine hohe Relevanz für die Diagnostik bzw. Kompetenzerfassung besitzen.

An dieser Schnittstelle erhält die zu entwickelnde START-Testbatterie für Berufsanfänger als auch der Intelligenz-Struktur-Test 2000 R eine außerordentliche Bedeutung. Anhand der Testinstrumente sollen die genannten Schlüsselkompetenzen erfasst und operationalisiert werden.

Der zu validierende Allgemeinwissenstest (START-W) wird in der Zukunft Bestandteil dieser umfassenden Testbatterie für Berufseinsteiger werden, die verschiedene, anerkannte Schlüsselkompetenzen (z.B. Fremdsprachenkenntnisse, Rechenfertigkeiten, Konzentrationsfähigkeit, Allgemeinwissen, etc.) integrieren soll. Der Wissenstest wird ein herausragendes Verfahren sein, das über die bestehenden Instrumente zur Erfassung des Allgemeinwissens hinausgeht. Neben der Erfassung relevanter Themenbereiche des Allgemeinwissens wird die verbale, numerische und figurale Inhaltsskala berücksichtigt und zudem bietet der Test die Möglichkeit, die kristallisierte Intelligenz abzubilden.

### **3 Psychologische Diagnostik**

In den folgenden Kapiteln soll die psychologische Diagnostik, genauer gesagt die berufsbezogene Eignungsdiagnostik oder Personaldiagnostik, in seinen Grundzügen erläutert werden. Den Ausgangspunkt bilden die Definitionen bzw. eine Abgrenzung des Gegenstandsbereiches. Danach werden die übergeordnete Zielsetzung, die Vorbedingungen und die personalpolitischen Teilschritte der Berufseignungsdiagnostik dargelegt.

Darauffolgend werden psychologische Testverfahren vorgestellt, die im Rahmen der Eignungsdiagnostik von besonderer Relevanz sind. Aufgrund der Vielfalt von Testinstrumenten beinhaltet dieser Gliederungspunkt eine übersichtliche Klassifizierung eignungsdiagnostischer Instrumente. Das Hauptaugenmerk liegt indessen auf der Kategorie

der Leistungstests und den dazugehörigen Gütekriterien. Daran anknüpfend werden Methoden und Konzepte zur Diagnostik von Schlüsselkompetenzen dargelegt, wobei der Fokus auf den Intelligenz- und Wissenstests liegt. Abschließend wird die Bedeutung der Eignungsdiagnostik für die Personalpolitik und im weiteren Sinne für die Organisation hervorgehoben.

### **3.1 Abgrenzung und Zielsetzung des Gegenstandsbereiches**

Für die Abgrenzung des Gegenstandsbereiches bietet sich zunächst ein etymologischer Einstieg an. Der Terminus Diagnostik leitet sich aus dem griechischen Wort „diagnóskein“ ab, welches im Allgemeinen verschiedene Elemente eines kognitiven Prozesses bezeichnet. In der wortgetreuen Übersetzung bedeutet das Verb „genau kennenlernen“, „entscheiden“ und „beschließen“. Aus den Grundbedeutungen lässt sich entnehmen, dass vom Psychologen oder Diagnostiker bestimmte Leistungen erwartet werden, die das „genaue Kennenlernen“ des menschlichen Verhaltens mit seinen Persönlichkeitsmerkmalen einschließen und vernünftige Schlussfolgerungen vom Diagnostiker abverlangen. In der klassischen Literatur wird Psychologische Diagnostik als Methodenlehre verstanden, die der Angewandten Psychologie zuzuordnen ist. Ringelband und Birkhan (2000) definieren die Diagnostik als „Vorgehen, in dem menschliche Verhaltensdaten erhoben und auf der Grundlage von theoretisch-psychologischen Annahmen so interpretiert werden, daß sie eine Erklärung für vergangene und eine Voraussage für zukünftige Verhaltensweisen erlauben“ (S. 796).

Die Berufseignungsdiagnostik, als Spezialgebiet der Psychologischen Diagnostik, wird hingegen als „Methodologie der Entwicklung, Prüfung und Anwendung psychologischer Verfahren zum Zwecke eignungsdiagnostischer Erfolgsprognosen und Entscheidungshilfen im beruflichen Kontext“ (Schuler & Höft, 2006, S. 102) verstanden. Berufsbezogene Eignungsdiagnostik bezeichnet demnach ein Kollektiv für Methoden und Verfahren zur Erfassung von Potentialen, Verhaltensweisen, Kompetenzen, Interessen und Kompetenzen, deren erklärtes Ziel die Prognose des Berufserfolges ist (Schuler, 2007; Kanning, 2004). Die dahinterstehende Grundüberlegung ist, dass Personen aufgrund ihrer Fähigkeiten für eine bestimmte Arbeitsstelle gut geeignet sind und die geforderte Tätigkeit erfolgreich ausführen können.

Die Aufgabe der Berufseignungsdiagnostik besteht nun darin, die Fähigkeiten aufzudecken und die Kompatibilität zwischen der Person und dem beruflichen Aufgabenbereich zu gewährleisten.

Um die Zweckmäßigkeit der beruflichen Eignungsdiagnostik begründen zu können, müssen daher zunächst die folgenden vier Voraussetzungen gegeben sein (Schuler, 2000):

1. Die zu besetzende Stelle bzw. Tätigkeit wird von verschiedenen Personen unterschiedlich gut ausgeführt.
2. Es besteht ein Zusammenhang zwischen den differenten Merkmalen der Personen und dem Ausprägungsgrad der beruflichen Leistung.
3. Die Merkmale einer Person sind entweder konstant oder können in ihrer Variabilität vorhergesagt bzw. gemessen werden.
4. Die Organisation kann eine Moderatorfunktion übernehmen: Anders ausgedrückt sollte die Organisation die Möglichkeit besitzen, eine passende Organisationsstruktur (Kommunikationsstruktur, Führungsverhalten, Organisationsklima, etc.) zu gewährleisten, um die Personen nachhaltig weiterentwickeln zu können.

Auf Basis dieser Voraussetzungen gliedert sich die Bildung eines zuverlässigen psychologischen Eignungsurteils in fünf personalpolitische Teilschritte (Brambring, 1983):

1. Bestimmung der Anforderungsmerkmale für den zu besetzenden Arbeitsplatz.
2. Bestimmung der individuellen Eignungsvoraussetzungen (Fähigkeiten, Fertigkeiten, Intelligenz, Wissen) für die definierten Anforderungsmerkmale.
3. Bestimmung der passenden psychologischen Verfahren.
4. Durchführung von Validitätsuntersuchungen hinsichtlich des Eignungsurteils und der Berufsbewährung.
5. Beurteilung der Effizienz der getroffenen Entscheidungen.

Eine ähnliche Herangehensweise verfolgen Schuler und Höft (2006). Sie schlagen vor, das Individuum mit seinen Persönlichkeitsmerkmalen und die Anforderungen der auszuführenden Tätigkeit auf drei Ebenen zu vergleichen. Die drei Ebenen werden indessen in tätigkeitspezifische Anforderungen und Qualifikationen, tätigkeitsübergreifende Qualifikationen sowie in Befriedigungsangebot und –bedarf unterteilt.

Demzufolge ergeben sich drei Aspekte (s. Tabelle 1):

1. Vergleich der tätigkeitsspezifischen Anforderungen mit den Kenntnissen, Fertigkeiten und Fähigkeiten der Person.
2. Ermittlung tätigkeitsübergreifender Anforderungen der Arbeit und deren Abgleich mit dem Entwicklungspotential und den allgemein erfolgsrelevanten Merkmalen.
3. Feststellung des Befriedigungspotentials der Position und dessen Vergleich mit den Bedürfnissen, Interessen und Werthaltungen der Person.

Hierzu werden wiederum auf drei Ebenen Informationen zur Berufstätigkeit kollektivierte:

1. Aufgaben- und Ergebnisanforderungen  
(Auflistung der Tätigkeiten und erforderlichen Resultate)
2. Verhaltensanforderungen  
(Angaben über Verhaltensweisen zur erfolgreichen Berufsausübung)
3. Eigenschaftsanforderungen  
(Notwendige Personenqualifikationen)

Zur Informationssammlung der stellenrelevanten Anforderungen werden in der Praxis die erfahrungsgeladete, intuitive Methode (Befragung von Mitarbeitern, Experten), die personenbezogene empirische Methode (Analyse von Personenmerkmalen erfolgreich arbeitender Menschen) und die arbeitsplatzanalytische empirische Methode (formalisierte Methode mittels Fragebögen, *Critical-Incidents-Technique*) favorisiert (Schuler, 2007; Eckhardt & Schuler, 1992). Die Anforderungsanalyse dient im Weiteren zur Entwicklung der eignungsdiagnostischen Instrumente (siehe darauf aufbauend Kapitel 3.2.1) und leistungsbezogenen Kriterien.

Eine detailliertere Darstellung über die unterschiedlichen Methoden der Arbeits- und Anforderungsanalyse gibt Schuler (2006a).

Zusammenfassend besteht das grundlegende Ziel einer guten Eignungsdiagnostik schließlich darin, die Fähigkeiten, Fertigkeiten, Kenntnisse, Bedürfnisse, Interessen und das (Entwicklungs-) Potential einer Person mit Hilfe geeigneter wissenschaftlicher, psychologischer Testverfahren zu bestimmen und diese in Einklang mit den definierten Anforderungen zu bringen (Schuler, 2006a; Frieling & Sonntag, 1999).



Mit anderen Worten sollen Zusammenhänge zwischen Persönlichkeitsmerkmalen und dem Berufserfolg anhand von adäquaten Messinstrumenten untersucht werden, um fundierte Wahrscheinlichkeitsaussagen abgeben zu können (siehe Tabelle 1).

*Tabelle 1: Trimodale Beziehung zwischen Anforderungen, Persönlichkeitsmerkmalen und Testverfahren (in Anlehnung an Schuler, 2007)*

Tätigkeit	Diagnostische Verfahren (z.B. Leistungstests zur Ermittlung der Kompatibilität)	Person
Tätigkeitsspezifische Anforderungen	↔	Fähigkeiten, Fertigkeiten, Kenntnisse
Tätigkeitsübergreifende Anforderungen	↔	Erfolgsrelevante Eigenschaften, Entwicklungspotential
Befriedigungspotential	↔	Interessen, Bedürfnisse, Werthaltungen

Das kurze Resümee stellt jedoch eine Idealvorstellung dar. In der Realität steht die Personal- bzw. Eignungsdiagnostik allerdings vor einer schwierigen Aufgabe und erfordert von allen Beteiligten eine hohe methodische Kompetenz.

Ausgangspunkt jeder guten berufsbezogenen Eignungsdiagnostik ist vorerst eine fundierte Anforderungsanalyse. Denn „jeder Personalauswahlprozess [ist] maximal so gut wie das Anforderungsprofil differenziert und realistisch ist“ (Müllerschön, 2005, S. 15). Das Anforderungsprofil determiniert somit die erforderlichen Fähigkeiten und gibt gleichzeitig Hinweise über ein geeignetes Testverfahren zur Erfassung der geforderten Persönlichkeitsmerkmale.

Eine Möglichkeit über die Eignung einer Person zu urteilen, bilden psychologische Testverfahren (z.B. Intelligenztests, Wissenstests, Konzentrationstests, Tests motorischer und sensorischer Leistung, etc.). Brandstätter (1979) bezeichnet die Tests im Rahmen der Eignungsdiagnostik als „standardisierte, routinemäßig anwendbare Verfahren zur Messung individueller Verhaltensmerkmale, aus denen Schlüsse auf Eigenschaften der betreffenden Person oder auf ihr Verhalten in anderen Situationen gezogen werden können“ (S. 82). Darüber hinaus werden Tests als Instrumente verstanden, mit denen psychometrisch gültige und vergleichbare Informationen gewonnen werden.

Die Variationsbreite von möglichen Testverfahren in der Personaldiagnostik ist dabei mannigfaltig. Die eignungsdiagnostischen Testverfahren können in diesem Zusammenhang bei allen berufsbezogenen bzw. personalpolitischen Arbeitsfeldern, wie der Auswahl, Platzierung und Entwicklung von Mitarbeitern sowie bei Berufsberatungen eingesetzt werden und eine wichtige Entscheidungshilfe darstellen. Insbesondere in der Personalauswahl gewinnt diese Selektionsstrategie an Bedeutung, um „für eine spezifische Bedingung aus einer Reihe von Personen passende auszusuchen oder aber für eine Person aus einer Reihe von Bedingungen die passende zu finden“ (Gebert & Rosenstiel, 2002, S. 260). An diesem Punkt wird auch die Bedeutung einer guten und wissenschaftlich fundierten Eignungsdiagnostik für die Organisation deutlich. Mithin kann eine „falsche“ Zuordnung von Personen zu einer Arbeitsstelle erfahrungsgemäß Frustrationen und mangelnde Motivation hervorrufen. Eine sorgfältig durchgeführte Diagnostik und die darauffolgende Besetzung eines vakanten Arbeitsplatzes mit der „richtigen“ Person (*person-job-fit*) induziert hingegen monetäre Vorteile und erhält vor allem eine übergeordnete immaterielle und soziale Relevanz. Aus Unternehmenssicht führt die optimale Korrespondenz zwischen Fähigkeiten und den Anforderungen beruflicher Tätigkeiten und die damit einhergehende Berücksichtigung von Erwartungen und Bedürfnissen der potentiellen Mitarbeiter nicht nur zu wirtschaftliche Vorteilen, sondern erhöht auch die Arbeitszufriedenheit, das *job-commitment* und die Arbeitsleistung (vgl. z.B. Schuler, 2000 2006a, 2007; Gnahs, 2007).

### **3.2 Testverfahren im Rahmen der personalpolitischen Eignungsdiagnostik**

Im Folgenden wird zunächst eine Klassifizierung eignungsdiagnostischer Testverfahren vorgestellt. Danach wird die Kategorie der Leistungstests im Besonderen dargestellt, da sie eine Relevanz für das Promotionsvorhaben bzw. für die Entwicklung des START-W Wissenstests (spezieller Leistungstest) besitzt. In diesem Zusammenhang werden die Begrifflichkeit und die dahinterstehende Testtheorie in ihren Grundzügen erläutert sowie ausgewählte Leistungstests dargelegt. Außerdem werden die allgemein anerkannten Testgütekriterien und Qualitätsstandards im Sinne der DIN Norm 33430 (Berufsverband Deutscher Psychologen, 2002) für berufsbezogene Eignungsbeurteilungen beschrieben.

### 3.2.1 Klassifizierung eignungsdiagnostischer Testverfahren

Hinsichtlich der Klassifikationsmöglichkeiten gibt es unterschiedliche Ansätze, die auf verschiedenen Zweckmäßigkeitsgesichtspunkten beruhen. Schumacher und Brähler (2005) begreifen beispielsweise die Klassifikationsvorschläge als Ordnungssysteme, die einerseits eine didaktische Funktion erfüllen und andererseits eine Orientierungshilfe für die große Bandbreite von eignungsdiagnostischen Verfahren bieten.

Auf einer allgemeinen Ebene können Testverfahren nach formalen und inhaltlichen Aspekten unterteilt werden (Brickenkamp, 1997; Brähler, Holling, Leutner & Petermann, 2002; Lienert & Raatz, 1998). Aus Gründen der praktischen Bedeutsamkeit werden von Lienert und Raatz (1998) differenziertere Kriterien zur formalen Klassifikation von Tests herangezogen:

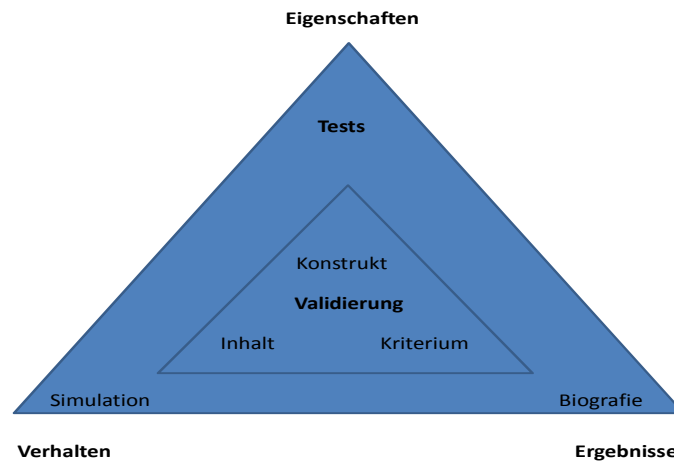
- Standardisierte Tests vs. nichtstandardisierte Tests: Standardisierte Tests bezeichnen dabei wissenschaftliche Tests, die in Standardsituationen durchgeführt werden, normiert sind und den Gütekriterien entsprechen.
- Art des zu untersuchenden Persönlichkeitsmerkmals: Das zu erfassende Persönlichkeitsmerkmal wird dabei in Intelligenztests, Leistungstests und Persönlichkeitstests untergliedert, wobei die Anzahl der Persönlichkeitsaspekte eine weitere Unterteilung in eindimensionale und mehrdimensionale Tests erlaubt.
- Eignungstests, die prognostische Absichten des Berufserfolges u.a. verfolgen.
- Zeitbemessung der Tests: Hier wird zwischen Speed-Tests, die die Schnelligkeit bei der Aufgabenlösung messen und Power-Tests, die das maximale geistige Niveau bei großzügiger Zeitvorgabe messen sollen, unterschieden.
- Psychometrische Tests (z.B. Leistungstests) versus projektive Verfahren.
- Objektive Tests vs. nichtobjektive Tests.
- Art der Durchführung: Hier wird zwischen Befragungstests, Bleistifttests und Materialbearbeitungstests unterschieden.
- Art der Testbeantwortung: Verbale vs. nonverbale Tests
- Individualtests vs. Gruppentests.
- Einfache Tests (Testergebnis ist ein Punktwert) vs. komplexe Tests (das Testergebnis setzt sich aus mehreren Punktwerten zusammen).
- Art der Bezugsgröße: Normorientierte Tests (der individuelle Testwert wird ins Verhältnis zum Normwert einer Gruppe gesetzt) vs. kriterienorientierte Tests (das Ergebnis wird mit der Gesamtzahl der Aufgaben in Relation gesetzt).

- Struktur der Aufgaben: Hochstrukturierte Tests (z.B. Intelligenztests mit einer richtigen Lösung) vs. nichtstrukturierte Tests (Kreativitätsmessungen).
- Art der Aufgabenbeantwortung: Freie vs. gebundene Aufgabenbeantwortung.

Im Gegensatz zu den formalen Kriterien umfassen die inhaltlichen Gliederungskriterien auch die individuellen Persönlichkeitsmerkmale und werden daher bevorzugt verwendet. Dementsprechend wird nach Brickenkamp (1997) eine grobe Dreiteilung vorgeschlagen, die (a) Leistungstests, (b) psychometrische Persönlichkeitstests und (c) Persönlichkeits-Entfaltungsverfahren unterscheidet (vgl. auch Bortz & Döring, 2006).

Schuler und Höft (2006) differenzieren drei methodische Ansätze zur Klassifikation eignungsdiagnostischer Instrumente. Zu denen gehören der Eigenschaftsansatz, der Simulationsansatz und der biografische Ansatz. Die Ansätze unterliegen dabei verschiedenen Methoden der Merkmalerfassung und folgen demnach einer eigenen Validierungslogik. Die Abbildung 2 veranschaulicht die Ansätze der Eignungsdiagnostik und deren Validierungslogik.

1. Der Eigenschaftsansatz bzw. Konstruktansatz bezieht sich auf die Erfassung von relativ stabilen Persönlichkeitsmerkmalen (z.B. Intelligenz, Wissen) mittels psychologischer Tests. Das Kriterium zur Überprüfung ist die Konstruktvalidität.
2. Der Simulationsansatz zielt auf die Erfassung menschlichen Verhaltens ab, wie es am Arbeitsplatz erwartet wird. Gängige Erhebungsformen sind die Arbeitsprobe oder Rollenspiele in Assessment Centern. Als Validitätsprinzip wird die Kontextvalidität genannt.
3. Der biografische Ansatz soll biografische Informationen anhand von Fragebögen, Interviews oder dem Lebenslauf erfassen. Die dem biografischen Ansatz inhärente Validitätslogik ist die prognostische Validität.



*Abbildung 2: Ansätze der Eignungsdiagnostik und deren Validierungslogik, nach Schuler und Höft, 2006*

Abschließend empfehlen die Autoren eine systematische Kombination unterschiedlicher Verfahren, um alle relevanten Facetten des Persönlichkeitsmerkmals oder des Verhaltens abzubilden. Für die berufliche Eignungsdiagnostik impliziert dieser Aspekt, dass eine hohe Prognosesicherheit des beruflichen Erfolges am besten durch den Einsatz unterschiedlicher psychologischer Methoden und Verfahren gewährleistet werden kann (Schuler und Höft, 2006; Sarges & Wottawa, 2001; Hossiep & Paschen, 1998, 2003; Schuler, 2000). Zusammengefasst proklamieren die Autoren ein multimethodales Auswahlverfahren. Der Vorteil liegt darin begründet, dass die Reliabilität, die Validität und im Gesamten die Qualität der Potentialdiagnose erhöht wird (Schuler & Höft, 2006).

In diesem Kontext zeigt auch Hossiep (1995, 1996) in seinen Forschungsbeiträgen auf, dass insbesondere die Verwendung einer umfassenden Testbatterie, die verschiedene Einzelverfahren wie beispielsweise Intelligenztests, Wissenstests oder Persönlichkeitstests integriert, den langfristigen Berufserfolg am zuverlässigsten prognostizieren kann. Hier gewinnt die zu entwickelnde START-Testbatterie für Berufseinsteiger als auch der darin enthaltende Wissenstest START-W an Bedeutung und kann als zukunftsweisendes, psychologisch fundiertes Verfahren im Besonderen hervorgehoben werden.

### 3.2.2 Der Testbegriff und die klassische Testtheorie

Eines der vorrangigen Ziele psychologischer Eignungsuntersuchungen ist es, anhand von wissenschaftlichen Methoden die Bewährungswahrscheinlichkeit einer Person für den schulischen, akademischen oder beruflichen Werdegang zu prognostizieren (Brambring, 1983). Eng verbunden mit den psychologischen Eignungsuntersuchungen steht der Begriff Test. In Anlehnung an Guthke (1972, zitiert nach Fisseni, 2004) gibt Fisseni die folgende Testbegriffsdefinition wieder: „Als Test soll gelten ein Prüfverfahren, bei dem in standardisierten Situationen Verhaltensmerkmale (Verhaltensstichproben) von Personen erfasst werden, die als Indikatoren für bestimmte Personeneigenschaften dienen sollen, und dessen Resultate eine Einordnung der Untersuchten in eine Klassifikation ermöglichen, die an einer Gruppe vergleichbarer Personen gewonnen wurde“ (S. 21). Einerseits erleichtert diese Definition das grundlegende Verständnis über die Bedeutung des Begriffes Test, andererseits wird die Bedeutung auch durch jene Definition eingeschränkt. Lienert und Raatz (1998) geben in diesem Kontext eine weitere Begriffsdefinition: „Ein Test ist ein wissenschaftliches Routineverfahren zur Untersuchung eines oder mehrerer empirisch abgrenzbarer Persönlichkeitsmerkmale mit dem Ziel einer möglichst quantitativen Aussage über den relativen Grad der individuellen Merkmalsausprägung“ (S. 1). Die Bedeutung des Testbegriffs wird in diesem Kontext auf die Bestimmung eines Persönlichkeitsmerkmals gelenkt. Es sei jedoch darauf hingewiesen, dass dem Begriff auch andere Bedeutungen inhärent sind. In der einschlägigen Literatur werden jedoch sechs grundlegende Charakterisierungen bzw. Leitvorstellungen von Tests anerkannt (Fisseni, 2004; Lienert & Raatz, 1998; Scheurer, 1999):

1. Der Test wird als routinemäßig anwendbares Prüfverfahren verstanden, das in seiner Durchführung, Auswertung und Interpretation standardisiert ist, um die Objektivität und Vergleichbarkeit der Resultate zu gewährleisten.
2. Der Test ist ein Prüfverfahren, das Personen hinsichtlich Merkmalsausprägungen differenziert. Abgrenzbare Fähigkeiten, Fertigkeiten, Verhaltensdispositionen, Eigenschaften oder Kenntnisse können demnach empirisch überprüft werden.
3. Der Test prüft Verhaltensmerkmale ab, die auf Personeneigenschaften übertragen werden können.
4. Die Testergebnisse eines eignungsdiagnostischen Verfahrens sollen einen Indikator für zukünftiges Verhalten darstellen.

5. Tests dienen als Hilfsmittel einer Querschnittsdiagnose (relative Positionsbestimmung eines Individuums innerhalb seiner Bezugsgruppe mittels geeigneter Norm- und Eichstichproben, Bestimmung der Gradausprägung einer Leistung zwischen Individuen und Gruppen, Ermittlung von Leistungsprofilen), Längsschnittsdiagnose (Detektion von Merkmalsveränderungen) und der Forschung (empirischer Nachweis lokaler und globaler Merkmalszusammenhänge).
6. Bei der Entwicklung von psychologischen Prüfverfahren müssen unabhängig von der gewählten Methode bestimmte, wissenschaftlich anerkannte Gütekriterien berücksichtigt werden.

Die Entwicklung von Tests bzw. die Konstruktion von Tests unterliegt indessen den Regeln der sogenannten Testtheorie. Damit einhergehend wird eine Theorie über psychologische Tests entwickelt, die dabei helfen soll, psychische Eigenschaften oder Merkmale von Individuen zu erfassen und Rückschlüsse auf zukünftiges Verhalten zulässt. Drei Aspekte spielen diesbezüglich eine Rolle (Rost, 2004, S.17):

1. Ein wissenschaftlich fundierter Test ist dadurch gekennzeichnet, dass eine Theorie darüber besteht, inwieweit und unter welchen Bedingungen aus den Testergebnissen Aussagen über die getestete Person generiert werden können.
2. Mithilfe von Tests werden relativ stabile und konsistente Persönlichkeitsmerkmale erfasst, das für das im Test gezeigte Verhalten verantwortlich ist.
3. Der Test sollte ein Routineverfahren darstellen, das an einer großen Stichprobe bereits erprobt, ausgewertet und detailliert beschrieben wurde und von anderen Testleitern umgesetzt werden kann.

Die Genese eines Tests und die damit zusammenhängende Testtheorie soll anhand der folgenden tabellarischen Übersicht verdeutlicht werden:

*Tabelle 2: Genese eines Tests  
(Fisseni, 2004, S. 27)*

Genese eines Tests - Übersicht
<b>Entwurf:</b>
<ul style="list-style-type: none"> <li>- Sichtung theoretischer Ansätze und empirischer Befunde, die zum Thema vorliegen;</li> <li>- Konzeptualisierung und Ausarbeitung der Fragestellung;</li> <li>- Festlegung der Merkmale, die erfasst (oder miterfasst) werden sollen:               <ul style="list-style-type: none"> <li>➤ Befragung von Experten</li> <li>➤ Befragung von Mitgliedern der Zielgruppe;</li> </ul> </li> <li>- Bestimmung der Testart, Entwurf von Testaufgaben (Items) und Instruktionen: Testvorform.</li> </ul>
<b>Erprobung:</b>
<ul style="list-style-type: none"> <li>- Vorlage der Testvorform bei einer Stichprobe, die der Zielgruppe möglichst ähnlich ist:               <ul style="list-style-type: none"> <li>➤ Zur Ermittlung missglückter Itemformulierungen,</li> <li>➤ Zur Analyse der Testaufgaben (Itemanalyse);</li> </ul> </li> <li>- Auswertung.</li> </ul>
<b>Revision:</b>
<ul style="list-style-type: none"> <li>- Selektion, Elimination oder Überarbeitung der Items aufgrund von Erprobungsergebnissen,</li> <li>- gegebenenfalls Wiederholung der Erprobung,</li> <li>- Erstellung eines revidierten Tests.</li> </ul>
<b>Endfassung:</b>
<ul style="list-style-type: none"> <li>- Ziehung einer angemessenen Stichprobe der Zielgruppe,</li> <li>- Vorlage und Auswertung des revidierten Tests,               <ul style="list-style-type: none"> <li>➤ Bestimmung von Standardisierung (Objektivität), Messgenauigkeit (Reliabilität) und Gültigkeit (Validität),</li> <li>➤ Festlegung genereller und, sofern möglich, spezieller Normen,</li> </ul> </li> <li>- Interpretation der Ergebnisse unter Berücksichtigung der methodischen Beschränkungen, Vergleich mit den Ergebnissen, die auf ähnlichen oder die auf andersartigen Methoden beruhen.</li> </ul>



### 3.2.3 Leistungstests in der Eignungsdiagnostik

Mit Blick auf spezielle Testverfahren, nämlich den Leistungstests, sollen entsprechend den obigen Ausführungen, Verhaltensmerkmale einbezogen werden, die beobachtbar und damit messbar sind. Gleichzeitig werden die zu messenden Konstrukte (z.B. Intelligenz und Wissen) im Rahmen der Eignungsdiagnostik bzw. Kompetenzerfassung einerseits operationalisiert und andererseits voneinander abgegrenzt, spezifiziert und definiert. Anders formuliert werden geeignete Tests zur systematischen Messung der Intelligenz und des Wissens konstruiert. Die Merkmale Intelligenz und Wissen werden mit Aufgabenstellungen verbunden, die immer eine richtige und eine falsche Lösung beinhalten. Das Ziel ist also die Gewinnung von messbaren Daten, um die einzelnen Individuen bezüglich der Merkmale miteinander vergleichen zu können bzw. die interindividuellen Unterschiede sichtbar zu machen. Die Eignungsdiagnostik kann somit als angewandter Teil der Differentiellen Psychologie verstanden werden (Lienert & Raatz, 1998; Fisseni, 2004, Amelang & Schmidt-Atzert, 2006). Die Voraussetzung dafür ist, dass die Leistungen an eine bestimmte Situation, Ort und Aufgabenstellung gekoppelt sind. Aufgrund der ermittelten Daten können dahinter liegende Persönlichkeitsmerkmale abgeleitet werden, die nicht direkt in der Testsituation erfasst werden. Es werden demnach Verhaltensausschnitte gemessen, die repräsentativ für das Verhalten außerhalb der Testsituation sind (Brickenkamp, 1997). Darauf aufbauend definiert Fisseni (2004) die Leistungstests als „eine Gruppe von Verfahren, die einem Probanden eine Leistung abverlangen, in der Anteile des Zielmerkmals „realisiert“ werden“ (S. 177). Auf Basis der ermittelten Leistungen können sodann Rückschlüsse auf zugrunde liegende Fähigkeiten und Fertigkeiten, Eignungen und Begabungen gezogen werden. Die Leistung bestimmt sich darüber hinaus durch die maximal erbrachte Arbeit in einer bestimmten Zeit. Im Rahmen von psychologischen Testverfahren drückt sich die Leistung bspw. in der Lösung von vorgegebenen Aufgabenstellungen aus. Letztendlich soll die Leistung „jene „Anstrengung“ charakterisieren, die ein Proband aufbringt, um das Zielmerkmal eines Tests zu „realisieren“ [Hervorhebung v. Verf.]“ (Fisseni, 2004, S.176).

#### *Überblick über ausgewählte Arten von Leistungstests*

Im Weiteren steht die Klassifikation von Leistungstests im Vordergrund der Betrachtung. Sarges und Wottawa (2001) unterteilen die Kategorie der Leistungstests beispielsweise in (a) Wissens- und Kenntnistests, (b) allgemeine Wissenstests, (c) kogniti-

ve Tests, (d) psycho- und sensomotorische Tests, (e) Aufmerksamkeits- und Konzentrationstests, (f) computergestützte Simulationen und (g) in Belastungstests. Nach der Einteilung von Fisseni (2004) werden die Leistungstests in allgemeine und spezielle Leistungstests unterteilt. Die allgemeinen Leistungstests sollen die Fähigkeiten messen, die Bestandteil jeder Leistung sind. Hierzu zählen bspw. die Konzentrationstests. Spezielle Leistungstests zielen auf die Identifikation bestimmter Persönlichkeitsmerkmale ab, die sich aus besonderen Anforderungen ergeben. Es sei angemerkt, dass die Intelligenzleistung aufgrund ihrer weitreichenden, geschichtlichen Entwicklung und ihrer Bedeutung als eigenständige Klasse aufgefasst werden, obwohl sie prinzipiell den allgemeinen Leistungstests zugeordnet werden können. Eine andere Unterteilung der Leistungstests wird in Form von *Speed- und Powertests* vorgenommen. Bei den Speed Tests wird durch eine begrenzte Zeitvorgabe sowohl die Qualität der erbrachten Leistung als auch die Geschwindigkeit gemessen. Die Power Tests erfassen hingegen nur, ob die Aufgaben richtig oder falsch gelöst wurden, und nicht wie hoch der Zeitaufwand für die Lösung der Aufgaben war (Rost, 2004; Amelang & Schmidt-Atzert, 2006). Im Folgenden wird jedoch der Einteilung nach Fisseni (2004) gefolgt.

### ***Intelligenztests***

Entsprechend der Ausführungen in Kapitel 4 wird die Intelligenz als Konstrukt verstanden, das durch die Theorie- und Forschungsentwicklung sukzessive erschlossen wird und zum Zeitpunkt des aktuellen Forschungsstandes seinen Bedeutungsgehalt entfaltet. Im Bezug auf den Bedeutungsgehalt des Konstrukts Intelligenz spielen die dargestellten Strukturtheorien eine maßgebliche Rolle (siehe Kapitel 4.4). Die Strukturtheorien bilden danach auch die Grundlage für die verschiedenen Testkonstruktionen. Darauf aufbauend unterscheiden sich die Tests hinsichtlich der Annahme verschiedener Faktoren, in der Annahme eines g-Faktors sowie in der Akzentuierung der beteiligten Prozesse. In diesem Kontext wird in der Eignungsdiagnostik zwischen Tests zur Erfassung der allgemeinen Intelligenz und Tests zur Erfassung spezieller Intelligenzkomponenten differenziert, wobei die zuletzt genannten Tests häufig eine Summierung der Einzelergebnisse zu einem einzigen Kennwert, im Sinne eines allgemeinen Intelligenzwertes, zulassen (Brambring, 1983; Fisseni, 2004). Aufgrund der Tatsache, dass nicht alle Testautoren dieselbe Auffassung über das Konzept der allgemeinen Intelligenz verfolgen, lassen sich nach Amelang und Schmidt-Atzert (2006) oder Bortz und Döring (2006)

grundsätzlich zwei Ansätze unterscheiden. Die einen Ansätze und die dazu gehörigen Tests beziehen sich auf den sogenannten Kernbereich der Intelligenz, nämlich das schlussfolgernde Denken (*Reasoning*) bzw. die von Bildungseinflüssen bereinigte fluide Intelligenz. Die anderen rekurren auf kulturgebundene Tests, die kulturspezifisches Hintergrundwissen sowie verbale Fähigkeiten verlangen. Eine typische Form zur Bestimmung der kristallisierten Intelligenz ist der Wissenstest. Darüber hinaus werden in der eignungsdiagnostischen Praxis überwiegend Strukturtests eingesetzt, die verschiedene Bereiche der Intelligenz abbilden können. Gemeint sind die beiden Generalfaktoren der Intelligenz und zwar die fluide und die kristallisierte Intelligenz. Diese Arten von Tests erlauben differenzierte Informationen über verschiedene Intelligenzkomponenten, wie z.B. verbales, numerisches und figürliches Denken und Wissen. Das Ziel ist die Erstellung eines „Intelligenzprofils“. Einen Überblick über verschiedene Intelligenztests geben zum Beispiel Brähler, Holling, Leutner und Petermann (2002).

Die allgemeinen Intelligenztests weisen bezüglich der Prognose von schulischen Leistungen besonders hohe Korrelationen auf und gelten demnach als gute Prädiktoren (zum Überblick, Jensen, 1980, 1998; Fisseni, 2004; Schuler & Höft, 2006; Schuler, 2000). Hinsichtlich des beruflichen Ausbildungserfolges wird ebenfalls ein eindeutiger Zusammenhang zur Intelligenz nachgewiesen. Zusammenfassend besitzen die Intelligenztests hohe prädiktive Validitäten bezüglich der Vorhersage des Ausbildungserfolges (Schmidt & Hunter 1998). Aus diesem Grund haben die allgemeinen Intelligenztests einen hohen Stellenwert im Rahmen der Eignungsdiagnostik und werden daher vielfältig eingesetzt (Schuler & Höft, 2006; Schuler, 2000). Die übergeordnete Bedeutung von Intelligenztests in der Eignungsdiagnostik wird im Kapitel 3.4 dargelegt. Eine detailliertere Vorstellung des Intelligenztests I-S-T 2000 R von Liepmann et al. (2007) wird im empirischen Teil der Ausführungen gegeben.

### ***Konzentrationstests***

Nach der Einteilung von Fisseni (2004) werden die Konzentrationstests unter den allgemeinen Leistungstests subsumiert. Die Bedeutung der Konzentrationstests offenbart sich nach Bartenwerfer (1983) in der Erfassung der Fähigkeit eines Individuums, „eine der richtigen Aufgabenlösung dienende angemessene „innere Grundlage“ zu schaffen und über die erforderliche Zeit hinweg aufrecht zu erhalten“ (S. 485).

Bartenwerfer geht es somit um die Messung der allgemeinen Voraussetzungen für erfolgreiches Verhalten in verschiedenen Situationen.

Als Beispiel wird der Aufmerksamkeits- Belastungs-Test- („d2“) von Brickenkamp (2002) genannt. Der Aufmerksamkeits-Belastungs-Test stellt ein Verfahren zur Erfassung der individuellen Aufmerksamkeit und Konzentrationsfähigkeit. Er findet Verwendung in nahezu allen psychologischen Arbeitsbereichen, beispielsweise in der Pädagogischen Psychologie (Erziehungsberatung), Schulpsychologie, Arbeits- und Betriebspsychologie (Berufsberatung, Eignungsdiagnostik), Verkehrspsychologie, Klinischen Psychologie etc.. Der Test d2 stellt eine standardisierte Weiterentwicklung der sogenannten Durchstreichtests dar. Er misst Tempo und Sorgfalt des Arbeitsverhaltens bei der Unterscheidung ähnlicher visueller Reize (Detail-Diskrimination) und ermöglicht damit die Beurteilung individueller Aufmerksamkeits- und Konzentrationsleistungen. Die vielfältige Absicherung der Testgütekriterien (Objektivität, Zuverlässigkeit und Gültigkeit) sind Ursache dafür, dass der Test d2 zu den am häufigsten verwendeten psychodiagnostischen Verfahren in Deutschland gehört. Gemäß der Handanweisung werden dem Probanden vierzehn Buchstabenreihen vorgegeben. Für jede Zeile wird dem Probanden 20 Sekunden zur Bearbeitung Zeit gegeben, um jedes „d“ in einer Reihe ähnlicher Zeichen durchzustreichen, das mit zwei Strichen gekennzeichnet ist.

### ***Wissenstests***

Gemäß der Klassifizierung von Brickenkamp (1997) lassen sich die Wissenstests in die Kategorie der speziellen Leistungstests einordnen. Angesichts der hohen Relevanz für den empirischen Teil der Arbeit sollen die Wissenstests an dieser Stelle besondere Erwähnung finden. Prinzipiell kann in diesem Bezugsrahmen zwischen Allgemeinwissenstests und Fachwissenstests differenziert werden. Bezüglich der Fachwissenstests werden drei grundlegende Verfahren unterschieden: (a) schriftliche Fachwissenstests, (b) handlungsbezogene Wissenstests und (c) situationsbezogene Wissenstests (Salgado, 1999). Der schriftliche Fachwissenstest kann bestimmte Wissensgebiete abprüfen, wie z.B. Biologie, Geschichte, Mathematik, Technik, Wirtschaft, oder auch bereichsspezifisches Wissen. Die Ermittlung des handlungsbezogenen Wissens umfasst bspw. Fragestellungen zu vorgegebenen (Text-) Aufgaben, die berufsbezogene (Problem-) Situationen darstellen bzw. beschreiben. Hinsichtlich der situationsbezogenen Wissenstests sollen die Probanden Lösungen für komplexe und komplizierte Situationsbedingungen

finden. Es wird jedoch angemerkt, dass die Abgrenzung zwischen Allgemeinwissen und Fachwissen äußerst schwierig ist, da die Definitionen in der Literatur uneinheitlich sind. Die Grenzen der Gegenstandsbereiche sind somit nicht trennscharf und sollten eher als Konstrukte verstanden werden, die sich gegenseitig bedingen. Die einschlägige Literatur ist sich indessen einig, dass wissenschaftliche Testverfahren einen Induktionsschluss über den Wissensstand eines Individuums zulassen (siehe z.B. Hossiep & Schulte, 2008). Darüber hinaus wird darauf verwiesen, dass der Wissenstest einerseits als eigenständiges Instrument zur Vorhersage von schulischen oder beruflichen Leistungen eingesetzt werden kann.

Als älteres Beispiel kann hier der Differentielle Wissenstest (DWT) angeführt werden (Jäger & Fürntratt, 1968). Der DWT kann sowohl die Allgemeinbildung einer Person bestimmen als auch Aussagen über Wissensschwerpunkte in einzelnen Themenbereichen machen. Der DTW setzt sich aus 11 Subtests zusammen, die jeweils 20 Fragestellungen zu den folgenden Wissensgebieten umfassen: Sport, Geldwesen, Politik, Technik, Physik und Chemie, Biologie, Erdkunde, Geschichte, Literatur, Kunst und Musik.

Als aktuelles Beispiel kann der Bochumer Wissenstest (BOWIT) von Hossiep und Schulte (2008) genannt werden, der seit 2008 in der praktischen Anwendung ist. Der BOWIT liegt in seiner aktuellen Version in zwei Formen vor (Form A und B). Aufgrund der Effizienz und dem Interesse der Praxis werden zusätzlich zwei Kurzversionen angeboten. Die Vollversion umfasst auf der Grundlage umfangreicher Voruntersuchungen (aus einem Pool von 792 Testitems) 154 Items, die den beiden Formen A und B zugeführt wurden. Die Kurzversionen enthalten 45 der trennschärfsten Items. Auf der Basis von Internetrecherchen, lexikalischen Analysen und Expertenbefragungen setzt sich die Vollversion aus 11 Facetten mit jeweils 14 Items zusammen: Bildende Kunst/ Architektur, Biologie/ Chemie, Ernährung/ Bewegung/ Gesundheit, Geographie/ Verkehr, Geschichte/ Archäologie, Gesellschaft/ Politik, Mathematik/ Physik, Philosophie/ Religion, Sprache/ Literatur, Technik/ EDV und Wirtschaft/ Recht. Für die 11 Facetten konnten anhand von Faktorenanalysen (Hauptkomponentenanalyse, oblimin) zwei Faktoren ermittelt werden (64% Prozent erklärte Varianz), die als „naturwissenschaftlich-technisches Wissen“ und „gesellschafts- und geisteswissenschaftliches Wissen“ benannt werden. Die Reliabilität des Gesamttests wird durchweg mit einem Wert größer als .90 angegeben. Die Retest-Reliabilität ergab für beide Formen einen Wert von .96 (N=122, Zeitintervall 6 Monate). Die Testhalbierungsmethode mithilfe der *odd-even-Methode* lieferte ähnlich hohe Werte.

Die Konsistenzanalysen zeigen ein Cronbachs  $\alpha$  von .95 auf, wobei für die einzelnen Formen ein Cronbachs  $\alpha$  zwischen .60 und .84 berichtet wird. Zur Bestimmung der internen Validität wurden Korrelationen zwischen den Facettenrohwerten und den individuellen Interessensgebieten durchgeführt. Der Korrelationskoeffizienten liegen zwischen .19 und .54. Hinweise für die Konstruktvalidität ergeben sich aus der diskriminanten und konvergenten Validität. Die Ermittlung der diskriminanten Validität wurde mithilfe des Zahlen-Verbindungstests (Oswald & Roth, 1987) erhoben. Die Zusammenhänge zwischen der *kognitiven Leistungsgeschwindigkeit*, als Konzept der fluiden Intelligenz, und dem BOWIT sind erwartungsgemäß schwach. Hinsichtlich der konvergenten Validität wird der Wissenstest des I-S-T 2000 R (Liepmann et al., 2007) zur Beurteilung mit einbezogen. Die Korrelation ergab einen Wert von .75 (Form A) und .70 (Form B). Zur Überprüfung der Kriteriumsvalidität werden verschiedene Außenkriterien angeführt: Abiturnote, Studiumsnote, beruflicher Erfolg, privater Erfolg, Entgelt, hierarchische Position. Die Korrelationen weisen jedoch nur eine schwache bis mittlere Stärke auf. Des Weiteren liegen für den BOWIT verschiedene Normen vor. Eine differenzierte Auswertung nach Geschlecht, Alter, Berufstätigkeit (Student, Fachkraft, Führungskraft, Vertrieb, Forschung, Personal/Lehre/Weiterbildung) ist daher möglich.

Im Hinblick auf die Testdurchführung, wird der BOWIT als Powertest ohne Zeitbeschränkung dargeboten. Das heißt, es werden Multiple-Choice-Aufgaben vorgegeben, die eine richtige und eine falsche Lösung beinhalten, wobei eine Antwortkategorie „keine der zuvor genannten Antworten trifft zu“ lautet.

Das Ziel des Verfahrens ist, laut den Autoren, die Erfassung des semantisch-deklarativen Wissens. Mithin soll der Allgemeinwissensstand einer Person zuverlässig und ökonomisch abgebildet werden. Die Zielgruppe schließt Studierende, Hochschulabsolventen als auch akademisch gebildete Führungs- und Fachkräfte ein. Aufgrund der Generalisierbarkeit wissensdiagnostischer Verfahren umfassen die praxisorientierten Anwendungsfelder neben der Wirtschaftspsychologie (Personalauswahl-, Personalentwicklungs- und Personalplatzierungsentscheidungen) auch die pädagogische Psychologie sowie die klinische Psychologie (Hossiep & Schulte, 2008).

Andererseits kann ein Wissenstest bspw. im Rahmen des I-S-T- 2000 R als Subtest zur Bestimmung der kristallisierten Intelligenz herangezogen werden (Amthauer et al., 2001, Liepmann et al., 2007). Außerdem werden Wissenstests auch in sogenannte Testbatterien integriert (siehe Kapitel 3.3).

Die in Kapitel 5.1 vorgestellten Wissensformen, das deklarative Wissen und das prozedurale Wissen, werden in diesem Zusammenhang häufig als Grundvoraussetzungen für die erfolgreiche Bearbeitung von Testaufgaben gesehen und sollen mittels der Leistungstests abgeprüft werden. Die Leistung bestimmt sich danach aus dem deklarativen Wissen, dem prozeduralem Wissen und der Motivation (Schüppel, 1996).

Abschließend werden die entscheidenden Beiträge der Leistungstest für die Eignungsdiagnostik zusammengefasst (Fisseni, 2004):

- Aufgrund der Standardisierung und der damit einhergehenden Kontrolle über die Durchführung und Anwendung der Tests können subjektive Beurteilungsfehler eingeschränkt werden.
- Die Erfüllung der Gütekriterien erlaubt dem Anwender zuverlässigere Aussagen über das zu messende Persönlichkeitsmerkmal machen zu können.
- Die Erfassung von Verhaltensausschnitten innerhalb der Tests bietet dem Anwender die Möglichkeit, Rückschlüsse auf das Verhalten außerhalb der Testsituation zu ziehen.
- Aufgrund der Normierung können Vergleiche zwischen Individuen und Gruppen gemacht werden.
- Die Operationalisierung der Persönlichkeitsmerkmale (bspw. Intelligenz, Wissen) kann einen entscheidenden Hinweis auf die Bedeutung der Konstrukte liefern. Leistungstests können danach auch operationale Definitionen darstellen.
- Leistungstests befähigen den Diagnostiker dazu, ein Persönlichkeitsmerkmal in einen theoretischen Kontext einzubetten.
- Im Rahmen der Eignungsdiagnostik können die Leistungstests bspw. dazu dienen, den Leistungsstand, das Entwicklungspotential, den Trainingsbedarf und den Trainingserfolg zu ermitteln.

### **3.2.4 Gütekriterien der formalen Diagnostik**

Wie in verschiedenen Arbeitsbereichen obligatorisch, unterliegt auch die personalpolitische Eignungsdiagnostik einer umfassenden Qualitätskontrolle. Diese ist von hoher Bedeutung, weil sie die personalpsychologische Tätigkeit hinsichtlich der Entscheidungsträger im Unternehmen legitimiert.

Daher müssen die Anwendung von psychometrischen Tests und die nachfolgenden diagnostischen Interpretationen mehreren Gütekriterien genügen.

Auf der Item- und Skalenebene soll geprüft werden, ob die Test-Items und die aggregierten Skalen der Test-Absicht entsprechen. Im Rahmen der Itemanalyse werden u.a. die Gütekriterien (a) Schwierigkeitsindex, (b) Trennschärfe und (c) Homogenität herangezogen (Fisseni, 2004, S. 32f). Die Itemkennwerte kennzeichnen den jeweiligen Test hinsichtlich ihrer Einzelaufgaben, wohingegen die Gesamtstruktur eines Tests anhand der Testgütekriterien (Haupt- und Nebengütekriterien) verifiziert wird. Die folgenden drei Grundvoraussetzungen oder Perspektiven bestimmen sozusagen, „wie angemessen das empirische Relativ im numerischen Relativ [abgebildet wird]“ (Fisseni, 2004, S.46; Lienert & Raatz, 1998):

Hauptgütekriterien:

1. Validität
2. Reliabilität
3. Objektivität und Standardisierung

Nebengütekriterien:

1. Normierung
2. Ökonomie
3. Nützlichkeit
4. Vergleichbarkeit

Eine umfassende Übersicht geben beispielsweise Amelang und Schmidt-Atzert (2006), Bühner (2006), Bundschuh (1999), Häcker, Leutner und Amelang (1988), Lienert und Raatz (1998), Fischer (1974), Fisseni (2004), Wottawa (1980).

Es muss jedoch darauf hingewiesen werden, dass der Katalog einschlägiger Gütekriterien nicht verbindlich definiert ist. Bei einer Beurteilung diagnostischer Methoden und Tests muss daher im Einzelfall geprüft und berücksichtigt werden, welche Bedingungen, Umstände und Zielsetzungen gegeben sind. Aus diesem Grunde kann es auch zu unterschiedlichen Gewichtungen der Kriterien kommen (Amelang & Schmidt-Atzert, 2006). Einen umfangreichen Überblick über die Anforderungen und Bewertungsmaßstäbe wurde im Testkuratorium der Föderation deutscher Psychologenverbände 1986), in den „Standards für pädagogisches und psychologisches Testen“ (Häcker, Leutner & Amelang, 1988) oder auf globaler Ebene in den „Standards of Psychological and Educational Testing“ (1995, American Psychological Association) festgelegt.



Bezogen auf die diagnostische Eignungsfeststellung mittels Leistungstests wurde im Weiteren die DIN Norm 33430 (Berufsverband Deutscher Psychologen, 2002) entwickelt, um einen einheitlichen Rahmen für die Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen zu schaffen.

### ***Anforderungen der DIN 33430 an die berufsbezogene Eignungsdiagnostik***

Mit der Konzeption der DIN –Norm 33430 (Berufsverband Deutscher Psychologen, 2002) wird die Sicherung von Qualität und Qualitätsentwicklung für die berufliche Eignungsdiagnostik oder Eignungsbeurteilung intendiert. Die Eignungsbeurteilungen sowie die damit zusammenhängenden Qualifikationsanforderungen beziehen sich hierbei auf die Bereiche der Berufswahl, Bewerberselektion, Personalplatzierung und Personalentwicklung. Vor allem in der Wirtschaft und Verwaltung kommt dem Qualitätsmanagement und der Qualitätsentwicklung eine große Bedeutung zu und wird von jenen gefordert. Denn insbesondere Dienstleistungen benötigen einen hohen Standard an Kompetenz und Sorgfalt. Für die Entwicklung der DIN-Norm waren in diesem Kontext mehrere Aspekte und Vorüberlegungen maßgebend (Hornke & Winterfeld, 2004, S. 16; Berufsverband Deutscher Psychologen, 2002, S. 3):

- Die DIN- Norm soll Anbietern von Dienstleistungen (organisationsintern, organisationsextern) als Leitfaden für die Planung, Erstellung und Durchführung von Eignungsbeurteilungen dienen.
- Sie soll als Maßstab zur Bewertung externer Angebote herangezogen werden. Die Betroffenen sollen demnach vor zu einseitigen und wenig fachgerechten Verfahren und Instrumenten bewahrt werden.
- Sie beabsichtigt, den Anwender oder Kandidaten vor unsachgemäßer Anwendung von Verfahren zu schützen und zielt im Umkehrschluss auf den sachgerechten Einsatz und die fachgerechte Entwicklung von Verfahren ab.
- Die DIN-Norm soll zur Qualitätssicherung und Qualitätsoptimierung beitragen und dem Anwender (psychologischen Diagnostiker, Personalberater, Personalabteilung, etc.) eine Entscheidungshilfe bei der Berufswahl, Bewerberselektion, Personalplatzierung oder Personalentwicklung geben.
- Eine berufliche Eignungsbeurteilung sollte so gestaltet werden, dass ihr Nutzen größer ist als die Kosten.

- Mit der DIN-Norm sollen der Stand der Wissenschaft und Technik zusammengefasst werden und fachlich fundierte Informationen verbreitet werden.

Vor diesem Hintergrund inkludiert die DIN-Norm 33430 wesentliche Leitsätze und Festlegungen, die sich auf drei grundsätzliche Segmente der diagnostischen Praxis beziehen (a.a.O.):

1. Planung von berufsbezogenen Eignungsbeurteilungen.
2. Auswahl, Zusammenstellung, Durchführung, Auswertung von Verfahren und Interpretation der Verfahrensergebnisse.  
(Die angegebenen Qualitätskriterien orientieren sich an der oben dargelegten klassischen Testtheorie und den Gütekriterien).
3. Anforderungen an die Qualifikation der an der Eignungsbeurteilung beteiligten Personen.

Die DIN-Norm richtet sich infolgedessen an (Ackerschott, 2004, S. 6):

- Unternehmer, die ihre Maßstäbe und Anwendungspraktiken verbessern und optimieren wollen.
- Führungskräfte, die gute Mitarbeiter bekommen, behalten und weiterentwickeln wollen.
- Personalverantwortliche, die in ihrer Organisation die besten und nutzenbringenden Vorgehensweisen, Instrumente und Prozesse etablieren, die effizientesten Berater einsetzen und die gut qualifizierte Mitarbeiter zur größtmöglichen Leistungsfähigkeit entwickeln möchten.
- Berater, Betriebsräte, die ihren Personalvertretungsauftrag sachkundig, kompetent erfüllen sowie Personalentscheidungen unter Berücksichtigung ökonomischer, individueller und sozialer Aspekte treffen möchten.

Die DIN-Norm bildet somit ein fachliches, wissenschaftliches Fundament für die berufsbezogene Eignungsdiagnostik, die dazu beitragen soll, die Kompetenzen von (potentiellen) Mitarbeitern realistisch darzustellen (Heyse & Kersting, 2004). Schließlich gehört es zu den Kernaufgaben jedes Unternehmens gute Mitarbeiter auszuwählen, verborgene Fähigkeiten aufzudecken sowie das Potential des Personals sichtbar zu machen, um eine wirtschaftliche und persönliche Verbesserung aller Betroffenen zu erreichen.

Das übergeordnete Ziel ist demnach die Erstellung eines Leitmodells, mit deren Hilfe die Prognose über die berufliche Bewährung einer Person bzw. die Passung eines

Bewerbers oder Mitarbeiters für eine zukunftsorientierte Stelle gewährleistet werden kann. Die DIN-Norm soll letztendlich einen Beitrag leisten, die Qualitätsstandards im Unternehmen, genauer gesagt, in der Personalarbeit sicherzustellen, um ziel-, kunden- und marktgerecht zu handeln (Hornke & Winterfeld, 2004). Die mit der DIN-Norm beabsichtigte Qualitätssicherung dient also dazu die Transparenz und die Optimierung von Beurteilungsverfahren voran zu treiben. Der Vorteil eines qualitätsgesicherten Eignungsurteils ist nicht nur für die Bewerber bzw. Mitarbeiter ersichtlich und von hoher Relevanz, sondern spielt auch für den wirtschaftlichen Erfolg der Organisation eine herausragende Rolle. Denn die Fähigkeit den richtigen Bewerber auszuwählen, ist unter dem Gesichtspunkt der Kosten-Nutzen-Rechnung eine Investition in die Wettbewerbsfähigkeit der Organisation (Kersting, 2004).

### **3.3 Methodische Konzepte zur Diagnostik von Schlüsselkompetenzen**

Die Ermittlung von relevanten, berufsbezogenen Schlüsselqualifikationen kann aus den oben dargestellten theorieorientierten und empirischen Ansätzen, den Klassifikationsmodellen und den damit zusammenhängenden Kategorisierungsvorschlägen abgeleitet werden. Insbesondere die praxisorientierten Konzepte von Stangel-Meseke (1994, in Anlehnung an Bunk et al., 1981, 1982, 1991) und Eilles-Mathiessen et al. (2002) liefern Hinweise und ein solides Grundgerüst zur Identifikation der geforderten Schlüsselqualifikationen. Darauf aufbauend werden durch die aktuellen Forschungsbeiträge aus dem Bereich der Kompetenzmessung weitere Anhaltspunkte für die Erfassung von Schlüsselkompetenzen gegeben (vgl. Erpenbeck & Rosenstil, 2003, 2007). Eine Übersicht über den derzeitigen Stand der wissenschaftlichen Forschung dokumentieren Erpenbeck und Rosenstil (2003, 2007) und Frey, Jäger und Renold (2005).

Obwohl die Forschung auf diesem Gebiet erhebliche Fortschritte gemacht hat, liegt bis heute noch kein wissenschaftlich fundiertes Standardwerk vor, das eine konkrete, anwendungsorientierte Zusammenstellung (Empfehlungen oder Anweisungen) über Testinstrumente zur psychologischen Diagnostik von Schlüsselkompetenzen enthält. Das liegt unter anderem darin begründet, dass der Schlüsselqualifikationsbegriff oder Kompetenzbegriff von Wissenschaftlern und Praktikern unpräzise definiert bzw. uneinheitlich verwendet wird, (Gnahs, 2007; Erpenbeck & Rosenstil, 2003, 2007; Lang v. Wins, 2003) und sich die diagnostische Erfassung von Schlüsselkompetenzen auf zum Teil

„wenig abgesicherte und empirische Instrumente und Ergebnisse“ stützt (Frey & Balzer, 2005, S. 31). Es herrscht jedoch Einigkeit darüber, dass kognitive Fähigkeiten (Intelligenz) und Wissen in nahezu allen Ansätzen als entscheidende Schlüsselkompetenzen hervorgehoben werden und als übergeordnete Dimensionen interpretiert werden können. Lang von Wins (2003) macht zugleich deutlich, dass die Konstrukte „als Voraussetzungen für die Herausbildung von [vielen] spezifischen Kompetenzen [aufgefasst werden können]“ (S. 589). Dementsprechend werden psychometrische Testverfahren, wie Intelligenz- und Wissenstests zur Erfassung dieser Merkmale bevorzugt eingesetzt und deren Bedeutung für die berufsbezogene Eignungsdiagnostik betont (Schuler & Höft, 2006; Schuler, 2000). Neben der Erfassung einzelner Komponenten erlangen jedoch vor allem multimodale Verfahren bzw. systematische Kombinationen verschiedener diagnostischer Methoden an Bedeutung (Hossiep & Paschen, 1998; Schuler, 2000; Sarges & Wottawa, 2001). Diesbezüglich stellte Hossiep (1995, 1996) in seinen Forschungen heraus, dass gerade eignungsdiagnostische Testbatterien, die unterschiedliche Einzelverfahren aus dem Bereich der Persönlichkeit, der Intelligenz und weiteren Leistungsbereichen inkludieren, ein realistisches Profil einer Person geben und den beruflichen und privaten Erfolg am besten prognostizieren können. Die Kombination der Einzeltests muss dabei so gewählt werden, dass es den Anforderungen der zu besetzenden Stelle bzw. dem jeweiligen praktischen Bedürfnis gerecht wird. Prinzipiell lassen sich zwei Arten von Testbatterien unterscheiden, wobei dem zweiten Verfahren hinsichtlich der zu entwickelnden START- Testbatterie für Berufseinsteiger (Lipmann, in Vorbereitung) eine höhere Bedeutung zukommt:

1. Homogene Testbatterien

(Diese Form ist dadurch gekennzeichnet, dass sie einen einzigen Aspekt der Persönlichkeit erfassen sollen. Die Einzeltests sollten dabei untereinander sowie mit dem Validitätskriterium hoch korrelieren.)

2. Heterogene Testbatterien

(Das Charakteristikum dieser Methode besteht darin, dass jeder Einzeltest ein anderes Merkmal der Persönlichkeit erfasst. Die Einzeltests korrelieren demnach untereinander niedrig, aber mit dem Kriterium hoch. Der Anwendungsbereich liegt vor allem in der beruflichen Eignungsvorhersage.)

Der Vorteil von Testbatterien liegt darin begründet, ein Profil zu erstellen, welches differenzierte Aussagen über die Persönlichkeitsmerkmale (hier z.B. Intelligenz- und Wissenstrukturen) eines Individuums erlaubt. Die Profildarstellung mittels der differentiellen Eignungstests ermöglicht außerdem eine Vergleichbarkeit der Merkmalsbereiche, sofern für die Einzelresultate entsprechende Norm- und Standardwerte sowie Prozentränge angegeben werden. Im Rahmen der psychologischen Diagnostik gelten Testbatterien aufgrund ihrer vielseitigen Anwendungsmöglichkeiten außerdem als adäquates Instrument, um dem neuzeitlichen, wirtschaftlichen und gesellschaftlichen Wandel und der damit zusammenhängenden Dynamik der beruflichen Anforderungen gerecht zu werden. Darüber hinaus konstatieren Lienert und Raatz (1998), dass Testbatterien gegenüber Elementartests eine wesentlich höhere Validität aufweisen. Wohingegen Einzeltests nur in Ausnahmefällen einen Validitätskoeffizienten von 0.6 übersteigen, können Testbatterien Korrelationen mit einem Außenkriterium bis zu 0.8 und mehr erreichen. Zusammenfassend stellen die Autoren folgendes fest: „Wo komplexe Persönlichkeitsmerkmale wie Berufseignung untersucht werden sollen, wird eine Testbatterie im allgemeinen mehr und besseres leisten als ein Elementartest“ (S. 338).

Im Weiteren können verschiedene (psychologische) Verfahren zur Diagnostik von Schlüsselkompetenzen herangezogen werden, die in Anlehnung an Schuler und Höft (2006) dem biografischen, Simulations- und Eigenschaftsansatz zugeordnet werden können. Daher werden nachfolgend ausgewählte Verfahren der beruflichen Eignungsbeurteilung überblicksartig dargestellt (eine umfassendere Auswahl mit detaillierten Informationen geben z.B. Eilles-Mathiessen et al., 2002; Lang v. Wins, 2003; Schuler, 2000, 2006a; Brähler, Holling, Leutner & Petermann, 2002). Die Eignung der jeweiligen Verfahren variiert jedoch je nach dem zu überprüfenden Anforderungstyp.

Bezogen auf den biografischen Ansatz (Schuler & Höft, 2006), wird naheliegenderweise der biografische Fragebogen oder die Auswertung von Bewerbungsunterlagen (Anschreiben, Lebenslauf, Zeugnisse, Spezialkenntnisse, etc.) angeführt. Sie liefern wichtige Hinweise über erforderliche Qualifikationen sowie einen Indikator für zukünftiges Verhalten und weitergehende Merkmale, wie z.B. Führungsverhalten, Motivation, Flexibilität. Die Analyse kann auch in Form eines Interviews stattfinden. Es wird angemerkt, dass insbesondere das Auswahlgespräch mit zu den verbreitetsten Verfahren gehört (Schuler, Frier & Kaufmann, 1993). Ein Prototyp der simulationsorientierten Prognosekonzepte ist die Arbeitsprobe, die erfolgsrelevantes berufliches Verhalten am Arbeitsplatz aufdecken soll. Alternative situative Verfahren umfassen beispielsweise

Präsentationsaufgaben, Gruppendiskussionen, computergestützte Simulationsverfahren (z.B. Schuler, 2000; Kreuzig & Enzenberger, 2000) oder Breitbandverfahren (z.B. Hogrefe TestSystem, 2001; Etzel & Küppers, 2000). Die Eigenheit dieser Verfahren besteht darin, dass sie verhaltensnahe und praxisorientierte Anhaltspunkte über die Qualifikation der Bewerber geben. Im Rahmen des Eigenschaftsansatzes gelten psychologische Tests als typisches Erhebungsinstrument. Anhand von Tests sollen primär Persönlichkeitsmerkmale bzw. Eigenschaften erfasst werden, die einen relativ stabilen Charakter haben. Hierzu zählen hauptsächlich Persönlichkeits-, Interessen- und Leistungstests (Schuler & Höft, 2006). Eine Möglichkeit zur Integration der verschiedenen Verfahren bietet das Assessment Center. Es bezeichnet eine multiple Verfahrenstechnik, bei der unterschiedliche eignungsdiagnostische Instrumente und leistungsrelevante Aufgaben kombiniert werden, um den zukünftigen Berufserfolg möglichst gut vorherzusagen zu können. Gleichzeitig wird die Methode auch als Beurteilungs- und Förderinstrument angewendet (vgl. Schuler, 2000; Fisseni & Fennekels, 1995).

Resümierend lässt sich festhalten, dass alle Verfahren einen wertvollen Beitrag zur praxisbezogenen Diagnostik von Schlüsselkompetenzen leisten. Im Besonderen können jedoch die psychologischen Tests hervorgehoben werden, da sie hinsichtlich der DIN-Norm 33430 überaus ökonomische, effiziente, objektive, vergleichsweise preiswerte und wissenschaftlich fundierte Verfahren darstellen.

### **3.4 Die Bedeutung von Leistungstests im Human Resource Management**

Die Bedeutung der psychometrischen Intelligenztests und Wissenstests in der Eignungsdiagnostik ist bis heute ungebrochen. In Nordamerika kann man bspw. davon ausgehen, dass nahezu jeder Jugendliche mindestens einmal mit einem Intelligenztest konfrontiert worden ist. Auch in der Bundesrepublik Deutschland wurden in der Zeit um 1970 bereits ca. 1 Million Intelligenztests durchgeführt (Conrad, 1983). Die Anwendung von Intelligenz- und Wissenstests in der Eignungsdiagnostik ist seitdem stetig gewachsen. Im Hinblick auf die umfassende Literatur zeigt sich, dass die Eignungsdiagnostik und damit die Prüfung von geistigen Leistungsfähigkeiten in Form von Testverfahren auf eine lange Historie zurückblicken kann (siehe Kapitel 4.1).

Demzufolge zählen die Intelligenztests auch zu den am weitesten fortentwickelten und zuverlässigsten Instrumenten der Eignungsdiagnostik (Brambring, 1983; Schuler, 2000). Im Zentrum der psychologischen Diagnostik stehen daher die psychometrischen Testverfahren zur Messung der Intelligenz, die häufig einen Wissenstest zur Erfassung der kristallisierten Intelligenz enthalten.

Im Vergleich zu Nordamerika fällt die Einsatzhäufigkeit von psychologischen Testverfahren im Rahmen von personalpolitischen Arbeitsfeldern in der Bundesrepublik Deutschland noch sehr gering aus.<sup>1</sup> Hinsichtlich der Personalauswahl zeichnet sich jedoch ein Aufwärtstrend ab. Das liegt nach Schuler und Höft (2006) vor allem darin begründet, „dass es praktisch keinen Beruf gibt, für den allgemeine Fähigkeitstests keine kriteriumsbezogene Validität für berufliche Leistungsvariablen aufweisen (...) und diese Verfahrensgruppe häufig der beste Prädiktor für Berufserfolg ist“ (S. 111). In diesem Zusammenhang zeigen auch die Ergebnisse einer Befragung der Industrie- und Handelskammer (2007), dass die Diagnose des Allgemeinwissens und der Intelligenz eine bedeutende Rolle für die Unternehmen spielt. Ihrer Ansicht nach stellen die beiden Faktoren Grundvoraussetzungen für die Bewältigung neuartiger beruflicher Anforderungen dar, wobei insbesondere das Wissen als Indikator für zukünftiges Lernverhalten gesehen wird.

Demzufolge soll die Bedeutung der Intelligenz und des Wissens für die Eignungsdiagnostik im Weiteren hervorgehoben werden. Hierfür wird zunächst die Einsatzhäufigkeit von Testverfahren im Rahmen der Personalauswahl, als Anwendungsgebiet der Eignungsdiagnostik, dokumentiert. Danach werden ausgewählte Befunde zur prädiktiven Validität von eignungsdiagnostischen Instrumenten vorgestellt. Daraus ergeben sich auch die Fragen, welche grundsätzlichen Vorteile die Intelligenz- und Wissenstests haben und wie zuverlässig die Tests den Berufserfolg vorhersagen können. Ein Vorteil der Tests liegt im Allgemeinen darin begründet, dass sie theoretisch und wissenschaftlich fundiert sind. Zudem ermöglichen die Tests, „sofern sie den Gütekriterien der psychologischen Diagnostik entsprechen, objektive, reliable, valide und zudem zeitökonomische Messungen“ (Holling, Preckel & Vock, 2004, S.57).

---

<sup>1</sup> Die Verwendungshäufigkeit von psychologischen Testverfahren in amerikanischen Großbetrieben (in den 60er Jahren) wird von Schuler und Höft (2006) mit einem Wert von >80% angegeben.

### 3.4.1 Einsatzhäufigkeiten von Leistungstests

Eine Darstellung der Verwendungshäufigkeit berufsbezogener Testverfahren spielt sowohl für die Personalpsychologie als Wissenschaft, Testanbieter, anwendungsorientierte Unternehmen und Psychologen eine besondere Rolle und ist von erhöhter Relevanz und Nutzen. Im Folgenden wird primär auf die Einsatzhäufigkeit von Leistungstests bzw. Tests zur Erfassung kognitiver Fähigkeiten innerhalb der externen Personalauswahl eingegangen. Eine Übersicht über die Befunde zur internen Personalauswahl geben Hell, Schuler, Boramir und Schaar (2006).

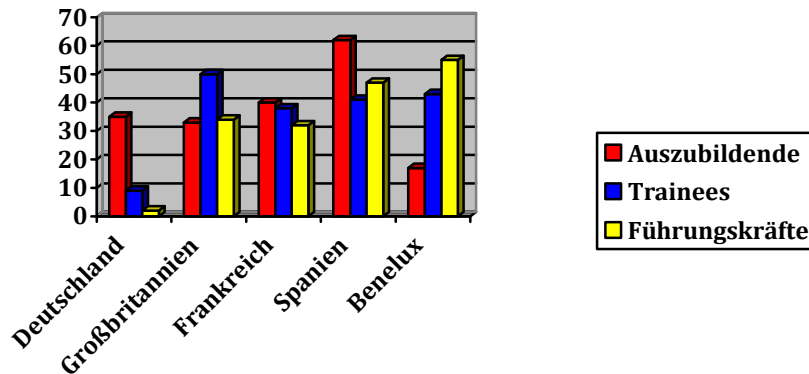
In einer Studie von Schuler, Frier und Kaufmann (1993) wird aufgezeigt, wie häufig die Leistungstests bzw. Intelligenztests bei der beruflichen Personalauswahl in Deutschland, Frankreich, Großbritannien, Spanien und in den Beneluxländern eingesetzt werden. Die Ergebnisse der Stichprobe dokumentieren, dass 47% der deutschen Unternehmen Leistungstests und 37% Intelligenztests verwenden. Eine genauere Darstellung bezüglich der Einsatzhäufigkeit der Leistungs- bzw. Intelligenztests zur externen Personalauswahl zeigt, dass Auszubildende im technischen Bereich in 39% der Fälle mit Leistungstests konfrontiert werden und zu 28% mit Intelligenztests. Im kaufmännischen Bereich liegen die Prozentsätze bei 40% bzw. 35%. In den restlichen Berufsgruppen (ungelernte Arbeiter, Facharbeiter, Angestellte ohne Führungsaufgaben, Trainees und Führungskräfte) liegt der Anteil der im Auswahlprozess eingesetzten Leistungstests oder Intelligenztests höchstens bei 11%. Ähnliche Ergebnisse zeigt auch Kanning (2004) in seinen Forschungsbeiträgen und stellt zudem heraus, dass die Sichtung von Bewerbungsunterlagen, Interviews und Assessment Centern von deutschen Unternehmen nach wie vor als bevorzugtes Kriterium zur Bewerberselektion herangezogen werden. Im Hinblick auf die Verwendung von psychologischen Testverfahren durch Organisations- und Arbeitspsychologen berichtet Schorr (1991) in seinen Studien, dass 46,8% der 661 befragten Psychologen in ihrer diagnostischen Praxis Intelligenztests und 14,8% allgemeine Leistungstests gebrauchen.

In Relation zu den oben genannten europäischen Staaten lässt sich jedoch im Weiteren feststellen, dass die Einsatzhäufigkeit von Leistungs- und Intelligenztests in Deutschland im Unterschied zu den anderen Ländern, in denen ca. 17-62% der Unternehmen Leistungstests in der Personalauswahl einsetzen, vergleichsweise gering ausfällt (Schuler et al., 1993).



Darüber hinaus lässt sich feststellen, dass psychologische Leistungstests hauptsächlich zur Selektion von Auszubildenden eingesetzt werden und weniger auf der Ebene der Führungskräfte (siehe Tabelle 3).

*Tabelle 3: Einsatzhäufigkeit von Leistungs- und Intelligenztests im europäischen Vergleich (vgl. Schuler, Frier & Kaufmann, 1993)*



Hinsichtlich der Gegenüberstellung zur Verwendungshäufigkeit von psychologischen Testverfahren im europäischen Raum weisen die Studien von Dany und Torchy (1994) in die gleiche Richtung. Auch hier zeigt sich, dass die Bundesrepublik Deutschland im Vergleich zu elf anderen Staaten das Schlusslicht bildet. Die Autoren dokumentieren, dass gerade einmal 6% der deutschen Unternehmen psychometrische Tests in Betracht ziehen, wohingegen Spanien und Finnland beispielsweise über 60% psychologische Testverfahren verwenden. Schackleton und Newell (1997) untersuchten die Personalauswahlpraktiken an 250 verschiedenen Unternehmen in Frankreich, Italien, Großbritannien, Belgien und Deutschland. Sie kamen ebenfalls zu demselben Ergebnis, dass deutsche Personalleiter kognitive Leistungstests nur „selten“ bis „manchmal“ (vorgegebene Antwortkategorien) einsetzen. 12 % der britischen und 30% der belgischen Unternehmen nutzen solche Testverfahren „immer“ und über 70% „manchmal“. In einer weiteren umfassenden internationalen Studie von Ryan, McFarland, Baron und Page (1999) konnten die Resultate repliziert werden.

Im Bezug auf den eingangs dargestellten Forschungsbeitrag von Schuler et al. (1993) und den Studien von Schulz, Schuler und Stehle (1985), ziehen Schuler, Hell, Trappmann, Schaar und Boramir (2007) in ihrer aktuellen Untersuchung, zur „Nutzung psychologischer Verfahren der externen Personalauswahl in deutschen Unternehmen“, einen Vergleich über 20 Jahre.

Für die Erhebungen wurden 125 deutsche Unternehmen unterschiedlicher Branche und Größe einbezogen. Die wichtigsten Ergebnisse lauten (Schuler et al., 2007, S. 61ff):

- Das Einstellungsgespräch bzw. das Interview ist branchenübergreifend das meist verwendete Verfahren zur Bewerberselektion und wird von jedem der befragten Unternehmen eingesetzt. Dahinter folgt die Auswertung der Bewerbungsunterlagen und das Schlusslicht bilden die Online Persönlichkeitstests sowie der biografische Online Fragebogen.
- Im Rahmen der Online- Verfahren dominieren die Online- Bewerbungsunterlagen, dicht gefolgt von dem Online-Personalfragebogen.
- Eine deutliche Erhöhung der Einsatzhäufigkeit zeigt das Assessment Center und das strukturierte Interview. Die anderen Verfahren weisen keine gravierenden Veränderungen zu den älteren Studien auf.
- Die vermutete Verbreitung psychologischer Testverfahren (vgl. Sackmann & Elbe, 2000; Göhs und Dick, 2001) bei ca. 30% der Unternehmen kann bestätigt werden, wobei Leistungstests verglichen mit Persönlichkeitstests häufiger angewendet werden.

Eine detailliertere Aufschlüsselung der Einsatzhäufigkeiten von Leistungs- und Intelligenztests kann der folgenden Tabelle entnommen werden:

*Tabelle 4: Einsatzhäufigkeit von Leistungs- und Intelligenztests  
In Deutschland  
(verkürzt nach Schuler et al., 2007)*

	Ausbildung		Trainees	Führungskräfte		
	technisch	kaufmännisch		untere	mittlere	obere
<b>Leistungstests</b>	33,0 (-6,0) [-28,7]	36,6 (-3,4) [-18,5]	4,2 (-6,8) [-14,6]	1,6 (-2,4) [-4,7]	1,7 (-1,3) [-2,8]	0,0 (-2,0) [-5,9]
<b>Intelligenztests</b>	23,4 (-4,6) [-26,6]	31,3 (-3,7) [-26,0]	4,2 (-3,8) [-16,6]	0,8 (-3,2) [-4,0]	1,7 (-0,3) [-4,3]	0,8 (+0,8) [-5,1]

Anhand der Tabelle ist es ersichtlich, dass die Verwendungshäufigkeiten von psychologischen Testverfahren im Vergleich zu der Vorgängerstudie von Schuler et al. (1993) nahezu unverändert geblieben ist, obgleich im Verlauf der 20 Jahre ein leichter Rückgang zu verzeichnen ist. Eine mögliche Erklärung dieser Entwicklung kann darin gesehen werden, dass dem Anwender eine fast unüberschaubare Bandbreite an Tests vorliegt. Die Studie von Jung (2000) bestätigt die Vermutung und dokumentiert, dass es 73% von 300 Großbetrieben schwierig finden, einen zweckmäßigen Test herauszufiltern. Außerdem wird oft die wissenschaftliche Seriosität von Testverfahren in Frage gestellt und die Intransparenz der dahinter stehenden Methoden und Theorien bemängelt. Damit zusammenhängend erscheint es manchen Praktikern überdies nicht nachvollziehbar, inwieweit einige Testverfahren einen Bezug zum beruflichen Erfolg besitzen. Gleichwohl wird die Praktikabilität und die Validität von den Personalverantwortlichen jedoch auf einem mittleren bis hohen Niveau eingestuft (Schuler et al., 2007). Ferner wird auch die ungenügende Akzeptanz der Verfahren seitens der Bewerber als Grund für die dürftige Einsatzhäufigkeit angeführt. Ein gangbarer Weg, dieser Entwicklung entgegenzutreten und neue Perspektiven zu eröffnen, stellt die vom Bund deutscher Psychologen (BdP) aufgestellte DIN-Norm 33430 (2002) über die Mindestanforderungen an Tests dar. Vor dem Hintergrund der neuen Herausforderungen im Zuge der Globalisierung, respektive steigender wirtschaftlicher Dynamik und Flexibilität, werden insbesondere die Aspekte der Objektivität, Effektivität, Vergleichbarkeit und Ökonomie von psychologischen Testverfahren in ihrer Bedeutung und Relevanz hervorgehoben. Darüber hinaus kann davon ausgegangen werden, dass durch die verstärkte internationale Ausrichtung der Organisationen und der damit einhergehenden Synchronisierung der Selektionsprozesse im Rahmen von Personalentscheidungen ein höherer Einsatz von psychometrischen Testverfahren zu erwarten ist.

Obwohl noch erheblicher Forschungsbedarf zur Verfeinerung der Messinstrumente besteht, wird die Bedeutung von Testverfahren meines Erachtens in den nächsten Jahren weiterhin ansteigen.

### **3.4.2 Internationale Forschungsbeiträge und Befunde zur Relevanz von Intelligenz- und Wissenstests**

Die Bedeutung psychologischer Testverfahren kann aus verschiedenen Blickwinkeln beurteilt werden. Zum Einen ergibt sich die Relevanz von Intelligenz- und Wissenstests aus den oben dargestellten Beiträgen der Schlüsselqualifikationsforschung (siehe Kapitel 2). Im Zuge des gesellschaftlichen und wirtschaftlichen Wandels erlangen extrafunktionale Qualifikationen wie Intelligenz und Wissen einen herausragenden Stellenwert, da sie dazu dienen die neuartigen Anforderungen des kontinuierlichen Wandels in vielfältiger Weise zu meistern (Liepmann, 1993; Bürgel & Zeller, 1998; Eilles-Matthiessen, el Hage, Janssen & Osterholz, 2002). Zum Anderen werden die Bedeutsamkeit und der Nutzen von psychologischen Testverfahren durch die prädiktive Validität bestimmt. Die prädiktive Validität gibt dabei an, wie zuverlässig die Ergebnisse den schulischen, akademischen oder beruflichen Erfolg vorhersagen können.

#### ***Befunde zur prädiktiven Validität von Intelligenztests***

Im Vorfeld können verschiedene allgemeine Vorteile von Intelligenztests angeführt werden (Schmidt & Hunter, 1998). Erstens besteht der Einsatz von kognitiven Leistungstests durch den breiten Anwendungsbereich (siehe unten). Zweitens überzeugen die Tests im Vergleich zu personalintensiven Arbeitsproben, Assessment Centern oder Interviews als kosteneffiziente und ökonomische Gruppenverfahren. Drittens wird der „wissenschaftliche Nachweis der Validität von Intelligenztests zur Vorhersage von Berufsleistungen besser geführt als bei irgendeiner anderen Methode“ (Schmidt & Hunter, 1981, zitiert nach Schmidt & Hunter, 1998, S. 264). Diese Tatsache lässt sich auch mit der langen historischen Entwicklungsgeschichte der Intelligenzforschung begründen. Folglich zeigt sich in einer Gegenüberstellung mit anderen Persönlichkeitsmerkmalen, dass die Intelligenz mit zu den besten Prädiktoren für Schul-, Ausbildungs- und Berufserfolg zählt (Schuler & Höft, 2006; Schmidt & Hunter, 1998). Viertens eignen sich Intelligenztests, im Sinne der Investmenttheorie (Cattell, 1987), als gute Prognoseinstrumente für berufsbezogenes Lernen und die Aneignung von beruflichem Wissen (Felser, 2005). Fünftens liegen die Vorzüge „kognitiver Fähigkeitstests bei der Einfachheit der Anwendung, der Generalisierbarkeit der Ergebnisse und der Parsimonität der Erklärung“ (Murphy, 1996, zitiert nach Schuler & Höft, 2006, S. 112).

Unter der Annahme, dass Intelligenz auch als Fähigkeit zum Erwerb von Wissen oder als „Fähigkeit zu hoher Bildung“ (Asendoorpf, 2004, S. 191) verstanden werden kann, lassen sich die nachfolgenden Befunde besser nachvollziehen.

Schon im vergangenen Jahrhundert wurde eine Vielzahl von Forschungsbeiträgen zur prädiktiven Validität unterschiedlicher Methoden publiziert. Die grundlegende Problematik bei der Bewertung der verschiedenen Studien zur prognostischen Befähigung von Testverfahren lag anfangs in erheblichen Validitätsschwankungen. Die Abweichungen wurden bis zum Anfang der 1980er Jahre durch die *Theorie der situationalen Spezifität* begründet, die sich auf den Einfluss der unterschiedlichen Berufsgruppen und den situativen Kontext bezog (Schmidt & Hunter, 1998; Schmidt & Hunter, 2000). Diese Auffassung wurde indessen revidiert und es stellte sich heraus, dass die Differenzen eher auf statistische Artefakte, wie ungenügende Stichprobengrößen oder durch Messfehler zurückgeführt werden. Um den unzureichenden Ergebnissen entgegenzuwirken, wurden darauf aufbauend innovative metaanalytische Methoden verwendet. Die Metaanalyse ist gekennzeichnet durch „eine systematische, nach transparenten Kriterien vollzogene Zusammenführung (=Sekundarstudie) von unterschiedlichen empirischen Einzelstudien zur gleichen Fragestellung (=Primärstudie)“ (Höft, 2006, S.769). Innerhalb der Metaanalysen hat die Methode der Validitätsgeneralisierung an Bedeutung gewonnen. Die Vorzüge der Metaanalysen liegen insbesondere in den genaueren Validitätsschätzungen und der Möglichkeit zum Vergleich der Validitäten unterschiedlicher eignungsdiagnostischer Testinstrumente für verschiedene Berufe (Schmidt & Hunter, 1998). Dementsprechend spielen die metaanalytischen Befunde im Folgenden eine herausragende Rolle.<sup>2</sup>

### ***Zum Zusammenhang zwischen Intelligenz und Schulerfolg***

Hinsichtlich der Vorhersage des Schulerfolges belegen eine große Anzahl von Studien, dass Intelligenztests eine hohe prognostische Validität besitzen (z.B. Jensen, 1980; Süß, 2001). In den Forschungsbeiträgen von Wechsler (1958) zur Prognose der Art des Schulabschlusses wurde eine Korrelation von  $r = .70$  ermittelt.

---

<sup>2</sup> Sofern in den Forschungsbeiträgen korrigierte Validitätskoeffizienten angegeben wurden, werden diese im Folgenden verwendet. Korrigierte bzw. „wahre“ Validitäten bieten die Möglichkeit verschiedene Fehlerquellen, wie Reliabilitätsmängel oder Stichprobenfehler auszuschließen und haben den Vorteil einen besseren Vergleich zwischen unterschiedlichen Berufsgruppen herzustellen.

In weiteren zusammenfassenden Studien werden Koeffizienten zwischen  $r = .34$  und  $r = .51$  berichtet (Steinkamp & Maehr, 1983; Hatie & Hansford, 1982, zitiert nach Süß, 2001). Eine differenziertere Herangehensweise zur Analyse des Zusammenhangs zwischen Intelligenz und Schulerfolg ermöglicht der Intelligenz-Struktur-Test 2000 R (Liepmann et al., 2007). Aufgrund der Unterteilung der allgemeinen Intelligenz in verbale, numerische und figurale Intelligenzfacetten sowie in schlussfolgerndes Denken können spezifischere Korrelationen angegeben werden, die einen hohen diagnostischen Wert in sich bergen. In diesem Kontext werden beispielsweise Korrelationen zwischen der numerischen Facette und der Mathematiknote von  $r = .40$ , der figuralen Facette und der Mathematiknote von  $r = .36$  oder zwischen dem schlussfolgerndem Denken und der Mathematiknote in Höhe von  $r = .45$  berichtet. Es zeigt sich, dass die stärksten Zusammenhänge zwischen dem numerischen Faktor und den mathematisch-naturwissenschaftlichen Fächern bestehen. Erwartungsgemäß niedrigere Zusammenhänge bestehen zwischen der numerischen bzw. figuralen Skala und der Deutsch- bzw. Englischnote (a.a.O.). In die gleiche Richtung weisen die Untersuchungen von Süß (2001), die sich auf das Berliner Intelligenzstrukturmodell (Jäger, 1982, 1984) beziehen. So konnte unter anderem belegt werden, dass die verbale Intelligenz naheliegender Weise ein sehr guter Prädiktor für sprachlich orientierte Fächer ist (bis zu  $r = .47$ ). Generell stellt der Autor heraus, dass die Teilkonstrukte auf der zweiten Ordnungsebene eine enge Beziehung zu den schulischen Leistungen aufweisen und dass multiple Prognosen die Vorhersagen optimieren. Jensen (1980) konstatiert im Weiteren, dass die prädiktive Validität jedoch mit steigendem Ausbildungslevel (Schule-Universität) sinkt. Die Ursache hierfür liegt womöglich in der unzureichenden Güte der Hochschulnoten, der geringeren Leistungsbreite oder den unterschiedlichen Einflüssen auf die Persönlichkeit (vgl. Giessen, Gold, Hummer & Jansen, 1986).

### ***Zum Zusammenhang zwischen Intelligenz und Ausbildungserfolg***

Im Hinblick auf den postulierten Zusammenhang zwischen Intelligenz und Ausbildungserfolg lässt sich ebenfalls feststellen, dass die Intelligenz eine hohe prädiktive Validität besitzt. Sofern die Kriterien des Ausbildungserfolges durch das Bildungsniveau oder den erzielten Bildungsabschluss substituiert werden, bekräftigen zum Beispiel die Forschungsbeiträge von Jensen (1980, 1998) die hohen Korrelationen mit der allgemeinen Intelligenz ( $r = .60$  –  $r = .70$ ).

Die Ergebnisse konnten auch anhand von umfassenden Metaanalysen repliziert werden. So geben Walberg, Strykowski, Rovai und Hung (1984) in ihrer Metaanalyse zum Zusammenhang zwischen Intelligenz und akademischen Erfolg (Aggregation von ca. 3000 Studien) einen Korrelationskoeffizienten von  $r = .71$  an, der auch in neueren Forschungsbeiträgen bestätigt wird (z.B. Gagné & St Pèrè, 2001). Unter Bezugnahme auf weitere Ausbildungserfolgskriterien, wie z.B. Zwischenergebnisse oder Leistungsbeurteilungen durch Lehrkräfte werden beispielsweise in einer Untersuchung von Ones, Viswesvaran und Dilchert (2004) Zusammenhänge in Höhe von  $r = .40$  dokumentiert. Darüber hinaus können auch Korrelationen zum Weiterbildungs- oder Trainingserfolg nachgewiesen werden. Hierbei wurden Leistungseinschätzungen oder gemessene Wissenszuwächse mit spezifischen oder allgemeinen kognitiven Fähigkeiten korreliert, die Koeffizienten zwischen  $r = .40$  und  $r = .70$  zeigten (Hunter & Hunter, 1984; Daumenlang, 1990, 1995; Ones et al., 2004). In diesem Rahmen ermittelten auch Schmidt und Hunter (1998) in ihren Metaanalysen einen Korrelationskoeffizienten von  $r = .56$  für arbeitsrelevantes Lernen, das als Ausbildungserfolg interpretiert werden kann. In Relation zu verschiedenen anderen eignungsdiagnostischen Instrumenten (z.B. Arbeitsproben, strukturierte Interviews, etc.) ist der berechnete Wert verhältnismäßig hoch. Hinsichtlich des deutschen Bildungssystems, welches eine duale Verknüpfung von Theorie und Praxis vorsieht, stellt sich zudem die Frage, inwieweit die obigen Resultate übertragen werden können. In den Untersuchungen von Schmidt-Atzert und Deter (1993a, 1993b) wurden demnach theoretische und praktische Abschlussnoten als Außenkriterium zur prädiktiven Validität herangezogen. Außerdem wurden betriebliche Beurteilungen des Arbeitsergebnisses und der Arbeitsweise berücksichtigt. Die Autoren fanden heraus, dass die theoretischen Abschlussklausuren generell besser prognostiziert werden konnten als die praktischen Prüfungen. Unter Verwendung des Intelligenz-Struktur-Tests 70 berichten Schmidt-Atzert und Deter (1993a) von einem Koeffizienten von  $.54$  mit der theoretischen Abschlussprüfung. Für die praktische Abschlussnote wird ein Wert von  $.27$  angegeben und für die betrieblichen Beurteilungen ein Koeffizient von  $.43$  bzw.  $.44$ . Vergleichsweise niedrigere Werte erbrachte beispielsweise das Testverfahren d2. Die metaanalytischen Befunde zur prädiktiven Validität von Intelligenztests konnte somit für den theoretisch ausgelegten Teil der dualen Ausbildungsvariante bekräftigt werden. Auf der Basis ihrer Forschungen wird resümierend der Schluss gezogen, dass Leistungstests gute Testinstrumente zur Vorhersage des Ausbildungserfolges darstellen.

### ***Zum Zusammenhang zwischen Intelligenz und Berufserfolg***

Bei der Sichtung verschiedener Studien fällt auf, dass die prädiktive Validität von Testverfahren bezüglich der Prognose des Ausbildungserfolges häufig höher ist, als die zur Vorhersage des Berufserfolges. Außerdem differieren die Validitätskoeffizienten gleichermaßen zwischen verschiedenen Berufsgruppen (z.B. Brambring, 1983; Schmidt-Atzert & Deter, 1993a, 1993b). Man spricht in diesem Zusammenhang auch von differentieller Validität.

Im Bezug auf die Forschungsbeiträge von Hunter (1980) stellen Hunter und Hunter (1984) heraus, dass die kognitiven Fähigkeiten gemessen mit dem GVN als Untertest des GATB bzw. *General Aptitude Test Batterie*) höher mit dem Ausbildungserfolg korrelieren ( $r=.57$ ) als mit dem Berufserfolg ( $r=.51$ ). Des Weiteren zeigen die Autoren in ihren Untersuchungen, dass allgemeine Intelligenztests durchschnittliche Korrelationskoeffizienten von  $r=.54$  mit dem Außenkriterium des Ausbildungserfolges sowie korrigierte Korrelationen von  $r=.45$  mit dem Berufserfolg aufweisen. In einer neueren europäischen Metaanalyse von Salgado, Moscoso, Fruyt, Anderson, Bertua und Rolland (2003) werden ähnliche Ergebnisse festgestellt. Darüber hinaus kristallisiert sich auch unter Berücksichtigung verschiedener Komplexitätsniveaus der einbezogenen Berufe heraus, dass sich die Verwendung von allgemeinen mentalen Fähigkeitstests (General Mental Ability- GMA) vor allem für die Vorhersage des Ausbildungserfolges (korrigierte Validität von  $r=.40$  bis  $r=.74$ ) sowie des Berufserfolges (korrigierte Validität von  $r=.45$  bis  $r=.67$ ) eignen. Ein Zwischenfazit der Autoren macht zudem deutlich, dass sich die prädiktive Validität mit zunehmender Komplexität des Berufes erhöht. In einer weitergehenden Untersuchung von Bertua, Anderson und Salgado (2005) wurden neben allgemeinen kognitiven Tests spezifische Leistungstests herangezogen, die eine Differenzierung in verbale, numerische, räumliche und wahrnehmungsbezogene Fähigkeiten ermöglichen. Die grundlegenden Resultate demonstrieren, dass der Berufserfolg am Besten mit einem allgemeinen kognitiven Test ( $r=.48$ , korrigiert) oder einem spezifischen Wahrnehmungstest ( $r=.50$ , korrigiert) vorhergesagt werden kann. Wohingegen der Ausbildungserfolg am deutlichsten durch die numerischen Fähigkeiten prognostiziert werden kann ( $r=.54$ , korrigiert), gefolgt von den wahrnehmungsbezogenen Fähigkeiten und den allgemeinen kognitiven Fähigkeiten ( $r=.50$ , korrigiert). Im Hinblick auf die zum Teil marginal differierenden Befunde zur Vorhersage des Ausbildungs- oder Berufserfolges legen die Forschungsbeiträge von Hühsheger, Maier und Stumpp (2007)



indessen nahe, dass allgemeine Intelligenztests und insbesondere spezifische kognitive Fähigkeitstests aufgrund ihrer hervorragenden Prognosemöglichkeiten geeignet sind, beide Bereiche im Kontext der Eignungsdiagnostik abzudecken.

