



universität
wien

Diplomarbeit

Titel der Arbeit

Machbarkeitsstudie zum Untertest 2
„Realitätssicherheit“
der Intelligenz-Testbatterie AID-Gruppe

Verfasserin

Angelika Erika Bichler

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, März 2013

Studienkennzahl: A 298

Studienrichtung: Psychologie

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

Danksagung

Großer Dank gilt meinem Betreuer, Herrn Univ.-Prof. Dr. Mag. Klaus D. Kubinger, für die Möglichkeit, meinen Beitrag zu diesem interessanten und wichtigen Großprojekt leisten zu dürfen. Die Bearbeitung dieses Themas war fordernd und lehrreich, ich würde diese Erfahrung nicht missen wollen!

Ein besonderer Dank gebührt auch Projektkoordinatorin Fr. Mag. Bettina Hagenmüller, die sprichwörtlich „rund um die Uhr“, mit großem Engagement während des gesamten Werdungsprozesses dieser Arbeit, mit ihrem fachlichen Rat zur Verfügung stand.

Ich bedanke mich bei sämtlichen Direktoren und Direktorinnen, der Lehrerschaft, den Eltern und nicht zuletzt den vielen Schülern und Schülerinnen, die an dem Projekt in Oberösterreich und Niederösterreich teilgenommen haben. Auch den vielen Testleitern und Testleiterinnen sei an dieser Stelle ausdrücklich gedankt.

Ich bedanke mich bei Serra Ünal und Mag. phil. Klaus Vogelauer für das Korrekturlesen dieser Arbeit.

Ich möchte ganz besonders meiner Mutter, Eva Bichler, danken, die mich während des ganzen Studiums und der Anfangsphase dieser Arbeit, emotional in unvergleichlicher Weise unterstützt hat. Voller Interesse für dieses Fach, Verständnis und Geduld, hat sie mich beinahe bis zur Vollendung dieser Arbeit begleiten können. Leider war es ihr nicht vergönnt, den Abschluss dieser Arbeit miterleben zu dürfen.

Ich bedanke mich bei meinem Vater Rudolf Bichler, meinem Bruder Andreas Bichler, und meiner Schwägerin Mag. phil. Barbara Bichler-Schätz, für die langjährige finanzielle Unterstützung während meiner Studienzeit.

Ich danke auch ganz besonders Nadja Trimmel für ihre emotionale und finanzielle Unterstützung in den letzten Jahren. Sie war mir eine gigantische Stütze in dem besonders schwierigen Jahr 2012.

Vorwort

Die vorliegende Diplomarbeit stellt einen Teil des Großprojekts AID-Gruppe (Kubinger & Hagenmüller, in Vorb.) des Instituts für Angewandte Psychologie: Gesundheit, Entwicklung und Förderung, im Arbeitsbereich Psychologische Diagnostik, der Universität Wien, dar. Die Projektleitung obliegt Herrn Univ.-Prof. Dr. Mag. Klaus D. Kubinger, wobei Fr. Mag. Bettina Hagenmüller als Projektkoordinatorin fungiert.

Die für diese Untersuchung relevanten 5 Untertests, der Intelligenz-Testbatterie AID-Gruppe, waren: *Alltagswissen*, *Realitätssicherheit*, *Angewandtes Rechnen*, *Antonyme Finden* und *Funktionen Abstrahieren*. Die Autorin war für die Konstruktion des Untertests *Realitätssicherheit* zuständig. Zeitgleich übernahm Fr. Lorenz (Lorenz, 2012) im Rahmen ihrer Diplomarbeit die Überarbeitung und Entwicklung der Items für den Untertest *Alltagswissen*. Teil der Untersuchung war außerdem die zusätzliche Vorgabe der entsprechenden Untertests des AID 3 (Adaptives Intelligenz Diagnostikum Version 3.1; Kubinger & Holocher-Ertl, 2013, in Druck), die parallel zu den Erhebungen der Kalibrierungsstichprobe des AID-Gruppe erfolgen konnte.

Um eine möglichst große Stichprobe von Testpersonen akquirieren zu können, fanden zeitgleiche Testungen der fünf Untertests sowohl in Oberösterreich als auch in Niederösterreich statt. Es sei hier darauf hingewiesen, dass sämtliche für die Datenerhebungen relevanten Materialien mit jenen von Fr. Lorenz ident sind. Dies betrifft vordergründig Elternbriefe, Lehrerbriefe, Projektbeschreibungen, offizielle Ansuchen etc. Sowohl die Datenerhebung und Dateneingabe als auch die Rückmeldung der Testergebnisse erfolgten zeitgleich mit Fr. Lorenz.

Da sich in den vergangenen drei Jahren mehrere Diplomarbeiten mit der Entwicklung des AID 3 und der Gruppenversion des AID beschäftigt haben, können zwischen den Arbeiten Ähnlichkeiten in Theorie- und Testbeschreibungen auftreten. Aufgrund der Verständlichkeit und der Nachvollziehbarkeit der wissenschaftlichen Überlegungen dieses Projekt betreffend, kann auf deren Darstellung jedoch nicht verzichtet werden.

Abschließend sei erwähnt, dass am 48. Kongress der Deutschen Gesellschaft für Psychologie im September 2012 bereits Teile der Ergebnisse öffentlich vorgestellt wurden.

Abstract (Deutsch)

Das Ziel dieser Arbeit ist, eine sprachliche Version des Untertests *Realitätssicherheit* aus dem Individualverfahren AID 3 für die in Entwicklung befindliche Intelligenz-Testbatterie mit dem Arbeitstitel „AID-Gruppe“ zu konstruieren. Im Gegensatz zu der bildhaften und weitestgehend sprachfreien Vorgabe im AID 3, erfolgt die Erfassung im AID-Gruppe durch eine verbalschriftliche Methode. Dabei haben Testpersonen im Alter von 8 bis 19 Jahren die Aufgabe, fehlende Wortteile von zusammengesetzten Nomen im freien Antwortformat zu ergänzen. Neben der Gruppenvorgabe des konstruierten Untertests wurde einem Teil der Stichprobe, durch Verwendung eines *balancierten* Versuchdesigns, zusätzlich der Untertest *Realitätssicherheit* des AID 3 vorgegeben. Insgesamt wurden 810 Testpersonen mit dem AID-Gruppenverfahren getestet, wovon 207 Testpersonen von der 3. bis zur 9. Schulstufe zusätzlich mit den entsprechenden Untertests des AID 3 getestet wurden, um die *konvergente* Validität zu bestimmen. Die Stichprobe wurde in Niederösterreich und in Oberösterreich akquiriert. In der darauf folgenden Datenanalyse wurde die Geltung des *Rasch*-Modells für den konstruierten Untertest überprüft. Nach dem schrittweisen Ausschluss von 13,1 % der konstruierten Items, konnte *a posteriori* Modellgültigkeit erreicht werden. Die nachfolgende Bestimmung der *konvergenten* Validität der beiden Versionen des Untertests, ergab einen niedrigen positiven Zusammenhang ($r = 0,317$). 10 % der Varianz des einen Tests konnte durch den anderen erklärt werden. Eine weitere Validierungsstudie zur Klärung des mit dem Untertest erfassbaren Konstrukts, insbesondere im Zusammenhang mit verbalem Aufgabematerial, wäre zu empfehlen. Für die Konstruktion eines geeigneteren Pendantes des Untertests *Realitätssicherheit* in Anlehnung an den AID 3 wäre die Entwicklung von bildhaftem Aufgabenmaterial anzudenken.

Schlüsselwörter: AID 3, AID-Gruppe, *Realitätssicherheit*, *Rasch*-Modell, *balanciertes* Versuchsdesign, *konvergente* Validität

Abstract (English)

This study aims to develop a verbal version of the subtest “*Realitätssicherheit*” from the individual test battery AID 3 for the intelligence test battery „AID-Gruppe”, which will be used in group settings. In contrast to the visual and – as far as possible - non-verbal presentation used in AID 3, AID-Gruppe uses a verbal-written method. Subjects between the ages of 8 and 19 are given the task of completing missing word pieces of complex nouns in open-ended questions. With the help of a *balanced* test design a part of the sample group was presented with the subtest *Realitätssicherheit* of AID 3 in addition to the group-version of the created subtest. The collected data allowed the calculation of the *convergent* validity for both subtests. A total of 810 subjects were tested with AID-Gruppe, of which 207 subjects between the 3rd and 9th grades of education were also tested with AID 3, to identify the *convergent* validity. The sample was acquired in Upper and Lower Austria. The applicability of the *Rasch*-model to the constructed subtest was examined according to the following data analysis. After the successive elimination of inappropriate items, comprising 13,1 % of the total items, the remaining itempool shows a fit to the *Rasch*-model. The following evaluation of *convergent* validity resulted in a low positive correlation between both subtests ($r = 0,317$). 10 % of empirical variance of one subtest was explained by the other subtest. Further validity assessments are suggested to clarify the underlying construct of the subtest, particularly in correlation with verbal test material. The use of visual stimuli should be considered for the development of a counterpart to the subtest *Realitätssicherheit* for the test battery AID-Gruppe according to AID 3.

