



universität
wien

Diplomarbeit

Titel der Arbeit

Machbarkeitsstudie zum Untertest "Alltagswissen"
des AID-Gruppe: Erste Befunde zur
psychometrischen Güte und konvergenten Validität

Verfasserin

Rafaela Lorenz

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, im September 2012

Studienkennzahl: 298

Studienrichtung: Psychologie

Betreuer: Univ.-Prof. Dr. Mag. Klaus D. Kubinger

DANKSAGUNG

An erster Stelle möchte ich mich bei Herrn Univ.-Prof. Dr. Mag. Klaus D. Kubinger für die hilfreiche Betreuung der Diplomarbeit bedanken.

Meiner Subbetreuerin Magistra Hagenmüller spreche ich großen Dank aus, ohne ihre tatkräftige Unterstützung, die sich in der dauernden Erreichbarkeit per Email oder auch Telefon zeigte, wäre diese Diplomarbeit noch lange nicht fertig gestellt.

Bei Claudia Gamsjäger möchte ich für die Begleitung während des gesamten 2. Abschnittes bedanken, ohne sie wäre dieser nie so aufregend und humorvoll gewesen. Des Weiteren ein großes Dankeschön für die täglichen Motivationsanrufe während der Test- und Schreibphase und für das Korrekturlesen meiner Arbeit.

Meinem Lebensgefährten möchte ich für seine tägliche Unterstützung, angefangen beim Neuaufsetzen meines Computers bis hin zum Aushalten vieler Verzweiflungsanfälle danken.

Meinen Freundinnen und Freunden möchte ich für Ihre emotionale Unterstützung und ihre Geduld danken.

Last but not least möchte ich meiner Familie, allen voran meinen Eltern, Gerhard und Eva Lorenz danken, dass sie mich überhaupt auf diesen Weg geleitet haben und mir ermöglicht haben, einen Beruf zu ergreifen, der mich auch wirklich interessiert und in dem ich mich voll entfalten kann. Sie haben mich durch meine gesamte Ausbildungsphase begleitet und haben immer Interesse an meinem Studium gezeigt.

Meinen drei Schwestern, Tanja, Nicole und Denise Lorenz, möchte ich ebenso danken. Auch sie haben mich während des gesamten Studiums mit Rat und Tat unterstützt. Danke für das gemeinsame Lernen, Prüfen, für die Unterstützung während den Testungen für die Diplomarbeit und für das Anspornen nach jeder geschriebenen Seite.

ABSTRACT (DEUTSCH)

Der neu entwickelte Untertest der AID-Gruppe, dessen Entwicklung, Durchführung und Analysen in dieser Arbeit beschrieben sind, soll das Konstrukt *Alltagswissen* bei Kindern und Jugendlichen im Alter von 8 - 18 Jahren eindimensional erfassen. Die Testpersonen müssen dabei je 20 Multiple-Choice Items und 4 Items im Kästchenformat zu unterschiedlichsten Themengebieten, wie zum Beispiel Ernährungswissenschaften, Sexualität usw. innerhalb von zehn Minuten bearbeiten. Die Erhebung der 814 Testpersonen fand in Volks- und Hauptschulen, Polytechnischen Schulen, Gymnasien und Berufsbildenden Höheren Schulen in Niederösterreich und Oberösterreich statt. 211 der getesteten SchülerInnen wurden neben der Gruppentestung auch einzeln getestet, um deren Leistungen im Gruppen- und im Einzelverfahren, hierzu wurde der AID 3 (Kubinger und Holocher-Ertl, in Vorbereitung) verwendet, miteinander zu vergleichen.

Eine hohe Übereinstimmungsvalidität $r=0,81$ ist gegeben. Dies zeigt, dass ein Zusammenhang zwischen den Leistungen der SchülerInnen in der Individualtestung und der Gruppentestung besteht.

Die Rasch-Modell-Analysen, deren Modellgültigkeit mittels dem *Likelihood-ratio*-Test nach Andersen(1973) und dem z-Test nach Fischer und Scheiblechner (1970) überprüft wurde, zeigte, dass nach sukzessivem Ausschluss von 12% der Items, von einer *a-posteriori* Modellgültigkeit ausgegangen werden kann.

Schlüsselwörter: *Alltagswissen*, Rasch-Modell, Gruppenverfahren, AID 2, AID 3, Übereinstimmungsvalidität

ABSTRACT (ENGLISH)

The new developed subtest of the AID-group, which is described in this paper, should measure "*Everyday Knowledge*" in children and teenager aged between eight and eighteen. This test has been designed to be taken by groups. Each testee has to solve 24 items from different fields of knowledge such as sexuality and nutritional science within ten minutes. The census of 814 test persons had taken place in elementary schools, secondary modern schools, grammar schools, polytechnic institutes and other institutions of higher education in Upper and Lower Austria. 211 of the students had received a single testing in addition, to be compared to their results of the group testing. The analysis shows that there is a high concurrent validity between the one to one test condition and the group testing. Equal results are being achieved in both test conditions. Rasch-model-analysis indicates that the new constructed test shows *a-posteriori* model-validity after excluding 12% of the items which implies a one dimensional measurement of the construct "*Everyday Knowledge*".

Schlüsselwörter: *Everyday Knowledge*, Rasch Modell, test for groups, AID 2, AID 3, concurrent validity

INHALTSVERZEICHNIS

I. Einleitung	17
II. Theorie	19
1. Intelligenz	19
1.1 Begriffsdefinition und Intelligenztheorien	19
1.2 Intelligenzquotient vs. Profilinterpretation	21
2. Adaptives Intelligenz Diagnostikum 2 (AID2)	23
3. Testentwicklung	25
3.1. Antwortformat	25
3.2 Speed-Powertest vs. Powertest	27
3.3 Gruppen- vs. Individualtestungen	27
3.5 Itemformulierung	28
4. Testtheoretische Grundlagen	29
4.1 Item-Response-Theorie	29
4.2 Rasch-Modell	30
III. Empirischer Teil	33
5. Forschungshypothesen und Ziel der Untersuchung	33
6. Konstruktion des Untertest 1 „Alltagswissen“	34
6.1 Itemkonstruktion und Itempool	34

6.2 Geltungsbereich und Testdesign	36
6.3 Aufgabenmaterial und Antwortformat	36
6.3 Instruktion, Testvorgabe und Kodierung	39
7. Methode	43
7.1 Untersuchungsplanung.....	43
7.2 Durchführung der Untersuchung.....	44
7.3 Stichprobenbeschreibung.....	45
8. Ergebnisse.....	49
8.1 Analyse der 1. Hypothese.....	49
8.2 Analyse der 2. Hypothese.....	57
9. Diskussion und Ausblick	59
10. Zusammenfassung	63
IV. Literaturverzeichnis.....	65
V. Anhang.....	67

TABELLENVERZEICHNIS:

TABELLE 1:	HÄUFIGKEITEN DER POSITIONEN DER LÖSUNGEN DER MULTIPLE-CHOICE AUFGABENUNTER DEN ANTWORTMÖGLICHKEITEN	38
TABELLE 2:	AUFTEILUNG DER TESTHEFTE JE SCHULSTUFE	41
TABELLE 3:	ANGEDACHTE VERTEILUNG DER STICHPROBE	45
TABELLE 4:	VERTEILUNG DER STICHPROBE AUF DIE BUNDESLÄNDER (N=814)	46
TABELLE 5:	VERTEILUNG DER GETESTETEN SCHÜLERINNEN AUF DAS GESCHLECHT (N=814)	46
TABELLE 6:	VERTEILUNG DER GETESTETEN SCHÜLERINNEN AUF DIE SCHULSTUFE (N=814)	47
TABELLE 7:	VERTEILUNG DER GETESTETEN SCHÜLERINNEN AUF DIE SCHULFORMEN(N=814)	47
TABELLE 8:	VERTEILUNG DER GETESTETEN SCHÜLERINNEN AUF DIE MUTTERSPRACHE (N=814)	48
TABELLE 9:	VERTEILUNG DER GETESTETEN SCHÜLERINNEN AUF DAS ALTER NACH JAHREN (N=814)	48
TABELLE 10:	DIE ERGEBNISSE DER LRTS VOR DEM AUSSCHLUSS EINES ITEMS	50
TABELLE 11:	SIGNIFIKANTE ITEMS VOR DEM AUSSCHLUSS EINES ITEMS FÜR DAS TK <i>SCORE</i>	53
TABELLE 12:	SIGNIFIKANTE ITEMS DES TK <i>GESCHLECHT</i> VOR DEM AUSSCHLUSS EINES ITEMS...	53
TABELLE 13:	SIGNIFIKANTE ITEMS DES TK <i>SPRACHE</i> VOR DEM AUSSCHLUSS EINES ITEMS	54
TABELLE 14:	VERTEILUNG DER MUTTERSPRACHEN AUF DIE SCHULSTUFEN	60
TABELLE 15:	VERTEILUNG DER MUTTERSPRACHEN AUF DAS ALTER	61

ABBILDUNGSVERZEICHNIS:

ABBILDUNG 1:	BEISPIEL FÜR EINEN GRAFISCHEN MODELLTEST	31
ABBILDUNG 2:	BEISPIEL EINER MULTIPLE-CHOICE AUFGABE IM NEU KONSTRUIERTEN UNTERTEST	37
ABBILDUNG 3:	BEISPIEL EINEE ERGÄNZUNGSAUFGABE IM NEU KONSTRUIERTEN UNTERTEST	37
ABBILDUNG 4:	ALLGEMEINE INSTRUKTION FÜR DEN GESAMTEN GRUPPENTEST	39
ABBILDUNG 5:	ANLEITUNG FÜR DEN UNTERTEST 1 <i>ALLTAGSWISSEN</i> , SEITE 1	40
ABBILDUNG 6:	ANLEITUNG FÜR DEN UNTERTEST 1 <i>ALLTAGSWISSEN</i> , SEITE 2	41
ABBILDUNG 8:	AUSCHNITT DES TESTDESIGNS DES UNTERTESTS <i>ALLTAGSWISSEN</i>	42
ABBILDUNG 9:	GRAFISCHER MODELLTEST DES TK <i>SCORE</i> VOR DEM AUSSCHLUSS EINES ITEMS.	51
ABBILDUNG 10:	GRAFISCHER MODELLTEST DES TK <i>ALTER</i> VOR DEM AUSSCHLUSS EINES ITEMS.	51
ABBILDUNG 11:	GRAFISCHER MODELLTEST DES TK <i>GESCHLECHT</i> VOR DEM AUSSCHLUSS EINES ITEMS.	52
ABBILDUNG 12:	GRAFISCHER MODELLTEST DES TK <i>SPRACHE</i> VOR DEM AUSSCHLUSS EINES ITEMS.	52
ABBILDUNG 13:	GRAFISCHER MODELLTEST DES TK <i>SCORE</i> NACH DEM AUSSCHLUSS VON 13 ITEMS.....	55
ABBILDUNG 14:	GRAFISCHER MODELLTEST DES TK <i>ALTER</i> NACH DEM AUSSCHLUSS VON 13 ITEMS.....	55
ABBILDUNG 15:	GRAFISCHER MODELLTEST DES TK <i>GESCHLECHT</i> NACH DEM AUSSCHLUSS VON 13 ITEMS.....	56
ABBILDUNG 16:	GRAFISCHER MODELLTEST DES TK <i>SPRACHE</i> NACH DEM AUSSCHLUSS VON 13 ITEMS.....	56
ABBILDUNG 17:	DARSTELLUNG DER PEARSON-KORRELATION DER PERSONENPARAMETER	58

I. EINLEITUNG

Die vorliegende Arbeit wurde im Rahmen eines Forschungsprojekts des Arbeitsbereichs Psychologische Diagnostik, Institut für Angewandte Psychologie: Gesundheit, Entwicklung und Förderung, Fakultät für Psychologie, Universität Wien, unter der Leitung von Herrn Univ.-Prof. Dr. Mag. Klaus D. Kubinger durchgeführt. Projektkoordination und Ansprechpartnerin am Arbeitsbereich war Frau Magistra Hagenmüller. Ziel des Forschungsprojekts ist die Entwicklung einer Intelligenz-Testbatterie für Kinder und Jugendliche, die in der Gruppe vorgegeben werden kann, da vielerorts eine Gruppenvorgabe einer Individualtestung aus ökonomischen Gründen gerechter wird. Diese soll an den AID3 (Kubinger und Holocher-Ertl, in Vorbereitung), der bisher nur als Individualtestung vorgegeben werden kann, angelehnt sein. Diese Diplomarbeit beschäftigt sich mit der Entwicklung des Untertests 1, *Alltagswisse*. Es soll gelingen, dass der Untertest für Gruppen so konstruiert wird, dass sich die Leistungen der Testpersonen nicht von den Leistungen in einer Einzeltestung unterscheiden. Ebenso soll der neu konstruierte Untertest das Konstrukt *Alltagswissen* eindimensional messen und somit *Rasch*-Modell konforme Items enthalten.

Zu Anfang sollten alle Items sowohl für ein freies Antwortformat, als auch für ein Multiple-Choice Antwortformat geeignet sein, um diese zwei Antwortformate zusätzlich gegenüberstellen zu können. Aus zeitlicher Ressourcenknappheit wurde dann jedoch das Konzept geändert und 20 der 24 Item als Multiple-Choice Items konstruiert, die übrigen vier Items wurden im Kästchenformat vorgegeben..

In der nachfolgenden Arbeit sollen zu Anfang die Definition und die Frage nach der Messbarkeit des Konstrukts Intelligenz im Vordergrund stehen. Anschließend wird auf die Konstruktion eines Untertests und die verwendbaren Gestaltungsweisen ein Augenmerk gelegt. Weiters wird der AID 2 (Kubinger, 2009a), an den der neu konstruierte Untertest angelehnt werden soll, ausführlich beschrieben, am Ende der theoretischen Fundierung werden kurz jene statistischen Verfahren dargestellt, die im empirischen Teil der Arbeit verwendet wurden.

Der zweite, sogenannte empirische, Abschnitt handelt von der Konstruktion des Untertests, angefangen bei der Itemkonstruktion bis zur Untersuchungsplanung, -

durchführung und Darstellung der Ergebnisse. Die Hypothesen werden mit dem dichotomen logistischen Modell nach Rasch und der Pearson-Korrelation überprüft. Anschließend werden die Ergebnisse diskutiert, zusammengefasst und ein Ausblick für weitere Forschungen gegeben.

II. THEORIE

Im folgenden Kapitel erwarten den/die LeserIn zuerst ein kurzer Überblick wie das Konstrukt *Intelligenz* überhaupt definiert wird und wie sich dieses Konstrukt messen lässt. Anschließend folgt ein Einblick in den AID 2 (Kubinger, 2009a), und welche Gestaltungsweisen bei einer Testkonstruktion möglich sind. Zum Schluss des Kapitels folgt eine kurze Darstellung der Item-Response-Theorie und des dichotomen logistischen Modells nach Rasch.

1. INTELLIGENZ

Weber und Rammsayer (2012) setzen sich mit dem Thema, was Intelligenz überhaupt ist, auseinander. Im Grunde wird Intelligenz sehr unterschiedlich definiert. Implizite Theorien, die von psychologischen Laien aufgestellt werden, sind sehr kulturabhängig. So wird in der westlichen Kultur Intelligenz eher mit verbalen Fähigkeiten und sozialen Kompetenzen gleichgesetzt, in Afrika oder Asien wird Intelligenz von Laien über andere Fähigkeiten definiert. Auch in der Differentiellen Psychologie herrschen sehr unterschiedliche Meinungen, welche Fähigkeiten zu dem Konstrukt Intelligenz gehören, und wie man die Intelligenz messen soll.

Das folgende Kapitel soll einige einschlägige Definitionen und Theorien in der Intelligenzforschung darstellen und sich anschließend mit dem Begriff Intelligenzquotient (IQ) auseinandersetzen.

1.1 BEGRIFFSDEFINTION UND INTELLIGENZTHEORIEN

Alfred Binet und Théodore Simon veröffentlichten 1905 den ersten Intelligenztest im heutigen Sinn. Sie definierten Intelligenz als die Fähigkeit „gut urteilen, verstehen und denken zu können“ (Weber & Rammsayer, 2012, S. 177).

David Wechsler (1964, S.13), der Autor des HAWIK und HAWIE, definiert Intelligenz als „die zusammengesetzte oder globale Fähigkeit des Individuums, zweckvoll zu handeln, vernünftig zu denken und sich mit seiner Umgebung wirkungsvoll auseinander zu setzen.“

Die im AID 2 zugrunde liegende Definition von Intelligenz meint, Intelligenz sei „die Gesamtheit aller kognitiven Voraussetzungen die notwendig sind, um Wissen zu erwerben und Handlungskompetenzen zu entwickeln“ (Kubinger, 2009a, S.23). Häcker und Stapf (2009, S.475) sehen die Gemeinsamkeit der unterschiedlichen Intelligenzdefinitionen darin, dass Intelligenz die Fähigkeit bezeichnet, „sich in neuen Situationen auf Grund von Einsichten zurechtzufinden, oder Aufgaben mit Hilfe des Denkens zu lösen,...“.

Da sich der AID 2, in Kapitel 2 näher beleuchtet, an die *Investmenttheorie* von Cattell anlehnt, sei das *Modell der fluiden und kristallisierten Allgemeinen Intelligenz* und die damit verbundene Investmenttheorie von Cattell nachstehend kurz beschrieben, interessierte LeserInnen seien auf Cattell (1997) verwiesen.

Das *Modell der fluiden und kristallisierten Allgemeinen Intelligenz* von Horn und Cattell (1966) befasst sich vorwiegend mit zwei Faktoren, der fluiden und der kristallisierten Intelligenz. Die fluide Intelligenz ist jene, welche vorwiegend biologisch determiniert ist, und somit nur wenig kulturell beeinflusst ist. Diese Fähigkeit dient dazu, sich neuen Situationen anzupassen, und abstrakte Probleme zu lösen. Die kristallisierte Intelligenz beinhaltet erworbenes, erlerntes Wissen, wodurch sie kulturabhängig ist (vgl. Cattell, 1961; Süß, 2003; Weber & Rammsayer, 2012; Stemmler et al., 2011). Die *Investmenttheorie* von Cattell besagt, dass die kristallisierte Intelligenz das Resultat der fluiden Intelligenz, in idealen Lerngelegenheiten, ist (Süß, 2003).

Von der Beschreibung anderer herkömmlicher Intelligenztheorien wird Abstand genommen, da sich die Faktorenstruktur des AID 2 in keiner bekannten Intelligenztheorie wiederfinden lässt. Auffällig ist, dass sich die 1-Faktorentheorie von Spearman nicht bestätigen lässt (Kubinger, 2009a).

Häcker und Stapf (2009, S.478) definieren den Intelligenzquotienten als das „Maß für intellektuelle Leistungsfähigkeit einer Person“. Die Anfänge des Intelligenzquotienten finden sich bei Alfred Binet und Théodore Simon, die das Intelligenzalter (IA) als Maß für die Intelligenz einer Testperson festlegten. Für die Interpretation subtrahierten sie vom Lebensalter (LA) der Testperson das erfasste IA (Herle, 2003). Diese Berechnung wurde ein Jahrzehnt später von Wilhelm Stern modifiziert. Da er einige Nachteile, unter anderem keine Berücksichtigung, dass die Differenz zwischen LA und IA in anderen Altersstufen eine unterschiedliche Gewichtung beigemessen wird, entdeckte. Er veränderte die Formel, um dem Problem entgegenzuwirken, dahingehend: $IQ = (IA/LA) * 100$ (Herle, 2003). Wenn heute vom Intelligenzquotienten die Rede ist, ist zumeist der IQ nach Wechsler gemeint, dieser berücksichtigte erstmals nicht das Konzept von Binet sondern definierte den IQ als Abweichungsmaß $IQ = 100 + 15 * (x - M) / S$ (vgl. Herle, 2003; Kubinger, 2009b).