Wie bereits verdeutlicht, weisen Intelligenztests auch gute bis sehr gute prädiktive Validitäten zur Vorhersage des beruflichen Erfolgs auf (z.B. Schmidt & Hunter, 1998; Schmidt & Hunter, 2000). Bereits 1980 ermittelte Jensen hohe Korrelationskoeffizienten ( $r = .50$  bis  $r = .70$ ) zwischen Intelligenztestwerten und dem Berufsniveau. Eine metaanalytische Zusammenfassung der Befunde, die eine Vielzahl unterschiedlicher Stichproben, Erfolgskriterien und Testverfahren berücksichtigen, geben Schmidt und Hunter (1998). In der folgenden Tabelle werden die Resultate der prädiktiven Validität ausgewählter eignungsdiagnostischer Testinstrumente illustriert.

*Tabelle 5: Metaanalytische Zusammenfassung der Befundlage zur Validität eignungsdiagnostischer Verfahren  
(verkürzt nach Schmidt & Hunter, 1998, S. 265)*

Prädiktor	Val (r)	inkrV	mR	%Zuw.
Allgemeine Intelligenztests	.51			
Arbeitsproben	.54	.12	.63	24%
Integritätstests	.41	.14	.65	27%
Gewissenhaftigkeitstests	.31	.09	.60	18%
Einstellungstest (strukturiert)	.51	.12	.63	24%
Einstellungstest (unstrukturiert)	.38	.04	.55	8%
Fachkenntnistests	.48	.07	.58	14%
Probezeit	.44	.07	.58	14%
Biografische Daten	.35	.01	.52	2%
Assessment Center	.37	.02	.53	4%
Interessen	.10	.01	.52	2%
Graphologie	.02	.00	.51	0%

*Anmerkung:* Val(r): korrigierte kriteriumsbezogene Validität (Kriterium: Leistungsbeurteilung); inkV: inkrementelle Validität (d.h. Validitätszuwachs durch Berücksichtigung eines zweiten Prädiktors); mR: multiple Korrelation von allgemeinen Intelligenztests als erster und mit jeweiligen Prädiktor als zweiter Variabel; %Zuw.: prozentualer Zuwachs der Validität durch Hinzunahme des zweiten Prädiktors.

Aus der Tabelle lässt sich entnehmen, dass allgemeine Intelligenztests neben Arbeitsproben und strukturierten Einstellungstests die höchsten prädiktiven Validitäten aufweisen. Obgleich die Testverfahren nahezu gleiche Kennwerte zeigen, erscheint es aus Gründen der Effektivität und unter ökonomischen Gesichtspunkten jedoch sinnvoll zu sein, primär allgemeine Intelligenztests für die eignungsdiagnostische Testpraxis zu empfehlen. Der Vorteil von standardisierten Intelligenztests kann insbesondere darin gesehen werden, dass sie bei Personengruppen sowie bei Berufseinsteigern, die kein berufstypisches Vorwissen mitbringen oder als Auswahlinstrument bei Berufen mit

breit gefächerten Anforderungen eingesetzt werden können. Außerdem stellen sie schnelle, objektive und kosteneffiziente Lösungen für die Eignungsdiagnostik dar. Obwohl die aufgeführten Untersuchungen und Befunde vorwiegend aus dem nordamerikanischen Raum stammen, können die Ergebnisse auch auf Europa bezogen werden. Demzufolge berichten Salgado et al. (2003) in ihren Untersuchungen von übereinstimmenden artefaktkorrigierten, prädiktiven Validitäten, die sich auf elf unterschiedliche Berufsgruppen beziehen. Zusätzlich stellen die Autoren heraus, dass eine ansteigende Beschäftigungskomplexität mit einer höheren Validität einhergeht. Hinsichtlich dessen zeigen die Autoren auf, dass die korrigierten Korrelationen zwischen .24 und .64 variieren. Für den deutschen Raum bestätigen zum Beispiel die Studien von Schmidt-Atzert, Deter und Jaeckel (2004) oder Hühsheger, Maier und Stump, 2007 die dargestellten Resultate. Sofern unter Berufserfolgskriterien beispielsweise das Einkommen, Vorgesetztenbeurteilungen, Produktivitätsdaten oder Berufsniveau subsumiert werden, geben Ones et al. (2004) für allgemeine und spezifische Intelligenztests eine durchschnittliche Validität zwischen  $r=.40$  und  $r=.60$  an, die gemeinhin als gut interpretiert werden kann. Im Hinblick auf den plausiblen Zusammenhang zwischen Intelligenz und erreichten Berufsniveau konnte Löhlin (2000) in seinen Studien nachweisen, dass der durchschnittliche Intelligenzquotient zwischen einfachen Arbeitern (93 Punkte,  $SD=14$ ) und hochqualifizierten Facharbeitern und Führungskräften (111 Punkte,  $SD=13$ ) erheblich voneinander abweicht. Zahlreiche andere Forschungsergebnisse weisen in die gleiche Richtung (z.B. Amelang & Bartussek, 1990, 1997, 2001; Bertua, Anderson & Salgado, 2005). Die Schlussfolgerung, dass sich die Mindestanforderungen an Intelligenz mit ansteigendem Berufsniveau grundsätzlich erhöhen, stößt in der wissenschaftlichen Auseinandersetzung jedoch auf erhebliche Kritik. Amelang und Bartussek (1990, 1997, 2001) stützen sich in ihrer Argumentation auf die vielfältigen, unübersichtlichen Umwelteinflüsse, die den postulierten Determinismus zwischen Ursache und Wirkung ausschließen. Aufgrund des mehrfach nachgewiesenen Zusammenhangs zwischen Intelligenz und Bildungsniveau bzw. Bildungsniveau und Berufsniveau ergibt sich vielmehr ein mittelbarer Einfluss der Intelligenz.

### ***Zum Zusammenhang zwischen Wissen und berufsbezogenen Erfolgskriterien***

Die Wissenstests als eigenständige Instrumente besitzen im Bezug auf die Eignungsdiagnostik eine ähnlich hohe prädiktive Validität wie allgemeine Intelligenztests. Darauf bezugnehmend ermöglichen einerseits (berufsbezogene) Fachwissenstests eine gute Prognose über den späteren Berufserfolg (Dye, Reck, & Mc Daniel, 1993). Dye et al. (1993) errechneten in diesem Kontext eine durchschnittliche Validität von  $r = .45$ . Schmidt und Hunter (1998) geben auf der Grundlage einer Metaanalyse einen Validitätskoeffizienten von  $.48$  an und bestätigen somit das Ergebnis. Hinsichtlich der Verwendung von Fachwissenstests wird jedoch der berechtigte Einwand erhoben, dass sie sich weniger für die Gruppe der Berufseinsteiger eignen, da diese nur über ein eingeschränktes Fachwissen und nicht über spezifische Berufserfahrungen verfügen (Keriting, 1999a; Schmidt & Hunter, 2000). Für diese Zielgruppe bieten sich andererseits eher Allgemeinwissenstests an. Sofern man eine fundierte Allgemeinbildung als Indikator für zukünftiges Lernen versteht, ergibt sich demnach ein breiterer Anwendungsbereich und die Diagnostik erfüllt im Rahmen von Personalauswahlverfahren, bei Personalentwicklungsentscheidungen oder Berufsberatungen seinen strategischen Zweck. Fürntratt (1969) macht zudem deutlich, dass eignungsdiagnostische Verfahren wie Allgemeinwissenstests insbesondere bei Berufen eingesetzt werden können, die hohe Anforderungen an „Allgemeinwissen und/oder breite, unspezialisierte Lernfähigkeit und geistige Regeamkeit“ (S.5) stellen. Hierzu zählen beispielsweise Berufe im Bereich Management, Beratung, Planung und Erziehung.

Im Hinblick auf den konkreten Zusammenhang zwischen Allgemeinwissen und Ausbildungserfolg belegt die Studie von Fürntratt (1969), dass Allgemeinwissenstests gute Prädiktoren darstellen. Unter Verwendung des DWT (Differentieller Wissenstests) ließen sich signifikante Zusammenhänge mit den gewählten Außenkriterien (hier: Durchschnittsnote der Berufsschule, Seminarleistungen, Ergebnis der Abschlussprüfung, Vorgesetztenurteile) nachweisen. In die Stichprobe wurden unter anderem die Berufssparten der Rechtspfleger und Regierungsinspektoren mit einbezogen. Aufgrund ähnlicher methodischer und konzeptioneller Voraussetzungen weist der Forschungsbeitrag von Althoff (1971) in die gleiche Richtung.

Anhand des PSB (Prüfsystem zur Schul- und Bildungsberatung von Horn (1969, zitiert nach Jungkunz & Bodinet, 1989) untersuchten Jungkunz und Bodinet (1989) ebenfalls den Zusammenhang zwischen Allgemeinwissen und Berufsabschlussnoten (Fertigkeits- und Kenntnisprüfungen). Die wichtigsten Ergebnisse waren, dass die PSB-Skalen 1 und 2, welche die Allgemeinbildung abbilden sollen, über alle berücksichtigten Berufsgruppen hoch mit den Berufsabschlussnoten korrelierten. Bezogen auf die Gesamtstichprobe zeigte sich, dass die theoretische Kenntnisprüfung mit den Allgemeinwissensskalen besonders gut prognostiziert werden konnte ( $r = -.60$ ). Innerhalb der einzelnen Berufsgruppen (KFZ-Mechaniker, Tischler, Friseur) variierte der Korrelationskoeffizient zwischen  $r = -.54$  und  $r = -.65$ . Die Zusammenhänge mit den Fertigkeitprüfungen fielen dagegen vergleichsweise gering aus ( $r = -.29$  bis  $r = -.37$ ). Eine aktuellere Studie von Kersting (1999b) bestätigt die bereits herausgestellten Befunde. Im Rahmen einer Untersuchung mit 135 Bewerbern für die Steuerverwaltung wurde der Zusammenhang zwischen dem Allgemeinwissen (gemessen mit dem Wissenstest BIS-r-DGP) und den Zwischenprüfungsergebnissen analysiert. Für den Gesamtwert des Wissenstests ergab sich eine signifikante Korrelation mit dem Außenkriterium in Höhe von  $r = .32$  (korrigiert). Für die einzelnen Themenbereiche des Wissenstests (Wirtschaft, Politik, Literatur, Gemeinschaftskunde) konnte nur für die Komponente „Wirtschaft“ ein signifikanter Zusammenhang aufgezeigt werden. Hossiep und Schulte (2008) stellten im Weiteren heraus, dass die individuellen Wissensausprägungen (Testinstrument: BOWIT) eng mit den Schul-, Universitäts- und Berufsleistungen verknüpft sind. Als Kriterium für die Schulleistungen wurde die Abiturnote herangezogen, für die universitäre Ausbildung die Abschlussnote und für den beruflichen Erfolg wurden die Indikatoren „Entgelt“ und „hierarchische Position“ verwendet. Hinsichtlich des Zusammenhangs zwischen BOWIT-Rohwerten und den durchschnittlichen Noten der Abiturfächer berichten die Autoren von einem Koeffizienten von  $r = -.32$  (Form A, Stichprobe der Berufstätigen). Für die Gruppe der Studenten wird ein Korrelationskoeffizient von  $r = -.40$  angegeben. Überdies ermittelten Hossiep und Schulte (2007) einen weiteren signifikanten Zusammenhang zwischen dem BOWIT und der Abschlussnote des Studiums. Im Bezug auf den Berufserfolg wurden sowohl für das „Entgelt“ als auch für die „hierarchische Position“ signifikante Korrelationen errechnet. Die Ergebnisse legen den Schluss nahe, dass die Höhe des Entgeltes und die Bekleidung von Führungspositionen oftmals mit einem „Wissensvorsprung“ (S. 41) einhergehen. Laut den Autoren gilt die Kriteriumsvalidität des Wissenstests somit als gesichert.

Obwohl dem Wissen eine hohe Wichtigkeit beigemessen wird, existieren im Vergleich zur Intelligenzforschung jedoch nur relativ wenige umfassende Befunde bezüglich des Zusammenhangs zwischen Allgemeinwissen und Ausbildungs- bzw. Berufserfolgs. Damit einhergehend sollen im Folgenden einige Aspekte verdeutlicht und ergänzt werden, die die Relevanz des Wissens für die personalpolitische Eignungsdiagnostik unterstreichen (siehe auch Kapitel 5).

Waldmann (1997) oder Kluwe (2000) betonen in ihren Forschungsbeiträgen, dass ein fundiertes und solides (Allgemein-) Wissen eine bedeutende Voraussetzung für den Neuerwerb und die Verarbeitung von Informationen darstellt, da das aufgenommene Wissen mit bereits vorhandenen Wissensstrukturen vernetzt werden kann. Das Allgemeinwissen dient demnach als Basis und Orientierungswissen zur Bewältigung einer großen Informationsmenge. Diese Ansicht wird auch von 500 Experten des Wissens- und Bildungsdelphie des Bundesministeriums für Bildung und Forschung (1998) vertreten (vgl. auch Wolf, 1999). Das Wissen hat folglich einen elementaren Einfluss auf die effiziente und effektive Lernleistung einer Person. Der Vorteil des Vorwissens macht sich demnach auch bei der Lösung komplexer Probleme bemerkbar (siehe Kapitel 5). In diese Richtung weisen beispielsweise die Untersuchungen von Süß (1998). Zusätzlich stellen Putz-Osterloh und Lemme (1987) heraus, dass neben dem Fachwissen vor allem das Allgemeinwissen, aufgrund des breiteren Anwendungsbereiches, einen positiven Effekt bei der Lösung verschiedener Probleme erzeugt. Außerdem geht Reetz (1990) davon aus, dass das Allgemeinwissen die soziale Kommunikation begünstigt und speziell in Dienstleistungsberufen oder generell in Berufen, die verstärkt durch zwischenmenschliche Interaktionen geprägt sind (z.B. im Bereich Vertrieb), einen entscheidenden Vorteil generiert. Vor dem Hintergrund, dass Organisationen kontinuierliches Lernen seiner Mitarbeiter voraussetzen und insbesondere Fähigkeiten verlangen, die den Erwerb von neuem Wissen erleichtern, gewinnt das Allgemeinwissen an Relevanz und entfaltet hier seine Bedeutung.

Resümierend zieht Kersting (1999a) den Schluss, dass „diese Überlegungen und Daten [dafür sprechen], Wissen nicht nur indirekt (durch die Messung der Intelligenz), sondern verstärkt auch direkt bei der Prognose beruflicher Leistungen zu berücksichtigen“ (S. 256).

### ***Zum Zusammenhang zwischen Wissen ( kristallisierter Intelligenz) und Berufserfolg***

Angesichts der guten prädiktiven Validitäten der Wissenstests und der Intelligenztests zur Vorhersage der beruflichen Leistungen erscheint eine Kombination dieser beiden Tests aber von großem, substantiellem Nutzen für die Eignungsdiagnostik zu sein, um differenzierte Aussagen über den Leistungsstand und das Leistungspotential einer Person zu geben. Entsprechend der vorangegangenen Ausführungen kann und wird der Wissenstest auch als Subtest zur Erfassung der kristallisierten Intelligenz herangezogen werden. Sofern das (Allgemein-) Wissen, im Sinne der Investmenttheorie von Cattell (1971a; 1971b; 1987) als Teil der allgemeinen Intelligenz begriffen wird, erscheint demnach der Einsatz von Allgemeinwissenstest in Kombination mit Intelligenztests auf personalpolitischen Arbeitsfeldern in hohem Maß zweckmäßig und wertvoll zu sein. Nach Kersting (2000) stellt die kristallisierte Intelligenz darüber hinaus einen Indikator dar, „wie gut es Personen gelingt, ihre Intelligenz erfolgreich zum Erwerb von Kenntnissen [und zur Aneignung von Fachwissen] zu nutzen“ (S.99).

In diesem Zusammenhang ermittelten Beauducel und Kersting (2002) einen Korrelationskoeffizienten in Höhe von  $r = .40$  zwischen der kristallisierten Intelligenz und den Zwischenprüfungsergebnissen von Bewerbern für die öffentliche Verwaltung.

Die fluide Intelligenz zeigte dagegen nur einen sehr geringen Zusammenhang mit dem Kriterium auf und wurde als unzureichender Prädiktor eingestuft. Der Forschungsbeitrag unterstützt auch die Schlussfolgerungen von Helmke und Weinert (1997), dass das Wissen zur Vorhersage des Schulerfolges besonders gut geeignet ist. Auch in den Untersuchungen von Goff und Ackerman (1992) wurden die Befunde bekräftigt. Die Autoren berechneten einen Korrelationskoeffizienten von  $r = .38$  zwischen kristallisierter Intelligenz und der Durchschnittsnote des Studiums. Die fluide Intelligenz erreichte nur einen Wert von  $r = .13$ . Demgegenüber stehen die Ergebnisse, welche an einer High School erhoben wurde. Hier zeigte sich ein stärkerer Zusammenhang zwischen der fluiden Intelligenz und dem Notendurchschnitt ( $r = .40$ ). Die Autoren argumentieren, dass die Bedeutung des Wissens wohl mit zunehmendem Ausbildungslevel ansteigt. In der Studie von Nettelstroth (2004) wurden die Korrelationen zwischen kristallisierter Intelligenz und unterschiedlichen beruflichen Erfolgskriterien analysiert. Die Stichprobe umfasste dabei 295 Probanden, die aus verschiedenen Berufsfeldern stammten.

Nettelstroth (2004) gibt in diesem Kontext eine Korrelation von  $r = .27$  zwischen der kristallisierten Intelligenz und der Einkommensentwicklung an, während die fluide Intelligenz einen Wert von  $r = .24$  erreicht. Im Weiteren dokumentiert der Autor, dass nur die kristallisierte Intelligenz klare Zusammenhänge zur Beförderungshäufigkeit und des *Commitments* erkennen ließ.

Auf der Basis der dargestellten Befunde lässt sich konstatieren, dass sowohl die Diagnose der allgemeinen Intelligenz sowie der kristallisierten Intelligenz (Wissen), aufgrund ihrer Interdependenz, einen fundierten Erkenntnisgewinn sicherstellen und somit einen wichtigen Faktor für die berufliche Eignungsdiagnostik darstellt.

Eine zusammenfassende Illustration der genannten Befunde gibt Kersting (1999a) mit dem Modell zur Beziehung zwischen Intelligenz, Wissen, Problemlösen und beruflicher Leistung (siehe Abbildung 3).

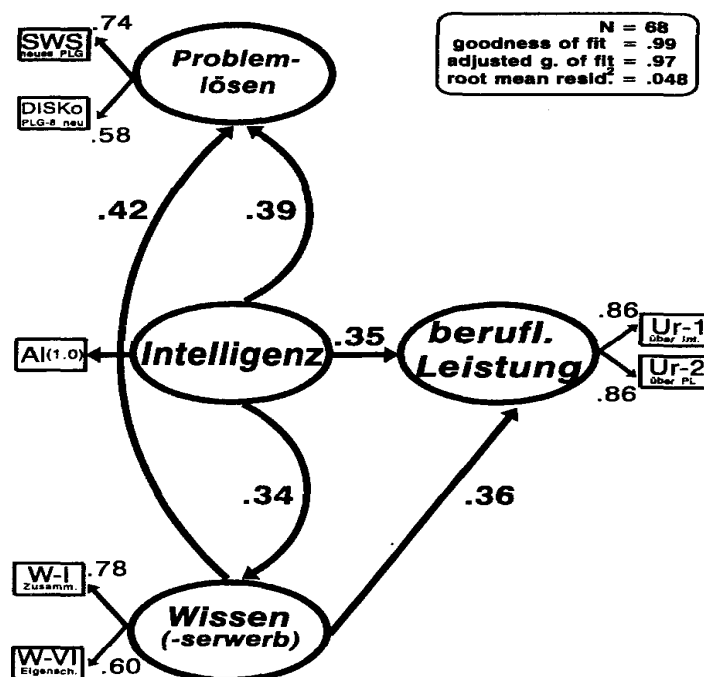


Abbildung 3: Modell zum Beziehungsgeflecht zwischen Intelligenz, Wissen, Problemlösung und beruflicher Leistung nach Kersting (1999a, S. 251)

Aus der mit LISREL erstellten Abbildung lässt sich entnehmen, dass die Einflussstärke der Intelligenz sowie des Wissens auf die berufliche Leistung nahezu identisch ist. Die Pfadkoeffizienten zeigen einen Wert von  $.35$  und  $.36$ . Gleichermäßen verhält es sich im Bezug auf das Problemlösen.

Darüber hinaus ist es ersichtlich, dass die Intelligenz neben dem direkten Einfluss auf die Berufsleistung zusätzlich einen mittelbaren Einfluss auf das Kriterium ausübt. Aufgrund der Dreiecksbeziehung bzw. durch die Aufteilung der (erklärten)Varianzen in die einbezogenen Variablen kann vermutet werden, dass die Korrelation zwischen Intelligenz und Berufsleistung letztendlich höher ist als zwischen Wissen und Kriterium. Das aufgestellte Strukturmodell bestätigt indessen den postulierten Zusammenhang und untermauert die Relevanz von Wissen und Intelligenz für die berufliche Eignungsdiagnostik. Kersting (1999a) konkludiert dementsprechend, dass „die Daten weder der Annahme [widersprechen], dass Intelligenz eine Voraussetzung zum Erwerb und zur Anwendung von Wissen ist noch der Annahme, dass Intelligenz und Wissen gemeinsam das problemlösende Handeln sowohl im Problemlöseszenario als auch im Berufsalltag bestimmen“ (S.251).

### ***Kontextabhängige Vor- und Nachteile der dargestellten Testverfahren***

Abschließend stellt sich die Frage, inwieweit eher allgemeine Intelligenztests, spezifische Intelligenztests oder differenzierte Testbatterien innerhalb der beruflichen Eignungsdiagnostik von Bedeutung sind. Generell ist die Wahl eines geeigneten Testinstruments abhängig von dem Anforderungsprofil der zu besetzenden Stelle. Sofern sich die beruflichen Anforderungen als vielseitig, heterogen und komplex darstellen, werden die Erfassung und die Vorhersage jener erschwert. Hierfür bietet sich die Verwendung eines allgemeinen Intelligenztests an, der durch seine Universalität dem Berufsbild am ehesten gerecht wird (Ackerman & Humphreys, 1991; Salgado et al., 2003). Ein sinnvolles Anwendungsfeld liegt unter anderem in der Berufsberatung, um eine groben Hinweis auf das zu erreichende Berufsniveau zu erhalten oder in den personalpolitischen Arbeitsbereichen (Schuler & Höft, 2006). Dem gegenüber stehen spezifische Tätigkeiten, die beispielsweise hohe Anforderungen an das räumliche Vorstellungsvermögen oder an verbale Fähigkeiten stellen (z.B. Beruf des Piloten). In diesem Kontext werden Intelligenztests präferiert, die eine differenzierte Aufschlüsselung verschiedener Intelligenzkomponenten beinhalten. Ein bewährtes Instrument in der eignungsdiagnostischen Praxis ist der Intelligenz-Struktur-Test 2000 R (Liepmann et al., 2007), der sowohl einzelne Facetten der Intelligenz berücksichtigt, die Berechnung eines sogenannten g-Faktors der Intelligenz zulässt sowie einen Wissenstest zur Bestimmung der kristallisierten Intelligenz enthält.



Solche Struktur-Tests werden auch von den Testpersonen mehrheitlich akzeptiert, da sie einen stärkeren Bezug zu den Ausbildungs- und Berufsanforderungen erkennen lassen (Holling, Preckel & Vock, 2004). Von Seiten der Organisation erscheint die Verwendung derartiger Tests ebenfalls von hoher Relevanz zu sein, um der zunehmenden Umweltdynamik und den damit einhergehenden Variabilität der Berufsbilder angemessen entgegenzutreten (a.a.O.). In dieselbe Richtung weisen auch die in Kapitel 3.3 angesprochenen Testbatterien, deren Testmodule entsprechend der eignungsdiagnostischen Anforderungen flexibel zusammengestellt werden können. Entsprechend den obigen Ausführungen wird angemerkt, dass vor allem Intelligenz- und Wissenstests einen integralen Bestandteil der Testbatterien bilden und demnach von großer Wichtigkeit für die Eignungsdiagnostik sind. Zudem ergibt sich die Bedeutung aus der sehr guten prognostischen Validität, hinsichtlich der Vorhersage von schulischen, akademischen und beruflichen Leistungen, so dass Intelligenz- und Wissenstests als erfolgsträchtige Prädiktoren für personalpolitische Entscheidungen genutzt werden können. Die differenzierten kognitiven Fähigkeitstests und die Testbatterien stellen unter diesen Gesichtspunkten zukunftsorientierte Verfahren dar, die als Wegweiser für die künftige praxisorientierte Eignungsdiagnostik fungieren.

### **3.4.3 Anwendungsfelder kognitiver Fähigkeitstests**

Aus den genannten Befunden und der hohen Generalität kognitiver und wissensdiagnostischer Verfahren lässt sich auch der breite Anwendungsbereich der psychologischen Testverfahren (hier Intelligenz- und Wissenstests) ableiten bzw. erklären. Die Anwendungsfelder erstrecken sich hierbei über die Wirtschaftspsychologie, die pädagogischen Psychologie und die klinische Psychologie. Bezugnehmend auf die Wirtschaftspsychologie können die Tests schon bei der Berufsberatung verwendet werden, um grundlegende Potentiale und Fähigkeiten einer Person aufzudecken. Des Weiteren treten vor allem die personalpolitischen Arbeitsfelder innerhalb von Unternehmen bzw. Organisationen in den Vordergrund der Betrachtung. Aufgrund der guten prädiktiven Validitäten eignen sich die Testverfahren insbesondere als Entscheidungshilfe bei der Personalauswahl, Personalplatzierung oder Personalentwicklung. Die Berufsbereiche, bei denen die Verfahren eingesetzt werden können, sind mannigfaltig.

Es werden jedoch speziell Tätigkeiten angesprochen, die angesichts ihrer Anforderungen einen hohen Anteil an Allgemeinbildung, Lernfähigkeit und geistiger Aktivität

erwarten (Hossiep & Schulte, 2008; Fisseni, 2004; Schuler & Höft, 2006). Vor dem Hintergrund des gesellschaftlichen Wandels hin zur Dienstleistungsgesellschaft stellen Hossiep und Schulte (2008) heraus, dass „eine fundierte Allgemeinbildung insbesondere bei Positionen mit einem hohen Ausmaß an sozialer Interaktion, wie beispielsweise im Vertrieb oder in der Beratung geradezu gefordert wird (...) [und] das Allgemeinwissen eine grundlegende Basis [darstellt], die eine angemessene und somit erfolgreiche Kommunikation in Bezug auf fachliche und überfachliche Themen erst sicherstellt“ (S. 7). Im Rahmen der pädagogischen Psychologie können die Leistungstests zur Ermittlung des Schul- oder Hochschulbildungsniveaus genutzt werden, um die Leistungsfähigkeit bzw. Leistungsunterschiede zu erfassen und Bildungsdefizite aufzuzeigen. Auf der Basis der eruierten Informationen können beispielsweise Unterrichtskonzeptionen evaluiert werden, Normwerte für Subgruppen erstellt werden oder die Verfahren werden als Kriterium für die Hochschulzulassung herangezogen (Hossiep & Schulte, 2008; Fisseni, 2004). Die Testergebnisse fungieren in diesem Zusammenhang als Indikatoren für den Lernerfolg und zukünftiges Lernverhalten (Bildungsmotivation). Auf dem Gebiet der klinischen Psychologie können die Tests in Kombination mit bestehenden Tests (wie z.B. dem Altgedächtnisinventar) zur Klärung von neurodiagnostischen Fragestellungen beitragen (Hossiep & Schulte, 2008).

Aufgrund der genannten Belege lässt sich festhalten, dass die psychologischen Leistungstests am häufigsten verwendet und in nahezu allen Berufsbereichen eingesetzt werden (Schuler & Höft, 2006).

Zusammenfassend können die folgenden Vorteile und die Bedeutung von psychologischen Testverfahren genannt werden:

- Breiter Anwendungsbereich.
- Gute prognostische Validitäten zur Vorhersage des beruflichen Erfolgs.
- Qualitätskontrollen durch die Haupt- und Nebengütekriterien sowie der DIN-Norm 33430. Dazu zählen insbesondere die Validität, Reliabilität, Objektivität der Tests: Standardisierung des Inhalts, der Durchführung und der Anwendung (siehe Kapitel 3.2.4).
- Vergleichbarkeit der Testergebnisse durch Norm- und Referenzwerte
- Möglichkeit zur Durchführung von Gruppentests.
- Kostengünstige und zeitökonomische Messungen (vs. zeitintensive und kostspielige Arbeitsproben und Assessment Center).

### 3.4.4 Exkurs: Betriebswirtschaftliche Bedeutung eignungsdiagnostischer Testverfahren

#### *Bedeutung des Produktionsfaktors „Wissen“ und die wissensorientierte Unternehmensführung*

Ein betriebswirtschaftlicher Blickwinkel, um die Bedeutung von Wissen und Testverfahren zu erschließen, besteht in der neuzeitlichen Akzentuierung des Produktionsfaktors Wissen und der Hinwendung zur wissensorientierten Unternehmensführung. Der Hintergrund für die Entwicklung wird in dem Wandel zur Wissensgesellschaft gesehen, der seit den sechziger Jahren erkannt und propagiert wird (z.B. Lane, 1966). Die damit einhergehenden Veränderungen von (globalen) Märkten mit schnellerer Innovationsgeschwindigkeit, kürzeren Produktlebenszyklen, Individualisierung von Kundenbedürfnissen, neuen Informationstechnologien, Projektarbeit und Teamarbeit sowie die Entstehung von neuen Dienstleistungssektoren erfordern ein Umdenken der Unternehmen. Zur Sicherstellung der Wettbewerbsfähigkeit kommt dem Produktionsfaktor Wissen, neben den klassischen Produktivkräften eine erhebliche Wichtigkeit zu. Hier stellt sich die Frage, welches Wissen im Unternehmen von Relevanz ist (North, 2005). Im Rahmen dieser wirtschaftswissenschaftlichen Betrachtungsweise definieren Probst, Raub und Romhardt (1998) Wissen als die „Gesamtheit der Kenntnisse und Fähigkeiten, die Personen zur Lösung von Problemen einsetzen. Dies umfasst sowohl theoretische Erkenntnisse als auch praktische Alltagsregeln und Handlungsanweisungen. Wissen stützt sich auf Daten und Informationen, ist im Gegensatz zu diesen jedoch immer an Personen gebunden. Es wird von Individuen konstruiert und repräsentiert deren Erwartungen über Ursache-Wirkungs-Zusammenhänge“ (S. 44). Drei wesentliche Aspekte können an dieser Stelle im Besonderen hervorgehoben werden. Erstens ist das Wissen immer an eine Person gebunden, das nicht ohne weiteres transferierbar und jedem zugänglich ist. Zweitens dient eine allgemeine, fundierte Wissensgrundlage zur Lösung neuartiger Probleme und kann als Indikator für zukünftiges Lernverhalten gesehen werden. Drittens bezeichnet Wissen nicht nur allgemeine Kenntnisse, sondern auch spezifische Kenntnisse über unternehmensinterne und unternehmensübergreifende Prozesse (*best practices*), Technologien, Märkte, etc., die in einen bestimmten Kontext eingebettet sind.

Dieses sogenannte *Wissenskapital* gilt es zu erkennen, zu erhalten und zu entwickeln (North, 2005). Eine sinnvolle Veranschaulichung des Sachverhaltes gibt North (2005) mit seiner Wissenstreppe.

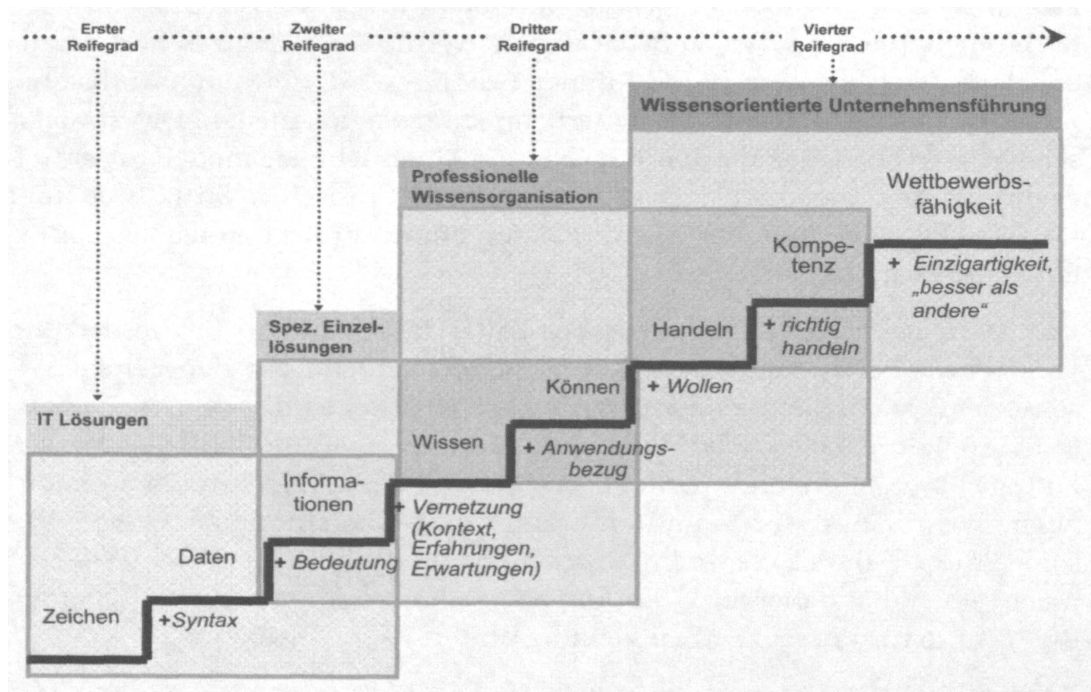


Abbildung 4: Die Wissenstreppe nach North (2005, S.32)

Ausgangspunkt der Treppe ist ein reichhaltiger Fundus an allgemeinem (Vor-) Wissen, das einer Person in Form von Zeichen (Buchstaben, Ziffern, Sonderzeichen) zur Verfügung steht. Anhand von Ordnungsregeln (Syntax) werden die Zeichen zu Daten umkodiert (z.B. 70% Eigenkapitalquote). Sobald die Daten in einem bestimmten Bezugsrahmen stehen, verdichten sie sich zu Informationen, die aus betriebswirtschaftlicher Sicht als Grundlage für Entscheidungen gebraucht werden. Durch die Vernetzung der Informationen mit bereits gespeichertem Wissen und die Einbindung in einen Bedeutungskontext wird sodann neues Wissen generiert. Die Interpretation des Informationsgehalts ist dabei abhängig von den Erfahrungen, Erwartungen und dem Vorwissen des Individuums. Das angeeignete Wissen (Wissen, dass...) entfaltet seinen Nutzen jedoch erst, wenn es zweckorientiert umgesetzt wird (Wissen, wie...) bzw. zu problembezogenen Handeln eingesetzt wird. Sofern das Wissen erfolgreich angewendet wird, kristallisieren sich Kompetenzen heraus, die die Wettbewerbsfähigkeit des Unternehmens langfristig sichern können.

Nachhaltige Wettbewerbsvorteile ergeben sich demzufolge dann, wenn das gebündelte Wissen des Unternehmens nicht oder nur begrenzt imitierbar, transferierbar, einzigartig und synergetisch mit anderen Kompetenzen verknüpft ist. Man spricht in diesem Zusammenhang von Kernkompetenzen, die „einen Verbund von Fähigkeiten und Technologien [darstellen], der auf expliziten und verborgenem Wissen beruht und sich durch zeitliche Stabilität und produktübergreifenden Einfluss kennzeichnet“ (North, 2005, S.34). Wissensorientierte Unternehmensführung bedeutet demzufolge, dass die Stufen der Wissenstreppe in der Organisation implementiert werden (zur Vertiefung: z.B. North, 2005). Daraus ergeben sich für das Wissensmanagement die nachstehenden Ziele (a.a.O, S. 3):

- Wissensaneignung: Sicherstellen, dass die Organisation und jeder ihrer Mitarbeiter lernfähig sind. Dazu gehört auch die Auswahl von kompetentem Personal (mittels geeigneter eignungsdiagnostischer Testverfahren), die sinnvolle Personalplatzierung und eine zukunftsorientierte Personalentwicklung.
- Wissensbeschaffung: Sicherstellen, dass für Geschäftsentwicklung und -prozesse benötigtes Wissen zur Verfügung steht.
- Wissensentwicklung: Sicherstellen, dass Wissen an der bestgeeigneten Stelle in oder außerhalb des Unternehmens entwickelt wird.
- Wissenstransfer: Sicherstellen, dass Wissen optimal nutzbar gemacht wird.
- Wissensweiterentwicklung: Sicherstellen, dass Wissen anwendungsbezogen aktualisiert, fortentwickelt wird.

Es ist ersichtlich, dass dem strategischem Management und im Besonderen dem Human Resource Management somit eine außerordentlich wichtige Aufgabe zukommt.

### ***Übergeordnete Bedeutung eignungsdiagnostischer Testverfahren im Rahmen des Human Resource Managements***

Sobald der personalpolitische Bereich der Eignungsdiagnostik in das Human Resource Management (HRM) einer Organisation eingebettet wird, entfaltet sich ein weiteres Bedeutungsspektrum für berufsbezogene Testverfahren.

Aus theoretischer Sicht wurden in den letzten Jahren verschiedene Ansätze entwickelt, um die Relevanz und den Stellenwert des HRM herauszustellen. Denn vor dem Hintergrund starker Umweltveränderungen und mangelnder Produktivität hat man erkannt,

dass vor allem das HRM im Zusammenspiel mit der Struktur und der Strategie einer Organisation effektiviert werden muss. Zwei klassische Konzepte sind der Michigan und der Harvard-Ansatz. Beim ersten Ansatz steht die integrative Verknüpfung von Organisationsstrategie, -struktur und HRM im besonderen Fokus. Hier wird also ein *best fit* zwischen den Komponenten angestrebt (Tichy, Fombrun & Devanna, 1982, 1984). Der Schwerpunkt des Harvard-Ansatzes beschäftigt sich mit der Nutzung und Identifizierung vorhandener Leistungspotentiale aller Organisationsmitglieder durch ein System aufeinander abgestimmter personalwirtschaftlicher Maßnahmen unter Berücksichtigung der Bedürfnisse der Beschäftigten, der Stakeholder, der gesellschaftlichen Wohlfahrt und anderer externer Umwelteinflüssen (Beer, Spector, Lawrence, Mills & Walton, 1985, S. 16f). Im Gegensatz zum Michigan-Ansatz kommen hier insbesondere verhaltenswissenschaftliche Aspekte zum Tragen und die Human Ressourcen werden noch stärker in die Entwicklung und Realisierung von strategischen Zielen mit eingebunden (Price, 2004). Letztendlich zeigt sich die Innovativität der Ansätze insoweit, als dass bislang getrennt betrachtete personalwirtschaftliche Teilfunktionen, wie Personalplanung, -beschaffung, -auswahl und -entwicklung in ihrem Zusammenhang erkannt und als integrative Problemlösungskonzepte behandelt werden. Deren Einbindung in zuvor übergeordnete Struktur- und Strategieentscheidungen werden hier vorgenommen. Die menschliche Ressource wird hier stärker aus einer „General Management-Perspektive“ und weniger aus einer speziellen Funktionsperspektive betrachtet. Das Management wird also in die Verantwortung über die menschlichen Ressourcen mit eingebunden. Angemerkt sei hier, dass das HRM die Unternehmensstrategie lediglich unterstützt. Die Organisationsstruktur und die Personalstrategie werden demnach aus der Unternehmensstrategie abgeleitet (Staehele, 1989, S. 391).

Dieser Argumentationsstruktur folgend wird im Weiteren der ressourcenbasierte Ansatz (RBV) hinzugezogen, der auf einer Stärken- und Schwächen-Analyse des Personals und des Personalmanagements beruht. Diese Sichtweise stellt die internen, vor allem die menschlichen Ressourcen einer Organisation in den Mittelpunkt der Betrachtung. Sobald die Ressourcen die folgenden Eigenschaften, wie selten, wertvoll, nicht imitierbar und nicht substituierbar aufweisen, können diese Ressourcen aufgrund von sozialer Komplexität, kausaler Ambiguität und historisch gewachsenen und daher schwer identifizierbaren Zusammenhängen, einen entscheidenden Wettbewerbsvorteil generieren. Als Quelle für wirtschaftlichen Erfolg wird also die Ressourcenstärke einer Organisation gesehen, die mit einer entsprechenden geeigneten Strategie und Struktur

Wettbewerbsvorteile sichern kann. Wright, Dunford & Snell (2001) stellen im Weiteren heraus, dass die Generierung von Wettbewerbsvorteilen allerdings nicht nur eine Funktion von isolierten und einzelnen Faktoren ist, sondern eher eine Kombination aus dem *Human Capital Pool (Stock of Skills*, die es ständig zu erhalten und zu entwickeln gilt), dem *Employee Relationships and Behaviours* (z.B. informale Strukturen) und dem *People Management Systems* (Bewertungssysteme, Kommunikationsregeln, Partizipationsmöglichkeiten, etc.).

Zusammenfassend lässt sich folgendes Charakteristikum für die Darstellung der HRM – Ansätze heranziehen: Menschen werden als Erfolgsfaktoren betrachtet, die zusammen mit den übrigen Ressourcen der Organisation so geführt, motiviert und entwickelt werden müssen, dass diese zum Erreichen von Organisationszielen beitragen (Handy, Burham, Panter & Winhard, A., 1989).

Aus praktischer Sicht wird demnach die wirtschaftliche und gesellschaftliche Bedeutung des Human Resource Managements (HRM) für die einzelnen Wirtschaftsunternehmen hervorgehoben. Hierzu gehören vor allem auch die personalpolitischen Teildisziplinen wie Personalauswahl, Personalplatzierung und Personalentwicklung. Es wird erkannt, dass die Menschen über den zukünftigen Erfolg oder Misserfolg eines Unternehmens entscheiden. Das Personal wird heute weniger als unliebsamer Kostenfaktor, sondern vielmehr als Leistungspotential des Unternehmens gesehen, das einen entscheidenden Wettbewerbsvorteil generieren kann. Dem Personalbereich kommt auf strategischer Ebene somit eine fundamentale Bedeutung zu, und zwar nicht nur in Bezug auf Neueinstellungen, sondern auch in der Ausschöpfung und Erweiterung der Potentiale des vorhandenen Mitarbeiterstammes. Dementsprechend gilt es im HRM vier grundsätzliche Herausforderungen zu meistern: Potentiale rekrutieren, Potentiale erkennen, Potentiale ausschöpfen und Potentiale erweitern. Wenn dieser Ansatz nicht berücksichtigt und hervorgehoben wird, werden Talente, Fähigkeiten und Engagement von Akteuren im betrieblichen Alltag womöglich übersehen und gehen danach für das Unternehmen verloren. Resignation und Fluktuation sind die Folge. Entweder bleiben für die Organisation daher Wertschöpfungspotentiale gänzlich ungenutzt oder es werden für anstehende Projekte erforderliche Talente und Fähigkeiten unnötig teuer eingekauft, obwohl sie im Unternehmen unerkannt vorhanden oder wenigstens leicht auszubilden sind. Einer solchen Entwicklung entgegenzusteuern ist vielleicht die größte und schwierigste Aufgabe des HRM.

In diesem Zusammenhang rückt eine geeignete Kompetenzfeststellung der Menschen in den Vordergrund der Betrachtung. Hierzu werden im Weiteren gute eignungsdiagnostische Instrumente verlangt, die grundlegende Begabungen und Talente aufdecken können. Die sogenannten Schlüsselqualifikationen sollen dazu dienen, neue Wissensgebiete rasch zu erschließen, um flexibel und selbständig auf die zukünftigen Veränderungen reagieren zu können. Dementsprechend gilt es, eignungsdiagnostische Instrumente zu entwickeln, die der Entwicklung und den Forderungen gerecht werden. Ein Schritt in diese Richtung ist die Entwicklung der umfangreichen START-Testbatterie (Liepmann et al., in Vorbereitung), die einerseits die erwähnten Kompetenzen erfassen kann, andererseits den anerkannten wissenschaftlichen Gütekriterien genügt sowie den obigen Forderungen nachkommt.

Es ist ersichtlich, dass eine gute Eignungsdiagnostik entscheidende Vorteile für den Menschen und das Unternehmen mit sich bringen kann. Aufbauend auf den oben genannten Herausforderungen gewährleisten elaborierte Testinstrumente, wie die START Testbatterie, dass die Fähigkeiten des Personals aufgedeckt werden können. Auf Basis des diagnostizierten Fähigkeitsprofils, kann das Personal entsprechend seiner Kompetenzen optimal ausgewählt, eingesetzt und weiterentwickelt werden.

Der positive Effekt kann darin gesehen werden, dass eine mögliche Unzufriedenheit des Menschen aufgrund von Fehlplatzierungen gemindert und die Motivation des Einzelnen zur Erreichung der Organisationsziele erhöht werden kann. Gleichzeitig profitiert auch das Unternehmen von dieser Entwicklung. Das Personal wird leistungsfähiger, die Fluktuation kann gesenkt werden, etc.. Dieser Argumentation folgend werden auch eindeutig ökonomische Aspekte angesprochen, die dem Unternehmen Wettbewerbsvorteile sichern können. Die Voraussetzung für den, an dieser Stelle, stark vereinfachten Sachverhalt ist natürlich, dass das gesamte HRM effektiver und effizienter gestaltet wird und in übergeordnete Struktur- und Strategieentscheidungen mit eingebunden wird. Aufgrund der Ausführungen lässt sich jedoch eindeutig festhalten, dass eine gute Eignungsdiagnostik einen wichtigen Grundstein für das Entwicklungspotential des Unternehmens darstellt und einen wesentlichen Beitrag zur gesamtwirtschaftlichen und gesellschaftspolitischen Wohlfahrt leisten kann.



## 4 Intelligenz- Theorien, Befunde und Bedeutung

Ausgehend von der Entwicklungsgeschichte der Intelligenzmessung wird in diesem Kapitel zunächst eine Begriffsbestimmung der Intelligenz avisiert. Damit einhergehend wird das Konstrukt-Verständnis näher erläutert und in einen sinnvollen Zusammenhang mit der Intelligenzmessung gebracht. Darauf aufbauend werden die Bedeutungsbereiche der Intelligenz beleuchtet, wobei die akademische Intelligenz im Vordergrund der Betrachtung steht. Danach werden die einschlägigen Strukturtheorien vorgestellt, die maßgeblich und konstituierend für das Verständnis der Intelligenz sind.

### 4.1 Intelligenzforschung im Überblick

Der Anfang der Intelligenzmessung bzw. -forschung lässt sich nicht auf einen konkreten Zeitpunkt beziehen oder an einer bestimmten Person festmachen. Die Intelligenzforschung ist das Ergebnis einer sukzessiven Entwicklung. Es lässt sich jedoch festhalten, dass verschiedene Personen, wie z.B. Darwin, Mendel, Wundt, Galton, Mc Keen Cattell, Binet und Simon und Stern entscheidende Beiträge für den Fortschritt der Intelligenzmessung geleistet haben. Im Folgenden werde ich die Intelligenzmessung in den psychologischen Kontext einordnen. Im Zentrum der Betrachtung steht hierbei die Differentielle Psychologie. Im Anschluss wird eine kurze Übersicht über die Forschungsfelder dieser Personen gegeben und es wird aufgezeigt, welchen Einfluss sie auf die Intelligenzforschung hatten bzw. noch immer haben.

#### 4.1.1 Eingliederung der Intelligenzmessung in den psychologischen Kontext

Unter Intelligenzmessung versteht man die Verwendung von Tests, die bestimmten Gütekriterien genügen. Im Rahmen der diagnostischen Technik wird die Intelligenzmessung unter dem Bereich der angewandten Psychologie subsumiert und entstammt vorwiegend der Differentiellen Psychologie. Hier liegen die meisten und am gründlichsten erforschten Befunde sowie die am weitesten fortentwickelten Instrumente zur Bestimmung individueller Differenzen (Groffmann, 1983). Ihr unmittelbares Ziel ist die „planmäßige psychologische Erforschung des Individuums“ (Groffmann, 1983, S. 2). Als Ursprung dieser Disziplin wird häufig die programmatische Schrift von Stern (1900) „Über Psychologie der individuellen Differenzen“ genannt.

Den Gegenpol der Differentiellen Psychologie bildet die Allgemeine Psychologie. Die Allgemeine Psychologie negiert vielmehr die Suche nach individuellen Differenzen und forscht stattdessen nach Eigenschaften, die verschiedenen Individuen gemein sind und daraufhin abstrahiert werden können. Mit anderen Worten untersucht die Allgemeine Psychologie vor allem die Auswirkungen experimentell variiertes (Stimulus-) Faktoren (z.B. Art und Inhalt der Instruktion, Dauer einer Beeinflussung, etc.), die für alle Individuen Gültigkeit besitzen. Die Gleichheit des Verhaltens der einzelnen Individuen wird hierbei vorausgesetzt. Die Differentielle Psychologie fokussiert sowohl die Unterschiede zwischen Personen zu einem speziellen Zeitpunkt (Querschnittsbetrachtung interindividueller Unterschiede), als auch die Differenzen innerhalb einer Person, die im Kontext von konkreten Situationen oder Zeitpunkten gesehen werden (Längsschnittsbetrachtung intraindividuelle Unterschiede) (Amelang & Bartussek, 2001).

Zusammenfassend lassen sich die folgenden fünf Forschungsfelder der Differentiellen Psychologie ausmachen (Amelang & Bartussek, 2001, S.4):

1. Die Beschaffenheit von Merkmalen, in denen es interindividuelle Differenzen gibt
2. Ausmaß dieser Differenzen
3. Wechselseitige Abhängigkeit solcher Merkmale
4. Ursachen der Differenzen
5. Beeinflussbarkeit durch Training, Umweltveränderungen, Medikamente, etc.

#### **4.1.2 Geschichtlicher Verlauf der Intelligenzmessung**

Die ersten „vorwissenschaftlichen diagnostischen Versuche zur Erfassung individueller Differenzen“ (Groffmann, 1983, S. 3) lassen sich schon in der Antike finden. Die Diagnostik umfasste jedoch zu dieser Zeit hauptsächlich physiologische Differenzen, weniger psychische Differenzen. Eine fruchtbare Basis für die Differentielle Psychologie und damit für die Intelligenzmessung bildete im 18.- und 19. Jahrhundert die Medizin, die Biologie und die Pädagogik. Hinsichtlich der medizinischen Forschung öffnete die Phrenologie (die Verknüpfung von Gehirn an intellektuelle Fähigkeiten) als auch die Arbeit von C.v. Linné (1707- 1778) bezüglich der Systematisierung von Geisteskrankheiten den Weg zur Differentiellen Psychologie und damit zur Intelligenzmessung (a.a.O).

Weitere Grundbausteine für die Differentielle Psychologie legten Darwin und Mendel. Darwin (1859) „On the origin of species by the means of natural selection“ stellte heraus, dass individuelle Differenzen nicht zufällig sind, sondern vielmehr durch die „natürliche Auswahl der im Kampf ums Überleben Bestangepassten“(Amelang & Bartussek, 2001, S. 19) bestimmt werden. Die Evolutionstheorie zeigte, dass die individuellen Differenzen vom Durchschnitt das Ergebnis der Interaktion zahlreicher Bedingungen ist. Auch Mendel postulierte- entgegen der teleologischen Erklärungsprinzipien- dass die „angeborene Eigenart des Individuums kausal durch die zufällige Kombination der Erbanlagen“(Amelang & Bartussek, 2001, S. 19) geformt wird.

Aufbauend auf diesen (und anderen) Erkenntnissen wurde im Rahmen der französischen Revolution das Schulwesen grundlegend reformiert, um den individuellen Differenzen der Individuen Rechnung zu tragen. Die Entwicklung äußerte sich in der Etablierung sogenannter allgemeiner Schulen und Hilfsschulen für Schwachbegabte. Ein weiterer Meilenstein in diesem Zusammenhang war der Auftrag des französischen Erziehungsministeriums an Binet und Simon, die ein Verfahren zur Klassifizierung von Begabungsunterschieden entwickeln sollten, das zum Einen zeitsparender und zum Anderen objektiver sein sollte als die vorherigen Prüfverfahren (Groffmann, 1983). Dieser Auftrag war der Anfang von intensiven Forschungsbemühungen von Binet und Simon (1905), die ihren Fokus bezüglich der Intelligenzmessung verstärkt auf die Normalpopulation legten, mit besonderer Akzentuierung von Minderbefähigten. Bevor auf Binet und Simon näher eingegangen wird, ist noch der Einfluss von W. Wundt (1832-1920) hervorzuheben, der die Experimentelle Psychologie konstituierte und entscheidende Beiträge zur Standardisierung der äußeren Bedingungen von Experimenten leistete. Eine weitere Erkenntnis seiner Arbeit war, dass sich die Ergebnisse der Probanden trotz verstärkter Bemühungen um die Kontrolle der Versuchsbedingungen nicht angleichen. Diese Feststellung rückte sodann die Berücksichtigung der individuellen Differenzen noch stärker in den Mittelpunkt der Betrachtung. Während W. Wundt jedoch noch als Begründer der Allgemeinen Psychologie gesehen wird, wird sein „Schüler“ Galton als Wegbereiter der Differentiellen Psychologie angesehen (Mehlhorn & Mehlhorn, 1981). In Anlehnung an Darwin und Mendel wird auch in Galtons (1869) umfassendem empirischen Datenmaterialien bzw. in seiner Publikation „Hereditary Genius“ die Vermutung geäußert, dass die Differenzen der intellektuellen Fähigkeiten vorwiegend erblich bedingt sind (Mehlhorn & Mehlhorn, 1981). In seinem Buch „Genie und Vererbung“ (Galton, übers. 1910) räumt er allerdings ein, dass soziale Faktoren, wie die Erziehung,

die Lebensumstände oder auch die persönlichen Lebensumstände des Menschen einen Einfluss auf die Höhe der Intelligenz haben können.

So formuliert Galton in seiner Einleitung, dass „ gegenwärtig soziale Faktoren alltäglicher Art, denen man solchen Einfluss nicht zuschreiben würde, wirken, und zwar die einen auf den Verfall, die anderen auf den Fortschritt der menschlichen Natur“ (Galton, 1910, zitiert nach Mehlhorn & Mehlhorn, 1981, S. 14). Mit seinen ersten systematischen Prüfverfahren wollte er die Variabilität physischer und psychischer Merkmale des Menschen sichtbar machen. In diesem Zusammenhang erschien erstmals der Testbegriff. Die von ihm entwickelten Tests sollten Leistungen, Fertigkeiten und Tätigkeiten, also ausschließlich beobachtbare und exakt kontrollierbare Verhaltensweisen quantitativ erfassen und Vergleiche zwischen Individuen ermöglichen. Mithilfe von einfachen statistischen Verfahren, dessen Begründer er ist, zeigte Galton auf, dass die Gauß'sche glockenförmige Normalverteilung auch für die geistigen Fähigkeiten bzw. die Intelligenz zutrifft. Darauf aufbauend konstruierte er ein simples Korrelationsverfahren (*index of co-relation*), um den Grad des Zusammenhangs zweier Variablen zu bestimmen (a.a.O.).

Unter dem Einfluss von Wundt und Galton stand auch J. Mc Keen Cattell (1860-1944). 1890 schrieb J. Mc Keen Cattell den berühmten Aufsatz „ Mental Test and measurement“. Durch die Verknüpfung von Geist, Begabung, Intelligenz (*mental*) mit dem von Galton entwickelten Testbegriff begründete er die eigentliche Testpsychologie. Die Testserie beinhaltete eine Reihe von zehn Tests, die körperliche, psychophysische und geistige Merkmale umfassten. Die Tests fanden vor allem durch die sehr pragmatische Herangehensweise in den USA Beachtung. Die Tests wurden indessen nicht wesentlich weiterentwickelt und stießen durch eine fehlende Validitätsprüfung- im Sinne von korrelationsstatistischen Daten zwischen Schulerfolg und Testergebnissen- an ihre Grenzen (Mehlhorn & Mehlhorn, 1981).

Auf der Grundlage einer statistischen und experimentalpsychologischen Methodik entwickelten Binet und Simon (1905) die ersten systematischen Intelligenzprüfungen, die eine hohe wissenschaftliche und praktische Relevanz hatten. Angestoßen wurden diese Forschungsanstrengungen, wie oben erwähnt, durch den Auftrag des französischen Erziehungsministeriums.

Die folgenden vier Kriterien wurden der ersten Binet- Skala zugrunde gelegt (Groffmann, 1983):

1. Die Tests sind in ihrer Anwendung und Auswertung standardisiert
2. Die objektive Schwierigkeitsanordnung der Aufgaben ist empirisch geprüft
3. Die Validität und Reliabilität werden als Gütekriterien vorausgesetzt
4. Die Stichprobe von Verhaltensweisen wird nicht in separate Teilfähigkeiten zergliedert, sondern zur allgemeinen Intelligenz vereinigt

Binet stellte mit diesen Annahmen die Weichen für die spätere Intelligenzforschung und verhalf dem Konzept der *Allgemeinen Intelligenz* zum Durchbruch. Binets Auffassung zufolge können die konstituierenden Einzelfähigkeiten der Intelligenz zu einem Globalindex zusammengefasst werden. Darüber hinaus erkannte Binet, dass die Intelligenzkapazität der Menschen individuell verschieden ist und vom Lebensalter abhängt (Binet, 1905, zitiert nach Amelang & Bartussek, 1990). Entsprechend dieser Erkenntnisse postulierte Binet, dass die Intelligenz hierarchisch aufgebaut ist bzw. in Form einer „*échelle métrique de l'intelligence*“ (Stufenleiter der Intelligenz) verbildlicht werden kann. In der Umsetzung wurde ein Fragenkatalog konzipiert, der von leichten zu schwierigeren Fragen ansteigt und an die entsprechenden Altersstufen angepasst wurde. Anhand der Fragen sollten der gesunde Menschenverstand, der praktische Sinn, die Initiative und die Fähigkeit sich anzupassen, gemessen werden. Als zentrale Bereiche der Intelligenz nennt Binet (1905, zitiert nach Amelang & Bartussek, 1990): „Gut urteilen, gut verstehen, gut denken“ (S. 177). Die Schwierigkeit der Aufgaben wurde so angelegt, dass sie von einer Höchstaltersstufe gerade noch gelöst werden können. Mithilfe der Stufenleiter ließ sich sodann eine Reihenfolge der Kinder hinsichtlich der gelösten Aufgaben bestimmen. Gleichzeitig konnten Durchschnittswerte der gelösten Aufgaben für verschiedene Altersstufen ermittelt werden. Daraufhin konnten die Leistungen der Kinder beurteilt werden, ob sie dem angemessenen Niveau ihres Alters entsprechen (*les réguliers*), es übertreffen (*les avances*) oder es unterschreiten (*les retards*). Aufgrund dieser Beurteilungen wurde das Intelligenz-Grundalter für jede Altersstufe definiert. Trotz der positiven Aufnahme und der schnellen Verbreitung der Tests zwischen 1908- 1912 gab es auch kritische Stimmen. Die Kritik richtete sich vor allem gegen die ungenügende Praktikabilität bei größeren Stichproben, da die Tests sehr umfassend und zeitaufwändig waren. Weiterhin wurde die Beeinflussung der Probanden durch die Testleiter (Testleitereffekt) und die fehlende Leistungsdifferenzierung zwischen verschiedenen

Handlungsbereichen bzw. die fehlende Berücksichtigung von Intelligenzstrukturen bemängelt. Daraufhin revidierten Binet und Simon die ersten beiden Skalen und entwickelten 1911 die dritte Binet-Skala (Groffmann, 1983). In Anbetracht der geleisteten Arbeit von Binet und Simon lässt sich festhalten, dass sie die Urväter der Testpsychologie sind und ihre Erkenntnisse bis heute Gültigkeit besitzen. Eine beachtliche Adaption des Stufentests von Binet und Simon (1905) stellen die bekannten Stanford- Binet- Revisionen dar, die bis zu einer dritten Version verfeinert wurden (Terman & Merrill, 1960). Terman und Merrill beachteten die oben angeführte Kritik und verbesserten den Test vor allem hinsichtlich seiner Reliabilität, Validität und Praktikabilität (Groffmann, 1983).

Stern (1871- 1938) verfeinerte das Verfahren von Binet und Simon ebenfalls, indem er den Intelligenzquotienten entwickelte, der nun einen Vergleich von Kindern unterschiedlichen Alters erlaubte. Auf dem Berliner Kongress für experimentelle Psychologie (1912) machte er den berühmt gewordenen Vorschlag, die Intelligenz und das Lebensalter in Relation zu setzen (Stern, 1912, S.28ff):

$$\text{Intelligenzquotient (IQ)} = \text{Intelligenzalter} / \text{Lebensalter}$$

Später wurde der IQ mit 100 multipliziert, um eine bessere Übersicht im Sinne von ganzen Zahlen zu gewährleisten. Das entscheidende Problem, mit dem der IQ behaftet war, zeigte sich in der ungenügenden Berücksichtigung des Zusammenhangs zwischen Intelligenz und Alterseinflüssen bei Erwachsenen. In diesem Kontext trifft die Grundannahme vom Anstieg der Intelligenz mit zunehmendem Alter nicht zu. Folglich zeigten sich uneinheitliche Testergebnisse (Groffmann, 1983). Darüber hinaus erschien in Sterns Publikation (1912) auch seine viel zitierte Definition der Intelligenz: „Intelligenz ist die allgemeine Fähigkeit eines Individuums, sein Denken bewusst auf neue Forderungen einzustellen; sie ist allgemeine geistige Anpassungsfähigkeit an neue Aufgaben und Bedingungen des Lebens“ (S.3). Die Definition ist auch deshalb bahnbrechend, weil sie erste Ansätze der von ihm entwickelten Konvergenztheorie beinhaltet, die das Zusammenwirken von Anlage und Umwelt thematisiert. „Intelligent ist nur derjenige, der unter verschiedensten Bedingungen und auf den verschiedensten Gebieten sich neuen Anforderungen geistig leicht anzupassen vermag“ (Stern, 1912, S.4).

In Anbetracht der problematischen Hypothese von Stern und Binet (1905), dass eine lineare Beziehung zwischen Intelligenz und Alter besteht, führte Wechsler (1958) den IQ als Abweichungsquotienten ein, um so auch ein vergleichbares Maß im Erwachsenenalter zu schaffen (Amelang & Bartussek, 1990, S. 181):

$$IQ = 100 + 15 * [(X - M) / \sigma]$$

( X= Messrohwert; M= Durchschnitt der Messrohwerte der Altersgruppe;  $\sigma$ = Standardabweichung der Messrohwertverteilung der Altersgruppe)

Die empirisch nachgewiesene Normalverteilung der Messrohwerte der jeweiligen Altersstufe ermöglicht es, die Messwerte jeder Altersstufe in Relation zu setzen und zu standardisieren. Unter Berücksichtigung der Streuung und Standardabweichung der verschiedenen Altersstufen können folglich auch die Durchschnittswerte bestimmt werden. Auf diese Weise kann der Mittelwert, der auf 100 normiert ist, mit anderen Leistungen verglichen werden (Groffmann, 1983).

Ein entscheidender Kritikpunkt der Tests wurde stets durch den immensen Zeitaufwand bei der Durchführung sowie bei der Auswertung begründet. Infolgedessen wurden in den USA -auch unter dem Druck des Ersten Weltkrieges- Gruppentestverfahren entwickelt, um die Wehrdienstleistenden schnell klassifizieren zu können (Alpha-Examen, 1932/33; Beta- Examen; Army General Classification Test- AGCT, 1940-1948) (a.a.O.).

Im Intelligenz-Struktur-Test-70 von Amthauer (1970) lassen sich einige Ansätze dieser Gruppentests wiederfinden.

Die Gruppentests haben u.a. folgende bedeutsame Vorteile (Groffmann, 1983):

- Sie werden unter gleichen Bedingungen für alle durchgeführt.
- Sie werden in gleicher Weise ausgewertet.
- Alle Personen bekommen die gleiche Anweisung.
- Sie lassen sich zum Teil in Multiple-Choice-Form konstruieren.

### 4.1.3 Abschlussbetrachtung und Ausblick der Intelligenzforschung

Die Vorgeschichte der Intelligenzmessung ist von zentraler Bedeutung für ein tieferes Verständnis der weiteren Fortschritte der Intelligenzforschung. In den dargestellten Zeitepochen wurden wichtige Grundlagen gelegt, die bis in die heutige Zeit Gültigkeit besitzen. Resümierend lassen sich folgende Erkenntnisse festhalten: Es wurden die Differenzen der einzelnen Individuen in den Mittelpunkt der Betrachtung gerückt und die vornehmlich teleologischen Ansichtsweisen von der Gleichheit aller Menschen zurückgedrängt. Auf dieser Basis schuf Galton, der als Begründer der Differentiellen Psychologie gesehen wird, den Testbegriff, um die individuellen Differenzen messbar zu machen. Schon zu diesem Zeitpunkt wurde neben den erblich bedingten Differenzen der Einfluss der Umwelt auf die geistigen Fähigkeiten gesehen. Die Gauß'sche glockenförmige Normalverteilung als auch weitere statistische Hilfsmittel wie der Korrelationskoeffizient und die Faktorenanalyse waren Meilensteine auf dem Gebiet der Intelligenzforschung (Groffmann, 1983). Sie bilden auch heutzutage die Grundlage für sämtliche statistische Berechnungen. Insbesondere die Faktorenanalyse leistete entscheidende Beiträge zur Datenreduktion und zur Interpretation der Messergebnisse. Binet und Simon (1905) verhalfen dem Konzept der Allgemeinen Intelligenz (später als *g-Faktor* bekannt) als übergeordnete Größe der einzelnen Teilfähigkeiten zum Durchbruch. Der von Stern (1912) entwickelte und von Wechsler (1958) überarbeitete Intelligenzquotient ist noch heute gebräuchlich. Darüber hinaus machte Stern in seiner Konvergenztheorie abermals deutlich, dass sowohl die Anlagen als auch die Umwelt einen entscheidenden Einfluss auf die Intelligenz haben. Diese Überlegung ist immer noch Gegenstand aktueller Diskussionen und wird im Kapitel 6.3 hinsichtlich der fluiden und kristallisierten Intelligenz wieder aufgegriffen. An dieser Stelle sei ein Zitat von Groffmann erwähnt, welches die Kontroverse treffend umschreibt: „Kultureller Wandel wandelt die Menschen; er wandelt auch deren Intelligenz“ (Groffmann, 1983, S. 76).

Letztendlich kann die Intelligenzdiagnostik auf mehr als ein Jahrhundert und 800.000 Publikationen (bis 1980) zurückblicken (Mehlhorn & Mehlhorn, 1981). Trotz der Fülle der Befunde und Entwicklung im Bereich der Intelligenzmessung werden weitere Forschungsanstrengungen verlangt, um das Verständnis über die Intelligenz zu erweitern.



## 4.2 Begriffsbestimmung der Intelligenz

Die Intelligenzforschung zählt, wie der Entwicklungsgeschichte zu entnehmen ist, zu den ältesten Themen der Psychologie. Obgleich in den letzten 100 Jahren die Intensität der Forschung weiter zunahm, lässt sich bis heute kein eindeutiger Intelligenzbegriff definieren. Sowohl die Laien als auch die Wissenschaftler finden hierbei keine präzise Antwort. Mit einer gewissen Ironie formuliert Dockrell (1969, zitiert nach Roth, Oswald & Daumenlang, 1980) die Situation anlässlich eines Symposiums über Fragen der Intelligenzforschung folgendermaßen: „Die Verschiedenheit der Intelligenzbegriffe, die von den Teilnehmern dieses Symposiums vertreten werden, reduziert die Gefahr, dass einer der Standpunkte als richtig erkannt wird, auf ein Minimum“ (S. 7).

Zumindest gibt es vier allgemein akzeptierte Feststellungen im Hinblick auf die Intelligenzforschung (Stern & Guthke, 2001, S. 9):

1. Die hohen Langzeitstabilitäten in der Intelligenztestleistung, die bereits in der frühen Kindheit um  $r = .50$  betragen und ab der späten Kindheit auf über  $r = .80$  ansteigen, sprechen für das Vorliegen eines stabilen Persönlichkeitsmerkmals.
2. In Kulturkreisen, in denen Kindern weitgehend alle Lerngelegenheiten offen stehen, können mindestens 50 Prozent der Varianz in der Intelligenztestleistung durch genetische Unterschiede erklärt werden.
3. Unterschiede in der Intelligenz haben eine zerebrale Grundlage, die prinzipiell aufgeklärt werden kann.
4. Es hängt entscheidend vom kulturellen und individuellen Kontext ab, wie kognitive Kompetenzen die Aneignung von Wissen beeinflussen.

Darüber hinaus lässt sich festhalten, dass die Intelligenz -wegen ihrer gesellschaftlichen Allgemeingültigkeit und Akzeptanz- von entscheidender Bedeutung für die westliche Kultur ist.

Im Folgenden werden mögliche Definitionen der Intelligenz und die Grenzen ihres Bedeutungsspektrums aufgezeigt. Danach wird das *definitive Verständnis* gegenüber dem *Konstrukt-Verständnis* der Intelligenz abgegrenzt und der hohe Stellenwert des letzteren innerhalb der Intelligenzforschung hervorgehoben. Des Weiteren werden die allgemeinen Bedeutungsbereiche und Bedeutungskomponenten der Intelligenz thematisiert. Auf dieser Grundlage werden letztlich ausgewählte Strukturtheorien vorgestellt.

Im Zentrum des Interesses steht hier das hierarchische Modell von Cattell (1971a, 1987), welches die Faktoren der fluiden und kristallisierten Intelligenz auf der zweiten Ebene abbildet. Die kristallisierte Intelligenz ist hier von besonderer Bedeutung, weil sie einen engen Zusammenhang mit dem (kulturabhängigen) Wissen aufweist. Das komplexe Beziehungsgeflecht zwischen der kristallisierten Intelligenz und dem Wissen wird im Kapitel 6 näher beleuchtet. Die übergeordnete Bedeutung der Intelligenz und des Wissens für schulische oder berufliche Leistungen (und damit für die Eignungsdiagnostik) soll in den nachfolgenden Kapiteln sukzessive herausgestellt werden.

#### 4.2.1 Definition von Intelligenz

Die Vorstellungen über die Intelligenz divergieren in der aktuellen Literatur erheblich. Wie selbstverständlich spricht man über die Erforschung der Intelligenz im Allgemeinen. Routinemäßig wird der spezifische Ausprägungsgrad der Intelligenz eines Individuums gemessen, und die meisten der traditionellen Definitionen sind häufig pragmatischer Natur: „Intelligenz ist...“ Es kann jedoch nicht davon ausgegangen werden, dass dem einheitlichen Begriff der Intelligenz ein ebenso einheitliches, reales Faktum entspricht. Die Intelligenz im engeren Sinne ist nicht beobachtbar, sondern vielmehr als hypothetisches Konstrukt zu verstehen (Brocke & Beauducel, 2001).

Im Rahmen des definitorischen Verständnisses wird der Versuch unternommen, die Intelligenz mit präzisen Definitionen zu beschreiben. Es sei jedoch angemerkt, dass die Schwerpunkte dieser Definitionen jeweils unterschiedliche Aspekte der Intelligenz betonen. Im Folgenden werden einige Beispiele von Explizitdefinitionen genannt:

*Tabelle 6: Intelligenzdefinition  
(Brocke & Beauducel, 2001, S.14)*

Autor	Definiendum	Definiens
Binet & Simon, 1905	Intelligenz ist...	...“die Art der Bewältigung einer aktuellen Situation“, genauer: „gut urteilen, gut verstehen und gut denken.“
Hofstätter, 1957	Intelligenz ist...	...“(...) das Ensemble von Fähigkeiten, das den innerhalb einer bestimmten Kultur Erfolgreichen gemeinsam ist.“

Groffmann, 1964	Intelligenz ist...	...“die Fähigkeit des Individuums, anschaulich oder abstrakt in sprachlichen, numerischen oder raumzeitlichen Beziehungen zu denken (...).“
Rohracher, 1965	Intelligenz ist...	...“der Leistungsgrad der psychischen Funktionen bei ihrem Zusammenwirken in der Bewältigung neuer Situationen.“
Stern, 1911	Intelligenz ist...	...“(...) eine durchaus formale Eigenschaft: sie bezieht sich auf eine Fähigkeit, die Geistesbewegung jeweiligen neuen Aufgaben anpassen zu können.“

Der zu erklärende erste Teil dieser Definition wird als *Definiendum* bezeichnet und der erklärende zweite Teil als *Definiens*. Die Problematik dieser Explizitdefinitionen liegt darin begründet, dass beide Teile gleichgesetzt werden. Entsprechend dieses Sachverhalts müsste jeder Intelligenztest bei Verwendung einer Explizitdefinition das Gleiche messen. Eine Alternative wäre hingegen, dass man für jede Definition ein neues Definiendum festlegt. Dadurch würden allerdings inflationär neue Bezeichnungen von Intelligenzen entstehen. Des Weiteren ist den Explizitdefinitionen inhärent, dass sie sehr umfangreich sein müssen, um die Intelligenz in ihren verschiedenen Facetten zu beschreiben. Infolgedessen werden die Explizitdefinitionen jedoch ungenauer und es entsteht eine Art Leerformelcharakter (Brocke & Beauducel, 2001). Dasselbe gilt, wenn man für die Termini „vernünftig“, „erfolgreich“, „produktiv“, usw. das Attribut „intelligent“ einsetzt. Eine sinnvolle empirische Überprüfung anhand von Hypothesen erscheint unter Berücksichtigung dieser Aspekte nicht möglich zu sein.

Zusammenfassend lässt sich sagen, dass die verbalen Definitionen für sukzessive Erweiterungen nicht offen sind (Amelang & Bartussek, 2001). Diese Problematik wird auch als „Dilemma des Operationalismus“ (Bridgeman 1927, zitiert nach Brocke & Beauducel, 2001, S.15) bezeichnet. Eine andere Möglichkeit zur Überwindung dieses Dilemmas könnten operationale Definitionen darstellen. Die dahinter stehende Idee ist, dass die Testverfahren die Definition des entsprechenden Merkmals (hier Intelligenz) darstellen. Mit anderen Worten „wäre Intelligenz dasjenige, was der Intelligenztest misst“ (Boring 1923, zitiert nach Amelang & Bartussek, 2001, S. 191).

„Die Definition der theoretischen Variablen erfolgt somit über die Detaillierung und Spezifizierung der zum Zwecke ihrer Messung ausgeführten empirisch-experimentellen Handlungen“ (Amelang & Bartussek, 2001, S.191). Der Nachteil dieser operationalen Definition ist die fehlende Erklärung, welches Verfahren die Intelligenz wirklich messen kann. Der Vorteil liegt in der detaillierteren Beschreibung der Tests begründet, wodurch das Merkmal Intelligenz genauer erfasst werden kann (a.a.O.).

#### 4.2.2 Intelligenz als Konstrukt

Intelligenz wird in der einschlägigen Literatur als komplexes Konstrukt und im engeren Sinne als Trait bzw. Disposition bezeichnet. Das Konstrukt-Verständnis ist gekennzeichnet durch eine Vielzahl von kognitiven Einzelfähigkeiten. Im Laufe der Zeit werden die Teilfähigkeiten durch Forschungsanstrengungen sukzessive erschlossen. Diese Teilfähigkeiten manifestieren sich zum Beispiel in den Begriffen wie verbale Intelligenz, numerische Intelligenz oder soziale und emotionale Intelligenz. Das Ergebnis der Forschungsanstrengungen ist ein sich ständig erweiterndes Spektrum von Intelligenzfaktoren. In diesem Zusammenhang wird das Intelligenzkonstrukt leistungsfähiger und entfaltet zum Zeitpunkt des Forschungsstandes seine Bedeutung. Mit anderen Worten ist das Konstrukt ständig erweiterungsfähig und lässt daher auch keinen Einheitstest zu. Demnach wird die Intelligenz auch als offenes Konstrukt bzw. hypothetisches Konstrukt bezeichnet (McCorquodale & Meehl, 1948, zitiert nach Brocke & Beauducel, 2001). Im Gegensatz dazu steht das definitorische Verständnis von Intelligenz, welches die gesamte Breite des Bedeutungsspektrums eines offenen Konstrukts in der Regel nicht erfassen kann (Brocke & Beauducel, 2001).

In der Psychologie liegt dem Konstrukt-Verständnis folgende Auffassung zugrunde: „Intelligenz als theoretischer Begriff bezeichnet eine latente Eigenschaft (Disposition), die selbst nicht beobachtbar ist und sich nur in beobachtbaren intelligenztypischen Verhalten manifestiert, zum Beispiel im Lösen von Testaufgaben“ (Brocke, 2000, S. 226). In diesem Zusammenhang spricht man wie oben erwähnt von einem hypothetischen Konstrukt oder Dispositionskonstrukt. Das Dispositionskonstrukt wird durch bestimmte *Zuordnungsregeln* festgelegt. Mit anderen Worten wird das Konstrukt (Definiendum, hier: *akademische Intelligenz*= P, siehe Tabelle 7: Beispiel A und B) mit den definierenden Beobachtungsvariablen (hier: Anzahl richtiger Lösungen= R) verkettet, wobei hier eine spezifische (Situations-) Bedingung (hier: HAWIE, Intelligenz-Struktur-Test

2000= S) vorangestellt wird (Brocke & Beauducel, 2001). Es wird also eine „Situationsklasse spezifiziert, bei deren Vorliegen die im Trait spezifizierte Disposition (Verhaltenstendenz) in einem beobachtbaren Verhalten manifestiert wird“ (Brocke & Beauducel, 2001, S.15). Im Rahmen der Operationalisierung wird demnach das nicht beobachtbare Konstrukt (hier: Intelligenz) mit beobachtbaren Situations- und Verhaltensvariablen in Beziehung gesetzt und damit auch partiell definiert (siehe Tabelle 7).

*Tabelle 7: Beispiele für Zuordnungsregeln  
(Brocke & Beauducel, 2001, S. 16)*

<b>Beispiele für Zuordnungsregeln</b>	
Beispiel (A)	Wenn Personen den IST mitmachen (S1), dann erzielen sie (oder genau dann, wenn sie...) über eine durchschnittliche akademische Intelligenz (P1) verfügen, eine festgelegte Testpunktzahl (R1).
Beispiel (B)	Wenn Personen den HAWIE mitmachen (S2), dann erzielen sie (oder genau dann, wenn sie...) über eine durchschnittliche akademische Intelligenz (P2) verfügen, eine festgelegte Testpunktzahl (R1).
Beispiel (C)	Aus A) und B) folgt (Konklusion): Wenn Personen den IST (S1) mitmachen und eine bestimmte Testpunktzahl (R1) erreichen, dann erzielen sie, wenn sie den HAWIE (S2) mitmachen, eine bestimmte andere Testpunktzahl (R2).
S= Situationskomponente; P= Dispositionskonstrukt; R= Verhaltensvariable	

Zuordnungsregeln können aufgrund der Ausführungen als Situations-Konstrukt-Verhaltensverknüpfungen bezeichnet werden. Es wird betont, dass die Zuordnungsregeln und die damit einhergehende Bedeutung des Konstrukts immer im Kontext des vorangestellten Situationstypus gesehen werden müssen. Im Unterschied zu den Explizitdefinitionen können in anderen Situationen durchaus zusätzliche Zuordnungsregeln eingeführt werden, wobei die Präzision des Konstrukts keinesfalls leiden muss und somit die Bedeutung erweiterungsfähig bleibt. In einer weiterentwickelten Form können die Zuordnungsregeln auch liberalisiert werden (a.a.O.). „Durch eine Liberalisierung können vor allem *Korrelationen* als Surrogate von bilateralen Reduktionssätzen verwendet werden. Dadurch können beispielsweise Validitätskoeffizienten (z.B. Binnen- oder der Kriteriumsvalidität), Item- Trait- Korrelationen (Trennschärfen) oder Ladungskoeffizienten als probabilistische Manifestationsgesetze bzw. Zuordnungsregeln behandelt und theoretisch eingeordnet werden“ (Brocke & Beauducel, 2001, S. 17).

Die Liberalisierung öffnet den Weg für den Einsatz statistischer Methoden, wie z.B. der Faktorenanalyse. Die zweifache Verknüpfung (siehe Abbildung 5) mit Situations- und Verhaltensvariablen ist bei Dispositionskonstrukten nach wie vor unerlässlich. Die Suche nach den Verhaltensindikatoren, die am besten dem Konstrukt der Intelligenz entsprechen, ist gemäß den dargelegten Ausführungen die entscheidende Aufgabe, um den Bedeutungsgehalt der Intelligenz zu spezifizieren. Die Summe der Verknüpfungen zwischen der Disposition und den entsprechenden Indikatoren gibt demnach den Bedeutungsgehalt wider (Brocke, 2000).

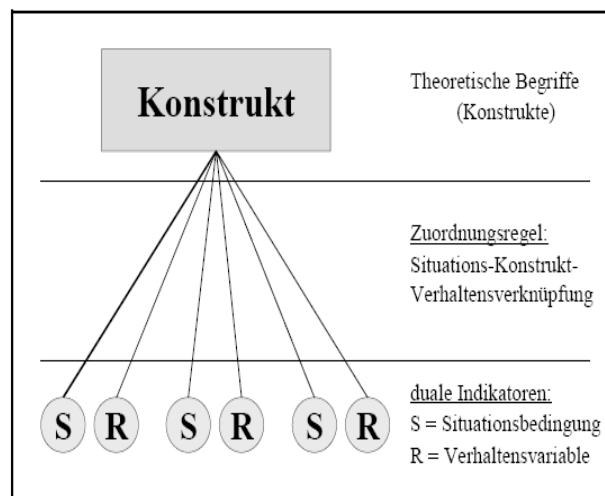


Abbildung 5: Konstrukte und Zuordnungsregeln (Brocke & Beauducel, 2001, S.17)

### **Konstrukte und Theorien**

Eine empirische Bestätigung (Validierung) erfährt das Konstrukt durch die Verwendung von mindestens zwei Zuordnungsregeln. In Anlehnung an das Beispiel lässt sich aus den beiden Zuordnungsregeln (siehe Tabelle 7: Beispiel A&B) eine testbare Implikation generieren, indem man die nicht beobachtbare Variable bzw. das Konstrukt (hier: akademische Intelligenz) durch die beiden beobachtbaren Variablen (Testpunktzahlen im HAWIE und I-S-T 2000) ersetzt (siehe Tabelle 7: Beispiel C). Demzufolge erlangen die Zuordnungsregeln bzw. Manifestationsgesetze, wie die Postulate einer Theorie, empirischen Gehalt (*Empirizitäts-Kriterium*). Darauf aufbauend kann das Konstrukt im Rahmen der Differentiellen Psychologie auch als Theorie verstanden werden, welches anhand von *Postulaten* (nomologische Aussagen) mit umfangreicheren nomologischen Netzen verknüpft wird. Charakteristische nomologische Aussagen im Rahmen der Intelligenzforschung sind u.a. Zusammenhänge zwischen Intelligenz und Alter, Intelligenz

und Bildungsniveau, Intelligenz und Berufserfolg, etc. Die Quintessenz ist, dass das Konstrukt Intelligenz durch die Einbindung in verschiedene nomologische Netze mit wiederum anderen Konstrukten weitere Bedeutungskomponenten erhält (Brocke & Beauducel, 2001). Diese nomologischen Netze sind im Weiteren häufig hierarchisch aufgebaut (u.a. Cattell, 1971a, 1987). Im Rahmen der hierarchischen Strukturtheorien, die in Kapitel 4.4 näher betrachtet werden, werden „Konstrukte einer höheren Generalitätsebene durch Teilkonstrukte auf der spezifischen Ebene strukturiert und in ihrer Beziehungskomponente differenziert, zum Beispiel „akademische Intelligenz“ durch verbale, numerische, figural-räumliche Intelligenz und weiteren Fähigkeitskomponenten“ (Brocke & Beauducel, 2001, S.18).

#### **4.2.3 Zusammenhang zwischen dem Konstrukt-Verständnis und der Intelligenzmessung**

Es stellt sich im Weiteren die Frage, wie die Intelligenz erfasst werden kann. Im Hinblick auf den oben aufgeführten Sachverhalt werden „Zuordnungsregeln für Intelligenz auf ein Individuum als Intelligenzmessung“ (Brocke & Beauducel, 2001, S.19) verstanden. Bezeichnend ist die in einem Test erreichte Punktzahl. Ein solcher Rohwert ist allerdings nicht von Nutzen, wenn dieser nicht mit anderen Rohwerten aus verschiedenen Tests verglichen werden kann. Aus diesem Grund benötigt man einen standardisierten bzw. normierten (z-) Wert, der mittels einer Eichstichprobe ermittelt werden kann. Der Wert gibt an, wo sich ein individueller Wert im Verhältnis zu seiner Vergleichspopulation bzw. zu seiner Bezugsgruppe befindet.

### **4.3 Bedeutungsbereiche der Intelligenz**

Zusammenfassend lässt sich vorerst festhalten, dass die Bedeutung von Intelligenz anhand von Zuordnungsregeln und Postulaten nach entsprechenden methodologischen Regeln festgelegt und sukzessive erschlossen werden kann (s.o.). Die Bedeutungsbereiche werden in diesem Zusammenhang als Teilkonstrukte verstanden, die einen Teil des übergeordneten Gesamtkonstrukts darstellen. Im Folgenden wird insbesondere die akademische Intelligenz fokussiert, weil sie eine hohe Relevanz für die vorliegende Arbeit hat. Bevor auf die akademische Intelligenz näher eingegangen wird ist es nützlich, zwischen inhaltlichen Leitvorstellungen, paradigmenspezifischen Bedeutungskomponenten

sowie theoriespezifischen und -übergreifenden Bedeutungsbereichen zu differenzieren (Brocke & Beauducel, 2001).

### 4.3.1 Methodische Ansätze der Intelligenzforschung

Unter inhaltlichen Leitvorstellungen lassen sich unter anderem die akademische, praktische, soziale Intelligenz oder die Kreativität subsumieren. Die oberflächlichen Unterscheidungen unterliegen aber häufig subjektiven Urteilen und werden hypothetisch unterstellt (Ford & Tisak, 1983). Inhaltliche Unterscheidungen sind hier nicht auszuschließen. Bisweilen erreichen die meisten Intelligenzbereiche noch keinen wissenschaftlichen Status. Dazu müssten die Unterscheidungen bzw. Abgrenzungen stärker präzisiert und empirisch fundiert werden (z.B. über Zuordnungsregeln). Die oben aufgeführten Bedeutungsbereiche können hingegen beim gegenwärtigen Forschungsstand und bei Verwendung eines konservativen Maßstabes differenziert werden. Bevor die inhaltliche Abgrenzung der akademischen Intelligenz erfolgt, werden im Folgenden die methodischen Paradigmen bzw. das *setting* dargelegt.

Im Bezug auf die Forschungsmethode ist es grundsätzlich möglich, zwischen Struktur- und Prozessforschung zu differenzieren (Amelang & Bartussek, 2001). In der weiteren Ausführung wird ausschließlich diese Unterscheidung fokussiert, weil sie hinsichtlich der in Kapitel 4.4 vorgestellten Strukturtheorien von besonderer Relevanz ist.

Innerhalb der Strukturforschung ist die Anzahl richtiger Lösungen in Intelligenzaufgaben von besonderer Bedeutung. In der Prozessforschung stehen zumeist die gemessenen Reaktionszeiten im Vordergrund der Analyse. Im Rahmen der Prozessforschung sollen demnach die kognitiven Prozesse herauskristallisiert werden, die für die Leistungen in den Intelligenztests verantwortlich sind (Brocke & Beauducel, 2001). Zum Einen kann dies durch eine Selbstbeschreibung der Probanden beim Lösen der Intelligenzaufgaben erfolgen (Ertel & Schindler, 1969). Zum Anderen sind die Ansätze der allgemeinspsychologischen Informationsverarbeitungsforschung (IPS- Ansatz) von großer Wichtigkeit. Die IPS- Ansätze unterstellen in diesem Zusammenhang eine Verknüpfung von Intelligenzleistungen und Qualität der Verarbeitung. Die individuellen Differenzen zwischen den Probanden äußern sich in der unterschiedlichen Leistungsfähigkeit im Hinblick auf das Aneignen, Speichern, Verbinden, Vergleichen und Erinnern von Informationen (Humphreys, 1979). Hierbei wird zwischen dem kognitiven Komponenten-Ansatz und dem kognitiven Korrelate-Ansatz differenziert.



Hinsichtlich des kognitiven Komponenten-Ansatzes sollen die kognitiven Teilprozesse ermittelt werden, die die interindividuellen Differenzen in den Intelligenzaufgaben erklären und die mit den Leistungen in den Intelligenztests korrelieren. In der Umsetzung werden die Intelligenzaufgaben nach bestimmten inhaltlichen Themengebieten (z.B. mathematische Aufgabenstellungen) zusammengefasst. Im weiteren Vorgehen werden die kognitiven Teilprozesse analysiert, die bei der Bearbeitung eines Themengebietes für den Leistungsunterschied verantwortlich sind (Brocke & Beauducel, 2001). Eine Form des kognitiven Komponenten-Ansatzes im Rahmen der Prozessforschung könnte z.B. ein Lerntest sein. Hier wird bei wiederholter Darbietung verschiedener Aufgaben die Veränderung der Bearbeitungszeiten erfasst.

Beim kognitiven Korrelate-Ansatz werden die Personen untersucht, die bei den Intelligenztests unterschiedliche Leistungen erbringen. Daraufhin werden die unterschiedlichen kognitiven Teilprozesse bzw. Vorgehensweisen zur Lösung der Aufgaben herauskristallisiert, die den Leistungsunterschied erklären können. Die Korrelate der psychometrischen Tests sind demnach die Prozesse, welche anhand von Reaktionszeiten bestimmt werden. Eine Form der kognitiven Korrelate-Ansätze ist u.a. der *Mental-Speed-Ansatz*, der den „speed“ der kognitiven Prozesse mit den Intelligenzleistungen in Verbindung bringt (Neubauer, 1995).

Bezogen auf den IPS- Ansatz hat Sternberg (1977) einige wichtige Forschungsergebnisse veröffentlicht: Dabei stellt Sternberg die Bearbeitung von Analogieaufgaben in den Vordergrund der Betrachtung und gliedert die dazu notwendigen kognitiven Teilprozesse in Merkmalsentdeckung (Enkodierung), Merkmalsvergleichsprozesse und Bewertungsprozesse. Diese kognitiven Teilprozesse werden dann in Relation zum benötigten Zeitaufwand gesetzt. Das Ergebnis zeigt, dass hohe Leistungen in Intelligenztests mit einem geringen Zeitaufwand in den Bereichen Merkmalsvergleich und Bewertung einhergehen. Mit anderen Worten brauchen intelligentere Personen weniger Zeit bei der Enkodierung von Analogieaufgaben. Eine mögliche Schlussfolgerung könnte hier sein, dass intelligente Personen in der Regel auch schneller in der Lage sind, Zusammenhänge zwischen Analogieaufgaben zu erkennen. Diese Vermutung ist jedoch nicht endgültig bewiesen und bedarf weiterer Forschungsanstrengungen.

Es wird angemerkt, dass die akademische Intelligenz, die im Weiteren fokussiert wird, beim gegenwärtigen Forschungsstand die zahlreichsten Befunde bezüglich der dargestellten Unterscheidung aufweist.

Das heißt, die akademische Intelligenz ist sowohl Gegenstand der Strukturforschung als auch der Prozessforschung (Brocke & Beauducel, 2001).

### **4.3.2 Die akademische Intelligenz als theorieübergreifender Bedeutungsbereich**

Im Hinblick auf die Verknüpfung der verschiedenen Bedeutungsbereiche der Intelligenz (s.o.) können die Konstrukte, Modelle, Tests insoweit systematisiert werden, als dass sie eher theorieübergreifend oder theoriespezifisch sind.

Die Forschung bezüglich der akademischen Intelligenz ist sehr umfassend und erfolgreich gewesen. Dieser Sachverhalt impliziert aber auch, dass diesbezüglich verschiedene Modelle bzw. Theorien entwickelt worden sind und wirft sogleich die Frage auf, inwieweit diese Konzepte in Verbindung stehen (Brocke & Beauducel, 2001). Es zeigt sich insbesondere bei dem Konstrukt der akademischen Intelligenz, dass sie Bestandteil mehrerer Intelligenzmodelle ist und somit als theorieübergreifend bezeichnet werden kann. In Abgrenzung zu diesem theorieübergreifenden Konstrukt stehen z.B. die Bedeutungsaspekte der fluiden und kristallisierten Intelligenz, die nur speziellen Modellen zugeordnet werden können (Brocke, 2000).

Vorab sei angemerkt, dass die in Kapitel 4.4 dargestellten Strukturtheorien das theoretische Fundament für die akademische Intelligenz bilden. Die akademische Intelligenz äußert sich verstärkt in Fertigkeiten und Fähigkeiten, die in der Schule und in anderen akademischen Bildungseinrichtungen gebraucht werden. Nach Neisser (1976) wird die Bedeutung der akademischen Intelligenz über Merkmale abgeleitet, die den Aufgaben zur Erfassung dieses Teilkonstrukts gemeinsam sind.

Die Aufgaben sind dadurch charakterisiert, dass sie...

- ... von anderen Personen (nicht vom Probanden) formuliert werden,
- ... meist nur von geringerem oder gar keinem intrinsischen Interesse sind,
- ... alle benötigten Informationen von Beginn an zur Verfügung stehen,
- ... von den allgemeinen Erfahrungen mehr oder weniger abgehoben sind,
- ... gut strukturiert sind,
- ... meist nur eine richtige Antwort haben,
- ... meist nur über einen angemessenen Lösungsweg verfügen.

(Brocke & Beauducel, 2001, S.27)

Die aufgeführten Merkmale bestimmen jedoch nicht die Bedeutung des Konstrukts der akademischen Intelligenz, sie liefern lediglich Hinweise. Die Bedeutung wird hauptsächlich anhand der Strukturtheorien entwickelt. Obgleich die akademische Intelligenz als eigenständiger Bedeutungsbereich innerhalb der Intelligenz aufgefasst werden kann, ist sie wie oben angedeutet durch eine Vielfalt von Faktoren, Modellen und Theorien gekennzeichnet. Aufgrund der Fülle von verschiedenen Integrationsmöglichkeiten zur Systematisierung der zahlreichen Modelle und Theorien werden daher im Folgenden lediglich zwei ausgewählte empirisch orientierte Arten und zwei formal theoretische Arten der akademischen Intelligenz vorgestellt (Brocke & Beauducel, 2001).

1. Der *Faktoren-Elektizismus* bezeichnet eine Sammlung von Intelligenzfaktoren in Form einer Aufgabenbatterie zum Zeitpunkt des aktuellen Forschungsstandes (French, Ekstrom & Price, 1963). Diese Form hat den Vorteil, dass sie ein hohes Maß an Transparenz bieten kann, so dass subjektive Urteile eingeschränkt werden können und die Bewertung der Befunde objektiver ist. Außerdem harmonisiert diese Form gut mit dem Verständnis des offenen Konstrukts. Der Nachteil liegt darin begründet, dass die Gewichtung der Teilaspekte der Intelligenz von der bis dato zusammengetragenen und dokumentierten Forschung abhängt. Außerdem ist die Abgrenzung zwischen modellspezifischen und modellübergreifenden Faktoren nicht eindeutig, so dass inhaltsgleiche Faktoren womöglich doppelt mit einbezogen werden. (Brocke & Beauducel, 2001).