Keywords: AID 3, AID-Gruppe, *Realitätssicherheit*, *Rasch*-model, *balanced* test design, *convergent* validity

meiner Mutter gewidmet

1947-2012

Inhaltsverzeichnis

I. Einleitung	1
II. Theoretischer Teil	3
1. Kurzdarstellungen der Intelligenz-Testbatterien	
AID 2, AID 3 und AID-Gruppe	3
1.1 Adaptives Intelligenz Diagnostikum 2 (AID 2, Version 2.2)	3
1.2 Adaptives Intelligenz Diagnostikum 3 (AID 3, Version 3.1)	5
1.3 AID-Gruppe (Kubinger & Hagenmüller, in Vorb.)	6
2. Das Konstrukt <i>Realitätssicherheit</i>	7
2.1 <i>Realitätssicherheit</i> der AID-Individualverfahren	7
3. Theoretische Fundierung zur Testkonstruktion	8
3.1 Rationale Testkonstruktion	8
3.2 Externale Testkonstruktion	8
3.3 Induktive Testkonstruktion	8
3.4 Prototypenansatz	9
3.5 Vergleich der Methoden	9
4. Techniken für die Indikatorensuche des Konstrukts	9
5. Regelgeleitete Itemkonstruktion	11
5.1 Richtlinien für den sprachlichen Aufbau von Aufgaben	11
6. Einzel- vs. Gruppentestung	12
6.1 Vor- und Nachteile einer Individualtestung	12
6.2 Vor- und Nachteile einer Gruppentestung	13
7. Das dichotome logistische Testmodell von Rasch	15
III. Empirischer Teil	18
8. Ziel der Untersuchung	18
9. Konstruktion des Untertests <i>Realitätssicherheit</i>	19
9.1 Umlegung der zu messenden Fähigkeit in verbales Material	19
9.2 Geltungsbereiche des AID 3 und des AID-Gruppe	20
9.3 Regeln für die Itemkonstruktion	20
9.3.1 Wortfindung	20
9.3.2 Satzbau und Inhalt	22
9.3.3 Inhaltliche Kategorien	23

9.3.4 Einschätzung der Schwierigkeitsstufen	24
9.4 Antwortformat	25
9.5 Itempool und Itemeigenschaften	26
9.6 Testdesign und Testformen (AID-Gruppe)	27
9.7 Balancierter Versuchsplan	29
9.8 Instruktion und Testvorgabe des AID-Gruppe.....	30
10. Datenerhebung	32
10.1 Akquirierung der Stichprobe	32
10.2 Durchführungszeitraum und Testdauer.....	33
10.3 Testleiter/innen	33
11. Stichprobenbeschreibung	35
12. Ergebnisdarstellung	38
12.1 Rasch-Modell-Analysen des Itempools	38
12.2 Itemschwierigkeitsparameter-Realitätssicherheit	46
12.3 Bestimmung der <i>konvergenten</i> Validität.....	48
13. Diskussion und Ausblick	50
14. Zusammenfassung	54
IV. Literaturverzeichnis.....	56
V. Anhang	60

Abbildungsverzeichnis

Abbildung 1:	Instruktionsbeispiel 1 – <i>Realitätssicherheit</i> – AID-Gruppe.....	20
Abbildung 2:	Beispielsätze für unterschiedliche Artikel der Wortteile	21
Abbildung 3:	Instruktionsbeispiel 2 – <i>Realitätssicherheit</i> – AID-Gruppe.....	22
Abbildung 4:	Beispielsatz einer Wortposition in der Mitte des Satzes	22
Abbildung 5:	Antwortfelder am Antwortbogen für den Untertest <i>Realitätssicherheit</i>	26
Abbildung 6:	Testleiteranweisung – <i>Realitätssicherheit</i> des AID-Gruppe	31
Abbildung 7:	Der grafische Modelltest des Teilungskriteriums <i>Score</i> vor dem Ausschluss eines Items.	40
Abbildung 8:	Der grafische Modelltest des Teilungskriteriums <i>Geschlecht</i> vor dem Ausschluss eines Items.	40
Abbildung 9:	Der grafische Modelltest des Teilungskriteriums <i>Schulstufe</i> (VS-HS) vor dem Ausschluss eines Items.	41
Abbildung 10:	Der grafische Modelltest für das Teilungskriterium <i>Score</i> nach dem Ausschluss von 8 Items.	44
Abbildung 11:	Der grafische Modelltest für das Teilungskriterium <i>Geschlecht</i> nach dem Ausschluss von 8 Items.	45
Abbildung 12:	Der grafische Modelltest für das Teilungskriterium <i>Schulstufe</i> (VS-HS) nach dem Ausschluss von 8 Items.	45
Abbildung 13:	Die Verteilung der Schwierigkeitsparameter der 53 <i>Rasch</i> -Modell- konformen Items des Untertests <i>Realitätssicherheit</i>	47
Abbildung 14:	Streudiagramm der Personenparameter für den Untertest <i>Realitätssicherheit</i> des AID 3 und des AID-Gruppe.	49

Tabellenverzeichnis

Tabelle 1:	Vorgabemethode je Untertest des AID 2	4
Tabelle 2:	Operationale Definitionen je Untertest des AID 3 (vgl. Kubinger, 2009b, S. 195).....	5
Tabelle 3:	Häufigkeiten der Charakteristika für 63 Items.....	27
Tabelle 4:	Vorgabe der Testhefte je Schulform und Schulstufe	28
Tabelle 5:	Balancierter Versuchsplan für NÖ und OÖ gesamt.....	29
Tabelle 6:	Die Verteilung der Testpersonen auf die Bundesländer Niederösterreich und Oberösterreich ($n=810$)	35
Tabelle 7:	Die Verteilung der Testpersonen auf die verschiedenen Schultypen ($n=810$).....	36
Tabelle 8:	Die Verteilung der Testpersonen auf das Alter in Jahren ($n=810$)	36
Tabelle 9:	Der balancierte Versuchsplan des AID 3 mit den Verteilungen auf die Schulstufen ($n=209$)	37
Tabelle 10:	Die Ergebnisse der LRTs vor dem Ausschluss eines Items	39
Tabelle 11:	Die Ergebnisse der LRTs nach dem Ausschluss von 8 Items	43
Tabelle 12:	Itemschwierigkeitsparameter der <i>Rasch</i> -Modell-konformen Items des Untertests <i>Realitätssicherheit</i> ; Rangreihung nach Schwierigkeit	46
Tabelle 13:	Die Rangkorrelation der Personenparameter für die Untertests <i>Realitätssicherheit</i> des AID 3 und des AID-Gruppe	48

I. Einleitung

Die vorliegende Arbeit beschäftigt sich mit einem Teilbereich des Projekts *AID-Gruppe*. Das Ziel dieses Projekts ist, für die bewährte Intelligenz-Testbatterie AID (bzw. dessen Folgeversionen) eine entsprechende Gruppenversion zu kreieren und diese umfassend auf dessen Tauglichkeit zu überprüfen.

Das Adaptive Intelligenz Diagnostikum 2 (AID 2, Version 2.2; Kubinger, 2009a) stellt ein bewährtes und anerkanntes Instrument der Intelligenz erfassung von Kindern und Jugendlichen im deutschsprachigen Raum dar. Leistungsuntersuchungen von Schülern und Schülerinnen zum Zwecke einer Bildungs- und Laufbahnberatung sowie das Erkennen von Förderbedarf in Teilbereichen der Leistungsfähigkeit, gewinnen immer mehr an Bedeutung. Ein Gelingen einer Umsetzung dieses umfassenden Individualverfahrens in eine qualitativ hochwertige Gruppenversion würde in seiner Gestalt als Intelligenz-Testbatterie eine Vorreiterfunktion einnehmen.