Wechsler definiert x als den resultierenden Rohwert einer Testperson, M sei der Mittelwert und SD die Standardabweichung, der in der zu vergleichenden Altersgruppe erreichten Rohwerte. Ein bedeutender Vorteil des IQ nach Wechsler liegt darin, dass er eine Klassifizierung nach ICD-10 der Weltgesundheitsorganisation erlaubt (Herle, 2003).

Kubinger (1996) beschreibt einige wichtige Kritikpunkte am IQ. Zum einen würde der IQ implizit davon ausgehen, dass eine mangelnd ausgebildete Fähigkeit, also Schwäche, durch eine leistungsfähigere Fähigkeit, gemeint ist damit eine beliebige Stärke, kompensiert werden kann. Zum anderen ist der IQ in der förderungsorientierten Diagnostik nicht zielführend.

Aus den eben genannten Gründen wird beim AID 2 kein Intelligenzquotient angegeben, sondern eine Profilinterpretation. Damit ist gemeint, dass die Leistungstiefen und Leistungshöhen einer Testperson für jeden Untertest in Beziehung gesetzt und interpretiert werden. Dadurch werden die oben genannten Kritikpunkte des IQ als Globalmaß umgangen. Durch die Profilinterpretation gibt es im Sinne der förderungsorientierten Diagnostik die Möglichkeit,

Maßnahmenvorschläge zur Nivellierung von Schwächen abzuleiten. (vgl. Kubinger, 2009b; Kubinger, 2003b; Bulheller, 2003)

Die folgende kurze Beschreibung des AID 2 sollte dem/der LeserIn lediglich einen Einblick gewähren, für eine detaillierte Beschreibung sei der/die LeserIn an Kubinger (2009a) verwiesen.

Der AID 2 ist eine Intelligenz-Testbatterie für Kinder und Jugendliche im Alter von 6,0 bis 15,11 Jahren. Sie soll die komplexen und basalen Kognitionen, also die Intelligenz, in einer Individualtestung erfassen (Kubinger, 2009a). Kubinger (2009a, S. 23) definiert Intelligenz als „die Gesamtheit aller kognitiven Voraussetzungen die notwendig sind, um Wissen zu erwerben und Handlungskompetenzen zu entwickeln“. Der Test setzt sich aus 11 Untertests und 3 Zusatztests zusammen, die in „manuell-visuellen“ und „verbal-akustischen“ Aufgabenstellungen unterschieden werden können. Die Testdauer ist zwischen 30 und 75 Minuten angesetzt. Für jeden Zusatztest werden zuzüglich 10-15 Minuten eingerechnet. Der AID 2 lehnt sich an das Testkonzept von Wechsler an, ein bedeutender Unterschied liegt jedoch in der adaptiven Testvorgabe. Bei acht der elf Unterstests (UT1, UT3, UT4, UT6, UT8, UT9, UT10, UT11) kommt es zur „adaptiven“ Aufgabenvorgabe und zwar gemäß dem sogenannten *branched-testing*. Die Vorteile des adaptiven Testens finden sich vor allem darin, dass die Testpersonen vorwiegend nur Aufgaben gestellt bekommen, die ihren Fähigkeiten entsprechen. Weitere Vorteile sind, dass die Messgenauigkeit trotz der geringeren Vorgabe von Aufgaben gleich hoch bleibt wie bei Tests mit konventioneller Aufgabenvorgabe und dass die Motivation des Kindes eher erhalten bleibt. Die Untertests sind in verschiedene Blöcke mit je fünf Items mit unterschiedlichem Schwierigkeitsgrad unterteilt. Jede Testperson startet bei einem altersadäquaten Block. Aufgrund der erbrachten Leistung wird dem Kind daraufhin ein Block mit leichteren, gleichschwierigen oder schwierigeren Items vorgegeben. Um auf der Schwierigkeit der Ausgangsposition zu bleiben müssen 2 bis 3 Items gelöst werden, bei 4 oder 5 Richtigen werden dem Kind schwierigere Items vorgegeben, bei keinem oder einem gelösten Item bekommt das Kind anschließend leichtere Items. Nach jedem Aufgabenblock der Untertests 1, 3, 6, 9 und 11 kann die Testperson noch befragt werden, ob diese nun leichtere, gleichschwierige oder schwierigere Items möchte. Damit ist eine „Polarisierung in die zwei typischen

Reaktionsweisen bei kognitiven Leistungsanforderungen nach Dweck (2000)“ möglich (Kubinger, 2009a, S.33).

Der AID 2 beinhaltet einige reine Power-Tests mit freiem Antwortformat. Die Auswertung des AID 2 erfolgt mittels eines detaillierten Intelligenzprofils, das die Leistungshöhen und -tiefen der Testperson in den einzelnen Untertests beschreibt. Ebenso kann der Range, die Streuung der Testergebnisse, und die Intelligenzquantität, die Testautoren definieren, hiermit die kognitive Mindestleistung, berechnet aus der niedrigsten Leistung der Untertests, dargestellt werden. Der AID2 ermöglicht zudem ein Screening von Teilleistungsstörungen. Des Weiteren ist eine Unterscheidung der Untertests in förderabhängige und förderunabhängige in Anlehnung an Cattells Investmenttheorie möglich. Zusätzlich gibt es noch ein Beiblatt, um die Arbeitshaltungen einer Testperson zu kategorisieren. Die Haupteinsatzmöglichkeiten sind in der Schulpsychologie, Klinischen Psychologie und Berufs- und Bildungsberatung (Kubinger, 2009a).

3. TESTENTWICKLUNG

Im nachstehenden Kapitel soll auf die verschiedenen Gestaltungsweisen, die bei einer Testkonstruktion beachtet werden müssen, eingegangen werden. Es sollen sowohl unterschiedliche Antwortformate mit Vor- und Nachteilen als auch unterschiedliche Vorgabearten (wie Gruppen-, Individualtestung, Speed- power-Tests, Papier-Bleistift- und Computertests) und allgemeinere Regeln zur Itemformulierung skizziert werden.

3.1. ANTWORTFORMAT

Das Antwortformat bestimmt wie eine Testperson auf ein Item reagieren, also antworten, kann. Prinzipiell wird zwischen einem freien Antwortformat und dem gebundenen Antwortformat unterschieden. Bei dem freien, auch offenes Antwortformat genannt, ist es der Testperson möglich die Antwort, zum Beispiel durch Schrift oder Zeichnungen, selbst zu formulieren (Seiwald, 2003).

Bei dem gebundenen Antwortformat kann man den Richtig-Falsch-Antworttyp, die Zuordnungsaufgabe, die Umordnungsaufgabe, die Check-List Aufgabe und die Mehrfach-Wahl-Aufgabe, oft auch Multiple-Choice Aufgabe genannt, unterscheiden. Ebenso gibt es bei der freien Aufgabenbeantwortung verschiedene Typen, die größte Bedeutung hat hier die Ergänzungsaufgabe (Lienert & Raatz, 1998).

Die nachfolgenden Darstellungen betreffen nur die Antwortformate, die zur Generierung des Untertests 1 herangezogen wurden, interessierte LeserInnen seien für genauere Informationen an Lienert & Raatz (1998) verwiesen.

3.1.1 Ergänzungsaufgaben

Beispielitem: Der wievielte Monat des Jahres ist der März? 3.

Der Ergänzungsaufgabentyp hat einen bislang noch nicht übertroffenen Vorteil, im Gegensatz zu den gebundenen Antwortformaten ist er fast frei vom Zufallseinfluss. Weiters wird durch diesen Aufgabentyp das Erraten der Lösung minimiert und es

genügt der Testperson kein reines Wiedererkennen der Lösung, sondern diese muss selbst produziert werden können. Nachteile liegen in der Verrechnungssicherheit, vor allem wenn es teilrichtige Antworten gibt, da sich hier jeder/jede TestleiterIn für „gelöst“ oder „nicht gelöst“ entscheiden muss. Ebenso sind Testpersonen, mit geringen Formulierungsfähigkeiten, benachteiligt (Lienert & Raatz, 1998).

3.1.2 Multiple-Choice-Aufgaben

Beispielitem: Von welchem Staat ist Paris die Hauptstadt?

1. Niederlande
2. England
3. Frankreich
4. Deutschland
5. Moldavien

Derzeit ist der MC-Aufgabentyp, der am häufigsten verwendet wird in der modernen Testtheorie (Lienert & Raatz, 1998). Die Verrechnungssicherheit ist bei MC-Aufgabentypen wesentlich höher, da hier der oder die TestleiterIn nur „gelöst“ oder „nicht gelöst“ verrechnet, und somit jeder Testleiter/ jede Testleiterin zum gleichen Ergebnis kommen sollte. Ein schwerwiegender Nachteil liegt allerdings in der Ratewahrscheinlichkeit. Mit der *a-priori* Ratewahrscheinlichkeit ist die Wahrscheinlichkeit gemeint, mit der eine Testperson die Lösung errät ohne seine/ihre Fähigkeiten zu benutzen, also im Grunde ohne die Frage gelesen zu haben. Bei der Annahme man gäbe der Testperson fünf Antwortmöglichkeiten vor, also eine Lösung und vier Distraktoren, so würde die *a-priori* Ratewahrscheinlichkeit bei 20% liegen. Diese könnte durch mehrere Lösungen pro Frage minimiert werden. Wenn zum Beispiel zwei aus fünf Antwortmöglichkeiten richtig sind, liegt die *a-priori* Ratewahrscheinlichkeit nur mehr bei 10%. Weiters spielt die Distraktorenerstellung eine große Rolle. Im Grunde sollen die Distraktoren jene Testpersonen, die nicht lösungskompetent sind, von der Lösung ablenken. Bei der Entwicklung der Distraktoren sollte deswegen immer darauf geachtet werden, dass sie teilweise auch aus richtigen Komponenten bestehen (Kubinger, 2009b).

3.2 SPEED-POWERTEST VS. POWERTEST

Bei einem Paper-Pencil-Gruppenverfahren ist eine zusätzliche Zeitbegrenzung der Aufgaben unbedingt notwendig, da es schier unmöglich ist, jeder Testperson die Zeit zu geben, die jede/jeder benötigt. Damit wird die Eindimensionalität des Tests eingeschränkt, da bei schlechten Testergebnissen nicht mehr eindeutig zu identifizieren ist ob es an einem langsamen Arbeitsstil oder zu geringen Fähigkeiten der Testperson gelegen ist. Die Zeitbegrenzung kann sowohl für jedes Item, als auch- wie bei den meisten Gruppentests- für eine Aufgabengruppe verhängt werden. Daraus entsteht der Nachteil, dass einige schnelle Testpersonen während der Testung immer wieder Wartezeiten ausgesetzt werden (Kubinger, 2009b).

Weiters stellt Kubinger (2009b) dar, dass das *Rasch*-Modell bei einem *Speed-and-Power*-Test sehr wohl gelten kann. In diesem Fall würde die Bearbeitungszeit adäquat gewählt worden sein, so dass trotz zeitlicher Beschränkung eindimensional gemessen wird.

3.3 GRUPPEN- VS. INDIVIDUALTESTUNGEN

Laut Kubinger (2009b) liegen die wichtigsten Vorteile einer Gruppentestung in der erhöhten Testleiterunabhängigkeit und der gesteigerten Ökonomie. Die Ökonomie betrifft vorwiegend das Testmaterial, die Anzahl der Testleiter, und die Zeitersparnis. Amelang und Schmid-Atzert (2006) sehen die Nachteile eines Gruppenverfahrens vor allem darin, dass ein gewisses Maß an Lesefertigkeit von den Testpersonen Voraussetzung für die Lösung eines Items ist. Somit werden Testpersonen, die Schwierigkeiten beim Lesen zeigen, automatisch benachteiligt. Des Weiteren führen sie als Nachteile an, dass sich Testpersonen gegenseitig stören könnten und dass die Gefahr besteht die Lösungen vom Nachbarn abzuschreiben. Als Vorteil der Einzeltestung führen sie die Möglichkeit einer differenzierteren Verhaltensbeobachtung an.

3.5 ITEMFORMULIERUNG

Laut Bühner (2011) sollte bei der Itemformulierung darauf geachtet werden, dass, im Sinne der Verständlichkeit, sowohl keine doppelten Verneinungen verwendet werden als auch keine Wörter verwendet werden, die der Zielpopulation nicht geläufig sind. Weiters soll darauf geachtet werden, dass jedes Item nur eine Frage enthält, also keine zwei sachlichen Inhalte in einem Item vermischt werden, ebenso sollten extreme Längen und Kürzel vermieden werden.

4. TESTTHEORETISCHE GRUNDLAGEN

Es folgt ein kurzer Überblick über die Item-Response-Theorie (IRT) und dem dichotomen logistischen Testmodell von Rasch, der Einfachheit halber als „Rasch-Modell“ bezeichnet wird.

4.1 ITEM-RESPONSE-THEORIE

In der Testtheorie gibt es zwei verschiedene Ansätze zur Konstruktion psychologischer Verfahren, die klassische Testtheorie und die probabilistische Testtheorie. Die meisten psychologischen Tests sind nach der klassischen Testtheorie konstruiert, wobei nach und nach immer mehr Tests nach der probabilistischen Testtheorie, also der Item-Response-Theorie (IRT) entwickelt werden. Die Vorteile der IRT im Vergleich zur klassischen Testtheorie sind zum einen, dass die Testkennwerte stichprobenunabhängig sind, zu anderen, dass die Eindimensionalität nachgewiesen werden kann und die essentiellen Grundannahmen empirisch überprüfbar sind und werden nicht einfach nur vorausgesetzt (vgl. Kubinger, 2003a; Mossbrugger & Hartig, 2003; Amelang, 2006; Rost, 2004).

Die Modelle der IRT beschäftigen sich damit, wovon die Lösungswahrscheinlichkeit einzelner Items abhängt. Es besteht die Annahme, dass den manifesten, also sichtbaren Verhaltensweisen der Testpersonen ein bis mehrere latente Eigenschaften zugrunde liegen. Wichtige Schlüsselbegriffe in der IRT sind der Itemparameter, welcher die Itemschwierigkeit angibt, der Personenparameter, welcher die Fähigkeitsausprägung der Testperson bezeichnet, der Trennschärfeparameter und der Rateparameter. Wichtige Modelle des IRT sind unter anderem das *1-Parameter-Logistische Modell (Rasch-Modell)*, das *Zwei-Parameter-Logistische-Modell*, das *Drei-Parameter-Logistische-Modell* und das *Linear-Logistische Testmodell*. Im nachfolgenden Abschnitt wird nur auf das *1-PL-Modell* eingegangen, interessierte LeserInnen seien auf Bühner, 2011 oder Kubinger, 2003a verwiesen (vgl. Kubinger, 2003a; Bühner, 2011; Kubinger, 1989).

4.2 RASCH-MODELL

Das *Rasch-Modell*(RM) liegt die Annahme zugrunde, dass die Lösungswahrscheinlichkeit eines Items, wenn man den Zufall außer Acht lässt, nur an der Itemschwierigkeit und der Fähigkeit der Personen liegt. Vorausgesetzt wird, dass die Lösungswahrscheinlichkeit unabhängig davon ist, ob vorangegangene Items gelöst oder nicht gelöst wurden. Dies wird als *lokale stochastische Unabhängigkeit* bezeichnet. Um das RM anwenden zu können müssen die bearbeiteten Aufgaben als „gelöst“ und „nicht gelöst“, also „0“ und „1“ kodiert sein. Wie die Formel 4.3 zeigt, beschreibt das RM die Wahrscheinlichkeit, dass die Testperson v ein bestimmtes Item i löst, unter Berücksichtigung des Itemparameters σ und des Personenparameters ξ (Kubinger, 2003a; Bühner, 2011; Kubinger, 1989).

$$P(1|\xi_v, \sigma_i) = \frac{e^{\xi_v - \sigma_i}}{1 + e^{\xi_v - \sigma_i}} \quad (4.3.1.)$$

Das bedeutet, je leichter ein Item ist, und je fähiger die Testperson, desto wahrscheinlicher wird die Testperson das Item lösen können. Aus der Formel kann ebenso abgeleitet werden, dass die Wahrscheinlichkeit, dass die Testperson zur Lösung gelangt, bei 50 liegt, wenn der Itemparameter und der Personenparameter gleich sind. Durch das RM wird gleichzeitig die Annahme, dass alle Items eine latente Eigenschaft (zum Beispiel: Reasoning) messen überprüft. Ein großer Vorteil des RM liegt darin, dass die Leistung zweier Testpersonen miteinander verglichen werden kann, ohne dass diese exakt die gleichen Items bearbeiten müssen, was wiederum eine wesentliche Prämisse für das adaptive Testen darstellt (Kubinger, 2003a; Bühner, 2011; Kubinger, 1989).

Die Gültigkeit des RM kann mit unterschiedlichen Methoden geprüft werden. Eine sehr einfache Art der Prüfung ist der *grafische Modelltest*, bei dieser grafischen Modellkontrolle werden die Itemparameter aus zwei Teilstichproben in einem Koordinatensystem aufgetragen, wenn das RM gültig ist, liegen alle Punkte auf

einer 45°-Geraden (Kubinger, 1989). Abbildung 1 zeigt einen grafischen Modelltest, bei dem einige Items deutlich von der 45°-Geraden abweichen.

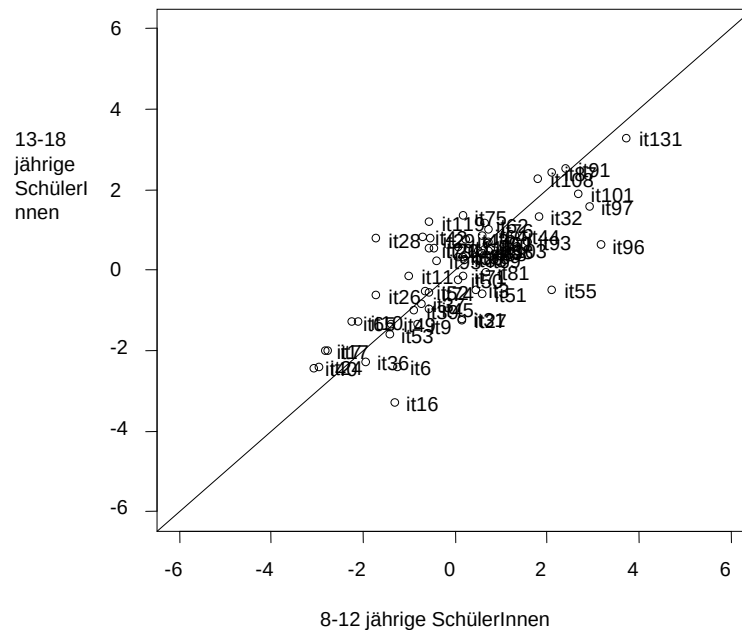


ABBILDUNG 1:BEISPIEL FÜR EINEN GRAFISCHEN MODELLTEST

Ein inferenzstatistische Modelltest ist der *Likelihood-Ratio-Test* nach Andersen (1973). Dieser dient zur Prüfung, ob die manifesten Daten „durch die in verschiedene Teilstichproben separat geschätzten Itemparameter“ signifikant besser erklärt werden können, „als durch die entsprechenden Schätzungen in der Stichprobe insgesamt“ (Kubinger, 2003a, S.418). Wichtig ist hier die Festlegung der Teilungskriterien, dabei wird zwischen einem internen Teilungskriterium (Score) und externen Teilungskriterien (zum Beispiel: Alter oder Geschlecht) unterschieden (Kubinger, 1989).