2. Die *Aufgaben-Integration* bezeichnet die Gewinnung eines Aufgabenpools, der für sämtliche in der Intelligenzforschung eingesetzten Aufgaben repräsentativ ist. In diesem Zusammenhang hat Jäger (1982; 1984) sich mit ca. 2000 verschiedenen Aufgaben auseinandergesetzt und die Mannigfaltigkeit dieses Aufgabenpools unter Berücksichtigung der Vielfalt des Materials und der Markiertvariablen für konkurrierende Strukturmodelle auf 191 Aufgabenblöcke und 98 Aufgabentypen reduziert. Durch den Einsatz von Strukturanalysen kristallisieren sich vier operative Fähigkeiten heraus (Bearbeitungsgeschwindigkeit, Merkfähigkeit, Einfallsreichtum und Verarbeitungskapazität). Aufgrund einer Kreuzklassifikation und einer weiteren Reduzierung der Aufgabenbereiche konnte die Existenz der inhaltsgebundenen verbalen, numerischen und figural-anschauungsgebundenen Fähigkeiten nachgewiesen werden. Die herausgefilterten operativen Fähigkeiten in Verbindung mit den inhaltsgebundenen Fähigkeiten bildeten sodann die Basis des Berliner Intelligenzstrukturmodells.

Im Gegensatz zum Faktoren-Elektizismus kann hier das Verhältnis von modellspezifischen und modellübergreifenden Faktoren analysiert werden, wie es unter anderem in der Arbeit von Cattell (1971a, 1987) praktiziert worden ist (v. Gilardi, Holling & Schmidt, 1983). Im Weiteren können subjektive Urteile und fehlende theoretische Begründungen bei der Variablenselektion und der Gewichtung der verschiedenen Fähigkeiten jedoch nicht ausgeschlossen werden. Im Zuge der Aufgabenselektion aus einem gegebenen Aufgabenpool sind darüber hinaus facettentheoretische Modellkonzepte mit eingeflossen, so dass das Ergebnis im Sinne einer möglichst umfassenden Strukturanalyse zahlreicher Aufgabentypen nicht induktiv erschlossen wurde. Das Berliner Intelligenzstrukturmodell ist somit vielmehr das Ergebnis von pragmatischen und inhaltlichen Entscheidungen. Die zwei aufgeführten Integrationsmöglichkeiten von Befunden in der Intelligenzstrukturforschung sind mit verschiedenen Vor- und Nachteilen behaftet. Eine ergänzende Betrachtung der verschiedenen Integrationsmöglichkeiten erscheint hier sinnvoll zu sein (Brocke & Beauducel, 2001).

Im Rahmen der theoretischen Strukturierung der Befunde und Modelle in der Intelligenzstrukturforschung ist zunächst das Evolutionsmodell der Strukturforschung zu nennen.

1. In dem von Sternberg und Powell (1982) entwickelten Evolutionsmodell wurden verschiedene Modelle zusammengeführt. Auf der ersten Stufe der Modellentwicklung wird eine monistische Sichtweise der Intelligenz einer pluralistischen Sichtweise gegenübergestellt. Die monistische Sichtweise geht von der Annahme aus, dass es einen einheitlichen, allgemeinen Intelligenzfaktor gibt. Die pluralistische Sichtweise hingegen begreift die Intelligenz als ein Bündel unabhängiger Bonds (Gewohnheiten, gelernte Assoziationen, etc.). Es wird jedoch nicht ausgeschlossen, dass verschiedenartige Verknüpfungen der Bonds einen gemeinsamen Intelligenzfaktor erklären können. Auf der zweiten Ebene wird zwischen einer hierarchischen Sicht und einer Überlappungssicht unterschieden. Die hierarchische Sicht wird hier eher der monistischen Sichtweise zugeordnet. Mit anderen Worten werden die vielfältigen Fähigkeiten der pluralistischen Sichtweise gebündelt und in übergeordnete Fähigkeiten eingebunden. In Cattells (1971a, 1987) Strukturtheorie (siehe Kapitel 4.4.4) findet man diese Sichtweise wieder. Die Überlappungssicht negiert einen generellen übergeordneten Intelligenzfaktor. Sie erklärt die Korrelationen zwischen den verschiedenen Fähigkeiten durch direkte Beziehungen, im Sinne von funktionalen, strukturellen und kausalen Überlappungen. Auf der dritten Ebene

wird die hierarchische Sichtweise mit der Überlappungssichtweise verkoppelt. Anders ausgedrückt können die Fähigkeiten miteinander korrelieren bzw. sich gegenseitig überlappen und gleichzeitig können sie hierarchisch aufgebaut sein (Brocke & Beauducel, 2001). In diesem Zusammenhang wird die überlappende Materialfacette (numerisch, verbal, figural) mit der hierarchisch aufgebauten Aufgabenfacette kombiniert (*Beispiel: Radex Modell*) (Guttman & Levy, 1991).

2. Die zweite mögliche theoretische Strukturierung ist das hierarchische Protomodell (HPI). Dieses Modell ist von besonderer Bedeutung, da es sich unter Berücksichtigung kleiner Variationen in den Arbeiten von Cattell (1971a, 1987) und dem Intelligenz-Struktur-Test 2000 R wiederfinden lässt, der im empirischen Teil der Dissertation die Grundlage für die Untersuchung darstellt (Amthauer et al., 2001; Liepmann et al., 2007).

Zwei grundsätzliche Fragen bezüglich der Konvergenzen zwischen den Strukturmodellen lassen sich hier unterscheiden (Brocke & Beauducel, 2001, S.34):

1. Die Frage nach dem Bedeutungsbereich bzw. dem Kontent-Bereich von Intelligenz (Kontent-Frage) und
2. Die Frage nach der Beziehung (Relation) der Teilkomponenten von Intelligenz, also nach der Struktur umfassender Intelligenzaspekte (Strukturfrage).

Hinsichtlich der Kontent- und Strukturfrage lassen sich in der Intelligenzstrukturforschung vielfältige Konvergenzen finden. Innerhalb der Kontent-Frage manifestieren sich fünf bis sieben modellübergreifende Intelligenzkomponenten, die mit Einschränkungen in den Modellen von Thurstone (1938), Cattell (1971a, 1987) und Jäger (1982, 1984) vorkommen. Die Intelligenzdimensionen umfassen dabei das schlussfolgernde Denken, verbale, numerische und räumlich-figurale Fähigkeiten, Ideenflüssigkeit bzw. Kreativität und mit Einschränkungen, Merkfähigkeit und wahrnehmungsbezogene Fähigkeiten. Die einzelnen Modelle werden im Weiteren durch zusätzliche Faktoren spezifiziert und ergänzt. Die Bezugnahme auf die genannten Primärfaktoren hat den Vorteil, dass ausschließlich modellübergreifende Faktoren einbezogen bzw. fokussiert werden, die in den bedeutendsten Intelligenzstrukturmodellen bereits etabliert und nachgewiesen sind. Hinsichtlich der Strukturfrage wird davon ausgegangen, dass jegliches intelligentes Verhalten- im Sinne des Konstrukt-Verständnisses- von mehreren Fähigkeiten gleichzeitig beeinflusst wird.

Das fundamentale Prinzip der Multi- Trait- Determination von intelligentem Verhalten bildet die Basis für die Annahme von Hierarchiestufen. Mit anderen Worten wird eine Fähigkeit, die mit anderen Fähigkeiten zusammenhängt, auf einer höheren Generalitäts- bzw. Hierarchieebene angeordnet, sodass die Intelligenz auf Basis von hierarchischen Modellen erschlossen werden kann (Brocke & Beauducel, 2001; Amthauer et al., 2001). Das oben angedeutete Radex- Modell findet sich lediglich in drei Strukturmodellen wieder, wobei das Berliner Intelligenzstrukturmodell eine herausragende Rolle einnimmt (Jäger, 1982, 1984). Das Zentrum der hierarchischen Modelle bilden somit die genannten inhaltlichen und formalen Konvergenzelemente bzw. die Primärfaktoren. Des Weiteren werden grundsätzlich zwei Hierarchie- bzw. Generalitätsebenen angenommen, die in ihrer Ausgestaltung jedoch offen konzipiert sind. Eine derartige Spezifikation auf der Ebene über den Primärfaktoren stellt beispielsweise die von Cattell (1971a, 1987) postulierte fluide und kristallisierte Intelligenz dar, die auch im I-S-T 2000 R berücksichtigt wird. Diese Unterscheidung ist in der Arbeit von großer Bedeutung und wird daher im Rahmen des theoretischen Modells von Cattell hervorgehoben und in Kapitel 6 vertieft.

#### **4.4 Ausgewählte Strukturtheorien der Intelligenz**

Kann die Intelligenz als eine einheitliche Fähigkeit aufgefasst werden oder ist sie vielmehr das Produkt aus einem Bündel von verschiedenen, unabhängigen Einzelfähigkeiten? Die Beantwortung dieser essenziellen Frage ist Gegenstand der Intelligenzforschung und wird im Rahmen der Strukturtheorien erörtert. In der dargestellten Entwicklungsgeschichte kann in diesem Zusammenhang insbesondere das Stufenleitermodell von Binet und Simon (1905) hervorgehoben werden. Binets Auffassung zufolge können die konstituierenden Einzelfähigkeiten der Intelligenz zu einem Globalindex zusammengefasst werden. Das heißt, das Ergebnis einer Testleistung kann durch einen Kennwert (z.B. IQ) ausgedrückt werden.

Die Wichtigkeit der Strukturtheorien in der Intelligenzforschung äußert sich einerseits darin, dass sie die Grundlage für den Bedeutungsgehalt der akademischen Intelligenz bilden und andererseits in jeglichen verwendeten Intelligenztests zum Tragen kommen. Das entscheidende Instrumentarium der Strukturtheorien ist die Faktorenanalyse. Die Faktorenanalyse dient dazu, aus empirischen Beobachtungen verschiedener manifester

Variablen auf zugrunde liegenden latenten Variablen („Faktoren“) zu schließen. Der Vorteil bzw. Mehrwert der Faktorenanalyse liegt darin begründet, die Vielzahl korrelierter Variablen auf wenige dahinter stehende Faktoren zu reduzieren. In erster Linie geht es um die Gewinnung von übersichtlichen, aussagekräftigen und gut interpretierbaren Faktorwerten.

Im Folgenden wird die Zwei-Faktoren-Theorie von Spearman (1904) dem Primärfaktorenmodell von Thurstone (1938) gegenübergestellt. Danach werden zwei Modelle dargestellt, welche die unterschiedlichen Ansätze vereinen.

#### **4.4.1 Die Zwei-Faktoren-Theorie von Spearman**

In Anlehnung an Galton (siehe Kapitel 4.1) übernahm Charles Spearman (1904) den Begriff der Intelligenz und die statistischen Messverfahren, insbesondere die Faktorenanalyse. Im Hinblick auf die Struktur der Intelligenz entwickelte Spearman daraufhin das Generalfaktor-Modell bzw. die Zwei-Faktoren-Theorie. Obgleich die Intelligenz als ein Bündel zahlreicher Einzelfähigkeiten gesehen wurde, setzt die Theorie einen allgemeinen *g*-Faktor (Generalfaktor) der Intelligenz voraus, auf den alle Intelligenzleistungen zurückgeführt werden können.

Mit der Formulierung der Zwei-Faktoren-Theorie wird zugleich unterstellt, dass jede Leistung genau genommen von zwei Faktoren abhängt. Der eine Faktor wird als allgemeine Intelligenz (entspricht seinem *g*-Faktor) bezeichnet und kommt in allen intellektuellen Leistungen zur Geltung. Der andere Faktor ist spezifischer Art und somit an den jeweiligen Test gekoppelt. Anders formuliert wird bei der Verwendung von Intelligenztests ein Großteil der Testleistungsvarianz durch den *g*-Faktor erklärt. Zur Erklärung der restlichen Varianz wird für den jeweiligen Test der spezifische Faktor herangezogen. Die Korrelation zwischen verschiedenen Leistungsmaßen kann allerdings nur durch denjenigen Faktor erklärt werden, der beiden Tests gemeinsam ist, nämlich dem *g*-Faktor. Dieser Argumentation folgend müsste eine Interkorrelationsmatrix von Intelligenztests stets positive Korrelationen zeigen, die durch den Generalfaktor erklärt werden können (Mehlhorn & Mehlhorn, 1981).

Im Bezug auf den Bedeutungsgehalt des Generalfaktors betrachtete Spearman zunächst nur allgemeine hypothetische Konstrukte. Später postulierte er, dass der Generalfaktor in der Qualität des gesamten menschlichen Nervensystems zum Ausdruck kommt. Eine Definition von Qualität im engeren Sinne wurde jedoch nicht gegeben. Im Weiteren

vertrat er die Auffassung, dass die Analyse von kognitiven Prozessen Aufschluss über die Bedeutung geben könne. Nach seiner Ansicht sind die bei der Bearbeitung von unterschiedlichen Tests übereinstimmenden kognitiven Prozesse dafür verantwortlich, dass diese Tests mit dem g-Faktor positiv korrelieren (Spearman, 1930, zitiert nach Putz-Osterloh, 1981). Die Untersuchung dieser Prozesse ist auch aktuell ein Bereich der Intelligenzforschung.

Zusammenfassend ist zu konstatieren, dass die Zwei-Faktoren-Theorie angesichts ihrer Einfachheit und Verständlichkeit einen hohen Stellenwert im Rahmen der Intelligenzstrukturforschung besitzt. Die Annahme eines g-Faktors als Kennwert der allgemeinen Intelligenz ist heute noch gängig und wird häufig als Gesamtergebnis in verschiedenen Intelligenztests zusätzlich ausgewiesen.

#### **4.4.2 Thurstones Primärfaktoren-Modell**

In Spearmans Zwei-Faktoren-Theorie werden alle Leistungen auf einen allgemeinen Intelligenzfaktor zurückgeführt, und neben diesem wird ein spezifischer Faktor ausgewiesen, der dem jeweiligen Test inhärent ist. Demgegenüber steht Thurstones Ansatz, der den vermeintlichen g-Faktor negiert und stattdessen sieben Primärfaktoren zur Erklärung der geistigen Fähigkeiten postuliert.

Thurstone (1938) entwickelte also die von Spearman begründete faktorenanalytische Methode und dessen Theorie weiter. Der Ausgangspunkt seiner Forschungsanstrengungen ist die unerklärte bzw. verbleibende Restvarianz. Für Thurstone war es aus allgemeinen-theoretischen und aus praktisch-diagnostischen Gründen nicht akzeptabel, die gemeinsame unerklärte Restvarianz, welche auch nach der Extraktion des g-Faktors und unter Berücksichtigung der spezifischen Faktoren bestehen blieb, zu vernachlässigen. Die Schlussfolgerung von Thurstone war, dass neben diesen Faktoren weitere Gruppenfaktoren (*primary mental abilities*) existieren müssen, die in unterschiedlichen Gewichtungverhältnissen in die Bearbeitung von Denkaufgaben eingehen (Jäger, 1967, 1970). Demzufolge konstatiert Thurstone (1938): „Performance of a task requires a certain number of fundamental or primary abilities“ (S. 2). Mit Hilfe multivariater Analysemethoden war es Thurstone sodann möglich, aus der Interkorrelation unterschiedlicher Tests, die verschiedenen Intelligenzleistungen auf relativ unabhängige Faktoren zurückzuführen (Thurstone, 1938).



Die folgenden zwei Kriterien bilden das Fundament für sein Modell (Amelang & Bartussek, 2001):

1. Die Verfahren (Tests) dürfen nicht nur die spezifischen Faktoren im Sinne von Spearman (1904) erfassen.
2. Die konkrete Testaufgabe soll nicht von allen Primärfähigkeiten gleichzeitig determiniert werden.

Die folgenden sieben Faktoren bzw. Faktorengruppen gelten in Thurstones Modell als gesichert und sind in die diagnostische Praxis eingegangen.

*Table 8: Die Primärfaktoren von Thurstone  
(Thurstone & Thurstone, 1941, zitiert nach Amelang & Bartussek, 2001, 208f)*

Primärfaktor	Beschreibung
V: Verbal Comprehension- -Sprachverständnis	Kenntnis von Wörtern und ihrer Bedeutung sowie angemessene Verwendung im Gespräch. Ladungen in Tests zu verbalen Analogien, Textverständnis, Rechtschreibung, richtige Reihung vertauschter Wörter und Sätze
W: Verbal fluency- -Wortflüssigkeit	Rasches Produzieren von Wörtern, die bestimmten strukturellen oder symbolischen Erfordernissen entsprechen. Tests: Anagramme, Reime, Benennungen, Wörter mit vorgegebenen Anfangsbuchstaben oder 4 Buchstaben.
N: Number - Rechenfertigkeit	Geschwindigkeit und Präzision bei einfachen arithmetischen Aufgaben. Tests: Addition, Subtraktion, Multiplikation.
S. Space - Raumvorstellung	Bewältigung von Aufgaben, die räumliches Vorstellen und Orientieren sowie das Erkennen von Objekten unter anderem Bezugswinkel erfordern (der Faktor untergliedert sich entsprechend dieser Beschreibung häufig in zwei oder drei speziellere S-Faktoren). Tests: Verfolgen von mechanischen Bewegungen, Vergleich von Würfeln aus verschiedener Perspektive (...)
M: Memory -Mechanisches Gedächtnis	Behalten paarweise gelernter Assoziationen. Tests: Wort – Zahl – oder Bild – Figuren – Paare und dergl.
P: Perceptual Speed Wahnehmungsgeschwindigkeit	Geschwindigkeit beim Vergleich oder der Identifikation visueller Konfigurationen. Tests: Anstreichen bestimmter Symbole, Erkennen von Gleichheiten bzw. Unterschieden
R: Reasoning - Schlussfolgern	Schlussfolgerndes Denken im Sinne des Auffindens einer allgemeinen Regel in einer vorgegebenen Abfolge von Zahlen und Symbolen und Anwendung derselben bei der Vorhersage des nächstfolgenden Elements (...) Tests: Reihen fortsetzen.

Die Existenz dieser sogenannten Primärfähigkeiten unterstreicht die Annahme, dass die Intelligenz durch ein Bündel von Einzelfähigkeiten gekennzeichnet ist und diese Faktoren in ihrer Summe die Grundvoraussetzung der menschlichen Leistungen darstellen (siehe auch Kapitel 6). Thurstones Auffassung über den Bedeutungsgehalt der Primärfähigkeiten war jedoch nicht einheitlich. In diesem Zusammenhang stellt Thurstone (1940, zitiert nach Jäger, 1967) einerseits die Behauptung auf: „ (...) it is a distortion to assume that they must be elemental (...) we know little about the determiners of human talent (...) we should not impose on our thinking an unnecessary rigid causal frame (...)“ (S.78). Gleichzeitig charakterisiert er die Faktoren als „ media for the expression of intellect (...) people differ markedly in the effectiveness with which they can express themselves in these different media“ (Thurstone, 1948, zitiert nach Jenkins & Paterson, 1961, S. 619). Trotz dieser zum Teil inkonsistenten Beschreibungen stellt er die Faktoren vielfach als funktionale Einheiten dar, die ähnliche Leistungen aufweisen und wahrscheinlich genetisch verankert sind. Ein weiteres Merkmal ist, dass die Faktoren gleichwertig nebeneinander stehen. Die Leistung der jeweiligen Person setzt sich daraufhin summativ aus den entsprechenden Primärfaktoren zusammen (Jäger, 1967).

Die Kritik an dem Faktorenmodell von Thurstone wird dahingehend geübt, dass die dem Modell zugrunde liegende Versuchspersonenstichprobe nicht repräsentativ war und nur College-Studenten mit einbezogen wurden. Wegen der geringen Leistungsvariabilität und den damit einhergehenden niedrigen Korrelationskoeffizienten zwischen den Messwertreihen wird die Chance, einen Generalfaktor zu finden, zusätzlich erschwert. Die Anwendung einer schief winkligen rotierten Faktorenlösung, um klare Einfachstrukturen zwischen den Primärfaktoren zu erhalten, führte zu Korrelationen innerhalb dieser Faktoren. Angesichts der positiven Korrelationen zwischen den Primärfaktoren in Höhe von ca.  $r = 0.35$  wird die Existenz des g-Faktors jedoch als bestätigt angesehen (Eyseneck, 1979). In diesem Zusammenhang räumt auch Thurstone ein, dass es einen g-Faktor geben könne und bezeichnet ihn als Faktor zweiter Ordnung. Weitere Untersuchungen bezüglich der Beziehungen zwischen den Primärfaktoren und dem g-Faktor hat er nicht durchgeführt (Thurstone, 1945, zitiert nach Jäger, 1967). Die scheinbare Unvereinbarkeit der Generalfaktor-Modelle und der Gruppenfaktoren-Modelle könnte vor diesem Hintergrund überwunden sein. Brody und Brody (1976) konstatieren: “Thurstone`s studies had not shown that g did not exist but rather that g would be distributed and divided into components that were themselves positively related“ (S.17).

### 4.4.3 Das Berliner Intelligenzstrukturmodell (BIS)

Das grundlegende Ziel des BIS ist eine Integration der verschiedenen Strukturtheorien. Das BIS stellt sich hauptsächlich als eine Zusammenführung der Kategorisierungssysteme von Spearman und Thurstone dar, welches durch spezielle empirische Erhebungen bestätigt wurde. Die Intelligenzstruktur bezeichnet Jäger (1970) als ein „System der theoretischen Begriffe, mit dem der im Menschen liegende Anteil der Bedingungen intelligenten Verhaltens erfasst wird. Bei unseren Untersuchungen ging es (...) um die Systemkomponenten, welche man als Fähigkeiten oder Intelligenzdispositionen bezeichnet, und um deren Ordnungsgefüge, nicht um die funktionalen Zusammenhänge dieser mit den weiteren, in der Persönlichkeit und in der Umwelt liegenden Bedingungen intelligenter Verhaltensweisen“ (S.172).

Mittels umfangreicher Analysen aus ca. 280 Variablen verschiedener Strukturmodelle entdeckte Jäger in seinem ersten Entwurf sechs Gruppen von Hauptfaktoren, die hierarchisch aufgebaut sind. Hierzu zählen: anschauungsgebundenes Denken, Einfallsreichtum und Produktivität, Konzentrationskraft und Tempomotivation, Verarbeitungskapazität, formallogisches Denken und Urteilsfähigkeit, zahlengebundenes Denken und sprachgebundenes Denken (Jäger, 1970).

Die folgenden Annahmen sind konstitutive Elemente des Berliner Modells (Jäger, Süß & Beauducel, 1997, zitiert nach Nettelnstroth, 2003, S.78f):

- An jeder Intelligenzleistung sind (neben anderen Bedingungen) alle intellektuellen Fähigkeiten beteiligt, allerdings mit deutlich unterschiedlichen Gewichten. Die Varianz jeder Leistung lässt sich in entsprechende Komponenten zerlegen.
- Intelligenzleistungen und Fähigkeitskonstrukte lassen sich unter verschiedenen - hier Modalitäten genannten- Aspekten klassifizieren. Vorerst wurde nur eine bimodale Klassifikation - Operationen/Inhalte - spezifiziert.
- Fähigkeitskonstrukte sind hierarchisch strukturiert, d.h. sie lassen sich unterschiedlichen Generalitätsebenen zuordnen.

Darüber hinaus ist das Konstrukt-Verständnis der Intelligenz bestimmend für das BIS. Die Basis für das BIS bildete die Analyse eines Aufgabenpools (ca. 2000 Item- Typen), der dem Stand der Forschung entsprach. Daran anknüpfend, wurde der Fundus an Aufgabenarten auf 191 Aufgabenblöcke reduziert, die 98 Aufgabentypen zugeordnet werden können. Die Verschiedenartigkeit der Aufgaben und die Repräsentation der konkurrierenden Strukturmodelle wurden hierbei stets berücksichtigt. In einer empirischen

Erstuntersuchung mit 545 Abiturienten im Alter von 16-21 Jahren wurden die restlichen 191 Leistungsvariablen geeicht bzw. falsifiziert. In einem weiteren Schritt wurden anhand von exploratorischen Faktoren- und Clusteranalysen vier operative Leistungsklassen herauskristallisiert: Einfallsreichtum, Verarbeitungskapazität, Gedächtnis und Bearbeitungsgeschwindigkeit (Jäger, 1982, 1984). Im Rahmen dieser strukturellen Analyse konnten jedoch nicht die ausgewiesenen inhaltspezifischen Faktoren( sprach-, zahlen-, und anschauungsgebundenes Denken) nachgewiesen werden. Die Problematik lag wahrscheinlich darin begründet, dass die Stichprobe zu homogen war und dadurch die inhaltspezifischen Faktoren überdeckt wurden (Amelang & Bartussek, 2001). Darauf aufbauend und gemäß der zweiten Annahme(s.o.) klassifizierte Jäger die Leistungsvariablen nach der Operations- und Inhaltsklasse (siehe Abbildung 6).

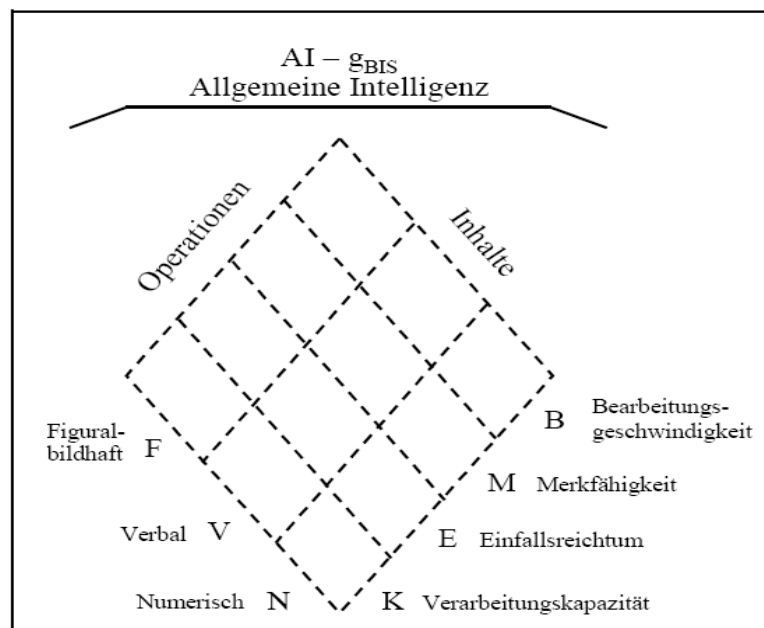


Abbildung 6: Das Berliner Intelligenzstrukturmodell (BIS), nach Jäger, 1982

Jäger (1984) fasste die operativen Faktoren als auch die inhaltspezifischen Faktoren zu homogenen Bündeln zusammen. Mithilfe einer kreuzklassifikatorischen Faktorisierung konnten sodann die inhaltspezifischen- und die operativen Faktoren empirisch und simultan bestätigt werden.

Die operativen und die inhaltsgebundenen Fähigkeiten werden wie folgt beschrieben:

*Tabelle 9: Fähigkeitsbündel im Berliner Intelligenzstrukturmodell  
(Jäger, Süß & Beauducel, 1997, zitiert nach Nettelnstroth, 2003, S.83f)*

Fähigkeit	Beschreibung
<b>Operative Fähigkeiten</b>	
K: Verarbeitungskapazität	Verarbeitung komplexer Informationen bei Aufgaben, die nicht auf Anhieb zu lösen sind, sondern Heranziehen, vielfältiges Beziehung stiften, formallogisch exaktes Denken und sachgerechtes Beurteilen von Informationen erfordern.
E: Einfallsreichtum	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 (G) 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.
<b>Inhaltsgebundene Fähigkeiten</b>	
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: Anschauungsgebundenes, figuralbildhaftes Denken	Einheitsstiftendes Merkmal scheint hier die Eigenart des Aufgabenmaterials zu sein, dessen Bearbeitung figural-bildhaftes und/oder räumliches Vorstellen erfordert.

Die dargestellte bi-modale Matrix lässt den Schluss zu, dass die operativen und die inhaltspezifischen Faktoren keine Primärfaktoren im Sinne von Thurstone sind, sondern

vielmehr multifaktoriell bedingte Leistungen. Hinsichtlich des relativ geringen Allgemeinheitsgrades können die Faktoren daher als Sekundärfaktoren aufgefasst werden. Aufgrund der Tatsache, dass zwischen den vier operativen und den drei inhaltsspezifischen Faktoren durchweg positive Korrelationen bestehen, erhält der g-Faktor seine Berechtigung. Der Unterschied zwischen der Allgemeinen Intelligenz und den Sekundärfaktoren besteht hier lediglich im Differenzierungsgrad (Amelang & Bartussek, 2001). Eine Bestätigung des Modells kann nach Schmidt (1984) nicht anhand des Kriteriums der Einfachstruktur erfolgen. Die Methode der Einfachstruktur ist so konzipiert, dass eine Variable möglichst nur auf einem Faktor hoch lädt. Die Bi-Modalität der Variablen verlangt jedoch, dass eine Variable sowohl auf einem inhaltsspezifischen als auch auf einem operativen Faktor hoch lädt. Die Lösung stellt hier die konfirmatorische Faktorenanalyse dar.

Es bleibt festzuhalten, dass das Berliner Intelligenzstrukturmodell einen bedeutenden Stellenwert in der aktuellen Eignungsdiagnostik einnimmt. Einerseits bietet die Aufgaben-Integration eine gute Möglichkeit zur Systematisierung der verschiedenen Strukturmodelle, andererseits ist die Integration der operativen und inhaltlichen Faktoren zur Beschreibung einer Variablen von großer Relevanz.

#### **4.4.4 Das Modell der fluiden und kristallisierten Intelligenz von Cattell**

Das Modell der fluiden und kristallisierten Intelligenz von Cattell (1971a, 1987) stellt ebenfalls eine Integration der Zwei-Faktoren-Theorie von Spearman und des Primärfaktoren-Modells von Thurstone dar (Brocke, 2000). Die methodische Basis des hierarchischen Strukturmodells beruht wiederum auf sukzessiven Faktorenanalysen. Im Unterschied zu der englischen Schule (z.B. Spearman) werden im Rahmen des hierarchischen Strukturmodells (amerikanische Schule) die Faktoren mit höherem Allgemeinheitsgrad durch aufwärtsgerichtete Faktorenanalysen erschlossen.

Auch in diesem Modell wird ein g-Faktor angenommen, der jedoch induktiv bestimmt wird. Die folgende Abbildung 7 illustriert das Modell:

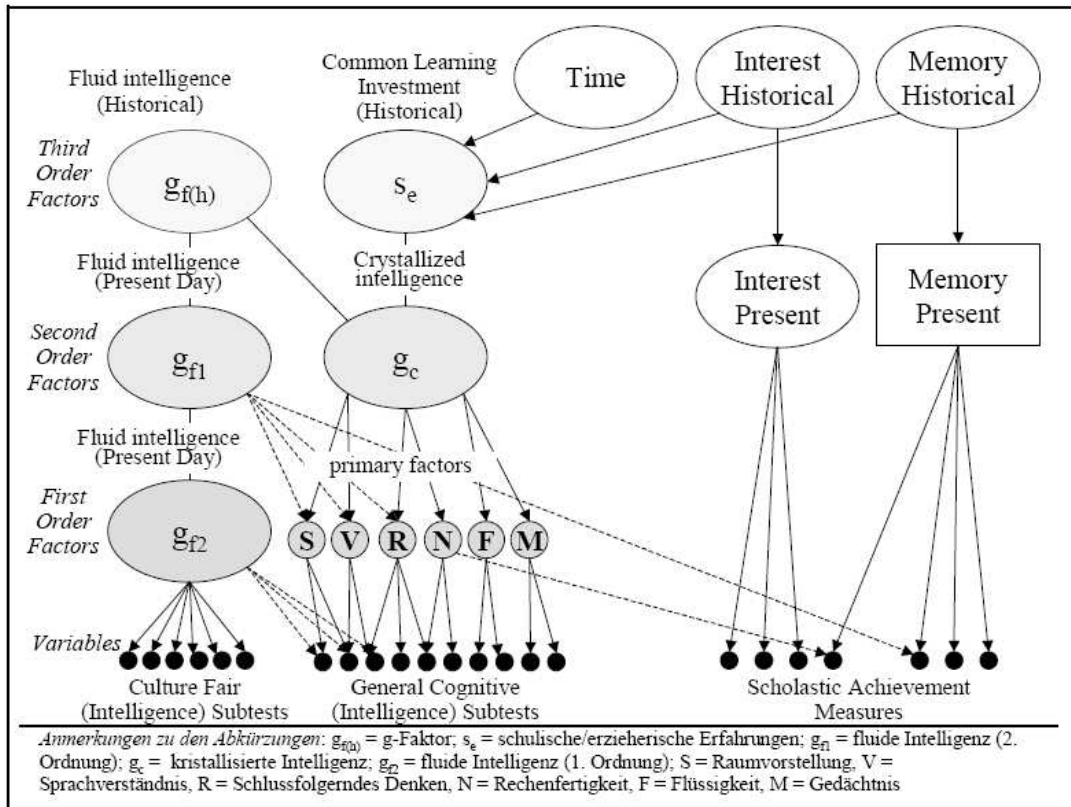


Abbildung 7: Das Modell der fluiden und kristallisierten Intelligenz von Cattell

Die methodische Vorgehensweise sieht in diesem Zusammenhang folgendermaßen aus: Aus den Tests werden durch die Korrelationen der Testitems die Faktoren erster Ordnung extrahiert (*first order factors*). Das Ergebnis zeigt sechs Primärfaktoren, die den von Thurstone (1938) ähneln und dem Faktor der fluiden Intelligenz ( $g_{f2}$ ). Dieser Faktor lädt am höchsten auf Erkenntnisleistungen, in denen individuelle Unterschiede der Lernerfahrungen weniger entscheidend sind. Cattell postuliert diesbezüglich die Verwendung von so genannten kulturfreien Tests.

Die Faktoren zweiter Ordnung (*second order factors*) umfassen wiederum den nahezu unveränderten Faktor der fluiden Intelligenz ( $g_{f1}$ ) und den Faktor der kristallisierten Intelligenz ( $g_c$ ), der sich aus der Korrelation der restlichen Primärfaktoren ergibt. Es sei an dieser Stelle angemerkt, dass Hakstian und Cattell (1978) in weiterführenden Analysen zusätzliche Faktoren zweiter Ordnung ermittelten: Die Visualisierungskapazität ( $g_v$ ), die in allen Primärfaktoren enthalten ist, die eine Form der visuellen Organisation

benötigen), die generelle Wahrnehmungsgeschwindigkeit (gps), die allgemeine Gedächtniskapazität (gm) und die *General Retrieval Capacity* (gr). Die gefundenen Faktoren verlangen jedoch noch weitere Replikationen.

Der Faktor dritter Ordnung (*third order factors*; hier gf(h)) wird durch die Korrelation von gf1 und gc eruiert. Dieser Faktor kann aufgrund des ihm inhärenten Allgemeingrades mit dem g-Faktor von Spearman gleichgesetzt werden. Die Faktoren gehen hier allerdings zusammen in den allgemeinen Faktor mit ein. Die Erklärung für die Extraktion eines allgemeinen Intelligenzfaktors wird in der bestehenden Interkorrelation der beiden Faktoren gf und gc gesehen. „This correlation, which hovers around  $r = 0.4-0.5$  (...) is that the acquisition of the crystallized ability skills (...) depends partly on the level of insightful „fluid“ ability and partly on hours spent in school, etc“ (Cattell, 1987, S.116). Die hohe Korrelation zwischen gf und gc als auch die starken Streuungen der Faktorladungen auf gf und gc beinhalten jedoch auch gleichzeitig die Kritik an der von Cattell (1971a, 1987) postulierten Trennung von fluider und kristallisierter Intelligenz (eine weiterführende Diskussion findet in Kapitel 6 und 11.4 statt).

Von besonderer Bedeutung im Rahmen des hierarchischen Strukturmodells von Cattell (1971a, 1987) ist die Differenzierung zwischen fluider und kristallisierter Intelligenz auf der zweiten Ordnungsebene. Im Folgenden wird lediglich eine kurze Einführung in die komplexe Thematik dieser Gegenüberstellung geben. In Kapitel 6 findet dann eine tiefgreifende Auseinandersetzung statt, welche die verschiedensten Einflussfaktoren und Bedeutungen der beiden Faktoren fokussiert und deren Interaktion in den Vordergrund der Betrachtung rückt.

Die fluide Intelligenz ist durch relativ kulturfreie (*culture fair* oder *culture reduced*) Faktoren gekennzeichnet. Die Auswahl der Test Items richtet sich danach, ob sie (a) gleichmäßig von allen gelernt sein sollte oder (b) gleichmäßig neuartig für alle sein sollte (Mehlhorn & Mehlhorn, 1981). Die fluide Intelligenz zeigt sich insbesondere in den Fähigkeiten zum schlussfolgernden Denken, des Problemlösens und des Abstrahierens. Im Weiteren wird die fluide Intelligenz unter anderem über Aufgaben zur Wahrnehmung von Beziehungen in Reizmustern oder durch Aufgaben zum räumlich-visuellen Vorstellungsvermögen operationalisiert (Horn, 1988).

Im Bezug auf die Unterscheidung zwischen fluider und kristallisierter Intelligenz schreibt Cattell (1973, zitiert nach Mehlhorn & Mehlhorn, 1981): „Während gf direkt physiologisch bestimmt ist, ist gc ein Produkt von umweltspezifischen variierenden erfahrungsgemäß bestimmten Ausfaltungen von gf“ (S.51). Aus dieser Definition lässt



sich auch entnehmen, dass die fluide Intelligenz vor allem genetisch verankert ist und daher auch vererbt werden kann. Die kristallisierte Intelligenz hingegen wird durch die Akkulturationseinflüsse der jeweiligen Gesellschaft geprägt.

Die kristallisierte Intelligenz ist nach Horn (1988) mit dem Bereich des Wissens, den Erfahrungen, der Kommunikations- und Urteilsfähigkeit, Raffinesse und mit dem Verständnis von gesellschaftlichen Gepflogenheiten eng verknüpft. Die Operationalisierung der kristallisierten Intelligenz manifestiert sich in den Testitems zum verbalen Verständnis, Begriffsbewusstsein und dem logischen, numerischen und generellen Schlussfolgern. Die Bezeichnung „acculturation knowledge“ (Horn, 1988, S.658) macht deutlich, dass die kristallisierte Intelligenz sowohl durch die spezifische Kultur eines Landes bestimmt wird als auch durch das in dieser Kultur vermittelte und aufgenommene Wissen. Zwei Wissensformen sind in diesem Zusammenhang von großer Bedeutung (Baltes, 1990):

1. Deklaratives bzw. explizites Wissen
2. Prozedurales bzw. implizites Wissen

In der Investmenttheorie von Cattell (1971a, 1987) wird die Beziehung zwischen fluider und kristallisierter Intelligenz (gf und gc) erläutert. In der Theorie werden die verschiedenartigen Einflüsse auf die Intelligenz analysiert und damit das Verhältnis zwischen biologisch vorgegebener Intelligenzausstattung (gf) und kulturell bedingtem Wissen und Verhalten (gc) thematisiert. In diesem Kontext postuliert Cattell, dass sich gc im Laufe der ontogenetischen Entwicklung auf Grundlage der investierten Basiskompetenzen (gf) zu spezifischen Fähigkeiten herauskristallisiert, die u.a. den Erwerb des Wissens steuern.

Die Verknüpfung von Intelligenz und Wissen ist in dieser Arbeit von großer Relevanz, da beide Konstrukte als extrafunktionale Qualifikationen gesehen werden, die den schulischen oder beruflichen Erfolg maßgeblich beeinflussen (z.B. Schmidt & Hunter, 1998; Süß, 2001). Das Konstrukt Wissen wird im nächsten Kapitel gesondert betrachtet in einen sinnvollen Zusammenhang mit der (kristallisierten) Intelligenz gebracht. Darauf aufbauend wird die Bedeutung der kristallisierten Intelligenz bzw. des Wissens für die Eignungsdiagnostik hervorgehoben (vgl. Kapitel 2, 3, 5 und 6).

Zusammenfassend ist zu konstatieren, dass das hierarchische Strukturmodell von Cattell (1971a, 1987) gut in einigen Strukturmodellen wiedergefunden werden kann. Darüber hinaus wurde das Modell auch in zahlreichen Studien weitestgehend empirisch

bestätigt, so dass es auch in einigen namhaften Intelligenztests berücksichtigt wird. Auf diesen Aspekt wird im empirischen Teil näher eingegangen (Amelang & Bartussek, 2001).

#### **4.4.5 Abschließende Betrachtung und Bewertung der Strukturtheorien**

Die eingangs gestellte Frage lässt sich vor dem Hintergrund der dargestellten Strukturtheorien nur schwer beantworten und bedarf weiterer Forschungsanstrengungen. Die Schwierigkeit in diesem Zusammenhang ergibt sich vor allem durch die angewendete Methode, die das Ergebnis maßgeblich determiniert. Gerade die verschiedenen faktoranalytischen Vorgehensweisen (orthogonal oder schiefwinklige Rotationen) oder auch das Kriterium der Einfachstruktur führen zu unterschiedlichen Ergebnissen. Darüber hinaus sind auch die vom Autor vorab getroffenen Setzungen und Annahmen von entscheidender Bedeutung für die Resultate. Werden beispielsweise bestimmte Merkmalsbereiche nicht mit einbezogen, kann dieser Umstand die Suche nach einem Generalfaktor erschweren bzw. unmöglich machen. Ein weiterer Faktor hinsichtlich einer zweckmäßigen Interpretation der Intelligenzstruktur ist die Bildung einer repräsentativen Stichprobe (Amelang & Bartussek, 2001). Die von Thurstone vorgenommene Stichprobenauswahl offenbart beispielsweise diese Problematik. Die ausschließliche Bezugnahme auf College-Studenten hat die Leistungsvariabilität eingeschränkt und damit auch die Chance einen G-Faktor zu finden. Darüber hinaus ist auch der Anspruch der Allgemeingültigkeit von Intelligenzstrukturen in Frage gestellt. In diesem Kontext stellt die Berücksichtigung der kognitiven Prozesse bei der Bearbeitung von Intelligenzaufgaben eine zusätzliche Herausforderung dar. Aufgrund der andersartigen Prozesse einer Vielzahl von Probanden können auch abweichende Intelligenzstrukturen auftreten. Außerdem können die Persönlichkeitsmerkmale der Probanden (z.B. Alter, Geschlecht, etc.) zu differenten Strukturen führen (a.a.O.).

Resümierend ist es weder Thurstone noch anderen Wissenschaftlern gelungen, den g-Faktor anhand von empirischen Analysen grundlegend zu negieren. Die positiven Korrelationen der theorieübergreifenden Primärfaktoren lassen vielmehr den Schluss zu, dass ein Generalfaktor der Intelligenz durchaus nachgewiesen werden kann. Im Weiteren kann sich der Ausweis eines Generalfaktors oder IQ-Wertes in der diagnostischen Praxis auch als sinnvoll erweisen. Er stellt einen groben Anhaltspunkt über das allgemeine Leistungspotential eines Menschen dar. Mit anderen Worten gibt er eine

zuverlässige Sicherheit für Entscheidungen, ob ein Schüler beispielsweise nur „faul“ ist oder ob er die intellektuellen Voraussetzungen nicht erfüllt. Für eine differenzierte Aussage ist die Prüfung bzw. Untersuchung der verschiedenen Subbereiche der Intelligenz natürlich unerlässlich. Die Strukturanalyse unterhalb des Generalfaktors bildet somit das Fundament für eine sinnvolle Interpretation der Intelligenz. Eine geeignete Weiterentwicklung neben dem Ausweis eines einzelnen IQ-Wertes kann sodann die Illustration von Fähigkeitsprofilen darstellen.

## **5 Wissen- Konzepte, Befunde und Bedeutung**

Die Wissenspsychologie steht in der Forschungstradition der Allgemeinen Psychologie. Im engeren Sinne kann sie als Kerngebiet der Kognitionspsychologie angesehen werden (Klix & Spada, 1998).

Ähnlich den Ausführungen über die Intelligenz kann das Wissen ebenfalls als kognitive Leistung verstanden werden. Die kognitive Leistung ergibt sich im Ergebnis aus dem Funktionskreislauf von Prozessen der Wahrnehmung, des Gedächtnisses, des Schlussfolgerns und des Handelns. Der Wesenszug der Kognition besteht darin, Informationen zu verarbeiten. Die stetige und alltägliche Informationsverarbeitung verfolgt das Ziel, Erkenntnisse über die Umwelt und die Realität zu gewinnen. Die gewonnenen Erkenntnisse dienen dann dazu, die situativen Anforderungen in der Gegenwart und in der Zukunft erfolgreich zu bewältigen (Mack, 1999). Die kognitive Leistung zeigt sich in der „Herausbildung, Speicherung und Verwendung einer Repräsentation, oder – wenn der Ausdruck nicht im Sinne einer naiv- realistischen „Bildertheorie“ missverstanden wird- eines inneren Abbildes der uns umgebenden Welt“ (Scheerer, 1978, S.2). Eine Schnittstelle zum Verständnis der Intelligenz besteht womöglich in der folgenden Definition: Intelligenz ist der „Leistungsgrad psychischer Funktionen bei ihrem Zusammenwirken in der Bewältigung neuer Situationen“(Rohracher, 1965, S.352). Auf Basis des gegebenen Leistungsvermögens psychischer Funktionen manifestiert sich das Wissen in dem Verständnis von Zusammenhängen, die aus der Umwelt generiert werden. Das Verständnis kann darüber hinaus unmittelbar erfahren werden oder es kann durch Medien vermittelt werden. Eine Erweiterung des Wissens erfährt der Mensch durch seine geistige Auseinandersetzung mit der Umwelt, also durch das Problemlösen, das schlussfolgernde Denken, etc. (Klix & Spada, 1998). Die generierten Wissensinhalte prägen im

Übrigen das Weltbild und die Identität des Menschen. Holyoak (1991) bringt die Bedeutung des Wissens auf den Punkt: „Knowledge is power“ (S.302). In diesem Kontext stellt Wissen die Voraussetzung für zielgerichtetes Handeln und die schnelle und adäquate Bearbeitung von Aufgaben und Problemen dar. Wissen ist demnach das Erfolgsrezept für gute Leistungen (Klix & Spada, 1998).

Das folgende Kapitel soll einen groben Überblick über das Themengebiet Wissen vermitteln. Dementsprechend werden zunächst die Taxonomie der Wissensarten dargestellt und grundlegende wissenspsychologische Aspekte herausgearbeitet. Die enge Beziehung zwischen Wissenserwerb und Lernen und die Einsicht, dass das Wissen in verschiedenen Bildungseinrichtungen gezielt gefördert werden kann, führt zum Unterkapitel „Wissenserwerb als Lernprozess“.

Es wird erneut betont, dass eine starke wechselseitige Beziehung zwischen dem Wissen und der von Cattell (1971a, 1987) konzipierten kristallisierten Intelligenz besteht. Seine Studien belegen, dass Wissensskalen oder auch Skalen zur Erfassung von verbalen Fähigkeiten oft hohe Zusammenhänge mit dem Faktor kristallisierte Intelligenz aufweisen. Dabei werden die verbalen Fähigkeiten sowie das Wissen vor dem Hintergrund der Lernumgebung bzw. der Kultur interpretiert.

Wissen wird zunächst unabhängig von der Intelligenz betrachtet, und steht im Weiteren als eigenständiges Konstrukt neben der Intelligenz. In Kapitel 6 werden die beiden Konstrukte dann zusammengeführt. Darauf aufbauend soll in den nachfolgenden Kapiteln auch die Bedeutung beider Faktoren für die schulische oder berufliche Leistung sukzessiv herausgearbeitet werden.

## **5.1 Die Taxonomie der Wissensarten**

Entsprechend den Ausführungen zur Intelligenz kann das Wissen ebenfalls als Konstrukt verstanden werden, welches durch die Einbindung in verschiedene nomologische Netze mit wiederum anderen Konstrukten weitere Bedeutungskomponenten erhält und durch weiterführende Forschungsanstrengungen sukzessive erschlossen wird. Die Verwendung des Begriffes Wissen variiert demnach zwischen den wissenschaftlichen Disziplinen und wird gemäß dem eigenen Gegenstandsbereich unterschiedlich definiert und präzisiert.

Anhand der folgenden Begriffsbestimmungen soll deutlich werden, dass das Konstrukt Wissen sehr komplex ist und durch verschiedene psychologisch relevante Facetten gekennzeichnet ist. Nach Reusser (1998) wird Wissen im Zusammenspiel mit allgemeinen kognitiven Fähigkeiten kulturspezifisch und „unter dem Einfluss anderer Menschen, d.h. im Rahmen eines sozialen Unterweisungs- und Interaktionsgefüges [sowie] einer an Bedeutung reichen Kultur“ (S.155) erworben. Zum Anderen umfasst der Wissensbegriff nach Krogh und Köhne (1998) „sämtliche Kenntnisse und Fähigkeiten, die Individuen zur Lösung von Aufgaben einsetzen und welche Handlungen sowie Interpretationen u.a. von Informationen ermöglichen; Wissen beinhaltet einen Sinngebungsprozess sowie normative und emotionale Elemente und ist sowohl kontext- als auch zeitabhängig“ (S.236).

Obwohl keine einheitliche Definition angegeben werden kann, wird das Wissen gemeinhin als interne Repräsentation der Umwelt bzw. Realität aufgefasst. Die Wissensdiagnose hat den Zweck, die Menschen bezüglich ihrer unterschiedlichen Merkmale und ihres Fundus an Wissen zu differenzieren. Der Gegenstand der Diagnose ist dementsprechend das Wissen, welches in der einschlägigen wissenschaftlichen Literatur in zwei Wissenskategorien untergliedert werden kann (Kluwe, 1990):

1. Das deklarative Wissen als erster Typ „bezeichnet die Kenntnisse eines Menschen über die Realität, das heißt über Sachverhalte, Vorgänge, Personen und Objekte“ (Kluwe, 1990, S.175). Das deklarative Wissen stellt somit eine Daten- bzw. Wissensbasis dar, auf der sämtliche Informationen gespeichert sind. Solches Wissen kann zwar interindividuell verschieden sein, gilt aber als verbalisierbar. Innerhalb des deklarativen Wissens wird noch zwischen episodischem Wissen und semantischem Wissen unterschieden. Das episodische Wissen ist hierbei durch individuell erfahrene Inhalte gekennzeichnet. Semantisches Wissen drückt sich hingegen in der Formulierung „Wissen, dass...“ aus (Kluwe, 1990). Im Rahmen der Kognitionswissenschaften werden vornehmlich Netzwerkmodelle verwendet, um das Wissen greifbar zu machen. Die dahinter stehende Annahme ist, dass die Speicherung von Informationen konkreten Ordnungsprinzipien folgt (Anderson, 1983). Dementsprechend wird die Struktur des deklarativen Wissens in Form von Wissensseinheiten illustriert. In diesem Zusammenhang spricht man auch von so genannten Schemata, deren Komponenten in komplexen Lernprozessen erworben, verknüpft und sukzessiv ausgebildet werden. Eine andere Darstellung der Wissensseinheiten zeigt sich in den mentalen Modellen. Ein mentales Modell

stellt in diesem Sinne eine spezifische interne Abbildung eines komplexen Sachverhalts dar. Das vorgefertigte Modell dient dem Menschen dazu, einen komplexen und wenig transparenten Sachverhalt zu verstehen und zu bearbeiten (Kluwe, 1990).

2. Das prozedurale Wissen als zweiter Typ beinhaltet „die verfügbaren geistigen Operationen und Prozesse zur Verarbeitung von Informationen“ (Kluwe, 1990, S.176). Auf Grundlage der Datenbasis drückt sich das prozedurale Wissen folglich in Form von geistigen Prozessen oder Handlungen aus, die sich in Produktionsregeln bzw. Produktionssystemen manifestieren (z.B. das Abstrahieren, Speichern, Verknüpfen, etc.). Die geistigen Prozesse sind jedoch im Gegensatz zum deklarativen Wissen nicht verbalisierbar (Anderson, 1983).

Deklaratives und prozedurales Wissen sind zwar unterschiedliche Wissensformen, können aber durchaus miteinander verflochten werden oder interdependent aufeinander bezogen sein. Prozedurales Wissen kann demnach auch verbalisiert werden, wenn sich zum Beispiel die Person bei ihrer Handlung auf deklaratives Wissen stützt oder durch Selbstbeobachtung den Informationsgehalt in deklaratives Wissen transformiert. Hinsichtlich beider Wissenskategorien werden die Menschen jedoch unterschieden (Süß, 1996).

Eine andere begriffliche Differenzierung von Wissen ist die Gegenüberstellung von implizitem und explizitem Wissen. Implizites Wissen wird als nicht verbalisierbares und nicht bewusstseinsfähiges Wissen verstanden. Die Aneignung des impliziten Wissens geschieht somit passiv und ist damit auch nicht zielgerichtet oder selektiv (Polany, 1985). Innerhalb des impliziten Wissens unterscheidet Berry (1987) noch zwischen echtem implizitem Wissen und implizitem Wissen. Ersteres wird grundsätzlich unbewusst aufgenommen bzw. gelernt. Die zweite Wissensform wird zunächst explizit und damit zielgerichtet erworben und sodann durch verschiedenartige miteinander verknüpfte geistige Prozesse zum impliziten Wissen transformiert. Das explizite Wissen hingegen ist verbalisierbar und wird bewusst erworben.

Zusammenfassend kann die Aussage getroffen werden, dass die Differenzierung zwischen implizitem und explizitem Wissen eine große Schnittmenge mit der von prozeduralem und deklarativem Wissen aufweist. Deklaratives Wissen ist in dem Sinne explizierbares Wissen. Das Verständnis von implizitem Wissen geht größtenteils mit der Auffassung über das prozedurale Wissen einher (Süß, 1996).

Im Alltag hat sich der Begriff des Allgemeinwissens durchgesetzt, das ein umfassendes Wissen verschiedener Gebiete repräsentieren soll. Die Basis bildet dabei das deklarative und prozedurale Wissen. Neidhard-Wilberg (2005) umschreibt das Allgemeinwissen wie folgt: „Das Allgemeinwissen kann als von mehreren Kohorten geteiltes Wissen definiert werden und umfasst sowohl Kenntnisse, die durch Erziehung und Bildung intentional vermittelt, als auch diejenigen, die durch Massenmedien verbreitet und inzidental aufgenommen werden“ (S.145). Hinsichtlich der Verwendung und Bedeutung erfüllt das Allgemeinwissen prinzipiell drei verschiedene Funktionen (Wolff & Stock, 2000, S.39):

1. Die *Einstiegsfunktion* des Allgemeinwissens- für die Aneignung von Spezialwissen im Sinne von Schlüsselkenntnissen, die den Zugang zu einem immer heterogener werdenden Spektrum des Wissens eröffnen.
2. Die *kommunikative Funktion* des Allgemeinwissens- für die notwendige gegenseitige Verständigung zwischen Experten und damit als Klammer zur Nutzung und Verknüpfung von Spezialwissen.
3. Die *Bewertungsfunktion* des Allgemeinwissens- für die Orientierung in der Flut an Informationen, die aufgegriffen, verknüpft und mit den eigenen Erfahrungswerten verbunden werden müssen.

Das Allgemeinwissen umfasst in diesem Zusammenhang jedoch nicht nur das inhaltliche Faktenwissen, sondern bezieht noch weitere Bedeutungskomponenten mit ein (Wolff & Stock, 2000; Wolf, 1999). Gemeint sind auch kommunikative Kompetenzen (z.B. Beherrschung der Sprache, Teamfähigkeit), persönliche Fähigkeiten (z.B. Flexibilität, Selbstmanagement), gesellschaftlich ethische Orientierungsmöglichkeiten (z.B. kultureller Austausch, Solidarität), Kulturtechniken (z.B. das Lesen, Schreiben, Rechnen, Fremdsprachenkenntnisse) und inhaltliche Grundlagen und Problemwissen (z.B. aus der Wirtschaft, Recht, Naturwissenschaft). Im Hinblick auf die Gemeinsamkeiten zwischen Allgemeinbildung und Allgemeinwissen bedient sich das ehemalige Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBWF) mit dem Titel „Vier Felder des modernen Allgemeinwissens als Basis für einen kompetenten Umgang mit Angeboten der Informations- und Wissensgesellschaft“ (1997) einer ähnlichen begrifflichen Annäherung.

Im Rahmen des integrierten Abschlussberichts der Wissens-Delphi-Befragung von 1996/1998 werden demnach vier Felder des „modernen Allgemeinwissens“ – welches als Allgemeinbildung aufgefasst werden kann- differenziert (Bundesministerium für Bildung und Forschung, 1999, S.125):

1. Inhaltliches Basiswissen (Wissen und Grundlagen über aktuelle Probleme, (Integration, Beruf, etc.), Wirtschaft Recht, Pädagogik, Politik, Ökologie, Geschichte, etc.)
2. Instrumentelle Fertigkeiten (Allgemeine Kulturtechniken, wie Fremdsprachenkenntnisse, Umgang mit Informationstechniken, Technikverständnis, etc.)
3. Personale Kompetenzen (Persönliche Fähigkeiten und Einstellungen im Umgang mit Wissen, wie z.B. Reflexionsfähigkeit, Empathie, Offenheit, Urteilsvermögen, persönliches Erfahrungswissen, etc.)
4. Soziale Fähigkeiten (Soziale Verantwortung, kommunikative Kompetenzen, etc.)

Der Aufzählung ist wiederum zu entnehmen, dass dem Allgemeinwissen neben dem reinen Basiswissen auch instrumentelle, personale und soziale Fähigkeiten und Fertigkeiten zugeordnet werden. Anders ausgedrückt bildet das Basiswissen ein Fundament für die Herausbildung und Entwicklung der weiteren Fähigkeiten, so dass Bildung als dynamischer Prozess zu verstehen ist.

Vor diesem Hintergrund wird die Bedeutung des START-W-Wissenstest und der START-Testbatterie für Berufseinsteiger erneut unterstrichen. Der START-W erfasst das sogenannte inhaltliche Basiswissen und die anderen Module fokussieren beispielsweise den Bereich der Persönlichkeit. Insofern werden wichtige Facetten bzw. Komponenten des modernen Allgemeinwissens berücksichtigt und der Komplexität des Konstrukts wird somit Rechnung getragen.

## **5.2 Wissenspsychologische Paradigmen und Aspekte**

Auf der einen Seite schreibt Spinner (1994): „ Der kognitive Kern allen Wissens, unbeschadet zusätzlicher epistemologischer Qualifikationen, wie Geltung, Gewissheit oder philosophischer Vollkommenheits- und Vernunftprädikate, ist (...) der aufgenommene Informationsgehalt“(S. 24f). Auf der anderen Seite wurde schon in der Taxonomie der Wissensbegriffe angedeutet, dass das Wissen nicht nur die reine Repräsentation von objektiven Informationen darstellt, sondern vielmehr als ein Netzwerk von Ursachen



und Wirkungen zu verstehen ist, die durch die Umwelt in mannigfaltiger Weise an die Person herangetragen werden und in unterschiedlichen geistigen Prozessen verarbeitet und interpretiert werden. Wissen und Informationen bedingen sich letztendlich gegenseitig. Es kann jedoch konstatiert werden, dass die Informationen das Rohmaterial des Wissens bildet (Schüppel, 1996). Der Mensch verarbeitet Informationen, „um Wissen zu erwerben, denkend über dieses Wissen hinauszugreifen, neue Einsichten zu gewinnen und schwierige Probleme zu lösen“ (Klix & Spada, 1998, S.1).

Im Rahmen der Wissenspsychologie lassen sich darauf aufbauend die Teilbereiche (a) Wissenserwerb, (b) Wissensrepräsentation, (c) Wissensanwendung und (d) Wissensveränderung unterscheiden. Es wird angemerkt, dass eine detaillierte Gesamtdarstellung möglicher Erklärungsmodelle bezüglich des Erwerbs von Wissen und der Repräsentationssysteme an dieser Stelle zu weit führen würde. Daher werden im Folgenden nur die Grundzüge dargelegt und zur weiteren Vertiefung auf den schematheoretischen Ansatz, den Ansatz der Produktionssysteme, den Ansatz mentaler Modelle und die skizzierten Modelle der Wissensrepräsentation verwiesen (zum Überblick Klix, 1988; Mandl & Spada, 1988; Mandl, Friedrich & Hron, 1988).

Der Wissenserwerb wird als Prozess verstanden, der die Informationssuche, -aufnahme und -verarbeitung umfasst. Die Kapazität der Wissensaufnahme ist jedoch begrenzt. Die Größe der Kapazität bestimmt sich durch Art und Umfang des vorhandenen Wissensbesitzes. Die Wissensaufnahme ist somit abhängig von der angelegten kognitiven Struktur des Wissens und dem Informationsgehalt der zusammenhängenden und interdependenten Wissensseinheiten. Mit anderen Worten übernimmt das Vorwissen hier eine wichtige Funktion zur effektiven und effizienten Aufnahme neuen Wissens. Hinsichtlich einer geeigneten Begriffsbestimmung bezeichnen Dochy und Alexander (1995) das Vorwissen als das zur Verfügung stehende gesamte Wissen, welches (a) dynamischen Charakter hat, (b) vor dem Lernprozess bereitgestellt wird, (c) in verschiedenen Ausprägungen vorhanden ist, (d) teils explizit und implizit ist, (e) metakognitive Elemente einschließt und (f) strukturiert ist. Eine sinnvolle Illustration wie Wissen entsteht geben Rehäuser und Krcemar (1996, S. 6; siehe auch North, 2005, S.32). In ihrem grafischen Modell, welches hierarchisch aufgebaut ist, wird auf der untersten Ebene ein Zeichenvorrat unterstellt, der dem Individuum zur Verfügung steht (z.B. „1“, „7“, „0“, „“ etc.). Diese informationstheoretischen Datenbestandteile sind jedoch zunächst zusammenhangslos und verdichten sich erst durch die Kombination anhand von Zuordnungsregeln. Sobald die gewonnenen Daten (z.B. „1,70“) in einen Problemkontext eingeordnet werden,

können die Daten als Information interpretiert werden (z.B. Devisenkurs: 1\$=1,70 DM). Durch die Vernetzung mit bereits vorhandenen Wissensstrukturen (z.B. Marktmechanismus des Devisenmarktes) entfaltet die Information seine Bedeutung und es entsteht neues Wissen. Der Wissensbegriff als solcher bezeichnet demnach „das Potential, Informationen in das Handeln zu integrieren“ (Dick & Wehner, 2002, S.15).

Hier werden nochmals zwei Aspekte deutlich: Ein fundiertes bzw. defizitäres Vorwissen fördert oder begrenzt die Aneignung von neuem Wissen und aus eignungsdiagnostischer Sicht erlaubt die Bestimmung des Wissensstandes einer Person (anhand eines Wissenstests) auch einen Rückschluss auf das zur Lösung gebrauchte Hintergrundwissen.

Die Leistung des Arbeitsgedächtnisses sowie der emotional-motivationale Aktivierungsgrad und das Interesse einer Person sind zudem weitere entscheidende Determinanten, die die Aneignung von Wissen beeinflussen (Schüppel, 1996). Anderson (1983) vergleicht die Begrenztheit des Wissenserwerbs treffend als einen Flaschenhals kognitiver Aktivitäten. Der Wissenserwerb ist demnach sehr eng an bestimmte Lernprozesse gekoppelt. In diesem Zusammenhang kann jeder Lernprozess „als Selbstmodifikation des Wissenszustandes eines Lerners verstanden werden (...) sie wird durch die Aufgaben oder Lernziele und die Interaktion mit internen oder externen Informationsquellen initiiert und in Gang gehalten“ (Möbus & Schröder, 1998, S. 415). Lern- und Wissenserwerbsprozesse initiieren danach den Aufbau neuer Wissensstrukturen und die Akkommodation bzw. Umstrukturierung bereits vorhandener Wissensstrukturen. Im Ergebnis entfaltet sich ein umfassendes Wissensnetzwerk, welches ständig erweiterungsfähig ist.

Vor diesem Hintergrund stellt sich die Frage, wie das erworbene Wissen, genauer das Faktenwissen, Handlungswissen und Metawissen (Wissen über die Planung und Steuerung von Handlungen), strukturiert und organisiert wird (Mandl & Spada, 1988). Unter der Prämisse, dass Wissen nicht nur reines Faktenwissen darstellt, sondern auch algorithmische Fertigkeiten, Wissen über Ursachen und Wirkungszusammenhänge, Heuristiken, etc. mit einschließt, werden zur Beantwortung dieser Frage verschiedene Repräsentationsmodelle bzw. Repräsentationssysteme herangezogen, die nachfolgend skizziert werden. Mandl und Spada (1988) unterscheiden drei elementare Arten, (a) analoge Repräsentation, (b) prozedurale Repräsentation und (c) propositionale Repräsentation.

Analoge Repräsentation stellen nach Steiner (1988) Abbildungen dar, „die die Eigenschaften (oder zumindest einige von ihnen) eines abzubildenden Objekts oder Umweltereignisses beibehalten, Abbildungen also, die in einer bestimmten Weise den äußeren Gegebenheiten ähnlich sind“ (S.99). So gesehen werden räumlich-visuelle als auch alle anderen Sinnesmodalitäten einbezogen, die ein reales Abbild einer Sache ermöglichen. Im Zuge der Forschungsanstrengungen wird demnach das Verhältnis zwischen innerer Repräsentation und Repräsentiertem, die Rahmenbedingungen sowie das Wirkungs- und Funktionsgefüge des Prozesses untersucht. Eine gängige Umsetzung der prozeduralen und propositionalen Konzepte (Repräsentationen) stellt vor allem das Modell der Produktionssysteme dar. Produktionssysteme werden dabei als Rahmenkonzeptionen verstanden, die „Strukturtheorien des Gedächtnisses mit Annahmen über Prozesse der menschlichen Informationsverarbeitung in integrierter Form miteinander [verbinden]“ (Opwis, 1988, S.74). Innerhalb der Modellarchitektur werden das angesprochene deklarative, prozedurale und das Kontrollwissen verknüpft und in drei Systemkomponenten berücksichtigt. Der *Datenspeicher* als erstes Element enthält demnach das deklarative Wissen einer Person. Der *Produktionsspeicher* verfügt über das prozedurale Wissen, welches sich in Form oder in Kenntnis von sogenannten Produktionsregeln manifestiert und zur Problemlösung beiträgt. Eine stark vereinfachte Produktion ist die „Wenn-Dann-Formel“, die eine Bedingung (erster Teil) und eine angemessene Aktion (zweiter Teil) enthält. Die Komponente *Interpreter* bezeichnet das Kontrollwissen, das dazu dient, den Informationsverarbeitungsprozess zwischen Datenspeicher und Produktionsspeicher zu steuern. Mit anderen Worten werden die Bedingungen unter Verwendung des Datenspeicherinhalts ausgewertet und daraus geeignete Produktionen generiert. Anschließend wird eine Auswahl über die zur Verfügung stehenden Produktionen getroffen und die Ausführung initiiert. Neben den Produktionssystemen werden im Weiteren auch sogenannte Netzwerkstrukturen als Modelle der Wissensrepräsentation verwendet. Eine überblicksartige Darstellung gibt zum Beispiel Anderson (2001).

Im Hinblick auf die Wissensanwendung und der damit einhergehenden Wissensveränderung ist nicht mehr der Erwerb und die Speicherung von besonderem Interesse, sondern vielmehr die Aktivierung und Wirkungsweise von Wissen. Das Wissen wird in diesem Kontext zum Werkzeug, um verschiedene Problemstellungen zu bewältigen. Ein Problem definiert sich dabei durch einen gegebenen (a) Anfangszustand, (b) einen erwünschten, aber noch nicht erreichten Zielzustand und (c) durch eine Barriere, die die Transformation des Anfangs- in den Zielzustand verhindert (Putz-Osterloh, 1988).

Das Problemlösen bezeichnet darauf aufbauend einen Suchprozess nach Lösungsstrategien, um die Barriere zu überwinden. Die eingangs dargestellten Wissensarten können diesbezüglich helfen, den Zielzustand zu erreichen bzw. das Problem zu lösen: Das deklarative Wissen dient einerseits dazu, Kategorien für den Problemzustand zu bilden, um den Anfangs- und Zielzustand erstmal erfassen zu können (Putz- Osterloh, 1988). Andererseits kann ein umfassendes deklaratives Wissen das Auftreten von Problemsituationen schon im Vorfeld unterbinden (Kluwe, 1990). Eine adäquate faktische Wissensbasis zur Behebung einer vermeintlichen Problemsituation steht nämlich bereits zur Verfügung. Das prozedurale Wissen im Sinne von geeigneten Heuristiken kann der Person eine ergänzende Hilfestellung geben, um ein aufgetretenes Problem zu beseitigen bzw. den Anfangszustand in den Zielzustand zu transformieren. Das entsprechende Lösungsverfahren kann demnach als gespeicherter Algorithmus aus dem Gedächtnis abgerufen werden (Putz-Osterloh, 1988). Letztendlich beruht die Lösung eines Problems auf zwei grundlegenden Faktoren: (a) umfangreiches und gut organisiertes Wissen; (b) Verfügbarkeit von Problemlösestrategien, die mit dem Faktenwissen verknüpft sind (Chi, Glaser & Rees, 1982, zitiert nach Kluwe, 1990). Die Lösung eines Problems bedeutet auch gleichzeitig, dass eine Wissenslücke geschlossen wird. Die kognitive Speicherung des Ergebnisses führt sodann zu einer Veränderung und Erweiterung des deklarativen und prozeduralen Wissens. Es sei angemerkt, dass die Wissensanwendung genauso wie der Wissenserwerb von emotionalen und motivationalen Faktoren abhängt (Putz-Osterloh, 1988).

### **5.3 Wissenserwerb als Lernprozess**

In Anlehnung an die Ausführungen des vorangegangenen Kapitels besteht zwischen dem Wissen und dem Lernen ein enger Zusammenhang. Lernen und Wissen stehen darüber hinaus in einem interdependenten Verhältnis. Im Hinblick auf die empirischen Forschungsfragen (Kapitel 8) und die zugrunde gelegte Stichprobe von Schülern, Auszubildenden und Studenten (Kapitel 9.3) wird im Folgenden das Thema Lernen und der Wissenserwerb in schulischen Bildungseinrichtungen überblicksartig erläutert.

Einleitend wird unter Lernen folgendes verstanden: „Wissenserwerb impliziert, dass die Lernenden neue Wissensstrukturen aufbauen, untereinander vernetzen, mit bestehenden Konzepten verknüpfen und immer wieder in verschiedenen Situationen verwenden

sowie mit neuen Kontexten verbinden“ (Reinmann-Rothmeier & Mandl, 1998, S. 466). Der Erwerb von Wissen und damit auch das Lernen sind darüber hinaus an bestimmte Bedingungen gebunden. Diese äußern sich in den folgenden fünf Charakteristika (Reinmann-Rothmeier & Mandl, 1998):

1. Extrinsische oder intrinsische Motivation und Interesse: Der Wissenserwerb setzt immer die aktive Beteiligung des Lernenden voraus.
2. Selbststeuerung: Der Wissenserwerb wird dabei selbstständig überwacht und gesteuert. Kognitive, emotionale und motivationale Komponenten der Selbststeuerung spielen hier ebenfalls eine entscheidende Rolle.
3. Konstruktivistischer Wissenserwerb: Der Erwerb von Wissen wird bei dieser Annahme nicht einfach rezipiert, sondern in bereits vorhandene Wissensstrukturen eingebunden und anhand der individuellen Erfahrungen interpretiert.
4. Situativer Wissenserwerb: Das Lernen findet immer in einem spezifischen Kontext statt, in dem der Lernende Wissen erwirbt und dessen Anwendungsmöglichkeit erfährt.
5. Wissenserwerb als sozialer Prozess: Jede Form des Wissenserwerbs findet vor dem Hintergrund sozialer Prozesse statt. Die soziale Interaktion beim Wissenserwerb und die sozio-kulturellen Einflussfaktoren auf die Prozesse und die Inhalte des Wissens sind in diesem Kontext von großer Wichtigkeit und Bewandnis.

Als theoretischer Hintergrund für das Themenfeld der Lernpsychologie werden hier zwei Erklärungsvarianten herangezogen:

1. Behavioristische Erklärungsperspektive
2. Kognitivistische Erklärungsperspektive

Zu 1) Der zentrale Untersuchungsgegenstand der Behavioristen ist das beobachtbare Verhalten und die daran anknüpfende intendierte Modifikation des Verhaltens. Das entscheidende Grundprinzip der klassischen Konditionierung sind bestimmte Reiz- Reaktions- Verbindungen. Die angestrebte Verhaltensänderung als Lernleistung ist dabei das Ergebnis von konkreten Stimulusbedingungen aus der Umwelt (Schüppel, 1996). In diesem Zusammenhang stehen die berühmt gewordenen Studien von Pawlow (1953). Aufgrund ihrer weiten Verbreitung werden die Studien an dieser Stelle nicht weiter vertieft. Eine weitere Erklärungsperspektive der Behavioristen besteht in der operanten Konditionierung. Die gewollte Verhaltensänderung bzw. Lernleistung

ergibt sich hierbei aus den antizipierten Konsequenzen einer Handlung. Die ebenfalls bedeutenden Studien von Skinner (1974) zeigen auf, dass eine gewünschte Verhaltensänderung durch die selektive Gabe von Futter anerzogen werden konnte. Die Basis der Behavioristen lässt sich prinzipiell auf den Empirismus zurückführen, wobei der Schlüssel für das Lernen die Erfahrung ist.

Zu 2) Während für die Behavioristen der Empirismus von entscheidender Bedeutung ist, wird von den Kognitivisten der Rationalismus bzw. die Introspektion in den Vordergrund gerückt. Auf dieser Grundlage wird das „Bewusstsein“ bzw. das Denken als Quelle des Lernens oder Wissens angesehen. Die aus der Umwelt gewonnenen erfahrungsbezogenen Informationen werden lediglich als „Rohmaterial“ in die kognitive Erklärungsperspektive mit einbezogen. Der theoretische Hintergrund dieses Ansatzes wird jedoch vielmehr durch die kognitiven Prozesse, wie Wahrnehmen, Denken, Urteilen, logisches Schlussfolgern, Problemlösen, etc. gebildet (Schüppel, 1996). Die folgenden fünf Grundannahmen sind in diesem Rahmen konstitutiv für die kognitivistische Erklärungsperspektive (Schönpflug & Schönpflug, 1989):

1. Prinzip der Erkenntnis: Die Quelle des psychischen Bewusstseins ist hierbei die Erkenntnis bzw. die Kognition, die sich auf die Welt und die eigene Person bezieht.
2. Prinzip der Bewusstheit
3. Prinzip der kognitiven Ordnung: Die Erkenntnis setzt sich hierbei aus den generierten Zusammenhängen unterschiedlicher wahrgenommener Daten zusammen. Der Vorgang umfasst auf der strukturellen Ebene die Generierung von begrifflichen Gruppen, und auf funktionaler Ebene die Generierung von Ursache- Wirkung- Verbindungen.
4. Prinzip des einsichtigen Handelns: Auf der Basis der kognitiven Ordnung wird es der Person möglich, Erwartungen über die Zukunft zu bilden und ihr Handeln dahingehend zu planen. Die Handlungswahl bestimmt sich darüber hinaus durch den Nutzen und die Erfolgserwartung.
5. Prinzip der Selbstverantwortung und Selbstregulation: Der Mensch ist in diesem Sinne frei und einsichtig und trägt damit die Verantwortung für sein Handeln.

Innerhalb der kognitivistischen Erklärungsperspektive ist das Lewin'sche Modell (1963) von besonderer Bedeutung.

In diesem Modell wird ein zirkulärer Verlauf des Lernens postuliert, der durch vier Phasen charakterisiert wird: (a) erfahrene Realität, (b) Reflexion und Analyse des Erfahrenen, (c) Entwicklung von Abstraktionen und Generalisierungen und (d) Überprüfung der gezogenen Schlussfolgerungen in neuen Handlungssituationen. Eine ausführliche Beschreibung würde hier zu weit führen, daher wird an dieser Stelle lediglich auf das Modell verwiesen.

Die Analyse und Interpretation der erhobenen Daten im empirischen Teil der Arbeit bezieht sich vor allem auf die Personengruppe von Auszubildenden, deren Wissenserwerb vor allem an schulische Bildungsangebote gebunden ist. Aufgrund dessen wird die behavioristische und kognitivistische Erklärungsperspektive des Lernens in den Kontext der *Instructional Design Modelle* integriert, um eine mögliche, kurze und anwendungsbezogene Verbindung des Wissenserwerbs zur Bildungseinrichtung Schule zu gewährleisten.

Die Instructional Design Modelle zeigen in diesem Bezugsrahmen einen gangbaren Weg zur Darstellung systemorientierter Lernumgebungen auf und stellen eine Möglichkeit zur Förderung des Lernens dar. Die ursprüngliche Idee des Ansatzes ist, dass der Wissenserwerb systematisch geplant und gesteuert werden kann. Der Lehrer präsentiert dabei bereits vorgefertigte Informationen, die der Schüler in einer eher passiven Rolle (rezeptiv) aufnimmt (Reinmann-Rothmeier & Mandl, 1998). Aus dem behavioristischen Blickwinkel kann Lernen in diesem Modell „nur durch äußere Manipulation, durch Zerlegung und dosierte Sequenzierung der Lerninhalte und Aufgaben gefördert werden“ (Reinmann-Rothmeier & Mandl, 1998, S.477). Die operationale Definition der einzelnen Lehrinhalte und die Vermittlung von primär reproduktiven Wissen oder automatisierten Fähigkeiten stehen dabei im Vordergrund. Aus der kognitivistischen Perspektive wird weniger eine Instruktion zur Steuerung des Lernens initiiert, sondern vielmehr eine Instruktion zur Erleichterung des Lernens gegeben. Die Lernziele untergliedern sich im Gegensatz zur behavioristischen Sichtweise in Wissen, Fertigkeiten und Einstellungen, wobei neben der reinen Informationsdarbietung auch kognitive Strategien beachtet werden.

Die typische Vorgehensweise der Instructional-Design-Modelle gliedert sich in vier Teilschritte (Reinmann-Rothmeier & Mandl, 1998):

1. Analyse der Anfangszustände der Schüler bezüglich ihres Vorwissensstandes und ihrem Fähigkeitspotential. Die behavioristischen Instructional-Design-Modelle untersuchen dabei vor allem die Quantität des Vorwissens und die kognitivistischen Instructional-Design-Modelle vorwiegend die Qualität des Vorwissens, die auch kognitive Strategien und motivationale Aspekte berücksichtigen.
2. Analyse der gewünschten Endzustände bzw. Lernziele
3. Analyse der Übergänge zwischen Anfangs- und Endzustand, um den Wissensbedarf zu ermitteln. Aufgrund der Analyse werden sodann Instruktionen zur Darbietung von systematischen Informationen abgeleitet.
4. Evaluation der Lernergebnisse, die auch eine Rückkopplung auf die ausgewählten Instruktionsstrategien- und -methoden zulassen.

Einige Aussagen innerhalb des Modells zeigen einen starken Bezug zu dem gegenwärtigen und kontrovers diskutierten Frontalunterricht auf. Insbesondere die passive Rolle des Schülers, die vorgefertigte Informationsdarbietung des Lehrers als auch der fehlende Anwendungsbezug des vermittelten Wissens wird in diesem Zusammenhang kritisiert. Daran anknüpfend kann vermutet werden, dass das volle Leistungspotential der Schüler möglicherweise nicht ausgeschöpft wird.

Es wird angemerkt, dass neben den systemtheoretischen Lernumgebungen noch weitere mögliche Ansätze zur Förderung des Lernens bestehen. Hierzu zählen bspw. die problemorientierte Lernumgebung und die adaptive Lernumgebung (Reinmann-Rothmeier, & Mandl, 1998).

#### **5.4 Ausblick- Wissen und seine strategische Relevanz**

Es lässt sich festhalten, dass das Konstrukt Wissen sehr komplex ist und durch verschiedene psychologische Facetten gekennzeichnet ist.

Ein besonderer Aspekt dabei ist, dass das Wissen mit dem Konstrukt der Intelligenz eng verknüpft ist. Auf diesen Zusammenhang wird im nächsten Kapitel detailliert eingegangen. Es wird jedoch an dieser Stelle noch einmal angemerkt, dass das Wissen nach



Lynn, Wilberg und Margraf-Stiksrud (2004) eine entscheidende Komponente der Intelligenz ist. Aus diesem Grund wird das Wissen sowohl in die verschiedenen Intelligenzmodelle integriert (z.B. Cattell, 1971a, 1987), in einigen Intelligenztests berücksichtigt (z.B. I-S-T-2000 R von Amthauer et al., 2001; Liepmann et al., 2007) oder als eigenständiger Prädiktor für den schulischen oder beruflichen Erfolg herangezogen (siehe Kapitel 3.4). In diesem Kontext kann auch die Bedeutung des Wissens hervorgehoben werden. Im Zentrum der heutigen „Informationsgesellschaft“ oder „Wissensgesellschaft“ steht „die bewusste Verarbeitung, Umwandlung, Bewertung und Verknüpfung von Informationen zu etwas Neuem, zu Ideen, Handlungen, zu Produkten und Dienstleistungen“ (Blümelhuber, 2005, S.145). Das Wissen kann dabei als Grundvoraussetzung verstanden werden, um neuartige (berufliche) Probleme zu bewältigen, fachübergreifend zu denken, zielgerichtet zu handeln und Zusammenhänge herzustellen (Klix & Spada, 1998). Darüber hinaus kann eine breite und fundierte (allgemeine) Wissensbasis dazu dienen, die Aufnahme von neuem Wissen effizient und effektiv zu gestalten. Das heißt, eine bereits angelegte und fundierte Wissensstruktur kann die Aufnahme von neuen Informationen begünstigen, weil die neuen Informationen mit bereits vorhandenen Wissensstrukturen vernetzt werden. Das Vorwissen übernimmt demnach eine Moderatorfunktion für das weitere Lernen und bestimmt damit gleichzeitig das Entwicklungspotential der Person (Schüppel, 1996). Aus diesen Gründen spielt das (Allgemein-)Wissen für den beruflichen Werdegang eine wichtige Rolle und wird von den Arbeitgebern als grundsätzliche Vorbedingung für den Einstieg in das Berufsleben gefordert (z.B. Industrie- und Handelskammer, 2007). Infolgedessen kann bereits im Vorgriff konstatiert werden, dass das Wissen heutzutage und in der Zukunft eine Schlüsselposition für den beruflichen und privaten Erfolg einnimmt bzw. einnehmen wird.

Vor diesem Hintergrund gewinnen eignungsdiagnostische Instrumente an Bedeutung, die den Wissensstand feststellen und das Entwicklungspotential prognostizieren können. Die Problematik einer repräsentativen Wissensdiagnose bezieht sich hier auf die unterschiedliche Art des Wissensbesitzes. Dieser lässt sich einerseits auf unbewusste, kognitive Vorgänge zurückführen (prozedurales Wissen) und andererseits auf bewusstseinsfähige, verbalisierbare Prozesse (deklaratives Wissen). Die herkömmlichen Tests zur Wissenserfassung verlangen in der Regel eine mündliche oder schriftliche Darstellung des Wissens. Die Tests erfassen damit aber hauptsächlich nur das deklarative Wissen und weniger das prozedurale Wissen (Kluwe, 1990). Hinweise über die Struktur des Wissens versucht man daher aus Verfahren abzuleiten, die spezifische Fähigkeiten

prüfen wie z.B. die Struktur-Lege-Technik (Reinmann-Rothmeier & Mandl, 1998). Eine vollständige Ermittlung des Wissens scheint zum gegenwärtigen Zeitpunkt noch nicht möglich zu sein und bedarf noch weiterer kognitionspsychologischer als auch neurophysiologischer Forschung. Im Endeffekt erscheint eine ganzheitliche Erfassung des Wissens –entsprechend den obigen Ausführungen- auch nicht zwingend notwendig zu sein, da die Diagnose von repräsentativen Wissensausschnitten gleichermaßen einen Rückschluss auf ein fundiertes Hintergrundwissen erlauben und einen zuverlässigen Indikator für zukünftiges Lernverhalten darstellen.

Einen zukunftsorientierten und innovativen Weg eröffnet in diesem Zusammenhang wiederum der zu entwickelnde Wissenstest START-W. Neben der Erfassung relevanter Themenbereiche des Allgemeinwissens wird die verbale, numerische und figurale Inhaltsskala berücksichtigt und zudem wird eine Interpretation des Faktorwertes kristallisierte Intelligenz ermöglicht.

## **6 Wissen und Intelligenz – Eine konzeptionelle Zusammenführung**

Bei Anwendung der „Alltagspsychologie“ wird der Wissende und Gebildete häufig auch als intelligent eingestuft. Anders ausgedrückt verwendet der Laie die Begriffe Intelligenz, Lernfähigkeit, Wissen, etc. wahrscheinlich synonym. Doch trotz bestehender Gemeinsamkeiten und Abhängigkeiten der Themengebiete Wissen und Intelligenz bestehen gleichermaßen erklärungsbedürftige Unterschiede.

Im Folgenden sollen die Unterschiede der Themengebiete skizziert werden und es soll aufgezeigt werden, inwieweit sie sich aufeinander beziehen. Im Hinblick auf die Strukturtheorie von Cattell (1971a, 1987) und dem Intelligenz-Struktur-Test 2000 R (Amthauer et al., 2001, Liepmann et al., 2007) soll darüber hinaus der Zusammenhang zwischen den so genannten Basiskompetenzen des Menschen (fluide Intelligenz) und dem kulturgebundenem Wissen (kristallisierte Intelligenz) näher betrachtet werden.

### **6.1 Grundlegende Zusammenhänge zwischen Intelligenz und Wissen**

Die Intelligenzforschung und die Wissenspsychologie entstammen zunächst unterschiedlichen Forschungstraditionen. Einerseits basiert der Untersuchungsgegenstand der Differentiellen Psychologie vorwiegend auf der Erforschung des Konstrukts Intelligenz.

Andererseits werden die Forschungsgegenstände Wissen und Lernfähigkeit eher der Allgemeinen Psychologie zugeordnet, wobei insbesondere die Lernfähigkeit unter ihrem Anwendungsaspekt auch in die Subdisziplin der Pädagogischen Psychologie eingegliedert werden kann. Dieser Sachverhalt impliziert gleichzeitig, dass verschiedene Interessenschwerpunkte bestehen und unterschiedliche konzeptuelle Methoden eingesetzt worden sind. Hinsichtlich dieser vermeintlichen Differenzierung erwägen Guthke und Beckmann (2001), die Intelligenzdiagnostik verstärkt kognitionspsychologisch zu fundieren, um die Erforschung des geistigen Leistungsgrades durch den gegenseitigen Austausch von Erkenntnissen voranzutreiben. Eine gute Möglichkeit zur Integration stellt der kognitive *Korrelate-Ansatz* und der *Komponenten-Ansatz* dar (siehe Kapitel 4.3).

Vor diesem Hintergrund wird im aktuellen wissenschaftlichen Diskurs die Frage thematisiert, welche Bedeutung das Wissen und die Intelligenz für berufliche oder akademische Leistungen hat (Fragestellung 1). Die andere grundlegende Frage behandelt darauf aufbauend das Verhältnis bzw. den Zusammenhang zwischen Wissen und Intelligenz (Fragestellung 2).

Zu 1) Im Hinblick auf das in diesem Zeitalter schnell veraltende Wissen sprechen einige Fachleute der allgemeinen geistigen Fähigkeit und damit der Intelligenz einen höheren Stellenwert zu. In diesem Zusammenhang werden Modebegriffe wie systemisches und vernetztes Denken, emotionale Intelligenz, Erfolgsintelligenz etc. verwendet. Die Wissenspsychologen betonen hingegen, dass gerade das domänenspezifische Wissen der beste Prädiktor für gute schulische oder berufliche Leistungen ist. Eine Bestätigung erfährt diese Annahme durch die Ergebnisse der Expertiseforschung hinsichtlich des Vergleichs von Experten und Novizen. Schneider (1993, zitiert nach Stern, 2001) zeigt in seinen Studien, dass vor allem die langjährige und effiziente Nutzung von Lern- und Übungsgelegenheiten eine Voraussetzung für geistige Höchstleistungen in den Bereichen wie z.B. Musik ist. Ericsson, Krampe & Tesch-Römer (1993) stellen des Weiteren heraus, dass es bei einer Überschreitung einer bestimmten Intelligenzschwelle weniger auf die Intelligenz ankomme, sondern vielmehr auf die Art und Intensität der Übung. Darüber hinaus gelangen Weinert und Helmke (1998) in ihren Untersuchungen zu dem Schluss, dass ein überdurchschnittliches Vorwissen Defizite in der allgemeinen Intelligenz ausgleichen kann und dass es keine eindeutigen Differenzen zwischen über- und unterdurchschnittlich intelligenten Personen gibt, die über ähnliches Vorwissen verfügen. In diesem Kontext konstatiert Weinert (1994), dass Wissen als Ergebnis

inhaltsbezogener Lernprozesse, mit stets begrenzter Verallgemeinerung und Abstraktion der erworbenen Kenntnisse, (...) zugleich die wichtigste Bedingung für zukünftiges Lernen innerhalb der gleichen Inhaltsdomäne [ist]“ (S. 197). Im Bezug auf eine repräsentative Befragung verschiedener Experten stellt Süß (2007) zudem die Annahme auf, „dass außergewöhnliche Leistungen im Wesentlichen auf bereichsspezifische Fertigkeiten zurückzuführen und nicht das Resultat interindividueller Unterschiede in basalen Fähigkeiten und Intelligenz sind“ (S.18). Sternberg (1985) kommt dagegen zu der Annahme, dass sich intelligente und weniger intelligente Experten womöglich bei der Bearbeitung von neuen Aufgaben unterscheiden, in denen sie kein direktes Vorwissen besitzen. Die Ergebnisse seiner Studie konnten die Vermutung jedoch nicht untermauern und zeigten eher niedrige Korrelationskoeffizienten zwischen der Höhe der Intelligenz und der Lösung neuartiger Aufgaben. Hinsichtlich der Vermutung, dass die Intelligenz die Ursache für interindividuelle Unterschiede sei, führte Stern (1994, 1998) eine weitere Studie durch. Auf Basis von Wissenstestergebnissen und Intelligenzmessungen zum Zeitpunkt der Erhebung und durch längsschnittliche Daten des zwei Jahre zurückliegenden Vorwissensstandes war es Stern möglich, die Einflussstärke der Intelligenz und des Vorwissens bezüglich der Lösung von mathematischen Textaufgaben gegeneinander abzuwägen. Das Resultat zeigte, dass Kinder mit kontinuierlichen, überdurchschnittlichen Mathematikleistungen (Vorwissen), unabhängig von ihrer Intelligenz, die besten Leistungen erbrachten und dass überdurchschnittlich intelligente Kinder mit unterdurchschnittlicher Mathematikleistung ähnlich schlechte Leistungen bei der Lösung von Textaufgaben erzielten wie Kinder mit unterdurchschnittlichen Mathematik- und Intelligenzleistungen. Entgegen der Vermutung stellte Stern (1994, 1998) fest, dass das Vorwissen und weniger die Intelligenz die entscheidende Leistungsdeterminante für den Lerntransfer in Textaufgaben ist. In die gleiche Richtung weisen auch die Forschungen von Weinert (1988). Hinsichtlich der Bearbeitung von leichten und schwierigen Mathematikaufgaben stellte Weinert heraus, dass das Vorwissen bei anspruchsvollen Aufgaben maßgeblich für die erfolgreiche Lösung war. Das Vorwissen übernimmt demnach eine kompensatorische Funktion bezüglich der kognitiven Fähigkeiten (Intelligenz). Bei leichten Fragestellungen konnte hingegen ein Kompensationseffekt in beide Richtungen ermittelt werden. Mit anderen Worten wächst die Relevanz des Vorwissens mit steigender Aufgabenkomplexität. Die aufgeführten Befunde konnten auch in weiteren zahlreichen Studien auf ähnliche Weise repliziert werden (z.B. Reinmann, 1988; Schneider, Körkel & Weinert, 1989; Waldmann & Weinert, 1990; Schneider & Bjorklund, 1992).

Eine Erklärung für die vorwissensgestützte Dominanz in der Experten-Novizen-Forschung wird letztendlich darin gesehen, dass durch das Vorwissen wichtige Informationen schneller erkannt, effizienter vernetzt, leichter gespeichert und abgefragt werden können. Aufgrund der automatisierten Aktivierung des Vorwissens können adäquate Problemlösestrategien zügig erarbeitet werden und die kognitive Bewusstseinskapazität gleichzeitig entlastet werden (z.B. Waldmann & Weinert, 1990).

Die Konklusion, dass Vorwissenstests oder Zensuren prinzipiell eine bessere Vorhersage garantieren können als Intelligenztests, muss an dieser Stelle indessen relativiert werden. Der Schulerfolg oder Berufserfolg wird nicht einzig allein durch die Intelligenz oder das Vorwissen bestimmt, sondern er ist vielmehr multifaktoriell determiniert. Die Forschungen von Guthke und Beckmann (2001) bestätigen den Befund, dass der Wissenserwerb in einer experimentell kontrollierten Lernsituation sowohl von der Intelligenz als auch durch das Vorwissen beeinflusst wird. Dieser Argumentationsstruktur folgend stellt sich auch Süß (1996) in seinen Studien die berechtigte Frage, wie Personen die Aneignung vollkommen unbekannter Informationen bewerkstelligen oder in fremden Situationen erfolgreich handeln können. Resümierend ziehen die Autoren den Schluss, dass die Informationsaufnahme und die anwendungsbereite Organisation der Wissensseinheiten im gleichen Maß von der Intelligenz und dem Vorwissen abhängen. Im Übrigen vermuten sie auch einen Zusammenhang zwischen der Intelligenzstruktur und dem erfolgreichen Erwerb von Expertise. In diesem Sinne kann die gegebene Intelligenz den Wissenserwerb und die Wissensanwendung begrenzen oder erleichtern. Die Intelligenz bildet folglich eine notwendige Voraussetzung für die Aneignung, Veränderung und Verwendung von Wissen. Ein logisch erscheinender Hinweis für das postulierte Zusammenspiel von Wissen und Intelligenz sind beispielsweise die Leistungsunterschiede bei gleich erfahrenen Experten oder die Tatsache, dass ein zehnjähriger Junge mit wenig Schacherfahrung erfahrene Schachspieler schlug (Guthke & Beckmann, 2001). In diesem Bezugsrahmen bekräftigen Gruber (1991) bzw. Gruber & Mandl (1996) die Resultate und konstatieren, dass Schachgroßmeister sowohl durch ihre exzellente Gedächtniskapazität und Kombinatorik auffallen als auch durch die auf Erfahrung basierende schnelle Identifikation und Wahrnehmung von Spielkonstellationen.

Schlussfolgernd erklärt neben dem bereichsspezifischen Vorwissen gleichermaßen die Intelligenz die bestehenden Leistungsunterschiede. Guthke und Beckmann (2001) sprechen in diesem Zusammenhang auch von intelligentem Wissen, das gut strukturiert, vernetzt und flexibel anwendbar ist. Im Hinblick auf die Experten-Novizen-Forschung

wird die Lösung von schwierigen mathematischen Textaufgaben bspw. durch die Kombination von Intelligenz, Wissen und Motivation determiniert. Die effektive und effiziente Verknüpfung und die Interdependenz zwischen der allgemeinen geistigen Grundausstattung (Intelligenz) und dem erworbenen Wissen stehen dabei im Vordergrund. Zusammenfassend kann man festhalten, dass das Wissen und die Intelligenz in möglicherweise variierender Einflussstärke mit dem Handeln verbunden sind und sich gegenseitig bedingen. Zu diesem Ergebnis kommt auch Schmidt (1992), der anhand seines pfadanalytischen Modells deutlich macht, dass beruflicher Erfolg sowohl von dem berufsspezifischen Wissen als auch von den kognitiven Fähigkeiten abhängt und somit nur unter Berücksichtigung multifaktorieller Moderatorvariablen begründet werden kann. Demgemäß lässt sich auch die allgemeine Leistungsvarianz nach Renkl und Stern (1994) auf drei Effekte zurückführen: „(a) den spezifischen Effekt des Vorwissens, in den der intelligenzunabhängige Übungseffekt eingeht; (b) den spezifischen Effekt der Intelligenz, in den Unterschiede in der allgemeinen Problemlösekompetenz eingehen; dieser kann in der Problemlösesituation zur Kompensation fehlenden Vorwissens genutzt werden; (c) den konfundierten Effekt aus Vorwissen und Intelligenz, in den die Auswirkungen der Intelligenz auf die Lerngeschichte eingehen“ (S.30).