In weiterer Folge der laufenden Untersuchungen zu diesem Thema soll die Gruppenvorgabe neben dem Papier-Bleistift-Verfahren auch am Computer anhand adaptiver Testform, dem sogenannten *tailored testing* (vgl. Kubinger, 2009b), vorgegeben werden können. Diesem Ziel gehen jedoch wiederholte Itementwicklungen und Voruntersuchungen voraus.

Gruppenvorgaben haben im Gegensatz zu Individualverfahren verschiedene Vor- und Nachteile. Eine genaue Aufschlüsselung dieser Argumente wird in der vorliegenden Arbeit behandelt. Der wohl gewichtigste positive Aspekt der Gruppentestung liegt in der finanziellen Ökonomie der Testvorgabe im Sinne der Wirtschaftlichkeit. Durch die gleichzeitige Testung einer größeren Gruppe an Testpersonen können durch einen verhältnismäßig deutlich geringeren Zeitaufwand sowie kalkulierbaren Kosten für psychologisches Fachpersonal flächendeckende Untersuchungen durchgeführt werden. Dabei sei noch ein wesentlicher Aspekt erwähnt, nämlich der der Sachlichkeit:

„Außer Frage steht, dass die Sachlichkeit vorrangig ist, die Aufwandsminimierung nachrangig. Die Wirtschaftlichkeit eines Tests darf bei einer konkreten Fragestellung erst dann zu Buche schlagen und mit

der eines anderen konkurrieren, wenn der Einsatz (auch) dieses Tests sachlich gerechtfertigt ist, er die gestellte Frage sachlich beantwortet.“
(Kubinger 2009b, S. 99)

Diese Arbeit beschäftigt sich ebenso mit der regelgeleiteten Entwicklung der Items und der Prüfung der Modellkonformität des Itempools. Außerdem soll die *konvergente* Validität als Teilbereich der Konstruktvalidität zwischen dem entsprechenden Untertest im Individualverfahren und dem Itempool des Gruppenverfahrens bestimmt werden. Das erklärte Ziel bestand demnach darin, Testaufgaben zu konstruieren, die nachweislich inhaltliche Gültigkeit aufweisen und dasselbe Konstrukt *eindimensional* erfassen.

Im nachfolgenden theoretischen Teil werden die verschiedenen Formen des Adaptiven Intelligenz Diagnostikums überblicksartig dargestellt sowie auf die theoretischen Überlegungen zur Konstruktion des Untertests *Realitätssicherheit* eingegangen. Des Weiteren werden die verschiedenen Aspekte von Gruppen- und Individualverfahren diskutiert sowie Erklärungen zum dichotom-logistischen Testmodell von Rasch geliefert.

Der empirische Teil skizziert den Vorgang der regelgeleiteten Itemkonstruktion des Untertests *Realitätssicherheit*, dessen Untersuchung anhand der akquirierten Stichprobe sowie deren Prüfung mittels *Rasch*-Modell-Analysen. Den Abschluss der statistischen Auswertungen der Daten stellt die Bestimmung der *konvergenten* Validität der beiden Versionen dieses Untertests dar.

II. Theoretischer Teil

1. Kurzdarstellungen der Intelligenz-Testbatterien AID 2, AID 3 und AID-Gruppe

Nachfolgend werden alle drei Verfahren überblicksartig dargestellt, wobei besonders auf Unterschiede eingegangen wird. Da sich der AID 3 (Adaptives Intelligenz Diagnostikum - Version 3.1; Kubinger & Holocher-Ertl, 2013, in Druck) gerade in der Version 3.1 in Druck befindet, werden inhaltliche und funktionelle Beschreibungen aus der Vorgängerversion AID 2 (Adaptives Intelligenz Diagnostikum - Version 2.2; Kubinger, 2009a) verwendet, da jene auch für diese Version Gültigkeit haben. Für eine vollständige inhaltliche Beschreibung des jeweiligen psychologisch-diagnostischen Verfahrens sei der/die interessierte Leser/in auf Kubinger (2009a), sowie auf das zukünftig publizierte Manual des AID 3 und jenes der Version AID-Gruppe verwiesen.

1.1 Adaptives Intelligenz Diagnostikum 2 (AID 2, Version 2.2)

Die erste Version dieses Verfahrens erschien im Jahre 1985 unter dem Titel AID. Danach folgte eine adaptierte, vollständig überarbeitete und neu normierte Fassung, nämlich das AID 2 (Kubinger & Wurst, 2000). Nach erneuter Eichung des Verfahrens entstand die Version AID 2 in der Version 2.2 im Jahre 2009 (vgl. Kubinger, 2009a).

Das AID 2 stellt eine Intelligenz-Testbatterie im Geltungsbereich zwischen 6;0 und 15;11 Jahren dar, dessen elf Untertests sowie drei fakultativ vorzugebende Zusatztests zum Screening ausgewählter Teilleistungsstörungen ihre Anwendung finden. Es handelt sich um ein Individualverfahren in Papier-Bleistift-Form. Das AID 2 eignet sich für die Berufs- und Ausbildungsberatung, für Fragestellungen der Klinischen Psychologie und die Testung fremdsprachiger Testpersonen, welche durch die sprachfreie Vorgabemöglichkeit bestimmter Untertests ermöglicht wird (Kubinger, 2009b).

„Inhaltlich ist die Testbatterie AID 2 innovatorisch am weltweit verbreiteten, bis auf das Jahr 1939 zurückgehende Testkonzept von *Wechsler* orientiert.“ (Kubinger 2009b, S. 194) Da bestimmte Untertests als förderabhängig und

andere als förderunabhängig zu bezeichnen sind, können diese im Sinne von *Cattell's* fluiden und kristallisierten Faktoren der Intelligenz verstanden werden (vgl. z.B. Cattell, 1963; Horn & Cattell, 1966).

Die explizit hervorzuhebende Methode dieser Intelligenz-Testbatterie betrifft das *adaptive Testen*, welches in 8 Untertests realisiert wird. Dies geschieht durch die verzweigte Vorgabe von Aufgabenblöcken, dem sogenannten „*branched testing*“. Beim *adaptiven Testen*¹ werden den Testpersonen nur jene Aufgaben vorgegeben, welche auch ihrem persönlichen Leistungsniveau entsprechen. Dadurch können Frustrationserlebnisse seitens der Testpersonen während der Testsituation minimiert werden (Kubinger & Wurst, 2000).

Tabelle 1: Vorgabemethode je Untertest des AID 2

Untertest	Vorgabemethode
1 Alltagswissen	adaptiv
2 Realitätssicherheit	konventionell
3 Angewandtes Rechnen	adaptiv
4 Soziale und sachliche Folgerichtigkeit	adaptiv
5 Unmittelbares Reproduzieren-numerisch	konventionell
6 Synonyme Finden	adaptiv
7 Kodieren und Assoziieren	konventionell
8 Antizipieren und Kombinieren-figural	adaptiv
9 Funktionen Abstrahieren	adaptiv
10 Analysieren und Synthetisieren-abstrakt	adaptiv
11 Soziales Erfassen und Sachliches Reflektieren	adaptiv
5a Unmittelbares Reproduzieren-figural/abstrakt	konventionell
5b Merken und Einprägen	konventionell
10a Strukturieren-visumotorisch	konventionell

¹ Neben dem motivationalen Problem ergibt sich noch der entscheidende Vorteil der Ökonomie während der Testung, da nur informative Aufgaben vorgegeben werden. Nämlich jene, die fein abgestufte Grade der zu messenden Fähigkeit unterscheiden können (Kubinger & Wurst, 2000).

1.2 Adaptives Intelligenz Diagnostikum 3 (AID 3, Version 3.1)

Statt einer weiteren Überarbeitung des AID 2, wurde im Jahre 2009 damit begonnen das AID 3 zu entwickeln. Neben der Konstruktion neuer Aufgaben je Untertest wurden auch neue Untertests in die Intelligenz-Testbatterie integriert. Dabei sei insbesondere auf die Untertests *Antonyme Finden* und *Formale Folgerichtigkeit*² hingewiesen.

Nachfolgend werden die gemessenen *operationalen Definitionen* jener fünf Untertests aufgelistet, welche im Zuge der vorliegenden Diplomarbeit vorgegeben wurden. Zusätzlich wird angegeben, ob verbal-akustische oder manuell-visuelle Fähigkeiten pro Untertest gemessen werden.