Der z-Test nach Fischer und Scheiblechner (1970, zitiert nach Kubinger 1989) wird ebenso zur Prüfung der Rasch-Modellgültigkeit herangezogen, dieser soll als Entscheidungshilfe dienen, welche Items aus dem Itempool ausgeschlossen werden sollen (Kubinger, 1989).

III. EMPIRISCHER TEIL

Im zweiten, empirischen Teil der Arbeit wird auf das Ziel der Untersuchung, die daraus abgeleiteten Hypothesen, die Konstruktion des Gruppentests, die Methode, gemeint sind hier Untersuchungsplanung, und – durchführung und Stichprobenbeschreibung eingegangen. Zudem werden die Ergebnisse der Auswertung, sowohl die *Rasch*-Modell Analysen als auch die Pearson-Korrelation dargestellt. Am Ende wird die Arbeit zusammengefasst und ein Ausblick für weitere Forschungen gegeben.

5. FORSCHUNGSHYPOTHESEN UND ZIEL DER UNTERSUCHUNG

Das Ziel der Untersuchung war es den Untertest *Alltagswissen* des AID 3 (Kubinger und Holoher-Ertl, in Vorbereitung) für die Gruppentestung geeignet zu konstruieren und dabei *Rasch*-Modell konforme Items zu erstellen, die dasselbe Konstrukt messen wie der bereits vorliegende Einzeltest. Im diesem Sinne wurden folgende Hypothesen aufgestellt:

H1₀: Die Items des Gruppenuntertest *Alltagswissen* sind (gegebenenfalls nach Ausschluss einzelner Items) konform mit dem dichotomen logistischen Testmodell von Rasch.

H2₀: Zwischen den Leistungen in der Gruppentestung und der Individualtestung besteht kein Zusammenhang.

6. KONSTRUKTION DES UNTERTEST 1 „ALLTAGSWISSEN“

Im nachstehenden Kapitel soll genauer auf die Itemkonstruktion, den Geltungsbereich und das Testdesign, Antwortformat und Aufgabenmaterial sowie Instruktion und Testvorgabe eingegangen werden.

6.1 ITEMKONSTRUKTION UND ITEMPOOL

Bei der Itemkonstruktion wurde zu Beginn der Begriff „Alltagswissen“ operationalisiert, dies geschah in Anlehnung an die Begriffsdefinition des AID 2. Laut Kubinger (2009a, S.9) ist Alltagswissen jene Fähigkeit, „sich Sachkenntnisse über Inhalte anzueignen, die in der heutigen Gesellschaft alltäglich sind.“

Die Itemkonstruktion des UT1 erfolgte zu Beginn intuitiv. Es wurden Themengebiete und Fragen gesammelt, die der Definition *Alltagswissen* nach Kubinger (2009b) entsprachen. Nach und nach wurde darauf aufbauend ein Regelkatalog erstellt. Die gesamte Itemkonstruktion erfolgte unter Beaufsichtigung und Hilfestellung von Magistra Hagemüller und Herrn Univ.-Prof. Dr. Mag. Klaus Kubinger.

Die Items wurden so konstruiert, dass jedes Item zum einen dem Gütekriterium Fairness entsprechen sollte, also zum einen keine Geschlechtsunterschiede zu erwarten sind, zum anderen, dass jedes Item 5-10 Jahre gültig sein sollte -vgl. hierzu die Deutsche Industrienorm DIN 33430 (vgl. Kubinger, 2009b). Ebenso sollten die Items relevantes Wissen, das den Alltag von Kindern und Jugendlichen prägt, beinhalten.

Folgende Themenbereiche sind in den Untertest 1 *Alltagswissen* für Gruppenvorgabe eingeflossen:

- Geschichte
- Geografie
- Sport
- Verkehr
- Wirtschaft

- neue Medien/ Computerfragen
- Biologie/ Ernährungswissenschaften
- Sexualität
- Chemie/ Physik/ Astronomie
- Kunst/ Musik/ Literatur
- Mathematik/ Finanzen/ Geometrie

Zusätzlich kam es zur Entwicklung eines Regelkatalogs, der neben den in Kapitel 4.6 erwähnten Punkten andere wichtige Regeln beinhaltet.

- Items sollten nicht nach Abkürzungen, Bezeichnungen und Jahreszahlen fragen.
- Zu vermeiden sind Items die mit „Wie heißt..“ beginnen, da dies reines „Nachschlagewissen“ darstellt und nicht zum relevanten Alltagswissen gezählt wird.
- Die Items dürfen nicht den Fremdwortschatz einer Testperson messen und sollten sich inhaltlich nicht mit den Items von UT 11 *Soziale und sachliche Folgerichtigkeit* des AID 2 (Kubinger, 2009a) oder AID 3 (Kubinger und Holocher-ertl, in Vorbereitung) überschneiden.
- Items sollten inhaltlich die Funktionalität von Dingen erfragen und öfter mit den Worten beginnen „Was bewirkt..?“, „Wer war..?“, „Wozu dienen..?“, „Zu welchem Zweck..?“.

Insgesamt entstand ein Itempool von 131 Items. 30 Items wurden ihrer Schwierigkeit betreffend als leicht, 23 als leicht-mittel, 35 als mittel, 19 als mittelschwierig und die letzten 24 als schwierig eingestuft. Von diesen 131 Items wurden 106 Items in das Testdesign eingebunden. Jede Testform sollte 24 Items beinhalten.

6.2 GELTUNGSBEREICH UND TESTDESIGN

Wie in Kapitel 2 beschrieben, hat der Untertest *Alltagswissen* im AID 2 einen Geltungsbereich von 6,0-15,9 Jahren. Da aber bei einem Gruppentest vorausgesetzt werden muss, dass die Testperson bereits lesen kann, wurde die untere Altersgrenze auf 8,0 Jahre angehoben. Im Zuge dessen wurde weiters beschlossen, die obere Altersgrenze auf 18 Jahre anzuheben, somit ergibt sich für den Gruppentest *Alltagswissen* ein Geltungsbereich von 8,0-18,11 Jahren. Um dem Gütekriterium der Zumutbarkeit und der Ökonomie gerecht zu werden, wurde der Gruppentest als Paper-Pencil-Verfahren konstruiert.

6.3 AUFGABENMATERIAL UND ANTWORTFORMAT

Es wurde die Entscheidung getroffen, dass die Items den Testpersonen in einem Testheft vorgelegt werden sollten, und die Testpersonen zusätzlich einen Antwortbogen zum Eintragen der Lösungen erhalten sollten. Die Testhefte gleichen sich alle im Aufbau, angefangen mit der allgemeinen Instruktion bis hin zu den einzelnen Untertests, denen allen eine kurze Anleitung samt Übungsaufgaben vorangeht. Insgesamt umfasst jedes Testheft 31 Seiten, wobei dem Untertest *Alltagswissen* elf Seiten eingeräumt werden. Der Antwortbogen beginnt mit der Erhebung der demographischen Variablen der Testpersonen. Angefangen mit dem Untertest *Alltagswissen* folgen die vier anderen Aufgabengruppen, am Ende liegt den Testpersonen noch ein kurzer Fragebogen vor.

Der Großteil der Items des Untertest *Alltagswissen*, je Testheft genau zwanzig, sind als Multiple-Choice-Aufgaben „1 aus 6“ gestaltet, das heißt, dass immer eine Lösung und fünf Distraktoren dargeboten werden. Dies erschien aus ökonomischen Gründen und zum Wohle der Verrechnungssicherheit am günstigsten. Die *a-priori*-Ratewahrscheinlichkeit liegt bei 16%. In Abbildung 2 ist dazu ein Beispielitem zu sehen.

Beispiel 1:

Wie nennt man eine männliche Ente?

- a) Eber
- b) Ochse
- c) Enter
- d) Erpel
- e) Este
- f) Wallach

ABBILDUNG 2: BEISPIEL EINER MULTIPLE-CHOICE AUFGABE IM NEU KONSTRUIERTEN UNTERTEST

Jeweils vier Items je Testheft sind vom Ergänzungsaufgabentypen, um die Verrechnungssicherheit dennoch zu gewährleisten, wurden nur Fragen ausgewählt, die mit Zahlen zu beantworten sind und nur eine Lösung zulassen. Den Testpersonen wurden diese Aufgaben wie in Abbildung 3 zu sehen im Kästchenformat vorgegeben.

Wie viele Tage hat eine Woche?

Tage

ABBILDUNG 3: BEISPIEL EINEE ERGÄNZUNGSAUFGABE IM NEU KONSTRUIERTEN UNTERTEST

Die Lösungen der Multiple-Choice-Items unter den Antwortmöglichkeiten sollten selten an der ersten Stelle positioniert sein, und sich über die anderen Positionen relativ gleichmäßig verteilen. In Tabelle 1 ist die Verteilung der Lösungen auf die Positionen ersichtlich.

TABELLE 1: HÄUFIGKEITEN DER POSITIONEN DER LÖSUNGEN DER MULTIPLE-CHOICE AUFGABENUNTER DEN ANTWORTMÖGLICHKEITEN

Testheft	1.Stelle(A)	2.Stelle(B)	3.Stelle(C)	4.Stelle(D)	5.Stelle(E)	6.Stelle(F)
1A,1B	2	3	5	4	4	2
2A,2B	1	3	5	4	3	4
3A,3B	1	3	4	4	4	4
4A,4B	2	2	5	3	4	4
5A,5B	2	3	4	4	4	3
6A,6B	1	5	3	4	4	3
7A,7B	1	3	3	4	5	4
8A,8B	2	4	4	2	5	3
Gesamt	12	26	33	29	33	27

6.3 INSTRUKTION, TESTVORGABE UND KODIERUNG

Die Testpersonen wurden instruiert einen Bleistift und Radiergummi zu verwenden, um Fehler korrigieren zu können. Vor Beginn der Testung wurde den Testpersonen eine allgemeine Anleitung vorgelesen, bei der die Testpersonen die Möglichkeit hatten diese selbstständig mitzulesen:

Allgemeine Anleitung
Dieses Testheft besteht aus verschiedenen Aufgabengruppen. Es gibt für jede Aufgabengruppe eine Anleitung und Beispielaufgaben, anhand derer dir erklärt wird, was zu tun ist.
Einige Aufgaben werden dir leicht fallen, andere schwer. Es ist nicht schlimm, wenn du eine Aufgabe nicht lösen kannst, versuche einfach die nächsten Aufgaben zu lösen.
Für jede Aufgabengruppe steht dir nur eine begrenzte Bearbeitungszeit zur Verfügung. Solltest du vor Ablauf der Zeit fertig sein, bleib bitte auf deinem Platz sitzen und störe die anderen nicht bei der Bearbeitung der Aufgaben. Du kannst die verbleibende Zeit auch dazu nutzen, deine Lösungen zu kontrollieren oder ausgelassene Aufgaben zu bearbeiten.
Bitte trage alle Lösungen auf dem Antwortbogen in die dafür vorgesehenen Kästchen ein oder kreuze die Lösung auf dem Antwortbogen pro Aufgabe an. Es wird dir vor jeder Aufgabengruppe genau erklärt, wie das funktioniert.
Solltest du eine Lösung korrigieren wollen, kannst du diese einfach ausradieren oder durchstreichen. Trage anschließend die neue Lösung ein. Sollte das nicht mehr klappen, so schreibe die Lösung neben die Aufgabennummer auf den Antwortbogen.
Bitte beachte , dass in diesem Testheft nichts geschrieben oder gezeichnet werden darf!
Solltest du während der Bearbeitung der Aufgaben Fragen haben oder die Anleitung nicht verstehen, dann melde dich bitte bei dem/der Testleiter(in) mit einem Handzeichen.
In diesem Testheft darfst du oft erst nach Aufforderung des/der Testleiter(in) umblättern. Achte auf die Kästchen am Ende jeder Seite. Steht in diesem Kästchen „Bitte Umblättern“, dann musst du nicht auf ein Zeichen des/der Testleiter(in) warten.
Fülle nun bitte deine persönlichen Daten (Alter, Muttersprache, Geschlecht) auf der ersten Seite des Antwortbogens aus. Danach erfolgt die Anleitung für die erste Aufgabengruppe.

ABBILDUNG 4: ALLGEMEINE INSTRUKTION FÜR DEN GESAMTEN GRUPPENTEST

Vor jeder Aufgabengruppe gab es eine Instruktion, die vom Testleiter vorgelesen und als Text vorgelegt wurde, zudem mussten die Testpersonen jeweils zwei

Beispielaufgaben bearbeiten, siehe dazu Abbildung 5 und 6. Die Aufgabengruppe des UT1 *Alltagswissen* umfasst pro Testversion 24 Items.

Anleitung: Untertest 1 – Alltagswissen
Auf den nächsten Seiten des Testhefts findest du einige Wissensfragen, die du beantworten sollst. Die Aufgaben werden dir dabei in zwei unterschiedlichen Arten gestellt, die dir nun in drei Beispielaufgaben erklärt werden.
Beispiel 1: Wie nennt man eine männliche Ente? a) Eber b) Ochse c) Enter d) Erpel e) Este f) Wallach
Du sollst bei dieser Art von Aufgaben die Lösung unter den 6 Antwortmöglichkeiten (a – f) finden. Beachte dabei, dass immer nur eine Antwortmöglichkeit richtig ist! In Beispiel 1 ist Erpel die Lösung. Man nennt eine männliche Ente Erpel. Daher ist auf dem Antwortbogen unter Beispiel 1 bereits der Buchstabe d als Lösung angekreuzt. Bitte kreuze bei solchen Aufgaben immer den Buchstaben der Lösung auf dem Antwortbogen an. Achte dabei auf die Nummerierung der Aufgaben und darauf, nicht die Zeile zu verlieren!
Beispiel 2: Wer ist kein Künstler? a) Maler b) Sänger c) Komponist d) Schriftsteller e) Musiker f) Pilot
Versuche das Beispiel 2 zu lösen, indem du auf dem Antwortbogen den Buchstaben der Lösung unter Beispiel 2 ankreuzt.

ABBILDUNG 5: ANLEITUNG FÜR DEN UNTERTEST 1 *ALLTAGSWISSEN*, SEITE 1

Anleitung: Untertest 1 – Alltagswissen

Beispiel 3:

Bei einigen Aufgaben stehen dir keine Antworten zur Auswahl bereit, sondern du musst die Lösung in ein Kästchen auf dem Antwortbogen eintragen.

Wie viele Tage hat eine Woche?

?

Tage

In **Beispiel 3** ist die Lösung 7. Daher ist in das Kästchen auf dem Antwortbogen unter **Beispiel 3** bereits die Zahl 7 eingetragen.

Bitte trage bei solchen Aufgaben immer die Lösung in das Kästchen auf dem Antwortbogen ein. Achte dabei auf die Nummerierung der Aufgaben und darauf, nicht das Kästchen zu vertauschen!

Hast du noch Fragen?

Auf den nächsten 4 Seiten findest du **24 Aufgaben**, die du genauso lösen sollst.

ABBILDUNG 6: ANLEITUNG FÜR DEN UNTERTEST 1 ALLTAGSWISSEN, SEITE 2

Insgesamt wurden 16 verschiedene Testhefte erstellt. Die nachstehende Tabelle stellt die Testhefte und deren Abfolge für die jeweilige Schulstufe dar. Die TestleiterInnen wurden angewiesen, die Testhefte in der vorgegebenen Abfolge auszuteilen, um ein Abschreiben seitens der SchülerInnen zu verhindern.

TABELLE 2: AUFTEILUNG DER TESTHEFTE JE SCHULSTUFE

Schulstufe	Abfolge der Testhefte
3.-4.	1a,1b,2a,2b
5.-6.	3a,3b,4a,4b
7.-9.	5a,5b,6a,6b
10.-11.	7a,7b,8a,8b

Die Abbildung 8 zeigt einen Ausschnitt des Testdesigns von Untertest *Alltagwissen*.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	24	25
													15		13	17	16	19	18				73	
	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
	K	K	K			K		K						K										K
	n	n	n	b	b	b	b	g	g	v	v	r	be	be	be	m	m	lit	lit	er	er	er	mat	mat
1a	5			1				10	2	3		4			6			13				14		17
2a		5			1				2		3	4	6	10		13			14		15			
3a			5			17			1	2						3				13			14	
4a							1		2		3						13							
5a									1											13				
6a									1								13							
7a									1															
8a									1															

ABBILDUNG 7: AUSCHNITT DES TESTDESIGNS DES UNTERTESTS *ALLTAGSWISSEN*

Zudem wurde aus den in Kapitel 3.2 genannten Gründen eine Zeitbeschränkung festgelegt. Um den Kindern und Jugendlichen genug Zeit zu lassen, wurde die Beschränkung auf 10 Minuten festgelegt. Die Kodierung der Antworten erfolgte mit „gelöst“ 1 und „nicht gelöst“ 0. Nach Ablauf der Bearbeitungszeit wurden die Testpersonen aufgefordert einen Strich unter die zuletzt bearbeitete Aufgabe zu machen. Die nicht bearbeiteten Items, also jene unterhalb des Strichs, wurden mit 99 als „fehlende Werte“ kodiert.

7. METHODE

7.1 UNTERSUCHUNGSPLANUNG

Nach Fertigstellung der Untersuchungsinstrumente wurde die Testung des UT *Alltagswissen* mit anderen Untertests der AID-Gruppe zusammengelegt. Die Gruppentestung sollte den UT 1 *Alltagswissen*, den UT 2 *Realitätssicherheit*, den UT 3 *Angewandtes Rechnen*, den UT 6 *Antonyme finden* und den UT 9 *Funktionen abstrahieren* enthalten. Die Stichprobenerhebung wurde mit der Erhebung einer Kollegin, die ebenso einen Untertest neu entwickelt hatte, zusammengelegt. Um zu verhindern, dass eine Speedkomponente auftaucht, musste jedem Kind ausreichend Zeit zur Verfügung gestellt werden. Zu diesem Zwecke wurde die Testdauer auf insgesamt 50 Minuten exklusive Instruktion festgelegt. Je Untertest wurde den Testpersonen eine Bearbeitungszeit von zehn Minuten zugesprochen. Da ebenfalls untersucht werden soll, ob bei einer Individualtestung und einer Gruppentestung dieselben Konstrukte erfasst werden, sollte ein Teil der Kinder und Jugendlichen zusätzlich mit dem AID 3 (Kubinger und Holocher-Erlt, in Vorbereitung) einzeln getestet werden. Die Vorgabe sollte balanciert erfolgen, damit ist gemeint, dass ein Teil der Kinder vor der Gruppentestung der Individualtestung unterzogen wird, der andere Teil nach der Gruppentestung. Für die Individualtestungen wurden 40 Minuten eingeplant. In dieser sollten der UT 1 *Alltagswissen*, der UT 2 *Realitätssicherheit*, der UT 3 *Angewandtes Rechnen*, der UT 6 *Synonyme finden*, UT 6a *Antonyme finden* und der UT 9 *Funktionen abstrahieren* durchgeführt werden. Gegen Ende Oktober sollten schriftliche Ergebnisberichte an die Erziehungsberechtigten der Kinder und Jugendlichen geschickt werden. Die Testungen sollten ca. nach 1 ½ Monate abgeschlossen werden können.