Zu 2) Auf Basis dieser Annahme rückt die zweite Frage in den Mittelpunkt der Betrachtung. Süß (1996) nimmt in diesem Zusammenhang zunächst eine begriffliche Differenzierung vor und versucht darauf aufbauend die Intelligenz und das Wissen im Sinne der Investmenttheorie in Beziehung zu setzen. Die Intelligenz wird demnach „als Voraussetzung für den Erwerb und die Anwendung von Wissen konzeptualisiert und Wissen, wie in der Investmenttheorie von Cattell (1987), als investierte Intelligenz begriffen“ (Süß, 1996, S.80). Die Intelligenz bezeichnet nach dieser Auffassung intellektuelle Fähigkeiten zu „induktivem und zu deduktivem Denken, zur Erinnerung und zur Aufmerksamkeitssteuerung (...), die den Erwerb und die Anwendung von Wissen begrenzen und individuell variieren“ (Süß, 1996, S.80). Die postulierte Beziehung lässt im Weiteren vermuten, dass in empirischen Untersuchungen positive Korrelationen zwischen den intellektuellen Fähigkeiten (Intelligenz) und dem gegenstandsspezifischen Wissen bestehen. Ein bestätigendes Ergebnis zeigten Hunter und Schmidt (1989), die eine durchschnittliche Korrelation von  $r = .80$  bezüglich des Zusammenhangs von Intelligenztestleistungen und beruflichem Fachwissen errechneten. Eine ähnliche Auffassung

vertritt auch Stern (2001), die die Intelligenz als Potential des Menschen versteht, „um Lern- und Bildungsangebote zur Aneignung von Wissen zu nutzen“(S.163).

Guthke und Beckmann`s (2001) Studien weisen in diesem Bezugsrahmen in eine ähnliche Richtung. Die folgende Grafik (siehe Abbildung 8) soll den grundlegenden Zusammenhang zwischen Wissen und Intelligenz verdeutlichen:

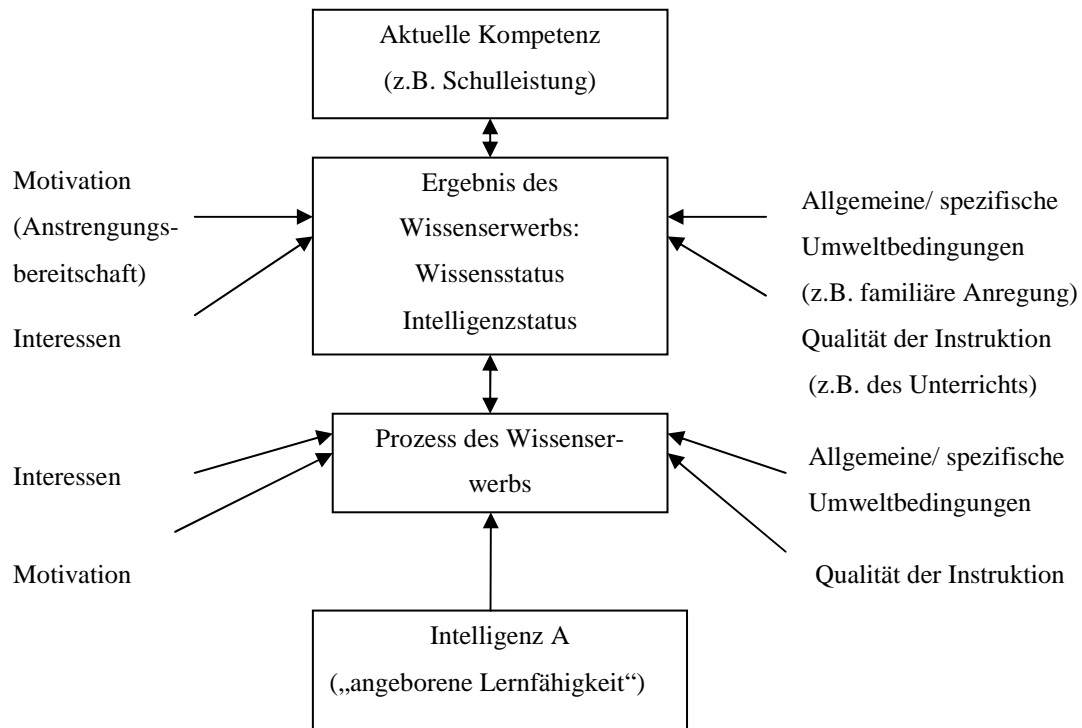


Abbildung 8: Die Interdependenz zwischen Wissen und Intelligenz  
( Guthke & Beckmann, 2001)

Die angeborene Intelligenz (Intelligenz A) bildet den Grundstock für den Wissenserwerb. Hinsichtlich der Investmenttheorie von Cattell (1971a, 1987) können hier konvergierende Elemente zur fluiden Intelligenz abgeleitet werden. Der Wissenserwerb, der unter verschiedenartigen Einflüssen stattfindet, bestimmt darauf aufbauend den Wissens- und Intelligenzstatus (Intelligenz B). Die Intelligenz B ist somit an den Wissenserwerbsprozess und an die angeborene Intelligenz A rückgekoppelt. Im Bezug zur Investmenttheorie weist die Intelligenz B einige Parallelen zu der kristallisierten Intelligenz auf. Die aufeinander bezogenen Intelligenzformen A und B äußern sich dann in der aktuellen Kompetenz (z.B. Schulleistung, IQ).

In einem ähnlichen Bezugsrahmen steht auch der Ansatz von Asendorpf (2004, siehe Abbildung 9). In diesem Rahmenmodell wird die Bedeutung der Intelligenz und des Wissens für die geistige Leistung ebenfalls hervorgehoben.

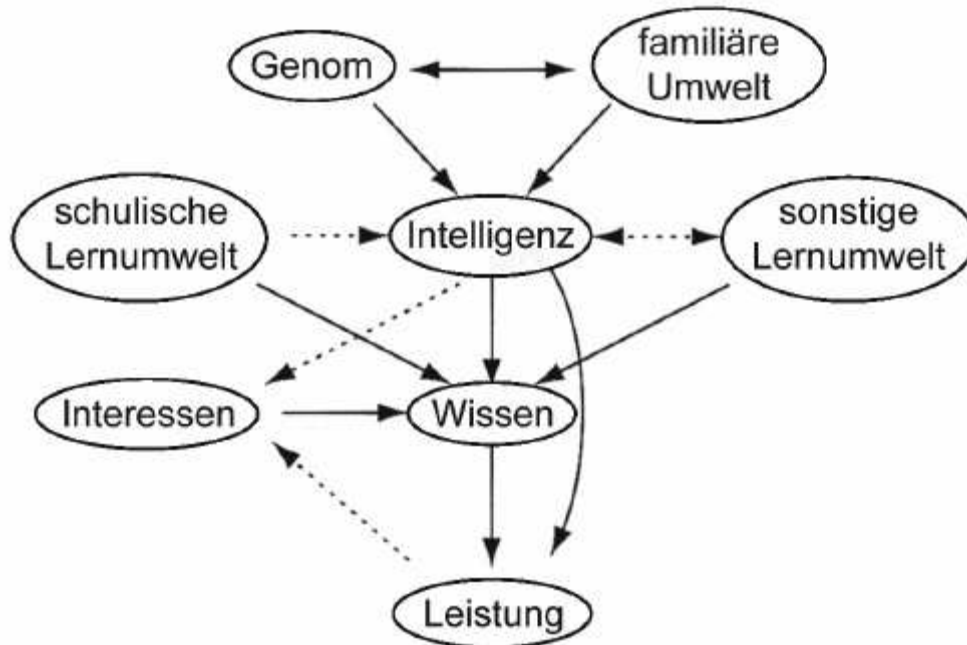


Abbildung 9: Das Rahmenmodell geistiger Leistung  
(Asendorpf, 2004, S. 351)

Die Intelligenz wird in diesem Modell durch verschiedenartige Faktoren direkt (durchgezogene Pfeile) und indirekt (gestrichelte Pfeile) beeinflusst. Die Haupteinflussfaktoren sind durch die genetische Grundausstattung und die familiäre Umwelt gekennzeichnet, die wiederum wechselseitig aufeinander bezogen sind. Im Bezug auf die Strukturtheorie von Cattell (1971a, 1987) kann die Intelligenz an dieser Schnittstelle als fluide Intelligenz charakterisiert werden. Unter dem Einfluss weiterer Umweltfaktoren kristallisieren sich auf Basis der latenten (fluiden) Intelligenz dann wissensabhängige Kompetenzen heraus, die als kristallisierte Intelligenz bezeichnet werden können. Die Intelligenz und das Wissen determinieren darauf aufbauend das geistige Leistungsverhalten der Person. Die Entfaltung des gegebenen Leistungspotentials ist jedoch an die jeweiligen Interessen und individuellen Umwelteinflüssen der Person (rück-) gekoppelt.



### ***Fazit***

Es lässt sich festhalten, dass die Intelligenz als Leistungsgrad psychischer Funktionen verstanden werden kann, der die Aneignung, Verarbeitung und Anwendung des Wissens in hohem Maße beeinflusst. Auf der Grundlage des gegebenen Leistungspotentials kann das generierte Wissen dann dazu dienen, die beruflichen und privaten Anforderungen durch ständige Lern- und Problemlöseprozesse zu bewältigen (Klix & Spada, 1998). Die Intelligenz und das Wissen stellen demnach entscheidende Determinanten für den beruflichen und privaten Erfolg dar (z.B. Süß, 2001; Schmidt & Hunter, 1998; siehe auch Kapitel 3.4), die aufgrund ihrer engen und wechselseitigen Verbindung stets simultan betrachtet werden sollten (Guthke & Beckmann, 2001). Das Modell von Cattell (1971a, 1987) stellt in der Theorie einen integrativen Ansatz dar, der hinsichtlich der kristallisierten Intelligenz einen engen Bezug zum Wissen aufweist. In der praxisorientierten Eignungsdiagnostik stellt der I-S-T 2000 R durch die Verknüpfung von Wissen und kristallisierter Intelligenz eine abgesicherte Möglichkeit dar, die Faktoren zu operationalisieren. Darüber hinaus besitzt auch der eingesetzte Wissenstest START-W, der weitgehend dem Erweiterungsmodul des I-S-T-2000 R entspricht, die Möglichkeit, die kristallisierte Intelligenz abzubilden. Das Beziehungsgeflecht zwischen der fluiden und der kristallisierten Intelligenz und dem Wissen wird im nächsten Unterkapitel näher betrachtet.

## **6.2 Das Beziehungsgeflecht zwischen fluider Intelligenz, kristallisierter Intelligenz und Wissen**

Ein sehr elaboriertes, psychometrisches Intelligenzmodell ist die Struktur- bzw. Investmenttheorie von Cattell (1971a, 1987). Innerhalb des hierarchischen Modells wird auf der zweiten Ebene zwischen fluider und kristallisierter Intelligenz unterschieden (siehe Kapitel 4.4.4). Diese Differenzierung ist hier von besonderem Interesse, weil die kristallisierte Intelligenz einen engen Bezug zum Wissenserwerb aufweist. Die fluide Intelligenz wird dabei als investierte Intelligenz begriffen. Im Folgenden werden die beiden Intelligenzkomponenten zunächst begrifflich voneinander abgegrenzt und im Zuge der Investmenttheorie (Cattell, 1971a, 1987) miteinander verbunden.

Die fluide Intelligenz beschreibt Cattell (1971b) wie folgt:

It is a capacity for insight into complex relations, accounting for that part of test variance which seems independent of the sensory or cultural area in which the test are expressed. The theory is that this capacity for relation education is neurological determined and a function of the size and functionality of the general association mass- i.e. the areas not devoted to specific sensory, motor, vegetative, and emotional control functions.(S.13)

In diesem Sinne wird die fluide Intelligenz als grundlegende biologische Lernkapazität verstanden, die sich in den ausgeprägten neuronalen Verknüpfungen des kognitiven Systems manifestiert. Die fluide Intelligenz wird damit als unspezifische Energie definiert, die in Form von Basisprozessen das Potential der geistigen Leistungsmöglichkeiten widerspiegelt. Darüber hinaus entspricht die fluide Intelligenz annähernd dem g- Faktor bzw. weist eine hohe Ladung zu diesem Faktor auf. Dementsprechend wird sie auch häufig als „*Mechanik der Intelligenz*“ (Baltes, 1990) oder als „Hardwarekomponente“ (Guthke & Beckmann, 2001) bezeichnet. Die ausgebildeten Fähigkeiten beruhen im Weiteren auf vorwiegend erbbedingten Einflüssen und weniger auf systematischen Akkulturationsprozessen. Horn (1988) führt dazu aus: „Individual differences in the knowledge of acculturation are not major determiners of successful performance on good measures of gf“(S.660). Die Messung der fluiden Intelligenz basiert demnach auf so genannten kulturfreien Tests, die Aufgaben zum schnellen Lernen, zur räumlich-visuellen Erfassung und zum schlussfolgernden Denken umfassen. Im Zuge der ontogenetischen Entwicklung kristallisieren sich aus den Basiskompetenzen (gf) dann spezifische Fähigkeiten (gc) heraus, die mit dem Wissenserwerb und der Wissensanwendung in Verbindung stehen (v.Gilardi, Holling & Schmidt, 1983). Im Rahmen der Investmenttheorie von Cattell (1971a, 1987) wird die fluide Intelligenz sozusagen in die kristallisierte Intelligenz investiert. „The term ‘crystallized’ is meant to imply this freezing in a specific shape of what was once fluid ability“(Cattell, 1987, S.140). Gf steckt folglich die Grenzen ab, innerhalb derer sich gc herauskristallisiert. Die fluide Intelligenz determiniert in diesem Sinne das Entwicklungspotential der kristallisierten Intelligenz (Stern, 2001).

Die kristallisierte Intelligenz stellt sich nach Horn (1988) wie folgt dar: „The measured factor is a fallible indicator of the extent to which an individual has incorporated, through the systematic influences of acculturation, the knowledge and sophistication that can be referred to as the intelligence of culture“(S.658).

Die kristallisierte Intelligenz kann somit als Reproduktion der aufgenommenen und organisierten Wissenssysteme aufgefasst werden, die in den jeweiligen kulturellen und gesellschaftlichen Kontext eingebunden sind (Kruse & Rudinger, 1997). Nach Gilardi et al. (1983) entspricht die kristallisierte Intelligenz „der Summe der Kenntnisse, die ein Individuum in einer bestimmten Kultur erworben hat“ (S.372). Horn (1988) bezeichnet die kristallisierte Intelligenz als *acculturation knowledge*. Danach zeigen sich die Fähigkeiten vor allem in der Lösung vertrauter kognitiver Aufgaben und in der Anwendung von gelerntem Wissen, das kulturell geprägt ist und dessen Erwerb häufig an institutionalisierte Bildungseinrichtungen gekoppelt ist. Des Weiteren umfasst die kristallisierte Intelligenz die gesammelten Erfahrungen, das Wissen, die Urteils- und Kommunikationsfähigkeit, das Verständnis von gesellschaftlichen bzw. sozialen Gepflogenheiten einer Person und die Befähigung zum sinnvollen Denken. Die kristallisierte Intelligenz wird unter anderem durch Aufgaben zum verbalen und numerischen Verständnis, logischen Schlussfolgern, etc. operationalisiert (Horn, 1988). Die Tests zur Überprüfung der kristallisierten Intelligenz beinhalten daher vorwiegend sprach- und zahlengebundene Aufgabenstellungen und allgemeine Wissenstests (Horn & Cattell, 1966). Im Bezug auf die in Kapitel 5.1 dargestellten Wissensformen lässt sich folgender Zusammenhang festhalten: Die kristallisierte Intelligenz repräsentiert (a) allgemeine Systeme des deklarativen und prozeduralen Wissens, (b) spezielle Systeme des deklarativen und prozeduralen Wissens (bereichsspezifisches Wissen), (c) Wissen über Fähigkeiten zur Ausübung verschiedener Handlungen (Baltes, 1990). Eine weitere empirische Bestätigung des Zusammenhangs zwischen kristallisierter Intelligenz und Wissen wird beispielsweise durch die Studien von Beier und Ackermann (2001) gegeben, die eine durchschnittliche Korrelation von  $r = .80$  ermittelten. Aufgrund der engen Verbindung zum Wissen und der damit einhergehenden ständigen Interaktion mit der Umwelt wird die kristallisierte Intelligenz auch oft als „Pragmatik der Intelligenz“ (Baltes, 1990) oder als „Softwarekomponente“ (Guthke & Beckmann, 2001) deklariert. Im Hinblick auf die entwicklungspsychologische Forschung zur Lebensspanne hält bspw. Lindenberger (2000) fest, dass die Leistungsfähigkeit der fluiden Intelligenz (z.B. die Merkfähigkeit oder die Geschwindigkeit der Informationsaufnahme und -verarbeitung) im Zuge der ontogenetischen Entwicklung nach dem 20. Lebensjahr kontinuierlich abgebaut wird, wohingegen die kristallisierte Intelligenz kaum ein Altersabbau aufweist. Im Gegensatz zur fluiden Intelligenz kann die kristallisierte Intelligenz durch die eigenständige Nutzung von Lerngelegenheiten und durch die Förderungsmöglichkeiten der Bildungseinrichtungen

im Laufe der Zeit sogar ausgebaut werden. In diesem Zusammenhang gewinnt auch das geflügelte Wort „lebenslanges Lernen“ an Bedeutung.

Angesichts der dargestellten Bewandnis des Wissens für das allgemeine Leistungsvermögen einer Person und der engen Beziehung zwischen Wissen und kristallisierter Intelligenz lässt sich abschließend auch der Bedeutungsgehalt der kristallisierten Intelligenz herleiten. Darauf aufbauend können das Wissen bzw. die kristallisierte Intelligenz als Schlüsselqualifikationen oder Grundvoraussetzungen für den beruflichen und privaten Erfolg gesehen werden (Kersting, 1999). Die Bestimmung der kristallisierten Intelligenz im Rahmen eines psychometrischen Testverfahrens kann demnach von großem Nutzen in der eignungsdiagnostischen Praxis sein (z.B. Beauducel & Kersting, 2002; Nettelnstroth, 2003; siehe auch Kapitel 3.4).

### **6.3 Die Interaktion zwischen fluider Intelligenz und kristallisierter Intelligenz – Der Einfluss von Anlage und Umwelt**

In der Einführung zum Thema Intelligenz wurde bereits festgestellt, dass die Intelligenz ein stabiles Persönlichkeitsmerkmal ist, welches über einen langen Zeitraum geringe Unterschiede in der Ausprägung aufweist. Die Intelligenz kann somit als Wegweiser für die langfristige Entwicklung gesehen werden. Als Grund für die Stabilität wurden anfangs genetische Erklärungen angeführt. Neuere Studien belegen hingegen, dass auch die kulturelle und gesellschaftliche Umwelt einen Einfluss auf die Intelligenz hat. Dieser Sachverhalt kommt auch in dem Modell der fluiden und kristallisierten Intelligenz von Cattell (1971a, 1987) bzw. in seiner Investmenttheorie zum Ausdruck. Die fluide Intelligenz wird dabei als Fähigkeit interpretiert, die außerhalb der Akkulturation gebildet wird. Auf der Grundlage der fluiden Intelligenz entfaltet sich die kristallisierte Intelligenz dagegen unter dem Einfluss der Akkulturation. In der Leistungsdiagnostik soll die fluide Intelligenz daher durch kulturfreie Tests (z.B. Tests zum räumlich visuellen Vorstellungsvermögen) und die kristallisierte Intelligenz durch kulturabhängige Wissens- und Sprachtests gemessen werden (Horn & Cattell, 1966).

Stern (2001) hält die Investmenttheorie von Cattell (1971a, 1987) prinzipiell für schlüssig, moniert aber gleichzeitig die Grundannahme, dass die fluide Intelligenz „die Grenzen absteckt, innerhalb derer Lerngelegenheiten zum Erwerb kristalliner Intelligenz genutzt werden können“ (S. 165). Die Kritik richtet sich damit vor allem gegen die

Annahme, dass die kristallisierte Intelligenz durch die Umwelt beeinflusst wird und die fluide Intelligenz hauptsächlich anlagebedingt ist. Folglich wird die Vermutung aufgestellt, dass die genetische Ausstattung einer Person nicht nur die fluide Intelligenz prägt, sondern auch direkt den Sprach- und Wissenserwerb und damit die kristallisierte Intelligenz beeinflusst. Umgekehrt kann auch die fluide Intelligenz durch die Umwelt beeinflusst werden (a.a.O.). Im Gegensatz zur Investmenttheorie stellt Stern (2001) fest, dass nicht-sprachliche Tests, die die fluide Intelligenz messen sollen, in demselben Ausmaß von der Dauer des Schulbesuchs abhängen, wie die sprachlichen Tests, die die kristallisierte Intelligenz messen. Dementsprechend wird die präzise Unterscheidung der beiden Faktoren bzw. die Kontamination fluider Intelligenz mit kulturfreien Fähigkeiten und kristallisierter Intelligenz mit kulturabhängigen Fähigkeiten infrage gestellt. Die damit einhergehende Konsequenz liegt sodann in der Auswahl der eignungsdiagnostischen Instrumente bzw. der verschiedenen Testarten zur Messung der Faktoren. Eine Veränderung der Aufgabenbatterie zur Erfassung der fluiden und kristallisierten Intelligenz wäre auf der Grundlage dieser Auffassung dann nicht auszuschließen. In diesem Kontext wird die Frage diskutiert, in welchem Ausmaß die fluide und kristallisierte Intelligenz und - damit zusammenhängend - auch die allgemeine Intelligenz durch die Umwelt oder die Anlage bedingt werden, bzw. inwieweit die Einflussfaktoren interagieren.

In diesem Bezugsrahmen werden drei Konzeptionen diskutiert (Hany, 2001):

1. Die *endogenistische Theorie* betrachtet die kognitive und psychische Entwicklung ausschließlich unter dem Aspekt der genetischen Vererbung.
2. Die *exogenistische Theorie* geht davon aus, dass die Umwelt einen maßgeblichen Einfluss auf die geistige Entwicklung hat. (Behaviorismus)
3. Die *interaktionistische Theorie* unterstellt, dass sowohl die Gene als auch die Umwelt in gleichem Maß für die geistige Entwicklung verantwortlich sind.

Die Grundlage für die erste Theorie bilden die Mendel'schen Forschungsanstrengungen bezüglich physischer und psychischer Merkmale. Die Forschungsergebnisse zeigen, dass verschiedene (stabile) menschliche Eigenschaften, wie z.B. die Augen-, Haarfarbe etc. vererbt werden können. Eine Erweiterung dieser Forschungen ist die Entdeckung der DNA, die als Träger für die Weitergabe der Erbinformationen verantwortlich ist. Heutzutage ist es unbestritten, dass alle stabilen Merkmale (s.o.) genetisch determiniert sind und dass für variierende Merkmale wie z.B. die geistige Leistungsfähigkeit, die Anlagen zumindest einen Einfluss haben müssen. Mendels Hauptaugenmerk lag vor

allem auf der Untersuchung von qualitativen Merkmalen. In Abgrenzung dazu wurden für quantitativ abgestufte Persönlichkeits- und Verhaltenseigenschaften polygene Vererbungen unterstellt (Hany, 2001). Man geht also davon aus, dass in einem Zusammenspiel von mehreren hundert Genen die Ausprägung kognitiver Merkmale bestimmt wird. Dieser Sachverhalt äußert sich darin, dass psychische Merkmale normalverteilt sind. Hany (2001) hält fest, dass aus ca. 80.000 Genen, die aus 3 Milliarden Basenpaaren der DNA auf den 46 Chromosomen des Menschen generiert werden, ca. 5% einen Einfluss auf die Produktion von Aminosäuren haben, die dann in Proteine umgewandelt werden und damit unseren Organismus gestalten. Ca. 1/3 dieser Gene sind für die Gehirnstruktur verantwortlich. Daher könnte man die Vermutung aufstellen, dass die verschiedenen Persönlichkeits- und Verhaltenseigenschaften in der Höhe dieser Maßangabe durch die Gene bzw. durch die Anatomie und Physiologie unseres Gehirns beeinflusst werden. Auf dieser Basis lässt sich aber zugleich festhalten, dass die Gene durch den verhältnismäßig geringen Anteil (1/3) an der Ausbildung der Gehirnstruktur keinen unmittelbaren Einfluss auf die Persönlichkeit haben. Sie bilden jedoch mit Hilfe unseres Nervensystems die physische Basis unserer Lernfähigkeit, unseres Gedächtnisses, unseres Temperaments und unserer neuronalen Erregbarkeit.

Die Bestimmung der biologischen Intelligenz gestaltet sich jedoch als schwierig. Es existieren zum gegenwärtigen Zeitpunkt noch keine humangenetischen Methoden zur Feststellung von „Intelligenzgenen“, so dass das fehlende Wissen einen Konsens über die Einflusshöhe der Gene bezüglich der Intelligenzleistung ausschließt. Stattdessen zieht man physiologische Methoden wie z.B. EEG- Messungen bei evozierten Potentialen oder den Glukoseverbrauch des Gehirns bei der Lösung von Intelligenzaufgaben heran. Des Weiteren dienen insbesondere psychologische Methoden (z.B. Testverfahren) dazu, die fluide Intelligenz messbar zu machen (Guthke & Beckmann, 2001). Die Tests zur Messung der erbbedingten (fluiden) Intelligenz haben demnach den Anspruch, möglichst bildungs- und kulturunabhängig zu sein. In diesem Zusammenhang sollen beispielsweise *Mental-Speed-Tests* (Reaktionszeitmessungen) eine Lösung darstellen, die nach Jensen (1993) Ausdruck neuronaler Effizienz sind.

Eine herausragende Rolle innerhalb der Vererbungswissenschaft kommt im Weiteren der Zwillingsforschung und den Adoptionsstudien zu. Im Rahmen der Zwillingsstudien sind die von Bouchard, Lykken, McGue, Segal, & Tellegan (1990) durchgeführten Untersuchungen von besonderem Interesse. Im Prinzip besagt die Hypothese folgendes: Je stärker die Intelligenzleistungen der eineiigen und getrennt aufgewachsenen Zwillinge

übereinstimmen, desto höher ist der Anteil der rein biologischen Ursachen an dieser Leistung. Die Stichprobe von Bouchard et al. (1990) umfasst dabei 56 eineiige erwachsene Zwillinge, die in den ersten Lebensjahren getrennt wurden, durchschnittlich 30 Jahre an anderen Orten aufwuchsen und erst als Erwachsene wieder persönlichen Kontakt hergestellt haben. Bouchard et al. (1990) errechneten einen durchschnittlichen IQ-Korrelationskoeffizienten bei eineiigen Zwillingen von  $r = .75$ . Das Ergebnis bedeutete, dass ein Großteil des geistigen Leistungspotentials genetischen Ursprungs sein muss. Die Adoptionsstudien stellten eine weitere Möglichkeit zur Erfassung von Umwelt- und Erbeeinflüssen auf die Intelligenz dar. Die Hypothese aus umwelttheoretischer Sicht ist, dass der Korrelationskoeffizient zwischen psychologischen Merkmalen (z.B. der IQ) der Kinder und der Adoptiveltern bei denen sie aufgewachsen sind, höher ist, als die Korrelationskoeffizienten zwischen den Kindern und den biologischen Eltern. Ein hoher Korrelationskoeffizient zwischen den Kindern und den biologischen Eltern bezüglich des IQ-Wertes würde umgekehrt für eine genetische Determination der geistigen Leistung sprechen (Amelang & Bartussek, 2001). Auf Basis der Daten von Scarr & Weinberg (1979) lagen die Korrelationskoeffizienten bezüglich der biologischen Familie stets über denjenigen der Adoptivfamilien. Die hohen Zusammenhänge innerhalb der biologischen Familien (im Durchschnitt ca.  $r = .45$ ) sprechen demnach für eine starke genetische Determination. Aufgrund der Tatsache, dass nicht die gesamte Varianz des Leistungsverhaltens erklärt wird, ist auch zu konstatieren, dass die vielfältigen Umweltbedingungen als Erklärung für die gemessene Leistung herangezogen werden können (Amelang & Bartussek, 2001). In welchem Ausmaß die Gene (Erbfaktoren) bzw. die Anlage die psychischen Merkmale wie die Intelligenz determinieren, lässt sich beim aktuellen Forschungsstand nicht eindeutig sagen und bleibt wahrscheinlich Gegenstand unterschiedener Kontroversen. Aufgrund der vielfältigen Probleme, die bspw. aus der Varianz von Umwelt- und Individualmerkmalen, der differenten Annahmen hinsichtlich der unterschiedlichen Modelle und der Vergleichbarkeit von Experimental- und Kontrollgruppe resultieren, erscheint es nicht sinnvoll, einen bestimmten Prozentsatz der Umwelt- und Erbvarianz anzugeben (Amelang & Bartussek, 1990). In einem kurzen Zwischenfazit kann man daher festhalten, dass sowohl die Anlagen als auch die unterschiedlichen Umweltfaktoren unsere kognitiven Fähigkeiten beeinflussen.

Obwohl die gegenwärtigen Befunde den Genen eine aktivere Rolle bezüglich der Intelligenzentwicklung zusprechen, gibt es auch gegenläufige Meinungen. Die Gegenseite verweist darauf, dass sich beständige Umwelteinflüsse auf Mutationen und auf die

Aktivierung von Genen hinsichtlich des Genotyps der Persönlichkeit und der Intelligenz auswirken (Gottlieb, 1991). Außerdem wird angeführt, dass ein Anstieg von drei IQ-Punkten pro Jahr auf soziale, kulturelle und bildungstechnische Ursachen zurückzuführen sei, weil das menschliche Genom durch Spontanmutation nicht in der Lage sei, sich so schnell zu verändern (Flynn, 1987). In eine ähnliche Richtung weist auch Hewitt (1990). Auf der Grundlage seiner Studien gibt er an, dass andauernde Umwelteinflüsse (z.B. Anforderungen des Berufs oder die schulische Ausbildung) die Intelligenzentwicklung stärker beeinflussen können als genetische bzw. anlagebedingte Ursachen. Eaves, Hewitt, Meyer und Neale (1990) halten dagegen fest, dass allein die genetische Basis des Organismus über die Lebensspanne eine größere Stabilität aufweist, als die sich ständig verändernde Umwelt und damit als Quelle für die Intelligenzentwicklung aufgefasst werden kann. In Bezugnahme auf das Konzept der fluiden und kristallisierten Intelligenz betont Hewitt (1990) den Effekt, dass durch andauernde und kumulative Umwelteinflüsse, wie bspw. das gewählte Bildungssystem, vor allem die kristallisierte Intelligenz nachhaltig beeinflusst werden kann und die Leistungsunterschiede hinsichtlich der Lese- und Konversationsgewohnheiten durch Training stabilisiert werden. Bezogen auf die entwicklungspsychologische Forschung zur Lebensspanne konnte der Effekt des Trainings auch bestätigt werden (Lindenberger, 2000). Auf diesen Zusammenhang wird auch im empirischen Teil der Arbeit Bezug genommen (siehe Kapitel 10.2).

Ausgehend von den dargestellten Kontroversen wird im Weiteren der Ansatz von Stern (2001) vorgestellt, der einige Analogien zu der Investmenttheorie von Cattell (1971a, 1987) erkennen lässt und an dieser Stelle als Beispiel für die Interaktion von fluider und kristallisierter Intelligenz hinsichtlich der beiden unterschiedlichen Einflussfaktoren dienen kann. Die Ergebnisse zur differentiellen Vorhersagbarkeit des Schriftspracherwerbs zeigen, dass eine Person, die gute phonologische Kompetenzen (tendenziell fluide und erbbedingte Kompetenzen) besitzt und in einem Kulturkreis mit Lautschrift lebt und aufgewachsen ist, weniger Schwierigkeiten damit haben wird, die Schriftsprache zu erlernen, als eine Person, die mit besseren grafisch-visuellen Fähigkeiten ausgestattet ist. Das Gleiche gilt bspw. auch für Personen, die gute grafisch-visuelle Fähigkeiten besitzen und in einem Kulturkreis mit Symbolschrift leben und aufgewachsen sind. Unabhängig von dem gegebenen Kulturkreis betont Stern (2001), dass die meisten anspruchsvollen akademischen Wissensgebiete sowohl auf sprachlichen, grafisch-visuellen als auch auf numerischen Zeichensystemen beruhen. Dies gilt für Naturwissenschaften



ebenso wie für die Ökonomie und die Sozialwissenschaften. Der Zugang zu diesen Gebieten setzt jedoch voraus, dass im Umgang mit allen drei Zeichensystemen hinreichende Kompetenzen entwickelt wurden. In diesem Zusammenhang spielt die Schule eine wichtige Rolle, die die Weiterentwicklung der Basiskompetenzen zu höheren geistigen Kompetenzen fördern kann. Dieser Sachverhalt verdeutlicht, dass die Grundausrüstung mit Basiskompetenzen durch die verschiedenen (kulturellen) Umwelteinflüsse nachhaltig beeinflusst wird bzw. erweitert werden kann. Intelligenzunterschiede im Sinne von sprachlichen, mathematischen und grafisch-visuellen Kompetenzen basieren demnach auf genetisch determinierten zerebralen Basiskompetenzen, die mit der Umwelt in vielfältiger Weise verknüpft sind (Stern, 2001).

### ***Wechselwirkung von anlagebedingten und sozokulturellen Einflussfaktoren***

Vor dem Hintergrund der Diskussion über den Einfluss von Anlage und Umwelt auf die intellektuellen Fähigkeiten stellt sich im Weiteren die Frage nach geschlechtsspezifischen Diskrepanzen.

In Anbetracht der intendierten empirischen Forschungsfragen (siehe Kapitel 8) wird zunächst der Fokus auf die Geschlechtsunterschiede hinsichtlich des Allgemeinwissens gelenkt. Anschließend wird die Befundlage über Ungleichheiten von kognitiven Fähigkeiten in das Zentrum der Überlegungen gerückt.

Aktuelle Forschungsbeiträge dokumentieren, dass Geschlechtsunterschiede evident sind und Anlass zu weiteren Untersuchungen geben. Zudem zeigte sich, dass die Disparitäten zugunsten des männlichen Geschlechts ausfallen. Darauf beziehend stellten Lynn, Wilberg und Margraf-Stiksrund (2004) in ihren Studien heraus, dass deutsche Gymnasiasten durchschnittlich mehr wissen als ihre Mitschülerinnen. In dieselbe Richtung weisen auch die Ergebnisse von Ackerman, Bowen, Beier und Kanfer (2001), die eine Stichprobe von amerikanischen Studienanfängern untersuchten oder Lynn und Irwing (2002) sowie Lynn, Irwing und Cammock (2002), die irische Studierende als Probanden mit einbezogen haben. Mit Rekurs auf den deutschen „Differenziellen Wissens-Test“ (DWT) berichtet Neidhardt-Wilberg (2005), dass die männlichen Probanden vor allem in naturwissenschaftlichen, technischen, sportlichen und politischen Wissensbereichen überwiegend überlegene Resultate erreichten als die weiblichen Probanden. Lediglich in den Fachgebieten Literatur, Kunst, Musik und Biologie fielen die Ergebnisse umgekehrt aus.

Die Unterschiede in den aufgelisteten Wissensdomänen konnten mit Einschränkungen auch in den Beiträgen von Wilberg und Lynn (1999), Evans, Schweingruber und Stevenson (2002) oder im Rahmen der PISA Studie 2000 (Prenzel, Carstensen, Rost & Senkbeil, 2002) festgehalten werden. Der Wissensvorsprung ließ sich ebenfalls in den drei kodierten Wissensskalen des I-S-T 2000 R (Amthauer et al., 2001) aufzeigen. Hinsichtlich der dargestellten Studien von Ackerman et al. (2001) als auch Lynn et al. (2002) kritisiert Neidhardt-Wilberg (2005) jedoch die zugrunde gelegte selektive Untersuchungsstichprobe und die Konfundierung mit heterogenen Bildungsklassen. Demzufolge lautete das erklärte Forschungsziel von Neidhardt-Wilberg (2005), die von Lynn et al. (2004) ermittelten Geschlechtsunterschiede anhand einer repräsentativen studentischen Stichprobe zu replizieren, um gehaltvollere und aussagekräftige Feststellungen treffen zu können. An der Untersuchung nahmen 64 männliche und 55 weibliche Studierende des Fachbereichs für Wirtschaftswissenschaften der Universität Lüneburg teil. Als Testinstrument wurde der „General Knowledge Test“ (GKT-D) eingesetzt, der 95 Testfragen aus 17 Themenbereichen inkludiert. Der Mittelwert der Gesamtstichprobe lag bei 43,9 Punkten ( $S=10,0$ ). Die Studenten konnten im Durchschnitt 46,6 ( $S=9,9$ ) Aufgaben richtig beantworten und die Studentinnen 41 Testitems ( $S=9,4$ ). Die Testergebnisse offenbarten weiterhin, dass die männlichen Probanden vor allem in den Themenbereichen Sport, Politik, Geschichte, Entdeckungen und Geographie bessere Testergebnisse erzielten. Obgleich das Gros der abgefragten Themengebiete weder weiblich noch männlich dominiert wird. Die Autorin konstatiert abschließend, dass die Geschlechtsunterschiede -angesichts der hohen Korrelation der Effektstärken einzelner Untertests als auch des Gesamttests- zufriedenstellend repliziert werden konnten. Resümierend zieht Neidhardt-Wilberg (2005) den Schluss, dass die aufgeführten Bereiche, wie Naturwissenschaften, Sport, etc. spezielle männliche Wissensdomänen zu sein scheinen. Der Wissensvorsprung konnte bereits in verschiedenen Ländern nachgewiesen werden. Ein genereller Vorteil seitens des männlichen Geschlechts (vgl. z.B. Baumert et al., 1998; Evans et al., 2002) muss indessen relativiert werden, da die Auswahl von Themenbereichen in den eingesetzten Testverfahren die männlichen Probanden von vornherein begünstigt. Vielmehr sollte bei zukünftigen Entwicklungen von Tests darauf geachtet werden, dass auch Wissensbereiche, wie z.B. Kunst oder Kultur berücksichtigt werden, um einen Ausgleich der vermeintlichen Differenzen herzustellen. Im Hinblick auf die eingangs gestellte Frage, ob das ungleiche Leistungsverhalten anlagebedingt oder soziokulturellen Ursprungs ist lässt sich auch im Rückgriff auf die Ausführungen

in Kapitel 6.2 sagen, dass die Ausprägung des Wissens vor allem durch die Umwelt und durch individuelle Fähigkeiten beeinflusst wird. Beispielsweise sorgen geschlechtsspezifische Stereotype dafür, dass Frauen intellektuelle Fragen oder Leitungsaufgaben nicht zugetraut werden und daher eine ambivalente Einstellung bezüglich ihrer eigenen Fähigkeiten aufbauen (vgl. Bischof-Köhler, 2002). Auf dieser Grundlage formuliert Neidhardt-Wilberg (2005) die folgende Definition von Allgemeinwissen, die die Schlussfolgerung erneut unterstreicht: „Das Allgemeinwissen kann als von mehreren Kohorten geteiltes Wissen definiert werden und umfasst sowohl Kenntnisse, die durch die Erziehung und Bildung intentional vermittelt, als auch diejenigen, die durch Massenmedien verbreitet und inzidentell aufgenommen werden (...). Aus allgemeinspsychologischer Sicht handelt es sich dabei um die semantischen Inhalte des deklarativen Gedächtnisses“ (S.145). Eine anlagebedingte Ursache für die Ungleichheiten wird auch darin gesehen, dass das männliche Geschlecht von Natur aus eher wettbewerbsorientiert und zielorientiert sind, wohingegen das weibliche Geschlecht tendenziell dazu neigt, das soziale Miteinander in das Zentrum des Interesses zu stellen (Lynn et al., 2002). Eine weiterführende Herangehensweise an die Problematik liefert die Intelligenzforschung, welche auf der Basis des Investmentmodells von Cattell (1987) davon ausgeht, dass die fluide Intelligenz, im Sinne eines kognitiven Leistungspotentials, den Aufbau von Wissensstrukturen (kristallisierte Intelligenz) steuert. Zuvor werden einige Studien dargestellt, die Hinweise auf die geschlechtsspezifischen Differenzierungen liefern. Amthauer et al. (1999) stellten in ihren Untersuchungen heraus, dass die Geschlechter hinsichtlich der verschiedenen Intelligenzfacetten des I-S-T 2000 klare Unterschiede aufweisen. Die weiblichen Probanden überzeugten bei den verbal kodierten Aufgabengruppen, wie Satzergänzungen, Wortauswahl oder Gemeinsamkeiten. Gleichzeitig dominierten die männlichen Probanden in der Bearbeitung von räumlich-visuellen oder mathematischen Aufgabenstellungen. Demnach räumten bereits Wottawa und Hossiep (1997) ein, dass die Berücksichtigung von speziellen Intelligenzaspekten von höherem Nutzen sind als ein Globalindex der Intelligenz. In Anlehnung an diverse Forschungsbeiträge aus den USA nimmt Asendorpf (2007) Bezug auf die geschlechtsspezifischen Differenzierungen in den genannten Domänen der Intelligenz und fasst die Ergebnisse der Studien zusammen. Es zeigte sich, dass im Bereich der räumlich-visuellen Kognition konsistente und frappierende Geschlechtsunterschiede zugunsten des männlichen Geschlechts bestehen. Gage und Berliner (1986) kommen in diesem Kontext zu der Erkenntnis, „dass es sich hierbei am ehesten um einen genetisch bedingten Unterschied handelt“ (S. 186).

Zu einer ähnlichen Konklusion kommen auch Linn und Peterson (1995), die in einer Metaanalyse entsprechende Ungleichheiten aufdeckten. Etwas anders sieht es bei den verbalen Fähigkeiten aus. Während die männlichen Probanden bei den räumlich-visuellen Aufgabenstellungen bessere Ergebnisse erzielen, schneiden die weiblichen Probanden bei verbal ausgelegten Fragestellungen besser ab (Gage & Berliner, 1986), wenngleich Hyde und Linn (1988) in einer Metaanalyse nur marginale Unterschiede entdeckten. Bei den Mathematikleistungen werden ebenfalls divergierende Befunde dargestellt, die keine endgültige Interpretation möglich machen. Auf der Grundlage einer Metaanalyse exponieren Hyde, Fennema und Lamon (1990), dass die Unterschiede zwischen männlichen und weiblichen Probanden sehr gering ausfallen und vielmehr altersbedingte Differenzierungen als sinnvoll erachtet werden sollten. Dennoch stellt sich die Frage, welche Faktoren die Vorteile in den verschiedenen Intelligenzbereichen erklären. Angesprochen werden beispielsweise neuropsychologische Effekte, die als *Lateralisierung* bezeichnet werden oder es werden Ursachen in den Geschlechtshormonen vermutet. Hierzu gibt es eine Vielzahl von Ansätzen, auf die im Einzelnen jedoch nicht weiter eingegangen werden kann. Stattdessen wird auf den Überblick von Bischof-Köhler (2006) verwiesen.

Festzuhalten bleibt, dass mannigfaltige Erklärungsansätze existieren, die zwar einen Erkenntnisfortschritt gewähren, aber kein abschließendes Urteil erlauben. Es gilt vielmehr zu konstatieren, dass Leistungsunterschiede multifaktoriell determiniert sind und Leistungsprognosen alleine auf der Basis der Geschlechtszugehörigkeit nicht zu vertreten sind.

Ein anderer Ansatzpunkt beschäftigt sich mit der Frage, ob das soziokulturelle Umfeld (wie z.B. migrationsbedingte Heterogenität und Mehrsprachigkeit) einen Einfluss auf die Ausprägung des Wissens besitzt. Insbesondere in der gegenwärtigen bildungspolitischen Debatte werden die Leistungsunterschiede von Schülern im Kontext der Migration und der Sprachenvielfalt betrachtet. Der Begriff Migrationshintergrund bedeutet in diesem Zusammenhang nicht zwangsläufig, dass ein oder zwei Elternteile im Ausland geboren wurden, sondern hinter der Größe „steht eine Vielzahl von Sprachen und damit eine breite kulturelle Vielfalt“ (Schwippert, Peek & Bolle, 2004, S. 253). Diese allgemein gehaltene Definition ist im Weiteren maßgeblich für die ausgewählten Studien, da die Operationalisierung bzw. die Erhebung des Migrationshintergrundes uneinheitlich gestaltet wird. Bei der Konstruktion von Fragebögen wird jedoch häufig die

Unterscheidung zwischen Deutsch als Muttersprache oder Fremdsprache gewählt, um somit einen Rückschluss auf die Migrationsgeschichte zu erhalten und um sensible persönliche Fragen zu umgehen.

Im Rahmen der MehrLe-Studie überprüften Schwippert und Peek (2003) das Leseverständnis von deutschen und russischen Schülern und stellten einen Zusammenhang zu schulischen Leistungen her. Das Leseverständnis wurde hierbei als Schlüsselqualifikation begriffen, welches für das Lernen in schulischen Bildungseinrichtungen sowie für das öffentliche und soziale Leben von großer Bedeutung ist. Hintergrund war die international angelegte Grundschul-Lese-Untersuchung IGLU/PIRLS (vgl. Bos, Lankes, Prenzel, Schwippert, Walther, & Valtin, 2003), die unzureichende Leistungen von Jugendlichen aus Migrationsfamilien dokumentieren. Innerhalb der Auswertungen wurden frappierende Leistungsunterschiede im Leseverständnis, in Mathematik und den Naturwissenschaften aufgedeckt. Insgesamt bekräftigten die darauf aufbauenden Analysen und Ergebnisse die eklatanten Unterschiede zwischen den beiden Gruppen und ließen den Schluss zu, dass die erworbenen Lesefähigkeiten die Schulleistungen beeinträchtigen. Die in der Sekundarstufe angesiedelte MARKUS-Studie (Helmke, Hosenfeld, Schrader & Wagner, 2002) konnte ebenfalls reduzierte Mathematikleistungen von Schülern mit Migrationshintergrund herausstellen. Ebenso konnten in der Hamburger Studie zur Lernauslage (LAU 5, Lehmann, Peek & Gänsfuß, 1997) verminderte Fähigkeiten in der Mathematik, im Sprach- und Leseverständnis sowie dem Schreiben (Rechtschreibung) bei Jugendlichen mit Migrationshintergrund festgestellt werden. In dieselbe Richtung weisen auch die Befunde der Hannoverschen Grundschulstudie, bei der ca. 1700 Schüler mit einbezogen wurden. Es zeigte sich, dass die Rechtschreibtest- und die Mathematikleistungen erheblich voneinander abwichen. Für die Ursachenbeschreibung der Schulleistungsvarianz wurden zusätzlich Regressionsberechnungen der Prädiktoren Fähigkeitsselbstkonzept, kognitive Fähigkeiten, Bildungsorientierung des Elternhauses, Familiensprache und Geschlecht analysiert. Insgesamt konnten 37% der Varianz aufgeklärt werden, wobei die kognitiven Fähigkeiten als auch die Bildungsorientierung der Eltern erklärungsstarke Prädiktoren darstellten (Tiedemann & Billmann-Mahecha, 2004). Die Autoren konstatieren, dass die Familiensprache „ein nicht unwesentliches Bestimmungsstück der Schulleistungen [ist]. Kinder, deren Familiensprache nicht Deutsch ist, zeigen gegenüber Kindern mit deutscher Familiensprache konsistent verminderte Leistungen“ (S. 275). Obgleich Tiedemann und Billmann-Mahecha (2004) einräumen, dass schlechtere Schulleistungen der Jugendlichen mit Migrationshinter-

grund nicht nur auf Sprachdefizite zurückzuführen seien, sondern vielmehr durch multiple und kumulative Beeinträchtigungen begründet werden müssen. Hierzu zählen unter anderem eine höhere Arbeits- und Perspektivlosigkeit der Eltern, eine verminderte Bildungsorientierung, ein geringerer sozioökonomischer Status aber auch Versäumnisse der Schul- und Bildungspolitik. Obwohl der Einwand erhoben werden kann, dass die aufgeführten Grundschulstudien vorerst keinen klaren Bezug zur Zielpopulation der Berufseinsteiger erkennen lassen, sind die Erkenntnisse der Primarstufe insoweit wegweisend, als dass die Grundschule eine Schlüsselstellung einnimmt, in der einerseits die Entwicklung von grundlegenden Lesekompetenzen stattfindet und andererseits ein eindeutiger Zusammenhang zwischen der sozialen Herkunft der Lesefähigkeiten und der Leistungen am Ende der Sekundarschulzeit hergestellt werden kann. Vergleichbare Erklärungsansätze in der Sekundarstufe I und II verfolgen Baumert und Schümer (2002), die im Zuge der PISA Studie 2000 die „familiären Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb im nationalen Vergleich“ untersuchen. Bereits im Vorgriff prononcieren die Autoren, dass „soziale Lebensverhältnisse (...) mit Bildungsaspirationen, Bildungsbeteiligung, den erbrachten Schulleistungen und Bildungsabschlüssen und schließlich auch mit Lebensplänen und Lebenschancen“ (S. 159) kovariieren. Zunächst wird die Aussage pointiert, dass gerade die Lesekompetenz als Basisqualifikation interpretiert wird, die in modernen Gesellschaften eine Grundvoraussetzung für das stetige und systematische Lernen darstellt und nicht durch Leistungsstärken in anderen Fächern ausgeglichen werden können. Ausgehend von dieser Annahme wird die Lesekompetenz von Jugendlichen mit und ohne Migrationshintergrund gegenüber gestellt. Die Ergebnisse zeigen, dass die Differenzen der Testwerte gravierend sind. Selbst in den Fächern, die als weniger „sprachlastig“ gelten, wie Naturwissenschaften oder Mathematik, wurden eindeutige Ungleichheiten gefunden. Folglich gehen Baumert und Schümer (2002) sowie Baumert und Artel (2002) davon aus, dass eine positive Korrelation zwischen den Sprachfähigkeiten und dem Erwerb von Kenntnissen in den genannten Fächern besteht. Die Resultate wurden grafisch visualisiert (S. 220). Aufgrund der Befundlage ist es ebenso nachvollziehbar, dass die Bildungsbeteiligung von Jugendlichen aus Migrationsfamilien an Gymnasien und höheren Bildungsschulen geringer ist (Baumert & Schümer, 2002, S. 195). Alleine die Lesekompetenzen oder den Schriftsprachenerwerb als Begründung für schlechte Schulleistungen anzuführen halten Baumert und Schümer indes für zu kurzichtig.

Daraufhin diskutieren sie über den Einfluss von soziokulturellen Disparitäten wie Sozialschichtzugehörigkeit und institutionelle Rahmenbedingungen, welche als Indikatoren für die Differenzen gesehen werden können. Die oben dargestellten einfachen bivariaten Zusammenhänge sind demnach irreführend und verschleiern ein komplexes Gefüge von Kontextbedingungen. Vielmehr ist der Kompetenzerwerb mit unterschiedlichen soziokulturellen, gesellschaftlichen und institutionellen Umweltfaktoren konfundiert.

### *Fazit*

Es lässt sich festhalten, dass sowohl die genetische Grundausstattung bzw. die Anlage einer Person und die Bedingungen der Umwelt aufgrund der zitierten Befunde als wirkungsvolle Faktoren auf die Entwicklung der Intelligenz angeführt werden können. Grundsätzlich herrscht auch Einigkeit darüber, dass die Anlage eines Menschen erst durch die Stimulation aus der Umwelt ihre Wirksamkeit entfaltet und damit besondere Fähigkeiten herauskristallisiert werden. In Anlehnung an Ceci (1996) konstatiert Stern (2001), „dass die kognitive Architektur des Menschen aus vielen Einzelfähigkeiten besteht, die sich im Laufe eines Entwicklungs- und Lernprozesses zu übergeordneten Fähigkeiten kristallisieren“ (S.166). Auf dieser Grundlage ist die Unterscheidung zwischen fluider und kristallisierter Intelligenz von großer Bedeutung, weil sie beide Einflussfaktoren berücksichtigt. Obgleich die begriffliche Unterscheidung den Eindruck macht, dass man die Intelligenzkomponenten sauber trennen könnte, gibt es vielfältige, gegenseitige Überlappungen, so dass die fluide Intelligenz und die kristallisierte Intelligenz nicht unabhängig voneinander betrachtet werden sollten. Dementsprechend spielt die Unterscheidung eine wesentliche Rolle, um differenzierte Aussagen über die Intelligenz treffen zu können. Eine ganzheitliche Sichtweise wird darüber hinaus durch die Zusammenfassung der fluiden und kristallisierten Intelligenz auf der dritten Ebene in Form der allgemeinen Intelligenz gewährleistet. Im Hinblick auf den empirischen Teil der Arbeit bietet der I-S-T 2000 R eine geeignete Möglichkeit, die fluide und die kristallisierte Intelligenz zu verbinden und zu operationalisieren.

Um die oben dargestellte Kontamination der fluiden Intelligenz mit vorwiegend figuralen Fähigkeiten und die kristallisierte Intelligenz mit verbalen Fähigkeiten zu reduzieren, wurde im I-S-T- 2000 R die Materialfacette für verbale, numerische, und figurale Fähigkeiten für beide Faktoren integriert (Amthauer et al., 2001).

Es wird erneut angemerkt, dass der eingesetzte Wissenstest START-W, der weitgehend dem Erweiterungsmodul des I-S-T 2000 R entspricht, ebenfalls die Materialfacetten enthält und die Möglichkeit besitzt, die kristallisierte Intelligenz abzubilden.

Abschließend wird noch einmal betont, dass auch künftig die Frage nach den Auswirkungen der verschiedenartigen Umwelteinflüsse auf das Wissen in der Wissenschaft von großer Bedeutung ist. Aus diesem Grund wird auf der Grundlage empirischer Daten, den noch nicht hinreichend geklärten Fragen nachgegangen, ob der gewählte Bildungsgang, das Alter, das Geschlecht, das Bildungsniveau oder der soziale Hintergrund etc. einen Einfluss auf die Wissensausprägung bzw. die kristallisierte Intelligenz haben (siehe Kapitel 10.2 und 11.2).

## **7 Überleitung zum empirischen Teil**

Im ersten Abschnitt des theoretischen Teils der Dissertation wurde aufgezeigt, dass Intelligenz und Wissen als Schlüsselqualifikationen gesehen werden, die gleichermaßen bedeutsam sind für die Wirtschaft wie auch für die Gesellschaft. Wirtschaftsunternehmen, Einrichtungen der Gesellschaft verlangen von den Mitarbeitern Kompetenzen, die weit über die Fachkenntnisse hinausgehen. Unter diesem Aspekt kommt der Auswahl von Berufseinsteigern und auch Mitarbeitern in Unternehmen allgemein eine hohe Wichtigkeit zu. Vor dem Hintergrund der Ansprüche verschiedener Institutionen, dem breiten Anwendungsbereich, der hohen Qualität psychologischer Testinstrumente, der mannigfaltigen Studien und den dargestellten Zusammenhängen zwischen Allgemeinwissen (kristallisierter Intelligenz) und Schul- Ausbildungs- sowie Berufserfolg wird die Funktionalität sowie die Forderung nach wissenschaftlich fundierten Testverfahren offensichtlich und im Besonderen hervorgehoben. Der START-W-Wissenstest bietet die Möglichkeit, das Allgemeinwissen einer Person abzubilden und liefert zudem einen Indikator für zukünftiges Lernverhalten, wie es von der Praxis gefordert wird.

Im zweiten Abschnitt des theoretischen Teils der Dissertation lagen die Konstrukte Intelligenz und Wissen im Zentrum der Betrachtung. Anhand der detaillierten Beschreibung der Konstrukte wurde das Ziel verfolgt, ein zweckdienliches, theoretisches Fundament für den empirischen Teil zu schaffen. Die Vermittlung eines tieferen Verständnisses über die zu messenden Konstrukte stand dabei im Vordergrund. Demnach wurde in Kapitel 4.2 das Konstrukt-Verständnis der Intelligenz hervorgehoben, dass durch



seine Erweiterungsfähigkeit die gesamte Breite des Bedeutungsspektrums berücksichtigt. Darauf aufbauend wurden die Bedeutungsbereiche in ihren zahlreichen Facetten spezifiziert und die akademische Intelligenz als theorieübergreifender Bedeutungsbe- reich fokussiert. Die dargestellten Strukturtheorien tragen in diesem Zusammenhang dazu bei, die breite Vielfältigkeit der Intelligenz zu erfassen. Darüber hinaus bilden sie auch die theoretische Grundlage für sämtliche Intelligenztests. Das Hauptaugenmerk lag auf der Strukturtheorie von Cattell (1971a, 1987), in der eine Unterscheidung zwischen fluider und kristallisierter Intelligenz postuliert wird und daher einen Bezugspunkt zum Wissen hergestellt wird. Im Hinblick auf die besondere Akzentuierung der kristallisier- ten Intelligenz wurde in Kapitel 5 das Themenfeld Wissen näher betrachtet. Hierbei wurden die beiden Wissensformen dargestellt, außerdem fanden grundlegende wissens- psychologische Aspekte zum Wissenserwerb, zur Wissensrepräsentation und der Wis- sensanwendung ihren Niederschlag in dieser Arbeit. Angesichts der Relevanz für den empirischen Teil wurde zusätzlich der Wissenserwerb bzw. das Lernen in schulischen Bildungseinrichtungen betrachtet und damit ein Bezug zur Stichprobenauswahl herges- tellt. Außerdem wurden verschiedene Einflussfaktoren auf die Wissensausprägung so- wie die Interaktion zwischen Umwelt, Anlage und Wissen dargestellt und diskutiert (vgl. Kapitel 5.3, 6.2 und 6.3). Aufgrund der Verkettung von kristallisierter Intelligenz mit dem Wissen im Kontext der Strukturtheorie von Cattell (1971a, 1987) und dem Testverfahren (I-S-T-2000 R) von Amthauer et al. (2001) bzw. Liepmann et al. (2007) wurde in Kapitel 6 das Beziehungsgeflecht und die Bedeutung gesondert hervorgeho- ben, erläutert und erörtert. In den Darstellungen wurde herausgestellt, dass die (kristalli- sierte) Intelligenz und das Wissen in einem komplexen und interdependenten Verhältnis stehen und in ihrem Zusammenspiel das Leistungsverhalten einer Person bestimmen. Auf der Grundlage einer Vorversion des Testverfahrens zur Erfassung des Allgemein- wissens (START-W), der zu einem nicht unerheblichen Teil aus Testitems des Erweite- ringsmoduls (I-S-T 2000 R) besteht, werden im empirischen Teil zunächst Item- und Reliabilitätsanalysen durchgeführt. Anhand der berechneten Kennwerte und der dazu- gehörigen Evaluierung sollen, entsprechend der unten aufgeführten Struktur, Skalen gebildet werden. Die Bildung von Skalen und die Selektion von geeigneten Testitems zur Bestimmung des verbal, numerisch und figural kodierten Wissens ist dabei eine wichtige Voraussetzung für die Prüfung der intendierten Struktur des Testverfahrens. Die gebildeten Skalen werden für weitere Berechnungen und deren Interpretationen herangezogen.

Im Rahmen des theoretischen Teils wurden die Auswirkungen von Umweltfaktoren auf das Wissen diskutiert und die Grundannahme bekräftigt, dass die Wissensentwicklung einer Person durch kulturelle Einflüsse geprägt wird. Die Forschungsannahmen in dieser Arbeit beziehen sich demnach auf die zugrunde gelegte erste Forschungsfrage, ob das Alter, die Muttersprache, das Geschlecht und das Bildungsniveau einen Einfluss auf das Wissen haben. Abschließend werden die Ergebnisse interpretiert, bewertet und diskutiert.

In einem zweiten Teil wird die Vorversion hinsichtlich ihrer kriterienbezogenen Validität untersucht, bei der die prognostische Validität von erhöhter Relevanz ist. Daran anschließend wird die Konstruktvalidität auf Basis von theoretischen, sachlogischen, begrifflichen und empirischen Untersuchungen überprüft. Zudem wird der Zusammenhang zwischen fluider Intelligenz und kristallisierter Intelligenz analysiert und diskutiert. Die dahinterstehende Intention dieses Vorgehens ist es, gemäß den oben aufgeführten Kriterien der DIN-Norm 33430, die Erfüllung der angegebenen Testgütekriterien zu evaluieren.

## **8 Empirische Forschungsfragen und Forschungsannahmen**

Vor dem Hintergrund der theoretischen Ausführungen zur eignungsdiagnostischen Bedeutung von Wissenstests, zu Intelligenzfacetten und dem Konstrukt Wissen werden die folgenden Forschungsfragen abgeleitet.

Zu Beginn der Formulierung von Fragestellungen und der Ergebnisdarstellung stehen im empirischen Teil Reliabilitätsanalysen auf Item- und Skalenebene im Vordergrund der Betrachtung. Im Zuge einer qualitativ hochwertigen Testentwicklung, die auf den Gütekriterien der DIN 33430 basiert, liegt das übergeordnete Ziel der Reliabilitätsschätzung darin begründet, die Zuverlässigkeit bzw. die Genauigkeit des Tests zu überprüfen, mit dem ein spezielles Merkmal erfasst wird. Hiermit soll eine Optimierung der Reliabilität und Normierung erreicht werden (Lienert & Raatz, 1998). Die Aufgabenanalyse bezeichnet in diesem Zusammenhang die Evaluation der psychometrischen Qualität der Testitems. Hierzu werden im Folgenden der Schwierigkeitsindex, die Trennschärfe und die Homogenität der Aufgaben untersucht. Auf der Grundlage der Aufgabenanalysen und der anschließenden Selektion von geeigneten Items werden inhaltliche Skalen (verbale, numerische und figurale Inhaltsskalen) gebildet, die wiederum auf ihre

Homogenität geprüft werden. Darauf aufbauend wird die Retest-Reliabilität und die innere Konsistenz nach der Testhalbierungsmethode (Split-half Reliabilität) kontrolliert.

Nach dieser Vorgehensweise ergeben sich die nachstehenden Fragestellungen:

Lassen sich auf Basis der intuitiven Zuordnung der Test-Items zu den verbalen, numerischen und figuralen Materialfacetten Skalen bilden? Genügen die Reliabilitätsschätzungen der Skalen den wissenschaftlich anerkannten Kriterien? Welche psychometrischen Kennwerte weisen die ausgewählten Test-Items auf?

## **Forschungsfragen- Teil 1: Einflussfaktoren auf die Wissensakkumulation**

### ***Forschungsfragen 1.1 bis 1.4***

Lassen sich unterschiedliche Ausprägungsgrade des Allgemeinwissens in Abhängigkeit des Alters (1.1), der Muttersprache (1.2), des Geschlechts (1.3) und des Bildungsniveaus (1.4) empirisch feststellen?

Cattell (1963) stellte in seiner Untersuchung fest, dass die Entwicklung der fluiden und kristallisierten Intelligenz über die Lebensspanne in unterschiedlicher Weise verläuft. Die fluide Intelligenz erreicht ihr Maximum etwa im Alter von ca. 25 Jahren und nimmt danach erkennbar ab. Die kristallisierte Intelligenz hingegen kann, abhängig von weiteren Erfahrungen und Lebensumständen, bis ins hohe Alter steigen. Auch bezüglich der entwicklungspsychologischen Forschung zur Lebensspanne nach Lindenberger (2000) wird vermutet, dass das Alter ebenfalls einen Einfluss auf die Höhe des Wissensstandes hat. Der Einfluss wird damit begründet, dass die Individuen im Laufe der Zeit mehr Wissen akkumulieren.

Darüber hinaus soll der Einfluss von verschiedenen Muttersprachen auf die Wissensausprägung untersucht werden. Vor dem Hintergrund, dass die kristallisierte Intelligenz als kulturell und gesellschaftlich geprägtes Wissen verstanden wird, liegt die Annahme nahe, dass der Umfang des Allgemeinwissens zwischen den Probanden differiert, die eine unterschiedliche Muttersprache haben. Aufgrund der Tatsache, dass das verbal ausgelegte Testverfahren auf Deutsch dargeboten wird und einige Testitems die deutsche Geschichte bzw. die deutsche Kultur behandeln, ist anzunehmen, dass die Probanden, die Deutsch als Muttersprache beherrschen, bessere Ergebnisse zeigen. Zum anderen kann auch der unterschiedliche kulturelle Hintergrund als Erklärung herangezogen werden.

Angesichts der beispielhaften Forschungsbeiträge von Baumert, Bos und Waterman (1998), Wilberg und Lynn (1999), Evans, Schweingruber und Stevenson (2002), Ackerman, Bowen, Beier und Kanfer (2001), Lynn, Wilberg und Margraf-Stiksrud (2004), argumentiert Neidhardt-Wilberg (2005), dass „Naturwissenschaften, Sport, Politik und Geschichte (...) dezidiert männliche Wissensdomänen [darstellen, und] der männliche Wissensvorsprung in diesen Bereichen bereits in mehreren Ländern dokumentiert [wurde]“ (S.154). Da der Großteil des Wissenstests sich aus den genannten Themenbereichen zusammensetzt, wird die Forschungsfrage aufgeworfen, ob unterschiedliche Wissensausprägungen auf der Skalenebene und im Gesamtscore in Abhängigkeit des männlichen und weiblichen Geschlechts bestehen.

Gemäß den Ausführungen im theoretischen Teil der Arbeit, wird die kristallisierte Intelligenz als investierte fluide Intelligenz verstanden, die aufgrund von verschiedenen Akkulturationseinflüssen einen engen Bezug zum Wissen aufweist (Cattell, 1987). Erwartet wird, dass vor diesem Hintergrund die Höhe des Ausprägungsgrades der kristallisierten Intelligenz bzw. des Wissens auch durch institutionalisierte Bildungseinrichtungen wie die Schule erklärt wird. Ein bestätigendes Ergebnis lieferten Beier und Ackerman (2001), die in ihren Studien eine Korrelation von  $r = .44$  zwischen dem Bildungsniveau und Wissen ermittelten. Das Allgemeinwissen gibt damit gleichermaßen einen Hinweis, inwieweit die Schüler die Bildungsmöglichkeiten bzw. die Lerngelegenheiten für sich nutzen. Auf einer höheren Generalitätsebene soll daher der Frage nachgegangen werden, welche Bedeutung die verschiedenen Umwelteinflüsse auf die kristallisierte Intelligenz bzw. das Wissen haben. Daraufhin können die folgenden Forschungsannahmen formuliert werden:

#### *Forschungsannahme 1.1*

Zwischen den verschiedenen Altersgruppen der Probanden besteht ein signifikanter Unterschied in der Ausprägung des Wissensstandes. Es wird angenommen, dass die Höhe des Wissensstandes mit steigendem Alter zunimmt.

#### *Forschungsannahme 1.2:*

Es wird erwartet, dass zwischen den Probanden, die Deutsch als Muttersprache beherrschen (hier: Muttersprachler-D) und den Probanden, die Deutsch als Fremdsprache gelernt haben (hier: Muttersprachler-F), ein signifikanter Unterschied in den Testergebnissen besteht. Es wird angenommen, dass die Muttersprachler-D im Durchschnitt bessere Ergebnisse erzielen.

*Forschungsannahme 1.3:*

Es lässt sich vermuten, dass zwischen den weiblichen und männlichen Probanden eine signifikante Differenz in der Wissensausprägung besteht. Es wird erwartet, dass die männlichen Testteilnehmer vorwiegend bessere Ergebnisse erreichen.

*Forschungsannahme 1.4:*

Es wird erwartet, dass die Ausprägung des Wissens zwischen den verschiedenen Bildungsgängen differiert. Es wird angenommen, dass die Höhe des Wissens mit dem Bildungsniveau ansteigt.

**Forschungsfragen- Teil 2: Kriteriumsvalidität*****Forschungsfrage 2.1: Übereinstimmungsvalidität***

Kann der verwendete Wissenstest hinsichtlich des Außenkriteriums der Schulnoten in den Fächern Mathematik, Deutsch, Englisch, Kunst sowie der angegebenen Durchschnittsnote als valide eingestuft werden?

Die kriterienbezogene bzw. prädiktive Validität zählt zu den wichtigsten Gütekriterien eines Testverfahrens, wobei die Güte von dem verwendeten Außenkriterium abhängt. In der Regel besitzen die Schulnoten eine hohe psychometrische Qualität und werden daher häufig als Kriterium herangezogen (Jensen, 1998). Es gilt jedoch zu beachten, dass die angeführten Noten keine einwandslosen objektiven Kriterien darstellen, die auch an soziale Bezugsnormen und anderen Einflussfaktoren orientiert sind (Trautwein, Köller, Lehmann & Lüdtke, 2007). Dennoch spiegeln die Schulnoten die Leistungen einer Person wider, die von qualifizierten Lehrkräften anhand von verschiedenen Leistungskriterien über eine längere Zeitspanne ermittelt worden sind. Außerdem stellen die Noten einen Indikator für die Allgemeinbildung dar, die in den Bildungsinstitutionen über die Ausbildungsdauer vermittelt werden soll. Aus diesem Grund können die Schulnoten als relevantes inhaltliches Kriterium zur Validierung des Wissenstests verwendet werden. Für diesen Zusammenhang wird die Forschungsannahme aufgestellt, dass gute Schulleistungen auch mit einem guten Testergebnis kovariieren. Aufgrund der plausiblen Annahme, dass im Fach Mathematik vor allem numerisches Wissen im Sinne eines guten Zahlenverständnisses von Bedeutung ist und für die Fächer Deutsch und Englisch vor allem verbale Fähigkeiten relevant sind, lassen sich die dritte und vierte Forschungsannahme ableiten. Es sei angemerkt, dass die Testergebnisse und die Schulnoten unterschiedlich gepolt sind und daher negative Zusammenhänge erwartet werden.

*Forschungsannahme 2.1.1:*

Es wird ein negativer Zusammenhang zwischen dem Gesamtscore, den einzelnen Skalenscores und den aufgeführten Schulnoten (Deutsch, Englisch, Mathematik, Kunst, Musik, Durchschnittsnote) in der Gesamtstichprobe erwartet.

*Forschungsannahme 2.1.2:*

Es wird überdies angenommen, dass der Gesamtscore des Wissenstests auf einem signifikanten Niveau negativ mit sämtlichen Noten korreliert. Außerdem wird erwartet, dass die aufgeführten Skalenscores mit der Durchschnittsnote signifikant negativ korrelieren.

*Forschungsannahme 2.1.3:*

Es wird die Hypothese aufgestellt, dass ein negativer Zusammenhang zwischen dem Testscore der numerisch kodierten Wissensskala und der Note im Fach Mathematik besteht.

*Forschungsannahme 2.1.4:*

Es wird erwartet, dass ein signifikanter, negativer Zusammenhang zwischen der verbal kodierten Wissensskala und den Noten im Fach Deutsch und Englisch besteht.

*Forschungsannahme 2.1.5:*

Es wird vermutet, dass ein signifikanter, negativer Zusammenhang zwischen der figural kodierten Wissensskala und der Note im Fach Kunst besteht.

***Forschungsfrage 2.2: Prognostische Validität***

Kann mit Hilfe des eignungsdiagnostischen Verfahrens bzw. des Wissenstests START-W der Erfolg auf der Zielposition (hier Ausbildungserfolg), für dessen Vorhersage er eingesetzt wurde, auch tatsächlich vorhergesagt werden?

Die prognostische Validität bzw. die Gültigkeit der Vorhersage spielt bei der Entwicklung von Testverfahren eine wichtige Rolle. Folglich soll eine Vorhersage zukünftigen Verhaltens bzw. Leistung ermöglicht und bestimmt werden. Insbesondere im Hinblick auf die eignungsdiagnostische Verwendung des Tests wird die Bedeutung dieses Gütekriteriums eminent. Die unten aufgeführten Forschungsannahmen stützen sich unter anderem auf die oben bereits dargestellten Studien verschiedener Autoren. Beispielsweise berichten Dye, Reck und Mc Daniel (1993) in diesem Kontext von einer durchschnittliche Validität in Höhe von  $r = .45$  zwischen Wissenstests und dem späteren Berufserfolg. Schmidt und Hunter (1998) geben auf der Grundlage einer Metaanalyse

einen Validitätskoeffizienten von .48 an und unterstützen somit diesen Befund. In dieselbe Richtung weisen auch die Befunde von Fürntratt (1969) Jungkunz und Bodinet (1989), Kersting (1999a,b), Schmidt und Hunter (2000), Hossiep und Schulte (2007), die in ihren Studien herausstellen, dass Allgemeinwissenstests sehr gute Prädiktoren für den Schul-, Ausbildungs- und Berufserfolg sind.

Als Außenkriterium (Ausbildungserfolg) werden die theoretischen Durchschnittsnoten des ersten Ausbildungsjahres im Bereich der Gesundheitsberufe herangezogen. Die psychometrische Qualität der verwendeten Noten ist vergleichbar mit den oben dargestellten Schulnoten. Demzufolge wird die nachstehende Forschungsannahme formuliert.

*Forschungsannahme 2.2:*

Es wird erwartet, dass ein signifikant negativer Zusammenhang zwischen den Testwerten und der Durchschnittsnote des Abschlusszeugnisses im ersten Ausbildungsjahr besteht.

### **Forschungsfragen- Teil 3: Konstruktvalidität**

*Forschungsfrage 3.1: Konvergente Validität*

Im Rahmen der Konstruktvalidität wird zunächst die konvergente Validität analysiert und im Anschluss die faktorielle Struktur des START-W abgebildet (faktorielle Validität). Zur Bestimmung der konvergenten Validität wird die etablierte Kurzform des Bochumer Wissenstests (BOWIT, Hossiep & Schulte, 2008) herangezogen. Das Testverfahren hat eine vergleichbare Theorieorientierung und wurde ebenfalls an den Gütekriterien der DIN-Norm 33430 orientiert. Es erfüllt somit einen ähnlichen Reliabilitäts- und Validitätsanspruch und ist damit ein adäquates Instrument für die Untersuchung. Folglich wird der Zusammenhang zwischen den konvergenten Konstrukten (erworbenes Wissen bzw. kristallisierte Intelligenz) überprüft und die Frage aufgeworfen, inwieweit die psychometrischen Kennwerte den Gütekriterien genügen.

*Forschungsannahme 3.1:*

Ausgehend von den theoretischen Vorüberlegungen und der darauf aufbauenden Konzeption des START-W Wissenstests ist davon auszugehen, dass die Korrelation mit dem BOWIT in substantiellem Umfang positiv ausfällt. Genügen die Validitätsschätzungen den wissenschaftlich anerkannten Kriterien?

***Forschungsfrage 3.2: Faktorielle Validität***

Lässt sich die intendierte Struktur des Allgemeinwissenstests START-W auf Basis von empirischen Daten aufzeigen?

Die Konstruktvalidität gibt an, ob das Verfahren auch tatsächlich das Konstrukt erfasst, welches es zu messen vorgibt. Die Bedeutung der Konstruktvalidierung liegt also weniger in der praktisch-diagnostischen Anwendbarkeit als vielmehr in der theoretischen, sachlogischen und begrifflichen Erklärung dessen, welche Aspekte das entwickelte Verfahren erfasst. Das Ziel der Konstruktvalidierung besteht demnach in der theoretischen und psychologischen Analyse der dem Test zugrunde liegenden Merkmale bzw. Konstrukte.

Das Erweiterungsmodul des I-S-T 2000 R besteht aus einem allgemeinen Wissenstest, der verschiedene Aspekte des erworbenen Wissens in unserer Kultur erfassen soll. In diesem Test werden Fragen zu den verschiedenen Themengebieten gestellt, die durch die Inhaltsfacetten (verbales, figurales und numerisches Wissen) kodiert werden. Demnach wird der Test zur Erfassung des verbal, numerisch und figural kodierten Wissens eingesetzt, wobei als Summenwert daraus das Wissen sowie über Faktorwerte das Wissen/gc abgebildet werden kann. Die Auswahl der Testitems des zu validierenden Allgemeinwissenstests START-W basiert einerseits auf dem oben dargestellten Erweiterungsmodul des I-S-T 2000 R und andererseits auf neu entwickelten Testitems. Aufgrund der nicht unerheblichen Überschneidungen ist davon auszugehen, dass der Wissenstest die kristallisierte Intelligenz in weiten Bereichen abbilden kann und die wisens-thematischen sowie die Inhaltsfacetten nachgewiesen werden können.

***Forschungsannahme 3.2:***

Es wird angenommen, dass der Allgemeinwissenstest über Faktorwerte die kristallisierte Intelligenz abbilden kann und dass die Struktur des elaborierten Erweiterungsmoduls (I-S-T 2000 R) im Allgemeinwissenstest START-W repliziert wird.

***Forschungsfrage 3.3: Zusammenhang zwischen fluiden Intelligenz (gf) und kristallisierten Intelligenz (gc) – Konvergente Validität***

Auf der Basis des entwickelten Strukturgleichungsmodells soll der Zusammenhang zwischen fluiden Intelligenz und Wissen untersucht werden. Vor dem Hintergrund des theoretisch fundierten Zusammenhangs wird der Frage nachgegangen, ob die postulierte Beziehung empirisch bestätigt werden kann und welche psychometrischen Kennwerte das Modell aufweist.



In Cattels Konzept der fluiden und kristallisierten Intelligenz (1987) wurde bereits der Zusammenhang zwischen fluiden Intelligenz und Wissen untersucht. So formuliert Cattell (1987): „This correlation, which hovers around  $r=0.4-0.5$  (...) is that the acquisition of the crystallized ability skills (...) depends partly on the level of insightful „fluid“ ability and partly on hours spent in school, etc“ (S.116). Diese Sichtweise wird in seiner Investmenttheorie detailliert dargestellt (nähere Ausführungen siehe Kapitel 4.4.4 und Kapitel 6). In einer Untersuchung von Amthauer, Brocke, Liepmann und Beauducel (2001) wurde der Zusammenhang zwischen dem Faktorwert der fluiden Intelligenz mit dem Wissenstest HAWIE-R (Tewes, 1991) untersucht. Die Autoren berichten von einem Korrelationskoeffizienten in Höhe von  $r=.31$ . In einer Untersuchung von Hunter und Schmidt (1989) konnte der Zusammenhang zwischen Intelligenztestleistungen und beruflichem Fachwissen (durchschnittliche Korrelation von  $r=.80$ ) ebenfalls bekräftigt werden. Der I-S-T 2000 R bietet die Möglichkeit, die fluide Intelligenz zu erfassen und ist damit geeignet den Zusammenhang zu analysieren. Darüber hinaus kann auf der Basis der Untersuchungen ein weiterer Hinweis auf die konvergente Validität des Wissenstests zu konstrukt-nahen Persönlichkeitsmerkmalen generiert werden. Zum Anderen wurde deklaratives Wissen, welches der Messung von gc im I-S-T 2000 R zugrunde liegt und dessen Abbildung auch im START-W Wissenstest ermöglicht wird, in unterschiedlichen Forschungskontexten und Wissenstests mit verschiedenen Konstrukten in Verbindung gebracht. So wird deklaratives Wissen beispielsweise im Bochumer Wissenstest (Hossiep & Schulte, 2008) zum Einen als eigenes Konstrukt behandelt und zum Anderen in enger Verbindung mit Interessen gesehen. Es stellt sich somit die Frage, ob die gc-Indikatoren in substantiellem Umfang auch andere Varianzen erfassen. Daher soll untersucht werden, in welchem Ausmaß der gc-Indikator deklaratives Wissen mit gf korreliert und inwiefern diese Korrelation vom Aggregations- bzw. Generalitätsniveau der Wissensvariablen abhängt.

### *Forschungsannahme 3.3*

Es wird erwartet, dass die Höhe der Korrelation, aufgrund des postulierten Zusammenhangs zwischen fluiden Intelligenz (gf) und Wissen (gc), substantiell ist. Darüber hinaus wird erwartet, dass Wissen auf einem geringeren Generalitätsniveau, bei dem vor allem die wissenthematischen Varianzen relevant sind, eher gering mit gf korreliert. Dagegen wird angenommen, dass die leichter generalisierbaren, inhaltsgebundenen Wissensvarianzen höher mit gf korrelieren.

## 9 Darstellung der Forschungsmethode

### 9.1 Beschreibung der eingesetzten Testinstrumente

#### 9.1.1 Intelligenz-Struktur-Test 2000 R (I-S-T- 2000 R)

In diesem Kapitel wird zunächst der gesamte I-S-T 2000 R von Amthauer et al. (2001) bzw. Liepmann et al. (2007) vorgestellt, wobei im Rahmen der Untersuchung lediglich das Grundmodul eingesetzt wurde.

Die Grundlage des I-S-T 2000 R ist der I-S-T 70 von Amthauer (1970), auf dessen Basis der I-S-T 2000 R revidiert und weiterentwickelt worden ist. Die aktuelle Version ermöglicht neben der Erfassung der Merkfähigkeit, der allgemeinen Intelligenz und der Intelligenzstruktur die Berechnung der Kennwerte zur fluiden und kristallisierten Intelligenz. Ein weiterer theoretischer Grundbaustein für den I-S-T 2000 R bildet das in Kapitel 4.4.4 dargestellte hierarchische Protomodell (HPI). Eine spezielle Form des HPI wird in dem Modell der fluiden und kristallisierten Intelligenz von Cattell (1971a, 1987) gesehen, welches bei der Entwicklung des I-S-T-2000 R eine bedeutende Stellung einnimmt. Die Entwicklung der vorliegenden Testkonstruktion beruht demnach auf der Differenzierung zwischen der fluiden und kristallisierten Intelligenz auf der zweiten Hierarchieebene, die im Rahmen des I-S-T 2000 R operationalisiert werden. Im Unterschied zu dem theoretischen Modell von Cattell (1971a, 1987) wird die *gf/gc* spezifische Konkretisierung auf der zweiten Ordnungsebene im Rahmen des I-S-T 2000 R jedoch durch eine Inhaltsfacette für verbale, numerische und figurale Fähigkeiten erweitert, um die Kontamination von *gf* mit figuralen Fähigkeiten und *gc* mit verbalen Fähigkeiten zu vermeiden. Darüber hinaus wird der Bedeutungsgehalt der Faktoren aus dem Modell von Cattell (1971a, 1987) abgeleitet. Die fluide Intelligenz wird als von den „Einflüssen der Akkulturation „bereinigtes“ schlussfolgerndes Denken“ (Amthauer et al., 2001, S.14), aufgefasst. Die kristallisierte Intelligenz ist dagegen ein kulturabhängiger Faktor, der mit der Leistungsdimension Wissen in Beziehung steht. In dem vorliegenden Test wird das schlussfolgernde Denken mit Wissensanteilen als schlussfolgerndes Denken aufgefasst, und das schlussfolgernde Denken ohne Wissensanteile als schlussfolgerndes Denken/*gf* bezeichnet. Umgekehrt beinhaltet die Leistungsdimension Wissen das schlussfolgernde Denken, das beim Wissen/*gc* herauspartialisiert wurde. Wissen/*gc* und schlussfolgerndes Denken/*gf* bilden dabei die Endpunkte eines Kontinuums, wobei das schlussfolgernde Denken und das Wissen aufgrund ihrer Überschnei-

dungen einen Übergangsbereich darstellen. Im Bezug auf die Investmenttheorie von Cattell (1971a, 1987) kann die Bezeichnung schlussfolgerndes Denken/gf und das Wissen/gc mit den Bezeichnungen fluide bzw. kristallisierte Intelligenz gleichgesetzt werden.

Zur Erfassung bzw. Messung der fluiden und kristallisierten Intelligenz auf der zweiten Hierarchieebene enthält der I-S-T 2000 R drei sogenannte Module, die entsprechend der diagnostischen Zielsetzung und aus ökonomischen Gesichtspunkten miteinander kombiniert werden können. Die Anwendung und Interpretation der Testmodule muss jedoch in einer von den Autoren festgelegten Form durchgeführt werden:

1. Grundmodul
2. Grundmodul- Kurzform
3. Erweiterungsmodul

Das Grundmodul wird aus neun Aufgabengruppen zusammengesetzt: Satzergänzung, Analogien, Gemeinsamkeiten, Rechenaufgaben, Zahlenreihen, Rechenzeichen, Figurenauswahl, Würfelaufgaben, Matrizen. Die Aufgabengruppen werden jedoch nicht einzeln interpretiert, sondern nur vor dem Hintergrund der gebildeten Skalen zur verbalen, numerischen und figuralen Intelligenz. Demnach entstehen 3x3 Aufgabengruppen, die jeweils 20 Aufgaben enthalten und die wie folgt unterteilt werden:

1. Verbale Intelligenz  
Satzergänzung (SE-A1), Analogien (AN-A2), Gemeinsamkeiten (GE-A3)
2. Numerische Intelligenz  
Rechenaufgaben (RE-A4), Zahlenreihen (ZR-A5), Rechenzeichen (RZ-A6)
3. Figurale Intelligenz  
Figurenauswahl (FA-A7), Würfelaufgaben (WÜ-A8) und Matrizen (MA-A9)

Daneben umfasst das Grundmodul noch das schlussfolgernde Denken, das sich aus dem Summenwert der gebildeten Skalen ergibt und zwei Aufgabenblöcke zur Merkfähigkeit (verbale und figurale Merkfähigkeit). Letztere werden jedoch getrennt von den Inhaltsklassen interpretiert. Zusammenfassend werden im Grundmodul die verbale, figurale, und numerische Intelligenz sowie die Merkfähigkeit und das schlussfolgernde Denken erfasst. Die Grundmodul-Kurzform kann aus ökonomischen Gesichtspunkten auch ohne die beiden Aufgabengruppen zur Merkfähigkeit eingesetzt werden.

Die Reliabilitätsschätzungen für die Skalen (verbale, figurale und numerische Intelligenz) und dem schlussfolgernden Denken sind dabei zufriedenstellend und liegen zwischen .87 und .96 (Cronbachs Alpha) bzw. zwischen .88 und .97 (*split half*).

Da die einzelnen Aufgabenblöcke in der diagnostischen Anwendung nicht einzeln interpretiert werden, sind die Angaben über die Inhaltsskalen maßgeblich. Hinsichtlich der getätigten Dimensionsanalysen lässt sich festhalten, dass die im Grundmodul vorgegebenen Inhaltsskalen anhand von multivariaten Analysen abgesichert werden konnten ( $N = 2.208$  bzw.  $N = 2.363$ ). Weitere Validitätsbelege konnten durch die Betrachtung der Zusammenhänge mit anderen Tests (kriterienbezogene Validität) ermittelt werden ( $N=1836$ ). Hinweise auf die diskriminante Validität ergab beispielsweise die erwartungsgemäß geringe Korrelation zwischen dem Grundmodul und dem D2 (Konzentrationsstest). Hohe Korrelationen werden hingegen zwischen den Matrizen des CFT 20 (Weiß, 1997), dem Raven (Raven, 1971) und dem Grundmodul berichtet (bis zu  $r = .69$ ). Hinsichtlich der Zusammenhänge mit dem Außenkriterium Schulnote werden die höchsten Korrelationskoeffizienten zwischen den numerischen und mathematisch-naturwissenschaftlichen Fächern angegeben. Darüber hinaus liegen Standardwert-Normen für die Formen A, B und C auf der Grundlage von mehr als 5800 Personen im Alter von 15 bis 60 Jahren vor (Gymnasiasten, Nicht-Gymnasiasten). Ausführlichere Informationen und Befunde werden im Manual des I-S-T 2000 R gegeben (Liepmann et al. 2007).

Das Erweiterungsmodul wird in Kombination mit dem Grundmodul zur Erfassung der fluiden und kristallisierten Intelligenz (schlussfolgerndes Denken/gf bzw. Wissen/gc) eingesetzt. In diesem Test werden jeweils 14 Fragen zu den Themengebieten Geografie /Geschichte, Wirtschaft, Kunst/ Kultur, Mathematik, Naturwissenschaften und Alltag gestellt, die durch die Inhaltsfacetten (verbales, figurales, und numerisches Wissen) kodiert werden. Die Test-Items des I-S-T 2000 R umfassen insgesamt 84 Fragen, von denen 28 Fragen durch die verbale, 28 Fragen durch die figurale und 28 Fragen durch die numerische Inhaltsfacette gekennzeichnet sind. Die Rohwerte des Wissenstests erhält der Anwender durch die Aufsummierung der richtigen Lösungen, die mit einem Punkt bewertet werden. Die Rohwerte der drei Wissensskalen und deren Aggregation, die den Gesamtwert des Wissens einer Person (gesamte Wissensskala) darstellt, werden zunächst in sogenannte Punktwerte (Multiplikation der Rohwerte mit ihren  $\beta$ -Gewichten und Summierung der Produkte) umgerechnet und anschließend werden anhand von

Normtabellen Standardwerte berechnet. Die Standardwerte können dann, unter Berücksichtigung der theoretischen Annahmen, interpretiert werden.

Der Summenwert aus den einzelnen Inhaltsfacetten bestimmt demgemäß das Wissen und den Faktorwert Wissen/gc. Diesbezüglich wurde darauf geachtet, dass die Themenbereiche und die Kodierungsart in einem ausbalancierten Verhältnis stehen. Im Prinzip wird der Test zur Bestimmung des aufgenommenen Wissens eingesetzt, womit gleichzeitig die Fähigkeit zum Wissenserwerb in der jeweiligen Kultur analysiert wird. Darüber hinaus bietet das Erweiterungsmodul die Möglichkeit, das schlussfolgernde Denken/ gf abzubilden. Zusammengefasst werden im Erweiterungsmodul das verbal, figural und das numerisch kodierte Wissen sowie das Wissen (als Summenwert der Kodierungsarten), das Wissen/gc und indirekt das schlussfolgernde Denken/gf (als Faktorwerte), mit einbezogen (siehe Abbildung 10). Im Sinne der Investmenttheorie von Cattell (1971a, 1987) können die Faktorwerte schlussfolgerndes Denken/gf und Wissen/gc auch als fluide Intelligenz und kristallisierte Intelligenz bezeichnet werden. Fluide Intelligenz wird in diesem Zusammenhang wie folgt beschrieben: „Der Faktor ist ein fehlbarer Indikator des formal-logischen Denkens verschiedener Arten, wie Abstrahieren und Problemlösen, wenn diese Fähigkeiten außerhalb des Akkulturationsprozesses, durch persönliche Erfahrung und nicht selektiv eingeschränktes Lernen, erworben werden“ (Horn, 1988, S.660, übersetzt nach Liepmann et al., 2007, S. 47). Kristallisierte Intelligenz wird hingegen folgendermaßen definiert: „Der gemessene Faktor ist ein fehlbarer Indikator des Ausmaßes, in dem ein Individuum durch die systematischen Einflüsse der Akkulturation das Wissen und die Bildung erworben hat, die als Intelligenz einer Kultur bezeichnet werden können“ (Horn, 1988, S. 658-659, übersetzt nach Liepmann et al., 2007, S. 47). Entsprechend den obigen Ausführungen (vgl. Kapitel 4.4.4 und Kapitel 6) wird darauf hingewiesen, dass die beiden Faktoren jedoch keine klare Einfachstruktur aufweisen, sondern sich gemäß der Investmenttheorie bedingen und interkorrelieren. Demnach wurde bei der Konstruktion des Tests zudem darauf geachtet, dass (a) die fluide Intelligenz und die kristallisierte Intelligenz simultan erfasst werden, (b) klassische Markiervariablen in die jeweiligen Skalen einbezogen werden und (c) die Kontamination fluider und kristallisierter Intelligenz mit figuralem Material bzw. verbalen Material vermieden wird. Es lässt sich somit festhalten, dass der I-S-T 2000 R explizit für die Erfassung der fluiden und kristallisierten Intelligenz konstruiert wurde, das im Vergleich zu anderen Testverfahren ein innovatives Novum für den deutschsprachigen Raum darstellt.

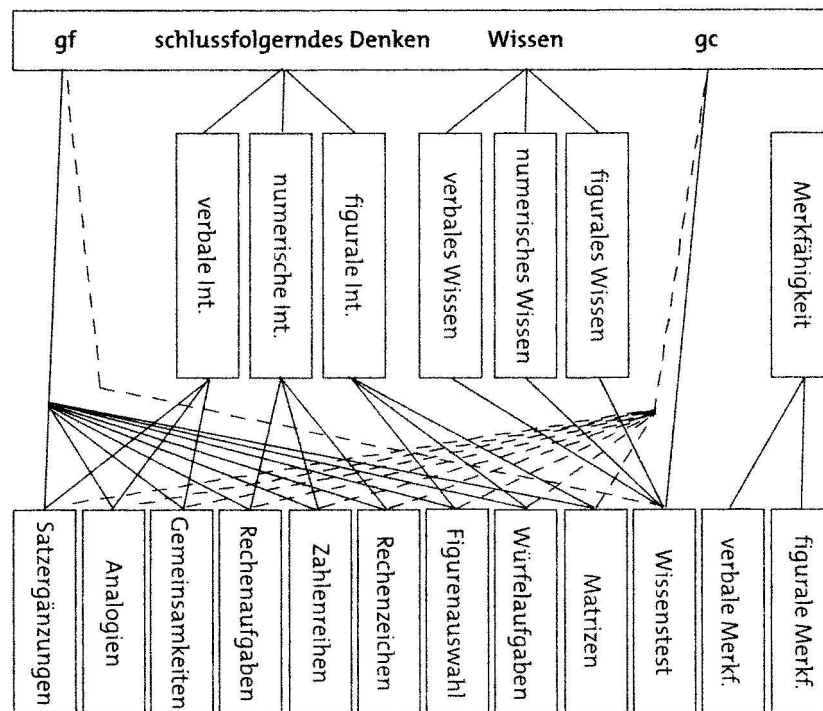


Abbildung 10: Fähigkeitsstruktur des I-S-T-2000 R

Die zugrunde liegende Stichprobe für die Berechnung der Kennwerte des Wissensmoduls umfasst insgesamt 661 Probanden, von denen 45% (N=295) männlich und 55% (N=366) weiblich sind. Das Alter der Probanden lag zwischen 15 und 60 Jahren und das durchschnittliche Alter betrug 28 Jahre (SD=9,6). Die stärkste Vertretung in der Normstichprobe stellen die Probanden mit Fach- oder Hochschulreife dar (72%). Im Verhältnis zu dieser Personengruppe sind die Volks-, Haupt- und Realschüler (21%), Schüler (6%) und Probanden ohne Schulabschluss eher unterrepräsentiert, so dass nur relativ vage Alters- und Bildungsdifferenzierungen möglich sind. Daher wurde das Verfahren der Gewichtung von Teilstichproben verwendet (vgl. Lynn, Hampson & Lee, 1988), um repräsentative Normen zu erhalten. Das Ziel von weiteren Untersuchungen ist es daher, die Normstichprobe zu erweitern.