Tabelle 2: Operationale Definitionen je Untertest des AID 3 (vgl. Kubinger, 2009b, S. 195)

Untertest	Operationale Definitionen (sinngemäß)
Alltagswissen Verbal-akustisch	Es wird die Fähigkeit gemessen, sich Sachkenntnisse über Inhalte anzueignen, die in der heutigen Gesellschaft alltäglich sind. Dabei werden der Testperson mündliche Wissensfragen gestellt.
Realitätssicherheit Manuell-visuell	Der Untertest Realitätssicherheit soll prüfen, inwieweit die Wirklichkeit um Dinge des Alltags verstanden wird, bzw. kontrolliert werden kann. Dabei sollen die Testpersonen fehlende Elemente auf Bildkarten erkennen.
Antonyme Finden Verbal-akustisch	Es wird die Fähigkeit gemessen, die Bedeutung sprachgebundener Begriffe zu erfassen, indem das Gegenteil zu vorgegebenen Begriffen gefunden werden soll. Es wird hierbei die Größe des Wortschatzes bzw. das Sprachvermögen erfasst.
Angewandtes Rechnen Verbal-akustisch	Es wird die Fähigkeit gemessen, (weitestgehend) unabhängig von schulischen Rechenfertigkeiten, alltägliche Problemstellungen durch Anwendung passender Rechenoperationen zu lösen.
Funktionen Abstrahieren Verbal-akustisch	Es wird die Fähigkeit erfasst, durch Abstraktion zu einer Begriffsbildung zu gelangen. Die Testperson soll aus zwei Begriffen deren gemeinsame Funktion erschließen.

² In den voran gegangenen Versionen des AID gab es keine integrierte Möglichkeit, die Fähigkeit „schlussfolgerndes Denken“ im figuralen Bereich zu untersuchen. Durch die Erweiterung um dieses Intelligenzkonstrukt kann zukünftig eine umfassendere Untersuchung bzw. Leistungsbeurteilung durch die alleinige Vorgabe des AID 3 erfolgen.

1.3 AID-Gruppe (Kubinger & Hagenmüller, in Vorb.)

Mit der Entwicklung der Intelligenz-Testbatterie AID-Gruppe wird angestrebt, eine qualitativ hochwertige Leistungsdiagnostik auch im Zuge einer Gruppenvorgabe realisieren zu können. Dabei sollen konstruktäquivalente Untertests des AID 3 (Kubinger & Holocher-Ertl, 2013, in Druck) für eine Gruppenversion entwickelt werden. Durch die allgemeinen Rahmenbedingungen schriftlicher Gruppentestungen ist die Einteilung in verbal-akustische bzw. manuell-visuelle Untertests nicht haltbar. Allen Untertests liegt eine verbale Grundkomponente zugrunde, da ohne ein altersgerechtes Maß an Wortschatz und Leseverständnis die Bearbeitung der Testaufgaben unmöglich ist. Einem Großteil der Untertests der AID-Gruppenversion wird am ehesten der Bezeichnung *verbal-visuell* gerecht. Allein die Instruktionen der Testleiter/innen werden mündlich-akustisch gegeben.

Dieses Verfahren soll die individuellen Fähigkeiten *eindimensional* im Sinne eines *Power-Tests* erfassen. Dennoch wurde zugunsten der zeitlichen Planbarkeit der Testungen, eine Zeitbeschränkung je Untertest vorgesehen. Durch diese standardisierte Vorgabe wird die vorzugebende Zeitbeschränkung der Endversion einschätzbar. Eine Vorgabe gänzlich ohne zeitliche Begrenzung stellt sich für die Praxis als unangemessen dar, da durch die klassenweise Vorgabe im Schulkontext eine planbare Zeiteinteilung von Nöten ist. In jedem Fall soll das Bearbeiten der Aufgaben unter Zeitdruck vermieden werden.

Die Testungen der vorliegenden Arbeit stellen Vorerhebungen für die Aufgabenanalyse der bereits entwickelten Testaufgaben für fünf Untertests dar, nämlich *Alltagswissen*, *Realitätssicherheit*, *Antonyme Finden*, *Angewandtes Rechnen* und *Funktionen Abstrahieren*.

Breit angelegte Normierungstestungen der gesamten Intelligenz-Testbatterie sind für das Jahr 2013 geplant. In weiterer Folge der Testentwicklung soll zukünftig eine adaptive Computerversion entstehen.

2. Das Konstrukt *Realitätssicherheit*

Im Folgenden wird das Konstrukt *Realitätssicherheit*, das sowohl im AID-Gruppe (Kubinger & Hagenmüller, in Vorb.) als auch in der Intelligenz-Testbatterie AID 3 (Kubinger & Holocher-Ertl, 2013, in Druck) und dessen Vorgängerversionen erfasst wird, skizziert und definiert.

2.1 *Realitätssicherheit* der AID-Individualverfahren

„Der Untertest 2 *Realitätssicherheit* soll prüfen, inwieweit die Wirklichkeit um Dinge des Alltags verstanden bzw. kontrolliert werden kann.“ (Kubinger & Wurst, 2000, S. 16)

Den Testpersonen werden nacheinander Bilder gezeigt, auf denen jeweils ein grundlegendes, funktionelles Detail fehlt bzw. nicht dargestellt ist. Das Fehlen des Bildteils kann entweder verbal benannt oder durch Hinzeigen identifiziert werden. Der Untertest misst nachweislich *eindimensional*. Die Aufgaben werden konventionell vorgegeben (vgl. Kubinger & Wurst, 2000).

Ein explizit zu erwähnendes Charakteristikum dieses Tests betrifft seine Sprachfreiheit. Sowohl die Instruktion als auch das Vorgeben und Beantworten der Items können ohne sprachliche Ausführungen auskommen. Zur Bearbeitung der Aufgaben ist ein entsprechendes Maß visueller Differenzierungsfähigkeit von Nöten.

3. Theoretische Fundierung zur Testkonstruktion

Nach Bühner (2011) können die Methoden der Testkonstruktion in nachfolgender Weise eingeteilt und beschrieben werden:

3.1 Rationale Testkonstruktion

Die rationale oder deduktive Methode der Testkonstruktion bedient sich, im Idealfall, einer gut ausgearbeiteten Theorie. Je genauer das zu untersuchende Konstrukt definiert ist, desto einfacher wird es entsprechende Aufgaben zur Erfassung des Konstrukts abzuleiten (Bühner, 2011).

3.2 Externale Testkonstruktion

Die externe oder kriteriumsorientierte Testkonstruktion dient dem Zweck, zwischen verschiedenen Gruppen zu trennen. Unabhängig vom zu untersuchenden Konstrukt sollen Items entwickelt werden, die diese Trennung besonders effektiv ermöglichen. Eine typische Konstruktion stellt z.B. ein Fragebogen zur Selbsteinschätzung dar. Der Nachteil dieser Methode liegt darin, dass zur Interpretation der Gruppenzugehörigkeit auf die Beantwortung einzelner Items zurückgegriffen werden muss (Bühner, 2011).

3.3 Induktive Testkonstruktion

Bei dieser Methode wird zumeist ein sehr großer Itempool generiert, der nach Vorgabe einer Stichprobe mittels exploratorischer Faktorenanalyse ausgewertet wird. Diese Methode bedient sich keinem vorausgehenden theoretischen Modell, sondern konstruiert ihr Modell anhand der Analyse der Daten. Es werden Dimensionen des Konstrukts anhand der gegebenen Itemantworten gebildet.

Ein Beispiel dieser Konstruktionsmethode liefert der Persönlichkeitsfragebogen von Costa & McCrae (1985) mit der Bezeichnung *Big 5*³ (Bühner, 2011).

³ In diesem Modell wurden anhand dieser Methode 5 grundlegende Faktoren zur Beschreibung der Persönlichkeit entwickelt. Eine genauere Beschreibung dazu findet sich z.B. bei John, Angleitner & Ostendorf, 1988.

3.4 Prototypenansatz

Mit bestimmten Verhaltensweisen und Fähigkeiten gehen meist prototypische Vorstellungen von deren Ausprägungen einher. Um verschiedene Items zu entwickeln, werden die Vorstellungen verschiedener Personen gesammelt, woraus letztendlich entsprechende Aufgaben konstruiert werden. Dieser Ansatz besitzt deutlich mehr Relevanz für die Itemgenerierung von Persönlichkeitsfragebogen als bei Leistungstests (Bühner, 2011).