7.2 DURCHFÜHRUNG DER UNTERSUCHUNG

Nachdem alle Vorbereitungen für die Testung abgeschlossen waren, wurde Kontakt zu Schulen aufgenommen, und um eine Bewilligung beim Landesschulrat angesucht, so dass ein schneller Start der Testungen möglich war. Nach telefonischen und persönlichen Gesprächen mit den Direktoren der Schule erklärten sich nachstehende Schulen bereit an der Untersuchung teilzunehmen:

Zwei Hauptschulen, zwei Volksschulen, eine HTL, ein Gymnasium und eine Polytechnische Schule in Oberösterreich. In Niederösterreich nahmen zwei Volksschulen, ein BORG und ein Gymnasium teil.

Vom 23. Mai bis zum 28. Juni 2012 fanden in Schulen in Ober- und Niederösterreich sowohl die Individual- als auch die Gruppentestungen statt. Als zusätzliche TestleiterInnen fungierten DiplomandInnen des Arbeitsbereiches Psychologische Diagnostik. Diese mussten vorab einen Einschulungstermin absolvieren und bekamen eine genaue Testanleiteranleitung. Des weiteren wurden nur TestleiterInnen herangezogen die AID 2 zertifiziert sind oder zumindest den Theoriekurs bereits belegt hatten.

Die geplante Testdauer, bei der Gruppentestung ohne Instruktion, war bei beiden Testbedingungen gut gewählt.

7.3 STICHPROBENBESCHREIBUNG

Da auf Grund von nur begrenzten zeitlichen und finanziellen Ressourcen nicht die Gesamtheit der interessierenden Population untersucht werden kann, wird anstelle dieser eine Stichprobe zur Untersuchung herangezogen, deren Ergebnisse die Gesamtpopulation darstellen soll (Schäffer, 1996). Da die Stichprobe repräsentativ für Kinder zwischen 8-18 Jahre sein soll, wurde eine Erhebung an Schulen angedacht, da hier auch die Erreichbarkeit gegeben ist. Hierzu werden Volksschulen, Hauptschulen, Allgemein Höhere Schulen und Berufsbildende Höhere Schulen herangezogen.

Die angedachte Stichprobengröße lag bei 800 Kindern, wobei 210 dieser Kinder sowohl einzeln als auch in der Gruppe getestet werden sollten, und 500-600 Kinder nur einer Gruppentestung unterzogen werden sollten. Folgende Tabelle soll die angedachte Verteilung der Stichprobe zeigen.

TABELLE 3: ANGEDACHTE VERTEILUNG DER STICHPROBE

Schulstufe	Anzahl angestrebter Tp	Einzeltestungen	vorher	nachher
3.	90	35	18	17
4.	90	35	17	18
5.	90	35	18	17
6.	90	35	17	18
7.	60	24	12	12
8.	60	24	12	12
9.	60	24	12	12
10.	90	0	0	0
11.	90	0	0	0
12.	60	0	0	0

Insgesamt wurden 815 SchülerInnen zwischen 8-18 Jahren in Oberösterreich und Niederösterreich getestet. Die Stichprobe wurde hinsichtlich des Bundeslandes, der Schulstufe, des Alters, der Muttersprache, und des Geschlechts analysiert. Es gelang annähernd gleich viele Kinder und Jugendliche in Oberösterreich und Niederösterreich zu testen, siehe hierzu Tabelle 3.

TABELLE 4: VERTEILUNG DER STICHPROBE AUF DIE BUNDESLÄNDER (N=814)

	Häufigkeiten	Prozent
<i>NÖ</i>	433	53,2
<i>OÖ</i>	381	46,8
<i>Gesamt</i>	814	100

Wie Tabelle 4 zu entnehmen ist setzt sich die Stichprobe aus 54,5% Mädchen und 45,5% Jungen zusammen.

TABELLE 5: VERTEILUNG DER GETESTETEN SCHÜLERINNEN AUF DAS GESCHLECHT (N=814)

	Häufigkeit	Prozent
<i>männlich</i>	370	45,5
<i>weiblich</i>	444	54,5
<i>Gesamt</i>	814	100

Die angedachte Stichprobengröße wurde auf Grund des nahenden Schuljahresende nicht ganz erlangt, wie der Tabelle 5 zu entnehmen ist, wurden von der 3. Schulstufe bis zur 10. Schulstufe die angedachte Stichprobengröße erreicht. Ab der 11. Schulstufe ist dies bedauerlicherweise nicht mehr gelungen, die 12. Schulstufe konnte auf Grund der Reifeprüfung nicht getestet werden.

TABELLE 6: VERTEILUNG DER GETESTETEN SCHÜLERINNEN AUF DIE SCHULSTUFE (N=814)

	Häufigkeit	Prozent
3.Schulstufe	112	13,76
4.Schulstufe	105	12,9
5.Schulstufe	91	11,2
6.Schulstufe	124	15,2
7.Schulstufe	83	10,2
8.Schulstufe	70	8,6
9.Schulstufe	71	8,7
10.Schulstufe	92	11,3
11.Schulstufe	66	8,1
Gesamt	814	100

Die Stichprobenverteilung in Bezug auf die Schulformen ist Tabelle 6 zu entnehmen.

Über die Hälfte der Testpersonen besuchten zum Zeitpunkt der Testungen eine Allgemein Höhere Schule, etwas über ein Viertel die Volksschule. 11% der Stichprobe wurde in Hauptschulen erhoben, 7% in einer HTL, 2% der Stichprobe besuchten eine Polytechnische Schule.

TABELLE 7: VERTEILUNG DER GETESTETEN SCHÜLERINNEN AUF DIE SCHULFORMEN (N=814)

	Häufigkeit	Prozent
Volksschule	217	26,7
Hauptschule	91	11,2
AHS	432	53
Polytechnische	16	2
HTL	58	7,1
Gesamt	814	100

75,7% der Testpersonen gaben als Muttersprache Deutsch an, demzufolge hatten 24,3% der SchülerInnen eine andere Muttersprache. 10,3% davon fallen unter sonstige Muttersprachen, am häufigsten unter den anderen Muttersprachen waren mit 6,6% Kroatisch/Serbisch/Slowenisch und mit 5,4% Türkisch zu verzeichnen, siehe dazu Tabelle 7.

TABELLE 8:VERTEILUNG DER GETESTETEN SCHÜLERINNEN AUF DIE MUTTERSPRACHE (N=814)

	Häufigkeit	Prozent
<i>Deutsch</i>	616	75,7
<i>Türkisch</i>	44	5,4
<i>Kroatisch/Serbisch/Slowenisch</i>	54	6,6
<i>Slowakisch/Tschechisch</i>	2	0,2
<i>Polnisch</i>	4	0,5
<i>Russisch</i>	2	0,2
<i>Englisch</i>	2	0,2
<i>Französisch/Italienisch/Spanisch</i>	1	0,1
<i>Arabisch</i>	5	0,6
<i>sonstige</i>	84	10,3
<i>Gesamt</i>	814	100

Tabelle 8 zeigt die Analyse der Stichprobe nach Alter in Jahren. Zwischen 9 und 17 Jahren sind die Testpersonen gut verteilt, die wenigen Testpersonen mit 8 bzw. 18 oder 19 Jahren ergeben sich durch den späten Testzeitpunkt.

TABELLE 9:VERTEILUNG DER GETESTETEN SCHÜLERINNEN AUF DAS ALTER NACH JAHREN (N=814)

Alter	Häufigkeit	Prozent
8	16	2
9	86	10,6
10	104	12,8
11	102	12,5
12	119	14,6
13	83	10,2
14	81	10
15	72	8,8
16	81	10
17	67	8,2
18	2	0,2
19	1	0,1
Gesamt	814	100

8. ERGEBNISSE

Im folgenden Kapitel werden zuerst die Ergebnisse der *Rasch*-Modell Analysen des UT1 *Alltagswissen* dargelegt, anschließend werden die Ergebnisse zur konvergenten Validität beschrieben.

8.1 ANALYSE DER 1. HYPOTHESE

Um die *Rasch*-Modell-Analysen durchführen zu können wurde das Programmpaket eRm (Mair, Hatzinger & Meier, 2010), der frei verfügbaren Statistiksoftware R, genutzt.

Die Gültigkeit des *Rasch*-Modells wurde mittels des *Likelihood-Ratio*-Test (LRT) nach Andersen (1973) überprüft, dazu wurden nachstehende Teilungskriterien (TK) festgelegt.

Internes Teilungskriterium:

- Score (Rohscore $\leq$ Median vs. Rohscore $>$ Median)

Externes Teilungskriterium:

- Alter (≤ 9 Jahre vs. > 9 Jahre)
- Geschlecht (männlich vs. weiblich)
- Sprache (deutsche Muttersprache vs. andere Muttersprache)

Das Alter wurde genauso geteilt wie es im AID 2 (Kubinger, 2009a) gehandhabt wurde.

Im Falle eines signifikanten LRTs in einem der Teilungskriterien sollten die grafische Modellkontrolle und der z-Test nach Fischer und Scheiblechner (1970), beide in Kapitel 4.3 beschrieben, herangezogen werden, mittels derer nicht *Rasch*-Modell-konforme Items ausgemacht und sukzessive ausgeschieden werden sollten. Die Irrtumswahrscheinlichkeit, also das Risiko 1.Art, wurde mit $\alpha = 0.05$

festgelegt. Aufgrund der Durchführung von vier Signifikanztests wurde eine α -Anpassung auf $\alpha = 0.01$ vorgenommen.

Wie in Tabelle 9 ersichtlich zeigten sich im ersten Berechnungsdurchgang die LRTs der Teilungskriterien *Score*, *Geschlecht* und *Sprache* signifikant, der LRT des TK *Alter* fällt nicht signifikant aus. Einige Items sind für die Teilungskriterien nicht schätzbar, und zwar jene, die in einer der zwei Teilstichproben entweder von allen oder von keiner Testperson gelöst wurden. Beim LRT des TK *Score* flossen aus diesem Grund 8 Items nicht mit ein (Item1, 4, 7, 17, 25, 45, 47, 80). Beim LRT des TK *Geschlecht* wurden ein Item (47) und beim LRT des TK *Sprache* 2 Items (47,121) nicht berücksichtigt. Sehr viele Items (50+) flossen beim LRT des Teilungskriteriums *Alter* nicht mit ein.

TABELLE 10: DIE ERGEBNISSE DER LRTS VOR DEM AUSSCHLUSS EINES ITEMS

Teilungskriterium	χ^2	df	χ^2_{krit}	p
<i>Score</i>	142,196	96	131,141	0,002
<i>Alter</i>	32,6	37	59,893	0,675
<i>Geschlecht</i>	221,728	103	139,297	0
<i>Sprache</i>	211,829	102	138,135	0

Die grafischen Modellkontrollen (siehe dazu Abbildung 9 - 12) lassen erkennen, dass einige Items auffällig sind. Im TK *Score* sind dies vor allem die Items 73, 15, 96. Grafisch fallen im TK *Alter* die Items 32, 81 und 25 auf. Item 25, 4, 38, 106 sind im TK *Geschlecht* auffällig, Item 80 und 108 im TK *Sprache*.

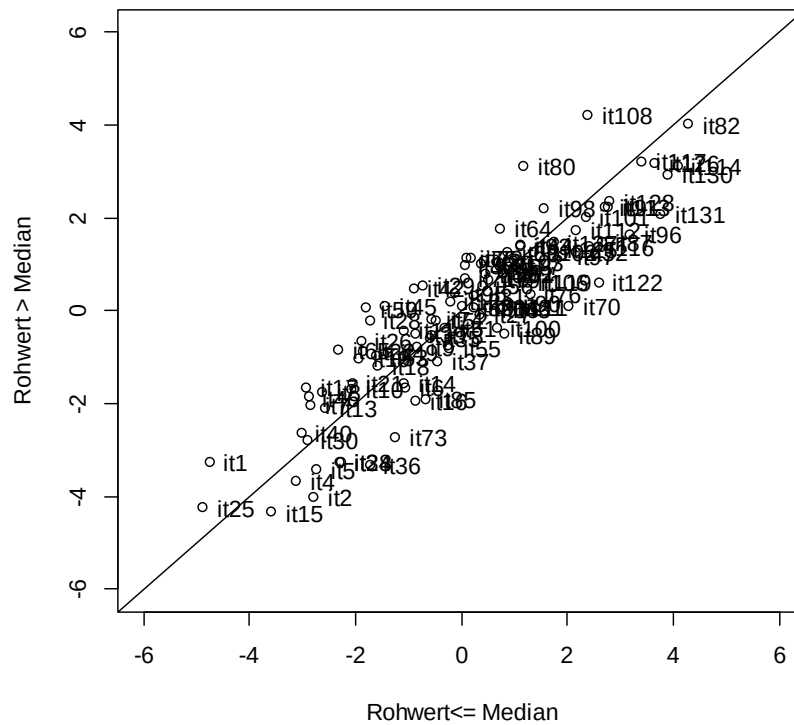


ABBILDUNG 8: GRAFISCHER MODELLTEST DES TK SCORE VOR DEM AUSSCHLUSS EINES ITEMS.

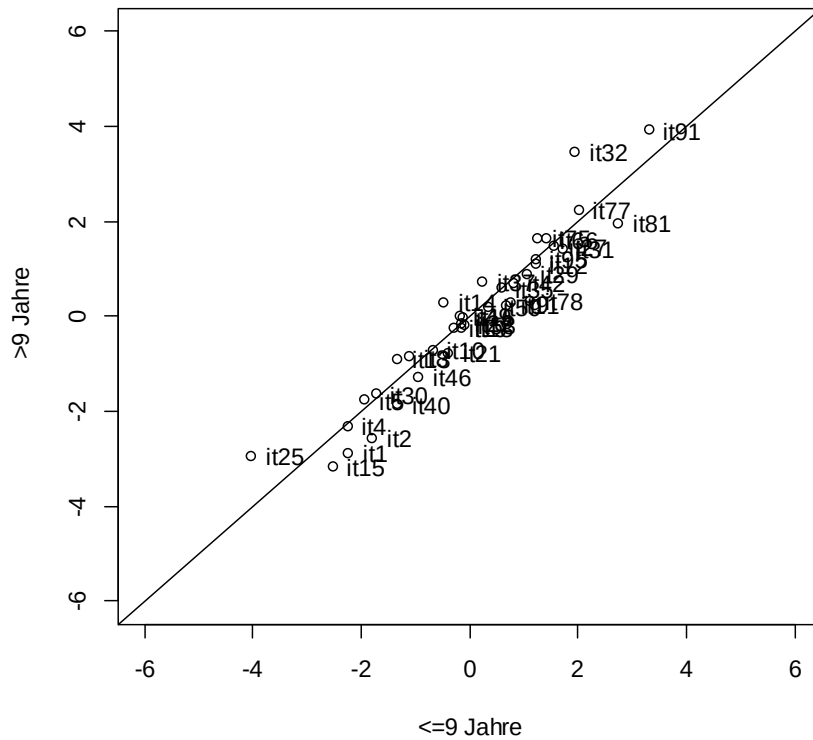


ABBILDUNG 9: GRAFISCHER MODELLTEST DES TK ALTER VOR DEM AUSSCHLUSS EINES ITEMS.

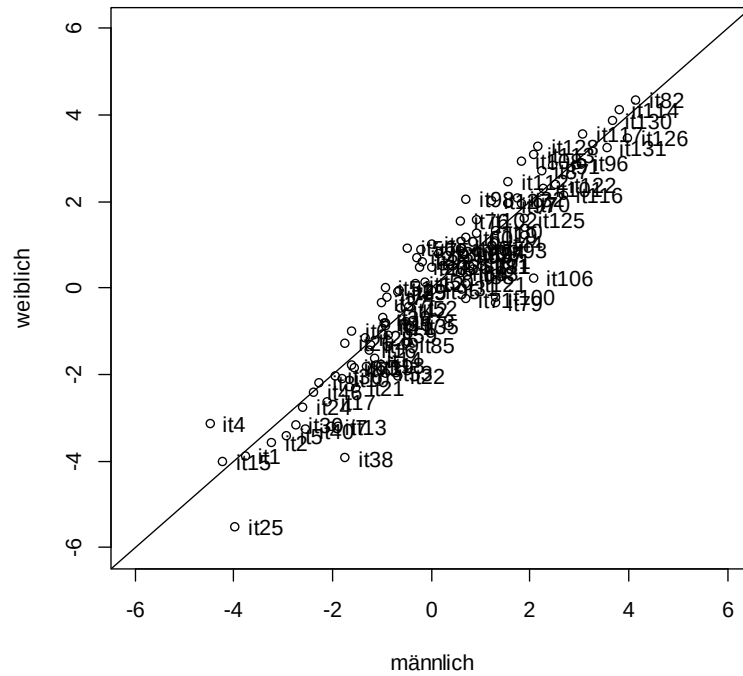


ABBILDUNG 10: GRAFISCHER MODELLTEST DES TK *GESCHLECHT* VOR DEM AUSSCHLUSS EINES ITEMS.

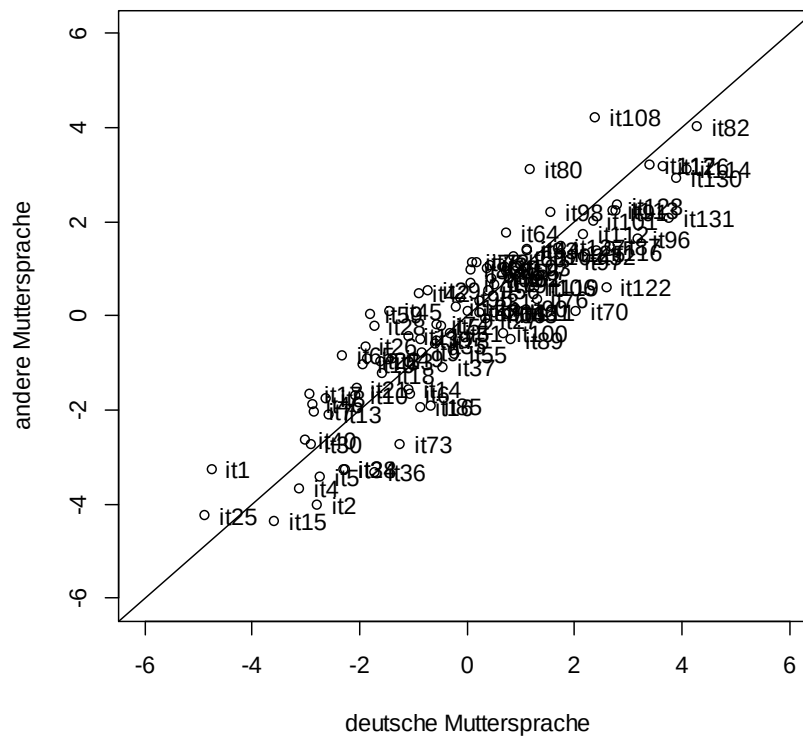


ABBILDUNG 11: GRAFISCHER MODELLTEST DES TK *SPRACHE* VOR DEM AUSSCHLUSS EINES ITEMS.