Hinsichtlich der Konstrukt-Validität konnte anhand von *multidimensionalen Ähnlichkeitsstrukturanalysen* (MDS, Borg & Lingoes, 1987) und *konfirmatorischen Faktorenanalysen* mithilfe von LISREL die intendierten Themenfacetten sowie die Kodierungsstruktur nachgewiesen bzw. differenziert werden. Zudem konnten auf der Basis des Grundmoduls und des Wissenstests die Faktoren Wissen/gc und das schlussfolgernde Denken/gf extrahiert bzw. abgebildet werden. Die Ermittlung der Faktorwerte erfolgte mithilfe von *exploratorischen Faktorenanalysen* (*Hauptachsenanalysen mit Kommunalitäteniteration*) und nachfolgender schiefwinkliger Rotation (oblimin), multidimensio-

nen Ähnlichkeitsstrukturanalysen und konfirmatorischen Faktorenanalysen. Darüber hinaus werden neben der faktoriellen Struktur noch weitere Validitätsbelege mit Kriteriumsvariablen hinzugezogen. Bezüglich der kriterienbezogenen Validität mit dem Außenkriterium Schulnote rangieren die Korrelationen zwischen dem Faktor Wissen/gc und ausgewählten Schulnoten (Geschichte, Erdkunde, Mathematik, Physik) zwischen  $r = .29$  und  $r = .38$  und bei dem schlussfolgernden Denken/gf zwischen  $r = .11$  und  $r = .43$ . Im Hinblick auf die Korrelationen mit anderen Tests (D2, Brickenkamp, 1997; MWT-B, Lehrl, 1995; HAWIE-R-Wissen, Tewes, 1991; CFT 20, Weiß, 1974, 1997) werden für den Faktor Wissen/gc Koeffizienten von  $-.09$  (CFT 20) bis  $.68$  (HAWIE-R) berichtet und für schlussfolgerndes Denken/gf von  $.16$  (MWT-B) bis  $.58$  (CFT 20). Die Kriteriumsvalidität wird insgesamt als zufriedenstellend bewertet. Die Reliabilitätsschätzungen der Wissensskalen liegen in dem Bereich zwischen  $0.82$  und  $0.93$  (Cronbachs Alpha) sowie zwischen  $.84$  und  $.93$  (split half). Für den Faktorwert Wissen/gc geben die Autoren eine Split-half Reliabilität von  $.91$  (Spearman-Brown korrigiert) an und für das schlussfolgernde Denken/gf eine Kennzahl von  $.96$  (Spearman-Brown korrigiert). Es lässt sich festhalten, dass die Validitäts- und Reliabilitätsschätzungen ausreichend sind und die Anforderungen grundsätzlich erfüllt werden. Detaillierte Angaben werden im Manual des I-S-T 2000 R (Liepmann et al., 2007) gemacht.

Resümierend lässt sich festhalten, dass der I-S-T 2000 R (Liepmann et al., 2007) ein eignungsdiagnostisches Instrument darstellt, das im Hinblick auf andere Intelligenztests den Leerraum zwischen fluider und kristallisierter Intelligenz schließt und damit dem vielschichtigen Konstrukt-Verständnis gerecht wird. Durch die Abbildung der verschiedenen Facetten und der Möglichkeit zur Erstellung eines Fähigkeitsprofils eignet sich der umfassende Test besonders gut zur flankierenden Absicherung zukünftiger Personalauswahl-, Personalplatzierungs- oder Personalentwicklungsentscheidungen.

### **9.1.2 START-W Wissenstest für Berufseinsteiger**

Wie schon angedeutet liegt das Hauptaugenmerk auf dem oben dargestellten Erweiterungsmodul. Der darin enthaltene Wissenstest bildet im Weiteren die Grundlage für den eingesetzten Wissenstest START-W.

Der Allgemeinwissenstest START-W (Liepmann & Beauducel, in Vorbereitung) soll in der Zukunft ein Bestandteil der START-Testbatterie für Berufseinsteiger werden.

Dieses Modul soll zur objektiven Erfassung des Allgemeinwissens integriert werden, um Auswahl- und Platzierungsentscheidungen bei potentiellen Berufsanfängern zu optimieren. Mit der Weiterentwicklung der START-Testreihe wird das Ziel verfolgt, ein eignungsdiagnostisches Instrument für Berufseinsteiger zu konzipieren, das fundamentale Kompetenzen erfassen soll. Die Basiskompetenzen beziehen sich dabei auf die Bereiche (a) Allgemeinwissen (START-W), (b) Konzentrationsfähigkeit (START-K), (c) Persönlichkeit (START-P), (d) Fremdsprachenkenntnisse (START-E), (e) Rechtschreibung (START-R), (f) EDV- Wissen (START-C) und (g) Mathematik (START-M). Die einzelnen Module sind, entsprechend der eignungsdiagnostischen Fragestellungen oder auch hinsichtlich der oben genannten Anwendungsfelder (siehe Kapitel 3.4.3), beliebig miteinander kombinierbar. Jedes Verfahren, der bereits etablierten Testinstrumente (START-K, START-P, START-E, START-C) wurde an mehr als 2000 Probanden normiert. Die Normtabellen beinhalten differenzierte Vergleichsdaten nach Schulbildung und Alter sowie eine Gesamtnorm. Darüber hinaus entsprechen sämtliche Module der vom Verband für Psychologen aufgestellten DIN-Norm 33430 und erfüllen in vollem Umfang die allgemeingeltenden wissenschaftlichen Gütekriterien.

Der hier verwendete START-W stellt eine Vorversion zur Erfassung des Allgemeinwissens dar und bildet im Weiteren die Grundlage für die empirische Untersuchung. Die Auswahl der Test-Items basiert einerseits auf dem oben dargestellten Erweiterungsmodul des I-S-T 2000 R und andererseits auf neu entwickelten Test-Items. Der Test umfasst insgesamt 134 Wissensfragen aus sieben Themengebieten: Kunst/Kultur, Alltag, Wirtschaft, Gesundheit, Mathematik, Naturwissenschaften und Geografie/Geschichte. Das Wissensgebiet „Naturwissenschaften“ ist dabei als Oberbegriff zu verstehen, das noch weitere Fragen zur Physik, Biologie und Chemie beinhaltet. Das gleiche gilt für die Aufgabengruppe „Kultur“, die die Themenblöcke Kunst, Literatur und Religion zusammenfasst. Darüber hinaus werden die Aufgabengruppen (Themenbereiche) durch die verbale, numerische und figurale Inhaltsfacette kodiert. Die Zuordnung der Test-Items auf die Inhaltsfacetten erfolgt in Kapitel 10.1 Aufgrund der nicht unerheblichen Überschneidungen mit dem Erweiterungsmodul des I-S-T 2000 R kann davon ausgegangen werden, dass der Wissenstest die kristallisierte Intelligenz abbilden kann.

Bezüglich der Anwendung des START-W werden den Probanden 134 Fragen vorgelegt, die fünf verschiedene Antwortmöglichkeiten enthalten und von denen jeweils eine Antwort richtig ist. Die Bearbeitungszeit beträgt 40 Minuten.



## 9.2 Darstellung des Untersuchungsablaufs

Vor dem Hintergrund der bisherigen Entwicklung, Normierung und der avisierten Validierung des START-W-Wissenstests wurden Datenerhebungen an verschiedenen Institutionen bzw. Bildungsträgern durchgeführt. Die Erfassung der Daten wurde hierbei in drei Phasen unterteilt.

Die gesamte Datenerhebung erstreckt sich über den Zeitraum von August 2008 bis Januar 2010.

Die erste Phase der zugrunde liegenden Datenerhebung wurde im Zeitabschnitt vom 12.08.2008 bis zum 19.03.2009 durchgeführt. In die Untersuchung wurden, entsprechend der angestrebten Zielpopulation der Berufseinsteiger, die folgenden Bildungseinrichtungen einbezogen:

1. Vivantes Netzwerk für Gesundheit GmbH, Institut für berufliche Bildung im Gesundheitswesen (IbBG)
2. Wannsee-Schule e.V. –Schule für Gesundheitsberufe
3. Charité Gesundheitsakademie- Universitätsmedizin Berlin
4. Freie Universität Berlin
5. Marie-Curie-Oberschule Gymnasium

An den genannten Institutionen wurde das Grundmodul des Intelligenz-Struktur-Tests 2000 R und der START-W Wissenstest eingesetzt.

Die zweite Phase bezieht sich auf den Zeitraum von März 2009 bis November 2009. In diesem Intervall wurden zusätzlich der Bochumer Wissenstest (BOWIT- Kurzform, Hossiep & Schulte, 2007) und der im Berliner Intelligenzstrukturtest (Jäger, 1997) enthaltene Konzentrations- bzw. Bearbeitungsgeschwindigkeitstest einbezogen. Außerdem beinhaltete diese Phase den wiederholten Einsatz des Wissenstests START-W (Retest). Der Ablauf der Untersuchung basierte auf einem mindestens zweimonatigen Zeitintervall.

Die dritte Datenerhebungsphase (Dezember 2009 bis Januar 2010) umfasste die Zusammenstellung der Noten der einzelnen Probanden nach Ablauf des Probehalbjahres in der beruflichen Ausbildung. Anhand dessen soll im empirischen Teil der Erfolg auf der Zielposition überprüft werden, für dessen Vorhersage das eignungsdiagnostische Instrument bzw. der Wissenstest START-W eingesetzt wurde (hier Ausbildungserfolg).

Aufgrund der Tatsache, dass einige Probanden schon im Vorfeld (nach dem Auswahlverfahren) keinen Ausbildungsplatz erhalten haben und aus Datenschutzgründen kein weiterer Kontakt zu den Bewerbern aufgebaut werden konnte, wird sich die Stichprobengröße der ersten und zweiten Datenerhebungsphase für diesen Analyseteil auf ca. 100 Probanden (vorwiegend Probanden der Vivantes-Schule) reduzieren.

Die Stichprobe umfasst insgesamt 765 Probanden im Alter von 15-46 Jahren (Durchschnittsalter 20,75 Jahre,  $SD= 4,12$ ).

Vor den Gruppenuntersuchungen wurden alle Testleiter detailliert über die Vorgehensweise in Kenntnis gesetzt, um standardisierte Testanweisungen, gleichbleibende Untersuchungsbedingungen und einen strukturierten Ablauf zu gewährleisten. Die Hinweise und Bedingungen zur Testdurchführung wurden gemäß dem Manual von Liepmann et al. (2007, S.75; siehe Anhang 1) angeleitet.

Der strukturelle Verlauf der Untersuchung wurde folgendermaßen arrangiert: Nach einer kurzen Vorstellung des Testleiters und des Fachbereichs für Wirtschafts- und Sozialpsychologie wurden die Schüler, Auszubildenden (Bewerber) und Studenten im Rahmen einer Unterrichtseinheit oder zu einem festgesetzten Termin zunächst über den Zweck der Untersuchung informiert. Für die Probanden des Vivantes Instituts sowie der Wannseeschule e.V. (Bewerber) waren die Tests integraler Bestandteil von Auswahlverfahren. Die Probanden der anderen Bildungseinrichtungen wurden über die wissenschaftliche Bedeutung aufgeklärt. Es wurde mitgeteilt, dass es bei dieser Untersuchung um eine bundesweite Entwicklung und Normierung dieser Testverfahren geht, wobei mehrere hundert Probanden mit einbezogen werden, um letztendlich repräsentative Aussagen treffen zu können. Mit dem Ziel, die Motivation, Konzentration und das Interesse zu erhöhen, wurden die Probanden darauf aufmerksam gemacht, dass es sich bei den ausgewählten Tests auch um eignungsdiagnostische Instrumente handelt, mit denen sie in der späteren Personalauswahl durchaus konfrontiert werden können. Demnach ergab sich der Nutzen aus berufspraktischer Perspektive dadurch, dass die Tests eine Übung für zukünftige Berufseinstellungsverfahren darstellen. Darüber hinaus wurde betont, dass insbesondere das erfasste Wissen im START-W (auch in Kombination mit Intelligenztests) eine zunehmende Relevanz bei der Beurteilung von Berufseinstiegern erhalten und der Test dabei helfen soll den eigenen Wissensstand und seine kognitiven Fähigkeiten zu überprüfen. Um die Testinstruktion und die Versuchsbedingungen konstant zu halten, war der Testleiter in der gesamten Unterrichtseinheit bzw. Testzeit

anwesend.<sup>3</sup> Als erstes wurde das Grundmodul des I-S-T 2000 R vorgelegt, der 120 Minuten dauerte. Nach einer dreiminütigen Pause und einer angemessenen Instruktionszeit wurde der Wissenstest START-W überreicht, der eine Zeit von 40 Minuten beanspruchte. Der Wissenstest, wie auch der I-S-T 2000 R, wurden zudem in zwei verschiedenen Versionen (Form A und Form B) ausgehändigt, um eventuelle Täuschungsversuche zu verhindern. Die zwei Versionen unterscheiden sich dabei lediglich durch die Reihenfolge der dargebotenen Testitems. Im Zuge der zweiten Datenerhebungsphase wurde der Bearbeitungsgeschwindigkeitstest vorangestellt, anschließend der START-W und abschließend der Bochumer Wissenstest (BOWIT-Kurzform) bearbeitet. Bei der Testdurchführung wurde stets auf die allgemeinen Standards der fachgerechten Testanwendung (z.B. Häcker, Leutner und Amelang, 1998) als auch auf die Vorgaben der DIN-Norm 33430 geachtet. Die Reihenfolge der Aufgabengruppen und die Bearbeitungsdauer der einzelnen Untersuchungssequenzen (Grundmodul-Kurzform und START-W) werden in der folgenden Tabelle 10 wiedergegeben.

*Tabelle 10: Reihenfolge und Bearbeitungszeit der eingesetzten Testverfahren*

Aufgabengruppe	Aufgabenzahl	Bearbeitungszeit (Min)
Soziodemografischer Fragebogen		20
<i>15 Min. Instruktionszeit</i>		15
SE- Satzergänzung	20	6
AN- Analogien	20	7
GE- Gemeinsamkeiten	20	8
RE- Rechenaufgaben	20	10
ZR- Zahlenreihen	20	10
RZ- Rechenzeichen	20	10
FA- Figurenauswahl	20	7
WÜ- Würfelaufgaben	20	9
MA- Matrizen	20	10
Zwischensumme	180	112
5 Min. Pause		5
<i>3 Min. Instruktionszeit</i>		3
START-W	134	40
Gesamtsumme	314	160

<sup>3</sup> Eine detailliertere Beschreibung des Untersuchungsablaufes wird im Anhang 1 gegeben.

Ab 15.01.2010		
2 Min. Instruktionszeit		2
Bochumer Wissenstest (BO- WIT)- Kurzform	45	ca. 30
2 Min. Instruktionszeit		2
Bearbeitungsgeschwindigkeitstest BIS-4	1	ca. 1
Gesamtsumme	360	195
Retest-Phase: START-W; BOWIT; BIS	180	78

### 9.3 Darstellung der Stichprobe

Hinsichtlich der empirischen Datenerhebung wurden die folgenden Vorüberlegungen getroffen: Entsprechend den vorangegangenen Ausführungen (s.Kap.1) wurde insbesondere durch die Praxis der Wunsch geäußert, eine Testbatterie zu entwickeln, die sogenannte Basiskompetenzen beim Übergang Schule – Ausbildung – Beruf erfasst und die mögliche individuelle Defizite aufzeigt. Der Anwendungsbereich der zukünftigen START-Testbatterie für Berufsanfänger bezieht sich demnach auch auf die berufliche Eignungsdiagnostik in Organisationen, wie sie beispielsweise im Rahmen von verschiedenen Auswahlverfahren praktiziert wird. Für die Entwicklung des Wissenstests START-W, der zukünftig ein Teil dieser Testbatterie sein soll, wird demzufolge eine Personengruppe verstärkt in die Untersuchung einbezogen, die sich genau an der Schnittstelle zwischen Schule und Berufsausbildung befindet und anhand derer ein enger Bezug zu den dargelegten Fragestellungen hergestellt werden kann. Darauf aufbauend wurden die weiteren empirischen Untersuchungen an den relevanten Bildungseinrichtungen bzw. Schulen und Auszubildendenstätten geplant, vorbereitet und durchgeführt. Um die eignungsdiagnostischen Fragestellungen (s.o.) adäquat zu bearbeiten, wird daher die folgende Vorgehensweise gewählt:

Im ersten Schritt werden die Daten einer groben Überprüfung unterzogen, um Probanden mit fehlenden oder stark abweichenden Angaben im Vorfeld zu eliminieren. Danach werden auf Basis der erhobenen Daten die Stichprobengröße bestimmt und die relevanten soziodemografischen Angaben beschrieben. Die deskriptive Statistik umfasst dabei vor allem die Darstellung der Geschlechterverhältnisse, Altersstruktur und Bildungsklassen.

### 9.3.1 Datenvorbereitung

Vor der Datenanalyse werden die Informationen und Angaben in angemessener Form aufbereitet. Dementsprechend werden im Folgenden die geplanten Modifikationen des Datensatzes, aufgrund der Forderung nach Transparenz und Verständlichkeit, präzise dokumentiert. Ein wichtiger Aspekt hierbei ist der adäquate Umgang mit fehlenden oder lückenhaften Werten. Daher werden der Datensatz bzw. die Rohwerte der einzelnen Substichproben einer genauen Analyse unterzogen, um Verzerrungen der Teststatistik zu vermeiden. Bedingt durch die Tatsache, dass sämtliche Probanden der Stichprobe entweder freiwillig oder im Rahmen der Auswahlverfahren an den Tests teilgenommen haben, lässt sich feststellen, dass die Testresultate von nur sechs Probanden Rohwerte von Null aufweisen. In dem vorliegenden Datensatz werden bereits im Vorfeld der Analysen die Probanden von der Grundgesamtheit ausgeschlossen, die entweder den soziodemografischen Fragebogen oder den Wissenstest komplett nicht ausgefüllt haben. Aus diesem Grund wurde der Stichprobenumfang von 765 Probanden auf 759 reduziert. Es sei jedoch angemerkt, dass einzelne soziodemografische Daten bei den 759 Probanden nach wie vor fehlen. Diese Daten sind jedoch bei der Stichprobenbeschreibung und der anschließenden Analyse vernachlässigbar.<sup>4</sup>

Ogleich die Probanden explizit auf eine vollständige Beantwortung der soziodemografischen Fragen und der Wissensfragen des START-W hingewiesen wurden, ließ es sich jedoch nicht vermeiden, dass lückenhafte Angaben in dem Datensatz vorkommen. Aufgrund dessen kann es bei der Datenanalyse zu Verzerrungen von Teststatistiken und Parameterschätzungen führen (Jones, 1996). Eine große Anzahl von unvollständigen Fragebögen birgt demnach das Risiko einer ineffizienten Nutzung der erhobenen Daten (Afif & Elashoff, 1966). Hinsichtlich des Umgangs mit lückenhaftem Datenmaterial werden in der einschlägigen Literatur unterschiedliche Verfahren angeboten, wobei der Ausschluss bestimmter Probanden eine leichte, aber keine probate Möglichkeit darstellt. Diese Methode kann ebenso zu Komplikationen führen, insbesondere wenn der Fragebogen bzw. der Test in die Kategorie der Leistungstests fällt. In Anbetracht dessen wird der Datensatz einerseits empfindlich gekürzt und andererseits gelten willentlich ausgelassene Fragen als relevante und bedeutende Informationsträger, die auch einen Indikator für aufgezeigte Wissenslücken darstellen.

---

<sup>4</sup> Aufgrund von fehlenden Angaben wird angemerkt, dass der Stichprobenumfang im Rahmen der deskriptiven Statistik variieren kann.

Mithin verzerrt die Eliminierung von unvollständigen Datensätzen die Stichprobe und reduziert den Erklärungsgehalt der Ergebnisse. Das lückenhafte Datenmaterial bleibt aus diesen Gründen unberücksichtigt.

### 9.3.2 Deskriptive Statistik

Die zugrunde liegende Gesamtstichprobe umfasst 759 Probanden. Die Größe der Stichprobe entspricht der von Lienert und Raatz (1998) angegebenen Anzahl von Probanden, die für statistische Berechnungen zugrunde gelegt werden sollten (hier:  $N > 200$ ). Der Anteil der weiblichen Probanden an der Gesamtstichprobe beträgt 78,00% ( $N = 592$ ). Im Verhältnis zu den männlichen Probanden mit 22,00% ( $N = 167$ ) stellen sie eine deutliche Mehrheit dar (siehe Abbildung 11). Aus der Tabelle 11 lässt sich ebenfalls entnehmen, dass das weibliche Geschlecht in den verschiedenen Altersstufen stärker vertreten ist.

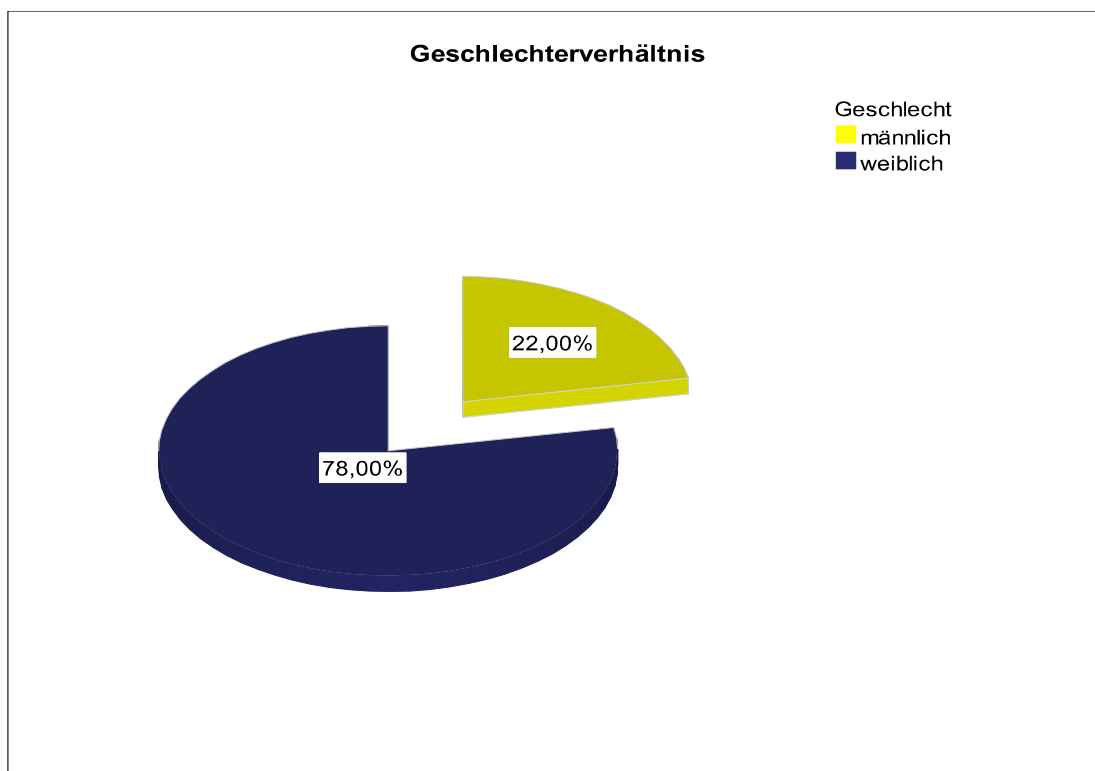


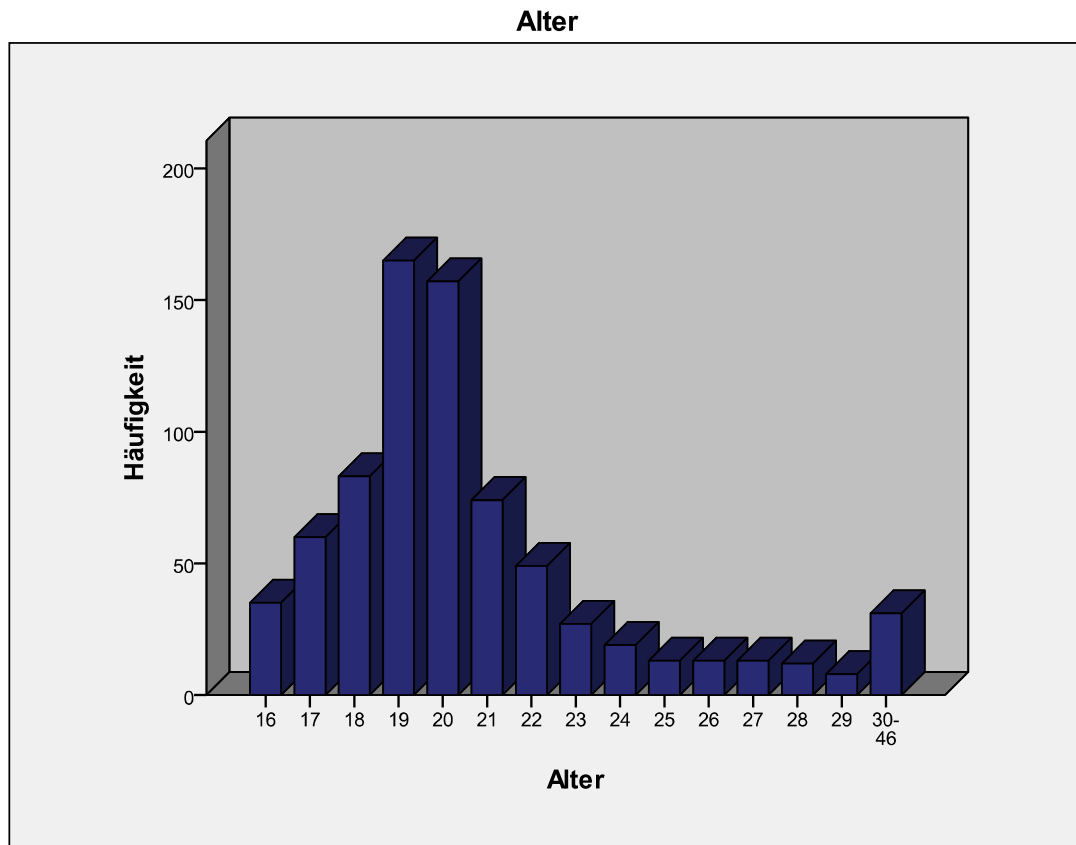
Abbildung 11: Geschlechterverhältnis

Das Alter der Probanden liegt zwischen 16 und 46 Jahren. Das Durchschnittsalter in der Gesamtstichprobe beträgt 20,74 Jahre ( $SD = 4,11$  Jahre), wobei 65,9% das 21. Lebensjahr noch nicht erreicht haben. Die 19- und 20-Jährigen stellen mit 21,7% bzw. 20,7% den Hauptteil der Gesamtstichprobe (siehe Tabelle 11 und Abbildung 12).

In der Teilstichprobe der Männer sind 14,4% 19 Jahre alt und 17,4% 20 Jahre alt. Im Bezug auf die Gruppe der Frauen stellen die 19- und 20-Jährigen ebenfalls die größte Altersgruppe dar (23,8% bzw. 21,6%).

*Tabelle 11: Altersverteilung getrennt nach Geschlecht*

Alter	Geschlecht		Anzahl N	Gesamt	
	Männlich Anzahl N %Zeile/%Spalte	Weiblich Anzahl N %Zeile/%Spalte		Prozentualer Anteil-Alter	Kumulierte Prozent
16	7 20,0% / 4,2%	28 80,0% / 4,7%	35	4,6%	4,6%
17	15 25,0% / 9,0%	45 75,0% / 7,6%	60	7,9%	12,5%
18	13 15,7% / 7,8%	70 84,3% / 11,8%	83	10,9%	23,5%
19	24 14,5% / 14,4%	141 85,5% / 23,8%	165	21,7%	45,2%
20	29 18,5% / 17,4%	128 81,5% / 21,6%	157	20,7%	65,9%
21	15 20,3% / 9,0%	59 79,7% / 10,0%	74	9,7%	75,6%
22	17 34,7% / 10,2%	32 65,3% / 5,4%	49	6,5%	82,1%
23	7 25,9% / 4,2%	20 74,1% / 3,4%	27	3,6%	85,6%
24	8 42,1% / 4,8%	11 57,9% / 1,9%	19	2,5%	88,1%
25	4 30,8% / 2,4%	9 69,2% / 1,5%	13	1,7%	89,9%
26	5 38,5% / 3,0%	8 61,5% / 1,4%	13	1,7%	91,6%
27	4 30,8% / 2,4%	9 69,2% / 1,5%	13	1,7%	93,6%
28	4 33,3% / 2,4%	8 66,7% / 1,4%	12	1,6%	94,9%
29	2 25,0% / 1,2%	6 75% / 1,0%	8	1,1%	95,9
30-46	13 41,9% / 7,8%	18 58,1% / 3,0%	31	4,1%	100%
<b>Gesamt</b>	167 22,0% / 100%	592 78% / 100%	759	100%	



*Abbildung 12: Altersverteilung in der Stichprobe*

Wie oben dargelegt, wurden die Daten der zugrunde gelegten Stichprobe an verschiedenen Bildungsinstitutionen erhoben, die in der Tabelle 12 wiedergegeben werden.

*Tabelle 12: Ausgewählte Bildungsinstitute*

Schulen für Gesundheitsberufe- Berlin	
Vivantes- Institut für berufliche Bildung Charité- Gesundheitsakademie Wannseeschule e.V.- Schule für Gesundheitsberufe	
Bildungsgang	Berufsausbildung
Duales Bildungssystem	Die Ausbildung dauert drei Jahre und gliedert sich in Phasen der theoretischen Ausbildung im Institut für berufliche Bildung - IbBG und Phasen der praktischen Ausbildung im klinischen und ambulanten Bereich.
Aufnahmevoraussetzungen	Realschulabschluss oder gleichwertiger Schulabschluss ,oder Fachhochschulreife, oder Hochschulreife ,oder Hauptschulabschluss und eine mindestens zweijährige abgeschlossene Berufsausbildung. Gutes Abschlusszeugnis. <b>Eignungstest und persönliches Interview</b> <b>Probhalbjahr</b>

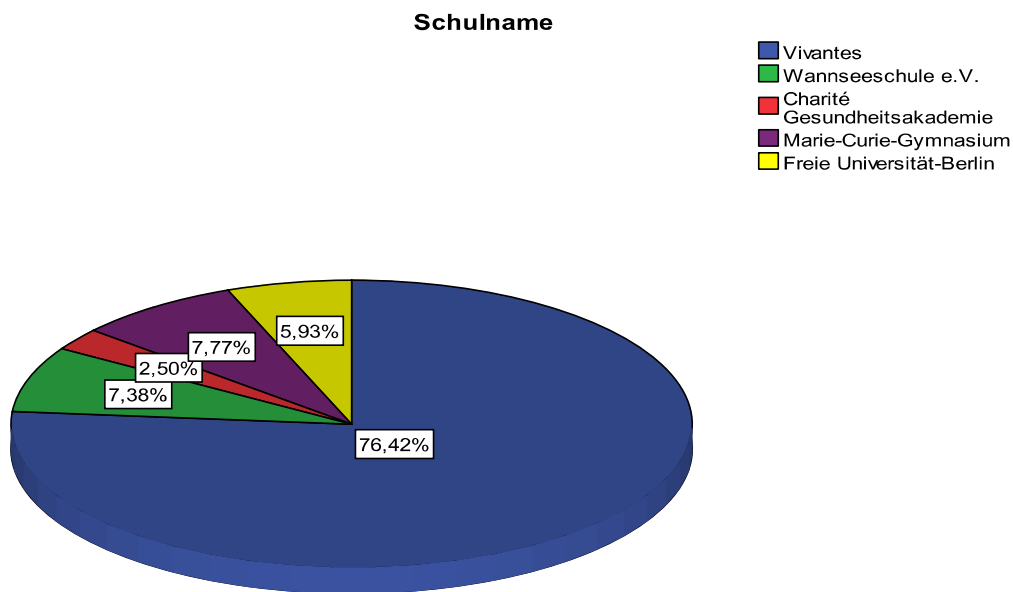


Berufe/ Ausbildungsziel	Ausbildungsziel:  Gesundheits- und Krankenpflege Altenpflege Gesundheits- und Kinderkrankenpflege, Operationstechnische Assistenten (OTA)
<b>Testdurchführung</b>	
Die Testdurchführung fand im Rahmen von Auswahlverfahren und in Form von Gruppenuntersuchungen an den Institutionen statt. Die Auswahlverfahren finden in einem einwöchigen Turnus statt. Die Auswahlverfahren sind für die Bewerber um einen Ausbildungsplatz obligatorisch. Zudem wurden Untersuchungen auf der Basis der freiwilligen Teilnahme durchgeführt.	
Stichprobengröße N= 655	

Marie Curie-Oberschule Gymnasium	
Bildungsgang	Gymnasiale Oberstufe
Aufnahmevoraussetzung	Sehr gutes Abschlusszeugnis an der Grundschule Empfehlungsschreiben des Klassenlehrers Probahalbjahr
Ausbildungsziel	Abitur / Allgemeine Hochschulreife
<b>Testdurchführung</b>	
Die Testdurchführung fand in Form von Gruppenuntersuchungen statt und basierte auf der freiwilligen Teilnahme der Schüler (11-13 Klasse).	
Stichprobengröße N= 59	

Freie Universität Berlin	
Bildungsgang	Studium/Psychologie und affine Fächer
Aufnahmevoraussetzungen	Sehr gutes Abschlusszeugnis (Abitur)
Ausbildungsziel	Hauptsächlich Bachelor/Master of Psychology
<b>Testdurchführung</b>	
Die Testdurchführung wurde in Form von Gruppenuntersuchungen durchgeführt und basierte auf der freiwilligen Teilnahme der Studenten bzw. auf der Ausgabe sogenannter Versuchspersonenstunden, die im Rahmen des Studiums obligatorisch sind.	
Stichprobengröße N=45	

Die Aufteilung auf die einzelnen Bildungsinstitute stellt sich wie folgt dar.



*Abbildung 13: Aufteilung der Bildungsinstitute in Prozent*

Abbildung 13 verdeutlicht, dass das Vivantes Institut für berufliche Bildung den größten Anteil an Probanden stellt (76,42%, N= 580). 7,38% (N= 56) besuchen die Wannseeschule e.V. und die Auszubildenden der Charité sind mit 2,50% (N= 19) vertreten. Die Bildungseinrichtungen für Gesundheitsberufe weisen somit den größten Anteil an der Stichprobe auf. Insgesamt 7,77% (N= 59) sind Schüler des Marie-Curie-Gymnasiums. Die Anzahl der Studenten der Freien Universität beläuft sich auf 45 (5,93%).

Es wird an dieser Stelle darauf verwiesen, dass nicht alle Probanden bzw. Bewerber der Schulen für Gesundheitsberufe letztendlich einen Ausbildungsplatz erhalten haben. Die Personengruppe wurde in der folgenden Abbildung 14 als „Bewerber ohne Ausbildungsplatz“ bezeichnet. Die Kriterien für die Selektion der Bewerber, die im Rahmen der Auswahlverfahren berücksichtigt wurden, werden in der Anlage 1 erwähnt. Dieser Aspekt ist von Bedeutung, weil demnach nicht alle Probanden der ersten Datenerhebungsphase in die zweite Phase mit einbezogen werden konnten und damit die Forschungsfragen 2 und 3 (siehe Kapitel 8) nur mit einem reduzierten Datensatz bearbeitet werden können.

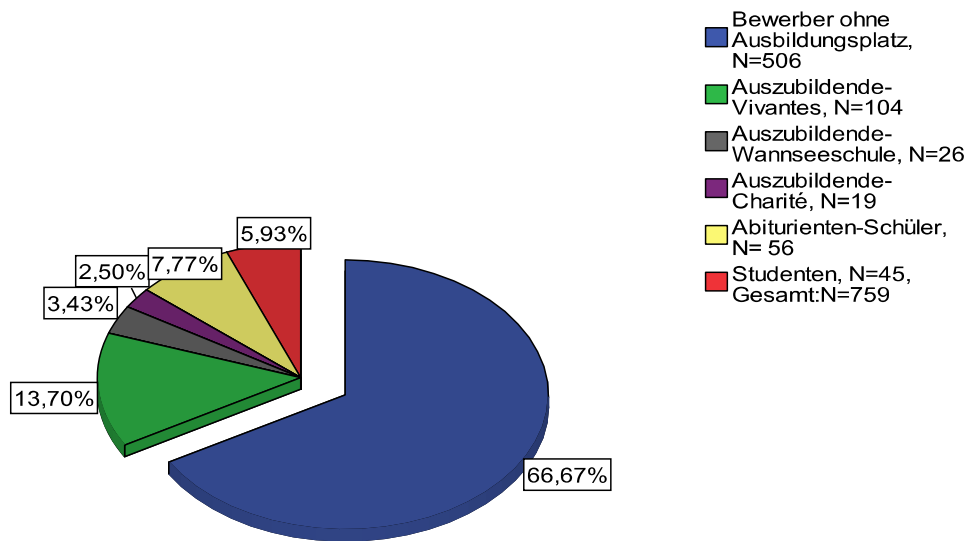


Abbildung 14: Differenzierung der Stichprobe

In der 13 Tabelle werden zudem die angestrebten Ausbildungsziele wiedergegeben und den entsprechenden Bildungseinrichtungen zugeordnet. Überdies wird das Geschlechterverhältnis für die jeweiligen Abschlüsse dargestellt.

Tabelle 13: Verteilung der Probanden über Ausbildungsziele und Bildungsinstitute getrennt nach Geschlecht

Ausbildungsziel		Schulname					Gesamt	Geschlecht
		Vivantes	Wannsee Schule	Charité	MCO	FU-Berlin		Männlich/ Weiblich
Abitur	Anzahl	0	0	0	59	0	59	31/28
	% Zeile	,0%	,0%	,0%	100,0%	,0%	100,0%	52,5%/47,5%
	% Spalte	,0%	,0%	,0%	100,0%	,0%	7,8%	18,6%/4,7%
	% der Gesamtzahl	,0%	,0%	,0%	7,8%	,0%	7,8%	4,1%/3,7%
Altenpflege	Anzahl	5	0	0	0	0	5	4/1
	% Zeile	100,0%	,0%	,0%	,0%	,0%	100,0%	80%/20%
	% Spalte	,9%	,0%	,0%	,0%	,0%	,7%	2,4%/0,2%
	% der Gesamtzahl	,7%	,0%	,0%	,0%	,0%	,7%	0,5%/0,1%

Gesundheits-u. Krankenpflege-1	Anzahl	396	56	19	0	0	471	99/372
	% Zeile	84,1%	11,9%	4,0%	,0%	,0%	100,0%	21%/79%
	% Spalte	68,3%	100,0%	100,0%	,0%	,0%	62,1%	59,3%/62,8%
	% der Gesamtzahl	52,2%	7,4%	2,5%	,0%	,0%	62,1%	13%/49%
Gesundheits-u. Krankenpflege -2	Anzahl	7	0	0	0	0	7	0/7
	% Zeile	100,0%	,0%	,0%	,0%	,0%	100,0%	0%/100%
	% Spalte	1,2%	,0%	,0%	,0%	,0%	,9%	0%/1,2%
	% der Gesamtzahl	,9%	,0%	,0%	,0%	,0%	,9%	0%/0,9%
Kinderkrankenpflege	Anzahl	96	0	0	0	0	96	2/94
	% Zeile	100,0%	,0%	,0%	,0%	,0%	100,0%	2,1%/97,9%
	% Spalte	16,6%	,0%	,0%	,0%	,0%	12,6%	1,2%/15,9%
	% der Gesamtzahl	12,6%	,0%	,0%	,0%	,0%	12,6%	0,3%/12,4%
Operationstechnische Assistenten	Anzahl	76	0	0	0	0	76	18/58
	% Zeile	100,0%	,0%	,0%	,0%	,0%	100,0%	23,7%/76,3%
	% Spalte	13,1%	,0%	,0%	,0%	,0%	10,0%	10,8%/9,8%
	% der Gesamtzahl	10,0%	,0%	,0%	,0%	,0%	10,0%	2,4%/7,6%
Bachelor/Master Studium	Anzahl	0	0	0	0	45	45	13/32
	% Zeile	,0%	,0%	,0%	,0%	100,0%	100,0%	28,9%/71,1%
	% Spalte	,0%	,0%	,0%	,0%	100,0%	5,9%	7,8%/5,4%
	% der Gesamtzahl	,0%	,0%	,0%	,0%	5,9%	5,9%	1,7%/4,2%
Gesamt	Anzahl	580	56	19	59	45	759	167/592
	% Zeile	76,4%	7,4%	2,5%	7,8%	5,9%	100,0%	22%/78%
	% Spalte	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100%/100%
	% der Gesamtzahl	76,4%	7,4%	2,5%	7,8%	5,9%	100,0%	22%/78%

Aus der Tabelle 13 lässt sich entnehmen, dass alle Schüler der Marie-Curie-Oberschule das Abitur anstreben. Das Geschlechterverhältnis ist hierbei nahezu ausgewogen (52,5% zu 47,5%), wobei die männlichen Probanden schon 18,6% des gesamten männlichen Geschlechts in der Stichprobe repräsentieren (4,1% im Verhältnis zu der Gesamtzahl der Probanden). Im Rahmen der beruflichen Ausbildung möchten lediglich fünf Probanden den Ausbildungsgang der Altenpflege am Vivantes Institut besuchen (0,7%), von denen vier männlichen Geschlechts sind (80%). Der Großteil der Probanden möchte den Abschluss zum (r) Gesundheits-und Krankenpfleger/in erwerben (N= 471). Die Probanden werden aus den drei genannten Schulen rekrutiert. Die Probanden bzw. die Bewerber um einen Ausbildungsplatz am Bildungsinstitut Vivantes (84,1%, N= 396) übersteigen hierbei deutlich die Probandengruppen der Wannseeschule (11,9%, N= 56)

und der Charité (4,0%, N= 19). Es fällt zudem auf, dass die Probanden der beiden zuletzt genannten Schulen ausschließlich das Ausbildungsziel der Gesundheits- und Krankenpflege verfolgen. Bezüglich des Verhältnisses zwischen den Geschlechtern zeigt sich, dass die weiblichen Probanden signifikant in der Überzahl sind (79% weibliche Probanden und 21% männliche Probanden). Unter Berücksichtigung der Vielzahl von Probanden in diesem Bildungsgang ist es im Vergleich zu den anderen Bildungsgängen jedoch auch ersichtlich, dass das männliche Geschlecht stärker vertreten ist (59,3% innerhalb des männlichen Geschlechts und 13% im Verhältnis zur Gesamtzahl der Probanden). Die nächste Gruppe hat sich am Vivantes Institut mit dem Ausbildungsziel zur Kinderkrankenpflege beworben (12,6%, N= 96). Das Geschlechterverhältnis offenbart eine auffällige Präsenz zu Gunsten des weiblichen Geschlechts (97,9%, N= 94 vs. 2,1%, N= 2). Hinsichtlich des Bildungsgangs zum Operationstechnischen Assistenten werden 76 Probanden vom Vivantes Institut in die Stichprobe inkludiert (10,0%). Auch hier überwiegt das weibliche Geschlecht (weiblich= 76,3%, N= 58 vs. männlich= 23,7%, N= 18). Die Gruppe der Bachelor- und Master-Studenten (das Hauptfach ist größtenteils Psychologie) umfasst 45 Probanden, die überwiegend weiblichen Geschlechts sind (71,1%, N= 32). Zusammenfassend lässt sich erkennen, dass die Majorität der angewendeten Gesamtstichprobe hauptsächlich aus dem Vivantes Institut für berufliche Bildung rekrutiert wurde, die Profession zum(r) Gesundheits- und Krankenpfleger(in) anstrebt und vorwiegend aus weiblichen Untersuchungsteilnehmerinnen besteht.

Im Hinblick auf die oben aufgeführten Forschungsfragen erscheint eine weitere Differenzierung bezüglich der unterschiedlichen Bildungsabschlüsse erstrebenswert. In der folgenden Abbildung 15 und Tabelle 14 wird das erreichte schulische Bildungsniveau wiedergegeben.

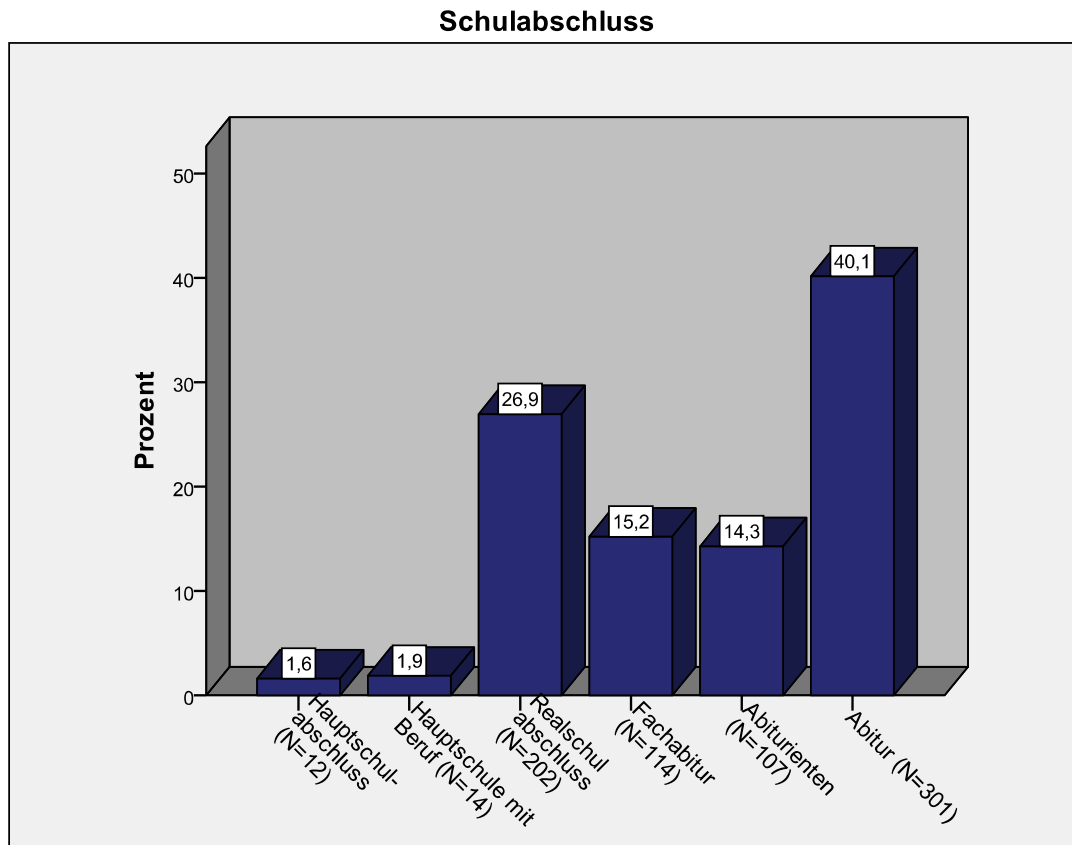


Abbildung 15: Verteilung der Probanden über die Bildungsstände in Prozent

Die Ergebnisse der Auswertungen zeigen, dass die Probanden mehrheitlich das Abitur erworben haben (40,1%, N= 301). 114 Teilnehmer haben das Fachabitur bestanden (15,2%). Insgesamt 107 (14,3%) Personen besuchen das Marie-Curie-Gymnasium mit dem Ziel des Erwerbs der allgemeinen Hochschulreife (Abitur). Ein weiterer Großteil, genauer 202 Probanden (26,9%) hat die mittlere Reife. Ein vergleichsweise geringer Anteil fällt auf die Personengruppen, die einen Hauptschulabschluss oder einen Hauptschulabschluss (1,6%, N= 12) mit abgeschlossener beruflicher Ausbildung (1,9%, N= 14) vorweisen können. In der nachstehenden Tabelle 14 wird aufgrund der im I-S-T 2000 R bewährten Clusterbildung zusätzlich eine Aggregation der oben angegebenen Abschlüsse in Gymnasiasten und Nicht-Gymnasiasten vorgenommen. Es ist offensichtlich, dass die Gymnasiasten in der vorliegenden Stichprobe die Mehrheit darstellen (69,6%, N= 522). Die Personengruppe der Nicht-Gymnasiasten umfasst 228 (30,4%) Probanden.

Des Weiteren lässt sich wie oben feststellen, dass die Relation zwischen den Geschlechtern innerhalb der Cluster „Bildungsabschlüsse“ zu Gunsten des weiblichen Geschlechts ausfällt. In einer Spannbreite von 66,4% bis 88,6% stellen sie in allen Untergruppen die Mehrheit dar.

Aufgrund der formulierten Forschungsannahmen werden in Tabelle 14 die verschiedenen Muttersprachen als Unterscheidungskriterium berücksichtigt. Die Differenzierung bezieht sich dabei auf die Probanden, die Deutsch als Muttersprache erlernt haben (Muttersprachler-D) und diejenigen, die Deutsch als Fremdsprache erlernt haben (Muttersprachler-F). Aus der Tabelle lässt sich entnehmen, dass die deutschen Muttersprachler in allen Bildungsgängen den Großteil darstellen (Muttersprachler-D: N= 661, 88,5%; Muttersprachler-F: N= 86, 11,5%). In den Kategorien Hauptschulabschluss (N= 3, 25%), Hauptschulabschluss mit Beruf (N= 5, 38,5%) und Realschulabschluss (N= 29, 14,4%) ist der Anteil an Probanden, die Deutsch als Fremdsprache gelernt haben, im Verhältnis am höchsten. Darüber hinaus ist zu konstatieren, dass sich der prozentuale Anteil der Muttersprachler-F in den genannten Kategorien schon auf 43% (kumulierte Prozent) bezogen auf die Gesamtstichprobe beläuft, obgleich nur 227 (30,4%) von 747 Probanden einen der aufgeführten Bildungsabschlüsse vorweisen. Bezüglich der Trennung nach der erlernten Muttersprache wird demnach deutlich, dass in den Klassen mit höheren Bildungsabschlüssen (Abiturienten, Fachabitur, Abitur) wesentlich weniger Muttersprachler-F als in den niedrigeren Bildungsklassen vertreten sind. Lediglich sieben von 107 Probanden mit dem Bildungsabschluss Fachabitur bzw. neun von 97 Abiturienten sind nicht mit Deutsch als Muttersprache aufgewachsen. 33 Probanden mit einer alternativen Muttersprache haben durch den Besuch des Gymnasiums die allgemeine Hochschulreife erworben. Insgesamt haben 71,3% der Muttersprachler-D einen höheren Bildungsabschluss erworben.

Tabelle 14: Verteilung der Probanden über die Bildungsabschlüsse getrennt nach Geschlecht und Muttersprache

Bildungsniveau		Häufigkeit Anteil %	Kumulierte Prozente	Geschlecht		Muttersprache	
				M (N) %Z/%S	W (N) %Z/%S	D (N) %Z/%S	F (N) %Z/%S
Gültig	Hauptschulabschluss	12 1,6%	1,6	3 25/1,8	9 75/1,5	9 75/1,4	3 25/3,5
	Hauptschulabschluss mit Beruf	14 1,8%	3,5	4 28,6/2,4	10 71,4/1,7	8 61,5/1,2	5 38,5/5,8
	Realschulabschluss	202 26,6%	30,4	38 18,8/22,9	164 81,2/28,1	173 85,6/26,2	29 14,4/33,7
	Fachabitur	114 15,0%	45,6	13 11,4/7,8	101 88,6/17,3	107 93,6/16,2	7 6,1/8,1
	Abitur	301 39,7%	85,7	72 23,9/43,3	229 76,1/39,2	267 89/40,4	33 11/38,4
	Abiturienten	107 14,1%	100,0	36 33,6/21,7	71 66,4/12,2	97 91,5/14,7	9 8,5/10,5
	Gesamt	750 98,8%		166 22/100	584 78/100	661 88,5/100	86 11,5/100
Fehlend	System	9 1,2%		9 1,2%		12 1,6%	
	Gymnasiasten	522 69,6%	69,6	121 23,2/72,9	181 80,3/31,3	471 90,6/71,3	49 9,4/57
	Nicht-Gymnasiasten	228 30,4%	100	45 19,7/27,1	401 76,8/68,7	190 83,7/28,7	37 16,3/43
Gesamt		759 100%		166 22/100	584 77,9/100	661 88,5/100	86 11,5/100

Anmerkung: Bedingt durch lückenhafte Angaben wurden für die Bildungsabschlüsse und die Kreuztabelle „Bildungsabschluss-Geschlecht“ lediglich 750 Probanden in die Kalkulation mit einbezogen. Für die Kreuztabelle „Bildungsabschluss-Muttersprache“ wurden 747 Probanden in die Berechnungen einbezogen.

D= Muttersprache „Deutsch“; F= Deutsch als Fremdsprache bzw. „Deutsch“ als Fremdsprache; %Z = prozentualer Anteil innerhalb der Zeile; %S= prozentualer Anteil innerhalb der Spalte.

## 9.4 Methodische Grundvoraussetzungen für die Datenanalyse

Das Skalenniveau einer Variablen ist in der Statistik und Empirie von zentraler Bedeutung. Eine sinnvolle Analyse und Interpretation von Variablen kann jedoch erst gewährleistet werden, wenn die erhobenen Daten ordinalskaliert sind. Hierbei können zumindest Rangordnungen getroffen werden (z.B. Schulnoten). Hinsichtlich der vorliegenden Daten des Wissenstests kann überdies davon ausgegangen werden, dass die Variablen



intervallskaliert sind. Anders ausgedrückt kann die Rangfolge und die Differenz von zwei Werten sachlich begründet werden (Backhaus, Erichson, Plinke & Weiber, 2003). Eine weitere Voraussetzung für statistische Berechnungen bildet eine genügend große Stichprobe, die in der einschlägigen Literatur vorwiegend mit einem Wert von  $N > 200$  angegeben wird (z.B. Lienert & Raatz, 1998). Die vorliegende Stichprobe umfasst 759 Probanden und erfüllt demnach das Kriterium.

Darüber hinaus sollten die zu analysierenden Variablen multivariat normalverteilt sein (Backhaus et al., 2003). Zunächst lässt sich bei einer genaueren Betrachtung der Gesamtrohwerte jedoch feststellen, dass weitere sechs Probanden als sogenannte „Ausreißer“ identifiziert werden. Die Ergebnisse der Testteilnehmer können sich im Weiteren verzerrend auf die Berechnungen auswirken (Bühner, 2006). Aus diesem Grund wurden die Probanden ebenfalls aus der Gesamtstichprobe ausgeschlossen (s. Kapitel 9.3.2). Die Gesamtstichprobe wurde daraufhin auf 753 Probanden reduziert. Im Hinblick auf die Verteilung der Gesamtrohwerte des START-W lässt sich erkennen, dass sich der Rohwert der einzelnen Probanden in einer Spanne zwischen 14 und 108 richtig gelöster Aufgaben bewegt (der selektierte Item-Pool umfasst 123 Testitems).<sup>5</sup> Der Reliabilitätskoeffizient für die gesamte Wissensskala weist einen Wert von .923 auf. Die Höhe des Koeffizienten lässt den Schluss zu, dass das Maß der Genauigkeit, mit der das Merkmal durch den Test erfasst wird, hoch ist (vgl. dazu Bühn, 2006, S. 477). Der Mittelwert der Gesamtstichprobe beträgt dabei 62,09 richtig gelöste Aufgaben und die Standardabweichung (SD) wird mit einem Wert von 17,37 angegeben. Bezüglich der Prüfung auf Normalverteilung wurden die Kennwerte *Schiefte* und *Exzess (Kurtosis)* herangezogen. Hinsichtlich der multivariaten Normalverteilung geben Lienert und Raatz (1998) einen Richtwert der Schiefe von +/- 0.5 an und für die Kurtosis sollen die Kennwerte in einem Bereich zwischen +1 und -1 liegen. Auf Basis dieser Richtwerte sind die Gesamtrohwerte des selektierten Item-Pools mit einer Schiefe von - 0.023 und einer Kurtosis von - 0.371 normalverteilt (siehe Abbildung 16).<sup>6</sup>

---

<sup>5</sup> Die Statistiken des unselektierten Item-Pools werden im Anhang A aufgeführt. Der selektierte Item-Pool wurde auf der Basis der folgenden Berechnungen und Ergebnissen generiert (siehe Kapitel 10.1).

<sup>6</sup> Der Kolmogorov-Smirnov Test zeigt ein deutlich nicht signifikantes Ergebnis ( $K-S-Z = .915$ ;  $p = .372$ ), so dass die Rohwerte in der Gesamtstichprobe hinreichend normalverteilt sind.

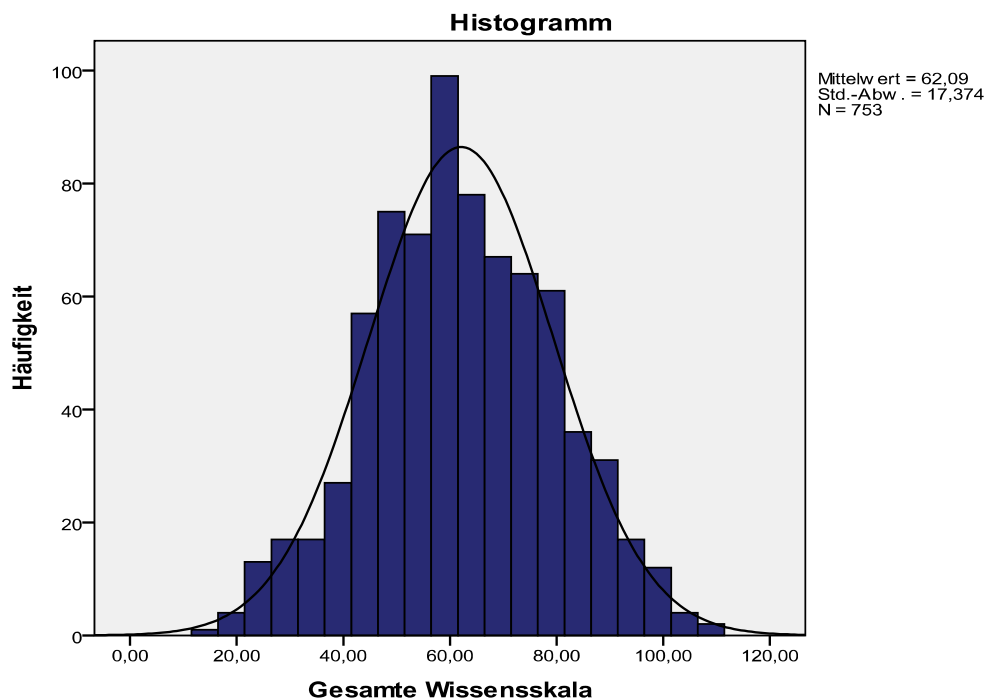


Abbildung 16: Multivariate Verteilung der Gesamtrohwerte

## 9.5 Methodische Vorgehensweise bei der Datenanalyse

Die Grundlage für die Erläuterung der methodischen Vorgehensweise bildet die deskriptive Statistik und die Prüfung der methodischen Grundvoraussetzungen (wie z.B. die Normalverteilung der Gesamtrohwerte)

Im Hinblick auf die Untersuchung der Forschungsfragen 2 und 3 und den dazugehörigen Forschungsannahmen werden zunächst Item- und Reliabilitätsanalysen der verbal, numerisch und figural kodierten Aufgabengruppen durchgeführt, um „schlechte“ Items mit geringen Trennschärfen zu selektieren. Bezüglich der vorgesehenen Itemanalyse werden angemessene Schwierigkeits- und Trennschärfeindizes berücksichtigt. (Cronbach, 1951). Es wird das Ziel verfolgt, möglichst homogene Aufgabengruppen bzw. Skalen zu erhalten. Als Kennwert für die interne Konsistenz wird *Cronbach's Alpha* zugrunde gelegt, um die Reliabilität der Skalen zu prüfen. Die interne Konsistenz, gilt als eine Erweiterung der Testhalbierungsmethode und führt gemäß den Ausführungen von Bortz und Döring (2002) zu stabileren Schätzungen. Anschließend werden die genannten Aufgabengruppen auf Normalverteilung geprüft. Darauf aufbauend wird zusätzlich die Methode der Split-Half -Reliabilität verwendet. Der Betrag der Reliabilitätskoeffizienten, der durch die Testhalbierung verloren geht, wird Spearman-

Brown korrigiert. Darüber hinaus soll die Reliabilität des Testverfahrens anhand eines Retests überprüft werden. Aufgrund der zeitlichen Stabilität der Personenmerkmale sind Querschnittszusammenhänge ausreichend und angemessen.

Für die Bearbeitung der folgenden Forschungsfragen werden multivariate Analysemethoden herangezogen. Im Rahmen des ersten Forschungsansatzes wird auf Basis der postulierten Interaktion zwischen Umwelt und Wissen der Frage nachgegangen, ob die Moderatorvariablen Alter, Muttersprache, Geschlecht und das erreichte Bildungsniveau einen Einfluss auf das Wissen haben. Um die unterschiedliche Wissensausprägung über die Bildungsklassen und das Alter zu überprüfen, wird einerseits eine parametrische Varianzanalyse vorgenommen und andererseits wird der Post-Hoc-Mehrfachvergleich nach Sheffé durchgeführt. Zur Absicherung der Ergebnisse wird zusätzlich der nichtparametrische H-Test (Kruskal & Wallis) angewendet. Es kann an dieser Stelle darauf hingewiesen werden, dass vor allem die Analyse über den Einfluss des Bildungsgangs auf die Wissensakkumulation bereits einen verwertbaren Hinweis auf die Validität des Wissenstests gibt. Die zweite und dritte Forschungsannahme, die den Einfluss der Muttersprache bzw. des Geschlechts auf die Wissensausprägung untersucht, wird mittels des T-Tests bei unabhängigen Stichproben überprüft. Die Berücksichtigung des anschließenden nichtparametrischen U-Tests (Mann & Whitney) dient wiederum der Absicherung. Abschließend werden die Befunde inhaltlich interpretiert, diskutiert und infolgedessen ergänzende Regressionsanalysen als Erklärung einbezogen.

Im Zuge der zweiten Forschungsfrage wird die kriterienbezogene Validität des Wissenstests ermittelt. Die angegebenen Schulnoten bilden in diesem Fall das Außenkriterium. Anhand von bivariaten Korrelationen werden demnach die Zusammenhänge zwischen Schulnoten und den Wissensskalen analysiert. Aufgrund der unterschiedlich „gepolten“ Bewertungsskalen werden des Weiteren negative Zusammenhänge angenommen. Hinsichtlich der Forschungsfrage 2.2 wird die prognostische Validität überprüft. Hierzu werden die Durchschnittsnoten des ersten Ausbildungsjahres im Bereich der Gesundheits- und Krankenpflege als Außenkriterium verwendet. In diesem Kontext kommen die gleichen Analysemethoden zur Anwendung.

Im Rahmen der Konstruktvalidität (Forschungsfrage 3) werden die konvergente und die faktorielle Validität des Wissenstests betrachtet. Zur Bestimmung der konvergenten Validität wird der Zusammenhang zwischen dem Bochumer Wissenstest (BOWIT-Kurzform; Hossiep & Schulte, 2008) und dem START-W mittels korrelationsstatistischen Methoden berechnet. Bezüglich der Untersuchung der faktoriellen Validität ist zu

beachten, dass der Wissenstest START-W auf Basis des Erweiterungsmoduls des I-S-T 2000 R weiterentwickelt wurde, aber weder in seiner Binnenstruktur noch in seinem Aufbau mit diesem identisch ist. Gleichwohl ist die Möglichkeit zur Berechnung fluider und kristallisierter Intelligenz in Kombination mit dem Grundmodul geblieben, die im I-S-T 2000 R allerdings als schlussfolgerndes Denken/gf (fluide Intelligenz) bzw. Wissen/gc (kristallisierte Intelligenz) bezeichnet werden. Außerdem kann man davon ausgehen, dass die drei Inhaltsfacetten (verbal, numerisch, figural) sowie die Themenbereiche ebenfalls repliziert werden können. Zur Überprüfung der Fragestellungen 3.2 und 3.3, bzw. um die intendierte Struktur zu eruieren und den postulierten Zusammenhang zwischen Wissen und Intelligenz zu untersuchen, werden ebenfalls multivariate Analysemethoden verwendet. Den Ansatzpunkt für die Überprüfung der Faktorenstruktur liefern die Interkorrelationen der sieben Themenbereiche. Dasselbe gilt für die Skalenebene bzw. für die verbale, numerische und figurale Inhaltsfacette. Die Kreuzkorrelationsmatrix der sieben Aufgabengruppen und der Inhaltsskalen bildet die Basis für die Ermittlung der faktoriellen Struktur des Wissenstests START-W. Die Struktur soll sowohl durch eine Dimensionsanalyse (Hauptachsenmethode; schiefwinkliges Rotationsverfahren, oblimin), multidimensionale Ähnlichkeitsstrukturanalysen (MDS) als auch durch eine konfirmatorische Faktorenanalyse (LISREL 8. 30) für die vorliegende Stichprobe bestätigt werden<sup>7</sup>.

## 10 Darstellung der Ergebnisse

### 10.1 Item- und Reliabilitätsanalysen

Der Wissenstest START-W wurde so konzipiert, dass neben dem Gesamtrohwert auch die Rohwerte für das verbale, numerische und figurale Wissen ermittelt werden können. Die Tabelle 15 gibt die Zuordnung der Testitems zu den Inhaltsfacetten verbales Wissen, numerisches Wissen und figurales Wissen sowie zu den postulierten Themenbereichen wieder. Die angegebenen Aufgabennummern beziehen sich in diesem Fall auf die Testversion Form A des START-W. Auf Basis dieser Zuordnung wird im Weiteren eine Aufgabenanalyse durchgeführt. Die Aufgaben- bzw. Skalenanalyse dient dazu,

---

<sup>7</sup> Es ist zu beachten, dass der Nachweis von wissensthemen Facetten und Facetten zur verbalen, figuralen und numerischen Kodierung (Dimensionen) nicht ohne weiteres mit exploratorischen Faktorenanalysen (Hauptachsenanalysen mit Kommunalitäteniteration) durchgeführt werden kann.

qualitativ „schlechte“ Items aus dem Wissenstest zu eliminieren, um die Reliabilität des Verfahrens zu optimieren. Die Güte einer Testaufgabe wird in diesem Zusammenhang durch das Kriterium der *Trennschärfe* bestimmt (Lienert & Raatz, 1998). Die Trennschärfe gibt im Grunde genommen an, wie „gut das gesamte Testergebnis aufgrund der Beantwortung eines einzelnen Items vorhersagbar ist“ (Bortz & Döring, 2002, S.219). „Gute“ Items sind demnach durch hohe Trennschärfen gekennzeichnet, die in einem Wertebereich zwischen +1 und -1 liegen. Die korrelationstypischen Trennschärfekoeffizienten, die  $< .30$  sind, liegen dabei im niedrigen Bereich und Trennschärfen, die  $> .50$  sind, werden im hohen Bereich angesiedelt. Aus diesen Gründen wurden sämtliche Testitems entfernt, die einen Trennschärfekoeffizienten von  $< .10$  aufwiesen. Im Weiteren wird, gemäß der klassischen Testtheorie, die psychometrische Qualität der Aufgabenzusammenstellung anhand der Aufgabenschwierigkeit evaluiert. Zur anschließenden Bestimmung der Reliabilität wird der Alpha-Koeffizient von Cronbach (1951) herangezogen, der als Maß für die Homogenität der Skalen verstanden werden kann. Cronbachs Alpha ist demnach ein Kennwert, der die innere Konsistenz der Skalen wiedergibt (Cronbach, 1951). Der Wertebereich liegt hierbei zwischen 0 und 1, wobei in der einschlägigen Literatur häufig ein Mindestwert von  $\alpha = .80$  angegeben wird.

*Tabelle 15: Zuordnung der Test-Items zu den Inhaltsfacetten und Themenbereichen*

Wissensbereich	Kodierungsform		
	Verbales Wissen	Numerisches Wissen	Figurales Wissen
Geographie/ Geschichte	1, 16, 20, 35, 56, 87, 127	14, 25, 33, 51, 63, 96, 98, 107, 119, 126	26, 55, 64, 100, 128
Wirtschaft	38, 68, 72, 75, 93, 108, 109, 122	11, 42	6, 46, 74, 110, 123, 134
Kunst/Kultur	10, 18, 29, 32, 50, 88, 113, 116,	36, 39, 54, 57, 60, 76, 79, 112, 129	9, 12, 28, 31, 67, 77, 81, 82, 97, 115, 125, 131
Mathematik	41, 47, 59, 62, 65, 90, 103	2, 8, 21, 27, 30, 45, 48, 66, 73, 89, 102, 130	15, 34, 37, 40, 43, 52, 118, 133
Naturwissenschaft	4, 7, 44, 78, 91, 117, 120	5, 53, 94, 105	22, 49, 58, 70, 92, 101, 121
Alltagswissen	13, 24, 80, 83, 84, 85, 114	17, 69, 71, 95, 104, 111, 124	3, 19, 23, 61, 86, 99, 106, 132
START-W	44 Testitems	44 Testitems	46 Testitems
Gesamt	134 Items		

**Item-Analyse der verbalen Wissensskala**

Im Anhang A werden die Item- bzw. Reliabilitätsanalysen des unselektierten Item-Pools dargelegt. Die folgende Tabelle 16 zeigt den selektierten Item-Pool (40 Testitems). Mit anderen Worten wurden die Items mit geringen Trennschärfen bereits entfernt, so dass die Reliabilitätsschätzung einen Wert von  $\alpha = .797$  erreicht. Die geringere Reliabilität ( $< .80$ ) kann damit begründet werden, dass die Aufgaben unterschiedlichen Themenbereichen entstammen, so dass eine hohe Homogenität der Aufgabengruppe nicht unbedingt gewährleistet werden kann. Aufgrund der Anzahl der Testitems liegt der Wert damit in einem akzeptablen Bereich. Die Trennschärfen rangieren zwischen .108 und .461 und liegen damit gemäß der oben angeführten Kriterien ebenfalls in einem positiven und annehmbaren Bereich. Man kann demnach davon ausgehen, dass die Aufgabengruppe „verbales Wissen“ annähernd homogen ist. Der Schwierigkeitsindex streut zwischen .1713 (Item 91) und .8977 (Item 83).

*Tabelle 16: Kennziffern der Aufgabengruppe „verbales Wissen“*

Item Nr.	Mittelwert Schwierigkeit	Korrigierte Item-Skala- Korrelation Trennschärfe	Cronbachs Alpha, wenn Item weggelassen
ri01	,3679	,218	,795
ri07	,6985	,291	,792
ri10	,3732	,196	,796
ri13	,5219	,153	,797
ri16	,3147	,166	,797
ri18	,4754	,265	,793
ri20	,2961	,182	,796
ri24	,8274	,314	,792
ri29	,4887	,181	,796
ri32	,4303	,160	,797
ri35	,7344	,178	,796
ri38	,5166	,330	,791
ri41	,3865	,327	,791
ri44	,5206	,311	,791
ri47	,6985	,388	,789
ri50	,7875	,287	,792
ri59	,3347	,126	,798
ri62	,6428	,196	,796
ri65	,4077	,239	,794
ri68	,7716	,389	,789

ri72	,6348	,280	,793
ri75	,5591	,147	,797
ri78	,5113	,461	,786
ri80	,7968	,328	,791
ri83	,8977	,318	,792
ri84	,5259	,195	,796
ri87	,4967	,305	,792
ri88	,3280	,291	,792
ri90	,3028	,208	,795
ri91	,1713	,115	,797
ri93	,2457	,108	,798
ri103	,7450	,291	,792
ri109	,5564	,340	,790
ri113	,2855	,367	,790
ri114	,7782	,373	,790
ri116	,2151	,341	,791
ri117	,5100	,440	,786
ri120	,7317	,430	,787
ri122	,3652	,260	,793
ri127	,2444	,266	,793
Anzahl der Items: 40			<b>Cronbachs Alpha: .797</b>

Die Rohwerte der verbalen Wissensskala liegen zwischen 4 und 37 richtig gelöster Aufgaben (von insgesamt 40 Testitems) und der Mittelwert beträgt  $M = 20,49$  ( $SD = 6,18$ ). Die Schiefe von  $.017$  gibt an, dass die Verteilung der Rohwerte leicht rechtssteil bzw. nahezu symmetrisch ist. Die Kurtosis (Exzess) weist einen Wert von  $-.371$  auf. Die Verteilung der Rohwerte kann gemäß den aufgeführten Kriterien als normalverteilt angenommen werden. (siehe Abbildung 17).<sup>8</sup>

<sup>8</sup> Der Kolmogorov-Smirnov-Test zeigt ein signifikantes Ergebnis und bestätigt somit nicht die Normalverteilung der Rohwerte in der verbalen Wissensskala:  $K-S-Z = 1,516$ ,  $Sig. = .020$ .

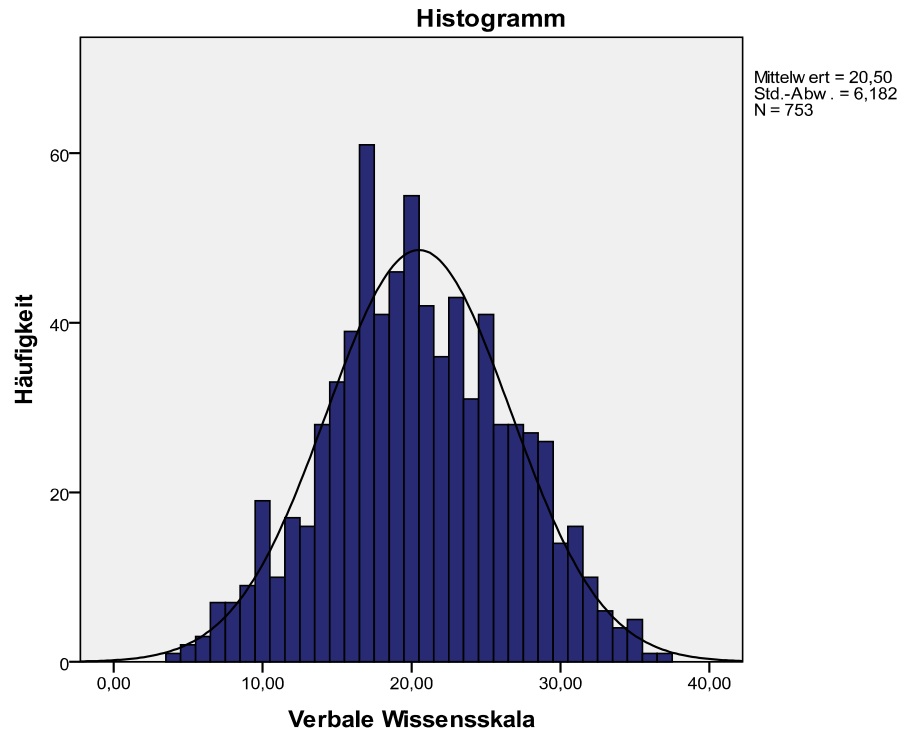


Abbildung 17: Rohwertverteilung der Aufgabengruppe „verbales Wissen“

#### ***Item-Analyse der numerischen Wissensskala***

In Tabelle 17 werden ebenfalls die bereits selektierten Items der Aufgabengruppe wiedergegeben (37 Testitems).<sup>9</sup> Aufgrund der Eliminierung der Items mit niedrigen Trennschärfen ergibt sich ein Cronbachs Alpha in Höhe von  $\alpha = .778$ . Der Wert erreicht wiederum nicht den oben angegebenen Mindestwert von  $\alpha = .80$ . Die Ursache liegt u.a. in der verbleibenden inhaltlichen Heterogenität der Testitems begründet. Eine weitere Reduzierung der Testitems würde in diesem Fall jedoch keine Erhöhung der Reliabilität nach sich ziehen. 13 der 28 Items weisen gemäß den Kriterien eine mittlere Trennschärfe auf. Die restlichen Items zeigen niedrige Trennschärfen (.107 bis 494). Der Schwierigkeitsindex, der den prozentualen Anteil der richtigen Antworten repräsentiert, bewegt sich im Bereich von 10,36% (Item 53) und 85,39% (Item 76).

---

<sup>9</sup> Im Anhang A werden die Kennzahlen des unselektierten Item-Pools dargestellt.



Tabelle 17: Kennziffern der Aufgabengruppe „numerisches Wissen“

Item Nr.	Mittelwert Schwierigkeitsindex	Korrigierte Item-Skala- Korrelation Trennschärfe	Cronbachs Alpha, wenn Item weg- gelassen
ri05	,3108	,211	,775
ri08	,1607	,253	,773
ri11	,2510	,149	,777
ri14	,4011	,227	,774
ri21	,3174	,331	,770
ri25	,5060	,156	,778
ri27	,6999	,339	,769
ri30	,2616	,339	,770
ri36	,5139	,241	,774
ri39	,4568	,207	,775
ri42	,3785	,256	,773
ri45	,8287	,306	,771
ri48	,4967	,417	,766
ri51	,3639	,224	,774
ri53	,1036	,175	,776
ri54	,2125	,136	,777
ri57	,6946	,262	,773
ri60	,4369	,178	,776
ri63	,4117	,172	,777
ri66	,3201	,252	,773
ri69	,7955	,273	,772
ri71	,7809	,171	,776
ri73	,4050	,284	,772
ri76	,8539	,180	,776
ri79	,6614	,306	,771
ri89	,7304	,309	,771
ri95	,4290	,107	,780
ri96	,5737	,293	,771
ri102	,6547	,494	,762
ri104	,5883	,310	,771
ri105	,1912	,112	,778
ri111	,5604	,421	,765
ri112	,7928	,332	,770
ri119	,2019	,184	,776
ri124	,3758	,142	,778

ri126	,2882	,356	,769
ri130	,5618	,451	,764
Anzahl der Test-items: 37			<b>Cronbachs Alpha: .778</b>

Die Anzahl der richtig gelösten Aufgaben rangiert in einem Bereich zwischen 3 und 34 Aufgaben (von 37 Testitems). Der Mittelwert beträgt  $M = 17,57$  und die Standardabweichung nimmt den Wert  $SD = 5,61$  an. Bezüglich der Prüfung der Rohwerte auf ihre Normalverteilung wurden auch hier die Kennwerte Schiefe und Kurtosis beachtet. Der Schiefewert beträgt  $.168$  und die Kurtosis zeigt einen Wert von  $-.307$ . Entsprechend der Beurteilungsrichtlinien lässt sich sagen, dass die Rohwertverteilung in der Inhaltsfacette normalverteilt ist (siehe Abbildung 18).<sup>10</sup>

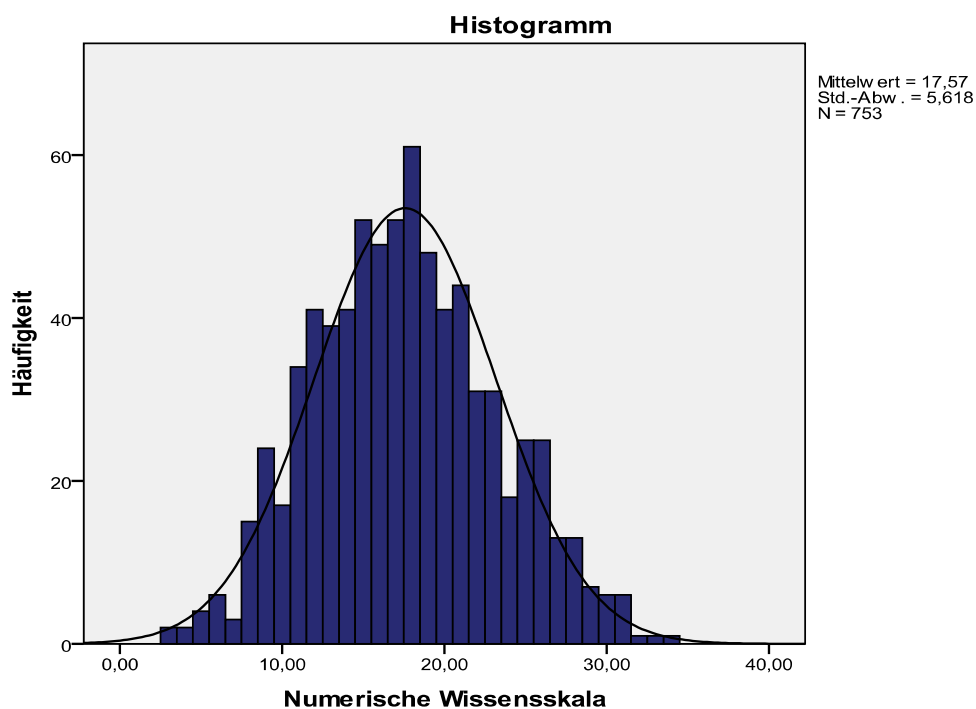


Abbildung 18: Rohwertverteilung der Aufgabengruppe „numerisches Wissen“

<sup>10</sup> Der Kolmogorov-Smirnov-Test zeigt ein signifikantes Ergebnis und bestätigt somit nicht die Normalverteilung der Rohwerte in der numerischen Wissensskala:  $K-S-Z = 1,552$ ,  $Sig. = .016$ .

**Item-Analyse der figuralen Wissensskala**

Im Rahmen der Aufgabenanalyse weisen die bereits selektierten Items (46 Testitems) der Aufgabengruppe „figurales Wissen“ eine interne Konsistenz von  $\alpha = .832$  auf. Die Skala weist im Vergleich zu den anderen Skalen den höchsten Reliabilitätskoeffizienten auf und übersteigt zudem den geforderten Wert von  $.80$ . Die Trennschärfen liegen größtenteils im niedrigen Bereich. 21 Items zeigen jedoch Trennschärfen auf, die  $> .30$  sind (siehe Tabelle 18). Der Mittelwert (Schwierigkeitsindex) der Items rangiert zwischen  $.1633$  (Item 09) und  $.9628$  (Item 23).

Tabelle 18: Kennziffern der Aufgabengruppe „figurales Wissen“

Item Nr.	Mittelwert Schwierigkeitsindex	Korrigierte Item-Skala- Korrelation Trennschärfe	Cronbachs Alpha, wenn Item weggelassen
ri03	,7357	,160	,832
ri06	,6892	,192	,831
ri09	,1633	,242	,830
ri12	,7238	,329	,828
ri15	,6388	,354	,827
ri19	,3174	,176	,831
ri22	,7849	,349	,827
ri23	,9628	,146	,831
ri26	,4914	,298	,828
ri28	,3639	,208	,831
ri31	,3174	,183	,831
ri34	,5219	,284	,829
ri37	,3479	,247	,830
ri40	,4210	,110	,833
ri43	,4967	,272	,829
ri46	,2085	,199	,831
ri49	,1713	,126	,832
ri52	,3984	,332	,827
ri55	,3878	,306	,828
ri58	,4874	,277	,829
ri61	,8778	,303	,829
ri64	,5724	,305	,828
ri67	,2869	,167	,832
ri70	,4661	,392	,826
ri74	,5750	,185	,831
ri77	,6534	,235	,830

ri81	,8114	,283	,829
ri82	,3134	,343	,827
ri86	,9256	,247	,830
ri92	,7185	,345	,827
ri97	,2364	,227	,830
ri99	,8579	,354	,827
ri100	,3572	,295	,828
ri101	,8858	,363	,828
ri106	,7981	,382	,826
ri110	,1899	,274	,829
ri115	,4382	,319	,828
ri118	,6467	,327	,828
ri121	,8127	,465	,825
ri123	,5378	,490	,823
ri125	,3665	,291	,828
ri128	,5259	,429	,825
ri131	,5113	,406	,825
ri132	,3825	,248	,830
ri133	,2696	,309	,828
ri134	,3785	,352	,827
Anzahl der Testitems: 46			Cronbachs Alpha: .832

Zur Beurteilung der Verteilung der Rohwerte wurden nochmals die Schiefe und die Kurtosis ermittelt. Der Kennwert für die Schiefe beträgt  $-.097$  (leicht linkssteil) und die Kurtosis nimmt einen Wert von  $-.449$  an.

Die Berechnung des Mittelwertes ergibt einen Wert von  $M = 24,02$  ( $SD = 7,04$ ). Die psychometrischen Kennzahlen liegen damit im akzeptablen Bereich und können als normalverteilt betrachtet werden (siehe Abbildung 19).<sup>11</sup>

<sup>11</sup> Der Kolmogorov-Smirnov-Test bestätigt die Normalverteilung der Rohwerte in der figuralen Wissensskala:  $K-S-Z = 1,188$  Sig. =  $.119$ .

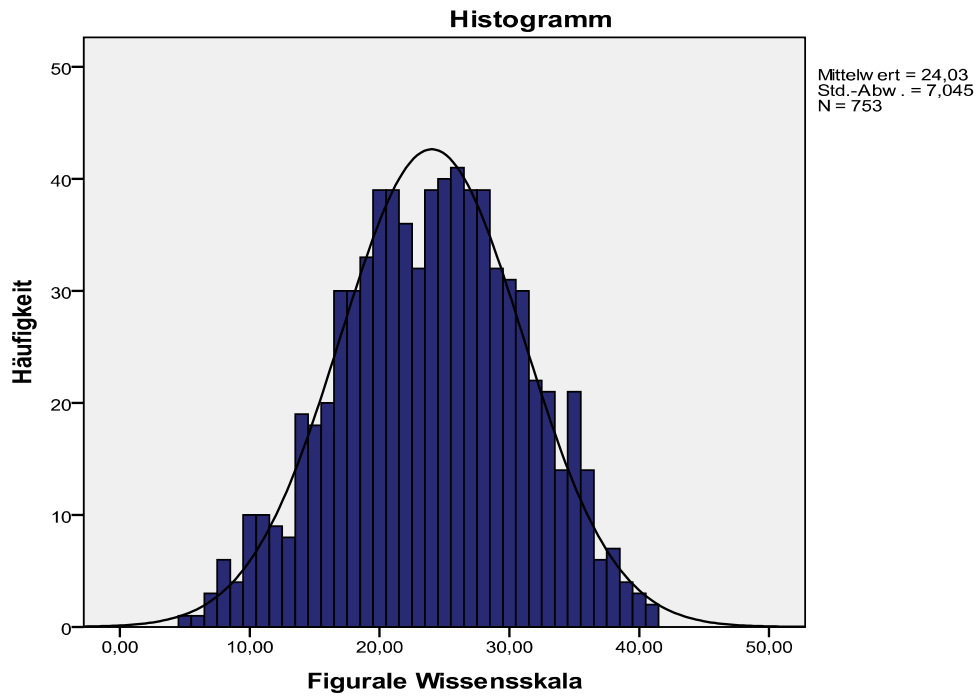


Abbildung 19: Rohwertverteilung der Aufgabengruppe „figurales Wissen“

### ***Testhalbierungsmethode***

Aufgrund der den Leistungstests inhärenten Testkonstruktion, genauer des ansteigenden Schwierigkeitsgrades der Testitems, wurde der Test nach der *odd-even-Methode* halbiert. Die Schätzwerte wurden anschließend *Spearman-Brown* korrigiert und werden in der nachfolgenden Übersicht (Tabelle 19) dargelegt. Die *Split-half-Reliabilitäten* liegen für die Wissensskalen zwischen .785 und .929 und liegen damit in einem akzeptablen Bereich.

Tabelle 19: Split-Half-Reliabilitätskoeffizienten

	Verbal	Numerisch	Figural	Gesamt
Cronbachs Alpha	.797	.778	.832	.923
Split Half	.809	.785	.823	.929

### ***Testwiederholungsmethode***

In der vorliegenden Studie basierten die Retest-Untersuchungen auf einem drei- bis sechsmonatigen Zeitintervall bei dem erneut der START-W eingesetzt wurde. Die Testergebnisse des Pretests und des Retests werden daraufhin korreliert und die Reliabilitätskoeffizienten bestimmt. Die Stichprobe umfasst insgesamt 101 Probanden (30 Männer, 71 Frauen im Alter von 16 bis 30 Jahren).<sup>12</sup> Die Retest-Reliabilitäten liegen zwischen  $r = .821$  und  $r = .888$  und sind auf dem Niveau von 0.01 (2-seitig) signifikant (siehe Tabelle 20). Vor dem Hintergrund der hohen Reliabilitätskoeffizienten kann davon ausgegangen werden, dass der START-W das Konstrukt „Wissen“ mit einer hohen zeitlichen Stabilität erfasst bzw. das Merkmal eine hohe Konstanz aufweist. Im Hinblick auf die Mittelwerte der beiden Testformen lässt sich zwar nur ein leichter Zuwachs des Wissensniveaus feststellen, der auf der Basis von T-Tests jedoch signifikant ausfällt.

*Tabelle 20: Retest-Reliabilität*

	Pretest und Retest		
	Korrelationsstatistik	Mittelwerte	
		Pretest	Retest
Verbale Wissensskala	$r = .821^*$	21,82	22,56
Numerische Wissensskala	$r = .835^*$	18,82	19,69
Figurale Wissensskala	$r = .835^*$	25,46	26,51
Gesamte Wissensskala	$r = .888^*$	66,1	68,77

*Anmerkung: Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant; Gesamtzahl der Testitems = 123; Testitems in der verbalen, numerischen und figuralen Skala: 40, 37, 46.*

<sup>12</sup> Der Kolmogorov-Smirnov-Test bestätigt die Normalverteilung der Rohwerte in allen Wissensskalen: verbale Skala: K-S-Z = .599 Sig. = .866; numerische Skala: K-S-Z = .609 Sig. = .852; figurale Skala: K-S-Z = .799 Sig. = .546; gesamte Wissensskala: K-S-Z = .616 Sig. = .843.

## 10.2 Darstellung der Ergebnisse zu den Forschungsfragen- Teil 1

### *Darstellung der Ergebnisse zur Forschungsfrage und Forschungsannahme- 1.1*

Hinsichtlich der Forschungsfrage 1.1 soll der Einfluss des Alters auf die Wissensausprägung untersucht werden. Die zugrunde liegende Forschungsannahme ist, dass der Wissensumfang mit steigendem Alter zunimmt. In einem ersten Schritt wurden daher die Probanden, in Anlehnung an die Einteilung des I-S-T 2000 R, auf fünf Altersgruppen verteilt (siehe Tabelle 21). Auf Basis der gebildeten Altersgruppen wurden im zweiten Schritt die Mittelwerte der Testergebnisse (hier: Mittelwert der richtig gelösten Aufgaben) verglichen, um mögliche Hinweise auf den vermuteten Einfluss zu finden. Aus der Tabelle 21 lässt sich entnehmen, dass sich die Altersgruppen 3 und 4 bezüglich der Wissensausprägung deutlich von den anderen drei Altersgruppen abhebt. Die erste und zweite Altersgruppe unterscheiden sich ebenfalls auffällig in den aufgeführten Wissensskalen. Die zweite und fünfte Altersgruppe weisen hingegen nur marginale Unterschiede auf. Dementsprechend können weitgehend drei homogene Gruppen gebildet werden. Die erste neu generierte Gruppe entspricht der Altersgruppe 1, die zweite umfasst die Altersgruppen 2 und 5 und die dritte Gruppe schließt die Altersgruppe 3 und 4 ein. In einer abschließenden Betrachtung der aufgezeigten Mittelwerte lässt sich festhalten, dass die Altersgruppe 3 in der verbalen, numerischen und figuralen Wissensskala sowie in der gesamten Wissensskala stets bessere Ergebnisse erzielen.

*Tabelle 21: Mittelwerte und Standardabweichungen der Testergebnisse getrennt nach Altersgruppen*

Altersgruppen	Gesamte W-Skala		Verbale Wissensskala		Numerische Wissensskala		Figurale Wissensskala	
	M	SD	M	SD	M	SD	M	SD
Altersgruppe 1 (16-Jährige; N= 35)	50,8	14,0	15,80	4,56	15,37	5,49	19,68	5,82
Altersgruppe 2 (17-18-Jährige; N=142)	58,0	14,4	18,33	5,40	16,35	4,73	23,34	6,08
Altersgruppe 3 (19-20-Jährige; N=321)	65,4	15,3	21,74	5,38	18,65	5,22	25,29	6,30
Altersgruppe 4 (21-25-Jährige; N=179)	63,9	18,0	21,65	6,24	17,81	5,74	24,51	7,26
Altersgruppe 5 (26-46-Jährige; N=76)	56,0	24,2	19,56	8,63	15,71	7,22	20,76	9,46
Gesamt (N= 753)	62,0	17,3	20,49	6,18	17,57	5,61	24,02	7,04

*Anmerkung:* M=Mittelwert; SD=Standardabweichung; Aufgrund von fehlenden Angaben wurden nur 753 Probanden in die Berechnung mit einbezogen.

Zur Überprüfung der Wissensunterschiede hinsichtlich der verschiedenen Altersgruppen wird im Weiteren eine univariate Varianzanalyse (bzw. einfaktorielle ANOVA) durchgeführt werden. Bevor die parametrische Varianzanalyse angewendet werden kann, werden mindestens zwei grundlegende Vorbedingungen analysiert:

1. Homogenität der Varianzen
2. Normalverteilung der Rohwerte in den Substichproben

Der Levene-Test auf Gleichheit der Varianzen erreicht in allen Fällen ein signifikantes Ergebnis. Das bedeutet, dass die Varianzhomogenität nicht gegeben ist. Der Kolmogorov-Smirnov-Test (K-S-Test) zeigt, dass die Normalverteilung der Werte in der gesamten sowie in den kodierten Wissensskalen für alle Altersgruppen gegeben ist ( $p > .05$ , siehe Anhang B1.1). Aufgrund des Ergebnisses des K-S-Tests und der anschließenden Kontrolle der Boxplots (Konfidenzintervalle), Standardabweichungen und Standardfehler, die augenscheinlich keine gravierenden Unterschiede in der Streuung aufweisen, ist es dennoch möglich und plausibel, den Testansatz für gleiche Varianzen als Kriterium heranzuziehen und zur Absicherung der Ergebnisse das Testverfahren ohne Annahme gleicher Varianzen (Tamhane-T2) zu berücksichtigen.

Das Ergebnis der parametrischen Varianzanalyse bestätigt den bereits dargestellten Befund, dass das Wissen in den Altersgruppen 3 und 4 stärker ausgeprägt ist, als in den anderen Altersgruppen (siehe auch Tabelle 21 und 22). Darüber hinaus ist es ersichtlich, dass die Unterschiede in der Wissensausprägung sogar signifikante Kennwerte aufzeigen. Die Gesamtskala weist einen Kennwert von  $F = 12,18$  ( $df = 4$ ) auf, der signifikant ist ( $p < .001$ ). Für die verbale, numerische und figurale Wissensskala resultieren in der angegebenen Reihenfolge auch hochsignifikante Kennwerte ( $p < .01$ ) von  $F = 14,69$  ( $df = 2$ ),  $F = 8,45$  ( $df = 2$ ) und  $F = 11,13$  ( $df = 2$ ).

Wenngleich der K-S-Test angibt, dass die Rohwerte der Wissensskalen innerhalb der Altersgruppen normalverteilt sind, überschreiten einige Schiefe- und Kurtosiskennwerte den von Lienert und Raatz (1998) angegebenen Bereich der Normalverteilung (vgl. dazu Anhang B1.1). Hinzu kommt, dass die Probandenanzahl in der Altersgruppe 1 sehr gering ist und die Probanden über die einzelnen Altersgruppen ungleich verteilt sind. Demzufolge kann der Annahme gefolgt werden, dass die Vorbedingung der Normalverteilung verletzt ist und die Teststärke der parametrischen Varianzanalyse beeinträchtigt wird. Hinsichtlich der anschließenden Interpretation ist es daher von Nutzen, die Anwendung des nichtparametrischen H-Tests (Kruskal & Wallis) durchzuführen.



Das Ergebnis des H-Tests bekräftigt den oben dargelegten Befund, dass sich die verschiedenen Altersklassen bezüglich des Allgemeinwissens signifikant ( $p < .05$ ) unterscheiden. Die Kennwerte für H-Test zeigen für die gesamte Wissensskala einen Chi-Quadrat-Wert von  $\chi^2 = 51,364$  ( $p = .000$ ), für die verbale Wissensskala  $\chi^2 = 61,709$  ( $p = .000$ ), für die numerische Wissensskala  $\chi^2 = 37,526$  ( $p = .000$ ) und für die figurale Wissensskala  $\chi^2 = 43,883$  ( $p = .000$ ).

Des Weiteren wurden der Scheffé-Test (Post-Hoc-Test) und der Tamhane-T2 (Testansatz für ungleiche Varianzen) durchgeführt, um die signifikanten Wissensunterschiede zwischen den Altersgruppen sichtbar zu machen.