3.5 Vergleich der Methoden

Es sei darauf hingewiesen, dass eine Rangreihung der Methoden kaum zielführend ist. Der Schlüssel zur sinnvollen Itemkonstruktion liegt in der Kombination geeigneter Methoden. Beispielsweise erscheint eine induktive Methode im Anschluss an eine rationale als zielführend um ungeeignete Items aufspüren zu können. Auch eine Kombination mit dem Prototypenansatz für beide zuvor genannten Methoden kann sich als sinnvoll erweisen (Bühner, 2011).

4. Techniken für die Indikatorensuche des Konstrukts

Bühner (2011) führt nachfolgende Techniken und Ansätze für das Eingrenzen des zu bestimmenden Merkmals an:

1. Top-Down-Techniken:

- Erfahrungsgeleitet-intuitiver Ansatz
- Sammlung und Analyse von Definitionen/Literaturrecherche

2. Bottom-Up-Techniken:

- Analytisch-empirischer Ansatz
- Personenbezogen-empirischer Ansatz

Beim *erfahrungsgeleiteten-intuitiven Ansatz* erfolgt die Eingrenzung des Messziels durch das detaillierte Wissen von Experten. Die Indikatoren des Konstrukts sollen aufgrund von bestehendem Wissen ermittelt werden.

Das Ziel der *Sammlung und Analyse von Definitionen* liegt darin, wiederkehrende Kernbereiche des Konstrukts zu identifizieren. Dazu können sämtliche bestehende Literaturdatenbanken und psychologisch-diagnostische Verfahren genutzt werden.

Beim *analytisch-empirischen Ansatz* sollen die Elemente des Messgegenstands durch standardisierte Beobachtungs- und Befragungsinstrumente erfolgen. Dabei kommen strukturierte Interviews, Fragebogen und Verhaltensbeobachtungen zum Einsatz.

Bei der *personenbezogenen-empirischen* Methode werden bereits bestehende empirische Befunde als Grundlage für die Indikatorensuche verwendet. Diese Methode ist jedoch nur dann ratsam, wenn bereits verlässliche psychologisch-diagnostische Verfahren zur Messung des Konstrukts bestehen.

5. Regelgeleitete Itemkonstruktion

Da die Konstruktion von Items idealerweise theoriegeleitet erfolgen sollte, empfiehlt sich eine regelgeleitete Itemkonstruktion. Diese wird dadurch erreicht, dass zur Lösung der Items seitens der Testpersonen bestimmte kognitive Operationen angewendet werden müssen. Durch diese festgelegten Problemlösestrategien wird schon bei der Konstruktion der Items das Ziel der zu messenden Fähigkeit miteinbezogen. Durch diese Vorgehensweise können die Konstruktvalidität und die inhaltlichen Gültigkeit eines Verfahrens bzw. dessen Items gewährleistet werden (Kubinger, 2009b).

Durch klare Festlegung der anzuwendenden kognitiv operationalen Regeln pro Item, sollten die Schwierigkeiten der Items durch die Kombination dieser Regeln entstehen. Die Prüfung dessen erfolgt mittels *Lineal Logistischem Test-Modell (LLTM)*. Sollte sich dieses für die generierten Items als gültig erweisen, könnten zukünftige Testaufgaben theoretisch mit zuvor festgelegter Schwierigkeit entwickelt werden (vgl. Fischer, 1974).

5.1 Richtlinien für den sprachlichen Aufbau von Aufgaben

Nach Lienert, G. A. & Raatz, U. (1998) lassen sich folgende Regeln für den sprachlichen Aufbau von Items zusammenfassen:

- Es sollen Begriffe vermieden werden, die multiple Bedeutungen haben.
- Es sind Begriffe und Formulierungen zu vermeiden, die nur einem Teil der zu testenden Personen geläufig sind.
- Jedem Item sollte nur ein sachlicher Inhalt zugrunde gelegt werden.
- Doppelte Verneinungen sind zu vermeiden, hingegen sollten Anweisungen, Fragen und Aussagen positiv formuliert werden.
- Verallgemeinerungen sollten vermieden werden, wobei besonders von jenen, die nur unter bestimmten Bedingungen gelten, Abstand genommen werden muss.
- Das Item sollte weder zu lang noch zu kurz formuliert sein. So soll einerseits eine unnötige Verkomplizierung und andererseits eine missverständliche Knappheit der Aufgabenstellung vermieden werden.

6. Einzel- vs. Gruppentestung

Nachfolgend werden die unterschiedlichen Aspekte von Einzeltestungen und Gruppentestungen erläutert und einander gegenübergestellt.

6.1 Vor- und Nachteile einer Individualtestung

Mittels einer Einzelfalldiagnostik können „spezifische Besonderheiten oder Veränderungen eines Probanden (oder einer Untersuchungseinheit)“ analysiert werden (Jäger, 1988, S. 169).

Die Einzeltestung ermöglicht „ein situationsabhängiges Eingehen auf die Testperson“ (Herle, 2003, S. 205) und führt dazu, dass Verständnisprobleme während der Instruktion leichter bemerkt werden (Herle, 2003). Nach Reindl (1998) ergibt sich daraus der Vorteil, genau so viele Instruktionsbeispiele zu liefern, bis die Aufgabenart verstanden wird. Jedoch bringt dieser Vorteil auch den Nachteil mit sich, dass die Testleiterunabhängigkeit⁴ sinkt, da die Instruktion stärker vom standardisierten Ideal abweicht. Dadurch sinkt die Objektivität.

Einen weiteren Vorteil stellt die Möglichkeit der Vorgabe besonderer Testmaterialien dar, wie z. B. Puzzles. Im Zusammenhang mit den verwendeten Testmaterialien werden Testpersonen mit geringer Lesefähigkeit nicht benachteiligt (vgl. Kubinger, 2009b).

„Im Rahmen einer Individualtestung ist eine Verhaltensbeobachtung möglich; so können z.B. Strategien einer Person beim Lösen einer Aufgabe, sowie Konzentrations- und Motivationseffekte Mitberücksichtigung finden.“ (Herle, 2003, S. 205)

Nach Reindl (1998) ergibt sich bei Individualtestungen außerdem der Vorteil, dass sich Testpersonen während der Bearbeitung nicht gegenseitig stören können. Demnach steigt die Wahrscheinlichkeit für eine höhere Konzentrationsleistung. Außerdem sei zu erwähnen, dass im Gegensatz zu Gruppentestungen keine Möglichkeit besteht von anderen Personen

⁴ „Sie ist für einen Test gegeben, wenn das Testverhalten der Tp und der damit für sie resultierende Testwert unabhängig von zufälligen oder systematischen Verhaltensvariationen ist, wie sie durch die Person des TI bedingt sind.“ (Kubinger & Jäger, 2003, S. 196)

abzuschreiben. Für Personen, die den Wettbewerbsdruck und das „sich beobachtet Fühlen“ innerhalb einer Gruppensituation als unangenehm empfinden, eignet sich die Testsituation während eines Individualverfahrens besser.

6.2 Vor- und Nachteile einer Gruppentestung

Einige Nachteile der Gruppentestung ergeben sich aus den bereits erwähnten Vorteilen der Individualtestung. Wie bereits erwähnt, ist eine detaillierte Verhaltensbeobachtung nur während einer Einzeltestung möglich. Außerdem kann aufgrund der Durchführbarkeit kaum mit besonderen Materialien gearbeitet werden. Da in der Regel Gruppenverfahren schriftlich zu bearbeiten sind, kann eine Benachteiligung für Testpersonen mit gering ausgeprägtem Leseverständnis entstehen und Verständnisprobleme während der Instruktion eventuell unentdeckt bleiben. Da Testungen in der Gruppe meist mit exaktem Zeitmanagement einhergehen (z.B. innerhalb des Schulbetriebes), erscheint eine Zeitbegrenzung pro Untertest als unumgänglich. Außerdem wäre eine ungestörte Aufgabenbearbeitung innerhalb der Gruppe durch individuelle Vergaben der Bearbeitungszeit unmöglich. Dies führt letztendlich zu einer Vermengung zweier Eigenschaftsdimensionen, auch bekannt als *Speed-and-Power-Problematik* (vgl. Kubinger 2009b; Kubinger & Jäger, 2003).

Ein in der Literatur oft erwähnter Vorteil von Gruppentestungen stellt die Testleiterunabhängigkeit dar. Diese ist Bestandteil des Hauptgütekriteriums Objektivität⁵. Nach Reindl (1998) ist die Bezeichnung der „Unabhängigkeit“ in Frage zu stellen, jedoch kann zumindest davon ausgegangen werden, dass sich der/die Testleiter/in gegenüber allen Mitgliedern der Gruppe gleich verhält.