Auf Grund der signifikanten LRTs in mindestens einem Teilungskriterium und um ein geeignetes Item zum Ausschluss zu finden wurde anschließend der z-Test nach Fischer und Scheiblechner (1970) für die Teilungskriterien *Score*, *Geschlecht* und *Sprache* durchgeführt.

Im z-Test nach Fischer und Scheiblechner (1970) für das Teilungskriterium *Score* fallen vier Items signifikant aus (vgl. Tabelle 10).

TABELLE 11:SIGNIFIKANTE ITEMS VOR DEM AUSSCHLUSS EINES ITEMS FÜR DAS TK *SCORE*

Item	z-Wert	p-Wert
14	2,608	0,009
60	2,806	0,005
73	2,891	0,004
90	3,202	0,001

Der z-Test nach Fischer und Scheiblechner (1970) für das Teilungskriterium *Geschlecht* zeigte 10 signifikante Items, die in Tabelle 11 veranschaulicht wurden.

TABELLE 121: SIGNIFIKANTE ITEMS DES TK *GESCHLECHT* VOR DEM AUSSCHLUSS EINES ITEMS

Item	z-Wert	p-Wert
22	-2,852	0,004
27	2,892	0,004
56	3,220	0,001
75	2,999	0,003
76	2,764	0,006
79	-3,996	0,000
89	3,435	0,001
106	-3,120	0,002
108	3,207	0,001
113	2,947	0,003

Der z-Test nach Fischer und Scheiblechner (1970) für das Teilungskriterium *Sprache* identifizierte 6 signifikante Items, die in Tabelle 12 zu sehen sind.

TABELLE 132: SIGNIFIKANTE ITEMS DES TK *SPRACHE* VOR DEM AUSSCHLUSS EINES ITEMS

Item	z-Wert	p-Wert
16	-2,771	0,006
29	3,361	0,001
42	3,195	0,001
59	3,494	0,000
65	2,724	0,006
89	-3,390	0,001

Anschließend wurden unter Hilfenahme der grafischen Modellkontrolle und des z-Tests sukzessive 13 Items ausgeschlossen, Item 90, 73, 79, 56, 89, 106, 108, 113, 75, 27, 76, 22, 25.

Dies entspricht in etwa 12,2% des Itempools.

Die Ergebnisse der LRTs nach dem Ausschluss von ca. 12% der Items ist Tabelle 13 zu entnehmen.

TABELLE 13: DIE ERGEBNISSE DER LRTS NACH DEM AUSSCHLUSS VON 13 ITEMS

Teilungskriterium	χ^2	df	χ^2_{krit}	p
Score	113,406	86	119,414	0,026
Alter	29,962	33	54,776	0,619
Geschlecht	123,324	90	124,1163	0,011
Sprache	178,997	89	122,942	0

Insgesamt kann von einer *a-posteriori* Modellgültigkeit ausgegangen werden, da das interne TK *Score* und die externen Teilungskriterien *Alter* und *Geschlecht* die H_0 bestätigen. Das LRT des TK *Sprache* fällt weiterhin signifikant aus, was jedoch nicht gegen die eindimensionale Messung des neu entwickelten Tests spricht, da vor allem bei schriftlichen Tests davon ausgegangen wird, dass sie in Bezug auf *Sprache* nicht fair messen können. Weitere Begründungen für das signifikante LRT des TK *Sprache* sind Kapitel 9 zu entnehmen.

Abbildung 13-16 zeigen die grafische Modellkontrolle der vier LRTs nach Ausschluss von 13 Items.

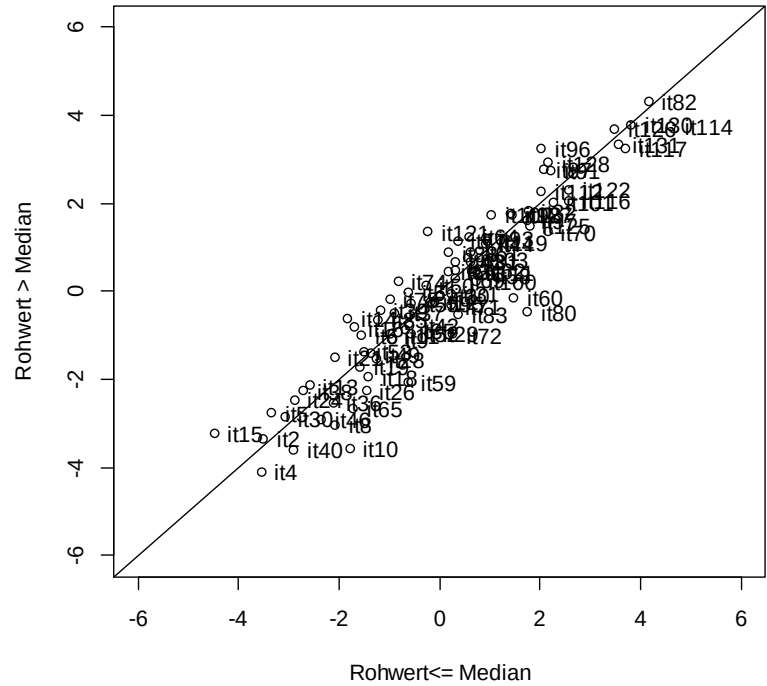


ABBILDUNG 12: GRAFISCHER MODELLTEST DES TK SCORE NACH DEM AUSSCHLUSS VON 13 ITEMS

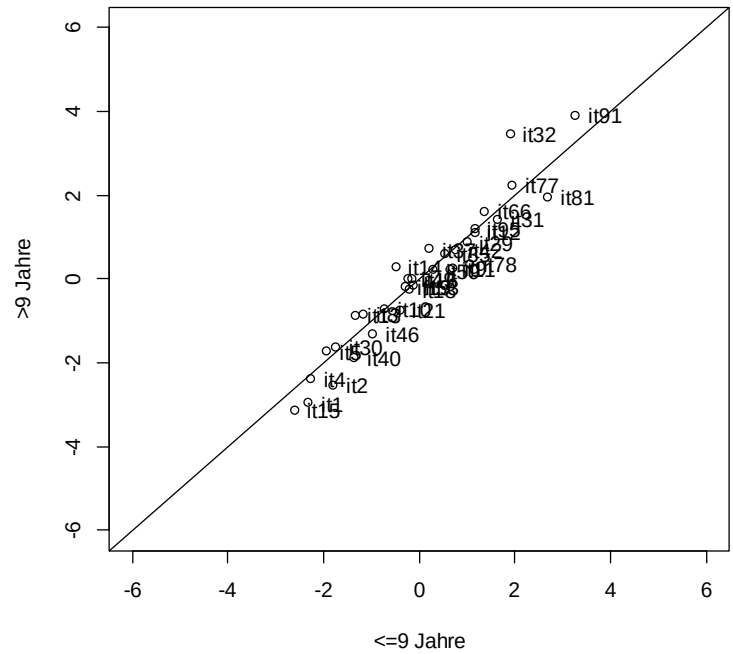


ABBILDUNG 13 :GRAFISCHER MODELLTEST DES TK ALTER NACH DEM AUSSCHLUSS VON 13 ITEMS

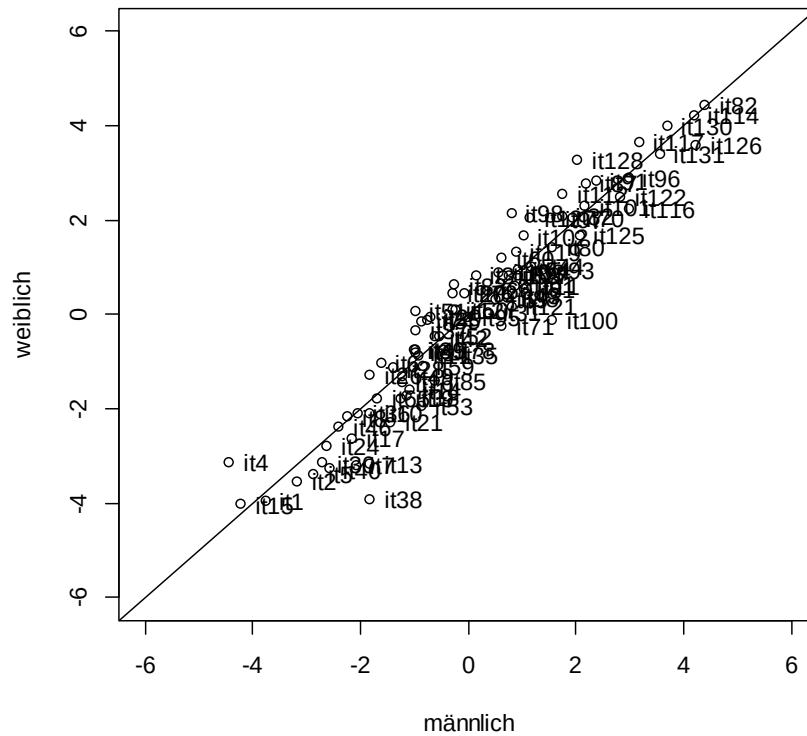


ABBILDUNG 14: GRAFISCHER MODELLTEST DES TK *GESCHLECHT* NACH DEM AUSSCHLUSS VON 13 ITEMS

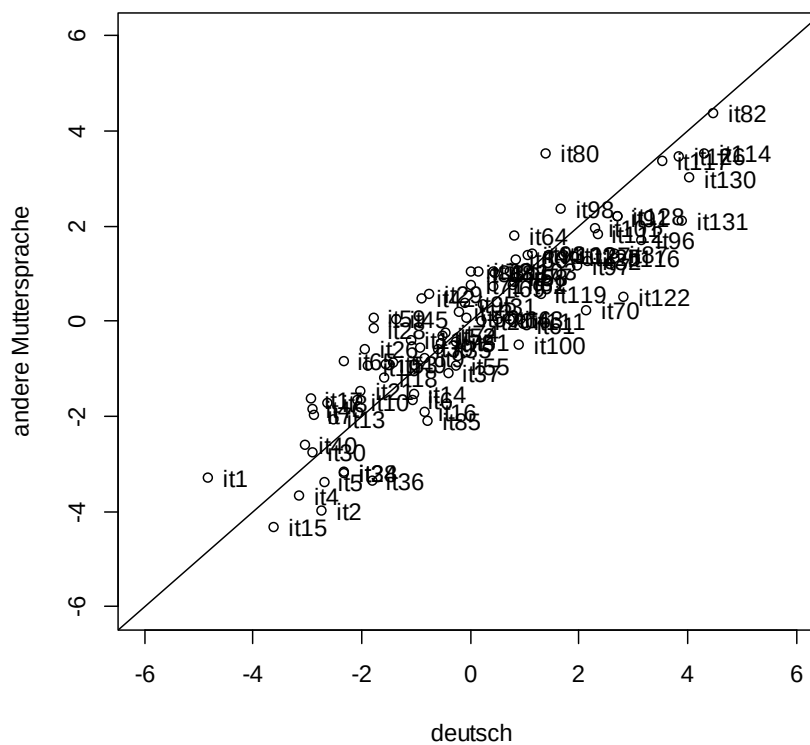


ABBILDUNG 15: GRAFISCHER MODELLTEST DES TK *SPRACHE* NACH DEM AUSSCHLUSS VON 13 ITEMS

Um der 2. Hypothese nachzugehen, nämlich ob es einen Zusammenhang zwischen den Leistungen der SchülerInnen im Einzelverfahren und den Leistungen der SchülerInnen im Gruppenverfahren gibt, kann eine Berechnung des Zusammenhangs der Personenparameter mittels Pearson-Korrelationskoeffizienten herangezogen werden, unter der Annahme einer gegebenen Normalverteilung und Intervallskalierung. Die Berechnung erfolgte wiederum unter der Verwendung der frei zugänglichen Software R.

211 der 814 Testpersonen wurden sowohl einzeln als auch in der Gruppe getestet, um Gewöhnungseffekte auszuschließen wurden 50% vor der Gruppentestung einzeln getestet während die andere Hälfte erst nach der Gruppentestung die Individualtestung hatte. Damit die Leistungen der SchülerInnen verglichen werden konnten, wurden zuerst mittels der open-source Software R die Personenparameter sowohl für die Individualtestung als auch für die Gruppentestung geschätzt.

Anschließend wurde eine Pearson-Korrelation durchgeführt. Der Korrelationkoeffizient beträgt $r=0.8099104$, damit belauft sich das Bestimmtheitsmaß auf $B= 0,6559$. Dies bedeutet, dass ca. 66% der Varianz der Leistungen in der einen Testbedingung durch das Resultat in der anderen Testbedingung erklärt werden. Damit ist eine hohe Übereinstimmungsvalidität gegeben. Die Abbildung 17 stellt die Korrelation grafisch dar.

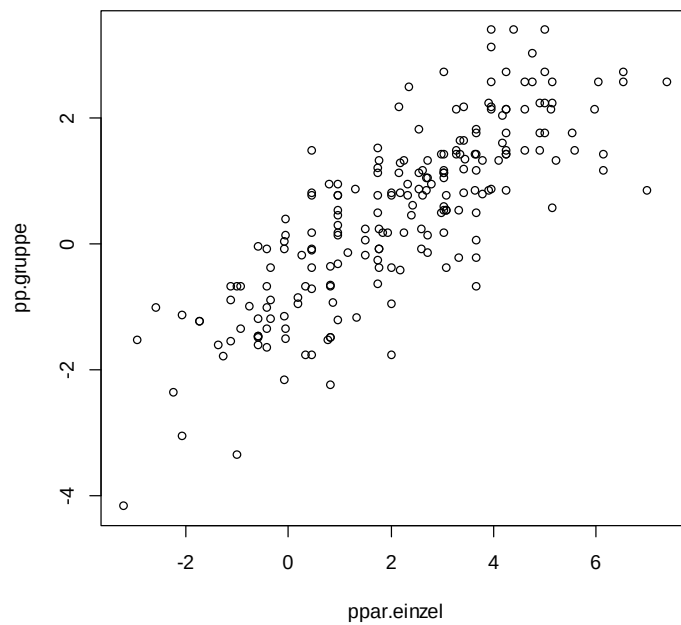


ABBILDUNG 16: DARSTELLUNG DER PEARSON-KORRELATION DER PERSONENPARAMETER

Dies bedeutet, dass ein Zusammenhang zwischen den Leistungen der SchülerInnen, trotz des unterschiedlichen Vorgabemodus, bei anzunehmenden gleichbleibenden Fähigkeiten besteht.

9. DISKUSSION UND AUSBLICK

Das Untersuchungsziel der vorliegenden Arbeit, einen *Rasch*-Modell konformen Gruppentest mit hoher Übereinstimmungsvalidität zu entwickeln, ist gelungen.

Die Konstruktion des Tests wurde nahe an den Untertest 1 des AID 2 (Kubinger, 2009a) und des AID 3 (Kubinger und Holocher-Ertl, in Vorbereitung) angelehnt, mit dem die Einzeltestungen durchgeführt wurden. Der neu konstruierte Unterest besteht aus je 20 Multiple-Choice Items, mit einer *a-priori*-Ratewahrscheinlichkeit von 16%, und 4 Items im Kästchenformat. Für die unterschiedlichen Schulstufen wurden acht Testhefte mit jeweils einem Parallelheft erstellt. Die Testpersonen wurden angehalten ihre Antworten auf dafür konzipierte Antwortbögen festzuhalten.

In Bezug auf das Testmaterial muss erwähnt werden, dass es für Volksschulkinder eher ungeeignet ist, da diese mit der Unterscheidung zwischen Testheft und Antwortbogen zum Teil große Probleme hatten. Dies könnte man jedoch mit der Einfärbung des Testheftes umgehen, sodass dann die Testleiteranweisung lauten würde: *In das rote Heft darf während der gesamten Testdauer nicht geschrieben werden.*

Des Weiteren wäre es sinnvoll das Testheft als Aufgabenheft zu benennen, dies würden sich die SchülerInnen möglicherweise besser merken, da sich dieses Heft ja aus den Aufgaben für die Testung zusammensetzt.

Die Stichprobe beinhaltet 814 Testpersonen im Alter von 8 - 18 Jahren aus Schulen in Ober- und Niederösterreich. Es ist zu empfehlen, bei weiteren Analysen eine altermäßig besser gestreute Stichprobe aus ganz Österreich zu wählen, da aufgrund des Schulschlusses nur wenige Testpersonen unter 9 Jahren bzw. über 17 Jahren zu verzeichnen sind.

Bei der *Rasch*-Modell Analyse konnte nach Ausscheiden von ca. 12% des Itempools eine *a-posteriori* Modellgültigkeit festgestellt werden. Trotz einem signifikanten *Likelihood-Ratio*-Test in dem Teilungskriterium *Sprache* kann davon ausgegangen werden, dass der neu konstruierte Gruppentest das Konstrukt *Alltagswissen* eindimensional misst. Die grafische Modellkontrolle des LRTs des Teilungskriteriums *Sprache* zeigt eine gute Modellgeltung. Kubinger (2009a) weißt

unter anderem darauf hin, dass bei großen Stichproben der Likelihood-Ratio-Test nach Andersen (1973) eher ein signifikantes Ergebnis nach sich zieht.

Das signifikante LRT im Teilungskriterium *Sprache* könnte darauf zurückzuführen sein, dass Kinder mit nicht deutscher Muttersprache wahrscheinlich einen geringeren Wortschatz ausgebildet haben wie ihre MitschülerInnen. Da keine Inhalte erklärt wurden, könnte in diesem Fall ein Nachteil für sie entstanden sein. Weiters wäre es möglich, dass Kinder mit nicht deutscher Muttersprache langsamer lesen und somit für die Itembearbeitung längere Zeit benötigen. Auffällig ist, dass in der erhobenen Stichprobe vorwiegend fremde Muttersprachen bei SchülerInnen bis zur 3. Klasse, also 7. Schulstufe, bzw. bis 13 Jahren vorliegen. (siehe dazu Tabelle 15 und 17)

TABELLE 14: VERTEILUNG DER MUTTERSPRACHEN AUF DIE SCHULSTUFEN

	D	Tür	Kro/Ser/SI	Slowk/Tsch	Pol	Russ	E	Fr/It/Sp	Arab	sonst
3.ST	48	20	12	1	1	1	0	0	1	28
4.ST	63	8	7	0	0	1	0	0	3	23
5.ST	70	4	8	0	1	0	1	0	0	7
6.ST	10 1	4	11	0	0	0	1	0	0	7
7.ST	63	2	7	1	1	0	0	0	0	9
8.ST	64	1	3	0	0	0	0	1	0	1
9.ST	60	2	2	0	0	0	0	0	1	6
10.ST	85	3	2	0	1	0	0	0	0	1
11.ST	62	0	2	0	0	0	0	0	0	2

TABELLE 15: VERTEILUNG DER MUTTERSPRACHEN AUF DAS ALTER

Alter	D	Türk	Kro/Ser/Slo	Slowak/Tsch	Pol	Russ	Engl	Fr/I/Sp	Arab	sonst
8	6	1	2	0	0	0	0	0	0	7
9	42	15	8	1	1	0	0	0	1	18
10	65	9	7	0	0	1	1	0	2	19
11	79	3	8	0	0	0	0	0	1	11
12	95	3	11	1	1	1	0	0	0	7
13	59	6	6	0	0	0	1	0	0	11
14	70	1	5	0	1	0	0	1	0	3
15	63	2	3	0	0	0	0	0	0	4
16	72	4	1	0	1	0	0	0	1	2
17	62	0	3	0	0	0	0	0	0	2
18	2	0	0	0	0	0	0	0	0	0
19	1	0	0	0	0	0	0	0	0	0

Es ergab sich eine hohe Übereinstimmungsvalidität zwischen dem neu konstruierten Untertest in gruppentauglicher Version und dem Untertest des AID3. Ungefähr 66 % der Varianz werden erklärt. Damit kann ausgesagt werden, dass trotz veränderte Versuchsbedingung ein Zusammenhang zwischen den Leistungen der Kinder und Jugendlichen besteht. Kinder und Jugendliche können ihre Fähigkeiten in einer Einzeltestung gleich gut entfalten, wie in der Gruppentestung.