Unter Verwendung des Scheffé-Tests zeigt sich hinsichtlich der gesamten Wissensskala, dass zwei homogene Untergruppen gebildet werden können, die sich signifikant unterscheiden. Die Altersgruppen 1 und 5 werden demnach der Untergruppe 1 zugeordnet und die Altersgruppe 4 und 3 der Untergruppe 2 (siehe Tabelle 22). Die Altersgruppe 2 kann sowohl der ersten als auch der zweiten Untergruppe zugeordnet werden, weil sie keinen signifikanten Unterschied zu den angegebenen Altersgruppen aufweist. In Anbetracht der Mittelwerte erscheint es jedoch sinnvoll zu sein, die Altersgruppe 5 und 2 eher in eine neu generierte bzw. ergänzende Untergruppe einzuordnen, da die Mittelwertdifferenzen zur Altersgruppe 1 (5,18/7,13) sowie zu den Altersgruppen 3 und 4 sehr deutlich sind. Demzufolge wird die Altersgruppe 1 als Untergruppe 1 klassifiziert, die Altersgruppen 2 und 5 werden zur homogenen Untergruppe 2 und die Altersgruppen 3 und 4 bilden die neue Untergruppe 3. Sofern man das 5% Signifikanzniveau vernachlässigt und die offensichtlichen Mittelwertunterschiede weiterhin als zusätzliches Unterscheidungsmerkmal heranzieht, wird es ebenso deutlich, dass die Erstellung von drei Untergruppen auch für die anderen Wissensskalen plausibel erscheint, obgleich in der numerischen Wissensskala wiederum nur zwei Untergruppen gebildet werden. Zur Absicherung dieser Lösung kann zudem der weniger konservative Student-Newman-Keul-Test (S-N-K) verwendet werden, der bis auf die figurale Wissensskala ebenfalls eine differenziertere Aufteilung in drei Untergruppen berechnet (siehe Anhang B1.1). In der gesamten Wissensskala kann dementsprechend die aufgestellte Einteilung beibehalten werden. Das Gleiche gilt für die verbale Wissensskala. Innerhalb der numerischen Skala werden die Altersgruppen 1, 5, 2 und 4 der ersten Untergruppe zugeordnet, wobei die Altersgruppe 2 zusammen mit der Altersgruppe 4 auch der Untergruppe 2 zugerechnet werden können. Gleichwohl weisen die Altersgruppen 3 und 4 keine signifikanten

Unterschiede auf und können, unter Berücksichtigung der obigen Annahmen (Mittelwertunterschiede), in eine eigenständige Untergruppe 3 eingegliedert werden. Bezüglich der figuralen Skala werden die Altersgruppe 1 und 5 unter der ersten Untergruppe subsumiert, wobei die Altersgruppe 5, aufgrund des nicht signifikanten Mittelwertunterschiedes, zusammen mit der Altersgruppe 2 auch in die Untergruppe 2 eingeordnet werden kann. Die restlichen beiden Altersgruppen bilden wiederum die Untergruppe 3. Aufgrund der dargestellten Ergebnisse des Levene-Tests empfiehlt es sich zudem, den Tamhane-T2-Test für ungleiche Varianzen heranzuziehen. Erwartungsgemäß führt dieser Test zu vergleichbaren Resultaten und unterstützt somit die obige Verfahrensweise. Zusammenfassend lässt sich festhalten, dass sich die Altersgruppen 3 und 4 fast durchgehend signifikant von den anderen Altersgruppen unterscheiden. Die Altersgruppe 1 weist konstant die geringste Wissensausprägung in allen Skalen auf. Die Altersgruppen 2 und 5 schwanken innerhalb der vier Skalen zwischen der Untergruppe 1 und 2 und zeigen demnach eine mittlere Wissensausprägung im Vergleich zu den anderen Altersstufen auf.

*Tabelle 22: Homogene Untergruppen bezüglich der Wissensausprägung in den Altersgruppen-Scheffé Prozedur*

Skala	Bildungsgruppe	Untergruppe 1	Untergruppe 2	Untergruppe 3
Wissen Gesamt	Altersgruppe 1	50,85	-	-
	Altersgruppe 5	56,03	-	-
	Altersgruppe 2	58,03	58,03	-
	Altersgruppe 4	-	63,98	-
	Altersgruppe 3	-	65,49	-
Verbales Wissen	Altersgruppe 1	15,80	-	
	Altersgruppe 2	18,33	18,33	
	Altersgruppe 5	-	19,56	19,56
	Altersgruppe 3			21,54
	Altersgruppe 4			21,65
Numerisches Wissen	Altersgruppe 1	15,37	-	
	Altersgruppe 5	15,71	-	-
	Altersgruppe 2	16,35	16,35	-
	Altersgruppe 4	17,81	17,81	-
	Altersgruppe 3	-	18,65	-

Figurales Wissen	Altersgruppe 1	19,68	-	
	Altersgruppe 5	20,76	20,76	
	Altersgruppe 2	-	23,34	23,34
	Altersgruppe 4			24,51
	Altersgruppe 3			25,29

*Anmerkung:* Der Scheffe Test verwendet Stichprobengrößen des harmonischen Mittels = 86,997; Unterguppen für Alpha = .05; Altersgruppe 1= 16-Jährige; Altersgruppe 2= 17-18-Jährige; Altersgruppe 3= 19-20-Jährige; Altersgruppe 4= 21-25-Jährige; Altersgruppe 5= 26-45-Jährige; Die grau unterlegten Altersgruppen können aufgrund des Signifikanzniveaus sowohl der einen als auch der anderen Gruppe zugeordnet werden.

### ***Darstellung der Ergebnisse zur Forschungsfrage und Forschungsannahme- 1.2***

Im Rahmen der Forschungsfrage 1.2 soll der Einfluss der verschiedenen Muttersprachen auf die Wissensausprägung untersucht werden. Es wird erwartet, dass die Probanden, die Deutsch als Muttersprache haben (Muttersprachler-D) bessere Testergebnisse erzielen als die Probanden, die Deutsch als Fremdsprache gelernt haben (Muttersprachler-F).

In einem ersten Schritt wurden die Mittelwerte der Muttersprachler-D und die der Muttersprachler-F verglichen. Es ist ersichtlich, dass die Muttersprachler-D sowohl in der gesamten Wissensskala als auch in der verbalen, numerischen und figuralen Wissensskala bessere Testergebnisse erreichen (siehe Tabelle 23). In der gesamten Wissensskala können die Muttersprachler-D im Durchschnitt ca. 14 Aufgaben mehr lösen als die Muttersprachler-F. In den kodierten Wissensskalen bewegt sich die Differenz in einem Bereich zwischen vier und fünf Aufgaben. Der Mittelwert-Vergleich liefert schon hier einige Hinweise auf den vermuteten Zusammenhang.

*Tabelle 23: Mittelwerte und Standardabweichungen der Testergebnisse getrennt nach den unterschiedlichen Muttersprachlern*

Gruppen	Wissen		W. verbal		W.numerisch		W.figural	
	M	SD	M	SD	M	SD	M	SD
Muttersprachler-D (Deutsch) N= 663	63,87	16,45	21,14	5,86	18,05	5,41	24,68	6,74
Muttersprachler-F N=87	49,06	18,38	15,78	6,38	14,02	5,86	19,26	7,38

*Anmerkung:* M=Mittelwert; SD=Standardabweichung; W=Wissen; Muttersprachler-D=Muttersprache ist Deutsch; Muttersprachler-F=Deutsch als Fremdsprache; Aufgrund von fehlenden Angaben wurden insg. nur 750 Probanden mit einbezogen.

Für eine weiterführende Analyse wurde der T-Test bei unabhängigen Stichproben herangezogen, um genauere Aussagen hinsichtlich signifikanter Unterschiede treffen zu können. Für die gesamte Wissensskala sowie für die verbale, numerische und figurale Wissensskala resultierten in der angegebenen Reihenfolge Levene-Kennwerte von  $F = 1.21$  ( $p=.27$ ),  $F = .146$  ( $p=.70$ ),  $F = .286$  ( $p=.59$ ) und  $F = 1.97$  ( $p=.16$ ). Die Signifikanzkennwerte der gesamten Wissensskala und der kodierten Wissensskalen sind größer als 0.05, so dass Varianzgleichheit (Homogenität der Varianzen) angenommen werden kann. Die Normalverteilung der Rohwerte in den beiden Gruppen -als weitere Voraussetzung parametrischer Testverfahren- ist ebenfalls gegeben (siehe Anhang B1.2). Auf Basis der berechneten T-Werte lässt sich feststellen, dass signifikante Unterschiede zwischen den Muttersprachlern-D und den Muttersprachlern-F bezüglich der Wissensausprägung in allen Skalen bestehen (siehe Tabelle 24).

Tabelle 24: T-Test bei unabhängigen Stichproben

Skalen	Levene-Test		T-Test		
	F	Signifikanz	T	df	Signifikanz
Gesamte Wissensskala	1,211	,271	7,781	748	,000
Verbale Wissensskala	,146	,702	7,929	748	,000
Numerische Wissensskala	,286	,593	6,464	748	,000
Figurale Wissensskala	1,976	,160	6,966	748	,000

Anmerkung: F = F-Wert; T=T-Wert; df=Freiheitsgrade.

Aus der deskriptiven Statistik lässt sich entnehmen, dass die Stichprobengrößen der beiden Gruppen sehr unterschiedlich sind. Die Stichprobengröße der Muttersprachler-D umfasst 663 Probanden, die Gruppe der Muttersprachler-F hingegen nur 87 Probanden. In der Regel wird eine Stichprobengröße von  $N > 30$  erwartet, um die Varianzhomogenität und die Normalverteilung zuverlässig prüfen zu können (Bortz, 2005). Es kann jedoch davon ausgegangen werden, dass die Teststärke des T-Tests, aufgrund der ungleichen Stichprobengrößen, womöglich gemindert wird. In diesem Kontext erscheint es im Weiteren sinnvoll, alternativ den nichtparametrischen U-Test (Mann & Whitney) zu berücksichtigen, um den bereits dargestellten Befund abzusichern. Es ist ersichtlich, dass der U-Test die geschilderten Ergebnisse ausnahmslos bestätigt (siehe Tabelle 25).

Tabelle 25: U-Test nach Mann und Whitney

	Gesamte Wissensskala	Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala
Mann-Whitney-U	15553,500	14899,000	17056,500	17185,000
Wilcoxon-W	19381,500	18727,000	20884,500	21013,000
Z	-6,995	-7,347	-6,211	-6,140
Asymptotische Signifikanz (2-seitig)	,000	,000	,000	,000

### **Darstellung der Ergebnisse zur Forschungsfrage und Forschungsannahme- 1.3**

Bezüglich der Forschungshypothese 1.3 soll der geschlechtsspezifische Einfluss auf die Wissensausprägung analysiert werden. Es wird angenommen, dass zwischen den männlichen und weiblichen Probanden signifikante Unterschiede in der Wissensausprägung bestehen.

Zunächst werden die Mittelwerte verglichen, um erste Hinweise auf den vermuteten Zusammenhang zu generieren. Anhand Tabelle 26 ist es ersichtlich, dass männliche Probanden in der gesamten Wissensskala sowie in den kodierten Wissensskalen stets bessere Testergebnisse erzielen. Im Bezug auf die gesamte Wissensskala zeigen sich größere Differenzen. Im Detail können die männlichen Probanden im Durchschnitt 13 Aufgaben mehr lösen als die weiblichen Teilnehmerinnen. Hinsichtlich der Inhaltsfacetten streuen die Mittelwertunterschiede zwischen vier und fünf Testitems.

Tabelle 26: Mittelwerte und Standardabweichungen der Testergebnisse getrennt nach dem Geschlecht

Gruppen	Wissen		W. verbal		W.numerisch		W.figural	
	M	SD	M	SD	M	SD	M	SD
Männlich N= 166	72,37	18,23	23,86	6,26	20,71	6,14	27,80	7,37
Weiblich N=587	59,18	15,97	19,54	5,81	16,68	5,12	22,95	6,57

Anmerkung: M=Mittelwert; SD=Standardabweichung; W=Wissen.

Eine systematischere Überprüfung der Unterschiede gewährleistet wiederum der T-Test. Die Kennzahlen des Kolmogorov-Smirnov-Tests bestätigen die Normalverteilung der Rohwerte über die einzelnen Skalen. Lediglich in der figuralen Skala wird ein erhöhter Kennwert der Schiefe ausgewiesen (vgl. dazu Anhang B1.3).

Für die gesamte Wissensskala und die numerisch kodierte Skala wurde mittels des Levene-Tests Varianzhomogenität ermittelt. In der verbalen und figuralen Skala liegt Varianzheterogenität vor, so dass ein korrigierter Wert berechnet wurde.

Die Ergebnisse des T-Tests belegen den obigen Befund und dokumentieren, dass sich die Geschlechter signifikant voneinander unterscheiden (siehe Tabelle 27).

*Tabelle 27: T-Test bei unabhängigen Stichproben*

Skalen	Varianzhomogenität	Levene-Test		T-Test		
		F	Signifikanz	T	df	Signifikanz
Gesamte Wissensskala	Ja	3,090	,079	8,293	751	,000
	Nein			7,956	251,05	,000
Verbale Wissensskala	Ja	12,12	,001	8,536	751	,000
	Nein			7,717	233,75	,000
Numerische Wissensskala	Ja	2,866	,091	8,154	751	,000
	Nein			7,648	244,03	,000
Figurale Wissensskala	Ja	5,957	,015	9,092	751	,000
	Nein			8,447	241,27	,000

Aufgrund der unterschiedlichen Stichprobengrößen wird zur Absicherung der dargestellten Resultate die Kalkulation des nichtparametrischen U-Tests (Mann & Whitney) herangezogen, der die Ergebnisse wiederum ausnahmslos bestätigt (siehe Tabelle 28).

*Tabelle 28: U-Test nach Mann und Whitney*

	Gesamte Wissensskala	Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala
Mann-Whitney-U	29764,500	30078,500	29396,500	28128,500
Wilcoxon-W	202342,500	202656,500	201974,500	200706,500
Z	-7,671	-7,545	-7,817	-8,324
Asymptotische Signifikanz (2-seitig)	,000	,000	,000	,000

#### ***Darstellung der Ergebnisse zur Forschungsfrage und Forschungsannahme- 1.4***

Im Hinblick auf die Forschungsfrage 1.4 wird der Frage nachgegangen, ob der Bildungsgang bzw. das Bildungsniveau der Probanden die Wissensausprägung beeinflusst. Hierzu wurden zunächst die Mittelwerte der einzelnen Bildungsgänge miteinander verglichen. Aus der Tabelle 29 lässt sich entnehmen, dass der Wissensumfang in der gesamten Wissensskala mit dem Bildungsniveau ansteigt. Die Ausnahme bilden die

Probanden, die einen Hauptschulabschluss mit abgeschlossener Berufsausbildung haben. Obwohl das Bildungsniveau höher ist als das der Hauptschulabgänger, ist die Wissensausprägung geringer. Die Abiturienten (Bildungsgang, der als Abschluss das allgemeine Abitur vorsieht) und die Probanden, die bereits das Abitur erworben haben, zeigen dabei die höchsten Mittelwerte (für eine detailliertere Beschreibung der Bildungsgänge, siehe Tabelle 29). Die Hauptschulabgänger und die mit abgeschlossener Berufsausbildung zeigen die schwächsten Testergebnisse. Die Probanden mit mittlerer Reife weisen auch eine mittlere Ausprägung des Wissens auf. Im Unterschied zu den Hauptschulabgängern können die Probanden mit Abitur im Durchschnitt 25 Aufgaben mehr lösen. Bezüglich des verbalen, numerischen und figuralen Wissens können ähnliche Ergebnisse berichtet werden. Innerhalb der kodierten Wissensskalen sind die Unterschiede zwischen den Hauptschulabgängern und den anderen Bildungsgruppen ebenfalls deutlich ausgeprägt. Die besten Testergebnisse weisen, wie erwartet, die Abiturienten und die Probanden mit Abitur auf. Die Differenz der richtig gelösten Aufgaben zwischen den Hauptschülern und den Probanden mit allgemeiner Hochschulreife bewegt sich in einem Bereich von 8-10 Aufgaben. Hinsichtlich des numerischen und verbalen Wissens zeigen die Realschüler lediglich einen marginalen Wissensvorsprung auf.

*Tabelle 29: Mittelwerte und Standardabweichungen der Testergebnisse getrennt nach Bildungsgruppen*

Bildungsgruppen (1-6)	W.gesamt		W. verbal		W.numerisch		W.figural	
	M	SD	M	SD	M	SD	M	SD
Hauptschulabschluss B1; N= 12	46,41	14,41	15,91	4,50	12,58	5,50	17,91	5,33
Hauptschulabschluss mit Beruf B2; N= 14	38,00	14,33	13,14	5,69	11,78	4,53	13,07	5,25
Realschulabschluss B3; N= 201	49,87	12,92	16,58	4,97	13,79	4,18	19,48	5,74
Abschluss Fachabitur B4; N= 112	57,10	12,53	18,79	4,91	15,83	4,11	22,47	5,57
Abiturienten B5; N= 107	67,16	12,64	21,30	4,75	19,29	4,53	26,56	5,72
Abschluss Abitur B6; N= 298	72,34	15,80	24,09	5,64	20,67	5,25	27,57	6,28
Gesamt N= 744	62,17	17,39	20,53	6,19	17,59	5,63	24,04	7,04

Anmerkung: M=Mittelwert; SD=Standardabweichung; W=Wissen; Hauptschulabschluss= Bildungsgruppe 1=B1; Hauptschulabschluss mit Beruf= Bildungsgruppe 2=B2; Realschulabschluss= Bildungsgruppe 3=B3; Fachabitur= Bildungsgruppe 4=B4; Abiturienten= Bildungsgruppe 5=B5; Abitur= Bildungsgruppe 6=B6; Aufgrund von fehlenden Angaben wurden insg. nur 744 Probanden mit einbezogen.

Für eine differenzierte Untersuchung der signifikanten Wissensdifferenzen, bezogen auf die verschiedenen Bildungsgruppen, wird im Weiteren eine Varianzanalyse durchgeführt werden. Der Levene-Test auf Gleichheit der Varianzen erreicht nur in der verbalen und figuralen Wissensskala ein nicht signifikantes Ergebnis. Das bedeutet, dass hier Varianzhomogenität gegeben ist. In den anderen Skalen liegt Varianzheterogenität vor, so dass im Rahmen der Post-Hoc-Tests zusätzlich Testverfahren ohne die Annahme gleicher Varianzen zur Anwendung kommen. Auf der Basis des K-S-Tests und der eingehenden Überprüfung der Standardabweichungen, Standardfehler, Boxplots und Konfidenzintervalle ist es gleichwohl von Nutzen, beide Testverfahren in die Betrachtung mit einzubeziehen. Der Kolmogorov-Smirnov-Test zeigt, dass die Normalverteilung der Werte in den kodierten Wissensskalen und der gesamten Wissensskala für alle Bildungsgruppen gegeben ist ( $p > .05$ , siehe Anhang B1.4). Lediglich in der Bildungsgruppe 2 und 3 liegen die Werte der Schiefe über den vorgegebenen Richtwerten und sind damit leicht linksschief bzw. rechtssteil. Anhand der Ergebnisse der Varianzanalyse lässt sich feststellen, dass die Wissensunterschiede zwischen den Gruppen signifikant sind. Für die Gesamtskala ergab sich ein F-Wert von 78,32 der signifikant ist ( $p < .001$ ). Für die verbale, numerische und figurale Wissensskala wurden in der angegebenen Reihenfolge auch signifikante Kennwerte ( $p < .001$ ) von  $F = 61,08$  ( $df=5$ ),  $F = 64,44$  ( $df=5$ ) und  $F = 62,43$  ( $df=5$ ) ermittelt.

Es wurde bereits erwähnt, dass die Normalverteilung der Werte in den einzelnen Substichproben gegeben sein muss, um die Anwendung der parametrischen Varianzanalyse zu rechtfertigen. Obgleich die Werte gemäß dem Kolmogorov-Smirnov-Test normalverteilt sind, überschreiten einige Schiefe- und Kurtosiswerte in den einzelnen Bildungsgruppen den von Lienert und Raatz (1998) angegebenen Bereich der Normalverteilung (vgl. Anhang B1.4). Hinzu kommt, dass die Probandenanzahl der Hauptschüler und der Hauptschüler mit abgeschlossener Berufsausbildung kleiner als 30 und demnach vergleichsweise gering ist. Demzufolge muss davon ausgegangen werden, dass die Vorbedingung der Normalverteilung (s.o.) verletzt ist und die Teststärke der parametrischen Varianzanalyse beeinträchtigt wird. Im Hinblick auf die anschließende Interpretation empfiehlt sich daher ein nichtparametrisches Verfahren (H-Test von Kruskal & Wallis). Das Ergebnis des H-Tests bestätigt den oben dargelegten Befund, dass sich die verschiedenen Bildungsklassen bezüglich des Allgemeinwissens signifikant ( $p < .05$ ) unterscheiden.



Die Kennwerte für den H-Test zeigen für die gesamte Wissensskala einen Chi-Quadrat-Wert von  $\chi^2 = 269,025$  ( $p = .000$ ), für die verbale Wissensskala  $\chi^2 = 222,400$  ( $p = .000$ ), für die numerische Wissensskala  $\chi^2 = 233,401$  ( $p = .000$ ) und für die figurale Wissensskala  $\chi^2 = 223,679$  ( $p = .000$ ).

Eine weitergehende Untersuchung der postulierten Wissensunterschiede zwischen den Altersgruppen wird wiederum durch die Post-Hoc-Tests gewährleistet. Im Bezug auf die Frage, zwischen welchen Bildungsgruppen signifikante Differenzen in der Wissensausprägung bestehen, wurde zunächst die homogene Untergruppenbildung des S-N-K-Tests betrachtet, der auf der Grundlage der Student-Verteilung paarweise Vergleiche zwischen den Gruppen berechnet und damit weniger konservative und eindeutigere Ergebnisse liefert als der auf der F-Verteilung basierende Scheffé Test. Darauf aufbauend wurden dennoch die Mittelwert-Mehrfachvergleiche des Scheffé-Tests zur Absicherung herangezogen und anschließend der Tamhane-T2-Test (ohne Annahme gleicher Varianzen) berücksichtigt.

Aus der Tabelle 31 lässt sich entnehmen, dass – bis auf die numerische Wissensskala – prinzipiell vier Untergruppen über sämtliche Skalen gebildet werden. Die erste Untergruppe wird durchgehend von den Hauptschülern mit abgeschlossener Berufsausbildung repräsentiert. Ausgenommen in der numerischen Skala zeigen die Hauptschüler und die Realschüler keine signifikanten Unterschiede zur Bildungsgruppe 2 auf.<sup>13</sup> Die zweite Untergruppe umfasst überwiegend die Haupt- und Realschüler, abgesehen von der verbalen Wissensskala, in der die Hauptschüler keine signifikanten Differenzen zu den Fachabiturienten aufweisen sowie von der numerischen Wissensskala, in der die Realschüler mit den Fachabiturienten eine homogene Untergruppe bilden. Die dritte Untergruppe inkludiert in der gesamten und figural kodierten Wissensskala die Fachabiturienten. Die Ausnahme bildet die verbale und numerische Wissensskala. In der verbalen Skala wurde kein signifikanter Unterschied zwischen den Fachabiturienten und der Bildungsgruppe 5 (Abiturienten) ermittelt, obschon der Scheffé-Test eine signifikante Mittelwertdifferenz berechnet. In der numerischen Skala werden die Abiturienten und die Probanden mit Abitur zu einer Gruppe zusammengefasst. Die vierte Untergruppe wird vorrangig von den Probanden mit abgeschlossenem Abitur vertreten. In der gesamten und figuralen Wissensskala bilden sie mit den Abiturienten eine homogene Gruppe.

---

<sup>13</sup> Auf Basis des Scheffé-Tests unterscheiden sich die Haupt- und Realschüler ebenfalls nicht signifikant von der Bildungsgruppe 2 (siehe Anhang B1.4).

Im Rahmen des Scheffé-Tests bilden die beiden genannten Bildungsklassen in alle Skalen eine homogene Untergruppe. Darüber hinaus ist zu konstatieren, dass der Tamhane-T2-Test vergleichbare Ergebnisse liefert. Sofern man das 5-Prozent-Signifikanzniveau vernachlässigt und die offensichtlichen Mittelwertunterschiede als Differenzierungsmerkmal analysiert, ist es ersichtlich, dass grundsätzlich vier Untergruppen gebildet werden können. Demnach können die Hauptschüler mit beruflicher Ausbildung in die Untergruppe 1 eingegliedert werden. Die Haupt- und Realschüler bilden die homogene Untergruppe 2, wobei die Hauptschüler, aufgrund der deutlichen Mittelwertunterschiede bzw. ihres Wissensstandes, eher der Untergruppe 1 zugeordnet werden können. Eine eigenständige Gruppe 3 bilden die Fachabiturienten, die eindeutige Wissensunterschiede zu den Abiturienten aufweisen.<sup>14</sup> Unter der letzten Gruppe können die Abiturienten und Probanden mit Abitur subsumiert werden.

In Anlehnung an die Differenzierung im I-S-T 2000 R, der eine Aufgliederung in „Gymnasiasten und Nicht-Gymnasiasten“ vornimmt, können auf einer höheren Aggregationsebene zwei Gruppen gebildet werden, die sich signifikant voneinander unterscheiden (siehe Tabelle 30). Das heißt, die Gymnasiasten zeigen eine deutlich höhere Wissensausprägung als die Nicht-Gymnasiasten. Es wird angemerkt, dass die Vorbedingung (Normalverteilung der Rohwerte) zur Durchführung des parametrischen Testverfahrens (T-Test bei unabhängigen Stichproben) gegeben ist (siehe Anhang B1.5). Für die gesamte, verbale und numerische Wissensskala wurde mittels des Levene-Tests Varianzheterogenität ermittelt, so dass ein korrigierter Wert berechnet wurde.

Tabelle 30: T-Test bei unabhängigen Stichproben

Skalen	Varianzhomogenität	Levene-Test		T-Test		
		F	Signifikanz	T	df	Signifikanz
Gesamte Wissensskala	Ja	9,317	,002	15,88	742	,000
	Nein			16,91	502,99	,000
Verbale Wissensskala	Ja	8,555	,004	13,68	742	,000
	Nein			14,38	486,57	,000
Numerische Wissensskala	Ja	11,16	,001	14,47	742	,000
	Nein			15,60	518,13	,000
Figurale Wissensskala	Ja	2,732	,099	14,65	742	,000
	Nein			15,09	463,00	,000

<sup>14</sup> Das Ergebnis des Scheffé Tests bekräftigt dieses Vorgehen auch auf dem Signifikanzniveau von  $p < .05$ . Das heißt, die Fachabiturienten unterscheiden sich signifikant von den Abiturienten (siehe Anhang B1.4).

Abschließend lässt sich festhalten, dass mit steigendem Bildungsniveau auch die Wissensakkumulation zunimmt, das heißt das Wissen der höheren Bildungsklassen ist stärker ausgeprägt. Die Fachabiturienten weisen zwar teilweise (verbale und numerische Wissensskala) keine signifikanten Unterschiede zu den Probanden der Bildungsgänge 1 und 3 (Haupt- und Realschüler) auf, können aber aufgrund der durchgehend höheren Mittelwerte eher den Bildungsgruppen 5 und 6 zugeordnet werden. Aus diesem Grund ist die Einteilung in „Nicht-Gymnasiasten und Gymnasiasten“ naheliegend.

Auf der Basis der Ergebnisse bzw. der eindeutigen Differenzierung zwischen den verschiedenen Bildungsklassen kann darüber hinaus bereits ein erster und zugleich positiver Hinweis auf die Validität des Testverfahrens START-W dokumentiert werden.

*Tabelle 31: Homogene Untergruppen bezüglich der Wissensausprägung in den Bildungsgruppen-S-N-K Prozedur*

Skala	Bildungsgruppe	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4
<b>Wissen Gesamt</b>	Hauptschule mit Beruf	38,0000			
	Hauptschulabschluss		46,4167		
	Realschulabschluss		49,8706		
	Fachabitur			57,1071	
	Abiturienten				67,1682
	Abitur				72,3490
	Signifikanz	1,000	,320	1,000	,136
<b>Verbales Wissen</b>	Hauptschule mit Beruf	13,1429			
	Hauptschulabschluss		15,9167		
	Realschulabschluss		16,5871		
	Fachabitur		18,7946	18,7946	
	Abiturienten			21,3084	
	Abitur				24,0973
	Signifikanz	1,000	,065	,051	1,000
<b>Numerisches Wissen</b>	Hauptschule mit Beruf	11,7857			
	Hauptschulabschluss	12,5833			
	Realschulabschluss	13,7960	13,7960		
	Fachabitur		15,8393		
	Abiturienten			19,2991	
	Abitur			20,6779	
	Signifikanz	,193	,078	,235	

<b>Figurales Wissen</b>	Hauptschule mit Beruf	13,0714			
	Hauptschulabschluss		17,9167		
	Realschulabschluss		19,4876		
	Fachabitur			22,4732	
	Abiturienten				26,5607
	Abitur				27,5738
	Signifikanz	1,000	,282	1,000	,487

*Anmerkung:* Die Mittelwerte für die in homogenen Untergruppen befindlichen Gruppen werden angezeigt. Verwendet ein harmonisches Mittel für Stichprobengröße = 33,082. Untergruppen für Alpha = .05; Die grau unterlegten Bildungsgruppen können aufgrund des Signifikanzniveaus sowohl der einen als auch der anderen Gruppe zugeordnet werden.

### 10.3 Darstellung der Ergebnisse zu den Forschungsfragen- Teil 2

#### *Forschungsfrage- 2.1:*

#### *Übereinstimmungsvalidität*

Entsprechend den oben dargelegten Ausführungen werden in der Tabelle 32, die für die Forschungshypothesen 2.1.1 bis 2.1.4 relevanten Korrelationskoeffizienten und alle zusätzlichen Zusammenhänge zwischen Testwerten und Zeugnisnoten angegeben. Zur Überprüfung der Forschungsannahme wurden die Testergebnisse der drei Wissensskalen und der Gesamtscore mit dem Außenkriterium der Schulnoten (Deutsch-, Englisch-, Mathematik-, Kunst-, und Durchschnittsnote) korreliert. Zur Überprüfung der Forschungsannahme 2.3 und 2.4 werden die Zusammenhänge zwischen dem numerischen Skalenergebnis und der Schulnote im Fach Mathematik sowie zwischen der verbalen Wissensskala und der Deutsch-bzw. Englischnote gesondert berechnet. Aufgrund der Tatsache, dass die Schulnoten mindestens auf einem ordinalskalierten Niveau eingestuft werden können und entsprechend des Kolmogorov-Smirnov-Tests nicht normalverteilt sind (siehe Anhang B2.1), bietet sich für die Analyse der Zusammenhänge der Spearman'sche Korrelationskoeffizient an.

Im Vorfeld wird darauf hingewiesen, dass die zugrundeliegenden Berechnungen nicht die Gesamtstichprobe umfassen. Bedingt durch fehlende Teilnehmerangaben konnten für die Berechnungen mit der Deutschnote nur 715, mit der Englischnote 708, mit der Mathematiknote 716, mit der Kunstnote 635 und mit der Durchschnittsnote 676 Probanden berücksichtigt werden.

### ***Darstellung der Ergebnisse zu den Forschungsannahmen- 2.1.1 und 2.1.2***

Aus der Tabelle 32 lässt sich entnehmen, dass gemäß der ersten formulierten Forschungsannahme alle Korrelationskoeffizienten negativ sind.

Bei einer differenzierteren Analyse lässt sich feststellen, dass die Testergebnisse der kodierten Wissensskalen sowie der gesamten Wissensskala mit der Deutsch-, Englisch-, Mathematik- und Durchschnittsnote - mit Ausnahme der Kunstnote - signifikant negativ korrelieren (Forschungsannahme 2.1.2). Die Korrelationskoeffizienten liegen im Bereich von  $r = -.190$  und  $r = -.291$  ( $p < .01$ ), die nach den Bewertungsrichtlinien zwar gering, aber dennoch signifikant sind. Den stärksten Zusammenhang weisen das numerische Wissen und die Mathematiknote auf ( $r = -.291$ ,  $p < .01$ ) auf. Die niedrigste Korrelation besteht zwischen dem numerischen Wissen und der Kunstnote ( $r = -.008$ ,  $p > .05$ ), wobei die Korrelationen zwischen der Kunstnote und den aufgeführten Wissensskalen als sehr gering interpretiert werden können und zudem nicht signifikant ausfallen. Hinsichtlich des postulierten Zusammenhanges zwischen den Wissensskalen und der Durchschnittsnote kann die aufgestellte Forschungsannahme mit kleinen Einschränkungen aufrecht erhalten werden. Die Tabelle 32 lässt zudem erkennen, dass die Koeffizienten durchgehend größer als  $r = -.20$  ( $p < .01$ ) sind und damit in einem akzeptablen Wertebereich liegen. Die Berechnungen zeigen durchgehend hochsignifikante und negative Zusammenhänge zwischen den einzelnen Skalen und der Durchschnittsnote sowie zwischen dem Gesamtwert und der Durchschnittsnote, die als gering eingestuft werden. Die höchsten Korrelationen bestehen zwischen dem Gesamtergebnis und der Durchschnittsnote ( $r = -.233$ ,  $p < 0.01$ ) sowie zwischen dem numerischen Wissen und der Durchschnittsnote ( $r = -.18$ ,  $p < 0.01$ ).

### ***Darstellung der Ergebnisse zu den Forschungsannahmen 2.1.3, 2.1.4 und 2.1.5***

Die Forschungsannahme 2.1.3 kann hingegen ohne Restriktionen aufrecht erhalten werden. Die Korrelation zwischen der numerischen Wissensskala und der Mathematiknote weist den höchsten Koeffizienten von  $r = -.291$  ( $p < .01$ ) auf.

Bezüglich der Forschungsannahme 2.1.4 ist es ersichtlich, dass die verbale Wissensskala mit der Deutschnote ( $r = -.190$ ) und der Englischnote ( $r = -.247$ ) signifikant negativ korreliert ( $p < .01$ ).

Die Forschungsannahme 2.1.5 kann hingegen nicht bestätigt werden. Der Koeffizient weist nur einen Wert von  $r = .013$  auf.

Der Vollständigkeit halber werden die postulierten Zusammenhänge ebenfalls für einzelne Subklassen (hier die Bildungsklassen „Nicht-Gymnasiasten und Gymnasiasten“) überprüft. In diesem Zusammenhang ist festzuhalten, dass ähnliche Ergebnisse ermittelt werden. Auch die Ergebnisse der Substichproben zeigen zwar überwiegend geringe, aber signifikante Korrelationen auf. Aufgrund der unterschiedlichen Anzahl von Probanden innerhalb der Substichproben werden die Ergebnisse an dieser Stelle nicht näher beschrieben. Eine abschließende Interpretation der Befunde findet in Kapitel 11.3 statt.

*Tabelle 32: Korrelation zwischen Testwerten und Schulnoten*

Wissensskalen	Schulnoten				
	Deutsch N=715	Englisch N= 708	Mathematik N= 716	Kunst N= 635	Durchschnitt N= 676
Verbales Wissen	-,190**	-,247**	-,241**	,009	-,210**
Numerisches Wissen	-,173**	-,247**	-,291**	,008	-,233**
Figurales Wissen	-,124**	-,258**	-,229**	,013	-,210**
Gesamtscore	-,174**	-,272**	-,273**	,009	-,233**

Anmerkung: \*.Die Korrelation ist auf dem 0.05 Niveau signifikant (zweiseitig);

\*\*..Die Korrelation ist auf dem 0.01 Niveau signifikant (zweiseitig).

### ***Forschungsfrage 2.2:***

#### ***Prognostische Validität***

Im Rahmen der prognostischen Validität ist es von besonderem Interesse den Zusammenhang zwischen dem individuellen Wissensstand und den Kriterien des Ausbildungserfolges abzubilden. Aufgrund der Annahme, dass das erfasste Allgemeinwissen ein Indikator für zukünftiges Lernverhalten darstellt, wird im Weiteren die Korrelation der Wissensskalen mit dem Außenkriterium der theoretischen Durchschnittsnote des ersten Ausbildungsjahres analysiert.<sup>15</sup>

Die theoriebasierte Durchschnittsnote<sup>16</sup> im Bereich der Gesundheits- und Krankenpfleger/innen umfasst fünf Lernbereiche: Lernbereich I „Pflegerische Kernaufgaben“, Lernbereich II „Ausbildungs- und Berufssituation von Pflegenden“, Lernbereich III

<sup>15</sup> Der Kolmogorov-Smirnov-Test bestätigt die Normalverteilung der Rohwerte der Durchschnittsnote in der Gesamtstichprobe: K-S-Z = 1,007 Sig. = .262.

<sup>16</sup> Im dualen Ausbildungssystem wird zwischen Theorie und Praxis unterschieden.

„Zielgruppen, Institutionen und Rahmenbedingungen pflegerischer Arbeit“, Lernbereich IV „Gesundheits- und Krankenpflege bei bestimmten PatientInnengruppen“, Lernbereich V „Gesundheits- und Kinderkrankenpflege bei bestimmten PatientInnengruppen“. Den Lernbereichen werden wiederum vier Wissensfacetten zugrunde gelegt, die kreuztabellarisch in der „Krankenpflege-Ausbildungsprüfungsverordnung“ (KrPflAPrV) wiedergegeben werden. Hierbei wird die Vermittlung von (a) pflegerelevanten Kenntnissen der Gesundheits- und Krankenpflege, der Gesundheits- und Kinderkrankenpflege sowie der Pflege- und Gesundheitswissenschaften, (b) pflegerelevanten Kenntnissen der Naturwissenschaften und der Medizin, (c) pflegerelevanten Kenntnissen der Geistes- und Sozialwissenschaften und (d) pflegerelevanten Kenntnissen aus Recht, Politik und Wirtschaft avisiert (Oelke, 1998, Anpassung von Hundeborn, Kühn, 2003, Ausbildungsrichtlinie für die staatlich anerkannten Kranken- und Kinderkrankenpflegesschulen). Die übergreifende Zielsetzung der Ausbildungsverordnung wird in den Richtlinien wie folgt dargestellt:

„Die bereits betonte Wichtigkeit einer übergreifenden Qualifizierung legt eine Zielorientierung am Konzept der Schlüsselqualifikationen nahe. Die an diesem Konzept ausgerichtete, gleichzeitig aber auch auf das pflegeberufliche Handeln bezogene (...) Ausbildungsrichtlinie lässt sich wie folgt charakterisieren: Es sollen sowohl fachliche als auch sozial-kommunikative, methodische und personale Kompetenzen vermittelt werden“ (S. 9).

Die theoriebasierte Durchschnittsnote im Bereich der Ausbildung zum Operationstechnischen Assistenten basiert grundsätzlich auf denselben Wissensfacetten, wohingegen das übergeordnete Ausbildungsziel auf den genannten Gesundheitsberuf ausgerichtet ist. Das erklärte Lernziel ist die Befähigung, die Aufgaben im Operationsdienst und in den Funktionsbereichen Ambulanz, Endoskopie und Zentralsterilisation sicher, selbständig, flexibel, teamorientiert und fachgerecht zu bewältigen (Vivantes- Institut für berufliche Bildung, 2010). Die Überprüfung der vermittelten Lehrinhalte wurde anhand von sieben Einzelklausuren vorgenommen. Aus den ermittelten Einzelklausuren wurde sodann die Durchschnittsnote berechnet.

### ***Darstellung der Ergebnisse zur Forschungsannahme 2.2***

Für die Auszubildenden am Vivantes Institut für Gesundheits- und Krankenpflege bzw. für die Operationstechnischen Assistenten wurden geringe bis mittlere und

hochsignifikante Korrelationen ermittelt. In der Substichprobe der Gesundheits- und Krankenpfleger bewegt sich der Koeffizient zwischen  $r = -.20$  und  $r = -.34$  (Korrelationskoeffizient nach Spearman), der bis auf die figurale Skala hochsignifikant ausfällt (siehe Tabelle 33). Unter der geprüften Voraussetzung, dass die Variablen intervallskaliert und normalverteilt sind (siehe Anhang B2.2), erhöht sich nach der Produkt-Moment-Korrelation von Pearson die Stärke des Zusammenhangs deutlich auf einen Kennwert von bis zu  $r = -.363$ . Bezogen auf die Gruppe der Operationstechnischen Assistenten werden mittlere Korrelationskoeffizienten berechnet. Der stärkste Zusammenhang zeigt sich mit der verbalen Skala in Höhe von  $r = -.583$  ( $p < .05$ ). Bezüglich der nichtsignifikanten Zusammenhänge lässt sich ergänzen, dass das Signifikanzniveau von  $p < .05$  nur marginal überschritten wird ( $p = .07$ - numerische Skala). Bei Verwendung des Korrelationskoeffizienten von Pearson steigt der Zusammenhang auf  $r = -.609$  ( $p < .01$ ). Im Hinblick auf die Gesamtstichprobe wurden durchgehend hochsignifikante Zusammenhänge berechnet, die im Bereich zwischen  $r = -.287$  und  $r = -.408$  liegen (Korrelationskoeffizient nach Spearman). Abschließend lässt sich feststellen, dass der stärkste Zusammenhang zwischen der numerischen Wissensskala und der Durchschnittsnote besteht, gefolgt von dem Gesamtscore Wissensskala und der verbalen Wissensskala. Bedingt durch die heterogene Zusammensetzung der Durchschnittsnote und die vielfältigen Einflussfaktoren können die hochsignifikanten Korrelationskoeffizienten als durchaus bedeutsam und zufriedenstellend bewertet werden.

*Tabelle 33: Korrelation zwischen Testwerten und der Durchschnittsnote des ersten Ausbildungsjahres*

Wissensskala	Durchschnittsnote des ersten Ausbildungsjahres					
	Gesundheits- u. Krankenpflege (N= 87)		Operationstechnische Assistenten (N= 18)		Gesamtstichprobe (N= 105)	
	Spearman	Pearson	Spearman	Pearson	Spearman	Pearson
Verbales Wissen	-,344**	-,363**	-,583*	-,609**	-,342**	-,345**
Numerisches Wissen	-,332**	-,359**	-,436	-,516*	-,408**	-,398**
Figurales Wissen	-,200*	-,252*	-,299	-,557*	-,287**	-,317**
Gesamtscore	-,305**	-,353**	-,517*	-,603**	-,368**	-,384**

Anmerkung: \*.Die Korrelation ist auf dem 0.05 Niveau signifikant (zweiseitig); \*\*.Die Korrelation ist auf dem 0.01 Niveau signifikant (zweiseitig).



## 10.4 Darstellung der Ergebnisse zu den Forschungsfragen- Teil 3

### *Forschungsfrage 3.1:*

#### *Konvergente Validität*

Im Rahmen der Konstruktvalidierung soll anhand von theoretischen und psychologischen Analysen geklärt werden, welche Merkmale der Test abbildet. Demzufolge werden zunächst Hinweise zur konvergenten Validität des START-W betrachtet.

#### *Darstellung der Ergebnisse zur Forschungsannahme 3.1*

Zur Überprüfung der intendierten Validitätsbestimmung wurde die Kurzform des Bochumer Wissenstests (kurz BOWIT; Hossiep & Schulte, 2008) herangezogen.<sup>17</sup> Aus der nachstehenden Tabelle 34 lässt sich entnehmen, dass die Korrelationskoeffizienten zwischen den einzelnen Wissensskalen des START-W und dem Gesamtrohwert in einem Wertebereich von  $r = .554$  und  $r = .678$  (Korrelation nach Spearman) liegen und hochsignifikant ausfallen. Die Produkt-Moment-Korrelation nach Pearson weist Werte bis zu  $r = .704$  auf. Bereits hier ist festzuhalten, dass die hohen Zusammenhänge einen eindeutigen Hinweis auf die konvergente Validität des START-W liefern und der Test die Erfassung des angeeigneten Wissens bzw. der kristallisierten Intelligenz ermöglicht.

*Tabelle 34: Korrelation der Wissensskalen des START-W mit dem Gesamtrohwert der BOWIT-Kurzform*

START-W	BOWIT-Kurzform	
Wissensskala Rohwerte	Gesamtrohwert	
	Spearman	Pearson
Verbales Wissen	,614**	,632**
Numerisches Wissen	,664**	,704**
Figurales Wissen	,554**	,604**
Gesamtscore	,678**	,704**

Anmerkung: \*\*. Die Korrelation ist auf dem Niveau von .01 (2-seitig) signifikant.

<sup>17</sup> Der Kolmogorov-Smirnov-Test bestätigt die Normalverteilung der Rohwerte in der gesamten Wissensskala des BOWIT: K-S-Z = 1,267 Sig. = .081.

### **Forschungsfrage 3.2 und 3.3**

#### **Darstellung der Ergebnisse zur Forschungsannahme 3.2**

Die Konzeption des START-W-Wissenstests basiert auf der von Cattell (1971a, 1987) aufgestellten Theorie der fluiden und kristallisierten Intelligenz bzw. dem hierarchischen Strukturmodell, deren Faktoren mit höherem Allgemeinheitsgrad durch sukzessive Faktorenanalysen erschlossen werden. Bei der Entwicklung des Instruments wurde darauf geachtet, unterschiedliche Wissensbereiche zu berücksichtigen, die durch die Inhaltsfacetten (verbales, figurales und numerisches Wissen) kodiert werden. Demnach wird der Test zur Erfassung des verbal, numerisch und figural kodierten Wissens eingesetzt. Ausgehend von der Annahme, dass das erworbene Wissen mit dem Konstrukt der kristallisierten Intelligenz (gc) verknüpft ist und der vorliegende Wissenstest verschiedene Aspekte des erworbenen Wissens in unserer Kultur erfasst, wird zudem davon ausgegangen, dass der START-W ebenfalls die Abbildung von gc ermöglicht. Es soll der Frage nachgegangen werden, ob die theoretisch fundierten Annahmen durch die Empirie bestätigt werden können. Die mit LISREL (Version 8.30; Jöreskog & Sörbom, 1993) durchgeführte konfirmatorische Faktorenanalyse testet danach die theoretisch angenommene Struktur und prüft, ob die empirisch erhobenen Daten dem theoretischen Modell entsprechen.

Die nachfolgende Tabelle 35 zeigt die zugrunde gelegte Testkonstruktion und gibt damit die Verteilung der Wissensfragen auf die Themenbereiche und die Inhaltskodierung (verbal, numerisch, figural) wieder.

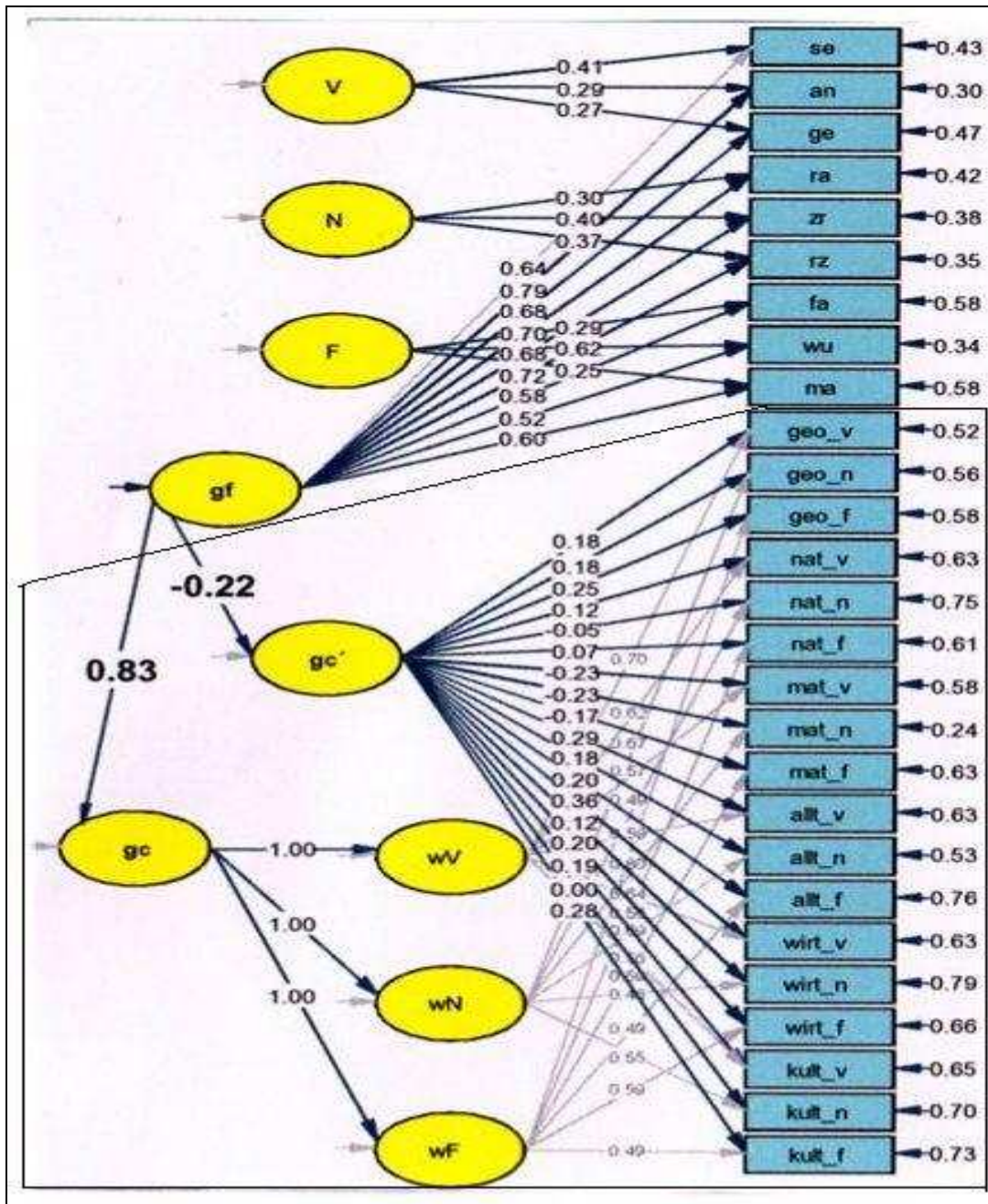
*Tabelle 35: Zuordnung der Wissensfragen zu Themenbereichen und Kodierungsarten*

Zuordnung der Wissensfragen zu Themenbereichen und Kodierungsarten (Kreuzklassifikationsmatrix * Themen * Kodierung)								
Themenbereich	Kodierungsform						I-S-T 2000 R	START-W
	verbal		numerisch		figural		Summe Testitems	
Geographie/Geschichte	5	7 (+2)	4	10 (+6)	5	5	14	22
Wirtschaft	4	8 (+4)	5	2 (-3)	5	6 (+1)	14	16
Kunst/Kultur	5	8 (+3)	5	9 (+4)	4	12 (+8)	14	29
Mathematik	5	7 (+2)	4	12 (+8)	5	8 (+3)	14	27
Naturwissenschaft	4	7 (+3)	5	4 (-1)	5	7 (+2)	14	18
Alltagswissen	5	7 (+2)	5	7 (+2)	4	8 (+4)	14	22
<b>Summe START-W</b>	44 Testitems		44 Testitems		46 Testitems		134	
<b>Summe I-S-T 2000 R</b>	28 Testitems		28 Testitems		28 Testitems		84	

Zur Überprüfung der intendierten Struktur wurde auf der Grundlage einer Korrelationsmatrix eine exploratorische Faktorenanalyse durchgeführt. Zur Extraktion der den beobachteten Variablen zugrunde liegenden Faktoren wurde die Hauptachsenanalyse verwendet. Anschließend wurde die schiefwinklige (*oblique*) Rotationsmethode eingesetzt, um eine *oblimine* Lösung zu erhalten. Im Rahmen der darauffolgenden konfirmatorischen Faktorenanalyse mit Hilfe von LISREL (Version 8.30), ergab sich auf der Basis der Strukturannahmen das nachstehende Modell (siehe Abbildung 20). Das Modell liefert einen ersten Hinweis auf die postulierten Zusammenhänge und Facettenstruktur<sup>18</sup>.

---

<sup>18</sup> Es wird angemerkt, dass das entwickelte Strukturgleichungsmodell sowohl für die Forschungsfrage 3.1 und die Forschungsfrage 3.2 konzipiert wurde. Hinsichtlich der erstgenannten Frage liegt das Hauptaugenmerk auf der Struktur des Wissenstest, die durch das innere schwarze Kästchen gekennzeichnet ist (Abbildung 20).



Anmerkung: SE= Satzergänzung; AN= Analogien; GE= Gemeinsamkeiten; RA= Rechenaufgaben; ZR= Zahlenreihen; RZ= Rechenzeichen; FA= Figurenauswahl; WA= Würfelaufgaben; MA= Matrizen; Geo.= Geografie/Geschichte; Nat.= Naturwissenschaften; Mat.= Mathematik; Allt.= Alltagswissen; Wirt.= Wirtschaft; Kult.= Kultur; V= Verbale Kodierung; N= Numerische Kodierung; F= Figurale Kodierung.

Abbildung 20: Strukturgleichungsmodell I-S-T 2000 R-Grundmodul  
& START-W Wissenstest

Im Hinblick auf die Schätzung der Ladungen der Variablen auf den *Faktoren erster Ordnung* (Themenbereiche) wurde ein Intervall zwischen .20 und .80 festgelegt.

Aus dem Pfaddiagramm lässt sich entnehmen, dass die Ladungen in einem Bereich von .24 und .79 rangieren. Dementsprechend erreicht die erste Ebene der Faktorenstruktur eine akzeptable Anpassung an die Daten. Ausgehend von den Wissensbereichen, die wiederum durch drei Facetten kodiert wurden, ergeben sich die *Faktoren zweiter Ordnung*, die als verbales Wissen (wV), numerisches Wissen (wN) und figurales Wissen (wF) interpretiert werden.<sup>19</sup> Bezüglich des Faktors wV liegen die Faktorladungen zwischen .56 und .70 für den Faktor wN innerhalb von .46 bis .80 und für wF zwischen .49 und .65 (siehe Anhang B3.2). Aufgrund der positiven Interkorrelationen zwischen den Kodierungsarten ist die Berechnung eines Gesamtwertes  $g_c$  gerechtfertigt. Bedingt durch die Komplexität des Modells, den multivariaten Einflussfaktoren und der damit zusammenhängenden Berechnung von Schätzwerten wurden die in diesem Modell ermittelten Ladungen auf  $g_c$  mit 1 fixiert. Es mussten die Pfade fixiert werden, um die Varianz von  $g_c$  bestimmen zu können. Andernfalls wäre die Anzahl der Freiheitsgrade derart gestiegen, dass die einzelnen Parameter nicht mehr geschätzt werden können. Der Faktor  $g_c$  repräsentiert somit die Wissensvarianzen auf einem höheren Generalitätsniveau. Ergänzend wurde in dem oben aufgeführten Modell ein ('nested') Faktor  $g_c'$  eingeführt, der die residualen Varianzen des Wissens darstellen soll. Dieser Faktor soll die spezifischen Wissensvarianzen repräsentieren, die sich auf ein geringeres Generalitätsniveau beziehen und die nicht von den inhaltsbezogenen Wissensfaktoren erklärt werden können. Die ausgewiesenen Faktorladungen liegen hingegen in einem niedrigen und teils negativen Bereich (-0.17 bis 0.36). Die geringen Ladungskoeffizienten können jedoch teilweise mit der Aufsplittung der Fehlervarianz auf  $g_c'$  und die kodierten Inhaltsskalen begründet werden.

Zur Beurteilung der Güte des Modells werden verschiedene *Fit-Indizes* herangezogen. Ein einfacher "Goodness-of-Fit"-Index ist der Chi-Quadrat-Test Wert ( $\chi^2$ -Test), der mittels der Freiheitsgrade des Strukturgleichungsmodells auf Signifikanz geprüft wird. Die Verwendung des  $\chi^2$ -Tests ist jedoch begrenzt, da er sehr sensitiv gegenüber der Stichprobengröße ist und mit steigender Stichprobengröße schon marginale Modellabweichungen signifikant werden. Daher sind alternative Fit-Indizes entwickelt worden. Ein gebräuchlicher Fit-Index ist demnach der *Comparative-Fit-Index* (CFI), der genau dann vorzuziehen ist, wenn die Abhängigkeit von der Stichprobengröße verringert werden soll und stabilere Schätzungen bzw. Ergebnisse angestrebt werden (Bentler, 1990).

---

<sup>19</sup> Der Faktor  $g_c'$  wird erst bei der Forschungsfrage 3.3 berücksichtigt und inhaltlich erklärt.

Den Erwartungen entsprechend fällt der  $\chi^2$ -Test (295, die Freiheitsgrade wurden für die 15 fixierten Ladungen korrigiert) mit einem Wert von 1117.07 signifikant aus ( $p < .001$ ). Der Comparative-Fit-Index weist einen Wert von .97 auf, der damit in einem sehr guten Bereich liegt (Pfeifer & Schmidt, 1987). Der *Root Mean Square Error of Approximation* (RMSEA) wird mit einem Wert von .066 angegeben und die *Standardized Root Mean Square Residual* (SRMR) mit einem Wert von 0.092. In dem separat betrachteten Strukturgleichungsmodell, in dem gc als Faktor dritter Ordnung aus den Faktoren verbales Wissen (wV), numerisches Wissen (wN) und figürliches Wissen (wF) ermittelt wurde, konnten folgende Kennzahlen berechnet werden:  $\chi^2(313) = 1605.92$  ( $p < .001$ ), CFI = .96, RMSEA = .073 und SRMR = .099.

### ***Darstellung der Ergebnisse zur Forschungsannahme 3.3***

Die Forschungsfrage 3.3 thematisiert einerseits den allgemeinen Zusammenhang zwischen fluider Intelligenz (gf) und kristallisierter Intelligenz (gc), der im Kapitel 6 ausführlich beschrieben wurde. Andererseits soll untersucht werden, in welchem Ausmaß der gc-Indikator deklaratives Wissen mit gf korreliert<sup>20</sup> und inwieweit diese Korrelation vom Aggregations- bzw. Generalitätsniveau der Wissensvariablen abhängt.

Anhand des aufgezeigten Pfaddiagramms ist es ersichtlich, dass die fluide Intelligenz (gf) in substantiellem Umfang mit der kristallisierten Intelligenz (gc) zusammenhängt. Die Korrelation erreicht einen Wert von .83, der signifikant ausfällt. Darüber hinaus ist festzustellen, dass der Faktor gf mehr Varianz an gc als an gc' erklärt (siehe Abbildung 20). In einem  $\chi^2$ -Differenzen-Test war der Unterschied zwischen den Pfadkoeffizienten signifikant ( $\chi^2_{\text{diff}}(1) = 135.45$ ;  $p < .001$ ). Darüber hinaus war der Pfad von gf nach gc' negativ. Offenbar hängen die nicht an die Inhaltsfaktoren gebundenen residualen Wissensvarianzen nur geringfügig und negativ mit gf zusammen.

---

<sup>20</sup> Der Faktor gf konnte aus den Daten mithilfe des elaborierten Grundmoduls des I-S-T 2000 R extrahiert werden.

## 11 Diskussion der Ergebnisse

### 11.1 Item- und Reliabilitätsanalysen

Mit der Binnenanalyse wurde geprüft, inwieweit die einzelnen Aufgaben der Testabsicht bzw. der Testkonstruktion gerecht werden. Hierzu wurden die Gütekriterien der (a) Schwierigkeit (Schwierigkeitsindex oder Index der kategorialen Häufigkeit), der (b) Trennschärfe und der (c) Homogenität herangezogen. Das vorrangige Ziel der Itemanalyse besteht darin, Aufgaben zu identifizieren, die den angegebenen Kriterien entsprechen, Items zu selektieren oder gegebenenfalls zu überarbeiten. Darüber hinaus gilt es zu berücksichtigen, dass die gesamte Wissensskala auf der Basis einer explorativen Herangehensweise in drei Wissensskalen unterteilt wurde. Die Skalenbildung wurde dabei auf Basis der verbalen, numerischen und figuralen Kodierungsarten vorgenommen. Es sei jedoch darauf hingewiesen, dass die gesamte Wissensskala auch anhand von Wissensgebieten aufgliedert werden kann. Die Skalenbildung würde demnach die Wissensgebiete Geografie/Geschichte, Naturwissenschaften, Kultur, Mathematik, Wirtschaft, Gesundheit und Alltagswissen umfassen (Amthauer et al., 2001). Die Entscheidung für die Skalenbildung auf Grundlage der Kodierungsarten wurde damit gerechtfertigt, dass die drei Wissensskalen aus Gründen der Übersicht und der inhaltlichen Relevanz eine höhere Relevanz für die weiteren zwei Forschungsannahmen besitzen. Die Berücksichtigung der verbalen, numerischen und figuralen Wissensskala bietet dabei die Möglichkeit, grundlegendere Aussagen über das Wissen treffen zu können.

Zu Beginn wird der Schwierigkeitsindex als wichtiges Maß zur Beurteilung der psychometrischen Eigenschaft eines Items betrachtet und interpretiert. Insbesondere bei der Konstruktion von Niveau- oder Leistungstests wird die Bedeutung des Kennwertes betont, um sicherzustellen, dass der Test zwischen Probanden mit unterschiedlichen (Wissens-) Merkmalsausprägungen differenziert. Hierzu ist eine größtmögliche Streuung der Schwierigkeit notwendig. Demzufolge werden Indizes zwischen 0.2 und 0.8 bevorzugt, „so dass Aufgaben von minimaler wie solcher von maximaler Schwierigkeit im Test vorkommen“ (Lienert & Ratz, 1998, S. 58). Zudem sollte bei der Entwicklung von Niveau- und Leistungstests darauf geachtet werden, die Aufgaben nach ihrer Schwierigkeit zu ordnen, um im Idealfall anhand der fortlaufenden Nummer der Aufgabe, welche der Proband als letztes bearbeitet hat, das Leistungsniveau des Probanden bestimmen zu können. Außerdem kann durch diese Vorgehensweise (Erstellung von

so genannten „Eisbrecherfragen“) einer vorzeitigen Demotivation oder gar Frustration von vermeintlich leistungsschwachen Testteilnehmern entgegengewirkt werden. Auf der Basis der vorliegenden Stichprobe konnten für die verbale Wissensskala Schwierigkeitsindizes berechnet werden, die in einem Wertebereich zwischen .17 und .89 liegen. Für die numerische Skala ergaben sich Werte zwischen .10 und .82 und für die figurale Inhaltsklasse zwischen .16 und .96. Der Großteil der Indizes für die einzelnen Aufgaben bewegt sich demnach in dem vorgegebenen Intervall. Obgleich eine abschließende Bewertung ohne die Betrachtung der Trennschärfe und der Homogenität nicht gegeben werden kann, ist festzustellen, dass die für die Zielsetzung des Verfahrens angemessenen Aufgaben im präferierten Wertebereich liegen.

In einem zweiten Schritt wird die Trennschärfe analysiert. Die Trennschärfe einer Aufgabe ist im Rahmen der Itemanalyse ein sehr aussagekräftiger Kennwert, da sie ein klares Differenzierungsmerkmal repräsentiert und damit für die Konstruktion eines Leistungs- oder Niveautests von entscheidender Bedeutung ist. Items mit geringer Trennschärfe sind demnach redundant und können ausselektiert werden. Die in Kapitel 10.1 dargestellten Tabellen zeigen den bereits selektierten Itempool über die drei verschiedenen Inhaltsklassen. Die Selektion einzelner Items erfolgte nach dem Kriterium, dass Items mit niedriger Trennschärfe ( $< .10$ ) eliminiert wurden. In der verbalen Skala wurden demzufolge vier Items (Item 04, 56, 85, 108) ausgesondert und in der numerischen Skala wurden sieben Items (Item 02, 17, 33, 94, 98, 107, 129) aus der Grundgesamtheit ausgeschlossen. Die figurale Skala blieb bestehen. Der Mittelwert der Trennschärfe-Koeffizienten in den einzelnen Kodierungsformen rangiert zwischen .26 und .30. Daraus lässt sich schließen, dass der Großteil der Koeffizienten im niedrigen bis mittleren Wertebereich liegt. Auch unter Verwendung strengerer Kriterien wird angemerkt, dass eine zusätzliche Eliminierung von Items keine weitere Erhöhung der Reliabilitätsschätzung nach sich zieht. Auf der Basis der berücksichtigten Beurteilungsbestimmungen liegen die Kennwerte damit in einem befriedigenden Bereich. Die durchgeführten Reliabilitätsanalysen der drei inhaltsbezogenen Klassen und der Gesamtskala weisen vergleichbare Ergebnisse auf.

Die Reliabilitätsschätzung der gesamten Wissensskala zeigt eine interne Konsistenz von  $\alpha = .923$ . Der Wert liegt in einem hohen und damit zufriedenstellenden Bereich. Innerhalb der drei inhaltlichen Kodierungsformen weist die figurale Skala den höchsten Wert der internen Konsistenz auf mit Cronbachs  $\alpha = .832$ . Die Reliabilitätsschätzungen der numerischen und verbalen Wissensskalen lassen nur unwesentliche Unterschiede



zwischen  $\alpha = .778$  und  $\alpha = .797$  erkennen, wobei die Homogenität in der verbalen Wissensskala stärker ausgeprägt ist. Zusammenfassend lässt sich festhalten, dass die Homogenitätskennwerte der verbalen und numerischen Skalen in einem niedrigen, aber dennoch annehmbaren Bereich liegen. Die figurale Wissensskala übertrifft indessen den vorgegebenen Wert von  $\alpha = .80$  und stellt somit eine passable und befriedigende Größe dar.

Die Itemanalysen der Vorversion des START-W offenbaren jedoch auch, dass die Kennwerte durchaus noch weiter optimiert werden können. Es muss im Rahmen dieser Erwägung berücksichtigt werden, dass die Ermittlung von optimalen itemspezifischen Kennwerten aus unterschiedlichen Gründen komplex ist und in der Praxis bzw. in der empirischen Forschung nicht zwangsläufig gegeben sein kann oder auch gefordert wird. Eine mögliche Erklärung für die zum Teil niedrigen Reliabilitätsschätzungen, Trennschärfen und Schwierigkeitsindizes kann in erster Line darin gesehen werden, dass das Wissen, entsprechend der vorangegangenen Ausführungen, ein umfangreiches und vielschichtiges Konstrukt darstellt. Selbst durch die Aufgliederung der Gesamtskala in die drei Wissensskalen oder auf Basis der sieben Themenbereiche (geringeres Generalitätsniveau) wird nur ein marginaler Einfluss auf die Verringerung der inhaltlichen Heterogenität des Konstrukts erreicht. Demnach beeinträchtigen die ausgewählten Aufgaben aus verschiedenen Wissensbereichen innerhalb einer Skala die interne Konsistenz des Tests. Darüber hinaus wird die Berechnung von einwandfreien psychometrischen Kennwerten auch durch die Wechselbeziehungen zwischen Homogenitätsindex, Trennschärfe und Schwierigkeit erschwert.

Einerseits wird der Homogenitätsindex durch die unterschiedlichen Schwierigkeitsgrade der Aufgaben bestimmt. Das heißt, eine perfekte Interkorrelation wäre nur denkbar, sofern die Items den gleichen Schwierigkeitsgrad aufweisen. Lienert und Raatz (1998) führen in diesem Zusammenhang an, dass „je größer die Schwierigkeitsunterschiede sind, um so geringer werden die Interkorrelationskoeffizienten von  $n$  Aufgaben sein“ (S. 108). Die Autoren relativieren demnach die Realisierung einer vollkommenen Homogenität, sobald die Konzeption eines Leistungstests, wie der START-W, auf der Grundlage von differierenden Aufgabenschwierigkeiten (als Differenzierungsmerkmal) beruht. Andererseits hängt der Trennschärfekoeffizient wiederum „von der Anzahl der möglichen interindividuellen Differenzierungen, die eine Aufgabe treffen kann[ab]“ (Lienert & Raatz, 1998, S. 107) und diese Anzahl ist durch den Index der Schwierigkeit funktional abhängig. Weiterhin steht die Trennschärfe in einem interdependenten Ver-

hältnis zur Homogenität des Tests, da beide Kriterien einen Bezug zum Gesamttest herstellen und demnach hohe Interkorrelationen hohe Trennschärfen implizieren.

Ein weiteres Argument zur Erklärung von qualitativ geringeren Kennwerten basiert auf der heterogenen Leistungsfähigkeit einzelner Teilstichproben der Gesamtstichprobe (z.B. Bildungsgruppen) (Bühner, 2006), die in Kapitel 10.2 ausführlich dargestellt worden sind. In diesem Kontext wurde bspw. ein Homogenitätsindex der verbalen Skala für die Teilstichprobe der Hauptschüler von  $\alpha = .580$  ermittelt und für die Abiturienten ein Wert von  $\alpha = .766$ . Wissensunterschiede in Abhängigkeit des männlichen und weiblichen Geschlechts könnten ebenfalls eine Begründung für die niedrigen Kennwerte sein (Neidhardt-Wilberg, 2005). Hinsichtlich des weiblichen Geschlechts liegt die innere Konsistenz der numerischen Skala bei  $\alpha = .73$  und beim männlichen Geschlecht bei  $\alpha = .81$ . Eine weitere Ursache kann auch in dem deutlichen Wissensunterschied zwischen den Muttersprachlern-D und den Muttersprachlern-F gesehen werden. Für die Substichprobe der Muttersprachler-D ergab sich ein Cronbachs Alpha von  $.77$  (verbale Skala) und für die Muttersprachler-F wurde ein Wert von  $\alpha = .80$  berechnet. Als Abschlussbemerkung sei nochmals darauf hingewiesen, dass die Bewertung der Reliabilität und der Trennschärfen auf allgemeinen Beurteilungsrichtlinien beruhen. Boyle, Stan-kow und Cattell (1995) kritisieren dagegen die Prämisse vieler Forscher, psychometrische Kennwerte zu maximieren. Die Autoren plädieren in diesem Zusammenhang für einen Homogenitätsindex, der im Bereich zwischen  $\alpha = .30$  und  $\alpha = .70$  liegt. Sie begründen ihren Standpunkt damit, dass mit höherer Homogenität auch gleichzeitig die Redundanz von Informationen der einzelnen Aufgaben zunimmt und aufgrund des starken Zusammenhangs der Nutzen zur Untersuchung eines Persönlichkeitsmerkmals verloren geht bzw. eine Differenzierung anhand des Merkmals kaum mehr möglich ist. Lienert und Raatz (1998) bestärken diese Sichtweise: „Tests, die einen umschriebenen Wissensstand prüfen sollen, sind grundsätzlich schwer auf Reliabilität zu prüfen. Innere Konsistenz kann, braucht aber nicht gegeben zu sein“ (S. 241). Dementsprechend würden die ermittelten Kennwerte im optimalen Bereich liegen.

Dafür sprechen auch die berechnete Halbierungsreliabilität (Split-Half-Reliabilität) und die Retest-Reliabilität. Im Zuge der Split-Half-Reliabilität wurden der Gesamttest sowie die Inhaltsskalen mithilfe der odd-even-Methode halbiert und die beiden Testhälften jeweils miteinander korreliert. Anschließend wurden die ermittelten Koeffizienten nach der Spearman-Brown korrigiert.

Innerhalb der Skalen liegen die Korrelationskoeffizienten in einem Bereich zwischen  $r = .785$  und  $r = .823$ . Für den Gesamttest ergibt sich ein Zusammenhang von  $r = .929$ .

Im Rahmen der Retest-Reliabilität sind, für die einzelnen Inhaltsskalen Korrelationskoeffizienten in einem Bereich von  $r = .821$  bis  $r = .835$  ( $p < .01$ ) festzuhalten. Der Gesamttest zeigt den höchsten Wert mit  $r = .88$  auf, der wie die anderen hochsignifikant ausfällt. In Anbetracht des langen bzw. sechsmonatigen Zeitabschnitts, in dem der Test erneut vorgelegt wurde, können die berichteten Befunde als beachtlich und zufriedenstellend bezeichnet werden.

## **11.2 Diskussion der Befunde zu den Forschungsfragen- Teil 1**

### ***Diskussion der Befunde zur Forschungsfrage und Forschungsannahmen- 1.1***

Als zentrale theoretische Grundlage für die entwicklungspsychologische Forschung im Lebenslauf wird häufig die Theorie der fluiden und kristallisierten Intelligenz von Cattell (1971a, 1987) herangezogen. Entsprechend der Ausführung im theoretischen Teil kristallisiert sich die fluide Intelligenz im Laufe der ontogenetischen Entwicklung zu spezifischen Fähigkeiten heraus. Die kristallisierte Intelligenz spiegelt demnach „die vom Individuum rezipierten und organisierten Wissensinhalte und -systeme wider, die für jene Gesellschaft und Kultur charakteristisch sind, in der es lebt“ (Kruse & Rudinger, 1997, S. 50) und die im Laufe des Lebens akkumuliert worden sind. Im Bezug auf die Intelligenzentwicklung im Lebenslauf zeigen zahlreiche Studien, dass die fluide Intelligenz ab dem Erwachsenenalter abgebaut wird, wohingegen die kristallisierte Intelligenz bis ins späte Erwachsenenalter zunimmt und ein höheres Maß an Stabilität aufweist (vgl. z.B. Lindenberger, 2000; Baltes, 1990). Diesbezüglich wird in der wissenschaftlichen Literatur die stärkere Wissensausprägung im Alter häufig mit der zunehmenden Lebenserfahrung und der kontinuierlichen Ausbildung von bereichsspezifischen Wissenssystemen begründet. Auch Baltes (1990) weist in seinen empirischen Beiträgen auf den Einfluss von Erfahrungen hin, die im Erwachsenenalter weitere Differenzierungen und Transformationen der Wissenssysteme indizieren und sich im Aufbau von Wissensstrukturen manifestieren. In diesem Zusammenhang wird auch die Bedeutung des lebenslangen Lernens hervorgehoben, das durch ständig neue Problemlöseprozesse in Gang gehalten wird. Hierbei spielen auch Trainings- und Übungeffekte eine entscheidende Rolle. Eine wichtige Voraussetzung bleibt jedoch, dass im jugendlichen Alter

bereits differenzierte faktische und prozedurale Wissensstrukturen aufgebaut wurden, um die zukünftige Wissensaufnahme zu erleichtern. Dabei kommt es nach Lindenberger (2000) aber auch auf bestimmte Sozialisationsinflüsse an (z.B. ein entsprechendes Bildungssystem), die den Erwerb von deklarativem und prozeduralem Wissen entweder fördern oder behindern. In diesem Kontext wird in der vorliegenden empirischen Untersuchung ebenfalls das Ziel verfolgt, einen Zusammenhang zwischen dem Alter und dem Wissen (kristallisierte Intelligenz) aufzudecken. Die daraus resultierende Forschungsannahme lautet: Der Wissensumfang steigt in Abhängigkeit des Alters. Ein geeignetes Messinstrument zur Überprüfung dieser Fragestellung ist der Wissenstest START-W, der die Möglichkeit besitzt, die kristallisierte Intelligenz abzubilden (siehe Kapitel 10.4.1 und 11.4).

Im Hinblick auf die dargestellten Ergebnisse ist bereits bei der Betrachtung der Mittelwerte aufgefallen, dass die Altersgruppe 3 (19-20-Jährige) und die Altersgruppe 4 (21-25-Jährige) im Vergleich zu den Altersgruppen 2 (17-18-Jährige) und 5 (26-46-Jährige) und vor allem gegenüber der Altersgruppe 1 (16-Jährige) in allen Wissensskalen wesentlich höhere Werte erreichen. Die zusammengefassten Altersgruppen zeigten dabei nur unerhebliche Differenzen auf, so dass eine dreiteilige homogene Untergruppenverteilung plausibel erscheint. Aufgrund von ersten Überlegungen kann dieser Effekt damit begründet werden, dass die 19- und 25-Jährigen (a) den Hauptteil der Altersgruppen bilden, (b) die beiden Altersstufen dicht beieinander liegen, (c) die Mehrheit in diesen Altersgruppen das allgemeine Abitur anstrebt oder besitzt und (d) die Probanden gerade erst die Schule verlassen haben. Die anschließende Varianzanalyse konnte die augenscheinliche Diskrepanz zwischen den genannten Altersgruppen bestätigen. Der F-Test sowie der H-Test (Kruskal & Wallis) machen deutlich, dass zwischen den Gruppen signifikante Unterschiede bestehen. Anhand der Post-Hoc-Tests (Scheffé-Test, S-N-K-Test und Tamhane T2) konnten, unter Berücksichtigung der offensichtlichen Mittelwertunterschiede, grundsätzlich drei Untergruppen gebildet werden, die sich im Wissen unterscheiden. Resümierend lässt sich konstatieren, dass sich die Altersgruppen 3 und 4 fast durchgehend signifikant von den anderen Altersgruppen differenzieren. Die Altersgruppe 1 weist stets den geringsten Wissensstand in allen Skalen auf. Die Altersgruppen 2 und 5 schwanken innerhalb der vier Skalen zwischen der Untergruppe 1 und 2 und zeigen demnach eine mittlere Wissensausprägung, im Vergleich zu den anderen Altersgruppen, auf.

Auf der Basis der angeführten Ergebnisse kann die aufgestellte Forschungsannahme nicht oder nur mit erheblichen Einschränkungen bestätigt werden. Hierfür sind diverse Gründe verantwortlich, die im Folgenden diskutiert werden. Die hinter der Konzeption des START-W liegenden Zielsetzungen beziehen sich auf die zweckbestimmte Entwicklung eines Tests für potentielle Berufseinsteiger, die sich an der Schnittstelle zwischen Schule, Ausbildung und Beruf befinden und die im Alter von 18 bis 28 Jahren sind. Außerdem wurde bei der Konstruktion darauf geachtet, dass der Inhalt weitgehend an dem vermittelten Lehrstoff der 11.-13. Klasse orientiert ist. Demzufolge kann davon ausgegangen werden, dass die Altersgruppe 1 (16-Jährige) und die Altersgruppe 2, auch im Sinne der formulierten Forschungsannahme, geringere Wissensausprägungen aufweisen. Entgegen der Annahme folgt die Altersgruppe 5 (26-46-Jährige). Eine Erklärung liegt womöglich darin, dass das in der Schule aufgenommene und eher allgemein gehaltene Wissen im Alter zunehmend verblasst und tendenziell spezifischeres Wissen aufgebaut wird. Die Altersgruppen 3 und 4 entsprechen genau der Zielgruppe und verfügen demnach über ein ausgeprägtes Wissensspektrum. Zudem besitzt ein Großteil der Probanden innerhalb der Altersgruppen das allgemeine Abitur, welches einen aussagekräftigen Indikator für die erhöhte Wissensakkumulation darstellt (siehe Tabelle 36), siehe auch Forschungsannahme 1.4). Bekräftigende Ergebnisse liefert auch die (Multiple) Regressionsanalyse (siehe Anhang C1.1). Es zeigt sich, dass das Alter nur einen marginalen Einfluss auf die Wissensausprägung besitzt, wohingegen das Bildungsniveau einen beachtlichen Teil zur Erklärung der Varianz beiträgt.

*Tabelle 36: Kreuztabelle mit Altersgruppen und Bildungsniveau*

Altersgruppen	Bildungsniveau	Mittelwert Gesamte Wissensskala	Häufigkeit N / %
Altersgruppe 1	Nicht-Gymnasiasten	44,33	24 / 68,57
	Gymnasiasten	65,09	11 / 31,42
Altersgruppe 2	Nicht-Gymnasiasten	50,47	61 / 43,26
	Gymnasiasten	64,01	80 / 56,73
Altersgruppe 3	Nicht-Gymnasiasten	52,20	55 / 17,29
	Gymnasiasten	68,27	263 / 82,7
Altersgruppe 4	Nicht-Gymnasiasten	51,02	50 / 28,57
	Gymnasiasten	69,45	125 / 71,4

Altersgruppe 5	Nicht-Gymnasiasten	41,83	37 / 49,33
	Gymnasiasten	70,18	38 / 50,06

Anmerkung: Altersgruppe 1= 16-Jährige; Altersgruppe 2= 17-18-Jährige; Altersgruppe 3= 19-20-Jährige; Altersgruppe 4= 21-25-Jährige; Altersgruppe 5= 26-45-Jährige.

### ***Diskussion der Befunde zur Forschungsfrage und Forschungsannahmen- 1.2***

Im Rahmen der Forschungsfrage 1.2 wurde der Einfluss der verschiedenen Muttersprachen auf die Wissensausprägung analysiert. Es wurde angenommen, dass die Muttersprachler-D (Deutsch als Muttersprache) bessere Testergebnisse erreichen als die Muttersprachler-F (Deutsch als Fremdsprache). Im Hinblick auf den Vergleich der Mittelwerte wurden die vermuteten Differenzen bereits angedeutet. Die Muttersprachler-D wiesen sowohl in der gesamten Wissensskala als auch in der verbalen, numerischen und figuralen Inhaltsskala eindeutig bessere Testergebnisse auf. Im Durchschnitt konnten die Muttersprachler-D in der gesamten Wissensskala ca. 14 Aufgaben mehr lösen und in den kodierten Inhaltsklassen annähernd fünf Items. Anhand des T-Tests für unabhängige Stichproben konnte der Unterschied ebenfalls bestätigt werden.<sup>21</sup> Die Ergebnisse zeigen signifikante Unterschiede zwischen den Gruppen auf. In diesem Kontext wird jedoch darauf hingewiesen, dass die Ergebnisse nur eingeschränkt interpretiert werden können. Die Problematik liegt darin begründet, dass die Größe der Substichproben erheblich voneinander abweichen. Demnach können die gefundenen Zusammenhänge lediglich als Hinweis verstanden werden. Für eine eindeutigere Aussage müsste der Stichprobenumfang der Muttersprachler-F deutlich erhöht werden. Trotz der Einschränkungen kann die Forschungsannahme in diesem Fall aufrecht erhalten werden.

Eine mögliche Ursache für die Diskrepanzen kann darin gesehen werden, dass der START-W ein verbal orientiertes Testverfahren ist und zudem in deutscher Schriftsprache vorgelegt wurde. Auch in der wissenschaftlichen Literatur wird diese Thematik aufgegriffen und gleichwohl bekräftigt, dass Personen, die Deutsch als Muttersprache gelernt haben gegenüber jenen, die Deutsch als Fremdsprache gelernt haben, eindeutige Vorteile im Hinblick auf die Bearbeitung von verbal ausgelegten Testmaterial besitzen (vgl. z.B. PISA 2006, 2007; Schwippert, Peek & Bolle, 2004; Tiedemann & Billmann-

<sup>21</sup> Aufgrund der unausgewogenen Verteilung der Substichproben wurde der U-Test (Mann & Whitney) herangezogen, der den signifikanten Unterschied zwischen den Muttersprachlern-D und den Muttersprachlern-F bestätigt und demnach eine zusätzliche Absicherung der Interpretation möglich macht.

Mahecha, 2004). Demnach hatten die Muttersprachler-F womöglich schon im Vorfeld eine schlechte Ausgangsvoraussetzung. Ein unzureichender Wortschatz und ein mangelndes Sinnverständnis könnten hier als Erklärungen für die geringeren Leistungen der Muttersprachler angeführt werden. In diesem Zusammenhang kann die Rolle der Sprache nicht hoch genug bewertet werden, da sie von großer Bedeutung für die Wissensaufnahme ist. Das bedeutet gleichzeitig, dass die sprachlich-symbolische Kodierung und Sortierung von Lehrinhalten zwischen den Muttersprachlern-D und Muttersprachlern-F variieren kann und den Aufbau von Wissen beeinflusst (Oerter, 1988). Insbesondere die Lehrinhalte, die in der Schule vermittelt werden, basieren auf sprachlicher Kommunikation, sodass die Muttersprachler-F womöglich Probleme mit dem schulischen Unterrichtsstoff haben. Darüber hinaus bezieht sich die inhaltliche Item-Auswahl teilweise auf den deutschen Kulturkreis (z.B. Fragen zur deutschen Geschichte). Unter der Annahme, dass das Wissen der Muttersprachler-F durch ein anderes soziokulturelles Umfeld geprägt worden ist und auf unterschiedlichen Normen- bzw. Wertesystemen beruht, kann man davon ausgehen, dass die Muttersprachler-F benachteiligt sind (Yönter, 1997). Das Wissen wird einerseits durch den Sozialisationsvorgang bestimmt und andererseits durch die kulturelle Bestimmtheit der Objekte selbst. Die Bedeutung und die Definition von Objekten sowie der damit zusammenhängende Wissensinhalt wird immer auch durch den Kontext der Gesellschaft und Kultur bestimmt (a.a.O.). Der Inhalt des Wissens und die Wissensausprägung können demnach auch zwischen den Muttersprachlern-D und den Muttersprachlern-F variieren. Vor einem ähnlichen Hintergrund wird auch die aktuelle bildungspolitische Debatte im Rahmen der PISA-Studien geführt. In diesem Zusammenhang geht es vor allen Dingen auch um die Thematisierung der multikulturellen und multilingualen deutschen Gesellschaft, die einen soziodemografischen Wandel durchläuft. Insbesondere innerhalb der Schulen und anderer Ausbildungseinrichtungen wird daher die Frage nach der Migration und dem Kompetenzerwerb als eminent wichtig erachtet.

Baumert und Schümer (2002) problematisieren in ihrem Forschungsbeitrag zu den „familiären Lebensverhältnissen, Bildungsbeteiligung und Kompetenzerwerb im nationalen Vergleich“ genau diesen Kernbereich und stellen heraus, dass erhebliche Leistungsunterschiede in den Testwerten der PISA-Studie 2000 zwischen Jugendlichen mit

Migrationshintergrund (Deutsch als Zweitsprache) und ohne Migrationshintergrund (Deutsch als Erstsprache) bestehen.<sup>22</sup>

In ihren Untersuchungen gehen sie von der Leitannahme aus, dass primär die Lese- und Sprachkompetenzen Schlüsselqualifikationen darstellen, die einen entscheidenden Einfluss auf das allgemeine Lernverhalten von Jugendlichen besitzen und diese nicht durch erhöhte Leistungsfähigkeiten in anderen Bereichen ausgeglichen werden können. Die Annahme wird insoweit relativiert, als dass sich Defizite in der Lese- und Sprachbeherrschung nicht auf weniger sprachlastige Fächer, wie z.B. Mathematik auswirken. Die Ergebnisse zeigen jedoch, dass in sämtlichen Fächern -entgegen der Annahme- die Kompetenzunterschiede deutlich sind, obgleich die Variabilität der Testwertdifferenzen von Land zu Land verschieden ist. Dementsprechend kann festgehalten werden, dass die sprachliche Disposition einen wesentlichen Einfluss auf die Lernfähigkeit und damit verbunden auf die Wissensakkumulation aufweist. In die gleiche Richtung weisen die Befunde von Baumert und Schümer (2002), die einen eindeutigen Zusammenhang zwischen den Lesekompetenzen und der Herausbildung naturwissenschaftlicher Kenntnisse, respektive mathematischer Kompetenzen ermittelten (S. 220). Darüber hinaus ist es in der vorliegenden Stichprobe ersichtlich, dass die Probanden mit Migrationshintergrund zu einem großen Anteil in die unteren Bildungsklassen eingestuft werden („Nicht-Gymnasiasten“). Demzufolge ergibt sich eine weitere Ursache für die geringen Leistungsindikatoren der Muttersprachler-F. Die Gymnasiasten erzielten in der Regel signifikant bessere Ergebnisse als die Nicht-Gymnasiasten (siehe Forschungsfrage 1.4). Im Zuge der Grundschul-Lese-Untersuchung (IGLU) machen auch die Befunde von Orth (2006) deutlich, dass Schüler mit Migrationshintergrund bereits in den Grundschuljahren erhebliche Leistungsdefizite zeigen. Vergleichbare Resultate liefern auch die Forschungsbeiträge von Schwippert, Peek und Bolle (2004) oder Tiedemann und Billmann-Mahecha (2004), die die Lese- und Sprachkompetenz „als ein nicht unwesentliches Bestimmungsstück“ (S. 275) definieren. Gleichzeitig wird jedoch darauf hingewiesen, dass die Leistungsunterschiede nicht nur auf der Unzulänglichkeit der Kenntnis der deutschen Sprache basieren, sondern auf kumulative Beeinträchtigungen zurückzuführen sind, wie zum Beispiel ein „geringerer sozioökonomischer Status, eine höhere

---

<sup>22</sup> Es wird darauf hingewiesen, dass der Begriff Migrationshintergrund in diesem Zusammenhang nicht-zwangsläufig bedeutet, dass ein oder zwei Elternteile im Ausland geboren wurden, sondern hinter der Größe „steht eine Vielzahl von Sprachen und damit eine breite kulturelle Vielfalt“ (Schwippert, Peek & Bolle, 2004, S. 253).



Betroffenheit von Arbeits- und Perspektivlosigkeit, eine geringere Bildungsorientierung und damit verbunden eine verminderte kognitive Förderung der Kinder“ (S. 277).

Für einen Ausgleich der Leistungsdifferenzen müssten genau an diesen sowie weiteren soziokulturellen und bildungspolitischen Schnittstellen gesellschaftliche und institutionelle Regelungen getroffen werden.

Abschließend lässt sich festhalten, dass die Muttersprachler-F zwar Wissenslücken aufweisen, die aber vor dem Hintergrund des jeweiligen kulturellen Umfeldes relativiert werden sollten.

### ***Diskussion der Befunde zur Forschungsfrage und Forschungsannahmen- 1.3***

Die übergeordnete Forschungsfrage bezieht sich auf den Einfluss des Geschlechts auf die Wissensausprägung. Auf der Basis der dargestellten Befunde (siehe Kapitel 6.3) wurde die Annahme aufgestellt, dass männliche Probanden bessere Testergebnisse erzielen.

Im Hinblick auf die Mittelwerte lässt sich die Annahme aufrecht erhalten. Das gilt sowohl für die gesamte Wissensskala als auch für die kodierten Inhaltsklassen. Die Analyse der Mittelwertdifferenzen (T-Test und U-Test) ergab, dass sich die Testleistungen der männlichen und weiblichen Probanden signifikant unterscheiden. Die ermittelten Befunde bestätigen somit die Forschungsannahme und dokumentieren Differenzen in Abhängigkeit des Geschlechts. In die gleiche Richtung weisen auch die ausgewählten Forschungsbeiträge im nächsten Abschnitt. Es wird darauf hingewiesen, dass die Resultate nur begrenzte Schlussfolgerungen zulassen, da die Stichprobengrößen deutlich voneinander abweichen.

Im Folgenden werden mögliche Erklärungsansätze für die eindeutigen Differenzen zwischen männlichen und weiblichen Probanden überblicksartig dargestellt. Als Grundlage für die weitere Diskussion wird auf die in Kapitel 6.2 und Kapitel 6.3 präsentierte Befundlage verwiesen. Ausgangspunkt der Kontroverse ist das in der Intelligenzforschung verbreitete Investmentmodell von Cattell (1987), welches die Interaktion von fluider Intelligenz und kristallisierter Intelligenz thematisiert. Eine umfassende Darstellung der Forschungsbefunde dieser Diskussion würde an dieser Stelle jedoch zu weit führen. Daher wird zunächst auf die einschlägigen Forschungsbeiträge von Levy (1974), Hyde und Lynn (1988), Halpern (1992), Lynn (1994), Allik, Must und Lynn (1999), Ackerman et al. (2001), Lynn und Irwing (2002), Nyborg (2003), Lynn, Irwing und Cammock (2002)

sowie Lynn, Wilberg und Margraf-Stiksrud (2004) verwiesen und ausgewählte Ansätze werden skizziert. Im Rahmen der Forschungen werden zur Erklärung der dokumentierten Wissensdifferenzen in Abhängigkeit des Geschlechts soziokulturelle sowie biologische Argumente zur Kontroverse mit angeführt. Eine herausragende Rolle nimmt hierbei die Analyse neuropsychologischer Effekte ein, die auch *Lateralisierung* genannt wird. In diesem Kontext wird die ungleiche Verteilung von Funktionen auf die linke, respektive rechte Gehirnhälfte thematisiert. Obwohl die Auslegungen nicht unbestritten sind, wird zugleich eingeräumt, dass durch die Hemisphärenasymmetrie die Geschlechtsunterschiede partiell begründet werden können. Vor allem für den Generalbefund, dass Männer im räumlich-visuellen, quantitativ-mathematischen und analytischen Bereich bevorteilt sind und Frauen dagegen über einen Vorsprung bei den verbalen Fähigkeiten verfügen, werden häufig anlagebedingte Argumente angeführt (vgl. Bischof-Köhler, 2006, S. 214ff). Darauf aufbauend wird zudem der Einfluss von Sexualhormonen und damit das Zusammenspiel organisierender pränataler und postnataler Hormonwirkungen auf die kognitive Leistung diskutiert. Diese Theorie vertritt der Mediziner Nyborg (1990) und bezeichnet sie als „General Trait Covariance-Androgen/Estrogen Balance Model“. Weitere Ursachen für die geschlechtstypischen Differenzen werden in der Persönlichkeit gesehen. Demgemäß werden spezifische Persönlichkeitsmerkmale für die Geschlechtsunterschiede verantwortlich gemacht (vgl. Rolfhus & Ackerman, 1999; Ackerman et al., 2001). Hierzu wird die Vermutung aufgestellt, dass differierende Merkmale wesensgemäße Handlungen unterstützen oder nicht und somit den Erwerb von Wissen steuern. Vor diesem Hintergrund zeigen beispielsweise die Befunde von Ackerman et al. (2001) auf, dass einerseits Eigenschaften wie Extraversion und Neurotizismus, die eher dem weiblichen Geschlecht zugeordnet werden, geringere Korrelationen mit dem Wissenserwerb bzw. mit Leistungsmerkmalen aufweisen. Andererseits kann nach Neidhardt-Wilberg (2005) das Individualitätsmerkmal der Wettbewerbsorientierung, welches vornehmlich dem männlichen Geschlecht zugerechnet wird, die Wissensaneignung begünstigen. Darüber hinaus werden auch komplementäre Denkstile als Begründung herangezogen (Bischof-Köhler, 2006). Auf der Grundlage der Forschungsergebnisse von Schwank (1990) werden die Denkstile in *prädiktives* und *funktionales* Denken unterteilt, wobei ersteres tendenziell dem weiblichen Geschlecht zugeschrieben wird. „Prädiktive Denker“ beschreiben eine Struktur in Form von *Prädikaten* und Relationen. Sie stellen begriffliche Beziehungen her, nehmen Abstraktionen vor, achten auf logische Verknüpfungen und ordnen in Klassen. Sie haben das ganze Bild im

Blickfeld (...) Ihr Denken hat einen eher *statischen* Charakter“ (Bischof-Köhler, S. 233). Der „funktionale Denker“ hingegen sieht „Strukturen in erster Linie unter dem Aspekt von *Wirkungsbeziehungen*. Sie beschreiben sie in Form von Operationen (...). Ihr Denken ist *prozessorientiert*“ (a.a.O., S. 234). Für Schwank (1990) erscheint es plausibel, dass gerade physikalisch-technische Probleme besser mit dem funktionalen Denkstil bearbeitet werden können. Des Weiteren hebt Bischof-Köhler (2006) vor allem auch motivationale Aspekte im Bezug auf die unterschiedlichen kognitiven Leistungen und den Wissenserwerb hervor. Daran anschließend ergibt sich auch ein anderer Zugang, der durch eine soziokulturelle Herangehensweise charakterisiert wird. Nach Neidhardt-Wilberg (2005) kann das Allgemeinwissen als das „von mehreren Kohorten geteiltes Wissen definiert werden und umfasst sowohl Kenntnisse, die durch Erziehung und Bildung intentional vermittelt, als auch diejenigen, die durch Massenmedien verbreitet und inzidentell aufgenommen werden“ (S. 145). Darauf bezugnehmend können auch durch die Sozialisation bedingte Interessensunterschiede und Rollenverständnisse für die Wissensdifferenzen verantwortlich sein. In ihrer Konklusion vertritt Neidhardt-Wilberg (2005) die Auffassung, dass Fachbereiche wie Sport, Naturwissenschaften -bis auf Biologie-, Politik und Geschichte entschieden männliche Wissensdomänen repräsentieren. Demgegenüber stehen beispielsweise die Fächer Kunst, Literatur oder Musik, die für das weibliche Geschlecht, auch hinsichtlich der obigen Ausführungen, leichter zugänglich zu sein scheinen. In Anbetracht der im START-W verwendeten Themenbereiche und Schwerpunkte ist es daher nachvollziehbar, dass die männlichen Probanden bessere Testleistungen zeigen.