Ein weiterer Vorteil betrifft das Gütekriterium Ökonomie. Nach Kubinger (2009b) betrifft Ökonomie in erster Linie den Kostenaufwand einer Untersuchung. Eine weitere Komponente betrifft die Zeit. Selbsterklärend bedeutet die gleichzeitige Testung einer Vielzahl an Testpersonen, im Gegensatz zu einer Individualtestung mit derselben Testlänge, einen erheblichen wirtschaftlich-ökonomischen Gewinn.

⁵ Weitere Teilaspekte der Objektivität stellen Verrechnungssicherheit und Interpretationseindeutigkeit dar.

Der Arbeitsaufwand für die einzelnen Teilnehmer und Teilnehmerinnen bleibt jedoch gleich. Trotzdem dürfen ökonomische Gesichtspunkte, unabhängig von bestehenden Ressourcen, nicht als Hauptentscheidungsgrund für den gewählten Vorgabemodus dienen. Die Entscheidung zugunsten einer ökonomischeren Variante, darf nicht auf Kosten der Qualität der psychologischen Untersuchung gefällt werden (vgl. dazu Kubinger, 2009b oder Kubinger & Jäger, 2003).

Abschließend sei noch der Umstand erwähnt, dass sich manche Personen während einer Testsituation behaglicher innerhalb einer Gruppe fühlen, da sie ein gesteigertes Anonymitätsempfinden erleben. Für diese Menschen stellt der Gruppenmodus einen erheblichen Vorteil für ihre individuelle Leistungsoptimierung dar (vgl. Reindl, 1998).

7. Das dichotome logistische Testmodell von Rasch

Das dichotome logistische Testmodell von Georg Rasch (1960) ist ein einparametrisches (1-PL-Modell) *IRT-Modell*⁶, welches häufig - auch in weiterer Folge der vorliegenden Arbeit -, als „*Rasch-Modell*“ bezeichnet wird. Nach Kubinger (1989) gehen alle IRT-Modelle von jenem Grundgedanken aus, dass sich nicht direkt beobachtbare Fähigkeiten oder Verhaltensweisen einer Personenstichprobe in den Items bzw. deren Bearbeitung ausdrücken.

Leistungstests betreffend wird die Beziehung zwischen dem Lösen und Nichtlösen eines Items durch die Person und der zu messenden Eigenschaft als wahrscheinlichkeitsfunktionaler Zusammenhang beschrieben. Die nicht direkt beobachtbaren Fähigkeiten bzw. Eigenschaften drücken sich durch Personenparameter aus, die Schwierigkeiten der Aufgaben durch Itemparameter (vgl. Kubinger 1989). Die Wahrscheinlichkeit eine Testaufgabe zu lösen wird unter Ausnahme des Zufalls, demnach nur durch diese beiden Parameter bestimmt. Die Kodierung der bearbeiteten Items kann nur auf zwei Arten geschehen, entweder 0 (nicht gelöst) oder 1 (gelöst). Die Verrechnung der Anzahl an gelösten Aufgaben beschreibt die sogenannte „*erschöpfende Statistik*“ (Fischer, 1974, zit. n. Kubinger 1989).

Durch die nachfolgende Formel des *Rasch-Modells* wird die Wahrscheinlichkeit beschrieben, dass T_p v Item i löst („+“), in Abhängigkeit des Personenparameters ξ_v , das ist die (wahre) Fähigkeit von v , und des Itemparameters σ_i , das ist die (wahre) Schwierigkeit von i (Kubinger 2009b, S. 89):

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

Im Gegensatz zur klassischen Testtheorie ermöglicht das *Rasch-Modell* der probabilistischen Testtheorie eine Prüfung des Gütekriteriums *Skalierung*. Für jeden Test, der als Verrechnungsmodus die Anzahl der gelösten Aufgaben

⁶ Anmerkung der Autorin: IRT steht für „Item Response Theory“. Eine im deutschsprachigen Raum geläufige Bezeichnung ist die „probabilistische Testtheorie“, die der sogenannten „klassischen Testtheorie“ bei der Konstruktion von psychologischen Tests gegenüber steht.

verwendet, muss notwendigerweise das *Rasch*-Modell gelten, damit dieser Verrechnungsmodus fair ist („*Notwendigkeitsbeweis*“ des *Rasch*-Modells). Eine wichtige Voraussetzung für diesen „*Notwendigkeitsbeweis*“ verkörpert die sog. „*lokale stochastische Unabhängigkeit*“. Diese besagt, dass das Lösen einer Aufgabe unabhängig davon sein muss, welche Items bereits bearbeitet wurden oder noch bearbeitet werden. Verfahren, für die das *Rasch*-Modell nicht gilt, sich aber an diesem Verrechnungsmodus bedienen, sind nicht verrechnungsfair (Kubinger, 2009b).

Das Modell geht von *Eindimensionalität* aus, das bedeutet, dass allen Items im modellkonformen Itempool dieselbe zu messende Fähigkeit zugrunde liegt (vgl. Kubinger, 2003).

Eine weitere mathematisch beweisbare Besonderheit des *Rasch*-Modells äußert sich in der Möglichkeit „*spezifisch objektive*“ Vergleiche durchzuführen. Dies bedeutet, dass Unterschiede zwischen den Fähigkeiten der Testpersonen unabhängig von den jeweils bearbeiteten Aufgaben bestimmt werden können. Außerdem sind Vergleiche zwischen den Aufgaben unabhängig von den Stichproben, aus denen die Parameter gewonnen wurden (vgl. Kubinger 2009b).

Die herausragende Eigenschaft des *Rasch*-Modells besteht darin, „dass es sich im Gegensatz zu anderen testtheoretischen Modellen um ein tatsächlich prüfbares handelt.“ (Kubinger, 2009b, S. 90)

Für die Prüfung mittels grafischen Modelltests, wird die Stichprobe in zwei Teilstichproben zerlegt und diese miteinander verglichen. Dabei wird die Stichprobe nach dem internen Teilungskriterium *Score* und variablen externen Teilungskriterien gespaltet. Als äußere Teilungskriterien können z.B. Alter, Geschlecht, Muttersprache, Schulstufe etc. fungieren. Die gewonnenen Parameterschätzungen werden für jedes Item in einem rechtwinkligen Koordinatensystem abgetragen. Die Parameter sollten möglichst nahe an der durch den Ursprung gehenden 45°-Geraden liegen. Sofern Items deutlich von dieser Geraden abweichen, kann keine Stichprobenunabhängigkeit und somit auch keine Verrechnungsfairness angenommen werden (vgl. Kubinger, 2009b, 2003, 1989).

Die Prüfung der Eindimensionalität der Items ist durch eine Vielzahl statistischer Modelltests möglich. Zum Beispiel kann die Prüfung mittels (bedingtem) *Likelihood-Ratio-Test* nach Andersen (1973) erfolgen.

Zur Überprüfung der Konformität einzelner Items kann, wie in der vorliegenden Arbeit, der *z-Test* nach Fischer und Scheiblechner (1970) berechnet werden. Dieser bedient sich ebenso des Vergleichs von Teilstichproben und ermöglicht einzelne abweichende Items herauszufiltern (vgl. Glas & Verhelst, 1995, zit. n. Weber, 2011).

Nicht gültige Items werden während der Testkonstruktion bzw. der Kalibrierung eines Tests schrittweise aus dem Itempool ausgeschieden. Nach jedem Ausscheiden eines Items werden erneut Modelltests berechnet, bis kein Item mehr der Modellvoraussetzung der Stichprobenunabhängigkeit zuwider läuft. Der reduzierte und modellangepasste Datensatz muss anschließend anhand einer gänzlich neuen Stichprobe geprüft werden (Kubinger, 2003).

III. Empirischer Teil

8. Ziel der Untersuchung

Diese Untersuchung hatte zum Inhalt, im Gegensatz zu der bildhaften und sprachfreien Version des AID 3, eine verbal-schriftliche Version des Untertests 2 „*Realitätssicherheit*“ für die AID-Gruppenversion zu entwickeln. Dazu sollte ein möglichst großer Itempool generiert werden.

Die Entwicklung der Items sollte regelgeleitet erfolgen und verschiedene Schwierigkeitsstufen beinhalten. Die Methode der Konstruktion für die einzelnen Aufgaben wird in Abschnitt 9 näher erläutert.

Die 61 final ausgewählten Items sollten mittels *Rasch*-Modell-Analyse auf *Eindimensionalität* geprüft werden und etwaige Ausschlüsse unpassender Items vollzogen werden.