Ausstehend ist eine sogenannte Kreuzvalidierung, darunter versteht man, den durch diese Arbeit veränderten Itempool mittels einer neuen Stichprobe auf die Gültigkeit des *Rasch*-Modells zu prüfen (vgl. Kubinger, 2005).

10. ZUSAMMENFASSUNG

Die vorliegende Arbeit hat sich mit der Konstruktion eines *Rasch*-Modell konformen gruppentauglichen Untertest zur Erfassung von *Alltagswissen* bei Kinder und Jugendliche im Alter von 8 - 18 Jahren befasst. Es wurde ein Itempool von 106 Items erstellt, die auf sechzehn verschiedene Testhefte verteilt wurden. Die Erhebung der Stichprobe fand in Schulen in Oberösterreich und Niederösterreich statt, insgesamt wurden 814 Kinder und Jugendliche getestet, 211 davon wurden zudem einer Einzeltestung mit einigen Untertests des AID 3 unterzogen.

Mittels dem *Rasch*-Modell Analysen sollte herausgefunden werden, ob der konstruierte Untertest das Konstrukt *Alltagswissen* eindimensional misst, zur Modellprüfung wurden der *Likelihood-Ratio*-Test nach Andersen (1973), die grafische Modellkontrolle und der z-Test nach Fischer und Scheiblechner (1973) herangezogen. Als Teilungskriterien fungierten *Score*, *Alter*, *Geschlecht* und *Sprache*. Bei den ersten drei Teilungskriterien zeigten die LRTs nach Ausschluss von 13 Items ein nicht signifikantes Ergebnis, trotz des übrigen signifikanten *Likelihood-Ratio*-Tests des Teilungskriteriums *Sprache*, kann davon ausgegangen werden, dass die Gültigkeit des *Rasch*-Modells belegt wurde. Somit kann ausgesagt werden, dass der neu konstruierte Untertest das Konstrukt *Alltagswissen* eindimensional misst.

Die Korrelation der Personenparameter der Individualtestung (AID 3 UT Alltagswissen) und Gruppentestung (AID-Gruppe, Alltagswissen) ergab eine hohe Übereinstimmungsvalidität. Der Korrelationskoeffizient beträgt $r=0.81$, und ergibt damit ein Bestimmtheitsmaß von $B=0,66$. Somit wird in etwa 66 % der Varianz der Leistung der einen Versuchsbedingung durch die Leistung der anderen Versuchsbedingung erklärt. Es kann davon ausgegangen werden, dass es zwischen den Leistungen der Kinder und Jugendlichen in Individual- und Gruppentestungen einen Zusammenhang gibt.

IV. LITERATURVERZEICHNIS

- Amelang, M., & Schmidt-Atzert, L. (2006). *Psychologische Diagnostik* (4 Aufl.). Heidelberg: Springer Medizin Verlag.
- Andersen, E. B. (1973). A goodness of fit test for the rasch model. *Psychometrika*, 38, 123-140.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (3 Aufl.). München: Pearson Studium.
- Bulheller, S. (2003). Normwert. In K. Kubinger & R. Jäger, *Schlüsselbegriffe der psychologischen Diagnostik* (S. 207-209). Berlin: Beltz.
- Cattell, R. B. (1961). Fluid and Crystallized Intelligence. In J. J. Jenkins & D. G. Paterson (Eds.), *Studies in individual differences. The search for intelligence* (pp. 738-746). London: Methuen & Co.
- Häcker, H. & Stapf, K. H. (2009). *Dorsch Psychologisches Wörterbuch* (15. überarbeitete und erweiterte Aufl.). Bern: Hans Huber.
- Herle, M. (2003). Intelligenzquotient. In K. Kubinger & R. Jäger, *Schlüsselbegriffe der psychologischen Diagnostik* (S. 207-209). Berlin: Beltz.
- Kubinger, K. D. (1989). Aktueller Stand und kritische Würdigung der Probabilistischen Testtheorie. In K. D. Kubinger (Hrsg.), *Moderne Testtheorie* (S. 19-83). Weinheim: Beltz.
- Kubinger, K. D. (1996). Methoden der Psychologischen Diagnostik. In E. Erdfelder, R. Mausfeld, T. Meiser, & G. Rudinger (Hrsg.), *Handbuch Quantitative Methoden* (S. 567-576). Weinheim: Beltz.
- Kubinger, K. D. (2003a). Testtheorie, Probabilistische. In K. D. Kubinger, *Schlüsselbegriffe der psychologischen Diagnostik* (S. 415-423). Berlin: Beltz.
- Kubinger, K. D. (2003b). Gütekriterien. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 195-204). Weinheim: Beltz.

- Kubinger, K. D. (2005). Psychological test calibration using the Rasch model- Some critical suggestions on traditional approaches. *International Journal of Testing*, 5, 377- 394
- Kubinger, K. D. (2009a). *Adaptives Intelligenz Diagnostikum - Version 2.2 (AID 2)samt AID 2-Türkisch*. Göttingen: Beltz Test.
- Kubinger, K. D. (2009b). *Psychologische Diagnostik. Theorie und Praxis psychologischen Diagnostizierens* (2 Aufl.). Göttingen: Hogrefe.
- Kubinger, K. D., & Holocher-Ertl S.(in Vorbereitung). *Adaptives Intelligenz Diagnostikum 3*.
- Lienert, G., & Raatz, U. (1998). *Testaufbau und Testanalyse* (6 Aufl.). Weinheim: Psychologie Verlags Union.
- Mair, P., Hatzinger, R. & Maier, M. (2010). *Extended Rasch Modeling. Package eRm* [Manual]. Verfügbar unter:<http://cran.r-project.org/web/packages/eRm/eRm.pdf>[20.08.2012]
- Moosbrugger, H., & Hartig, J. (2003). Testtheorie, Klassische. In K. Kubinger & R. Jäger, *Schlüsselbegriffe der psychologischen Diagnostik* (S. 408-415). Berlin: Beltz.
- Rost, J. (2004). *Testtheorie-Testkonstruktion* (2 Aufl.). Bern: Hans Huber.
- Seiwald, B. (2003). Antwortformat. In K. Kubinger & J. Reinhold, *Schlüsselbegriffe der psychologischen Diagnostik* (S. 23-29). Berlin: Beltz.
- Stemmler, G., Hagemann, D., Amelang, M., & Bartussek, D. (2011). *Differentielle Psychologie und Persönlichkeitsforschung* (7 Aufl.). Stuttgart: Kohlhammer.
- Süß, H.-M. (2003). Intelligenztheorien. In K. Kubinger & R. Jäger, *Schlüsselbegriffe der psychologischen Diagnostik* (S. 217-225). Berlin: Beltz.
- Weber, H., & Rammsayer, T. (2012). *Differentielle Psychologie-Persönlichkeitsforschung*. Göttingen: Hogrefe Verlag.
- Wechsler, D. (1964). *Die Messung der Intelligenz Erwachsener*. Bern: Huber.

Testdesign der Testhefte 1a – 8a des Gruppentest *Alltagswissen*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	24	25	26	27	28	29	30	31	32	35	36	37	38	39	
	L	L	L	L	L	L	L	L	L	L	L	L	L	L	15	13	17	16	19	18			L	L	L	L	L	L	L	L	M	58	L	M	L	M	128
	K	K	K			K		K						K									L	K	K	K	K			K				K			
	n	n	n	b	b	b	b	g	g	v	v	r	be	be	be	m	m	lit	er	er	er	er	mat	mat	mat	sp	sp	sp		n	n	b	b	g	g	v	
1a	5			1				10	2	3	4			6				13				14		17				15	16	22	7	8					
2a		5			1				2		3	4	6	10		13		14		15						17			16	22				7			
3a			5			17			1	2					3				13				14				10			22	4			6		15	
4a							1		2		3						13								5		17	14						10		4	
5a									1										13														5			2	
6a								1									13																				
7a									1																												
8a								1																											2		

	40	42	43	44	45	46	47	49	50	51	52	53	55	56	58	59	60	61	62	64	65	66	69	70	71	72	73	74	75	76	77	78	79	80	81	82
	L	M	L	M	L	M	L	M	L	M	L	M	L	M	M	M	M	M	63	M	M	M	M	M	M	M	24	72	M	M	M	M	M	M	M	
				K								K	K	K												K		K						K		
	v	m	m	lit	LIT	er	fi	spr	spr	spr	spr	n	n	n	b	b	b	b	g	g	g	er	er	er	mat	mat	mat	mat	sp	fi	ku	ku	kun	kun	spr	
1a	18					19		20								9															21				23	
2a		8					18	19				20				7													11		21	23				
3a		16					18											8			9	19					20							21		
4a	6			22	15			16	18								7				8	9			19				20							
5a					3			14	15		10		17					4				6	7				22			16						
6a	2			5		3		14		15	10		17						4	6	7							22				16				
7a								2				13	5							3	4		6			10						14				
8a								13						5	3			4				14						10	15				16		17	

	83	85	87	88	89	90	91	93	95	96	97	98	100	101	102	103	106	108	110	111	112	113	114	116	117	119	121	122	125	126	127	128	130	131
	M	M	M	M	M	S	M	S	M	S	M	S	M	S	M	S	M	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
	spr	sex	pc	psy	n	n	n	r	m	er	mat	mat	ku	wi	wis	wis	psy	n	n	b	v	r	r	r	r	m	lit	mat	spr	pol	pol	ge	wi	psy
1a							12		24																									
2a							12		24																									
3a					11			12	23							24																		
4a	21					11			12	23					24																			
5a		18	19			8	9		20						21			11				12									23	24		
6a				18	8			9	19		20				21					11						12							23	24
7a			15	16				7	8			18			19					22	9			11		12			20	21	23		24	
8a					6	7			8				18				19	9		11	22		12	20				21	23	24				

Testdesign der Testhefte 1b – 8b des Gruppentest *Alltagswissen*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	24	25	26	27	28	29	30	31	32	35	36	37	38	39
													15	13	17	16	19	18					73							58				128		
	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	M	L	M	L	M	L	M
	K	K	K			K			K					K									K	K	K	K				K			K			
	n	n	n	b	b	b	b	g	g	v	v	r	be	be	be	m	m	lit	lit	er	er	er	mat	mat	mat	sp	sp	sp		n	n	b	b	g	g	v
1b	17			13				22	14	15		16		18			1		2	3		2		5				3	4	10	19	20				
2b	17			13				14	15	16	18	22		1													5			4	10			19		
3b			17		5			13	14					15				1				2				22				10	16			18		3
4b						13		14	15								1								17		5	2						22		16
5b								13												1													17			14
6b								13									1																			
7b								13																												
8b								13																											14	

	40	42	43	44	45	46	47	49	50	51	52	53	55	56	58	59	60	61	62	64	65	66	69	70	71	72	73	74	75	76	77	78	79	80	81	82
	L	M	L	M	L	M	L	M	L	M	L	M	L	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	
				K								K	K	K												K	M	M	K					K		
	v	m	m	lit	LIT	er	fi	spr	spr	spr	spr	n	n	n	b	b	b	b	g	g	g	er	er	mat	mat	mat	mat	sp	fi	ku	ku	kun	kun	spr	spr	
1b	6					7		8							21																9			11		
2b		20					6	7				8																	23		9	11				
3b		4						6							19				20			21	7					8						9		
4b	18		10	3				4	6							19			20	21				7					8							
5b				15				2	3		22		5				16			18	19					10				4						
6b	14		17		15			2		3	22			5					16	18	19								10				4			
7b								14					1	17						15	16		18		22							2			5	
8b								1							17	15			16			2						22		3				4		5

	83	85	87	88	89	90	91	93	95	96	97	98	100	101	102	103	106	108	110	111	112	113	114	116	117	119	121	122	125	126	127	128	130	131	
	M	M	M	M	M_S	M_S	M_S	M_S	M_S	M_S	M_S	M_S	M_S	M_S	M_S	M_S	M_S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	
																				K													38		
	spr	sex	pc	psy	n	n	n	r	m	er	mat	mat	ku	wi	wis	wis	psy	n	n	b	v	r	r	r	r	m	lit	mat	spr	pol	pol	ge	wi	psy	
1b							24		12																										
2b							24		12																										
3b					23			24	11							12																			
4b	9					23			24	11					12																				
5b		6	7				20	21		8					9				23				24										11	12	
6b				6	20				21	7		8				9				23						24								11	12
7b			3	4				19	20			6			7					10	21					23		24		8	9	11		12	
8b					18	19			20				6				7	21		23	10		24	8				9	11	12					

**Z-TEST NACH FISCHER UND SCHEIBLECHNER (1970) FÜR DAS TK
SCORE VOR DEM AUSSCHLUSS VON ITEMS**

Item	z-Wert	p-Wert
<i>it2</i>	0,000	1,000
<i>it3</i>	-0,163	0,871
<i>it5</i>	0,524	0,600
<i>it6</i>	0,688	0,492
<i>it8</i>	-1,280	0,200
<i>it9</i>	-1,222	0,222
<i>it10</i>	-1,447	0,148
<i>it11</i>	0,700	0,484
<i>it12</i>	0,217	0,828
<i>it13</i>	0,409	0,683
<i>it14</i>	2,608	0,009
<i>it15</i>	1,214	0,225
<i>it16</i>	1,899	0,058
<i>it18</i>	-1,597	0,110
<i>it19</i>	-0,997	0,319
<i>it20</i>	0,846	0,398
<i>it21</i>	0,820	0,412
<i>it22</i>	-0,505	0,614
<i>it24</i>	0,251	0,802
<i>it26</i>	-0,841	0,400
<i>it27</i>	-0,217	0,828
<i>it28</i>	-0,456	0,648
<i>it29</i>	-2,314	0,021
<i>it30</i>	0,955	0,340
<i>it31</i>	-0,114	0,909
<i>it32</i>	-0,039	0,969
<i>it35</i>	-0,059	0,953
<i>it36</i>	1,052	0,293
<i>it37</i>	-1,302	0,193
<i>it38</i>	1,477	0,140
<i>it39</i>	2,156	0,031
<i>it40</i>	-0,087	0,930
<i>it42</i>	-0,143	0,886
<i>it43</i>	0,742	0,458
<i>it44</i>	1,409	0,159
<i>it46</i>	-1,238	0,216
<i>it49</i>	-0,373	0,709
<i>it50</i>	0,143	0,886
<i>it51</i>	0,487	0,626
<i>it52</i>	-1,676	0,094
<i>it53</i>	-0,041	0,968
<i>it55</i>	0,409	0,682

<i>it56</i>	-0,054	0,957
<i>it58</i>	-0,702	0,483
<i>it59</i>	-1,682	0,093
<i>it60</i>	-2,806	0,005
<i>it61</i>	0,613	0,54
<i>it62</i>	-0,275	0,783
<i>it64</i>	0,380	0,704
<i>it65</i>	-0,924	0,355
<i>it66</i>	0,469	0,639
<i>it69</i>	-0,448	0,654
<i>it70</i>	-1,938	0,053
<i>it71</i>	-1,246	0,213
<i>it72</i>	-2,206	0,027
<i>it73</i>	2,891	0,004
<i>it74</i>	1,686	0,092
<i>it75</i>	-0,309	0,757
<i>it76</i>	0,828	0,407
<i>it77</i>	1,910	0,056
<i>it78</i>	1,191	0,234
<i>it79</i>	0,835	0,403
<i>it81</i>	-0,523	0,601
<i>it82</i>	0,036	0,972
<i>it83</i>	-1,823	0,068
<i>it85</i>	1,832	0,067
<i>it87</i>	1,052	0,293
<i>it88</i>	1,342	0,179
<i>it89</i>	1,800	0,072
<i>it90</i>	3,202	0,001
<i>it91</i>	2,361	0,018
<i>it93</i>	1,154	0,249
<i>it95</i>	-0,919	0,358
<i>it96</i>	2,193	0,028
<i>it97</i>	-0,548	0,583
<i>it98</i>	-0,115	0,909
<i>it100</i>	-0,646	0,518
<i>it101</i>	-0,061	0,951
<i>it102</i>	0,611	0,541
<i>it103</i>	0,534	0,594
<i>it106</i>	-0,442	0,659
<i>it108</i>	-0,235	0,814
<i>it111</i>	-0,966	0,334
<i>it112</i>	-0,088	0,930
<i>it113</i>	0,425	0,671
<i>it114</i>	-1,064	0,287
<i>it116</i>	-1,337	0,181
<i>it117</i>	-1,238	0,216

<i>it119</i>	-0,305	0,761
<i>it121</i>	1,085	0,278
<i>it122</i>	-1,264	0,206
<i>it125</i>	-1,184	0,236
<i>it126</i>	-0,479	0,632
<i>it127</i>	0,676	0,499
<i>it128</i>	1,582	0,114
<i>it130</i>	-0,515	0,607
<i>it131</i>	-0,068	0,946

**Z-TEST NACH FISCHER UND SCHEIBLECHNER(1970) FÜR DAS TK
GESCHLECHT VOR DEM AUSSCHLUSS VON ITEMS**

Item	z-Wert	p-Wert
<i>it1</i>	-0,115	0,909
<i>it2</i>	-0,410	0,682
<i>it3</i>	1,532	0,126
<i>it4</i>	1,595	0,111
<i>it5</i>	-0,648	0,517
<i>it6</i>	1,069	0,285
<i>it7</i>	-1,216	0,224
<i>it8</i>	0,172	0,863
<i>it9</i>	0,827	0,408
<i>it10</i>	-0,641	0,522
<i>it11</i>	0,191	0,849
<i>it12</i>	1,216	0,224
<i>it13</i>	-1,944	0,052
<i>it14</i>	-0,842	0,400
<i>it15</i>	0,230	0,818
<i>it16</i>	-0,432	0,666
<i>it17</i>	-0,762	0,446
<i>it18</i>	-1,430	0,153
<i>it19</i>	-0,981	0,327
<i>it20</i>	2,145	0,032
<i>it21</i>	-1,505	0,132
<i>it22</i>	-2,852	0,004
<i>it24</i>	-0,129	0,897
<i>it25</i>	-1,264	0,206
<i>it26</i>	0,838	0,402
<i>it27</i>	2,892	0,004
<i>it28</i>	0,421	0,674
<i>it29</i>	1,853	0,064
<i>it30</i>	-0,893	0,372
<i>it31</i>	-1,171	0,241
<i>it32</i>	0,839	0,401
<i>it35</i>	-0,888	0,374
<i>it36</i>	-0,056	0,955
<i>it37</i>	1,998	0,046
<i>it38</i>	-1,863	0,062
<i>it39</i>	0,668	0,504
<i>it40</i>	-1,447	0,148
<i>it42</i>	0,448	0,654
<i>it43</i>	-1,164	0,245
<i>it44</i>	-0,675	0,500
<i>it45</i>	0,234	0,815
<i>it46</i>	0,003	0,998
<i>it49</i>	-0,285	0,776
<i>it50</i>	0,779	0,436
<i>it51</i>	1,609	0,108
<i>it52</i>	-0,002	0,998
<i>it53</i>	-2,049	0,041
<i>it55</i>	1,211	0,226
<i>it56</i>	3,220	0,001