Auf der Basis der dargestellten Befundlage bleibt abschließend festzuhalten, dass verschiedene Faktoren, ob biologischen Ursprungs oder soziokultureller Assimilation, einen Einfluss auf die Geschlechtsunterschiede haben. Unbestritten bleibt, dass Differenzen bestehen und diese multifaktoriell determiniert sind. Folglich besteht noch erheblicher Forschungsbedarf, welche Bedingungen bzw. Ursachen die Geschlechtsunterschiede bedingen.

#### ***Diskussion der Befunde zur Forschungsfrage und Forschungsannahmen- 1.4***

Anhand der geschilderten Ergebnisse ist festzustellen, dass die Forschungsannahme aufrecht erhalten werden kann. Es konnte ein Einfluss der Bildungsklassen auf die Wissensausprägung nachgewiesen werden. Einerseits differieren die Testergebnisse in

Abhängigkeit des erreichten Bildungsniveaus, andererseits konnte durch den Vergleich der Mittelwerte auch schon ein erster Hinweis auf den vermuteten Zusammenhang gefunden werden, dass der Wissensumfang in Abhängigkeit des Bildungsniveaus ansteigt. Die Bildungsgruppen, die als Abschluss das allgemeine Abitur anstreben bzw. bereits besitzen, weisen in allen Wissensskalen bessere Testergebnisse auf als diejenigen mit Haupt- oder Realschulabschluss. Eine Ausnahme bilden die Fachabiturienten, deren Testergebnisse teils im unteren und teils im oberen Bereich liegen. Die Ergebnisse der Varianzanalyse (F-Test) als auch des H-Tests (Kruskal & Wallis) bestätigen im Weiteren die Annahme, dass signifikante Wissensunterschiede zwischen den Bildungsgruppen bestehen. Entsprechend den obigen Ausführungen konnten mithilfe der Post-Hoc-Tests, unter den dort genannten Bedingungen, grundsätzlich vier Untergruppen gebildet werden, die sich im Wissensumfang unterscheiden. Die Untergruppe 1 umfasst demnach die Hauptschüler mit Beruf, die Untergruppe 2 schließt die Bildungsgänge der Haupt- und Realschüler ein, die Fachabiturienten werden in eine eigenständige Untergruppe 3 eingeordnet und die gymnasiale Oberstufe sowie die Probanden mit Abitur werden der Untergruppe 4 zugeteilt. Es wird jedoch angemerkt, dass einige Bildungsgruppen nicht immer auf Basis des erforderlichen Signifikanzniveaus der einen oder anderen Untergruppe zugeordnet wurden. Die Entscheidung für die dargelegte Untergruppenbildung wurde damit gerechtfertigt, dass zwischen den Fachabiturienten und den Bildungsgruppen der Haupt- und Realschüler durchweg eindeutige Mittelwertunterschiede bestehen.

Trotz dieser marginalen Einschränkung kann die Forschungsannahme dahingehend aufrecht erhalten werden, dass zwischen den Nicht-Gymnasiasten und Gymnasiasten stets bedeutende Wissensunterschiede bestehen. Demzufolge kann abschließend davon ausgegangen werden, dass der Wissensumfang in Abhängigkeit des Bildungsniveaus ansteigt. Folglich kann auch ein positiver Zusammenhang zwischen dem Bildungsniveau und der Wissensausprägung abgeleitet werden.

Eine Ursache für die bestehenden Wissensunterschiede liegt annahmegemäß in dem unterschiedlichen Hintergrund der Bildungsgruppen begründet. Infolgedessen liegen die signifikanten Wissensunterschiede womöglich in der inhaltlichen Ausgestaltung des Unterrichtsstoffes begründet. Es kann davon ausgegangen werden, dass der zugrunde gelegte Lehrinhalt, entsprechend der gewählten schulischen Laufbahn, differiert. Im Rahmen der Bildungsstandards der Kultusministerkonferenz werden die Schulformen

der Hauptschule und der Realschule als berufsqualifizierende Bildungsgänge bezeichnet, die tendenziell praxisbezogen, methoden- und handlungsorientiert aufgebaut sind. Das heißt, in den Bildungsgängen werden verstärkt fachspezifische Themen vermittelt, die auf den zukünftigen Berufseinstieg abzielen. Daraus ergibt sich auch der Schwerpunkt des schulischen Lernstoffs. Neben der Vermittlung von klassischen Fächern wie Deutsch oder Mathematik wird insbesondere die Problematik der Berufsorientierung thematisiert. Als zentrales Motiv der Lehrplangestaltung wird die Berufswahlvorbereitung in die unterschiedlichen Unterrichtseinheiten eingeflochten und durch spezielle Fächerkombinationen gefördert. Dem gegenüber steht die gymnasiale Oberstufe, in der vor allem allgemeinbildende Inhalte durch angemessene didaktische und pädagogische Konzepte in den Vordergrund gerückt werden. Demzufolge liegt es nahe, dass die Probanden mit gymnasialer Schulausbildung in einem Test, der die Allgemeinbildung abprüfen soll, auch bessere Testergebnisse erzielen. Unter der Annahme, dass man durch eine längere Konfrontation mit dem schulischen Lehrstoff auch mehr Wissen generiert, kann die Forschungsannahme gerechtfertigt werden und in diesem Fall aufrecht gehalten werden. Eine ergänzende Erklärung für die Wissensunterschiede zwischen Berufsfachschülern (Nicht-Gymnasiasten) und Gymnasiasten kann darin gesehen werden, dass die Probanden mit gymnasialer Ausbildung schon ab der Grundschule selektiert und entsprechend gefördert werden. Dieser Argumentation folgend kann man auch davon ausgehen, dass die kognitiven Wissensstrukturen besser ausgebildet wurden. Aufgrund der Tatsache, dass neue Informationen mit bereits vorhandenen Wissensstrukturen besser verknüpft werden können, wird die zukünftige Wissensaufnahme und damit die Lernleistung effektiver gestaltet (Kruse & Rudinger, 1997; siehe auch Kapitel 5).

Eine weiterführende Ursache für die Leistungsunterschiede sehen Baumert und Schümer (2002) in dem Zusammenhang zwischen Merkmalen der sozialen Herkunft und der Bildungsbeteiligung. Die Autoren postulieren, dass der „sozial diskriminierende Effekt“ (S. 163) bzw. die sozialen Disparitäten einen nicht unerheblichen Einfluss auf die Entscheidungsalternative sowie auf die frühzeitige Ausbildung von Kompetenzen haben, so dass der Besuch einer höheren Schule oftmals verwehrt bleibt. In Anlehnung an Boudon (1974) unterscheiden sie hierbei zwischen primären und sekundären sozialen Ungleichheiten. „Unter primären Ungleichheiten verstehen sie Unterschiede in den bis zu einer Übergangsschwelle erworbenen und für die nächste Etappe vorausgesetzten Kompetenzen, die in der Regel von der sozialen Herkunft nicht unabhängig sind. Als sekundäre Ungleichheiten bezeichnen sie soziale Disparitäten, die bei gleichen

Kompetenzen aus einem je nach sozialer Lage der Familie unterschiedlichen Entscheidungsverhalten entstehen“ (Baumert & Schümer, 2002, S. 163). Im Bezug auf die primären Ungleichheiten wird die Sozialschichtzugehörigkeit diskutiert, die die Ausbildung von Fähigkeiten und Qualifikationen für die eine oder andere Schule begünstigt bzw. hemmt. In diesem Kontext wird zunächst dokumentiert, dass über 50% der Schüler aus Familien der „oberen Dienstklasse“ (Klassifikation von Berufen durch das internationale Arbeitsamt- sogenannte EGP-Klassenraster I-II) die gymnasiale Oberstufe besuchen, wohingegen der Prozentsatz der Abiturienten mit niedriger werdenden Sozialschicht auf unter 10% sinkt (EGP-Klassen V-VII). Dieser Sachverhalt lässt sich auch in dem vorliegenden Datensatz replizieren. Aus der Tabelle 37 lässt sich entnehmen, dass Probanden aus Familien mit höherem Bildungsniveau vergleichsweise häufig das Gymnasium besucht haben (bis zu 93,6%, wenn die Bildung der Mutter einen Doktorgrad erreicht hat), wohingegen das Verhältnis bei niedrigem Bildungsniveau nahezu ausgeglichen ist. Dieser Befund wird gleichermaßen durch die anschließenden multiplen Regressionsanalysen bekräftigt. In der Modellzusammenfassung wird dementsprechend ein Korrelationskoeffizient von  $r = .347$  ausgewiesen, der hochsignifikant ausfällt.

*Tabelle 37: Kreuztabelle: Bildungsniveau der Eltern & Schulabschluss der Kinder*

Bildungsabschluss der Mutter	Probanden der Stichprobe (in %)	
	Nicht-Gymnasiasten	Gymnasiasten
Kein Schulabschluss	52,8%	47,2%
Hauptschulabschluss	45,5%	54,5%
Fachhochschulstudium	21,4%	78,6%
Hochschulstudium	8,9%	91,1%
Doktorgrad	6,7%	93,3%
Bildungsabschluss des Vaters		
Kein Schulabschluss	47,8%	52,2%
Hauptschulabschluss	44,9%	55,1%
Fachhochschulstudium	21,4%	78,6%
Hochschulstudium	9,1%	90,9%
Doktorgrad	5,0%	95,0%

Unter der Annahme, dass das Wissen und weitere Leistungsmerkmale auch durch das soziokulturelle Umfeld beeinflusst bzw. geprägt werden, können die Wissensdifferenzen begründet werden. Im Weiteren wurde der Leistungsunterschied zwischen Schülern mit Migrationshintergrund bzw. ohne Migrationshintergrund erörtert und in die Analyse mit einbezogen. Es stellte sich zum Einen heraus, dass Schüler und deren Familien mit Migrationshintergrund vermehrt der EGP-Klassen V-VII zugeordnet werden. Zum Anderen konnten eindeutige positive Zusammenhänge zwischen der Lesekompetenz und dem Sozialschichtindex aufgedeckt werden<sup>23</sup> sowie zwischen der Lesekompetenz und verschiedenen Schulfächern. Die Schüler mit Migrationshintergrund erzielten stets schlechtere Leistungen, welche unter anderem durch den Sozialschichtindex, die Lesekompetenz und die gewählte Schulform erklärt werden können. Anhand der vorliegenden Stichprobe wird ebenfalls ein Hinweis auf den Zusammenhang zwischen sozialer Herkunft und der optierten Bildungseinrichtung gegeben. Demzufolge besuchten 71,2% der Probanden ohne Migrationshintergrund (Muttersprachler-D) die gymnasiale Oberstufe.

Anhand der Ausführungen bleibt festzuhalten, dass das Bildungsniveau ein entscheidender Indikator für die Allgemeinbildung ist und ein Prädiktor für die spätere Lernleistung sein kann. Die Annahme lässt sich somit ohne Einschränkungen aufrecht erhalten werden. Darüber hinaus kann der Befund als Indikator für die Validität gewertet werden, da der Wissenstest bzw. Leistungstest eine eindeutige Differenzierung zwischen verschiedenen Bildungsklassen erlaubt.

### **11.3 Diskussion der Befunde zu den Forschungsfragen- Teil 2**

#### ***Diskussion der Befunde zur Forschungsfrage und Forschungsannahmen- 2.1: Übereinstimmungsvalidität***

Hinsichtlich der aufgestellten Forschungsfrage lässt sich feststellen, dass die Ergebnisse, in Anbetracht der psychometrischen Qualität von Schulnoten und den vielfältigen Einflussfaktoren, differenzierte Schlussfolgerungen zulassen. Die ermittelten Zusammenhänge weisen gemäß den oben aufgeführten Kriterien zwar geringe, aber

---

<sup>23</sup> Die Resultate zeigen, dass bei Veränderung der Sozialschicht um eine Standardabweichung die Lesekompetenz um etwas mehr als eine halbe Kompetenzstufe steigt oder sinkt (Baumert & Schümer, S. 182).

überwiegend hochsignifikante Korrelationen auf. Ausgenommen zwischen den Skalenergebnissen und der Kunstnote konnten nur sehr geringe und zudem nicht signifikante Korrelationen ermittelt werden.

Bezogen auf die erste Forschungsannahme ist zu konstatieren, dass –bis auf die Kunstnote– sämtliche Koeffizienten negativ sind und die Hypothese somit aufrecht erhalten werden kann. Die Analyse der zweiten Forschungsannahme zeigt erwartungsgemäß signifikante, sogar hoch signifikante Zusammenhänge zwischen den Skalenscores und der Durchschnittsnote sowie zwischen dem Gesamtscore und den aufgeführten Noten auf. Die Koeffizienten liegen im Bereich von  $r = -.174$  bis  $r = -.273$  ( $p < .01$ ). Die Ausnahme bildet wiederum die Kunstnote. Bei Vernachlässigung der Kunstnote, liegen die Kennwerte folglich in einem akzeptablen Bereich. Die dritte Forschungshypothese kann ohne Einschränkungen bestätigt werden. Die Korrelation zwischen der Mathematiknote und der numerischen Wissensskala beträgt  $r = -.291$  ( $p < .01$ ). Dieser Kennwert markiert gleichzeitig auch den stärksten Zusammenhang. Im Rekurs auf die vierte Forschungsannahme lässt sich feststellen, dass die verbal kodierte Inhaltsklasse hoch signifikant mit der Deutschnote ( $r = -.190$ ;  $p < .01$ ) sowie der Englischnote ( $r = -.247$ ;  $p < .01$ ) korreliert. Die berechneten Werte liegen zwar nach den Beurteilungskriterien nur im niedrigen Bereich, befinden sich jedoch auf einem höheren als das vorgegebene Signifikanzniveau von  $p < .05$ . Aufgrund der Prämisse, dass der erfolgreiche Umgang mit figuralbildhaftem Testmaterial auch mit einer guten Note im Fach Kunst einhergeht, wurde die fünfte Forschungsannahme formuliert. Entgegen der Hypothese weisen die Resultate keine bzw. unbedeutende Korrelationen auf, so dass die Annahme nicht aufrecht erhalten werden kann.

Insgesamt können die Zusammenhänge als akzeptabel und positiv beurteilt werden. Gleichzeitig lassen sie nur eingeschränkte Aussagen bzw. Interpretationen zu, weil die Korrelationen ausschließlich in einem niedrigen Bereich liegen.

Angesichts der dargestellten Studien und Befunde bezüglich des Zusammenhangs zwischen Wissenstestwerten und Schulnoten (vgl. z.B. Jensen, 1998; Amthauer et al., 2001) wären stärkere Zusammenhänge zu erwarten gewesen. Auch auf Basis der Annahme, dass das abgeprüfte Allgemeinwissen und die Schulnoten Ausdruck kognitiver Leistung und Lernfähigkeit sind, wären höhere Korrelationen denkbar gewesen. Dennoch konnten nur partiell bestätigende Ergebnisse ermittelt werden. Ein entscheidender Grund für die niedrigeren Korrelationen ist, dass die Schulnoten in der Gesamtstichprobe – im Gegensatz zu den Rohwerten der Wissensskalen – teilweise nicht normalverteilt sind. Eine



weitere mögliche Erklärung für die teils schwachen Korrelationen ist, dass der schulische Lehrstoff nur in einem geringen Ausmaß mit der inhaltlichen Ausgestaltung des Wissenstests konvergiert. Das heißt, der START-W prüft tendenziell allgemeines Wissen ab, welches aus verschiedenen Themenbereichen stammt. Die einzelnen Schulfächer sind dagegen eher spezifisch ausgelegt und erfassen nicht die Bandbreite des Tests. Darüber hinaus können auch unterschiedliche und subjektive Bewertungskriterien als Grund für die geringen Zusammenhänge angeführt werden. Die Leistungsbeurteilung der Lehrer beruht nicht nur auf dem abgeprüften faktischen Wissen der Schüler, sondern auch auf anderen Leistungsindikatoren, so dass die Entstehung von Zensuren multifaktoriell bedingt ist. In die Notengebung fließen demnach die Motivation bzw. die Beteiligung am Unterrichtsgeschehen, das Sozialverhalten, Fehlzeiten, etc. mit ein. Im Gegensatz dazu werden im Rahmen des Wissenstests nur die objektiv erzielten Testergebnisse zur Beurteilung der Leistung herangezogen. Eine weitere Ursache besteht womöglich auch in der schlechten Vergleichbarkeit der Probanden, die aus verschiedenen Bildungsgruppen kommen und deren Noten auf Grundlage unterschiedlicher Anforderungen zustande kommen. Demzufolge kann man davon ausgehen, dass der inhaltliche Schwerpunkt des Lehrstoffs von der spezifischen Fachrichtung abhängig ist und die Bewertung einzelner Fächer auf Basis anderer Kriterien bestimmt wird. In diesem Kontext setzt sich bspw. die Durchschnittsnote der Hauptschüler oder für Abiturienten auch aus unterschiedlichen Fächern zusammen. Unter Berücksichtigung der differierenden schulischen Anforderungen bzw. Schwerpunktsetzungen und der damit verbundenden Zusammensetzung der Durchschnittsnote erscheinen die geringen Korrelationen nicht nur nachvollziehbar, sondern können auch als beachtlich eingestuft werden. Als weitere alternative Begründung für die Verzerrungen bzw. Diskrepanzen können auch lückenhafte Erinnerungen an die Noten des letzten Abschlusszeugnisses angeführt werden. Weiterhin kann auch eine unzureichende Homogenität der kodierten Wissensskalen als Erklärung mit einbezogen werden. Bei einer Untersuchung der kodierten Testitems fällt auf, dass einige Fragen aus inhaltlich verschiedenen Themengebieten kommen. Sofern beispielsweise nur die Aufgaben zur Rechenfertigkeit betrachtet werden, die auch im Fach Mathematik von Bedeutung sind, reduziert sich die Anzahl der Testitems auf acht Aufgaben. Folglich muss eine hohe Korrelation zwischen der Mathematiknote und der numerischen Wissensskala nicht zwangsläufig gegeben sein.

Vor dem Hintergrund der Ausführungen bleibt abschließend festzuhalten, dass die dokumentierten Befunde als zufriedenstellend, annehmbar und aussagekräftig bewertet werden können und einen eindeutigen Hinweis auf die Kriteriumsvalidität liefern.

### ***Diskussion der Befunde zur Forschungsfrage und Forschungsannahmen- 2.2:***

#### ***Prognostische Validität***

Im theoretischen Teil wurde bereits herausgestellt, dass das Allgemeinwissen als Indikator für zukünftiges Lernverhalten gesehen wird. In diese Richtung weisen auch die vorgestellten Forschungsergebnisse verschiedener Autoren (vgl. z.B. Schuler & Höft, 2006). Im Hinblick auf die praxisorientierten Verwendung des Testverfahrens zählt die prognostische Validität zu den wichtigsten Gütekriterien, da im Rahmen der Analysen ein Bezug zwischen Testergebnissen und dem zukünftigen Schul-, Ausbildungs- oder Berufserfolg hergestellt wird. Aufgrund der dem Testverfahren inhärenten Ziel, nämlich der Prognose des Erfolges, erlangt die Vorhersagevalidität somit einen höheren Stellenwert als die Übereinstimmungsvalidität (Lienert & Raatz, 1998; Fisseni, 2004, Amelang & Schmidt-Atzert, 2006). Infolge dessen wird auf der Grundlage des vorliegenden Datensatzes der Zusammenhang zwischen den Testwerten des START-W-Wissenstests und dem Ausbildungserfolg für Gesundheitsberufe untersucht.

Das grundlegende Problem der kriterienbezogene Validierung drückt sich in der Auswahl eines geeigneten Kriteriums aus, welches seinerseits den Gütekriterien genügen sollte. Außerdem sollte ein Kriterium in Betracht gezogen werden, das repräsentativ und relevant für das zu erfassende Merkmal bzw. die auszuführende Tätigkeit ist (Michel & Conrad, 1982). Hinzu kommt, dass zwischen dem Prädiktor und dem Kriterium im Zeitverlauf Veränderungen zu berücksichtigen sind. Letztendlich hängt die Güte der Validität in erheblichen Maß von dem Kriterium ab. In der vorliegenden Studie wurde der Wissenstests auch als Prädiktor für den späteren Ausbildungserfolg eingesetzt. Für die Bewertung der Leistung im Rahmen der beruflichen Ausbildung wurden die Noten des ersten (Probe-) Jahres erfasst. Die daraus resultierende Durchschnittsnote wurde im Zuge von Lehrerkonferenzen und nach Absprache mit der Institutsleitung ermittelt. Es kann davon ausgegangen werden, dass die psychometrische Qualität der Durchschnittsnote auf einem akzeptablen Niveau ist und sich als Außenkriterium zur Beurteilung der Vorhersagevalidität eignet. Die berechneten Korrelationen bewegen sich in der Gesamtstichprobe zwischen  $r = -.287$  und  $r = -.408$  (Korrelationskoeffizient nach Spearman), die

allesamt hoch signifikant ausfallen. Bezüglich der Substichprobe der Gesundheits- und Krankenpfleger wurden Koeffizienten im Bereich von  $r = -.200$  bis  $r = -.344$  berechnet, die wiederum signifikant sind ( $r = -.252$  bis  $r = -.363$  nach der Produkt-Moment-Korrelation von Pearson). Innerhalb der Gruppe der Operationstechnischen Assistenten liegen die Kennwerte zwischen  $r = -.299$  und  $r = -.583$  ( $p < .05$ ;  $r = -.516$  und  $r = -.609$  nach Pearson). Die Korrelationen liegen somit in einem niedrigen bis mittleren Bereich. Die ermittelten Werte können, in Anbetracht der obigen Ausführungen, als hinreichend beurteilt bzw. interpretiert werden. Aufgrund der Höhe der Zusammenhänge kann zudem konstatiert werden, dass der Wissenstest einen aussagekräftigen Indikator für den zukünftigen Ausbildungserfolg darstellt. Dennoch befinden sich einige Korrelationskoeffizienten noch im niedrigen Bereich, obwohl höhere Zusammenhänge zu erwarten gewesen wären. Eine mögliche Erklärung bezieht sich auf die inhaltliche Heterogenität des Außenkriteriums. Die Durchschnittsnote setzt sich hierbei aus unterschiedlichen Faktoren zusammen. Maßgeblich sind die Bewertungskriterien der entsprechenden Bildungseinrichtung. Korrespondierend mit den Erläuterungen zu den Schulnoten fließen auch in diese Notengebung verschiedene Leistungsmerkmale mit ein. Zudem umfasst der Lehrinhalt der Schulen für Gesundheitsberufe sehr spezifische sowie allgemeine Themengebiete (wie z.B. Recht und Wirtschaft), die wenig Übereinstimmung mit der inhaltlichen Ausgestaltung des Wissenstest aufweisen. Unter dieser Voraussetzung ist es ein positiver Aspekt, dass das erfasste Allgemeinwissen mit dem START-W einen Hinweis bezüglich der zukünftigen Erfolgchancen auf dem Gebiet der Gesundheitsberufe liefert. In Anlehnung an Schmidt und Hunter (1998) gehen Schuler und Höft (2006) überdies davon aus, dass „es praktisch keinen Beruf gibt, für den allgemeine Fähigkeits-tests keine kriteriumsbezogene Validität für berufliche Leistungsvariablen aufweisen. Da diese Verfahrensgruppe häufig der beste Prädiktor für Berufserfolg ist, legen sie in ihrer Übersicht [auch] besonderen Wert auf den Informationszuwachs, den andere Verfahren in Ergänzung liefern können“ (S. 106). In diesem Bezugsrahmen wird gleichermaßen die Bedeutung der START-Testbatterie hervorgehoben.

Daraufhin kann die Kriteriumsvalidität des START-W als hinreichend bezeichnet werden.

## 11.4 Diskussion der Befunde zu den Forschungsfragen- Teil 3

Die Konstruktvalidität stellt eine Synthese der dargelegten inhaltlichen und kriterienbezogenen Validität dar. Nach Lienert und Raatz (1998) geht man bei „einer Konstruktvalidierung (...) von dem Konstrukt und der entsprechenden Theorie aus und leitet daraus Hypothesen ab, die unter Einsatz des zu validierenden Tests empirisch überprüft werden“ (S. 226). Folglich soll mit der Konstruktvalidität anhand von sachlogischen Überlegungen, psychologischen (korrelationsstatistischen und faktoranalytischen) Analysen und empirischen Untersuchungen bestimmt werden, inwieweit der Test ein konkretes, theoretisches Konstrukt bzw. die zugrunde liegenden Eigenschaften erfassen kann. Zudem leitet sich der Bedeutungsgehalt eines Konstrukts auch aus den Beziehungen zu anderen theoretisch ähnlichen oder fremden Konstrukten ab (nomologisches Netzwerk), die ebenfalls empirisch überprüft werden sollen. Der Test bzw. das Testverhalten fungiert dabei als Parameter für das zu messende Konstrukt. Folglich wurde im Rahmen der Fragestellung 3 die konvergente Validität anhand des Bochumer Wissenstests ermittelt. Auf der Grundlage der ausführlich dargelegten Beziehung zwischen fluider und kristallisierter Intelligenz wird darüber hinaus jener Zusammenhang erforscht, um theoriekonforme Hinweise zu generieren und einen weiteren Indikator für die konvergente Validität zu erhalten.

In der einschlägigen Literatur werden jedoch keine konkreten Handlungsanweisungen zur Durchführung der Validierung gegeben. Das liegt unter anderem darin begründet, dass der Bedeutungsgehalt der Validierung weniger als abgeschlossener Zustand begriffen wird, sondern vielmehr als Prozess. Dennoch können vier allgemein akzeptierte Aspekte bzw. Vorgehensweisen der Konstruktvalidierung angegeben werden (Fisseni, 2004, S. 68):

1. Das Testkonstrukt wird aus einer Theorie abgeleitet oder einer Theorie zugeordnet.
2. Die Items, die das Konstrukt repräsentieren, werden sowohl semantisch als auch statistisch analysiert.
3. Der Zusammenhang des Testkonstrukts mit anderen Konstrukten („konstrukt-nah“ oder „konstruktfern“) wird untersucht bzw. analysiert.
4. Die Faktorenanalyse gilt als klassisches Instrument der Konstruktvalidierung.  
(Die Faktorenanalyse ist ein heuristischer Ansatz, der Strukturen innerhalb von Testverfahren aufdecken kann. Die zusammengefassten gleichartigen Items

können dabei als Indikatoren für eine gemeinsame Eigenschaftsdimension gedeutet werden. Anders ausgedrückt kann der Faktor, auf dem die Items hoch laden als Konstrukt interpretiert werden. Fischer (1974) drückt es folgendermaßen aus: „Da die Faktorenanalyse eine der gängigsten Methoden zur Auffindung von latenten Eigenschaftsdimensionen ist, und da es sich bei ihr zudem um eine Verallgemeinerung der klassischen Testtheorie handelt, [lässt sich] das faktorenanalytische Modell zur Präzisierung des Begriffes der Konstruktvalidität“ (S.77) heranziehen.

Die Konstruktvalidierung kann als Oberbegriff verstanden werden, die verschiedene Analysemethoden umfasst und unterschiedliche Validitätsarten konstruktzentriert zusammenführt. Bei dieser Vorgehensweise spielen, wie oben angedeutet, zwei Validitätsarten eine besondere Rolle (Lienert & Raatz, 1998; Fisseni, 2004, Amelang & Schmidt-Atzert, 2006; Rost, 2004):

1. Konvergente Validität
2. Faktorielle Validität

### ***Diskussion der Befunde zur Forschungsfrage und Forschungsannahmen- 3.1:***

#### ***Konvergente Validität***

Bezugnehmend auf die konvergente Validität (Forschungsfrage 3.1) wird der Zusammenhang zwischen dem START-Wissenstest und dem Bochumer Wissenstest (BOWIT Kurzform, Hossiep & Schulte, 2008) analysiert. Der BOWIT ist geeignet, die konvergente Validität zu bestimmen, da der Test auf einer vergleichbaren Theorieorientierung und inhaltlichen Konzeption basiert. Das Testverfahren bietet die Möglichkeit das erworbene Wissen bzw. die kristallisierte Intelligenz abzubilden.

Die Berechnungen zeigen, dass zwischen den Testergebnissen des START-W und dem BOWIT-Kurzform eindeutige Zusammenhänge bestehen. Hinsichtlich des Gesamtscores wird eine Korrelation von  $r = .673$  ( $p < .01$ ) ermittelt. Unter der Annahme, dass die Rohwerte intervallskaliert und normalverteilt sind erhöht sich der Korrelationskoeffizient nach Pearson auf  $r = .704$ , der ebenfalls signifikant ist. Substantielle Kennwerte ergeben sich auch zwischen den einzelnen Inhaltsskalen und dem Gesamtscore des BOWIT, die in einem Bereich von von  $r = .554$  bis  $r = .678$  liegen ( $p < .01$ ;  $r = .604$  bis  $r = .704$  nach der Produkt-Moment Korrelation von Pearson). Die Werte können als „mittel“ bis „hoch“ interpretiert und als zufriedenstellend bewertet werden. Es wird jedoch angemerkt, dass der von Hossiep & Schulte (2008) ermittelte Zusammenhang

zwischen der Vollversion des BOWIT und dem Wissensmodul des I-S-T 2000 R (Liepmann et al., 2007) mit  $r = .75$  ( $p < .01$ ;  $N = 68$ ) höher ausgefallen ist. Zumal der START-W zu einem erheblichen Anteil auf den Test-Items des Erweiterungsmoduls des I-S-T 2000 R basiert, wären zumindest gleich hohe Korrelationskoeffizienten denkbar gewesen. Mögliche Erklärungsansätze liegen darin, dass die aus ökonomischen Gründen reduzierte Kurzform des BOWIT nicht alle aufgeführten Facetten berücksichtigt und nur die trennschärfsten Items der beiden aufgeführten Faktoren einbezieht. Außerdem konstatieren die Autoren, dass „eine gesonderte Validierung der beiden Kurzversionen“ (S. 55) noch aussteht und zu einer gleichwertigen bzw. intensiveren Auseinandersetzung mit dem Wissensstand der Probanden die Vollversion herangezogen werden sollte. Aufgrund dessen erscheinen die berechneten Kennwerte als positive und befriedigende Hinweise auf die konvergente Validität.

Obwohl die Konstruktvalidierung bzw. die konvergente Validität wichtige wissenschaftstheoretische Hinweise zur Verknüpfung von theoriegestützten Konzepten und praktisch-psychologischen Maßnahmen liefert, bemängeln Kritiker die fehlende methodische Stringenz und die damit verbundene präzise Elaboration der Konstrukte (Ameilang & Schmidt-Atzert, 2006; Fisseni, 2004; Michel & Conrad, 1982). Demgegenüber steht jedoch die dem Konstrukt-Verständnis zugrunde gelegte Prämisse, dass das zu identifizierende Konstrukt ständig erweiterungsfähig ist und erst zum Zeitpunkt des aktuellen Forschungsstandes seine Bedeutung entfaltet. Die dem Konstrukt inhärente Eigenschaft der Erweiterungsfähigkeit bzw. die offene Konzeption eines Konstrukts lässt folglich eine präzise Elaboration per Definition nicht zu. Hinsichtlich des erstgenannten Kritikpunktes kann in Anbetracht der gründlichen, wissenschaftlichen und an den Gütekriterien sowie der DIN-Norm 33430 orientierten Verfahrensweise festgehalten werden, dass die vorgegebene methodische Systematik eingehalten und umgesetzt wurde. Resümierend deuten die Zusammenhänge somit darauf hin, dass beide Testverfahren konvergente Konstrukte darstellen und das erworbene Wissen bzw. die kristallisierte Intelligenz partiell abbilden können.

### ***Diskussion der Befunde zur Forschungsfrage und Forschungsannahmen- 3.2:***

#### ***Faktorielle Validität***

Ogleich die Faktorladungsmatrix nur geringe bis mittlere Kennwerte aufweist – insbesondere die Ladungen auf den „nested“ Faktor  $gc^{\prime}$  – spricht der CFI

(*Comparative Fit Index*) für eine sehr gute Passung des gesamten Modells. Auch die Kennziffern des RMSEA und des SRMR liegen in einem zufriedenstellenden Bereich.

Die nachfolgenden Ausführungen beziehen sich zunächst auf das gesonderte Strukturgleichungsmodell, in dem gc als Faktor dritter Ordnung aus den Faktoren wV, wN und wF ermittelt wurde. Auf der Basis von multivariaten Analysemethoden, zu denen auch die konfirmatorische Faktorenanalyse zählt, wurde in einem ersten Schritt geprüft, ob sich die intendierten Themenbereiche empirisch differenzieren lassen. Entsprechend der in Tabelle 35 illustrierten Kreuzklassifikationsmatrix (3 Kodierungsformen x 6 Themenbereiche) wurde ein Aggregat der zugewiesenen Wissensfragen erstellt. Analog setzt sich das verbal kodierte Wissen aus verbal kodierten Fragen zu Wirtschaft, Naturwissenschaften, Kultur, Geografie/Geschichte und Alltag zusammen. Es zeigte sich, dass die Schätzung der Zellenaggregate auf die sechs Wissensbereiche (Faktoren) in dem vorgegebenen Intervall von .20 und .80 liegen. Folglich konnte die Identifikation der themenspezifischen Faktoren bestätigt werden und das Modell ermöglicht eine akzeptable Angleichung an die Daten auf der ersten Generalitätsebene. Darüber hinaus konnte nachgewiesen werden, dass die kodierte Wissensaggregate auf einem mittleren bis hohen Niveau mit den inhaltsbezogenen Wissensfaktoren korrelieren. Demgemäß konnte eine weitere Dimensionsreduktion der wissensthematischen Facetten in drei inhaltsgebundene Facetten dargestellt werden.

Aufgrund der Komplexität des Modells und der damit im Zusammenhang stehenden Schätzung einzelner Parameter wurden die Ladungskoeffizienten auf gc mit 1 fixiert. Das heißt zur Varianzbestimmung von gc mussten die Pfade auf 1 fixiert werden, da sonst die Freiheitsgrade derart gestiegen wären, dass die einzelnen Parameter nicht mehr geschätzt werden können. Die Vorgehensweise und die Interpretation des Faktors als kristallisierte Intelligenz kann unter anderem damit begründet werden, dass bedeutende Aspekte der kristallisierten Intelligenz durch Wissenstests abgebildet werden (Horn, 1988). Es kann festgehalten werden, dass das Strukturgleichungsmodell alle Annahmen über den START-W Wissenstest abbildet und danach die intendierte Struktur adäquat darstellt. Das erfasste deklarative Wissen lässt sich sowohl nach den Themenbereichen als auch nach den Inhaltsklassen strukturieren und bietet zudem die Möglichkeit, die kristallisierte Intelligenz zu repräsentieren.

In Anlehnung an die im I-S-T 2000 R (Liepmann et al., 2007) aufgestellten Interpretationen der Skalen und Faktorwerte wird die nachstehende Tabelle 38 hinzugezogen (S. 81f):

*Tabelle 38: Interpretation der Skalen und des Faktorwertes Wissen (gc)*

Skalen & Faktorwert	Interpretation
Gesamtskala Wissen	Mit dieser Skala werden wesentliche Ausschnitte des Wissens erfasst, das jemand im Laufe seines Lebens in unserem Kulturkreis erworben hat. Diese Wissensskala deckt einen recht breiten Bereich von Themen ab (Geographie/ Geschichte, Wirtschaft, Kunst/ Kultur, Mathematik, Naturwissenschaften, Alltagswissen).
Inhaltsklassen	Mit den Skalen verbal, numerisch und figural kodiertes Wissen werden wesentliche Ausschnitte des in der jeweiligen Kodierungsart (verbal, numerisch, figural) vorliegenden Wissens abgebildet. Auch diese Skalen umfassen jeweils die Themenbereiche Geographie/ Geschichte, Wirtschaft, Kunst/ Kultur, Mathematik, Naturwissenschaften und Alltagswissen.
Verbale Wissensskala	Mit dieser Skala wird die Fähigkeit zum Umgang mit sprachlichem Material im Rahmen von wissensbasierten Leistungen erfasst.
Numerische Wissensskala	Diese Skala erfasst im Kontext wissensbasierter Kompetenzen, die Rechenfertigkeit sowie den Umgang und den Erwerb von numerischen Wissensbestandteilen
Figurale Wissensskala	Es wird die Fähigkeit zum Umgang mit figural-bildhaften bzw. die Aneignung von bildhaftem Material erfasst. Diese Fähigkeit umfasst sowohl den Umgang mit zweidimensionalen und dreidimensionalen Figuren sowie die Fähigkeit Proportionen von Flächen und Gegenständen zu erfassen.
Kristallisierte Intelligenz gc	Probanden mit einer hohen Ausprägung auf dem Faktor sind in der Lage, in unserer Kultur ein umfangreiches Wissen zu erwerben. Das mit diesem Faktor erhobene Wissen wird meist in Institutionen der Gesellschaft wie Schulen, Hochschulen und anderen Bildungseinrichtungen erworben und als Leistung gefordert. Dieser Faktor stellt eine neue Operationalisierung der kristallisierten Intelligenz im Sinne Cattels (1987) dar.

Abschließend wird darauf hingewiesen, dass das ermittelte Pfaddiagramm bzw. Strukturgleichungsmodell zwar theoriekonforme Ergebnisse liefert, aber aufgrund der teils geringen Schätzparameter nur als erster Indikator auf die Struktur zu verstehen ist.



***Diskussion der Befunde zur Forschungsfrage und Forschungsannahmen- 3.3:******Zum Zusammenhang zwischen fluider Intelligenz und kristallisierter Intelligenz (Wissen)***

Die Forschungsfrage 3.3 thematisiert den Zusammenhang zwischen fluider Intelligenz (gf) und kristallisierter Intelligenz (gc). Hierzu wurde ein Strukturgleichungsmodell geschätzt, in dem gc als Faktor dritter Ordnung aus den Faktoren verbales Wissen, numerisches Wissen und figurales Wissen bestimmt wurde. Dieser Faktor gc soll die Wissensvarianzen auf einem höheren Generalitätsniveau repräsentieren. Gleichzeitig wurde ein sogenannter `nested` Faktor eingeführt, der die spezifischen Wissensvarianzen darstellt, die nicht von den inhaltsbezogenen Wissensfaktoren erklärt werden können. Der Faktor wurde aus den themenbezogenen Faktoren ermittelt und besitzt demnach ein geringeres Generalitätsniveau. Aus dem Modell lässt sich entnehmen, dass der Faktor gf mehr Varianz an gc erklärte als an gc' (siehe Abbildung 20). In einem  $\chi^2$ -Differenzentest war der Unterschied zwischen den Pfadkoeffizienten signifikant ( $\chi^2_{\text{diff}}(1) = 135.45$ ;  $p < .001$ ). Darüber hinaus war der Pfad von gf nach gc' negativ. Offenbar hängen die nicht an die Inhaltsfaktoren gebundenen residualen Wissensvarianzen nur geringfügig und negativ mit gf zusammen. Eine Erklärung für den hohen Pfadkoeffizienten zwischen gf und gc (.83) ist in der Investmenttheorie von Cattell (1971a, 1987) zu sehen. Nach der Theorie wird die fluide Intelligenz in die kristallisierte Intelligenz investiert. „The term `crystallized` is meant to imply this freezing in a specific shape of what was once fluid ability“ (Cattell, 1987, S.140). Die fluide Intelligenz, welche als unspezifische Energie definiert wird und die in Form von Basisprozessen das Potential der geistigen Leistungsmöglichkeiten widerspiegelt, steckt demnach die Grenzen ab, innerhalb derer sich gc herauskristallisiert. Die fluide Intelligenz determiniert in diesem Sinne das Entwicklungspotential der kristallisierten Intelligenz (Stern, 2001). Bedingt durch zeitlich nachgelagerte systematische Akkulturationsprozesse kristallisieren sich auf der Basis fluider Fähigkeiten dezidierte Kompetenzen heraus, die eng mit dem Wissen verknüpft sind (für eine detailliertere Darstellung siehe Kapitel 6.2). Im Hinblick auf die Beziehung zwischen gf und gc scheint die dargelegte Theorie von Cattell (1987) auch anhand der vorliegenden Daten bestätigt zu werden und erlaubt einen Rückschluss auf den postulierten Zusammenhang. In eine andere Richtung weist jedoch der Pfadkoeffizient zwischen gf und gc', der sogar einen negativen Kennwert von -.22 aufweist. Die geringen, negativen Zusammenhänge können so interpretiert werden, dass die spezifischen Wissensanteile nicht nur Fähigkeitsvarianz erfassen. Der Befund stellt zunächst ein

widersprüchliches Ergebnis dar. Einerseits sind die ausgewiesenen Faktorladungen auf  $gc^{\circ}$  sehr niedrig und erschweren somit eine Interpretation. Andererseits liegt  $gc^{\circ}$  auf einem geringeren Generalitätsniveau und erfasst demnach nur die spezifischen Wissensvarianzen, die nicht an die Inhaltsfaktoren gebunden sind. Gemäß einer theoriekonformen Auslegung wird das Wissen jedoch eher als allgemeines Wissen verstanden, welches als Summe von Kenntnissen definiert wird, das eine Person im Laufe der Zeit akkumuliert (Gilardi et al., 1983; Kruse & Rudinger, 1997). Auch der Ausdruck *acculturation knowledge*, den Horn (1988) für kristallisierte Intelligenz benutzt, bekräftigt diese Auffassung. Dieser Argumentation folgend liegt es nahe, dass der Zusammenhang zwischen  $gf$  und  $gc$  höher ausfällt als zwischen  $gf$  und  $gc^{\circ}$ . Die geringe Korrelation lässt vielmehr den Schluss zu, dass die fluide Intelligenz den Aufbau von allgemeineren Wissensstrukturen (Allgemeinwissen) fördert. Für den Erwerb von bereichsspezifischem Wissen hingegen spielen fluide Fähigkeiten offenbar keine bzw. nur eine untergeordnete Rolle.

Zusammenfassend lassen die Befunde den Schluss zu, dass die  $gc$ -Indikatoren in substantiellem Umfang auch andere Varianzen erfassen und diese vom Aggregationsniveau abhängen. Überdies lassen sich erste Hinweise auf den postulierten Zusammenhang zwischen fluider Intelligenz und kristallisierter Intelligenz finden und legitimieren damit weitere Forschungsbemühungen. Aufgrund der Beziehung zwischen den Konstrukten können zusätzliche Anhaltspunkte über die konvergente Validität des START-W Wissenstests abgeleitet werden.

## 12 Fazit und Ausblick

Mit der Arbeit wird das Ziel verfolgt, die Bedeutung des Wissens bzw. der kristallisierten Intelligenz für die Wirtschaft, Gesellschaft sowie für den schulischen, akademischen und beruflichen Werdegang herauszustellen. Wissen und Intelligenz werden in diesem Zusammenhang als Schlüsselkompetenzen begriffen, die den zukünftigen Wissenserwerb fördern und dazu dienen, flexibel auf die dynamischen Veränderungen beruflicher Anforderungen reagieren zu können. Hinsichtlich der eingangs dargestellten Studien zeigt sich, dass viele Schüler und Berufseinsteiger Defizite in der Allgemeinbildung aufweisen. Unternehmen und berufsberatende Institutionen fordern demzufolge eine geeignete Kompetenzfeststellung (Eignungsdiagnostik) und die Weiterentwicklung so-

lider Basiskompetenzen. Vor diesem Hintergrund wird insbesondere die Relevanz von Testverfahren hervorgehoben, die das Allgemeinwissen erfassen und die zukünftige Eignung für den Beruf diagnostizieren und prognostizieren können. Einige Studien bestätigen bereits, dass die Tests zur Erfassung der kristallisierten Intelligenz bzw. des Wissens gute prädiktive Validitäten bezogen auf den beruflichen Erfolg aufweisen (z.B. Beauducel & Kersting, 2002; Nettelnstroth, 2003, 2004; Schmidt & Hunter, 1998; siehe Kapitel 3.4). Die Validierungsstudie trägt dazu bei, einen Wissenstest (START-W) weiter zu entwickeln, der den geltenden wissenschaftlichen Gütekriterien genügt und an den Anforderungen der DIN-Norm 33430 orientiert ist.

Ausgehend von der klassischen Testtheorie wurde, darauf aufbauend, das Konstrukt Intelligenz in seinen verschiedenen Facetten dargestellt, wobei die Akzentuierung auf der kristallisierten Intelligenz lag. Es wurde an die bereits vorhandenen Theoriemodelle angeknüpft, insbesondere an das Strukturmodell von Cattell (1971a, 1987). Auf Basis des engen Zusammenhangs zwischen kristallisierter Intelligenz und Wissen wurde das Konstrukt Wissen näher betrachtet und das Beziehungsgeflecht gesondert erörtert, um ein theoretisches Fundament für das zu messende Merkmal Wissen zu schaffen. Außerdem wurden mögliche Einflussfaktoren auf die Wissensausprägung diskutiert und die Interdependenz von Anlage, Umwelt und Wissen thematisiert.

Als Grundlage für die Validierungsstudie wurde eine repräsentative Stichprobe von potentiellen Berufseinsteigern avisiert, die dem zukünftigen Anwendungsbereich der START-Testbatterie entspricht. Die empirischen Ergebnisse lieferten in diesem Zusammenhang eindeutige Hinweise auf die Reliabilität der Wissensskalen und die Validität der vorläufigen Testversion. Anhand von Item-Analysen wurde zunächst geprüft, ob die einzelnen Aufgaben die Kriterien der Testkonstruktion erfüllen. Bezogen auf die Schwierigkeitsindizes stellte sich heraus, dass die Kennwerte prinzipiell in einem zufriedenstellenden Bereich von .10 bis .96 liegen. Die Skalen beinhalten demnach die für die Zielsetzung des Verfahrens angemessenen Aufgaben. Die Trennschärfekoeffizienten in den einzelnen Inhaltsklassen des selektierten Itempools rangieren im Mittel zwischen .26 und .30. Nach den berücksichtigten Beurteilungsbestimmungen liegen die Kennwerte in einem niedrigen bis mittleren Wertebereich und können als akzeptabel beurteilt werden. Die Reliabilitätsschätzung der gesamten Wissensskala weist eine interne Konsistenz von  $\alpha = .923$  auf. Für die verbale, numerische und figurale Inhaltsskala wurde in der angegebenen Reihenfolge eine interne Konsistenz von  $\alpha = .797$ ,  $\alpha = .778$  sowie  $\alpha = .832$  berechnet. Die Werte liegen damit in einem annehmbaren und zufriedenstellenden

Bereich. Im Hinblick auf die Ausführungen in Kapitel 10.1 ist die Güte der Itemkennwerte vertretbar. Obgleich eine weitere Optimierung der Skalen- und Itemkennwerte nicht ausgeschlossen werden kann. Die Split-half-Reliabilität weist in den Inhaltsklassen einen Korrelationskoeffizienten zwischen  $r = .785$  und  $r = .835$  ( $p < .01$ ) auf. Für den Gesamttest wird ein Zusammenhang von  $r = .929$  ( $p < .01$ ) ermittelt (odd-even-Methode). Die Retest-Reliabilitätsschätzungen für die Inhaltsskalen liegen in einem Bereich von  $r = .821$  bis  $r = .835$  ( $p < .01$ ). Der Gesamttest zeigt einen Wert von  $r = .88$  auf ( $p < .01$ ). Auf der Basis der dargestellten Befundlage ist der START-W hinreichend reliabel.

Im theoretischen Teil der Arbeit wurde herausgestellt, dass die kristallisierte Intelligenz, die vom Individuum aufgenommenen und organisierten Wissenssysteme einer Kultur repräsentiert und dass die Wissensausprägung von verschiedenen Umwelteinflüssen bedingt wird. Darauf aufbauend wurden die Forschungsfragen Teil 1 abgeleitet. Unter Verwendung des START-W, der weitgehend die kristallisierte Intelligenz abbildet, werden die Fragen untersucht, ob Wissensdifferenzen in Abhängigkeit des Alters, der Muttersprache, des Geschlechts und des Bildungsniveaus bestehen.

Die erste Annahme bezieht sich auf die zunehmende Wissensakkumulation im Lebenslauf. Anhand der Ergebnisse zeigt sich, dass die Hypothese nicht aufrecht erhalten werden konnte. Der F-Test weist zwar signifikante Mittelwertunterschiede auf, die bei einer differenzierteren Betrachtung mittels des Scheffé-Tests keine annahmegemäßen Schlussfolgerungen zulassen. Unter den genannten Voraussetzungen können drei homogene Untergruppen gebildet werden. Die 16-Jährigen können der Untergruppe 1 zugeordnet werden, die 17-18-Jährigen sowie die 26-46-Jährigen werden der Untergruppe 2 zugewiesen und die 19-25-Jährigen weisen die höchste Wissensausprägung auf (Untergruppe 3). Vielmehr scheint es zu sein, dass die Ausprägung des Wissens stärker vom Bildungsniveau und vom zeitlichen Abstand des Schulabgangs abhängt.

Unter der Annahme, dass das Wissen von Personen zum Teil durch unterschiedliche kulturelle Einflüsse geprägt worden ist und aufgrund der Tatsache, dass der START-W ein verbal orientiertes Testverfahren mit teils kulturspezifischen Inhalten ist, wurde erwartet, dass sich die verschiedenen Muttersprachler differenzieren. Die Befunde zeigen, dass die Muttersprachler-D (Deutsch als Muttersprache) signifikant bessere Testergebnisse aufweisen als die Muttersprachler-F (Deutsch als Fremdsprache).

Die Annahme, dass männliche Probanden höhere Testwerte erzielen als die weiblichen Probanden kann ebenfalls verfestigt werden. Eine Analyse der Mittelwertdifferenzen zeigt, dass signifikante Unterschiede in Abhängigkeit des Geschlechts bestehen.

Als nächstes wurde die Hypothese untersucht, ob das erreichte Bildungsniveau einen Einfluss auf die Wissensausprägung hat. Die Ergebnisse der Varianzanalyse haben deutlich gemacht, dass die Wissensunterschiede zwischen den Bildungsgruppen signifikant sind. Des Weiteren konnte anhand von Post-Hoc-Tests nachgewiesen werden, dass der Wissensumfang in Abhängigkeit des Bildungsniveaus ansteigt. Dementsprechend kann die Annahme bekräftigt werden. Hinsichtlich personalpolitischer Auswahlentscheidungen sind die Resultate insoweit von Relevanz, als dass man davon ausgehen kann, dass die Bildung der Probanden einen Erklärungsansatz für die Wissensdifferenzen liefert. Aufgrund dessen ergibt sich zudem ein erster Hinweis auf die Validität des Gesamttests. Obgleich die Untersuchungen zunächst aussagekräftige Ergebnisse liefern, wird gleichermaßen darauf hingewiesen, dass einzelne Substichprobengrößen sowie die Stichprobenauswahl nicht ausreichen, um allgemeingültige Aussagen bzw. Interpretationen zu rechtfertigen. Weiterführende Forschungen sollten in die Richtung gehen, eine größere und repräsentativere Stichprobe zu erfassen, um entsprechende alters-, geschlechts- und bildungsspezifische Normen zu erstellen. Zur Erklärung von anlage- und umweltbedingten Wissensdifferenzen sollten überdies auch soziologische, anthropologische oder neurophysiologische Ansätze herangezogen werden. Die ermittelten Befunde sind demzufolge als Indikatoren zu verstehen. Dennoch lassen sie die Konklusion zu, dass das Wissen nicht nur durch kognitive Fähigkeiten oder die Anlage geprägt wird. Stattdessen kann festgehalten werden, dass kontinuierliche Akkulturationseinflüsse einen bedeutenden Effekt auf den Wissenserwerb haben und letztendlich beide Faktoren das Leistungspotential einer Person determinieren. Unter Berücksichtigung dieser Anmerkung lässt sich abschließend konstatieren, dass das Wissen einer Person im Laufe der Zeit vor allem durch institutionelle Bildungseinrichtungen und andere kulturelle Umweltfaktoren nachhaltig beeinflusst wird und die Höhe der Wissensausprägung durch das jeweilige kulturelle Umfeld weitgehend erklärt werden kann. Daraus kann abgeleitet werden, dass die Ergebnisse von Wissenstests immer im Kontext der vielfältigen Einflussfaktoren betrachtet, interpretiert und beurteilt werden sollten, um zukünftige Leistungsmöglichkeiten von Berufseinsteigern besser prognostizieren zu können.

Zur Validität des START-W-Wissenstests:

Zur Überprüfung der Kriteriumsvalidität wurden die Übereinstimmungsvalidität sowie die prognostische Validität untersucht. Hinsichtlich der Übereinstimmungsvalidität wurden die einzelnen Wissensskalen mit dem Außenkriterium der Schulnoten korreliert. Die Korrelationen sind –bis auf die Kunstnote- annahmegemäß negativ und signifikant.

Die Koeffizienten zwischen dem Gesamtscore, den Skalenscores und den Schulnoten liegen im Bereich von  $r = -.174$  bis  $r = -.273$  ( $p < .01$ ). Die dritte Forschungshypothese (Zusammenhang zwischen numerischer Skala und Mathematiknote) weist konforme Ergebnisse auf. Gleiches gilt für die vierte Forschungsannahme. Lediglich die fünfte Forschungsannahme konnte nicht aufrecht erhalten werden. Mögliche Ursachen für die niedrigen Korrelationen wurden in Kapitel 11.2 diskutiert. Die Korrelationen zeigen zwar niedrige, aber dafür durchweg signifikante und negative Koeffizienten auf. Zusammenfassend lässt sich festhalten, dass die Übereinstimmungsvalidität anhand des Außenkriteriums der Schulnoten als ausreichend bewertet werden kann und einen eindeutigen Hinweis auf eine praktische Bewährung liefert. Außerdem verweisen Lienert und Raatz (1998) darauf, dass einzelne Testverfahren, die Teil einer Testbatterie sind, auch mit signifikanten und niedrigen Validitätskoeffizienten ( $< .30$ ) von Bedeutung sind. Hinsichtlich der praxisorientierten Verwendung des Testverfahrens kommt der prognostischen Validität eine erhöhte Relevanz zu. Hierzu wurden die Testwerte mit der Durchschnittsnote des ersten Ausbildungsjahres (Probahalbjahr) korreliert. Die Ergebnisse sind theoriekonform und zeigen Korrelationskoeffizienten bis zu  $r = -.609$  ( $p < .01$ ) auf. Die Resultate können als zufriedenstellend beurteilt werden und erlauben zudem einen Rückschluss auf den Ausbildungserfolg an den Schulen für Gesundheitsberufe. Auf der Basis der Befunde ist die Kriteriumsvalidität hinreichend.

Im Rahmen der Konstruktvalidität wurde zunächst die konvergente Validität mit dem Bochumer Wissenstest (Hossiep & Schulte, 2008) überprüft. Der BOWIT stellt ein geeignetes Instrument dar, weil das Testverfahren eine vergleichbare Theorieorientierung aufweist. Für den Zusammenhang zwischen dem Gesamtscore des START-W und dem BOWIT ergibt sich eine Korrelation von  $r = .673$  ( $p < .01$ ;  $r = .704$  nach Pearson). Substantielle Korrelationen werden auch zwischen den Inhaltsskalen und dem Gesamtscore des BOWIT berechnet. Die hochsignifikanten Zusammenhänge zeigen auf, dass beide Verfahren konvergente Konstrukte darstellen und das erworbene Wissen bzw. die kristallisierte Intelligenz partiell abbilden können. Weiterhin wurde die faktorielle Validität des START-W durch eine konfirmatorische Faktorenanalyse abgebildet. Das Strukturgleichungsmodell, welches mithilfe von LISREL erstellt wurde, macht deutlich, dass sowohl die intendierten Themenbereiche, die verbal, numerisch und figural kodierten Inhaltsskalen sowie der Faktor gc extrahiert werden können. Der CFI als auch der RMSEA und der SRMR sprechen für eine gute Passung des gesamten Modells. Folglich kann davon ausgegangen werden, dass das Modell die intendierte Struktur adäquat ab-

bildet. Darüber hinaus wurde der Zusammenhang zwischen der fluiden Intelligenz (gf) und der kristallisierten Intelligenz (gc) untersucht. Die Ergebnisse zeigen theoriekonforme Zusammenhänge zwischen gf und gc, die jedoch vom Generalitätsniveau des Wissens abhängen. Die leichter generalisierbaren inhaltsgebundenen Wissensvarianzen hängen stärker mit gf zusammen als die spezifischen Wissensvarianzen (gc'). Bei der Konstruktion von Wissenstests sollten diese komplexen Zusammenhänge berücksichtigt werden.

Der START-W kann hinsichtlich der aufgeführten Kriterien als hinreichend reliabel und valide bezeichnet werden. In Bezug auf die Konstruktvalidität wird eine überzeugende Bestätigung der dreifaktoriellen Struktur abgebildet, mit dem eine Annäherung an das komplexe, theoretische Konstrukt „Allgemeinwissen“ möglich erscheint. Für eine differenziertere Schätzung der Validität ist es jedoch auch sinnvoll, den Wissenstest im Rahmen der gesamten START-Testbatterie zu betrachten. Des Weiteren ist eine Erstellung von spezifischen Normen erstrebenswert. Insgesamt stellt der START-W ein wissenschaftlich fundiertes Verfahren dar, welches das Allgemeinwissen zuverlässig und ökonomisch erfassen und als Indikator für die Bildungs- bzw. Lernmotivation interpretiert werden kann.

Es wurde vielfach in der Literatur hervorgehoben, dass Testbatterien sowie das Konzept der fluiden und kristallisierten Intelligenz eine erhöhte Relevanz für die praxisorientierte Eignungsdiagnostik haben müssten. Die im Zuge der Arbeit untersuchten Forschungsfragen unterstreichen die Annahme, dass eine Integration von Wissenstests und Intelligenzstrukturtests einen Erkenntnisgewinn für die berufliche Kompetenzfeststellung bedeutet. Der START-W befähigt unter anderem dazu, verschiedene Bildungsgruppen voneinander zu trennen, erlaubt eine Einschätzung des zukünftigen Schul- oder Ausbildungserfolges und gewährleistet in Kombination mit der START-Testbatterie differenzielle Aussagen. Mit der Erfassung der kristallisierten Intelligenz stellt auch Nettelnstroth (2003) heraus, dass „wissensrelevante Merkmale (kristallisierte Intelligenz) generell besser zwischen Berufstätigen differenzieren, als Merkmale des schlussfolgernden Denken (fluide Intelligenz, *Reasoning*) (...) und damit eine größere Bedeutung für berufliche Aspekte und Prognosen aufweist“ (S.395). Die Forschung sollte demnach bemüht sein, weiterführende Validierungen an Außenkriterien, wie Schul-, Ausbildungs- und Berufserfolg durchzuführen. Vor dem Hintergrund der hohen psychometrischen Qualität und des breiten Anwendungsbereiches, ist der START-W ein innovatives Verfahren und eröffnet der Praxis zusätzliche Perspektiven.

## Literaturverzeichnis

- Ackerman, P., Bowen, K., Beier, M. & Kanfer, R. (2001). Determinants of individual differences and gender differences in knowledge. *Journal of Educational Psychology, 93*, 797-825.
- Ackerman, P. L. & Humphreys, L. G. (1991). Individual differences theory in industrial and organizational psychology. In M. D. Dunette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (S. 223-282). Palo Alto, CA: Consulting Psychologists Press.
- Ackerschott, H. (2004). Die neue DIN 33430- Qualität ist die Antwort auf Fragen des Marktes. In L. Hornke, & U. Winterfeld. *Eignungsbeurteilungen auf dem Prüfstand: DIN 33430 zur Qualitätssicherung* (S.3-8). Heidelberg: Spektrum Akademischer Verlag.
- Afifi, A. & Elashoff, R. (1966). Missing observations in multivariate statistics. Part I review of the literature. *Journal of the American Statistical Association, 61*, 595-604.
- Allik, J., Must, O. & Lynn, R. (1999). Sex differences in general intelligence in high school students: Some results from Estonia. *Personality and Individual Difference, 26*, 1137-1141.
- Althoff, K. (1971). Die Vorhersage des Berufserfolgs von Regierungsinspektoren und Revierförstern: Zwei Bewährungskontrollen. *Diagnostica, 17*, 145-149.
- Amelang, M. & Bartussek, D. (1990). *Differentielle Psychologie und Persönlichkeitsforschung* (3. überarbeitete und erweiterte Aufl.). Stuttgart: Kohlhammer.
- Amelang, M. & Bartussek, D. (1997). *Differentielle Psychologie und Persönlichkeitsforschung* (4. überarbeitete und erweiterte Aufl.). Stuttgart: Kohlhammer.
- Amelang, M. & Bartussek, D. (2001). *Differentielle Psychologie und Persönlichkeitsforschung* (5. überarbeitete und erweiterte Aufl.). Stuttgart: Kohlhammer.
- Amelang, M. & Schmidt-Atzert, L. (2006). *Psychologische Diagnostik und Intervention* (4. Überarbeitete und erweiterte Aufl.). Heidelberg: Springer.
- Amthauer, R. (1970). *Intelligenz-Struktur-Test. I-S-T 70*. Göttingen: Hogrefe.
- Amthauer, R., Brocke, B., Liepmann, D., & Beauducel, A. (1999). *Intelligenz-Struktur-Test 2000*. Göttingen: Hogrefe.
- Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (2001). *Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R)*. Göttingen: Hogrefe.
- Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.



- Anderson, J. R. (2001). *Kognitive Psychologie*. Heidelberg: Spektrum Akad. Verlag.
- Arbeitsgemeinschaft betriebliche Weiterbildungsforschung e.V. –ABWF (2006). Herausforderungen der Zukunft. Beschluss der Mitgliederversammlung der ABWF vom 14.09.2006. *Quem (Qualifikations- und Entwicklungs-Management) Bulletin*, H. 6, 1-3.
- Asendorpf, J.B. (2004). *Psychologie der Persönlichkeit* (3.Auflage). Berlin: Springer.
- Asendorpf, J. B. (2007). *Psychologie der Persönlichkeit*. Berlin: Springer.
- Backhaus, K., Erichson, R., Plinke, W. & Weiber, R. (2003). *Multivariate Analyseverfahren- Eine anwendungsorientierte Einführung* (10 Auflage). Berlin: Springer.
- Baltes, P. B. (1990). Entwicklungspsychologie der Lebensspanne: Theoretische Leitsätze. *Psychologische Rundschau*, 41, 1-24.
- Bartenwerfer, H. (1983). Allgemeine Leistungsdiagnostik. In K. J. Groffmann & L. Michel (Hrsg.), *Enzyklopädie der Psychologie Themenbereich B Methodologie und Methoden, Serie II Psychologische Diagnostik, Band 2 Intelligenz- und Leistungsdiagnostik* (S. 482-502). Göttingen: Hofgrefe.
- Barz, A., Hopbach, A. & Meyer-Guckel, V. (2004). *Schlüsselkompetenzen und Beschäftigungsfähigkeit*. Essen: Stifterverband für die deutsche Wissenschaft.
- Baumert, J. & Artelt, C. (2002). Bereichsübergreifende Perspektiven. In J. Baumert u.a. (Hrsg.), *PISA 2000 - Die Länder der Bundesrepublik Deutschland im Vergleich* (S. 219-236) Opladen: Leske & Budrich.
- Baumert, J. & Schümer, G. (2002). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb im nationalen Vergleich. In J. Baumert u.a. (Hrsg.), *PISA 2000 - Die Länder der Bundesrepublik Deutschland im Vergleich* (S. 156-202) Opladen: Leske & Budrich.
- Baumert, J., Bos, W. & Waterman, R. (1998). *TIMSS/III Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Beauducel, A. & Kersting, M. (2002). Fluid and crystallized intelligence and the Berlin Model of Intelligence Structure. *European Journal of Psychological Assessment*, 97-112.
- Beer, M., Spector, B., Lawrence, P.R., Mills, D.Q. & Walton, R.E. (1985). *Human Resource Management. A General Manager`s Perspective*. New York: Free Press.
- Beier, M.E. & Ackermann, P. L. (2001). Current-events knowledge in adults: An investigation of age, intelligence, and nonability detremnants. *Psycholgy and Aging*, 16, 615-628.

- Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Berry, D.C. (1987). The problem of implicit knowledge. *Expert Systems*, 4, 144- 151.
- 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, 387-410.
- Berufsverband Deutscher Psychologen (2002). *DIN 33430- Berufsbezogene Eignungsdiagnostik*. Berlin: Beuth Verlag
- 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, 387-410.
- Binet, A. & Simon, T. (1905). Méthodes nouvelles pour le diagnostique du niveau intellectuel des anormaux. *Année Psychologique*, 11, 191-244.
- Bischof-Köhler, D. (2006). *Von Natur aus anderes: Die Psychologie der Geschlechterunterschiede*. Stuttgart: Kohlhammer.
- Blaschke, D. (1987). Soziale Qualifikation am Arbeitsmarkt und im Beruf. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 19, 536-552.
- Blümelhuber, C. (2005). Informationsüberlastung. In D. Frey, L.v.Rosenstiel & C. Graf Hoyos (Hrsg.) , *Wirtschaftspsychologie* (S. 143-148). Weinheim: Beltz.
- Borg, I. & Lingoes, J.C. (1987). *Multidimensional similarity structure analysis*. New York: Springer.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler* (6. Auflage). Heidelberg: Springer.
- Bortz, J. & Döring, N. (2002). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (3.Auflage). Heidelberg: Springer.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4., überarbeitete Auflage). Heidelberg: Springer Medizin Verlag.
- Bos, W., Lankes, E.M, Prenzel, M., Schwippert, K., Walther G., & Valtin, R. (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Bouchard, T.J., Jr., Lykken, T.D., McGue, M., Segal, N.L., & Tellegan, A. (1990). Sources of human psychological differences: The Minnesota study of twins reared apart. *Science*, 250, 223- 228.
- Boudon, R. (1974). *Education, opportunity and social inequality*. New York: Wiley.

- Boyle, G.J., Stankow, L. & Cattell, R.B. (1995). Measurement and statistical models in the study of personality and intelligence. In D. H. Saklofske & M. Zeidner (Hrsg.), *International handbook of personality and intelligence* (S. 414-446). New York: Plenum Press.
- Brähler, E., Holling, H., Leutner, D. & Petermann, F. (2002). *Brickenkamp Handbuch psychologischer und pädagogischer Test* (3., überarbeitete und erweiterte Aufl.). Göttingen: Hogrefe.
- Brambring, M. (1983). Spezielle Eignungsdiagnostik. In K. J. Groffmann & L. Michel (Hrsg.), *Enzyklopädie der Psychologie Themenbereich B Methodologie und Methoden, Serie II Psychologische Diagnostik, Band 2 Intelligenz- und Leistungsdiagnostik* (S. 414-481). Göttingen: Hofgreffe.
- Brandstätter, H. (1979). Die Ermittlung personaler Eigenschaften kognitiver Art. In G. Reber (Hrsg.), *Personalinformationssysteme* (S. 74-95). Stuttgart: Poeschel.
- Brickenkamp, R. (Hrsg.) (1997). *Handbuch psychologischer und pädagogischer Tests*. Göttingen: Hofgreffe.
- Brickenkamp, R. (2002). *Test d2: Aufmerksamkeits-Belastungstest-Test* (9. Überarbeitete und neu normierte Aufl.). Göttingen: Hogrefe.
- Brocke, B. (2000). Intelligenz: Struktur und Prozess. In W. Sarges (Hrsg.), *Management-Diagnostik* (S.225- 240). Göttingen: Hogrefe.
- Brocke, B. & Beauducel, A. (2001). Intelligenz als Konstrukt. In E. Stern & J. Gutke (Hrsg.), *Perspektiven der Intelligenzforschung* (S.13-43). Lengerich: Pabst.
- Brody, E. B. & Brody, N. (1976). *Intelligence: Nature, Determinants and Consequences*. New York: Academic Press.
- Bühl, A (2006). *SPSS 14. Einführung in die moderne Datenanalyse* (10. Auflage).München: Pearson Studium.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- Bürgel, H. D. & Zeller, A. (1998). Forschung und Entwicklung als Wissenscenter. In H. D. Bürgel (Hrsg.), *Wissensmanagement. Schritte zum intelligenten Unternehmen* (S. 53-65). Berlin: Springer.
- Buck, P.N. (2005). *Knowledge management and intellectual capital: Establishing a field of practice*. Basingstoke, Hampshire u.a. : Palgrave Macmillan.
- Bundesanstalt für Arbeit (2002). Kompetenzen klären. Kompetenzfeststellung in der Benachteiligtenförderung. *Direkt: Fördern und qualifizieren, 15*, 1- 36 [Internet]. Verfügbar unter [http://www.ausbildungsvorbereitung.de/download/2430\\_pub\\_fb\\_direkt\\_nr15.pdf](http://www.ausbildungsvorbereitung.de/download/2430_pub_fb_direkt_nr15.pdf) [10.3.2009].

- Bundesministerium für Bildung und Forschung (1998). *Delphi-Befragung 1996/1998-Integrierter Abschlussbericht* [Internet]. Verfügbar unter: [www.bmbf.de](http://www.bmbf.de) [10.10.2010].
- Bundesministerium für Bildung und Forschung-BMBF (1999). *Bildungsbericht 1999* [Internet], (S. 1-275). Verfügbar unter: <http://www.bmbf.de/pub/bbb1999.pdf> [10.5.2010].
- Bundesministerium für Bildung und Forschung-BMBF (2001). *Aktionsprogramm „Lebensbegleitendes Lernen für alle“- Bonn* [Internet]. Verfügbar unter: [http://www.bmbf.de/pub/aktionsprogramm\\_lebensbegleitendes\\_lernen\\_fuer\\_alle.pdf](http://www.bmbf.de/pub/aktionsprogramm_lebensbegleitendes_lernen_fuer_alle.pdf) [7.2.2010].
- Bundesministerium für Bildung und Forschung-BMBF (2006). *Exzellenz in Bildung und Forschung- mehr Wachstum durch Innovationen. Leitlinien der Bildungs- und Forschungspolitik-Bonn* [Internet]. Verfügbar unter: <http://www.bmbf.de/press/1921.php> oder [http://www.bmbf.de/pub/mehr\\_wachstum\\_durch\\_innovation.pdf](http://www.bmbf.de/pub/mehr_wachstum_durch_innovation.pdf) [7.2.2010].
- Bundschuh, K. (1999). *Einführung in die sonderpädagogische Diagnostik*. München: E. Reinhardt.
- Bunk, G.P. (1981). Technologischer Wandel und antizipative Berufsbildung. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 77, 266-257.
- Bunk, G.P. (1982). *Einführung in die Arbeits-, Berufs- und Wirtschaftspädagogik*. Heidelberg: Quelle und Meyer.
- Bunk, G.P. (1991). Schlüsselqualifikationen- Ein Strukturelement beruflicher Bildung. *Bildung real- Veränderte Anforderungen und neue Chancen in der Arbeitswelt*, 35, 29-37.
- Bunk, G.P., Kaiser, . & Zedler, R. (1991). Schlüsselqualifikationen- Intention, Modifikation und Realisation in der beruflichen Aus- und Weiterbildung. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 24, 365-374.
- Campbell, D.T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod-matrix. *Psychological Bulletin*, 81, 81-105.
- Cattell, R.B. (1963). Theory of fluid and crystallized intelligence: a critical experiment. *Journal of Educational Psychology*, 54, 1-22.
- Cattell, R. B. (1971a). *Abilities: Their structure, growth and action*. Boston: Mifflin.
- Cattell, R.B. (1971b). The structure of intelligence in relation to the nature-nurture controversy. In R. Cancro (Hrsg.), *Intelligence* (S. 3-30). New York: Grune & Stratton.
- Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. Amsterdam: Elsevier Science Publishers e.V..

- Ceci, S.(1996). *On intelligence*. Cambridge, MA: University Press.
- Chi, M. , Glaser, R. & Rees, E. (1982). Expertise in problem solving. In R.J. Sternberg (Hrsg.), *Advances in the psychology of human intelligence* (S. 7-76). Hillsdale: Erlbaum.
- Chomsky, N. (1973). *Aspekte der Syntax-Theorie*. Frankfurt (Main): Suhrkamp.
- Conrad, W. (1983). Intelligenzdiagnostik. In K. J. Groffmann & L. Michel (Hrsg.), *Enzyklopädie der Psychologie Themenbereich B Methodologie und Methoden, Serie 2 Psychologische Diagnostik, Band 2 Intelligenz- und Leistungsdiagnostik* (S. 104-201). Göttingen: Hogrefe.
- Cronbach, L.J. (1951). Coefficient Alpha and the internal structure of tests. *Psychometrika*, 16,297-334.
- Dahrendorf, R. (1956). Industrielle Fertigkeiten und soziale Schichtung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 12, 540-568.
- Dany, F. & Torchy, V. (1994). Recruitment and selction in Europa: Policies, practices and methods. In C. Brewster & A. Hegewisch (Hrsg.), *Policy and Practice in European Human Resource Management: The Price Waterhouse Cranfield Survey*. (S. 68-88). London: Routeledge.
- Daumenlang, K. (1990). Intelligenztests. In W. Sarges (Hrsg.), *Management-Diagnostik*. Göttingen: Hogrefe.
- Daumenlang, K. (1995). Intelligenztests. In W. Sarges (Hrsg.), *Management-Diagnostik* (2. vollständig überarbeitete und erweiterte Aufl.). Göttingen: Hogrefe.
- Dick, M. & Wehner, T. (2002). Wissensmanagement zur Einführung: Bedeutung, Definition, Konzepte. In W. Lüthy, E. Voit & T. Wehner (Hrsg.), *Wissensmanagement – Praxis* (S. 7-27). Zürich: vdf. Hochschulverlag AG.
- Dochy, F. J. R. C. & Alexander, P. A. (1995) Mapping for prior knowledge. A framework discussion among researchers. *European Journal of Psychology of Education*, 10, 225-242.
- Drexel, I. (1997): Die bilans de compétence- Ein neues Instrument der Arbeits- und Bildungspolitik in Frankreich. In Arbeitsgemeinschaft Qualifikations-Entwicklungs-Management (Hrsg.), *Kompetenzentwicklung 1997: Berufliche Weiterbildung in der Transformation-Fakten und Visionen* (S.197-249). Münster: Waxmann.
- Dye, D.A., Reck, M. & Mc Daniel, M.A.(1993). The validity of job knowledge measures. *International Journal of Selection and Assessment*, 1, 153-157.

- Eaves, L.J., Hewitt, J.K., Meyer, J. & Neale, M.C. (1990). Approaches to the quantitative genetic modeling of development and age-related changes. In M.E. Hahn, J.K. Hewitt, N.D. Henderson & R.H. Benno (Hrsg.), *Development behavior genetics: Neural, biometrical, and evolutionary approaches* (S. 266- 280). New York: Oxford University Press.
- Eilles-Matthiessen, C., el Hage, N., Janssen, S. & Osterholz, A. (2002). *Schlüsselqualifikationen in Personalauswahl und Personalentwicklung. Ein Arbeitsbuch für die Praxis*. Bern: Huber.
- Ericsson, K.A., Krampe, R. & Tesch-Römer, C.. (1993). The role of deliberate practice in the aquisition of expert performance. *Psychological Review*, 100, 363-406.
- Erpenbeck, J. & Heyse, V. (2007). *Die Kompetenzbiographie: Wege der Kompetenzentwicklung*. Münster: Waxmann.
- Erpenbeck, J. & Rosenstil, L.v. (2003). Einführung. In J. Erpenbeck & L. Rosenstil (Hrsg.), *Handbuch Kompetenzmessung: Erkennen, verstehen, und bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis* (S. IX-XL). Stuttgart: Schäffer-Poeschel.
- Erpenbeck, J. & Rosenstil, L.v. (2007). Einführung. In J. Erpenbeck & L. Rosenstil (Hrsg.), *Handbuch Kompetenzmessung: Erkennen, verstehen, und bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis* (S. XVII-XLVI). Stuttgart: Schäffer-Poeschel.
- Ertel, S. & Schindler, U. (1969). Intelligenzleistungen und Testverhaltensdifferential. *Diagnostica*, 15, 103-117.
- Europäischer Rat (2000). *Schlussfolgerungen des Vorsitzes. Lissabon 23. Und 24. März 2000* [Internet]. Verfügbar unter: [http://www.consilium.europa.eu/ueDocs/cms\\_Data/docs/pressdata/de/ec/00100-r1.d0.htm](http://www.consilium.europa.eu/ueDocs/cms_Data/docs/pressdata/de/ec/00100-r1.d0.htm) [7.2.2010].
- Europäisches Parlament und Europäischer Rat (2006). *Empfehlung des Europäischen Parlaments und des Rates vom 18.12.2006 zu Schlüsselkompetenzen für lebensbegleitendes Lernen (2006/962/EG)*. In *Amtsblatt der Europäischen Union L 394 vom 30.12.2006* (S. 10-18) [Internet]. Verfügbar unter: [http://www.consilium.europa.eu/ueDocs/cms\\_Data/docs/pressdata/de/ec/00100-r1.d0.htm](http://www.consilium.europa.eu/ueDocs/cms_Data/docs/pressdata/de/ec/00100-r1.d0.htm) [7.2.2010].
- Etzel, S. & Küppers, A. (2000). *Pro facts*. Nürnberg: Pro facts Assessment und Training.
- Evans E., Schweingruber, H. & Stevenson, H. (2002). Gender differences in interests and knowledge acquisition: The United States, Taiwan an Japan. *Sex Roles*, 47, 153-167.
- Eysenck, H. J. (1979). *The structure and measurement of intelligence*. New York: Springer.

- Felser, G. (2005). Differentielle Wirtschaftspsychologie. In D. Frey, L. v. Rosenstiel & C. Graf Hoyos (Hrsg.), *Wirtschaftspsychologie* (S. 49-55). Weinheim: Beltz.
- Fischer, G. (1974). *Einführung in die Theorie psychologischer Tests*. Bern: Huber.
- Fisseni, H.J. (1997). *Lehrbuch der psychologischen Diagnostik* (2. Auflage). Göttingen: Hogrefe.
- Fisseni, H.J. (2004). *Lehrbuch der psychologischen Diagnostik* (3.Auflage). Göttingen: Hofgrete.
- Fisseni, H. J. & Fennekels, G. P. (1995). *Das Assessment Center*. Göttingen: Hogrefe.
- Flynn, J.R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychology Bulletin*, 101, 171- 191.
- Ford, M. E. & Tisak, M. S. (1983). A further search of social intelligence. *Journal of Educational Psychology*, 75 (2), 196-206.
- French, J. W., Ekstrom, R. B. & Price, L. A. (1963). *Manual and kit of reference tests for cognitive factors*. Princeton, NJ: Educational Testing Service.
- Frey, A. & Balzer, L. (2005). Der Beurteilungsbogen smk: Ein Messverfahren für die Diagnose von sozialen und methodischen Fähigkeitskonzepten. In A. Frey, R. S.Jäger & U. Renold (Hrsg.), *Kompetenzdiagnostik - Theorien und Methoden zur Erfassung und Bewertung von beruflichen Kompetenzen* (S. 31-57). Landau: Verlag Empirische Pädagogik.
- Frey, A., Jäger, R. S. & Renold, U. (2005). *Kompetenzdiagnostik - Theorien und Methoden zur Erfassung und Bewertung von beruflichen Kompetenzen*. Landau: Verlag Empirische Pädagogik.
- Fricke, E. & Fricke, W. (1976). Berufsbildung und Beschäftigungssystem-Eine empirische Analyse der Vermittlung und Verwendung von Qualifikationen in fünf Großbetrieben der Metallindustrie. In Minister für Wissenschaft und Forschung-Johannes Rau im Auftrag des Ministerpräsidenten Heinz Kühn (Hrsg.), *Forschungsbericht des Landes Nordrhein-Westfalen, Nr. 2537/ Wirtschafts- und Sozialwissenschaften* (S.1-690). Opladen: Westdeutscher Verlag.
- Frieling, E. & Sonntag, K. (1999). *Lehrbuch Arbeitspsychologie* (2., vollständig überarbeitete Auflage). Bern: Huber.
- Fürntratt, E. (1969). Differentieller-Wissens-Test DWT. Ein Test zur differenzierten Erfassung des Allgemeinwissens. Handanweisung. In A. O. Jäger (Hrsg.), *Differentieller Wissenstest D-W-T*. Göttingen: Hogrefe.
- Gage, N. L. & Berliner, D. C. (1986). *Pädagogische Psychologie*. Weinheim: Beltz.
- Gagné, F. & St Père, F. (2001). When IQ is controlled, does motivation still predict achievement? *Intelligence*, 30, 71-100.

- Galton, F. (1869). *Hereditary genius*. New York: Macmillan.
- Gebert, D. & Rosenstil, L.v. (2002). *Organisationspsychologie* (5.Aufl.). Stuttgart: Kohlhammer.
- Giesen, H., Gold, A., Hummer, A. & Jansen, R. (1986). *Prognose des Studienerfolgs. Ergebnisse aus Längsschnittuntersuchungen*. Frankfurt/Main: Universität Frankfurt, Institut für Pädagogische Psychologie.
- Gillardi, R., Holling, H. & Schmidt, J. U. (1983). Replikationsanalyse zum Cattellschen Intelligenzkonzept. *Psychologische Beiträge*, 25, 369-384.
- Gnahs, D. (2007). *Kompetenzen- Erwerb, Erfassung, Instrumente. Studentexte für Erwachsenenbildung*. Bielefeld: Bertelsmann Verlag.
- Göhs, N. & Dick, J. (2001). Testverfahren bei der Personalauswahl- Qualitätssuche im intransparenten Markt. *Personal- Zeitschrift für Human Resource Management*, 53, 46-48.
- Goff, M. & Ackerman, P. L. (1992). Personality-intelligence relations: Assessment of typical intellectual engagement. *Journal of Educational Psychology*, 84, 537-553.
- Gottlieb, G. (1991). Experimental canalization of behavioral development: Theory. *Development Psychology*, 27, 4- 13.
- Grabowski, S. (1990). *Fachübergreifende Qualifikationen in der CNC-Ausbildung aus der Sicht der betrieblichen Ausbilder*. Unveröff. Dipl.Arbeit, Ruhr-Universität, Bochum.
- Groffmann, K. J. (1983). Die Entwicklung der Intelligenzmessung. In K. J. Groffmann & L. Michel (Hrsg.), *Enzyklopädie der Psychologie Themenbereich B, Methodologie und Methoden, Serie II Psychologische Diagnostik, Band 2 Intelligenz- und Leistungsdiagnostik* (S.2-76). Göttingen: Hogrefe.
- Gruber, H. (1991). *Qualitative Aspekte von Expertise im Schach*. München: Ludwig-Maximilian-Universität, Dissertation.
- Gruber, H. & Mandel, H (1996). Das Entstehen von Expertise. In J. Hoffmann & W. Kintsch (Hrsg.), *Enzyklopädie der Psychologie, Serie Kognition, Band 7: Lernen*. (S. 583-615). Göttingen: Hogrefe.
- Guthke, J. & Beckmann, J. F. (2001). Intelligenz als Lernfähigkeit. In E. Stern & J.Guthke (Hrsg.), *Perspektiven der Intelligenzforschung* (S. 137- 163). Lengerich: Pabst.
- Guttman, L. & Levy, S. (1991). Two Structural Laws for Intelligence Tests. *Intelligence*, 15, 79-103.



- Häcker, H. (1998). *Committee to Develop Standards for Educational and Psychological Testing : Standards für pädagogisches und psychologisches Testen* (ausgearbeitet vom "Committee to Develop Standards for Educational and Psychological Testing" of the American Educational Research Association). Bern: Huber.
- Häcker, H., Leutner, D. & Amelang, M. (1998). *Standards für pädagogisches und psychologisches Testen* (Suppl. der Diagnostica und der Zeitschrift für Differentielle und Diagnostische Psychologie). Bern: Huber.
- Hakstian, R. & Cattell, R.B. (1978): Higher stratum ability factors on a basis of twenty primary abilities. *Journal of Educational Psychology*, 70, 657-669.
- Haller, G. (2003). Südtiroler Bildungspass und Kom(petenz)Pass. In Ministerium für Bildung, Kultur und Wissenschaft des Saarlandes (Hrsg.), *Bildungspässe-Machbarkeit und Gestaltungsmöglichkeiten. Tagungsband zum internationalen Fachkongress am 21./22.01.2003 in Saarbrücken* (S.77-80) [Internet]. Verfügbar unter: <http://www.bildungsbuero-koeln.de/wp-content/uploads/2007/06/bildungspaesse.pdf> [7.2.2010].
- Handy, L., Burham, K., Panter, S. & Winhard, A.(1989). Beyond the Personel Function. The strategic Management of Human Resources. *Journal of European Industrial Training*, 13 (1), 13- 18
- Hany, E.A. (2001). Die Vererbung der Intelligenz unter der Entwicklungsperspektive. In E. Stern & J. Guthke(Hrsg.), *Perspektiven der Intelligenzforschung* (S.69- 89). Lengerich: Pabst.
- Hartig, J. & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schwizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 127-143). Heidelberg: Springer.
- Heinrich-Böll Stiftung (2009). *Auf dem Weg zur Wissens- und Kultargesellschaft* [Internet]. Verfügbar unter: <http://www.boell.de/bildungskultur/bildung-kultur.html> [20.9.2009].
- Hell,B, Schuler, H. Boramir, I. & Schaar, H. (2006). Verwendung und Einschätzung von Verfahren der internen Personalauswahl und Personalentwicklung im 10 Jahres-Vergleich. *Zeitschrift für Personalforschung*, 20, 58-78.
- Helmke, A., Hosenfeld, I., Schrader, F.W. & Wagner, W. (2002). Sozialer und sprachlicher Hintergrund. In A. Helmke & R.S. Jäger (Hrsg.). *Das Projekt MARKUS. Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext* (S. 71-153). Landau: Empirische Pädagogik.
- Helmke, A. & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie: Themenbereich D, Praxisgebiete: Serie 1 Pädagogische Psychologie, Band 3 Psychologie des Unterrichts und der Schule*. Göttingen: Hogrefe.

- Hewitt, J.K. (1990). Changes in genetic control during learning, development, and aging. In M.E. Hahn, J.K. Hewitt, N.D. Henderson & R.H. Benno (Hrsg.), *Development behavior genetics: Neural, biometrical and evolutionary approaches* (S.217-235). New York: Oxford University Press.
- Heyse, H. & Kersting, M. (2004). Anforderungen an den Prozess der Eignungsbeurteilung. In L. Hornke & U. Winterfeld. *Eignungsbeurteilungen auf dem Prüfstand: DIN 33430 zur Qualitätssicherung* (S.29-41). Heidelberg: Spektrum Akademischer Verlag
- Höft, S. (2006). Erfolgsüberprüfung personalpsychologischer Arbeit. In H. Schuler (Hrsg.), *Lehrbuch der Personalpsychologie* (S. 761-796). Göttingen: Hogrefe.
- Hogrefe TestSystem. (2001). In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren* (S. 583-590). Lengerich: Pabst Science Publishers.
- Holling, H. & Liepmann, D. (2007). Personalentwicklung. In H. Schuler (Hrsg.), *Lehrbuch der Organisationspsychologie* (S.345-383). Bern: Huber.
- Holling, H., Preckel, F. & Vock, M. (2004). *Intelligenzdiagnostik*. Göttingen Hogrefe.
- Holyoak, K.J. (1991). Symbolic connectionism. Toward third generation theories of expertise. In K.A. Ericsson & J. Smith (Hrsg.), *Toward a general theory of expertise* (S. 301-335). Cambridge: University Press.
- Horn, J. L. (1988). Thinking about human abilities. In J. R. Nesselroade & R. B. Cattell (Hrsg.), *Handbook of multivariate experimental psychology* (S. 645- 681). New York: Plenum.
- Horn, J.L. & Cattell (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology*, 57, 253- 270.
- Hornke, L. & Winterfeld, U. (2004). *Eignungsbeurteilungen auf dem Prüfstand: DIN 33430 zur Qualitätssicherung*. Heidelberg: Spektrum Akademischer Verlag
- Hossiep, R. (1995). *Berufseignungsdiagnostische Entscheidungen*. Göttingen: Hogrefe.
- Hossiep, R. (1996). Psychologische Test- Die vernachlässigte Dimension in Assessment Centern. In W. Sarges (Hrsg.), *Weiterentwicklung der Assessment Center- Methode* (S.53-67). Göttingen: Hogrefe.
- Hossiep, R. & Paschen, M. (1998). *Das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP)*. Göttingen: Hogrefe.
- Hossiep, R. & Paschen, M. (2003). *Das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP)*. Göttingen: Hogrefe.
- Hossiep, R. & Schulte, M. (2008). *BOWIT- Bochumer Wissenstest-Manual*. Göttingen: Hogrefe.

- Hülshager, U.E., Maier, G.W. & Stumpp (2007). Validity of General Metal Ability for the Prediction of Job Performance and Training Success in Germany: A metaanalysis. *International Journal of Selection and Assessment*, 15, 1-18.
- Humphreys, L. G. (1979). The construct of general intelligence. *Intelligence*, 3, 105-120.
- Hunter, J.E. & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulltin*, 96, 72- 98.
- Hunter, J.E. & Schmidt, F.L. (1989). Meta-Analysis: Facts and theories. In M. Smith & I.T. Robertson (Hrsg.), *Advances in selection and assessment* (S.203-216). New York: Wiley.
- Hyde, J. S. & Linn, M. C. (1988). Gender differnces in verbal ability: A meta-analysis. *Psychological Bulletin*, 104, 53- 69.
- Hyde, J. S., Fennema, E. & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107, 139-155.
- Industrie- und Handelskammer Berlin (Hrsg.) (2007). *IHK Umfrage: Unternehmen kritisieren Bildungsniveau* [Internet]. Verfügbar unter: [http://www.berlin.ihk24.de/servicemarken/presse/presseinfo/pressemeldungen/IHK\\_Umfrage\\_Bildungsniveau\\_kritisiert.jsp](http://www.berlin.ihk24.de/servicemarken/presse/presseinfo/pressemeldungen/IHK_Umfrage_Bildungsniveau_kritisiert.jsp) [09.05.2009] oder <http://www.ihk-berlin.de/linkableblob/824802/data/IHK-Jahresbericht-2007data.pdf> [14.12.2010].
- Industrie- und Handelskammer Berlin (Hrsg.) (2010). *IHK Ausbildungsumfrage* [Internet]. Verfügbar unter: [http://www.ihk-berlin.de/linkableblob/810602/data/Ausbildungsumfrage\\_2010\\_-\\_Bundesergebnisse-data.pdf](http://www.ihk-berlin.de/linkableblob/810602/data/Ausbildungsumfrage_2010_-_Bundesergebnisse-data.pdf) [14.12.2010].
- Jäger, A. O. (1967). *Dimensionen der Intelligenz*. Göttingen: Hogrefe.
- Jäger, A.O.(1970). *Dimensionen der Intelligenz*. Göttingen: Hogrefe.
- Jäger, A. O. (1982). Mehrmodale Klassifikation von Intelligenzleistungen: Experimentell kontrollierte Weiterentwicklung eines deskriptiven Intelligenzstrukturmodells. *Diagnostica*, 23, 195-225.
- Jäger, A. O. (1984). Intelligenzstrukturforschung: Konkurrierende Modelle, neue Entwicklungen, Perspektiven. *Psychologische Rundschau*, 35 (1), 21-35.
- Jäger, A. O. & Fürntratt, E. (1968). *DWT-Differentieller Wissenstest*. Göttingen: Hofgreffe.
- Jäger, A. O., Süß, H.-M. & Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test. BIS-Test, Form 4* (vergriffen). Göttingen: Hogrefe.

- Jenkins, J.J. & Paterson, D.G. (1961). *Studies in individual differences. The search for intelligence*. New York: Appleton- Century Crofts.
- Jensen, A. R. (1980). *Bias in mental testing*. London: Methuen.
- Jensen, A.R.(1993). Why is reaction time correlated with psychometric g? *Current Directions in Psychological Science*, 2, 53-56.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport: Praeger.
- Jones, M.P. (1996). Indicator and stratification methods for explanatory analysis in multiple linear regression. *Journal of the American Statistical Association*, 91, 222-230.
- Jöreskog, K.G. & Sörbom, D. (1993). *LISREL 8-Users reference guide*. Chicago, IL: Scientific Software International.
- Jung, H. (2000). *Persönlichkeitspsychologie: Instrument der Mitarbeiterführung- mit Persönlichkeitstest* (2. erw. Aufl.). Oldenbourg: Wissenschaftlicher Verlag.
- Jungkunz, D. & Bodinet, K. (1989). Korrelative Bedeutung von Testergebnissen, schulischer Vorbildung, Berufsschulnoten und Fähigkeitseinschätzungen für Berufsabschlussnoten. *Zeitschrift für Pädagogik*, 35, 711-730.
- Kanning, U.P. (2004). *Standards der Personaldiagnostik*. Göttingen: Hogrefe.
- Kemmet, C., Linke, H. & Wolf, A. (1982). *Studium und Berufschancen- Eine Analyse der Wirtschaft zur Situation der Hochschulabsolventen in den Wirtschafts- und Ingenieurwissenschaften*. Herford: Maximilian.
- Kersting, M (1999a). *Diagnostik und Personalauswahl mit computergestützten Problemlöseszenarien? Zur Kriteriumsvalidität von Problemlöseszenarien und Intelligenztests*. Göttingen: Hogrefe.
- Kersting, M. (1999b). Personalauswahl mit dem BIS-r-DGP Test: Evaluation der prognostischen Validität. *DGP-Informationen*, 56, 43-60.
- Kersting, M. (2006). *"DIN Screen" : Leitfaden zur Kontrolle und Optimierung von Verfahren und deren Einsatz bei beruflichen Eignungsbeurteilungen*. Berlin: Papst.
- Kersting, M (2000). Instrumente der Arbeits- und Organisationspsychologie: Rezension des Intelligenz-Struktur-Tests-2000 von R. Amthauer, B. Brocke, D. Liepmann, A. Beauducel. *Zeitschrift für Arbeits- und Organisationspsychologie*, 44, 96-101.
- Kersting, M. (2004). Kosten und Nutzen beruflicher Eignungsbeurteilungen. In L. Hornke, & U. Winterfeld. *Eignungsbeurteilungen auf dem Prüfstand: DIN 33430 zur Qualitätssicherung* (S.55-81). Heidelberg: Spektrum Akademischer Verlag
- Klix, F. (1988). Gedächtnis und Wissen. In H. Mandl & H. Spada (Hrsg.), *Wissenspsychologie* (S.19-55). München: Psychologie Verlags Union.