Im weiteren Arbeitsschritt war geplant, einer Teilstichprobe neben der AID-Gruppenversion auch das AID 3 im *balancierten Versuchsdesign*⁷ vorzugegeben. Die Vorgabe aus dem AID 3 umfasste jene Untertests, die auch in der Gruppenversion vorgegeben wurden (*Alltagswissen, Realitätssicherheit, Angewandtes Rechnen, Antonyme Finden, Funktionen Abstrahieren*). Die daraus gewonnen Daten sind dazu geeignet, die *konvergente* Validität zwischen den Untertests zu bestimmen.

⁷ Näheres zum verwendeten Testdesign findet sich in Abschnitt 9.7.

9. Konstruktion des Untertests *Realitätssicherheit*

In diesem Abschnitt werden die Teilschritte des Konstruktionsvorgangs des Untertests *Realitätssicherheit* für die Gruppenversion des AID dargestellt. Beschrieben werden die verwendeten Regeln zur Itemkonstruktion und die Umlegung der interessierenden Fähigkeit in verbales Material, das Antwortformat, der Itempool und die gewählten Testdesigns, sowohl für die Individual- als auch für die Gruppentestung.

Um die Urheberrechte zu schützen, können exemplarisch nur Instruktionsbeispiele angeführt werden, jedoch keine kompletten Wort- oder Itemlisten dargestellt werden. Der/Die geschätzte Leser/in sei um Verständnis gebeten.

9.1 Umlegung der zu messenden Fähigkeit in verbales Material

Realitätssicherheit soll auch im Gruppenverfahren gemäß der Definition in Punkt 2.1 erfasst werden. Durch die sprachlich-schriftliche Ausführung der Testaufgaben, werden zusätzlich vermutlich tiefgreifende Fähigkeiten der Sprachlogik erfasst (Wortschatz, Sprachverständnis, Grammatik etc.). Inwieweit die Konstruktvalidität in Bezug auf die *konvergente* Validität dieses Untertests zwischen der Einzelversion und der Gruppenversion gegeben ist, soll anhand dieser Arbeit geklärt werden.

Da in den bestehenden Versionen des AID die beschriebene Teilfähigkeit durch das Erkennen von fehlenden, semantisch funktionalen Bildteilen erfasst wird, lag die Überlegung nahe, diese Methode in ähnlicher Weise in verbales Material zu transformieren. Dazu sollten Sätze konstruiert werden, die sprachlich ihren Inhalt erst dann korrekt wiedergeben, sobald ein Wortteil eines zusammengesetzten Nomens ergänzt wird. In beiden Fällen, bildhaft und sprachlich, müsste die „Wirklichkeit der Dinge“ auf Vollständigkeit geprüft und der unvollständige Teilaspekt erkannt werden. Abgesehen von den Unterschieden, die sich durch die Verwendung von sprachlichem Material - im Gegensatz zu bildhaft-sprachfreien Testaufgaben – ergeben, genügt es in diesem konstruierten Untertest nicht, das Fehlen alleine zu erkennen. Das fehlende Element muss zusätzlich selbständig produziert und ergänzt werden.

Zur Veranschaulichung der unvollständigen Nomen wird nachfolgend ein Instruktionsbeispiel angeführt. Die Regeln für die Itemkonstruktion sowie genauere Erklärungen zu dem angeführten Beispiel, werden in Abschnitt 9.3 erläutert.

Abbildung 1: Instruktionsbeispiel 1 – *Realitätssicherheit* – AID-Gruppe

DIE SCHWESTER NIMMT IHM BLUT AB.

Lösung: KRANKENSCHWESTER

9.2 Geltungsbereiche des AID 3 und des AID-Gruppe

Der Geltungsbereich für das AID 3 liegt zwischen 6;0 und 15;11 Jahren. Das entspricht in der Regel einem Einsatzbereich von der 1. bis zur 9. Schulstufe. Die Gruppenversion des AID legt ihren Einsatzbereich etwas breiter an, nämlich beginnend mit der 3. Schulstufe bis hin zur 12. Schulstufe.

Aufgrund der unterschiedlichen Geltungsbereiche der beiden Testversionen konnten für die *balancierte* Testvorgabe der beiden Verfahren nur die Schulstufen 3 bis 9 herangezogen werden. Für die Testungen des AID-Gruppe wurden die 3.-11. Schulstufen akquiriert, da aufgrund laufender Vorbereitungen und Prüfungen, Testungen innerhalb der Maturaklassen organisatorisch nicht möglich waren.

9.3 Regeln für die Itemkonstruktion

Die Vorgehensweise während der Aufgabenkonstruktion erfolgte unter Berücksichtigung der Regeln für den sprachlichen Aufbau von Aufgaben nach Lienert & Raatz (1998). Nachfolgend sollen Ablauf und Methode der Wortfindung, die Gestaltung des Satzbaus, die Findung der Kategorien und die Einschätzung der Schwierigkeitsstufen näher beschrieben werden.

9.3.1 Wortfindung

Zu Beginn des Konstruktionsvorgangs stand eine Zusammenstellung geeigneter zusammengesetzter Nomen. Als geeignet wurden jene Begriffe eingeschätzt, die im Alltag häufig und gebräuchlich sind. Dies betraf insbesondere Nomen, deren

Vorkommen im Alltag eines Kindes bzw. eines/einer Schülers/Schülerin des festgelegten Einsatzbereichs zu vermuten sind. Für diesen Zweck wurden Wörterbücher (Duden, österreichisches Wörterbuch) hinsichtlich passender zusammengesetzter Nomen durchsucht. Ein zusammengesetztes Hauptwort sollte jedoch nicht nur inhaltlich passend erscheinen, sondern auch hinsichtlich der Möglichkeiten, die Wortteile getrennt in einen Satz verbauen zu können, ausgewählt werden. Aus der dadurch entstandenen Wortliste wurden jeweils entsprechende Sätze konstruiert.

Zur Veranschaulichung des entstehenden Problems während der Satzkonstruktion - mit ungeeigneten zusammengesetzten Nomen - wird das zusammengesetzte Hauptwort des Titels dieses Abschnitts exemplarisch verwendet.

Abbildung 2: Beispielsätze für unterschiedliche Artikel der Wortteile

Beispielwort: WORTFINDUNG
Beispielsatz 1:
DIE FINDUNG FÜR DEN UNTERTEST REALITÄTSSICHERHEIT ERFOLGTE ANHAND EINES WÖRTERBUCHS.
Beispielsatz 2:
DAS WORT FÜR DEN UNTERTEST REALITÄTSSICHERHEIT ERFOLGTE ANHAND EINES WÖRTERBUCHS.

Wie anhand der beiden Beispielsätze ersichtlich wird, können durch zusammengesetzte Nomen, deren Wortteile unterschiedliche Artikel tragen, nur eingeschränkt grammatikalisch richtige Sätze gebildet werden. Um das Lösungswort „Wortfindung“ zu erhalten, kann die Aufgabe nur in der Version des 1. Beispielsatzes konstruiert werden, da im 2. Fall der bestehende Artikel falsch wäre. Eine Verwendung ausschließlich solcher zusammengesetzter Nomen würde dazu führen, dass die Lösung immer den ersten Wortteil eines zusammengesetzten Hauptworts beträfe. Da dies zu einer wesentlichen Simplifizierung der Items führen würde, wurde versucht, möglichst wenige solcher

Nomen zu verwenden. Hingegen wurden Nomen präferiert, deren Wortteile denselben Artikel tragen (z.B. ZAHNERSATZ).

Die zusammengestellte Wortliste wurde deshalb bereits in diesem Arbeitsschritt deutlich reduziert.

9.3.2 Satzbau und Inhalt

Für die Konstruktion der Sätze wurden folgende Regeln festgelegt:

1. Ergänzung des Lösungswortes durch den ersten oder den zweiten Wortteil

Wie bereits im vorhergehenden Abschnitt erläutert, sollten die gesuchten Wortteile entweder den ersten oder den zweiten Wortteil betreffen. Diese Methode diene dazu, den gesuchten Wortteil des Zielwortes weniger offensichtlich zu gestalten. Abbildung 3 demonstriert das 2. Instruktionsbeispiel mit gesuchtem 2. Wortteil.

Abbildung 3: Instruktionsbeispiel 2 - *Realitätssicherheit* - AID-Gruppe

MIT DER LUFT KANN MAN AUF DEM WASSER LIEGEN.