<i>it58</i>	0,226	0,821
<i>it59</i>	-0,454	0,650
<i>it60</i>	1,120	0,263
<i>it61</i>	-1,136	0,256
<i>it62</i>	0,362	0,717
<i>it64</i>	-0,143	0,886
<i>it65</i>	-0,253	0,800
<i>it66</i>	1,531	0,126
<i>it69</i>	-0,709	0,479
<i>it70</i>	0,238	0,812
<i>it71</i>	-1,875	0,061
<i>it72</i>	1,162	0,245
<i>it74</i>	1,302	0,193
<i>it75</i>	2,999	0,003
<i>it76</i>	2,764	0,006
<i>it77</i>	-0,184	0,854
<i>it78</i>	-0,467	0,641
<i>it79</i>	-3,996	0,000
<i>it80</i>	-0,077	0,939
<i>it81</i>	-1,393	0,164
<i>it82</i>	0,271	0,786
<i>it83</i>	1,756	0,079
<i>it85</i>	-1,221	0,222
<i>it87</i>	1,505	0,132
<i>it88</i>	-1,112	0,266
<i>it89</i>	3,435	0,001
<i>it91</i>	1,051	0,293
<i>it93</i>	-1,715	0,086
<i>it95</i>	-0,336	0,737
<i>it96</i>	-0,142	0,887
<i>it97</i>	0,873	0,383
<i>it98</i>	2,278	0,023
<i>it100</i>	-2,317	0,021
<i>it101</i>	0,218	0,827
<i>it102</i>	1,105	0,269
<i>it103</i>	0,832	0,406
<i>it106</i>	-3,120	0,002
<i>it108</i>	3,207	0,001
<i>it111</i>	-0,769	0,442
<i>it112</i>	2,262	0,024
<i>it113</i>	2,947	0,003
<i>it114</i>	0,321	0,748
<i>it116</i>	-1,164	0,245
<i>it117</i>	0,802	0,423
<i>it119</i>	0,750	0,453
<i>it121</i>	-0,871	0,384
<i>it122</i>	-0,389	0,697
<i>it125</i>	-0,846	0,398
<i>it126</i>	-1,152	0,249
<i>it127</i>	2,513	0,012
<i>it128</i>	2,404	0,016
<i>it130</i>	0,459	0,646
<i>it131</i>	-0,481	0,630

**Z-TEST NACH FISCHER UND SCHEIBLECHNER FÜR DAS TK *SPRACHE*
VOR DEM AUSSCHLUSS VON ITEMS**

Item	z-Wert	p-Wert
it1	1,338	0,181
it2	-1,371	0,170
it3	1,081	0,280
it4	-0,755	0,450
it5	-0,958	0,338
it6	-0,937	0,349
it7	0,920	0,358
it8	1,493	0,135
it9	0,319	0,750
it10	0,987	0,324
it11	1,709	0,088
it12	1,141	0,254
it13	0,761	0,447
it14	-0,881	0,378
it15	-0,782	0,434
it16	-2,771	0,006
it17	1,900	0,057
it18	0,855	0,392
it19	1,790	0,073
it20	-0,335	0,738
it21	0,976	0,329
it22	2,135	0,033
it24	-0,960	0,337
it25	0,532	0,595
it26	1,923	0,055
it27	-1,092	0,275
it28	2,419	0,016
it29	3,361	0,001
it30	0,274	0,784
it31	0,305	0,761
it32	-1,965	0,049
it35	0,031	0,976
it36	-1,491	0,136
it37	-1,734	0,083
it38	-0,852	0,394
it39	0,819	0,413
it40	0,839	0,402
it42	3,195	0,001
it43	-1,661	0,097
it44	0,676	0,499
it45	2,236	0,025
it46	1,794	0,073
it49	2,166	0,030
it50	0,207	0,836
it51	-0,089	0,929
it52	0,553	0,580
it53	1,281	0,200
it55	-0,800	0,423
it56	0,362	0,718

it58	0,434	0,664
it59	3,494	0,000
it60	0,605	0,545
it61	-1,504	0,132
it62	0,042	0,966
it64	2,042	0,041
it65	2,724	0,006
it66	-1,955	0,051
it69	0,798	0,425
it70	-1,530	0,126
it71	1,102	0,270
it72	1,670	0,095
it73	-1,842	0,065
it74	0,742	0,458
it75	2,155	0,031
it76	-1,717	0,086
it77	0,125	0,900
it78	0,082	0,935
it79	0,323	0,746
it80	1,650	0,099
it81	0,099	0,921
it82	-0,235	0,814
it83	1,498	0,134
it85	-1,091	0,275
it87	-2,049	0,040
it88	1,305	0,192
it89	-3,390	0,001
it90	-1,749	0,080
it91	-1,010	0,313
it93	0,765	0,444
it95	2,165	0,030
it96	-1,996	0,046
it97	-1,274	0,203
it98	0,624	0,533
it100	-0,863	0,388
it101	-0,604	0,546
it102	-0,075	0,940
it103	0,853	0,394
it106	-0,562	0,574
it108	1,734	0,083
it111	-1,382	0,167
it112	-0,593	0,553
it113	-0,752	0,452
it114	-0,779	0,436
it116	-1,104	0,270
it117	-0,150	0,881
it119	-1,095	0,273
it122	-2,028	0,043
it125	-0,572	0,567
it126	-0,520	0,603
it127	-0,460	0,646
it128	-0,568	0,570
it130	-1,486	0,137
it131	-2,452	0,014

**PERSONENPARAMETER DER TESTPERSONEN DIE SOWOHL EINE
EINZEL- ALS AUCH EINE GRUPPENTESTUNG HATTEN:**

tp	pp gruppe	pp einzel
AB15	0,61606885	2,41749726
AB16	0,94050719	2,7772386
AB25	-1,52716236	0,75773539
AB3	-0,08108728	-0,41234066
AB30	-0,08108728	-0,07984588
AB35	0,2303142	2,59002535
AB41	-0,37739397	1,99684456
AB43	-1,33947288	-0,06986859
AB6	-0,6651641	-0,94474629
AB8	-1,23539567	-1,73979651
rl004	-1,16599644	1,32296928
rl012	0,94050719	0,79440284
rl027	-0,6651641	-0,99598139
rl028	-1,55006769	-1,13638977
rl053	-0,66511182	-0,41234066
rl224	-0,08108728	0,45331567
rl231	-0,08108728	1,76583979
rl236	-0,08108728	2,59002535
rl241	-2,15100106	-0,07984588
rl243	-0,0324192	-0,58814555
AB17	-0,85235055	0,17501354
AB2	-1,59464723	-0,58814555
AB29	1,18176866	3,42623941
AB37	-1,47423815	-0,58814555
AB42	0,44875267	2,3953081
AB5	0,44875267	0,94971983
rl033	-0,13573169	1,1611064
rl052	-1,77152985	-1,27201502
rl226	-1,19918457	0,97208335
rl228	-0,13573169	2,71896714
rl229	0,14611039	-0,06986859
rl235	2,48286791	2,34625463
AB136	1,32557839	1,76583979
AB140	1,64837209	3,42623941
AB146	0,77716799	0,97208335
AB153	0,52923349	0,97208335
AB158	1,32557839	3,7837648
AB188	1,64837209	3,34659871
AB192	1,03973859	3,03176578
AB197	1,32557839	4,08920263
AB201	0,77716799	2,62206329
AB239	1,32557839	2,246291
rl105	0,77716799	1,74821863
rl107	-0,66897027	3,64955974
rl111	-1,51784613	-2,96005111

<i>rl286</i>	0,77716799	0,97208335
<i>rl288</i>	-0,4210872	2,18708914
<i>rl294</i>	-0,18134508	1,48798005
<i>rl295</i>	0,38848157	-0,06986859
<i>rl307</i>	0,52923349	3,06795862
<i>rl311</i>	1,32557839	5,21062489
<i>AB142</i>	0,81000721	2,18708914
<i>AB145</i>	0,81000721	1,99684456
<i>AB164</i>	-0,09824859	0,45331567
<i>AB180</i>	0,48538861	3,64955974
<i>AB186</i>	0,48538861	1,74821863
<i>AB187</i>	2,16539481	2,15428086
<i>AB206</i>	1,17495225	2,62206329
<i>AB236</i>	0,18584943	2,246291
<i>AB241</i>	2,16539481	3,42623941
<i>rl089</i>	0,18584943	3,03176578
<i>rl091</i>	-1,48286591	0,82185125
<i>rl094</i>	-1,48286591	0,82185125
<i>rl096</i>	-0,63145989	1,74821863
<i>rl110</i>	0,18584943	1,82477972
<i>rl287</i>	1,28286682	2,18708914
<i>rl290</i>	-1,19558055	-0,58814555
<i>rl314</i>	-0,37361688	3,06795862
<i>AB104</i>	2,56916596	3,9482178
<i>AB267</i>	1,12541325	3,03176578
<i>AB285</i>	1,42969831	6,13397518
<i>AB290</i>	2,13306449	3,9482178
<i>AB305</i>	1,12541325	3,03176578
<i>AB307</i>	2,13306449	5,96979011
<i>rl213</i>	0,83963764	2,69545787
<i>rl216</i>	2,13306449	5,12698105
<i>rl250</i>	1,12541325	2,53720028
<i>rl261</i>	1,76115649	4,89395306
<i>rl330</i>	2,56916596	4,60625465
<i>rl372</i>	1,12541325	2,15428086
<i>rl379</i>	1,76115649	5,52091001
<i>AB110</i>	2,23654062	4,89395306
<i>AB118</i>	0,86968804	3,9482178
<i>AB122</i>	3,39088417	4,38515004
<i>AB258</i>	1,47632762	4,60625465
<i>AB275</i>	2,23654062	5,00385359
<i>AB318</i>	3,39088417	5,00385359
<i>AB339</i>	1,47632762	3,26861971
<i>rl210</i>	1,16136722	3,64955974
<i>rl244</i>	1,16136722	6,13397518
<i>rl245</i>	1,47632762	0,45331567
<i>rl344</i>	1,47632762	5,57363676
<i>rl374</i>	2,73663872	5,00385359
<i>rl375</i>	1,21183534	1,74821863
<i>AB10</i>	-3,34517787	-0,99598139
<i>AB12</i>	-1,35491392	-0,41234066

AB22	-0,94961446	2,01236464
AB26	-0,88711264	-1,13638977
AB27	0,2303142	1,48798005
AB32	-0,71912377	0,45331567
AB34	0,04668191	-0,07984588
AB36	-0,6651641	-1,13638977
AB436	-1,34099743	-0,94474629
rl001	-1,23539567	-1,73979651
rl015	-1,12360865	-2,0678776
rl029	-1,63445793	-0,41234066
rl043	-4,16060527	-3,22713874
rl044	-0,94961446	0,17501354
rl048	-2,34559856	-2,24735937
rl058	-1,76977044	0,45331567
rl225	0,94050719	0,97208335
rl230	0,94050719	2,33137502
rl239	-0,37739397	0,45331567
rl240	-0,08108728	1,76583979
rl242	-0,37739397	1,76583979
AB1	0,14611039	0,97208335
AB14	-1,76249336	0,32238718
AB19	-1,60313562	-1,36194857
AB20	-0,6693315	0,32238718
AB23	-1,76249336	1,99684456
AB31	-0,31128286	0,97208335
AB33	-0,98099261	-0,75450732
AB44	-0,93239186	0,86074053
AB7	-1,00770045	-2,59921633
rl011	-1,14773611	-0,07984588
rl025	-3,04379699	-2,0678776
rl051	-1,00770045	-0,41234066
rl227	0,78579389	3,7837648
rl233	0,14611039	2,70513915
rl237	-1,47423815	-0,58814555
AB129	1,32557839	2,70513915
AB131	0,05436752	1,48798005
AB135	0,77716799	1,99684456
AB149	0,77716799	0,45331567
AB156	0,52923349	3,03176578
AB159	1,03973859	2,71896714
AB162	0,52923349	3,06795862
AB185	0,77716799	2,32340311
AB204	1,51230642	1,74703446
AB205	2,0328515	4,16100231
AB240	1,03973859	2,69545787
rl087	-0,36402287	0,82185125
rl093	-0,89843587	-0,3503964
rl097	-2,23351339	0,82185125
rl098	-0,18134508	0,25759781
rl101	-0,66897027	0,82185125
rl284	0,24385477	1,76583979

<i>rl289</i>	-1,49578056	-0,06986859
<i>rl291</i>	1,33550902	3,42917826
<i>rl306</i>	0,28988357	0,97208335
<i>rl309</i>	0,77716799	3,07515455
<i>rl310</i>	0,52923349	3,30622858
<i>AB151</i>	0,81000721	0,45331567
<i>AB166</i>	0,18584943	0,97208335
<i>AB167</i>	0,18584943	1,93806189
<i>AB198</i>	2,16539481	3,9482178
<i>AB248</i>	0,48538861	2,97132444
<i>AB252</i>	3,02476937	4,74173165
<i>rl092</i>	-1,19558055	-0,3503964
<i>rl102</i>	1,17495225	3,03176578
<i>rl106</i>	-0,37361688	-0,3503964
<i>rl108</i>	-0,6453977	0,82185125
<i>rl109</i>	-0,25246012	1,74821863
<i>rl292</i>	0,18584943	0,45331567
<i>rl296</i>	-1,48286591	-0,58814555
<i>rl312</i>	1,6074344	4,16100231
<i>rl313</i>	0,81000721	3,42623941
<i>AB105</i>	1,42969831	2,97132444
<i>AB109</i>	2,13306449	4,60625465
<i>AB117</i>	1,42969831	4,24961911
<i>AB274</i>	0,83963764	4,24961911
<i>AB283</i>	0,83963764	3,893331
<i>AB288</i>	1,42969831	3,27052107
<i>AB296</i>	0,83963764	3,62319433
<i>AB297</i>	1,12541325	1,74821863
<i>AB312</i>	2,56916596	7,36932606
<i>AB315</i>	1,42969831	3,64955974
<i>AB320</i>	1,42969831	3,34659871
<i>AB330</i>	2,13306449	3,27052107
<i>AB336</i>	2,56916596	6,03552131
<i>rl193</i>	2,13306449	4,2413484
<i>rl198</i>	1,42969831	4,24961911
<i>rl199</i>	1,76115649	4,2413484
<i>rl200</i>	0,56601557	5,1286955
<i>rl215</i>	2,13306449	4,2413484
<i>rl219</i>	2,56916596	5,1286955
<i>rl222</i>	1,76115649	3,64955974
<i>rl251</i>	1,76115649	5,00385359
<i>rl254</i>	0,83963764	6,98417835
<i>rl256</i>	1,42969831	3,62319433
<i>rl337</i>	2,56916596	4,74173165
<i>rl346</i>	3,12209459	3,9482178
<i>rl356</i>	2,56916596	6,52889252
<i>rl373</i>	1,42969831	3,03176578
<i>AB107</i>	3,39088417	3,9482178
<i>AB111</i>	2,23654062	3,893331
<i>AB115</i>	1,82783177	2,53720028
<i>AB119</i>	1,82783177	3,64955974

<i>AB123</i>	2,73663872	3,03176578
<i>AB271</i>	0,59216918	3,03176578
<i>AB277</i>	0,86968804	2,53720028
<i>AB308</i>	2,73663872	6,52889252
<i>AB322</i>	-0,21824243	3,30622858
<i>AB329</i>	1,47632762	4,24961911
<i>rl204</i>	2,73663872	4,24961911
<i>rl217</i>	1,47632762	4,89395306
<i>rl248</i>	0,86968804	1,30847148
<i>rl259</i>	2,23654062	5,1416521
<i>rl376</i>	0,05350364	3,64955974
<i>rl382</i>	-0,21824243	3,64955974



universität
wien



Test- und
Beratungsstelle

Fakultät für Psychologie

Institut für Angewandte Psychologie:
Gesundheit, Entwicklung und
Förderung
Arbeitsbereich Psychologische Dia-
gnostik
samt Test- und Beratungsstelle
Liebiggasse 5
A-1010 Wien

T +43 (1) 4277-472 74
F +43 (1) 4277-479 06
bettina.hagenmüller@univie.ac.at
Wien, 18. April 2012

An AR Ulrike Wagner
Landesschulrat Oberösterreich
Sonnensteinstraße 20
4040 Linz

**Betrifft: Schüler(innen)-Testung mit dem AID-Gruppe – einer Gruppenversion der Intelli-
genz-Testbatterie AID 3 (Kubinger & Holocher-Ertl, in Vorb.) zur Erfassung von Intelli-
genz bei Kindern und Jugendlichen**

Bei der Intelligenz-Testbatterie AID 2 (*Adaptives Intelligenz Diagnostikum: in der aktuellen Fassung AID 2, Version 2.2*, Kubinger, 2009) handelt es sich um ein im deutschen Sprachraum bereits seit vielen Jahren sehr etabliertes psychologisches Verfahren zur Erfassung der intellektuellen Fähigkeiten von Kindern und Jugendlichen im Alter von 6 bis 15 Jahren. Insbesondere in der Schulpsychologie und Bildungsberatung wird es sehr häufig eingesetzt. Es dient vor allem der Beratung hinsichtlich schulischer Leistungsschwierigkeiten oder der Schullaufbahnberatung und ist insofern eine wesentliche Entscheidungshilfe für Schüler(innen) und deren Eltern.

Um mit dem AID tatsächlich jenes Wissen zu überprüfen, welches in heutiger Zeit von Relevanz ist, befindet sich gerade die Version AID 3 (Kubinger & Holocher-Ertl) in Vorbereitung. Eine entsprechende Eichung fand im letzten Schuljahr an Schüler(innen) im Alter von 6 bis 15 Jahren in Österreich und Deutschland statt.

Da es zur Beantwortung mancher psychologisch-diagnostischer Fragestellungen ökonomischer scheint, eine Intelligenz-Testbatterie in der Gruppe vorzugeben und somit mehrere Kinder zur gleichen Zeit testen zu können, wurde im Rahmen eines Forschungsprojekts des Arbeitsbereichs Psychologische Diagnostik, der Fakultät für Psychologie, unter der Leitung von Herrn Univ.-Prof. Dr. Mag. Klaus D. Kubinger, eine gruppentaugliche Intelligenz-Testbatterie – mit dem Arbeitstitel AID-Gruppe – für Kinder und Jugendliche der 3. bis 12. Schulstufe

entwickelt. Eine erste Version bestehend aus fünf Untertests liegt nun vor und bedarf einer ersten Erprobung im Rahmen von Schüler(innen)-Testungen in Österreich (Wien, Niederösterreich und Oberösterreich).