- Klix, F. & Spada, H. (1998). Einführung. In F.Klix & H. Spada (Hrsg.), *Enzyklopädie der Psychologie Themenbereich C Praxisgebiete, Serie II Kognition, Band 6 Wissen* (S.1- 13). Göttingen: Hogrefe.
- Kluwe, R. (1990). Wissen. In W. Sarges (Hrsg.), *Management-Diagnostik* (S. 174-181). Göttingen: Hofgreffe.
- Kluwe, R. H. (2000). Wissen. In W. Sarges (Hrsg.), *Management-Diagnostik* (3. Aufl., S. 218-225). Göttingen. Hogrefe.
- Kommission der Europäischen Gemeinschaften (2001). *Mitteilung der Kommission: Einen europäischen Raum des lebenslangen Lernens schaffen- KOM (2001) 678 endgültig- Brüssel (21.11.01)* [Internet]. Verfügbar unter: [http://eur-lex.europa.eu/LexUriServ/site/de/com/2001/com2001\\_0678de01.pdf](http://eur-lex.europa.eu/LexUriServ/site/de/com/2001/com2001_0678de01.pdf) [7.02.2010].
- Kommission der Europäischen Gemeinschaften (2006). *Vorschlag für eine Empfehlung des Europäischen Parlaments und des Rates zur Einrichtung eines Europäischen Qualifikationsrahmens für Lebenslanges Lernen- KOM (2006) 479-Brüssel* [Internet]. Verfügbar unter: [http://ec.europa.eu/education/policies/educ/eqf/com\\_2006\\_0479\\_de.pdf](http://ec.europa.eu/education/policies/educ/eqf/com_2006_0479_de.pdf) [7.2.2010].
- Kreuzig, H. & Enzenberger, P. (2000). *Manage!* Hamburg: MD – Gesellschaft für Management-Diagnostik.
- Krogh, G. & Köhne M. (1998). Der Wissenstransfer in Unternehmen. Phasen des Wissenstransfers und wichtige Einflussfaktoren. *Die Unternehmung*, 52, 235- 252.
- Krüger, H. (1988). *Organisation und extra-funktionale Qualifikationen*. Fankfurt (Main): Lang.
- Kruse, A. & Rudinger, G. (1997). Lernen und Leistung im Erwachsenenalter. In F. E. Weinert & H. Mandl (Hrsg.), *Enzyklopädie der Psychologie Themenbereich D Praxisgebiete, Serie I Pädagogische Psychologie, Band 4 Psychologie der Erwachsenenbildung* (S. 48-85). Göttingen: Hofgreffe.
- Kultusministerkonferenz KMK (1991). Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Hrsg.), *Rahmenvereinbarung über die Berufsschule*. Bonn: KMK.
- Landsberg, G. von (1986). Qualifikation und Berufserfolg. In H.J. Bödenhöfer, F. Offner, M. Kaiser, M. Otto & G. von Landsberg (Hrsg.), *Ingenieure in der Wirtschaft (Beiträge zur Gesellschafts- und Bildungspolitik des Instituts der deutschen Wirtschaft* (S.6-32). Köln: Deutscher Instituts-Verlag.
- Lane, R. E. (1966). The decline of politics and ideology in a knowledgeable society. *American Sociological Review*, 21, 649-662.

- Lang-von Wins, T. (2003). Die Kompetenzhaltigkeit von Methoden moderner psychologischer Diagnostik-, Personalauswahl- und Arbeitsanalyseverfahren sowie aktueller Management-Diagnostik-Ansätze. In J. Erpenbeck & L. v. Rosenstiel (Hrsg.), *Handbuch Kompetenzmessung. Erkennen, verstehen und bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis* (S. 585-618). Stuttgart: Schäffer-Poeschel.
- Laur-Ernst, U. (1990). Schlüsselqualifikationen- Innovative Ansätze in den neugeordneten Berufen und ihre Konsequenzen für Lernen. In L. Reetz & Th. Reitmann (Hrsg.), *Schlüsselqualifikationen. Dokumentation des Symposiums in Hamburg „Schlüsselqualifikationen- Fachwissen in der Krise?“* (S.36-55). Hamburg: Feldhaus.
- Lehmann, R. Peek, R., & Gänsfuß, R. (1997). *Aspekte der Lernausgangslage von Schülerinnen und Schüler der 5. Klassen an Hamburger Schulen (LAU 5)*. Hamburg: Behörde für Schule, Jugend und Berufsbildung.
- Lehrl, S. (1995). *Mehrfachwahl-Wortschatz-Intelligenztest (MWT-B)* (3. Auflage). Göttingen: Hogrefe.
- Lewin, K. (1963). *Die Feldtheorie in den Sozialwissenschaften*. Bern: Huber.
- Levy, J. (1976). Cerebral lateralization and spatial ability. *Behavior Genetics*, 6, 171-188.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Psychologie Verlags Union.
- Liepmann, D. (1993). Betriebliche Qualifizierung: Notwendigkeiten, Möglichkeiten, Perspektiven. In D. Liepmann (Hrsg.), *Qualifizierungsmaßnahmen als Konzepte der Personalentwicklung* (S. 11-26). Frankfurt/Main: Lang.
- Liepmann, D. & v. Gilardi, R. (1993). Zum Konzept extrafunktionaler Qualifikationen. In D. Liepmann (Hrsg.), *Qualifizierungsmaßnahmen als Konzepte der Personalentwicklung* (S. 89-102). Frankfurt/Main: Lang.
- Liepmann, D. , Beauducel, Brocke, B. & Amthauer, R. (2007). *Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R)*. Göttingen: Hogrefe.
- Lindenberger, U. (2000). Intellektuelle Entwicklung über die Lebensspanne: Überblick und ausgewählte Forschungsbrennpunkte. *Psychologische Rundschau*, 3, 135-145.
- Linn, M. C. & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: a meta-analysis. *Cild Development*, 56 (6), 1479-1498.
- Löhlin, J.C. (2000). Group differences in intelligence. In R.J. Sternberg (Hrsg.), *Handbook of intelligence* (S.176-196). Cambridge: Cambridge University Press.
- Lynn, R. (1994). Sex differences in intelligence and brain size: A paradox resolved. *Personality and Individual Differences*, 17 (2), 257-271.

- Lynn, R. & Irwing, P. (2002). Sex differences in general knowledge, semantic memory and reasoning ability. *British Journal of Psychology*, 93, 545-556.
- Lynn, R., Irwing, P. & Cammock, T. (2002). Sex differences in general knowledge. *Intelligence*, 30, 27-40.
- Lynn, R., Hampson, S. & Lee, M. (1988). The intelligence of chinese children in Hong Kong. *School Psychology International*, 9, 29-32.
- Lynn, R., Wilberg, S. & Margraf-Stiksrud, J. (2004). Sex differences in general knowledge in german high school students. *Personality and Individual Differences*, 37, 1643-1650.
- Mack, W. (1999). Intelligenz und Wissen. In H. Gruber, W. Mack & A. Ziegler (Hrsg.) , *Wissen und Denken. Beiträge aus Problemlösepsychologie und Wissenspsychologie* (S.119-151). Wiesbaden: DUV, Dt. Univ.-Verlag.
- Mandl, H., Friedrich, H.F. & Hron, A. (1988). Theoretische Ansätze zum Wissenserwerb. In H. Mandl & H. Spada (Hrsg.), *Wissenspsychologie* (S.123-160). München: Psychologie Verlag- Union.
- Mandl, H. & Spada, H. (1988). *Wissenspsychologie*. München: Psychologie Verlags Union.
- McClelland, D.C. (1973). Testing for competence rather than for „intelligence“. *American Psychologist*, 28, 1-14.
- Mehlhorn, G. & Mehlhorn H.G. (1981). *Intelligenz: Zur Erforschung und Entwicklung geistiger Fähigkeiten*. Berlin: VEB Deutscher Verlag der Wissenschaften.
- Mertens, D. (1974). Schlüsselqualifikationen. Thesen zur Schulung für eine moderne Gesellschaft. *Mitteilungen aus Arbeitsmarkt- und Berufsforschung*, 7, 36-43.
- Michel, L. & Conrad, W. (1982). Theoretische Grundlagen psychometrischer Tests. In K.J. Groffmann & L. Michel (Hrsg.), *Enzyklopädie der Psychologie, Themenbereich B: Methodologie und Methoden, Serie II Psychologische Diagnostik Band 1 Grundlagen psychologischer Diagnostik* (S.11-29). Göttingen: Hogrefe.
- Möbus, C. & Schröder, O. (1998). Zur Modellierung des Wissenserwerbs als deduktive und induktive Wissensveränderung. In F. Klix, & H. Spada (Hrsg.), *Wissen. Enzyklopädie der Psychologie Themenbereich C Praxisgebiete, Serie II Kognition Band 6 Wissen* (S.403- 450). Göttingen: Hogrefe.
- Müllerschön, A. (2005). *Bewerber professionell auswählen: Handbuch für Personalverantwortliche*. Weinheim: Beltz.
- Neidhardt-Wilberg, S. (2005). Weibliches und männliches Allgemeinwissen? In S. R. Schilling, J. R. Sparfeldt & C. Pruisken (Hrsg.), *Aktuelle Aspekte pädagogisch-psychologischer Forschung* (S. 145-159). Münster: Waxmann.

- Neisser, U. (1976). General, academic, and artificial intelligence. In L. Resnick (Hrsg.), *The nature of intelligence* (S. 135-147). Hillsdale, NJ: Erlbaum.
- Nettelstroth, W. (2003). *Intelligenz im Rahmen der beruflichen Tätigkeit*. Unveröffentlichte Dissertation, Freie Universität Berlin.
- Nettelstroth, W. (2004). Intelligenz und Berufserfolg. Eine Validierung der fluiden und kristallisierten Intelligenz an subjektiven und objektiven Berufserfolgskriterien. In W. Bungard, B. Koop & C. Liebig (Hrsg.), *Psychologie und Wirtschaft leben. Aktuelle Themen der Wirtschaftspsychologie in Forschung und Praxis* (S.415-420). München: Rainer Hampp.
- Neubauer, A. C. (1995). *Intelligenz und Geschwindigkeit der Informationsverarbeitung*. Wien: Springer.
- North, K. (2005). *Wissensorientierte Unternehmensführung. Wertschöpfung durch Wissen*. Wiesbaden: Gabler.
- Nyborg, H. (1990). Sex hormones, brain development, and spatio-perceptual strategies in Turner syndrome. In D.B. Berch & B.G. Bender (Hrsg.), *Sex chromosomes abnormalities and human behavior* (S. 100- 128). Washington: Am. Ass. For the Advancement of Science.
- Nyborg, H. (2003). *The scientific study of general intelligence: Tribute to Arthur R. Jensen*. Amsterdam: Pergamon Press.
- OECD-Organisation for Economic Co-operation and Development (2003). *Definition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo). Summary of the final report „Key Competencies for a successful Life and a Well-Functioning Society“* [Internet]. Verfügbar unter: [www.oecd.org/dataoecd/47/61/35070367.pdf](http://www.oecd.org/dataoecd/47/61/35070367.pdf) [2.2.2010].
- OECD-Organisation for Economic Co-operation and Development (2005). *Programme for international Assessment of Adult Competencies* [Internet]. Verfügbar unter: [http://www.oecd.org/document/35/0,3343,en\\_2649\\_201185\\_40277475\\_1\\_1\\_1\\_1\\_00.html](http://www.oecd.org/document/35/0,3343,en_2649_201185_40277475_1_1_1_1_00.html) oder [http://www.oecd.org/document/7/0,3343,en\\_2649\\_201185\\_44378247\\_1\\_1\\_1\\_1\\_00.html](http://www.oecd.org/document/7/0,3343,en_2649_201185_44378247_1_1_1_1_00.html) [7.2.2010].
- Oelke, U. (1998). *Ausbildungsrichtlinie für die staatlich anerkannten Kranken- und Kinderkrankenpflegeschulen* (Anpassung Hundenborn, G. & Kühn, C., 2003; Deutsches Institut für angewandte Pflegeforschung e.V.). Im Auftrag des Ministeriums für Gesundheit, Soziales, Frauen und Familien des Landes Nordrhein-Westfalen (S. 1-140).
- Oerter, R. (1988). Wissen und Kultur. In H. Mandl & H. Spada (Hrsg.), *Wissenspsychologie* (S.333-356). München: Psychologie Verlags Union.



- Offe, C. (1970). *Leistungsprinzip und industrielle Arbeit*. Frankfurt/Main: Europäische Verlagsanstalt.
- Ones, D. S., Viswesvaran, C. & Dilchert, S. (2004). Cognitive ability in selection decisions. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 431-468). Thousand Oaks, CA: Sage Publications.
- Opwis, K. (1988). Produktionssysteme. In H. Mandl & H. Spada (Hrsg.), *Wissenspsychologie* (S.74-99). München: Psychologie Verlags Union.
- Orth, G. (2006). Bestandsaufnahme zu den Kompetenzen deutscher Schulkinder. *Schulverwaltung NRW* (6).
- Pawlow, I.P. (1953). *Die gemeinsamen Typen der höheren Nerventätigkeit der Tiere und des Menschen*, Sämtliche Werke, Band 3. Berlin: Akademischer Verlag.
- Picot, A. & Scheuble, S. (2000). Die Rolle des Wissensmanagements in erfolgreichen Unternehmen. In H. Mandl & G. Reinmann- Rothmeier (Hrsg.), *Wissensmanagement* (S.19-37). München: Oldenburg.
- PISA- Program for international Student Association (2000). *Ergebnisse und Berichte*. In OECD- Organisation for economic co-operation and development (Hrsg.) [Internet]. Verfügbar unter: <http://www.mpib-berlin.mpg.de/pisa/> [10.10.2007].
- PISA- Program for international Student Association (2006). *Ergebnisse und Berichte*. In OECD- Organisation for economic co-operation and development (Hrsg.) [Internet]. Verfügbar unter: [http://www.oecd.org/document/37/0,3343,de\\_34968570\\_34968855\\_39066085\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/37/0,3343,de_34968570_34968855_39066085_1_1_1_1,00.html) oder <http://www.bmbf.de/de/6624.php> [10.10.2009].
- PISA- Program for international Student Association (2007). *PISA 2006 - Schulleistungen im internationalen Vergleich* (1.Aufl.). Bielefeld: Bertelsmann Verlag.
- Polanyi, M. (1985). *Implizites Wissen*. Frankfurt am Main: Suhrkamp.
- Prenzel, M., Carstensen, C.H., Rost, J. & Senkbeil, M. (2002). Naturwissenschaftliche Grundbildung im Ländervergleich. In J. Baumert u.a. (Hrsg.), *PISA 2000 - Die Länder der Bundesrepublik Deutschland im Vergleich* (S. 129-158) Opladen: Leske & Budrich.
- Price, A. (2004). *Human Resource Management in a business context*. London: Thomson.
- Probst, R., Raub, S. & Romhardt, K. (1998). *Wissen Managen. Wie Unternehmen ihre wertvollste Ressource optimal nutzen*. Wiesbaden: Gabler.
- Putz-Osterloh, W. (1981). *Problemlöseprozesse und Intelligenztestleistung*. Bern: Huber.

- Putz-Osterloh, W. (1988). Wissen und Problemlösen. In H. Mandl, H. Spada (Hrsg.), *Wissenspsychologie* (S.247- 263). München: Psychologie Verlag- Union.
- Putz-Osterloh, W. & Lemme, M. (1987). Knowledge and its intelligent application to problem solving. *The German Journal of Psychology*, 11, 286-303.
- Raven, J.. (1971). *The Advanced and Standard Progressive Matrices*. London: Lewis & Co.
- Reetz, L. (1989). Das Konzept der Schlüsselqualifikationen in der Berufsbildung (Teil 1). *Berufsbildung in Wissenschaft und Praxis*, 18 (5), 3-10.
- Reetz, L. (1990). Zur Bedeutung der Schlüsselqualifikationen in der Berufsbildung. In L. Reetz & Th. Reitmann (Hrsg.), *Schlüsselqualifikationen. Dokumentation des Symposiums in Hamburg „Schlüsselqualifikationen-Fachwissen in der Krise?“* (S. 16-35). Hamburg: Feldhaus.
- Rehäuser, J. & Krcmar, H. (1996). Wissensmanagement im Unternehmen. In G. Schreyögg & P. Conrad (Hrsg.), *Managementforschung 6: Wissensmanagement* (S. 1-40). Berlin: De Gruyter.
- Reinmann, P (1988). Novizen- und Expertenwissen. In F. Klix & H. Spada (Hrsg.), *Enzyklopädie der Psychologie: Themenbereich C, Theorie und Forschung: Ser. II, Kognition; Bd. 6: Novizen- und Expertenwissen* (S. 335-367). Göttingen: Hogrefe.
- Reinmann-Rothmeier, G. & Mandl, H. (1998). Wissensvermittlung: Ansätze zur Förderung des Wissenserwerbs. In F. Klix, H. Spada (Hrsg.), *Enzyklopädie der Psychologie: Themenbereich C Praxisgebiete, Serie II Kognition Band 6 Wissen* (S. 457-491) Göttingen: Hogrefe.
- Renold, U., Frey, A. & Balzer, L. (2002). Reform der kaufmännischen Grundbildung: Engineering von Innovationen in einem komplexen Berufsfeld. *Die Volkswirtschaft*, 7, 52-57.
- Renkl, A. & Stern, E. (1994). Die Bedeutung von kognitiven Eingangsvoraussetzungen und schulischen Lerngelegenheiten für das Lösen von einfachen und komplexen Testaufgaben. *Zeitschrift für Pädagogische Psychologie*, 8, 27-29.
- Reusser, K. (1998). Denkstrukturen und Wissenserwerb in der Ontogenese. In F. Klix & H. Spada (Hrsg.), *Enzyklopädie der Psychologie, Themenbereich C: Theorie und Forschung: Ser. II. Kognition; Bd. 6: Wissen* (S. 115-166). Göttingen: Verlag für Psychologie, Dr. J. C. Hogrefe.
- Ringelband, J. O. & Birkhan, G. (2000). Rückmeldung der Eignungsbeurteilung an den Kandidaten und diskursive Abstimmung. In W. Sarges (Hrsg.), *Management-Diagnostik* (3., unveränderte Aufl., S.796-802). Göttingen: Hogrefe.
- Rohracher, H. (1965). *Einführung in die Psychologie* (9. Aufl.). Wien: Urban & Schwarzenberg.

- Rolfhus, E. L. & Ackerman P. L. (1999). Assessing individual differences in knowledge. Knowledge, intelligence, and related traits. *Journal of Educational Psychology*, 91 (3), 511-526.
- Rost, J. (2004). *Lehrbuch. Testtheorie-Testkonstruktion* (2. Überarbeitete und erweiterte Aufl.). Bern: Verlag Hans Huber.
- Roth, E. (1966). *Pädagogische Anthropologie* (Bd. 1). Hannover: Schroedel.
- Roth, E. (1971). *Pädagogische Anthropologie* (Bd. 2). Hannover: Schroedel.
- Roth, E., Oswald, W.D. & Daumenlang, K. (1980). *Intelligenz: Aspekte, Probleme, Perspektiven* (4.Aufl.). Stuttgart: Kohlhammer.
- Ryan, A. M., Farland, L., Baron, H. & Page, R. (1999). An international look at selection practices.: Nation and culture as explanations for variability in practices. *Personnel Psychology*, 52, 359-391.
- Sackmann, S. & Elbe, M. (2000). Tendenzen und Ergebnisse empirischer Personalforschung der 90er Jahre in West-Deutschland. *Zeitschrift für Personalforschung*, 14, 131-157.
- Salgado, J. F. (1999). Personnel selection methods. *International Review of Industrial and Organizational Psychology*, 14, 1-54.
- Salgado, J. F., Anderson, N., Mosocso, S., Bertua, C., de Fruyt, F. & Rolland, J. P. (2003). A meta-analytic study of general mental ability validity for different occupations in the European Community. *Journal of Applied Psychology*, 88, 1028-1081.
- Sarges, W. & Wottawa, H. (2001). *Handbuch wirtschaftspsychologischer Testverfahren*. Lengerich: Pöpst Science Publisher.
- Scarr, S. & Weinberg, R.A (1979). Intellectual similarities in adoptive and biologically related families of adolescents. In L. Willermann & R.B. Turner (Hrsg.), *Readings about individual and group differences* (S.61-73). San Francisco: Freeman.
- Scheerer, E. (1978). *Probleme der Modellierung kognitiver Prozesse. Von der Funktionsanalyse zur genetischen Analyse* (Arbeitseinheit Kognitionspsychologie, Bericht Nr.4). Bochum: Psychologisches Institut der Ruhr- Universität.
- Scheurer, H. (1999). Diagnostik als Testung. In Jäger R.S. & Petermann (Hrsg.), *Psychologische Diagnostik* (S. 257-263). Weinheim: Psychologie Verlags Union.
- Schmidt, J. U. (1984). Simultane Überprüfung der Zweimodalität im Berliner Intelligenzstrukturmodell. *Diagnostica*, 30, 93-103.
- Schmidt, F. L. (1992). What do data really mean. Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist*, 47, 1173-1181.

- Schmidt-Atzert, L. & Deter, B. (1993a). Intelligenz- und Ausbildungserfolg: Eine Untersuchung zur prognostischen Validität des IS-T 70. *Zeitschrift für Arbeits- und Organisationspsychologie*, 37, 52-63.
- Schmidt-Atzert, L. & Deter, B. (1993b). Die Vorhersage von Ausbildungserfolg bei verschiedenen Berufsgruppen durch Leistungstests. *Zeitschrift für Arbeits- und Organisationspsychologie*, 37, 191-161.
- 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, F. L. & Hunter, J. E. (1998). 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. (2000). Messbare Personenmerkmale: Zur Stabilität, Variabilität und Validität zur Vorhersage zukünftiger Berufsleistung und berufsbezogenes Lernens. In M. Kleinmann & B. Strauß (Hrsg.), *Potentialfeststellung und Potentialentwicklung* (S. 15-41). Göttingen: Verlag für angewandte Psychologie.
- Schmiel, M. (1988). Schlüsselqualifikationen als Lernziele in der beruflichen Aus- und Weiterbildung. In S. Rosa, D. Schart & K.H. Sommer (Hrsg.), *Fachübergreifende Qualifikationen und betriebliche Aus- und Weiterbildung*. Stuttgarter Beiträge zur Berufs- und Wirtschaftspädagogik (Bd. 5, S. 51-80). Esslingen: Deugro.
- Schneider, W. & Bjorklund, D. F. (1992). Expertise, aptitude and strategic remembering. *Child Development*, 63, 461-473.
- Schneider, W., Koerkel, J. & Weinert, F. E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low aptitude children. *Journal of Educational Psychology*, 81, 306-312.
- Schönpflug, W. & Schönpflug, U. (1989). *Psychologie: Allgemeine Psychologie und ihre Verzweigungen in die Entwicklungs-, Persönlichkeits- und Sozialpsychologie*, 2. Auflage. München: Psychologie Verlags Union.
- Schorr, A. (1991). Diagnostische Praxis in der Arbeits- und Organisationspsychologie. Teilergebnisse aus einer repräsentativen Umfrage zur diagnostischen Praxis. In H. Schuler & U. Funke (Hrsg.), *Eignungsdiagnostik in Forschung und Praxis. Psychologische Information für Auswahl, Beratung und Förderung von Mitarbeitern, Band 10 Beiträge zur Organisationspsychologie* (S. 6-14). Stuttgart: Verlag für Angewandte Psychologie.
- Schüppel, J. (1996). *Wissensmanagement: Organisatorisches Lernen im Spannungsfeld von Wissen- und Lernbarrieren*. Wiesbaden: Gabler.
- Schuler, H. (2000). *Psychologische Personalauswahl*. Göttingen: Hofgreffe.

- Schuler, H. (2006a). *Lehrbuch der Personalpsychologie* (2. Aufl.). Göttingen: Hogrefe.
- Schuler, H. (2006b). Gegenstandsbereich und Aufgaben der Personalpsychologie. In H. Schuler (Hrsg.), *Lehrbuch der Personalpsychologie* (2. Aufl., S. 3-15). Göttingen: Hogrefe.
- Schuler, H. (2006c). Arbeits- und Anforderungsanalyse. In H. Schuler (Hrsg.), *Lehrbuch der Personalpsychologie* (2. Aufl., S. 45-69). Göttingen: Hogrefe.
- Schuler, H. (2007). *Lehrbuch Organisationspsychologie*. Bern: Verlag Hans Huber.
- Schuler, H., Frier, D. & Kaufmann, M. (1993). *Personalauswahl im europäischen Vergleich*. Göttingen: Verlag für angewandte Psychologie.
- 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*, 2, 60- 70.
- Schuler, H. & Höft, S. (2006). Konstruktorientierte Verfahren der Personalauswahl. In H. Schuler (Hrsg.), *Lehrbuch der Personalpsychologie* (S. 101-144). Göttingen: Hogrefe.
- 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). Stuttgart: Verlag für angewandte Psychologie.
- Schumacher, J. & Brähler, E. (2005). Psychodiagnostische Testverfahren. In P.L. Jannsen, P. Joraschki & W. Tress (Hrsg.), *Leitfaden psychosomatischer Medizin und Psychotherapie* (S. 86-92). Deutscher Ärzteverlag.
- Schwank, I. (1990). *Untersuchungen algorithmischer Denkprozesse von Mädchen. Abschlussbericht Band I*. Osnabrück: Forschungsinstitut für Mathematikdidaktik.
- Schwippert, K. & Peek, R. (2003). *Mehrsprachigkeit und Leseverständnis russlanddeutscher Grundscülerinnen und Grundschüler (MehrLe)- Pilotuntersuchung* (Sozialwissenschaftliche Forschungsdokumentation, Bd.18). Münster: Institut für sozialwissenschaftliche Forschung e.V.
- Schwippert, K., Peek, R. & Bolle, B. (2004). Mehrsprachigkeit und Leseverständnis russlanddeutscher Grundschüler. In W. Bos, E.M- Lankes, N. Plaßmeier & K. Schwippert (Hrsg.), *Heterogenität- Eine Herausforderung an die empirische Bildungsforschung* (S.253-269). Münster: Waxmann.
- Shackelton, V. & Newell, S. (1997). International selection and assessment. In N. Anderson & P. Herriet (Hrsg.), *International Handbook of Selection and Assessment* (S. 81-95). Chichester: John Wiley & Sons.
- Skinner, B.F. (1974). *About Behaviorism* (2.Aufl.). New York: Knopf.

- Sonntag, Kh. (1996). *Lernen im Unternehmen*. München: C.H. Beck.
- Sonntag, Kh. & Schaper, N. (2006). Wissensorientierte Verfahren der Personalentwicklung. In H. Schuler (Hrsg.), *Lehrbuch der Personalpsychologie* (2. Aufl., S. 255-280). Göttingen: Hogrefe.
- Spada H. & Mandl H. (1988). Einführung. In H. Mandl & H. Spada (Hrsg.), *Wissenspsychologie* (S.1-19). München: Psychologie Verlags Union.
- Spearman, C. (1904). „General Intelligence“, objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Spinner, H.F. (1994). *Die Wissensordnung. Ein Leitkonzept für die dritte Grundordnung des Informationszeitalters*. Opladen: Leske und Budrich.
- Spreter-Müller, B. (1988). Außerfachliche Qualifikation in der Wirtschaft, Anforderungen für den beruflichen Ein- und Aufstieg. In Bundesminister für Bildung und Wissenschaft (Hrsg.), *Studien zu Bildung und Wissenschaft Band 62* (S.1-135). Bad Honnef: Bock.
- Staehe, W.H. (1989). Human Resource Management und Unternehmensstrategie. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 22 ( 3), 388- 396.
- Stangel-Meseke, M. (1994). *Schlüsselqualifikation in der betrieblichen Praxis. Ein Ansatz in der Psychologie*. Wiesbaden: Deutscher Universitäts-Verlag.
- Staudt, E., Kriegesmann, B. & Muschik, C. (2003). IAI Corecard of Competence. In J. Erpenbeck & L.v. Rosenstil (Hrsg.), *Handbuch Kompetenzmessung*. (S. 160-168). Stuttgart:Schäffer-Poeschel.
- Steiner, G. (1988). Analoge Repräsentation. In H. Mandl & H. Spada (Hrsg.), *Wissenspsychologie* (S.99-123). München: Psychologie Verlags Union.
- Steinkamp, M. W. & Maehr, M. L. (1983). Affect, ability, and science achievement: A quantitative synthesis of correlational research. *Review of Educational Research*, 53, 369-396.
- Stern, W. (1900). *Über Psychologie der individuellen Differenzen: Ideen zu einer „Differentiellen Psychologie“*. Leipzig: Barth.
- Stern, W. (1912): Die psychologischen Methoden der Intelligenzprüfung. In F. Schumann (Hrsg.), *Bericht über den 5. Kongress für Experimentelle Psychologie* (S.1-109). Leipzig: Barth.
- Stern, E. (1994). Wie viele Kinder bekommen keinen Mohrenkopf? Zur Bedeutung der Kontexteinbettung beim Verstehen des quantitativen Vergleichs. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 24, 79- 94.

- Stern, E. (1998). Die Entwicklung schulbezogener Kompetenzen: Mathematik. In F.E. Weinert (Hrsg.), *Entwicklung im Kindesalter- Bericht über eine Längsschnittstudie* (S. 95- 113). Weinheim: Psychologie Verlags Union.
- Stern, E. (2001). Intelligenz, Wissen, Transfer und der Umgang mit Zeichensystemen. In E. Stern & J. Guthke (Hrsg.), *Perspektiven der Intelligenzforschung* (S.163-205). Lengerich: Pabst.
- Stern, E. & Guthke, J. (2001). Vorwort. In E. Stern & J. Guthke (Hrsg.), *Perspektiven der Intelligenzforschung*. Lengerich: Pabst.
- Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning*. Hillsdale, NJ: Erlbaum.
- Sternberg, R. J. & Powell, J. S. (1982). Theories of intelligence. In R. J. Sternberg (Hrsg.), *Handbook of human intelligence* (S. 975-1007). Cambridge, UK: Cambridge University Press.
- Sternberg, R.J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.
- Süß, H. M. (1996). Intelligenz, Wissen und Problemlösen. Kognitive Voraussetzungen für erfolgreiches Handeln bei computersimulierten Problemen. In D. Albert & K. Pawlik, K. H. Stapf & W. Stroebe (Hrsg.), *Lehr- und Forschungstexte Psychologie*. Göttingen: Hogrefe.
- Süß, H.-M. (1998). Diagnostik der operativen Intelligenz oder: Brauchen wir ein neues Fähigkeitskonstrukt? *Zeitschrift für Differentielle und Diagnostische Psychologie*, 19, 261-288.
- Süß, H.M. (2001). Prädiktive Validität der Intelligenz im schulischen und außerschulischen Bereich. In E. Stern & J. Guthke (Hrsg.), *Perspektiven der Intelligenzforschung* (S.109-136). Lengerich: Pabst Science Publishers.
- Süß, H. M. (2007). Eine Intelligenz – viele Intelligenzen. Neuere Intelligenztheorien im Widerstreit. *News@science. Begabtenförderung und Begabtenforschung*, 16, 17-34.
- Süß, H.M., Weiß, S. & Seidel, K. (2005). Soziale Kompetenzen. In H. Weber & T. Rammsayer (Hrsg.), *Handbuch der Persönlichkeitspsychologie und Differentiellen Psychologie* (S. 350- 361). Göttingen: Hogrefe.
- Teichler, U., Bugereit, M. & Holtkamp, R. (1984). Hochschulzertifikate in der betrieblichen Einstellungspraxis. In Bundesminister für Bildung und Wissenschaft (Hrsg.), *Studien Bildung und Wissenschaft Band 6* (S. 1-208). Bad Honnef: Bock.
- Terman, L. & Merrill, M.A. (1960). *Stanford Binet Intelligence Scale: Manual for the third revision*. Boston: Houghton Mifflin Co.

- Tewes, U. (1991). *Hamburg-Wechsler-Intelligenztest für Erwachsene-Revision (HA-WIE-R)*. Bern: Huber.
- Thurstone, L. L. (1938). *Primary and mental abilities*. Chicago: The University of Chicago Press.
- Tichy, N.M./ Fombrun, C.J./ Devanna, M.A. (1982). Strategic Human Resource Management. *Sloan Management Review*, 23 (2), 47- 61.
- Tichy, N.M., Fombrun, C.J. & Devanna, M.A. (1984). A Framework for strategic Human Resource Management. In N.M. Tichy, C.J. Fombrun, C.J. & Devanna, M.A. (Hrsg.), *Strategic Human Resource Management* (S.33- 51). New York: Wiley.
- Tiedemann, J. & Billmann-Mahecha (2004). Migration, Familiensprache und Schulerfolg– Ergebnisse aus der Hannoverischen Grundschulstudie. In W Bos, E. M. Lankes, N. Plaßmeier & K. Schwippert (Hrsg.), *Heterogenität eine Herausforderung an die empirische Bildungsforschung* (S. 269-279). Münster: Waxmann.
- Trautwein, U. Köller, O., Lehmann, R. & Lüdtke, O. (2007). Öffnung von Bildungswegen, erreichtes Leistungsniveau und Vergleichbarkeit von Abschlüssen. In U. Trautwein, O. Köller, R. Lehmann & O. Lüdtke (Hrsg.), *Schulleistungen von Abiturienten: Regionale, schulformbezogene und soziale Disparitäten* (S. 11-31). Münster: Waxmann Verlag.
- Vivantes- Netzwerk für Gesundheit (2010). *Ausbildungsberufe im IBBG (Institut für berufliche Bildung im Gesundheitswesen* [im Internet]. Verfügbar unter: <http://www.vivantes.de/vivantes-das-unternehmen/karriere/ausbildung/berufe/?0=> [29.10.2010].
- Walberg, H. J., Strykowski, B. F., Rovai, E. & Hung, S. S. (1984). Exceptional performance. *Review of Educational Research*, 54, 87-112.
- Waldmann, M. R. (1997). Wissen und Lernen. *Psychologische Rundschau*, 48, 84-100.
- Waldmann, M. & Weinert, F. E. (1990). *Intelligenz und Denken*. Göttingen: Hogrefe.
- Wechsler, D. (1958). *The measurement and appraisal of adult intelligence*. Baltimore: Williams & Wilkins.
- Wechsler, D. (1964). *Die Messung der Intelligenz Erwachsener* (3. Unveränderte Aufl.). Bern: Huber.
- Weinert, F. E. (1988). Jenseits des Glaubens an notwendige und hinreichende Bedingungen des schulischen Lernens. In J. Lompscher, W. Jantos & S. Schönian (Hrsg.), *Psychologische Methoden der Analyse und Ausbildung der Lerntätigkeit*. (S. 116-139). Berlin: Gesellschaft für Psychologie der DDR.
- Weinert, F. E. (1994). Lernen lernen und das eigene Lernen verstehen. In K. Reusser & M. Reusser-Weyeneth (Hrsg.), *Verstehen. Psychologischer Prozess und didaktische Aufgabe*. (S. 183-205). Bern. Huber.



- Weinert, F.E. & Helmke, A. (1998). The neglected role of individual differences in theoretical models of cognitive development. *Learning and Instruction*, 8 (4), 309-323.
- Weinert, F.E. (2001). Vergleichende Leistungsmessung in Schulen - eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessung in Schulen* (S. 17-33). Weinheim und Basel: Beltz-Verlag.
- Weiß, R.H. (1974). *Grundintelligenztest CFT 2, Skala 2. Handanweisung für die Durchführung, Auswertung und Interpretation*. Braunschweig: Westermann.
- Weiß, R.H. (1997). *Grundintelligenztest Skala 2 (CFT 20) mit Wortschatztest (WS) und Zahlenfolgetest (ZF)* (4. Auflage). Göttingen: Hogrefe.
- Wenninger, G. (2001). *Lexikon der Psychologie: In fünf Bänden*. Heidelberg: Spektrum Akademischer Verlag.
- White, R.W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, 66 (5), 297-333.
- Wilberg, S. & Lynn, R. (1999). Sex differences in historical knowledge and school grades: A 26-nations study. *Personality and Individual Differences*, 27, 1221-1229.
- Wolff (1999). Potentiale und Dimensionen der Wissensgesellschaft – Ergebnisse aus dem “Wissens- Delphi“. In B. von Rosenblatt (Hrsg.): *Bildung in der Wissensgesellschaft* (S. 11-18). Münster: Waxmann.
- Wolff, H. & Stock, J. (2000). Allgemeinwissen – die Herausforderung der Wissensgesellschaft. In G. de Haan, J. Hamm-Brücher & N. Reichel (Hrsg.), *Bildung ohne Systemzwänge. Innovationen und Reformen* (S. 23-44). Neuwied: Luchterhand.
- Wottawa, H. (1980). *Grundriß der Testtheorie*. München: Juventa.
- Wottawa, H. & Hossiep, R. (1997). *Anwendungsfelder der Diagnostik*. Göttingen: Hogrefe.
- Wright, P.M., Dunford, B. & Snell, S.A. (2001). Human Resources and the Resource Based View of the Firm. *Journal of Management*, 27 (6), 701- 721.
- Yönter, Isil (1997) Eltern- und Familienarbeit. Familienarbeit als Bestandteil interkultureller Mädchenarbeit. In J. Ehlers, A. Bentner & M. Kowalczyk (Hrsg.), *Mädchen zwischen den Kulturen. Anforderungen an eine Interkulturelle Pädagogik* (S. 117-130). Frankfurt am Main: IKO – Verlag für Interkulturelle Kommunikation.

## Abbildungsverzeichnis

Abbildung 1: Kompetenz und Performanz, nach Gnahn (2007) .....	33
Abbildung 2: Ansätze der Eignungsdiagnostik und deren Validierungslogik .....	45
Abbildung 3: Modell zum Beziehungsgeflecht zwischen Intelligenz & Wissen .....	79
Abbildung 4: Die Wissenstreppe nach North (2005) .....	84
Abbildung 5: Konstrukte und Zuordnungsregeln .....	102
Abbildung 6: Das Berliner Intelligenzstrukturmodell (BIS), nach Jäger (1982) .....	116
Abbildung 7: Das Modell der fluiden und kristallisierten Intelligenz von Cattell .....	119
Abbildung 8: Die Interdependenz zwischen Wissen und Intelligenz .....	143
Abbildung 9: Das Rahmenmodell geistiger Leistung .....	144
Abbildung 10: Fähigkeitsstruktur des I-S-T-2000 R .....	174
Abbildung 11: Geschlechterverhältnis .....	182
Abbildung 12: Altersverteilung in der Stichprobe .....	184
Abbildung 13: Aufteilung der Bildungsinstitute in Prozent .....	186
Abbildung 14: Differenzierung der Stichprobe .....	187
Abbildung 15: Verteilung der Probanden über die Bildungsstände in Prozent .....	190
Abbildung 16: Multivariate Verteilung der Gesamtrohwerte .....	194
Abbildung 17: Rohwertverteilung der Aufgabengruppe „verbales Wissen“ .....	200
Abbildung 18: Rohwertverteilung der Aufgabengruppe „numerisches Wissen“ .....	202
Abbildung 19: Rohwertverteilung der Aufgabengruppe „figurales Wissen“ .....	205
Abbildung 20: Strukturgleichungsmodell I-S-T 2000 R-Grundmodul .....	228

## Tabellenverzeichnis

Tabelle 1: Trimodale Beziehung .....	41
Tabelle 2: Genese eines Tests.....	48
Tabelle 3: Einsatzhäufigkeit von Leistungs- und Intelligenztests .....	65
Tabelle 4: Einsatzhäufigkeit von Leistungs- und Intelligenztests .....	66
Tabelle 5: Metaanalytische Zusammenfassung der Befundlage zur Validität eignungsdiagnostischer Verfahren .....	73
Tabelle 6: Intelligenzdefinition .....	98
Tabelle 7: Beispiele für Zuordnungsregeln .....	101
Tabelle 8: Die Primärfaktoren von Thurstone.....	113
Tabelle 9: Fähigkeitsbündel im Berliner Intelligenzstrukturmodell.....	117
Tabelle 10: Reihenfolge und Bearbeitungszeit der eingesetzten Testverfahren.....	179
Tabelle 11: Altersverteilung getrennt nach Geschlecht.....	183
Tabelle 12: Ausgewählte Bildungsinstitute .....	184
Tabelle 13: Verteilung der Probanden über Ausbildungsziele und Bildungsinstitute getrennt nach Geschlecht.....	187
Tabelle 14: Verteilung der Probanden über die Bildungsabschlüsse getrennt nach Geschlecht und Muttersprache .....	192
Tabelle 15: Zuordnung der Test-Items zu den Inhaltsfacetten und Themenbereichen	197
Tabelle 16: Kennziffern der Aufgabengruppe „verbales Wissen“ .....	198
Tabelle 17: Kennziffern der Aufgabengruppe „numerisches Wissen“ .....	201
Tabelle 18: Kennziffern der Aufgabengruppe „figurales Wissen“ .....	203
Tabelle 19: Split-Half-Reliabilitätskoeffizienten .....	205
Tabelle 20: Retest-Reliabilität .....	206
Tabelle 21: Mittelwerte und Standardabweichungen der Testergebnisse getrennt nach Altersgruppen .....	207
Tabelle 22: Homogene Untergruppen bezüglich der Wissensausprägung in den Altersgruppen-Scheffé Prozedur .....	210
Tabelle 23: Mittelwerte und Standardabweichungen der Testergebnisse getrennt nach den unterschiedlichen Muttersprachlern.....	211
Tabelle 24: T-Test bei unabhängigen Stichproben.....	212
Tabelle 25: U-Test nach Mann und Whitney .....	213

Tabelle 26: Mittelwerte und Standardabweichungen der Testergebnisse getrennt nach dem Geschlecht.....	213
Tabelle 27: T-Test bei unabhängigen Stichproben.....	214
Tabelle 28: U-Test nach Mann und Whitney .....	214
Tabelle 29: Mittelwerte und Standardabweichungen der Testergebnisse getrennt nach Bildungsgruppen.....	215
Tabelle 30: T-Test bei unabhängigen Stichproben.....	218
Tabelle 31: Homogene Untergruppen bezüglich der Wissensausprägung in den Bildungsgruppen-S-N-K Prozedur .....	219
Tabelle 32: Korrelation zwischen Testwerten und Schulnoten.....	222
Tabelle 33: Korrelation zwischen Testwerten und der Durchschnittsnote des ersten Ausbildungsjahres.....	224
Tabelle 34: Korrelation der Wissensskalen des START-W mit dem Gesamtrohwert der BOWIT-Kurzform .....	225
Tabelle 35: Zuordnung der Wissensfragen zu Themenbereichen und Kodierungsarten.....	226
Tabelle 36: Kreuztabelle mit Altersgruppen und Bildungsniveau .....	237
Tabelle 37: Kreuztabelle: Bildungsniveau der Eltern & Schulabschluss der Kinder...	246
Tabelle 38: Interpretation der Skalen und des Faktorwertes Wissen (gc).....	256

## Verzeichnis des Anhangs

Anhang 1 .....	294
Anhang A: Empirischer Teil- Statistiken.....	297
Anhang B: Ergebnisse der Forschungsfragen- Teil 1 .....	302
Anhang B 1.1.....	302
Anhang B 1.2.....	309
Anhang B 1.3.....	310
Anhang B 1.4.....	311
Anhang B 1.5.....	318
Anhang B: Ergebnisse der Forschungsfragen- Teil 2 .....	319
Anhang B 2.1.....	319
Anhang B 2.2.....	320
Anhang B: Ergebnisse der Forschungsfragen- Teil 3 .....	321
Anhang B 3.2.....	321
Anhang C: Diskussion der Forschungsfragen- Teil 1 .....	323
Anhang C 1.1: .....	323

## Anhang 1

Am 12.8.2008 startete die erste Datenerhebungsphase mit der Untersuchung an der Charité Gesundheitsakademie (N= 19) statt. Die Probanden wurden durch die Fachbereichsleitungen, Lehrer, Poster und Flyer über das Forschungsprojekt informiert. Im Zeitraum vom 25.09.2008 bis zum 24.11.2008 wurden die Forschungsarbeiten am Marie-Curie-Gymnasium durchgeführt (N= 62). In der Zeitspanne vom 15.01.2009 bis zum 15.03.2009 wurden die Untersuchungen an der Freien Universität Berlin aufgenommen (N= 46).

Die Untersuchungen bei Vivantes (N = 581) und der Wannsee-Schule (N= 57) erstreckten sich über die gesamte Periode und wurden im Rahmen von internen Auswahlverfahren von potentiellen Bewerbern durchgeführt. Die Bewerbungsrunden setzten sich jeweils aus 10-20 Bewerbern zusammen und fanden in einem einwöchigen Turnus statt. Die Auswahlverfahren sind obligatorisch, um einen beruflichen Ausbildungsplatz zu erhalten und wurden wie folgt aufgebaut: Auf der ersten Stufe haben die Bewerber schriftliche Bewerbungsunterlagen eingereicht, die ein Anschreiben bzw. Motivations schreiben, Lebenslauf und ein Bewerbungsfoto enthielten. Nach einem positiven Bescheid durch das zentrale Bewerbungsbüro und der Institutsleitung wurden die Bewerber zu einer Aufnahmeprüfung eingeladen, die in den Örtlichkeiten des Instituts stattfand und von den Mitarbeitern des zentralen Bewerbungsbüros koordiniert und angeleitet wurde. In dem Zulassungsbescheid wurden die Bewerber vorab über den Ablauf und den Zeitaufwand (8.00 – 14.00 Uhr) des sogenannten Testtages informiert. Im Vorfeld der internen Aufnahmeprüfung wurden der Konzern, das Profil bzw. der Aufbau der berufsbezogenen Ausbildung, der mögliche Tätigkeitsbereich und der Tagesablauf vorgestellt. Die Aufnahmeprüfung des Instituts umfasst die Bereiche Rechtschreibung, Texterfassung, Mathematik, logisches Denken, Wahrnehmungsfähigkeit und Allgemeinwissen. Anschließend wurde der individuelle Test des jeweiligen Instituts bearbeitet. Die Bearbeitungsdauer betrug ca. zwei Stunden. Zum Abschluss bearbeiteten die Bewerber die Grundmodul-Kurzform des I-S-T 2000R und den START-W Wissenstest. Es wird nochmals darauf hingewiesen, dass bei der Testdurchführung sämtliche Richtlinien der DIN-Norm 33430 eingehalten worden sind (siehe Kapitel 3.2.4) und die Anwendung entsprechend der Hinweise und Bedingungen zur Testdurchführung von Liepmann et al. (2007, S.75) vorgenommen wurden (siehe Tabelle):

## Vorbereitung der Testdurchführung

- Untersuchungsteilnehmer bei der Einladung auf Vorbereitung hinweisen (nicht von den üblichen Ess- und Schlafgewohnheiten abweichen, keine stark belastenden Tätigkeiten vor der Prüfung, etc.,
- max. 30 Personen für jeweils zwei Untersuchungsleiter vorsehen; für entsprechende Räumlichkeiten ist zu sorgen,
- Testhefte auf Eintragungen, Markierungen, etc. durchsehen,
- Gegebenenfalls die Form (A,B,C) auf dem Antwortbogen markieren

## Testdurchführung: Allgemeine Hinweise

- Vorstellung der Untersuchungsleiter/innen, Dank für die Teilnahmebereitschaft,
- Sicherstellung, dass die drei Grundvoraussetzungen für die Durchführung von Testverfahren gegeben sind: (a) hinreichende Motivation zur (wahrheitsgemäßen) Bearbeitung der Fragen; (b) Erfahrung im Umgang mit umfassendem, gedruckten Material; (c) Beherrschung der deutschen Sprache.
- Teilnehmern Vorausschau auf die folgende Testaufgaben geben (unterschiedliche Aufgabengruppen mit jeweils aufsteigendem Schwierigkeitsgrad...),
- Informationen über Umfang: drei (zwei/ein) Aufgabenhefte sind zu bearbeiten; zwischen der Bearbeitung jeweils ca. 5-10 Min. Pause.

## Testdurchführung: Technische Hinweise

- Bedeutung der Funktion der Beispielaufgaben: die Beispielaufgaben dienen dazu, das Grundverständnis für die jeweilige Aufgabe herzustellen,
- Rückfragen nach Erläuterung der Beispielaufgaben sind zugelassen, während der Bearbeitung nur im Notfall und ohne Störung der anderen,
- Beginn der Bearbeitung der Aufgabengruppen jeweils erst nach Aufforderung,
- Verweis auf Zeitbegrenzung der einzelnen Aufgabengruppen geben; zügiges Arbeiten erforderlich,
- Überspringen der Aufgaben, sofern Schwierigkeiten auftreten,
- Beenden der Bearbeitung nach Ablauf der Zeit bzw. Hinweis des(r) Untersuchungsleiter(s),
- Wenn man in einer Aufgabengruppe weniger Aufgaben bearbeitet hat, bedeutet dies nichts für die Bearbeitung der nächsten Aufgabengruppe,
- Notizen können auf den freien Flächen des Antwortbogens gemacht werden, nicht im Aufgabenheft,
- bei fehlerhaften Markierungen sind richtige Lösungen eindeutig zu kennzeichnen,
- Lösungen nur auf dem Antwortbogen eintragen,
- vor Beginn der Testbearbeitung sind der Name und/oder Teilnehmercode und die weiteren Angaben einzusetzen; wenn Rückmeldung gewünscht, sollte der Teilnehmer einen Code eintragen, der eine Identifizierung möglich macht,
- das Testheft erst öffnen, wenn dazu aufgefordert wurde,
- nicht zurückblättern,
- alle Anweisungen, Erläuterungen und Beispiele vollständig lesen,

### **START-W: Besondere Hinweise**

„Für die Bearbeitung des Wissenstest haben Sie 40 Min. Zeit. Sie können innerhalb des Wissenstests selbständig vor- und zurückblättern. Markieren Sie die richtigen Lösungen immer auf dem Antwortbogen. Es ist immer nur eine Lösung richtig.“

Der strukturelle Ablauf am Untersuchungstag stellte sich wie folgt dar:

Nach den einführenden Worten und der Klärung von Fragen, seitens der Probanden, konnte sicher gestellt werden, dass die drei Grundvoraussetzungen für die Durchführung von Testverfahren gegeben waren. Das heißt, die Probanden verfügten über (a) eine hinreichende Motivation zur (wahrheitsgemäßen) Bearbeitung der Fragen, (b) hatten Erfahrung im Umgang mit umfassenden, gedruckten Material und (c) und beherrschten die deutsche Sprache. Im Anschluss wurde der demografische Fragebogen erläutert und gemeinsam ausgefüllt. Es folgte eine Instruktion über die Vorgehensweise bezüglich der verwendeten Testverfahren, die in Form einer Papier-und-Bleistift-Version dargeboten wurden. Die Anweisung lautete, die Fragestellungen aus den obigen Gründen ehrlich, ernsthaft und angemessen zu beantworten, bei Unsicherheit die schlüssigste Antwort zu wählen und den Test schnellstmöglich zu bearbeiten.

Im Unterschied zu den anderen Bildungseinrichtungen diente der I-S-T 2000 R und der START-W am Vivantes Institut und an der Wannseeschule e.V. als Entscheidungshilfe zur Bewerberselektion und demnach war die Teilnahme vorgegeben. Aus diesem Grund kann davon ausgegangen werden, dass die Motivation der Probanden von vornherein entsprechend hoch war. Die Ergebnisse, der namentlich bekannten Bewerber, wurden folglich an die Institutsleitung zurückgemeldet, wobei die Personen die Resultate zu einem späteren Zeitpunkt ebenfalls einsehen konnten. Nach bestandener Prüfung (ausschlaggebend war die interne Aufnahmeprüfung, die mindestens mit der Note 4,0 bestanden werden musste) wurden die Bewerber erneut zu einem persönlichen Gespräch eingeladen, um die besten Bewerber auszusuchen. An der Charité, dem Marie-Curie Gymnasium sowie an der Freien Universität wurde auf den Datenschutz verwiesen und dementsprechend ein persönliches Kennwort generiert, das die Anonymität der Probanden garantieren sollte. Das Kennwort ermöglichte den Schülern die veröffentlichten Auswertungen bzw. die Ergebnisse im Internet nachzuvollziehen.



## Anhang A: Empirischer Teil- Statistiken

### I.

Statistiken und Überprüfung der Skalen auf Normalverteilung:  
Kolmogorov-Smirnov-Test & Schiefe- und Kurtosiskennwerte

### II.

Reliabilitätsanalysen der Wissensskalen

### I.

#### Statistiken der gesamten Wissensskala- nicht selektiert

##### Zusammenfassung der Fallverarbeitung

		N	%
Fälle	Gültig	753	100,0
	Ausgeschlossen <sup>a</sup>	0	,0
	Gesamt	753	100,0

a. Listenweise Löschung auf der Grundlage aller Variablen in der Prozedur.

#### Statistiken

##### Skala\_Gesamt

N	Gültig	753
	Fehlend	0
Mittelwert		65,4781
Standardabweichung		17,69714
Schiefe		,041
Standardfehler der Schiefe		,089
Kurtosis		-,275
Standardfehler der Kurtosis		,178
Minimum		18,00
Maximum		112,00

##### Kolmogorov-Smirnov-Anpassungstest

	Skala_Gesamt
Kolmogorov-Smirnov-Z	,931
Asymptotische Signifikanz (2-seitig)	,351

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

#### Reliabilitätsstatistiken

Cronbachs Alpha	Anzahl der Items
,920	134

**Statistiken der verbalen Wissensskala- nicht selektiert****Statistiken****Wissen\_Verbal**

N	Gültig	753
	Fehlend	0
Mittelwert		21,8619
Standardabweichung		6,39706
Schiefe		,045
Standardfehler der Schiefe		,089
Kurtosis		-,366
Standardfehler der Kurtosis		,178
Minimum		6,00
Maximum		39,00

**Reliabilitätsstatistiken**

Cronbachs Alpha	Anzahl der Items
,789	44

	Mittelwert Schwierigkeitsindex	Korrigierte Item-Skala- Korrelation Trennschärfe	Cronbachs Alpha, wenn Item weggelassen
ri01	,3679	,226	,785
ri04	,3891	,109	,790
ri07	,6985	,280	,784
ri10	,3732	,187	,787
ri13	,5219	,163	,788
ri16	,3147	,179	,787
ri18	,4754	,262	,784
ri20	,2961	,181	,787
ri24	,8274	,312	,783
ri29	,4887	,174	,787
ri32	,4303	,164	,788
ri35	,7344	,186	,787
ri38	,5166	,319	,782
ri41	,3865	,326	,782
ri44	,5206	,309	,782
ri47	,6985	,384	,780
ri50	,7875	,284	,784
ri56	,4489	,051	,792
ri59	,3347	,122	,789

ri62	,6428	,192	,787
ri65	,4077	,242	,785
ri68	,7716	,378	,781
ri72	,6348	,271	,784
ri75	,5591	,159	,788
ri78	,5113	,458	,777
ri80	,7968	,323	,782
ri83	,8977	,314	,784
ri84	,5259	,188	,787
ri85	,2829	,059	,791
ri87	,4967	,305	,783
ri88	,3280	,290	,783
ri90	,3028	,204	,786
ri91	,1713	,121	,788
ri93	,2457	,115	,789
ri103	,7450	,277	,784
ri108	,2444	,103	,789
ri109	,5564	,337	,781
ri113	,2855	,368	,781
ri114	,7782	,371	,781
ri116	,2151	,341	,782
ri117	,5100	,437	,777
ri120	,7317	,419	,779
ri122	,3652	,263	,784
ri127	,2444	,269	,784

### Statistiken der numerischen Wissensskala- nicht selektiert

#### Statistiken

##### Wissen\_Numerisch

N	Gültig	753
	Fehlend	0
Mittelwert		19,5910
Standardabweichung		5,81177
Schiefe		,188
Standardfehler der Schiefe		,089
Kurtosis		-,177
Standardfehler der Kurtosis		,178
Minimum		4,00

## Statistiken

## Wissen\_Numerisch

N	Gültig	753
	Fehlend	0
Mittelwert		19,5910
Standardabweichung		5,81177
Schiefe		,188
Standardfehler der Schiefe		,089
Kurtosis		-,177
Standardfehler der Kurtosis		,178
Minimum		4,00
Maximum		38,00

## Reliabilitätsstatistiken

Cronbachs Alpha	Anzahl der Items
,764	44

	Mittelwert Schwierigkeitsindex	Korrigierte Item-Skala- Korrelation Trennschärfe	Cronbachs Alpha, wenn Item weggelassen
ri02	,2975	,108	,765
ri05	,3108	,215	,760
ri08	,1607	,252	,759
ri11	,2510	,158	,762
ri14	,4011	,224	,760
ri17	,1886	,043	,766
ri21	,3174	,327	,755
ri25	,5060	,154	,763
ri27	,6999	,322	,756
ri30	,2616	,328	,756
ri33	,9602	,074	,764
ri36	,5139	,233	,759
ri39	,4568	,196	,761
ri42	,3785	,255	,758
ri45	,8287	,303	,757
ri48	,4967	,413	,751
ri51	,3639	,217	,760
ri53	,1036	,166	,762
ri54	,2125	,125	,763
ri57	,6946	,249	,759
ri60	,4369	,172	,762

ri63	,4117	,177	,762
ri66	,3201	,259	,758
ri69	,7955	,268	,758
ri71	,7809	,162	,762
ri73	,4050	,276	,757
ri76	,8539	,177	,761
ri79	,6614	,289	,757
ri89	,7304	,308	,756
ri94	,1461	,042	,766
ri95	,4290	,116	,765
ri96	,5737	,292	,757
ri98	,2072	,035	,767
ri102	,6547	,488	,748
ri104	,5883	,318	,756
ri105	,1912	,114	,764
ri107	,0664	-,028	,766
ri111	,5604	,418	,751
ri112	,7928	,345	,755
ri119	,2019	,193	,761
ri124	,3758	,151	,763
ri126	,2882	,355	,754
ri129	,1541	,062	,765
ri130	,5618	,444	,750

### Statistiken der figuralen Wissensskala- nicht selektiert

#### Statistiken

##### Wissen\_Figural

N	Gültig	753
	Fehlend	0
Mittelwert		24,0252
Standardabweichung		7,04549
Schiefe		-,097
Standardfehler der Schiefe		,089
Kurtosis		-,449
Standardfehler der Kurtosis		,178
Minimum		5,00
Maximum		41,00
Cronbachs Alpha: .832; Anzahl der Items: 46		

## Anhang B: Ergebnisse der Forschungsfragen- Teil 1

### Anhang B 1.1

Vorbedingungen einer parametrischen Varianzanalyse:

- I. Test auf Homogenität der Varianzen: Levene Test
- II. Statistiken und Überprüfung der Daten auf Normalverteilung:  
Kolmogorov-Smirnov-Test & Schiefe- und Kurtosiskennwerte

Post-Hoc-Mehrfachvergleiche:

- III. Mehrfachvergleich nach der Scheffé Prozedur und homogene Untergruppenbildung nach Student-Newman-Keul

#### I.

##### Test auf Homogenität der Varianzen

Test der Homogenität der Varianzen

	Levene-Statistik	df1	df2	Signifikanz
Verbale Wissensskala	13,932	4	748	,000
Numerische Wissensskala	4,886	4	748	,001
Figurale Wissensskala	9,991	4	748	,000
Gesamte Wissensskala	12,510	4	748	,000

#### II.

Statistiken und Überprüfung der Daten auf Normalverteilung

##### Altersstufe 1: 16-Jährige

Statistiken

		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	35	35	35	35
	Fehlend	0	0	0	0
	Mittelwert	15,8000	15,3714	19,6857	50,8571
Schiefe		,218	1,361	1,022	1,148
Standardfehler der Schiefe		,398	,398	,398	,398
Kurtosis		1,231	1,947	,617	1,121
Standardfehler der Kurtosis		,778	,778	,778	,778
Minimum		5,00	8,00	10,00	28,00
Maximum		27,00	31,00	35,00	87,00
Kolmogorov-Smirnov-Z		1,037	,920	1,376	1,229

**Statistiken**

		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	35	35	35	35
	Fehlend	0	0	0	0
	Mittelwert	15,8000	15,3714	19,6857	50,8571
Schiefe		,218	1,361	1,022	1,148
Standardfehler der Schiefe		,398	,398	,398	,398
Kurtosis		1,231	1,947	,617	1,121
Standardfehler der Kurtosis		,778	,778	,778	,778
Minimum		5,00	8,00	10,00	28,00
Asymptotische Signifikanz (2-seitig)		,232	,366	,045	,097

**Altersstufe 2: 17-18-Jährige****Statistiken**

		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	142	142	142	142
	Fehlend	0	0	0	0
Mittelwert		18,3310	16,3592	23,3451	58,0352
Standardabweichung		5,40780	4,73308	6,08690	14,43912
Schiefe		,291	,125	,206	,281
Standardfehler der Schiefe		,203	,203	,203	,203
Kurtosis		-,192	-,176	-,325	,003
Standardfehler der Kurtosis		,404	,404	,404	,404
Minimum		7,00	4,00	10,00	21,00
Maximum		34,00	28,00	40,00	100,00
Kolmogorov-Smirnov-Z		1,158	1,011	1,144	,870
Asymptotische Signifikanz (2-seitig)		,137	,258	,146	,436

**Altersstufe 3: 19-20-Jährige****Statistiken**

		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	321	321	321	321
	Fehlend	0	0	0	0
Mittelwert		21,5421	18,6511	25,2991	65,4922
Standardabweichung		5,38101	5,22342	6,30805	15,30443
Schiefe		-,125	,136	-,176	-,053
Standardfehler der Schiefe		,136	,136	,136	,136
Kurtosis		-,131	-,210	-,159	,058
Standardfehler der Kurtosis		,271	,271	,271	,271
Minimum		4,00	5,00	5,00	14,00
Maximum		35,00	34,00	40,00	105,00
Kolmogorov-Smirnov-Z		1,086	,946	1,044	,760
Asymptotische Signifikanz (2-seitig)		,189	,332	,226	,610

**Altersstufe 4: 21-25-Jährige****Statistiken**

		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	179	179	179	179
	Fehlend	0	0	0	0
Mittelwert		21,6536	17,8156	24,5140	63,9832
Standardabweichung		6,24883	5,74305	7,26803	18,04099
Schiefe		-,023	,017	-,271	-,117
Standardfehler der Schiefe		,182	,182	,182	,182
Kurtosis		-,507	-,560	-,521	-,581
Standardfehler der Kurtosis		,361	,361	,361	,361
Minimum		8,00	3,00	7,00	24,00
Maximum		36,00	31,00	39,00	103,00
Kolmogorov-Smirnov-Z		,836	,738	,949	,589
Asymptotische Signifikanz (2-seitig)		,487	,647	,329	,878



**Altersstufe 5: 26-46-Jährige****Statistiken**

		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	76	76	76	76
	Fehlend	0	0	0	0
Mittelwert		19,5658	15,7105	20,7632	56,0395
Standardabweichung		8,63688	7,22923	9,46413	24,23575
Schiefe		,146	,538	,449	,389
Standardfehler der Schiefe		,276	,276	,276	,276
Kurtosis		-1,030	-,460	-,784	-,827
Standardfehler der Kurtosis		,545	,545	,545	,545
Minimum		5,00	3,00	6,00	21,00
Maximum		37,00	32,00	41,00	108,00
Kolmogorov-Smirnov-Z		,971	1,104	,962	,789
Asymptotische Signifikanz (2-seitig)		,302	,175	,313	,563

**III.**

Post-Hoc-Mehrfachvergleiche: Scheffé Prozedur und homogene Untergruppenbildung nach Student-Newman-Keul

Abhängige Variable: Verbale Wissensskala						
(I) Altersgruppen_New		(J) Altersgruppen_New		Mittlere Differenz (I-J)	Standardfehler	Signifikanz
	1,00	—	2,00	-2,53099	1,12631	,283
			3,00	-5,74206*	1,06240	,000
			4,00	-5,85363*	1,10305	,000
			5,00	-3,76579*	1,21918	,050
	2,00	—	1,00	2,53099	1,12631	,283
			3,00	-3,21107*	,60151	,000
			4,00	-3,32265*	,67070	,000
			5,00	-1,23480	,84825	,714
	3,00	—	1,00	5,74206*	1,06240	,000
			2,00	3,21107*	,60151	,000
			4,00	-,11158	,55674	1,000
			5,00	1,97627	,76135	,152

	4,00		1,00	5,85363 <sup>+</sup>	1,10305	,000
			2,00	3,32265 <sup>+</sup>	,67070	,000
		–	3,00	,11158	,55674	1,000
			5,00	2,08784	,81712	,164
	5,00		1,00	3,76579 <sup>+</sup>	1,21918	,050
			2,00	1,23480	,84825	,714
		–	3,00	-1,97627	,76135	,152
			4,00	-2,08784	,81712	,164
Abhängige Variable: Numerische Wissensskala						
	1,00		2,00	-,98773	1,03980	,924
			3,00	-3,27966 <sup>+</sup>	,98080	,025
		–	4,00	-2,44421	1,01833	,219
			5,00	-,33910	1,12554	,999
	2,00		1,00	,98773	1,03980	,924
			3,00	-2,29194 <sup>+</sup>	,55531	,002
		–	4,00	-1,45649	,61919	,238
			5,00	,64863	,78310	,953
	3,00		1,00	3,27966 <sup>+</sup>	,98080	,025
			2,00	2,29194 <sup>+</sup>	,55531	,002
		–	4,00	,83545	,51398	,620
			5,00	2,94056 <sup>+</sup>	,70287	,002
	4,00		1,00	2,44421	1,01833	,219
			2,00	1,45649	,61919	,238
		–	3,00	-,83545	,51398	,620
			5,00	2,10512	,75436	,101
	5,00		1,00	,33910	1,12554	,999
			2,00	-,64863	,78310	,953
		–	3,00	-2,94056 <sup>+</sup>	,70287	,002
			4,00	-2,10512	,75436	,101
Abhängige Variable: Figurale Wissensskala						
	1,00		2,00	-3,65936	1,29516	,093
			3,00	-5,61335 <sup>+</sup>	1,22167	,000
		–	4,00	-4,82825 <sup>+</sup>	1,26841	,006
			5,00	-1,07744	1,40196	,964
	2,00		1,00	3,65936	1,29516	,093
			3,00	-1,95399	,69168	,093
		–	4,00	-1,16890	,77125	,681

	3,00	-	5,00	2,58191	,97542	,137			
			1,00	5,61335*	1,22167	,000			
			2,00	1,95399	,69168	,093			
			4,00	,78510	,64021	,826			
	4,00	-	5,00	4,53591*	,87549	,000			
			1,00	4,82825*	1,26841	,006			
			2,00	1,16890	,77125	,681			
			3,00	-,78510	,64021	,826			
	5,00	-	5,00	3,75081*	,93962	,003			
			1,00	1,07744	1,40196	,964			
			2,00	-2,58191	,97542	,137			
			3,00	-4,53591*	,87549	,000			
				4,00	-3,75081*	,93962	,003		
				Abhängige Variable: Gesamte Wissensskala					
				1,00	-	2,00	-7,17807	3,18524	,280
						3,00	-14,63507*	3,00450	,000
4,00	-13,12610*	3,11947	,002						
5,00	-5,18233	3,44790	,688						
2,00	-	1,00	7,17807	3,18524	,280				
		3,00	-7,45700*	1,70109	,001				
		4,00	-5,94803*	1,89677	,044				
		5,00	1,99574	2,39890	,952				
3,00	-	1,00	14,63507*	3,00450	,000				
		2,00	7,45700*	1,70109	,001				
		4,00	1,50897	1,57449	,922				
		5,00	9,45274*	2,15313	,001				
4,00	-	1,00	13,12610*	3,11947	,002				
		2,00	5,94803*	1,89677	,044				
		3,00	-1,50897	1,57449	,922				
		5,00	7,94377*	2,31084	,019				
5,00	-	1,00	5,18233	3,44790	,688				
		2,00	-1,99574	2,39890	,952				
		3,00	-9,45274*	2,15313	,001				
		4,00	-7,94377*	2,31084	,019				

**Homogene Untergruppen nach der S-N-K- Prozedur**

Gesamte Wissensskala							
	Altersgruppen_New		N	Untergruppe für Alpha = 0.05.			
				1	2	3	
Student-Newman-Keuls-Prozedur <sup>a,b</sup>	-	1,00	35	50,8571			
		5,00	76		56,0395		
		2,00	142		58,0352		
		4,00	179			63,9832	
		3,00	321			65,4922	
		Signifikanz			1,000	,436	,556

Verbale Wissensskala							
	Altersgruppen_New		N	Untergruppe für Alpha = 0.05.			
				1	2	3	
Student-Newman-Keuls-Prozedur <sup>a,b</sup>	-	1,00	35	15,8000			
		2,00	142		18,3310		
		5,00	76		19,5658	19,5658	
		3,00	321			21,5421	
		4,00	179			21,6536	
		Signifikanz			1,000	,173	,055

Numerische Wissensskala							
	Altersgruppen_New		N	Untergruppe für Alpha = 0.05.			
				1	2	3	
Student-Newman-Keuls-Prozedur <sup>a,b</sup>	-	1,00	35	15,3714			
		5,00	76	15,7105			
		2,00	142	16,3592	16,3592		
		4,00	179		17,8156	17,8156	
		3,00	321			18,6511	
		Signifikanz			,464	,082	,318

Figurale Wissensskala							
	Altersgruppen_New		N	Untergruppe für Alpha = 0.05.			
				1	2	3	
Student-Newman-Keuls-Prozedur <sup>a,b</sup>	-	1,00	35	19,6857			
		5,00	76	20,7632			
		2,00	142		23,3451		
		4,00	179		24,5140		
		3,00	321		25,2991		
		Signifikanz			,301	,146	

## Anhang B 1.2

### I.

Vorbedingungen einer parametrischen Varianzanalyse:

Statistiken der Gruppen (Muttersprachler D & F) und Überprüfung der Daten auf Normalverteilung Kolmogorov-Smirnov-Test & Schiefe- und Kurtosiskennwerte

Muttersprachler-D		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	663	663	663	663
	Fehlend	0	0	0	0
Mittelwert		21,1433	18,0543	24,6802	63,8778
Standardabweichung		5,86857	5,41650	6,74214	16,45736
Schiefe		,054	,188	-,048	,095
Standardfehler der Schiefe		,095	,095	,095	,095
Kurtosis		-,367	-,297	-,456	-,362
Standardfehler der Kurtosis		,190	,190	,190	,190
Minimum		5,00	3,00	6,00	21,00
Maximum		37,00	34,00	41,00	108,00
Kolmogorov-Smirnov-Anpassungstest					
Kolmogorov-Smirnov-Z		1,581	1,676	1,127	1,034
Asymptotische Signifikanz (2-seitig)		,014	,007	,158	,236

Muttersprachler-F		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	87	87	87	87
	Fehlend	0	0	0	0
Mittelwert		15,7816	14,0230	19,2644	49,0690
Standardabweichung		6,38389	5,86273	7,38077	18,38970
Schiefe		,601	,631	,185	,465
Standardfehler der Schiefe		,258	,258	,258	,258
Kurtosis		,162	,236	-,816	-,105
Standardfehler der Kurtosis		,511	,511	,511	,511
Minimum		4,00	3,00	5,00	14,00
Maximum		33,00	31,00	36,00	97,00
Kolmogorov-Smirnov-Anpassungstest					
Kolmogorov-Smirnov-Z		,956	1,259	,629	,686

Muttersprachler-F		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	87	87	87	87
	Fehlend	0	0	0	0
Mittelwert		15,7816	14,0230	19,2644	49,0690
Standardabweichung		6,38389	5,86273	7,38077	18,38970
Schiefe		,601	,631	,185	,465
Standardfehler der Schiefe		,258	,258	,258	,258
Kurtosis		,162	,236	-,816	-,105
Standardfehler der Kurtosis		,511	,511	,511	,511
Minimum		4,00	3,00	5,00	14,00
Asymptotische Signifikanz (2-seitig)		,320	,084	,824	,735

## Anhang B 1.3

### I.

Vorbedingungen einer parametrischen Varianzanalyse:

Statistiken der Gruppen (Männlich und Weiblich)

Überprüfung der Daten auf Normalverteilung Kolmogorov-Smirnov-Test & Schiefe- und Kurtosiskennwerte

Mann		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	166	166	166	166
	Fehlend	0	0	0	0
Mittelwert		23,8614	20,7108	27,8012	72,3735
Standardabweichung		6,26767	6,14675	7,37335	18,23098
Schiefe		-,310	-,212	-,618	-,433
Standardfehler der Schiefe		,188	,188	,188	,188
Kurtosis		-,404	-,526	-,009	-,127
Standardfehler der Kurtosis		,375	,375	,375	,375
Minimum		7,00	4,00	6,00	21,00
Maximum		37,00	34,00	41,00	108,00

	Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
Kolmogorov-Smirnov-Z	1,022	1,220	1,200	,818
Asymptotische Signifikanz (2-seitig)	,247	,102	,112	,515

Frau		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	587	587	587	587
	Fehlend	0	0	0	0
Mittelwert		19,5451	16,6831	22,9574	59,1857
Standardabweichung		5,81946	5,12657	6,57354	15,97976
Schiefe		,008	,102	-,083	-,015
Standardfehler der Schiefe		,101	,101	,101	,101
Kurtosis		-,268	-,253	-,380	-,257
Standardfehler der Kurtosis		,201	,201	,201	,201
Minimum		4,00	3,00	5,00	14,00
Maximum		36,00	33,00	40,00	103,00
Kolmogorov-Smirnov-Z		1,228	1,140	1,134	,704
Asymptotische Signifikanz (2-seitig)		,098	,149	,152	,705

## Anhang B 1.4

Vorbedingungen einer parametrischen Varianzanalyse:

- I. Test der Homogenität der Varianzen: Levene Test
- II. Statistiken getrennt nach Bildungsgruppen und Überprüfung der Daten auf Normalverteilung Kolmogorov-Smirnov-Test & Schiefe- und Kurtosiskennwerte
- III. Post-Hoc-Mehrfachvergleiche:  
Scheffé Prozedur

### I.

Test auf Varianzhomogenität

**Test der Homogenität der Varianzen**

	Levene-Statistik	df1	df2	Signifikanz
Verbale Wissensskala	1,673	5	738	,139
Numerische Wissensskala	3,591	5	738	,003
Figurale Wissensskala	,763	5	738	,576
Gesamte Wissensskala	2,856	5	738	,015

**II.**

Statistiken getrennt nach Bildungsgruppen

Überprüfung der Daten auf Normalverteilung: Kolmogorov-Smirnov-Test &amp; Schiefe- und Kurtosiskennwerte

**Bildungsgruppe 1: Hauptschulabschluss**

		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	12	12	12	12
	Fehlend	0	0	0	0
Mittelwert		15,9167	12,5833	17,9167	46,4167
Standardabweichung		4,50168	5,50138	5,33357	14,41248
Schiefe		-,688	,262	-,695	-,609
Standardfehler der Schiefe		,637	,637	,637	,637
Kurtosis		-,162	,575	-,561	-,027
Standardfehler der Kurtosis		1,232	1,232	1,232	1,232
Minimum		8,00	4,00	9,00	21,00
Maximum		22,00	24,00	25,00	69,00

**Kolmogorov-Smirnov-Anpassungstest**

	Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
Kolmogorov-Smirnov-Z	,584	,641	,630	,635
Asymptotische Signifikanz (2-seitig)	,885	,806	,823	,814

**Bildungsgruppe 2: Hauptschulabschluss mit abgeschlossener Berufsausbildung**

		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	14	14	14	14
	Fehlend	0	0	0	0
Mittelwert		13,1429	11,7857	13,0714	38,0000
Standardabweichung		5,69557	4,83860	5,25451	14,39551
Schiefe		,502	,319	,666	,639
Standardfehler der Schiefe		,597	,597	,597	,597
Kurtosis		-,824	-1,488	-,351	-1,069
Standardfehler der Kurtosis		1,154	1,154	1,154	1,154
Minimum		6,00	5,00	6,00	21,00
Maximum		24,00	19,00	23,00	64,00



**Kolmogorov-Smirnov-Anpassungstest**

	Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
Kolmogorov-Smirnov-Z	,549	,806	,570	,876
Asymptotische Signifikanz (2-seitig)	,924	,535	,902	,427

**Bildungsgruppe 3: Realschulabschluss**

		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	201	201	201	201
	Fehlend	0	0	0	0
Mittelwert		16,5871	13,7960	19,4876	49,8706
Standardabweichung		4,97832	4,18726	5,74640	12,92839
Schiefe		,066	,299	,204	,171
Standardfehler der Schiefe		,172	,172	,172	,172
Kurtosis		,108	,512	,043	,111
Standardfehler der Kurtosis		,341	,341	,341	,341
Minimum		4,00	3,00	5,00	14,00
Maximum		31,00	28,00	37,00	89,00

**Kolmogorov-Smirnov-Anpassungstest**

	Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
Kolmogorov-Smirnov-Z	1,063	1,104	1,008	,980
Asymptotische Signifikanz (2-seitig)	,209	,175	,261	,292

**Bildungsgruppe 4: Fachabitur**

		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	112	112	112	112
	Fehlend	0	0	0	0
Mittelwert		18,7946	15,8393	22,4732	57,1071
Standardabweichung		4,91576	4,11666	5,57821	12,53417
Schiefe		,234	-,004	,163	,030
Standardfehler der Schiefe		,228	,228	,228	,228
Kurtosis		,421	,067	-,172	,780
Standardfehler der Kurtosis		,453	,453	,453	,453
Minimum		7,00	4,00	10,00	21,00
Maximum		32,00	26,00	36,00	92,00

**Kolmogorov-Smirnov-Anpassungstest**

	Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
Kolmogorov-Smirnov-Z	,832	,657	,891	,909
Asymptotische Signifikanz (2-seitig)	,493	,781	,405	,381

**Bildungsgruppe 5: Abiturienten**

		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	107	107	107	107
	Fehlend	0	0	0	0
Mittelwert		21,3084	19,2991	26,5607	67,1682
Standardabweichung		4,75516	4,53958	5,72347	12,64649
Schiefe		,200	,165	-,156	,102
Standardfehler der Schiefe		,234	,234	,234	,234
Kurtosis		-,648	-,259	-,363	-,626
Standardfehler der Kurtosis		,463	,463	,463	,463
Minimum		11,00	9,00	14,00	38,00
Maximum		34,00	31,00	40,00	100,00

**Kolmogorov-Smirnov-Anpassungstest**

	Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
Kolmogorov-Smirnov-Z	1,106	,864	,998	,915
Asymptotische Signifikanz (2-seitig)	,173	,445	,272	,373

**Bildungsgruppe 6: Abitur**

		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	298	298	298	298
	Fehlend	0	0	0	0
Mittelwert		24,0973	20,6779	27,5738	72,3490
Standardabweichung		5,64618	5,25353	6,28902	15,80220
Schiefe		-,389	-,128	-,443	-,369
Standardfehler der Schiefe		,141	,141	,141	,141
Kurtosis		,018	-,208	,221	,257
Standardfehler der Kurtosis		,281	,281	,281	,281
Minimum		7,00	5,00	8,00	25,00
Maximum		37,00	34,00	41,00	108,00

	Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
Kolmogorov-Smirnov-Z	1,154	1,001	1,191	,735
Asymptotische Signifikanz (2-seitig)	,139	,269	,117	,652

### III.

Post-Hoc-Mehrfachvergleiche nach der Scheffé Prozedur und homogene Untergruppenbildung nach Student-Newman-Keul

Bildungsgruppe (i)	Bildungsgruppe (j)	Mittlere Diffe- renz (I-J)	Standardfehler	Signifikanz
<b>Verbale Wissensskala</b>				
Hauptschulabschluss	Hauptschule mit Beruf	2,77381	2,05554	,873
	Realschulabschluss	-,67040	1,55273	,999
	Fachabitur	-2,87798	1,58711	,656
	Abitur	-8,18065*	1,53843	,000
	Abiturienten	-5,39174*	1,59069	,044
Hauptschule mit Beruf	Hauptschulabschluss	-2,77381	2,05554	,873
	Realschulabschluss	-3,44421	1,44428	,339
	Fachabitur	-5,65179*	1,48118	,013
	Abitur	-10,95446*	1,42889	,000
	Abiturienten	-8,16555*	1,48502	,000
Realschulabschluss	Hauptschulabschluss	,67040	1,55273	,999
	Hauptschule mit Beruf	3,44421	1,44428	,339
	Fachabitur	-2,20758*	,61611	,026
	Abitur	-7,51025*	,47691	,000
	Abiturienten	-4,72135*	,62529	,000
Fachabitur	Hauptschulabschluss	2,87798	1,58711	,656
	Hauptschule mit Beruf	5,65179*	1,48118	,013
	Realschulabschluss	2,20758*	,61611	,026
	Abitur	-5,30267*	,57912	,000
	Abiturienten	-2,51377*	,70634	,028
Abitur	Hauptschulabschluss	8,18065*	1,53843	,000
	Hauptschule mit Beruf	10,95446*	1,42889	,000
	Realschulabschluss	7,51025*	,47691	,000
	Fachabitur	5,30267*	,57912	,000
	Abiturienten	2,78890*	,58887	,000
Abiturienten	Hauptschulabschluss	5,39174*	1,59069	,044
	Hauptschule mit Beruf	8,16555*	1,48502	,000

	Realschulabschluss	4,72135*	,62529	,000
	Fachabitur	2,51377*	,70634	,028
	Abitur	-2,78890*	,58887	,000
<b>Numerische Wissensskala</b>				
Hauptschulabschluss	Hauptschule mit Beruf	,79762	1,85486	,999
	Realschulabschluss	-1,21269	1,40114	,980
	Fachabitur	-3,25595	1,43216	,397
	Abitur	-8,09452*	1,38823	,000
	Abiturienten	-6,71573*	1,43539	,001
Hauptschule mit Beruf	Hauptschulabschluss	-,79762	1,85486	,999
	Realschulabschluss	-2,01031	1,30328	,794
	Fachabitur	-4,05357	1,33657	,103
	Abitur	-8,89214*	1,28939	,000
	Abiturienten	-7,51335*	1,34004	,000
Realschulabschluss	Hauptschulabschluss	1,21269	1,40114	,980
	Hauptschule mit Beruf	2,01031	1,30328	,794
	Fachabitur	-2,04327*	,55596	,020
	Abitur	-6,88183*	,43035	,000
	Abiturienten	-5,50305*	,56424	,000
Fachabitur	Hauptschulabschluss	3,25595	1,43216	,397
	Hauptschule mit Beruf	4,05357	1,33657	,103
	Realschulabschluss	2,04327*	,55596	,020
	Abitur	-4,83857*	,52258	,000
	Abiturienten	-3,45978*	,63738	,000
Abitur	Hauptschulabschluss	8,09452*	1,38823	,000
	Hauptschule mit Beruf	8,89214*	1,28939	,000
	Realschulabschluss	6,88183*	,43035	,000
	Fachabitur	4,83857*	,52258	,000
	Abiturienten	1,37879	,53138	,243
Abiturienten	Hauptschulabschluss	6,71573*	1,43539	,001
	Hauptschule mit Beruf	7,51335*	1,34004	,000
	Realschulabschluss	5,50305*	,56424	,000
	Fachabitur	3,45978*	,63738	,000
	Abitur	-1,37879	,53138	,243
<b>Figurale Wissensskala</b>				
Hauptschulabschluss	Hauptschule mit Beruf	4,84524	2,33274	,506
	Realschulabschluss	-1,57090	1,76212	,977
	Fachabitur	-4,55655	1,80113	,271
	Abitur	-9,65716*	1,74589	,000
	Abiturienten	-8,64408*	1,80520	,000
Hauptschule mit Beruf	Hauptschulabschluss	-4,84524	2,33274	,506

	Realschulabschluss	-6,41613*	1,63905	,010
	Fachabitur	-9,40179*	1,68092	,000
	Abitur	-14,50240*	1,62158	,000
	Abiturienten	-13,48932*	1,68527	,000
Realschulabschluss	Hauptschulabschluss	1,57090	1,76212	,977
	Hauptschule mit Beruf	6,41613*	1,63905	,010
	Fachabitur	-2,98565*	,69920	,003
	Abitur	-8,08626*	,54123	,000
	Abiturienten	-7,07319*	,70961	,000
Fachabitur	Hauptschulabschluss	4,55655	1,80113	,271
	Hauptschule mit Beruf	9,40179*	1,68092	,000
	Realschulabschluss	2,98565*	,69920	,003
	Abitur	-5,10061*	,65722	,000
	Abiturienten	-4,08753*	,80160	,000
Abitur	Hauptschulabschluss	9,65716*	1,74589	,000
	Hauptschule mit Beruf	14,50240*	1,62158	,000
	Realschulabschluss	8,08626*	,54123	,000
	Fachabitur	5,10061*	,65722	,000
	Abiturienten	1,01308	,66828	,806
Abiturienten	Hauptschulabschluss	8,64408*	1,80520	,000
	Hauptschule mit Beruf	13,48932*	1,68527	,000
	Realschulabschluss	7,07319*	,70961	,000
	Fachabitur	4,08753*	,80160	,000
	Abitur	-1,01308	,66828	,806
<b>Gesamte Wissenskala</b>				
Hauptschulabschluss	Hauptschule mit Beruf	8,41667	5,55147	,806
	Realschulabschluss	-3,45398	4,19350	,984
	Fachabitur	-10,69048	4,28634	,287
	Abitur	-25,93233*	4,15487	,000
	Abiturienten	-20,75156*	4,29602	,000
Hauptschule mit Beruf	Hauptschulabschluss	-8,41667	5,55147	,806
	Realschulabschluss	-11,87065	3,90061	,101
	Fachabitur	-19,10714*	4,00026	,000
	Abitur	-34,34899*	3,85905	,000
	Abiturienten	-29,16822*	4,01063	,000
Realschulabschluss	Hauptschulabschluss	3,45398	4,19350	,984
	Hauptschule mit Beruf	11,87065	3,90061	,101
	Fachabitur	-7,23650*	1,66395	,002
	Abitur	-22,47835*	1,28801	,000
	Abiturienten	-17,29758*	1,68873	,000
Fachabitur	Hauptschulabschluss	10,69048	4,28634	,287

	Hauptschule mit Beruf	19,10714*	4,00026	,000
	Realschulabschluss	7,23650*	1,66395	,002
	Abitur	-15,24185*	1,56405	,000
	Abiturienten	-10,06108*	1,90764	,000
Abitur	Hauptschulabschluss	25,93233*	4,15487	,000
	Hauptschule mit Beruf	34,34899*	3,85905	,000
	Realschulabschluss	22,47835*	1,28801	,000
	Fachabitur	15,24185*	1,56405	,000
	Abiturienten	5,18077	1,59039	,061
Abiturienten	Hauptschulabschluss	20,75156*	4,29602	,000
	Hauptschule mit Beruf	29,16822*	4,01063	,000
	Realschulabschluss	17,29758*	1,68873	,000
	Fachabitur	10,06108*	1,90764	,000
	Abitur	-5,18077	1,59039	,061

## Anhang B 1.5

### I.

Vorbedingungen einer parametrischen Varianzanalyse:

Statistiken der Gymnasiasten und Nicht Gymnasiasten und Überprüfung der Daten auf Normalverteilung: Kolmogorov-Smirnov-Test & Schiefe- und Kurtosiskennwerte

Nicht-Gymnasiasten		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	227	227	227	227
	Fehlend	0	0	0	0
Mittelwert		16,3392	13,6079	19,0088	48,9559
Standardabweichung		5,04795	4,31611	5,88608	13,35916
Schiefe		,033	,241	,151	,084
Standardfehler der Schiefe		,162	,162	,162	,162
Kurtosis		-,026	,330	-,045	-,032
Standardfehler der Kurtosis		,322	,322	,322	,322
Minimum		4,00	3,00	5,00	14,00
Maximum		31,00	28,00	37,00	89,00

### Kolmogorov-Smirnov-Anpassungstest

	Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
Kolmogorov-Smirnov-Z	1,090	1,027	,905	,876
Asymptotische Signifikanz (2-seitig)	,186	,242	,386	,426

Gymnasiasten		Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
N	Gültig	517	517	517	517
	Fehlend	0	0	0	0
Mittelwert		22,3714	19,3443	26,2592	67,9749
Standardabweichung		5,73793	5,24011	6,34952	15,72256
Schiefe		-,067	,064	-,210	-,079
Standardfehler der Schiefe		,107	,107	,107	,107
Kurtosis		-,379	-,271	-,252	-,106
Standardfehler der Kurtosis		,214	,214	,214	,214
Minimum		7,00	4,00	8,00	21,00
Maximum		37,00	34,00	41,00	108,00

#### Kolmogorov-Smirnov-Anpassungstest

	Verbale Wissensskala	Numerische Wissensskala	Figurale Wissensskala	Gesamte Wissensskala
Kolmogorov-Smirnov-Z	1,134	1,277	1,191	,805
Asymptotische Signifikanz (2-seitig)	,153	,077	,117	,537

## Anhang B: Ergebnisse der Forschungsfragen- Teil 2

### Anhang B 2.1

#### I.

Statistiken und Prüfung der Daten auf Normalverteilung: Kolmogorov-Smirnov-Test & Schiefe- und Kurtosiskennwerte

	N	Mittelwert	Schiefe		Kurtosis	
	Statistik	Statistik	Statistik	Standardfehler	Statistik	Standardfehler
Note Deutsch	715	2,619	,071	,091	-,024	,183
Note Englisch	708	2,710	,010	,092	-,200	,183
Note Mathematik	716	2,960	-,084	,091	-,587	,182
Note Kunst	635	2,008	,522	,097	,874	,194
Note Musik	529	1,963	,461	,106	,052	,212
Notendurchschnitt	676	2,555	-,418	,094	-,224	,188
Gültige Werte (Listenweise)	461					

**Kolmogorov-Smirnov-Anpassungstest**

	Note Deutsch	Note Englisch	Note Mathematik	Note Kunst	Note Musik	Notendurchschnitt
Kolmogorov-Smirnov-Z	6,238	5,634	4,379	6,956	5,899	2,412
Asymptotische Signifikanz (2-seitig)	,000	,000	,000	,000	,000	,000

**Anhang B 2.2****I.**

Statistiken und Prüfung der Daten auf Normalverteilung: Kolmogorov-Smirnov-Test

**Kolmogorov-Smirnov-Anpassungstest**

Gesamte Stichprobe		Note	V. Wissen	N.Wissen	F.Wissen	Gesamte Skala
N		105	105	105	105	105
Parameter der Normalverteilung <sup>a,b</sup>	Mittelwert	2,3430	21,7810	18,6571	25,1905	65,6286
	Standardabweichung	,62899	6,04467	5,77100	6,83190	17,06056
Extremste Diffe- renzen	Absolut	,111	,065	,076	,054	,048
	Positiv	,111	,050	,076	,054	,048
	Negativ	-,059	-,065	-,041	-,045	-,042
Kolmogorov-Smirnov-Z		1,135	,663	,782	,553	,492
Asymptotische Signifikanz (2-seitig)		,152	,771	,574	,920	,969
Gesundheits-und Krankenpflege		Note	V. Wissen	N.Wissen	F.Wissen	Gesamte Skala
N		87	88	88	88	88
Parameter der Normalverteilung <sup>a,b</sup>	Mittelwert	2,4691	21,7614	18,2273	24,7045	64,6932
	Standardabweichung	,60397	5,98656	5,81462	6,63179	16,82524
Extremste Diffe- renzen	Absolut	,077	,064	,096	,050	,061
	Positiv	,077	,064	,096	,050	,061
	Negativ	-,057	-,059	-,046	-,044	-,039
Kolmogorov-Smirnov-Z		,722	,601	,902	,473	,570
Asymptotische Signifikanz (2-seitig)		,674	,863	,390	,979	,901



Operationstechnische Assistenten		Note	V. Wissen	N.Wissen	F.Wissen	Gesamte Skala
N		18	18	18	18	18
Parameter der Normalverteilung <sup>a,b</sup>	Mittelwert	1,7333	21,9444	20,7222	27,6667	70,3333
	Standardabweichung	,31623	6,33823	5,02705	7,27607	17,44234
Extremste Differenzen	Absolut	,194	,115	,156	,102	,145
	Positiv	,194	,086	,091	,078	,145
	Negativ	-,125	-,115	-,156	-,102	-,117
Kolmogorov-Smirnov-Z		,824	,486	,661	,431	,614
Asymptotische Signifikanz (2-seitig)		,505	,972	,774	,992	,846

## Anhang B: Ergebnisse der Forschungsfragen- Teil 3

### Anhang B 3.2

#### I.

Konfirmatorische Faktorenanalyse (LISREL 8.30) mit dem Grundmodul und sechs Wissensaggregaten (completely standardized solution)- Faktorladungsmatrix

Aggregat	Verbal	Numerisch	Figural	GF	GC`	Wissen Verbal	Wissen Numerisch	Wissen Figural	GC
SE	.41			.64					
AN	.29			.79					
GE	.27			.68					
RA		.30		.70					
ZR		.40		.68					
RZ		.37		.72					
FA			.29	.58					
WA			.62	.52					
MA			.25	.60					
Geo_V					.18	.70			
Geo_N					.18		.67		

Geo_F					.25			.65	
Nat_V					.12	.62			
Nat_N					-.05		.49		
Nat_F					.07			.64	
Mat_V					-.23	.57			
Mat_N					-.23		.80		
Mat_F					-.17			.56	
Allt_V					.29	.59			
Allt_N					.18		.69		
Allt_F					.20			.49	
Wirt_V					.36	.56			
Wirt_N					.12		.46		
Wirt_F					.20			.59	
Kult_V					.19	.59			
Kult_N					.00		.55		
Kult_F					.28			.49	
GC						1	1	1	

Anmerkung: SE= Satzergänzung; AN= Analogien; GE= Gemeinsamkeiten; RA= Rechenaufgaben; ZR= Zahlenreihen; RZ= Rechenzeichen; FA= Figurenauswahl; WA= Würfelaufgaben; MA= Matrizen; Geo.= Geografie/Geschichte; Nat.= Naturwissenschaften; Mat.= Mathematik; Allt.= Alltagswissen; Wirt.= Wirtschaft; Kult.= Kultur; V= Verbale Kodierung; N= Numerische Kodierung; F= Figurale Kodierung.

## Anhang C: Diskussion der Forschungsfragen- Teil 1

### Anhang C 1.1:

#### I.

Regressionsanalyse- Alter, Bildungsniveau und Wissen

#### Regressionanalyse: Einflußvariable Altersgruppen

##### Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,060 <sup>a</sup>	,004	,002	17,35377

a. Einflußvariablen : (Konstante), Altersgruppen

##### Koeffizienten<sup>a</sup>

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Sig.
		RegressionskoeffizientB	Standardfehler	Beta		
1	(Konstante)	58,788	2,104		27,947	,000
	Altersgruppen	1,047	,635	,060	1,647	,100

a. Abhängige Variable: Gesamte Wissensskala

#### Regressionanalyse: Einflußvariable Schulabschluss

##### Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,519 <sup>a</sup>	,269	,268	14,88329

a. Einflußvariablen : (Konstante), Schulabschluss

##### Koeffizienten<sup>a</sup>

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Sig.
		RegressionskoeffizientB	Standardfehler	Beta		
1	(Konstante)	28,604	2,102		13,609	,000
	Schulabschluss	7,749	,469	,519	16,538	,000

a. Abhängige Variable: Gesamte Wissensskala

**Multiple Regressionsanalyse****Regressionanalyse: Einflußvariable Bildungsniveau (G & N-G), Altersgruppen****R<sup>2</sup>=1=vollständige Erklärung der Varianz der abh.V.**

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,505 <sup>a</sup>	,255	,253	15,04049

a. Einflußvariablen : (Konstante), Bildungsniveau-G &amp; N-G, Altersgruppen

**Koeffizienten<sup>a</sup>**

Modell	Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Sig.	
	Regressionskoeffizient	Standardfehler	Beta			
1	(Konstante)	28,183	2,661		10,591	,000
	Altersgruppen	,597	,554	,034	1,076	,282
	Bildungsniveau Gymn. & N-Gymn	18,944	1,200	,502	15,793	,000

a. Abhängige Variable: Skala\_Gesamt\_New