Es besteht zusätzlich das Forschungsinteresse, ob mittels Einzeltestung mit dem AID 3 und Gruppentestung mit dem AID-Gruppe dieselben Fähigkeiten erfasst werden können bzw. wodurch Leistungsunterschiede von Kinder und Jugendlichen in den beiden psychologisch-diagnostischen Verfahren bedingt sein könnten.

Ziel der Studie ist es somit, neben der Entwicklung eines psychologisch-diagnostischen Verfahrens (dem AID-Gruppe), herauszufinden unter welchen Rahmenbedingungen psychologische Testungen stattfinden müssen, damit Kinder und Jugendliche ihr Leistungsoptimum zeigen können.

Wir ersuchen daher Schulen in Wien, Niederösterreich und Oberösterreich um deren Unterstützung bzw. Mitwirkung bei unserem Projekt.

Die Schüler(innen)-Testung wird von Wien aus koordiniert (Projektleitung: Univ.-Prof. Dr. Mag. Klaus D. Kubinger, Projektkoordination/Ansprechpartnerin: Mag. Bettina Hagenmüller), es stehen darüber hinaus stets Ansprechpartner(innen) vor Ort zur Verfügung (Frau Rafaela Lorenz und Frau Angelika Bichler, beides Diplomandinnen des Arbeitsbereichs).

Wir bemühen uns, den Aufwand für die jeweilige Schule so gering wie möglich zu halten. Die teilnehmenden Schulen tragen ganz wesentlich zur Etablierung einer fundierten Leistungsdiagnostik bei. Damit sollen die Schüler(innen) in Zukunft optimal hinsichtlich verschiedener leistungsbezogener Aspekte beraten werden können, was wiederum den Schulen zugute kommt.

Es besteht die Möglichkeit – sollte das gewünscht sein –, den Schüler(innen) bzw. deren Eltern als Dankeschön einen kurzen schriftlichen Ergebnisbericht über kognitive Stärken und Schwächen des Kindes auszuhändigen.

Mit der Bitte um Unterstützung und freundlichen Grüßen,

Mag. Bettina Hagenmüller

Anlagen:

Ansprechpartnerin
Informationen zum AID 3 und AID-Gruppe
Untersuchungsziel und -ablauf
Stichprobenzusammensetzung
Elternbrief
Lehrer(innen)brief

Ansprechpartnerin:

Projektleitung:

Projektleiter ist **Univ.-Prof. Dr. Mag. Klaus D. Kubinger**, Leiter des Arbeitsbereichs Psychologische Diagnostik des Instituts für Angewandte Psychologie: Gesundheit, Entwicklung und Förderung, Fakultät für Psychologie, Universität Wien

Projektkoordination:

Als **Ansprechpartnerin** steht Ihnen die Projektkoordinatorin, **Mag. Bettina Hagenmüller**, zur Verfügung. Sie beantwortet alle Fragen bezüglich Inhalt und Ablauf der Schüler(innen)-Erhebung. Bitte entnehmen Sie nachfolgend die Kontaktdaten:

Mag. Bettina Hagenmüller
Institut für Angewandte Psychologie: Gesundheit, Entwicklung und Förderung
Arbeitsbereich Psychologische Diagnostik
Fakultät für Psychologie
Universität Wien
Liebiggasse 5
1010 Wien
Tel: 01 - 4277 47274
Fax: 01 - 4277 47206
E-mail: bettina.hagenmueller@univie.ac.at
Homepage: <http://psychologie.univie.ac.at/diagnostik>

Informationen zum AID 3:

Der AID 3 (Kubinger & Holocher-Ertl, in Vorb.) stellt eine überarbeitete und um einige Untertests erweiterte Version der Intelligenz-Testbatterie AID 2 (Kubinger, 2009) dar. Einerseits wurden die Aufgaben aktualisiert, um jene Fähigkeiten und jenes Wissen zu überprüfen, welches in der heutigen Zeit tatsächlich von Relevanz ist. Andererseits wurde die Intelligenz-Testbatterie, um drei neue Untertests zur Erfassung der visuellen Merkfähigkeit, des Wortschatzes und des logisch-schlussfolgernden Denkens ergänzt, um in Zukunft noch spezifischer über Stärken und Schwächen in wesentlichen Teilleistungsfähigkeiten von Kindern und Jugendlichen beraten zu können.

Informationen zum AID-Gruppe:

Der AID-Gruppe stellt eine gruppentaugliche Intelligenz-Testbatterie für Kinder und Jugendliche der 3. bis 12. Schulstufe dar. Es sollen im Wesentlichen jene Fähigkeitsbereiche des AID 2 (Kubinger, 2009) bzw. seiner noch unpublizierten Nachfolgeversion AID 3 (Kubinger & Holocher-Ertl, in Vorb.) erfasst werden.

Die aktuell zu erprobende erste Version des AID-Gruppe besteht aus fünf Untertests zur Erfassung des Alltagswissens, des Wortschatzes, der Rechenfertigkeit, des sprachlichen Abstraktionsvermögens und der Realitätssicherheit. Unter dem Punkt „Beispielaufgaben des AID-Gruppe“ wird jeweils eine Beispielaufgabe pro Untertest samt der zu messen intendierten Fähigkeit vorgestellt.

Untersuchungsablauf und -ziel:

Das Forschungsprojekt besteht aus zwei Teilen: Einerseits dient die Datenerhebung der Entwicklung einer gruppentauglichen Version des AID 3 – dem AID-Gruppe –, andererseits der Untersuchung von Leistungsunterschieden in Einzel- und Gruppentests und der Prüfung der Äquivalenz von zu messen intendierten Fähigkeiten in Einzel- vs. Gruppenvorgabe.

Um beide Fragestellungen entsprechend untersuchen zu können, bedarf es neben Gruppentestungen mit dem AID-Gruppe, noch zusätzlicher Einzeltestungen eines Teils der Kinder und Jugendlichen mit dem AID 3 und weiterer psychologisch-diagnostischer Verfahren. Beide Erhebungsarten werden mit den Bezeichnungen Teil 1 und Teil 2 nachfolgend detailliert beschrieben. Anschließend werden für beide Erhebungsarten zusätzlich relevante Informationen bezüglich des Untersuchungsablaufs beschrieben.

Die genaue Anzahl zu testender Kinder und Jugendlicher, ist dem Abschnitt „Stichprobenszusammensetzung“ zu entnehmen.

Teil 1 der Datenerhebung:

Bei Teil 1 der Datenerhebung handelt es sich ausschließlich um **Gruppentestungen** mit fünf Untertests des AID-Gruppe. Die Kinder/Jugendlichen werden von einem Testleiter bzw. einer Testleiterin im Klassenverband (in der Gruppe) getestet.

Jedes Kind bzw. jeder Jugendliche erhält ein Testheft und einen Antwortbogen, auf welchem die Lösungen zu vermerken sind.

Die Gruppentestung findet **während der Unterrichtszeit** des Kindes/Jugendlichen statt, es wird nicht in der Freizeit (Mittagspause, nach dem Unterricht etc.) des Kindes/Jugendlichen getestet. Es kommen dafür sowohl Vormittags- als auch Nachmittagstermine in Frage.

Die **Dauer der Testung entspricht in etwa 70 Minuten**, um den Kindern/Jugendlichen genügend Zeit zur Beantwortung der Aufgaben zu geben und zu verhindern, dass Kinder/Jugendliche mit einem langsamen, sorgfältigen Arbeitsstil benachteiligt und unter Zeitdruck gesetzt werden.

Die teilnehmenden Schüler(innen) sollten zwischen 9 bis 18 Jahre alt sein (siehe die genauere Stichprobenszusammensetzung weiter unten).

Teil 2 der Datenerhebung:

Um überprüfen zu können, ob mittels AID 3 und AID-Gruppe tatsächlich dieselben Fähigkeiten je Untertest erfasst werden, ist vorgesehen einen Teil der Kinder zusätzlich **vor oder nach** der AID-Gruppentestung, **einzel**n mit dem AID 3 zu testen.

Es wurde dabei die minimale Anzahl zu testender Kinder/Jugendlichen vorab berechnet, um noch statistisch gültige Aussagen treffen zu können und gleichzeitig den Schulen und deren Unterricht möglichst wenig zu stören (siehe dazu die Stichprobenszusammensetzung weiter unten).

In der Einzeltestung werden den Kindern folgende sechs Untertests des AID 3 vorgegeben: *Alltagswissen, Realitätssicherheit, Angewandtes Rechnen, Synonyme Finden, Antonyme Finden und Funktionen Abstrahieren*.

Zusätzlich soll die Lern- und Leistungsmotivation, das Selbstkonzept und andere leistungsbezogene Persönlichkeitseigenschaften mit folgenden psychologisch-diagnostischen Verfahren bzw. Teilen davon erfasst werden: SELLMO (*Skalen zur Erfassung der Lern- und Leistungsmotivation*; Spinath, Stiensmeier-Pelster, Schöne & Dickhäuser, 2002), SESSKO (*Skalen zur*

Erfassung des schulischen Selbstkonzepts; Schöne, Dickhäuser, Spinath & Stiensmeier-Pelster, 2002), DISK-Gitter (*Differentielles Schulisches Selbstkonzept-Gitter*; Rost, Sparfeldt & Schilling, 2007), vier Skalen des PFK 9-14 (*Persönlichkeitsfragebogen für Kinder und Jugendliche zwischen 9 und 14 Jahren*; Seitz & Rausche, 2004) und drei Skalen des DAI (*Differentielles Leistungsangst Inventar*; Rost & Schermer, 1997).

Den Kindern/Jugendlichen wird in Abhängigkeit ihres Alter und der besuchten Schulstufe ein Teil dieser Verfahren im Anschluss an die Testung mit den sechs Untertests des AID 3 zusätzlich vorgegeben; mit dem Ziel, Rückschlüsse über mögliche Leistungsunterschiede in Einzel- und Gruppentestung zu erhalten.

Für die **Einzeltestungen** würde ein Teil der Schüler(innen) von einem Testleiter bzw. einer Testleiterin (Diplomand(innen) des Arbeitsbereichs, die zuvor eine umfangreiche Schulung absolviert haben) gesondert getestet werden. Das Kind/der Jugendliche wird dabei von der Klasse abgeholt und in einem von der Schule zur Verfügung gestellten Raum (freies Klassenzimmer, Bibliothek, Arzttraum etc.) getestet. Die Testung findet **während der Unterrichtszeit** statt und dauert je nach Alter und Arbeitsstil des Kindes/Jugendlichen zwischen **30 und maximal 50 Minuten**.

Während dieser Zeit ist es den Kindern/Jugendlichen selbstverständlich möglich – je nach individuellem Bedarf – **kurze Pausen** einzulegen. Nach Abschluss der Testung werden die Kinder wieder in ihre Klasse zurückgebracht. Die teilnehmenden Schüler(innen) sollten zwischen 9 und 15 Jahre alt sein (siehe die genauere Stichprobenzusammensetzung weiter unten).

Zusätzliche Informationen zum Ablauf der Testungen, die für Teil 1 und 2 ident sind:

Natürlich werden nur Kinder/Jugendliche getestet, deren **Eltern ihr Einverständnis** gegeben haben. Der Elternbrief ist in Deutsch abgefasst und enthält wichtige Informationen an die Erziehungsberechtigten.

Wenn klar ist, welche Kinder/Jugendlichen der jeweiligen Schule das Einverständnis der Eltern haben, werden **Termine fixiert**, an denen Testleiter(innen) zum Testen kommen dürfen. Bei der Terminvereinbarung legen wir großen Wert darauf, den Ablauf des regulären Unterrichts so wenig wie möglich zu unterbrechen. Wir werden versuchen – im Rahmen der Möglichkeiten – Rücksicht auf die Wünsche der einzelnen Schulen zu nehmen.

Testleiter(innen) sind mehrere Diplomanden(innen) des Arbeitsbereichs Psychologische Diagnostik der Fakultät für Psychologie, die an der Erstellung der Aufgaben für den AID-Gruppe selbst mitgearbeitet haben und sowohl theoretisch als auch aufgrund ihrer praktischen Erfahrung besondere Eignung für die Testung der Kinder und Jugendlichen aufweisen.

Natürlich werden die gewonnenen Daten im Sinne des **Datenschutzes** ausschließlich für wissenschaftliche Zwecke verwendet. Daher sind die Testleiter(innen) angewiesen, jedem Kind einen Probandencode, der sich aus den Initialen des/der Testleiter(in) und einer Laufnummer zusammensetzt zu **anonymisieren**.

Wünschen die Eltern als Dankeschön einen kurzen schriftlichen Ergebnisbericht über Stärken und Schwächen der getesteten Schüler(innen), erfolgt die Anonymisierung erst nach der Erstellung dieses Befundes, ansonsten gleich nach Beendigung der Testung.

Stichprobenzusammensetzung:

Teil 1 der Datenerhebung:

Gruppentestungen

Alter	Schulform	Schulstufe	Anzahl
8-9 Jahre	Volksschule	3	30
9-10 Jahre	Volksschule	4	30
10-11 Jahre	Gymnasium oder Hauptschule	5	30
11-12 Jahre	Gymnasium oder Hauptschule	6	30
12-13 Jahre	Gymnasium oder Hauptschule	7	20
13-14 Jahre	Gymnasium oder Hauptschule	8	20
14-15 Jahre	Gymnasium	9	20
15-16 Jahre	Gymnasium	10	45
16-17 Jahre	Gymnasium	11	45
Insgesamt			270

Teil 2 der Datenerhebung:

Einzel- und Gruppentestungen (Kinder und Jugendliche, die vor oder im Anschluss an die Gruppentestung auch einzeln getestet werden)

Alter	Schulform	Schulstufe	Anzahl
8-9 Jahre	Volksschule	3	18
9-10 Jahre	Volksschule	4	18
10-11 Jahre	Gymnasium oder Hauptschule	5	18
11-12 Jahre	Gymnasium oder Hauptschule	6	18
12-13 Jahre	Gymnasium oder Hauptschule	7	12
13-14 Jahre	Gymnasium oder Hauptschule	8	12
14-15 Jahre	Gymnasium	9	12
Insgesamt			108



Frau
Mag. Bettina Hagenmüller
Universität Wien – Fakultät für Psychologie
Institut für angewandte Psychologie
Liebiggasse 5
1010 Wien

Bearbeiterin:
Ulrike Wagner
Tel: 0732/7071-2321
Fax: 0732/7071-2330
E-mail: lsr@lsr-ooe.gv.at

Ihr Zeichen

vom
24. 4. 2012

Unser Zeichen
B5 – 14/16 – 2012

vom
27. 4. 2012

**Schüler/innen-Testung mit dem AID-
Gruppe und dem AID 3 im Rahmen
eines Forschungsprojektes**

Sehr geehrte Frau Mag. Hagenmüller!

Sie haben beim Landesschulrat für OÖ. um Genehmigung der gegenständlichen Untersuchung an Pflichtschulen und allgemeinbildenden höheren Schulen in Oberösterreich angesucht.

Nach Prüfung der Untersuchungsunterlagen genehmigt der Landesschulrat für OÖ Ihre Testung unter den üblichen Bedingungen:

- Freiwilligkeit der Teilnahme
- Information der Erziehungsberechtigten
- Einhaltung der Datenschutzbestimmungen
- Übermittlung des Endberichtes an den Landesschulrat für OÖ.

Bei Ihrer Kontaktaufnahme mit den Schulen verweisen Sie bitte auf diese Genehmigung.

Mit freundlichen Grüßen

Für den Amtsführenden Präsidenten
Mag. Girzikovsky eh.



LEBENS LAUF

Rafaela Lorenz
Erne-Seder-Gasse 8/2/2.102
1030 Wien
geboren am 08.06.1987
+43 / 664/ 2276516
österreichische Staatsbürgerin
lorenz.rafaela@gmail.com
Familienstand: ledig

BERUFSERFAHRUNG

10/2012	Mystery Shopper, AC Nielson Freie Dienstnehmerin
09/2010 – 10/2012	Einzel und Familienbetreuerin, Mag^a Eva Lorenz Angestellte Fälle der Jugendwohlfahrt Ried im Innkreis - Betreuung von Kindern und Jugendlichen - Ordnung der Lebensverhältnisse - Unterstützung beim Erwerb von social skills - Selbstfindung und Selbststärkung - Hilfe bei Problembewältigung - Unterstützung bei schwerwiegenden psychischen Störungen
07/2010 – 09/2010	Promoterin, Public Realtions GmbH Angestellte - Recyclingpromotion in Schwimmbäder - Zielgruppe: Kinder und Jugendliche
03/2010 – 08/2010	Praktikantin, Kinderschutzzentrum „die möwe“ Angestellte - Anamnesegespräche - Durchführung, Auswertung und Interpretation von psychologischen Testverfahren - Verfassen von psychologischen Befunden - Protokollierung von Telefongesprächen der Hotline - Prozessbegleitung - Betreuung von Kindern und Jugendlichen
2008 – 2009	Dateneingabe, XCOM-Datensysteme GmbH Angestellte - Eingabe von Krankenscheinen diverser Ärzte
2007	Nachhilfe und Nachmittagsbetreuung bei Kindern und Jugendlichen
2006	Videotheksmitarbeiterin, rent-a-movie Angestellte

2008 – 2009

Promoterin, Easystaff

Freie Mitarbeiterin

- Eventbetreuung
- Flyerverteilung

AUSBILDUNG

2005-2012

Diplomstudium Psychologie an der Universität Wien

- Schwerpunkt: Klinische Psychologie
- Diplomprüfungsfächer: Diagnostik und Differentielle Psychologie
- Diplomarbeit: Machbarkeitsstudie zum Untertest „Alltagswissen“ der AID-Gruppe: Erste Befunde zur psychometrischen Güte und konvergenten Validität (Betreuer: Univ.-Prof. Dr. Mag. Klaus D. Kubinger)

1997 – 2005

Bundesgymnasium Schärding

- Sprachen: Englisch, Latein, Französisch
- Wahlfächer: Geschichte, Psychologie
- Reifeprüfung schriftlich: Englisch, Mathematik, Deutsch
- Reifeprüfung mündlich: Latein, Musik, Englisch, Deutsch

1993-1997

Volksschule, 4875 Suben/Inn

SPRACHKENNTNISSE

Deutsch	Muttersprache
Englisch	fließend in Wort und Schrift
Französisch	Grundkenntnisse

PC KENNTNISSE

SPSS	sehr gute Anwenderkenntnisse
R	gute Anwenderkenntnisse
MS Office	gute Anwenderkenntnisse

SKILLS

Kreativität, hohe soziale Kompetenz, Engagement, Einfühlungsvermögen, zielorientiert, koordiniert, loyal, teamfähig, genau, verantwortungsbewusst, pünktlich

Seminarteilnahmen

- Kongress: Psychosomatik, Psychiatrie, Psychologie & Psychotherapie 2011-Psychosomatisches Zentrum Waldviertel

- AID-Zertifizierungskurs (theoretischer Teil) -2011
- Powerful Girls go politics (Rhetoriktraining) -2001

Führerschein

- Klasse B

Publikation

- Hagenmüller, B., Kubinger, K. D., Bichler, A. E. & Lorenz, R. (2012). *Zur Dimensionalität von und Leistungsunterschieden in Einzel- und Gruppentests*. 48. Kongress der Deutschen Gesellschaft für Psychologie (DGPs), Bielefeld, 23.September-27.September 2012. (Vortrag)