Lösung: LUFTMATRATZE

2. Das Lösungswort kann an unterschiedlichen Stellen im Satz stehen

Neben den zwei Möglichkeiten der gesuchten Wortteile sollten die zusammengesetzten Nomen außerdem durch ihre Position innerhalb des Satzes variiert werden. Nachfolgend wird anhand des Beispiels in Abbildung 2 ein veränderter Satzbau demonstriert. Im Gegensatz zur Satzposition des gesuchten Wortes in Abbildung 2 (Stelle 2), befindet sich das Lösungswort (Abbildung 4) in der Satzmitte.

Abbildung 4: Beispielsatz einer Wortposition in der Mitte des Satzes

Beispielwort: WORTFINDUNG

Beispielsatz einer anderen Wortposition im Satz:

ANHAND EINES WÖRTERBUCHS KONNTE DIE FINDUNG FÜR DEN UNTERTEST
REALITÄTSSICHERHEIT ERFOLGEN.

3. Satzlänge:

Die Satzlänge sollte pro Item zwischen 4 und 12 Wörtern betragen.

4. Grammatikalische Korrektheit der Sätze:

Eine weitere Bedingung während der Konstruktion der Items war die grammatikalische Richtigkeit der Sätze. Diese sollte sowohl mit nur einem Wortteil als auch mit dem gesamten Lösungswort gegeben sein. Beispielsatz 2 (Abbildung 2) wäre demnach ungültig, da sich der Artikel „das“ nicht auf das Lösungswort „Wortfindung“ übertragen ließe.

5. Beobachtbarer oder abstrakter Inhalt:

Die zusammengesetzten Nomen sollten nicht nur aus beobachtbaren oder gegenständlichen Bereichen entnommen werden, sondern auch abstrakte Begriffe beinhalten (z.B. BANDBREITE).

6. Lösungsmöglichkeiten:

Die Lösung sollte sich weitestgehend auf eine gültige Möglichkeit beschränken, jedoch konnten für bestimmte Items mehrere Antwortmöglichkeiten in den Lösungsschlüssel aufgenommen werden.

9.3.3 Inhaltliche Kategorien

Die gesuchten zusammengesetzten Nomen konnten größtenteils thematisch zehn verschiedenen inhaltlichen Kategorien zugeordnet werden:

- Alltag (beobachtbar/abstrakt)
- Technik (beobachtbar)
- Natur (beobachtbar)
- Medizin (beobachtbar)
- Finanzen (beobachtbar)
- Verkehr und Sicherheit (beobachtbar/abstrakt)
- Sport (beobachtbar)
- Freizeit (beobachtbar)
- Arbeit (beobachtbar)
- Nahrung (beobachtbar)

9.3.4 Einschätzung der Schwierigkeitsstufen

Die Schwierigkeitsstufen („einfach“, „einfach-mittel“, „mittel“, „mittel-schwierig“, „schwierig“) der einzelnen Items wurden auf Grundlage der unten angeführten Kriterien *a priori* eingeschätzt.

1. Satzlänge:

Je mehr Nomen sich in dem Aufgabensatz befinden, desto schwieriger wird es, das zu bearbeitende Wort herauszufiltern. Nach diesem Kriterium steht eine Satzlänge mit 4 Wörtern für ein besonders leichtes Item, eine Satzlänge mit 12 Wörtern für ein schwieriges.

2. Anzahl der Lösungen:

Items, in denen nur eine Lösung als richtig zu verrechnen ist, gelten tendenziell als schwieriger, im Gegensatz zu Aufgaben, in denen mehrere Lösungen gültig sind. Die Anzahl der Lösungen variierte zwischen 1 und 4, wobei am häufigsten nur eine richtige Antwort gesucht war.

3. Eindeutigkeit des Fehlens eines Wortteils:

Um zu verdeutlichen, wie unterschiedlich offensichtlich das Fehlen eines Wortteils ausfallen kann, seien erneut die beiden Instruktionsbeispiele herangezogen (vgl. dazu die Abbildungen 1 und 3).

Im ersten Beispiel (Abbildung 1) ist der fehlende Wortteil weniger ersichtlich als im zweiten Instruktionsbeispiel (Abbildung 3). Dies liegt zum einen darin, dass umgangssprachlich eine Krankenschwester oft als „Schwester“ bezeichnet wird und zum anderen daran, dass auch eine leibliche Schwester, sofern sie über eine geeignete medizinische Qualifikation verfügt, Blut abnehmen könnte.

Im zweiten Beispiel erscheint das Fehlen eines Wortteils deutlicher, da er ohne diesen Wortteil zwar grammatikalisch richtig, jedoch inhaltlich völlig absurd anmutet.

4. Die Geläufigkeit der Begriffe

Es ist davon auszugehen, dass die ausgewählten zusammengesetzten Nomen umgangssprachlich in unterschiedlicher Weise gebräuchlich sind.

Manche Begriffe finden im Alltag häufiger Verwendung als andere. Wie häufig Begriffe gehört, geschrieben oder gesprochen werden, kann mit den gebildeten Kategorien der Items in Verbindung gebracht werden. Es ist anzunehmen, dass Kategorien wie „Alltag“ und „Freizeit“ geläufigere Begriffe hervorbringen als z.B. „Technik“ und „Natur“.

5. Beobachten und Erlernen:

Die Geläufigkeit eines Begriffs wird auch dadurch bestimmt, wie wahrscheinlich es ist, den damit beschriebenen Sachverhalt regelmäßig zu beobachten (z.B. am Schulweg). Außerdem beeinflusst auch die Wahrscheinlichkeit, das gesuchte Vokabel standardmäßig zu erlernen (z.B. in der Schule), den Bekanntheitsgrad des Begriffs. Ersteres kann *a priori* eher eingeschätzt werden als Letzteres.

Unter Berücksichtigung dieser Kriterien und der erfahrungsgeleiteten Expertenmeinung von Fr. Mag. Bettina Hagenmüller wurden die entwickelten Items *a priori* in 5 Schwierigkeitsstufen eingeteilt.

9.4 Antwortformat

Die Verwendung des freien Antwortformats bei Leistungstests birgt verschiedene Vor- und Nachteile. Die Entscheidung, welches Antwortformat für den jeweiligen Test angemessen ist, erfolgt anhand unterschiedlicher Kriterien. Nachfolgend werden sowohl die Beschaffenheit als auch die Entscheidungsgrundlage für das hier verwendete Format skizziert. Auf eine umfassende Beschreibung der verschiedenen Antwortformate wird jedoch verzichtet. Für nähere Beschreibungen sei der/die interessierte Leser/in z.B. auf Lienert & Raatz (1998) oder Rost (2004) verwiesen.

Der Vorteil des freien Antwortformats für den Untertest *Realitätssicherheit* liegt zum einen darin, dass die Möglichkeit für die Testpersonen die Lösung zu erraten minimiert wird, zum anderen wird dadurch der zukünftige Geltungsbereich des Tests berücksichtigt. Für Kinder kann das freie Antwortformat leichter zu bewältigen sein als andere Antwortformate (vgl. Rost, 2004).

Durch die Art der Konstruktion der Aufgaben ergibt sich eine spezielle Version des freien Antwortformats, das - ohne Antwortmöglichkeiten vorzugeben -,

trotzdem eine „gebundene Komponente“ beinhaltet. Die gesuchten Antworten erfahren eine Einschränkung durch die in der Aufgabe enthaltenen Wörter. Da die Lösung einen fehlenden Wortteil aus dem Satz in der Aufgabenstellung darstellt, kann das Format als atypische Ergänzungsaufgabe angesehen werden. Je nach Satzlänge der Fragestellung ergeben sich unterschiedliche Schwierigkeitsstufen⁸ der Aufgaben. Da die exakten Lösungen vorab klar definiert wurden, kann hohe Verrechnungssicherheit angenommen werden.

In Abbildung 5 ist das zusammengesetzte Lösungswort exemplarisch bereits eingetragen. In dieser Weise sollten die Lösungswörter von den Testpersonen innerhalb der Kästchen am Antwortbogen handschriftlich eingetragen werden.

Abbildung 5: Antwortfelder am Antwortbogen für den Untertest *Realitätssicherheit*

Untertest 2 - Realitätssicherheit

Beispiel 1 Krankenschwester

Beispiel 2

9.5 Itempool und Itemeigenschaften

Unter Berücksichtigung der festgelegten Regeln für die Itemkonstruktion wurden zahlreiche Items entwickelt. Es wurde darauf geachtet, unterschiedliche Schwierigkeitsgrade zu generieren. Außerdem sollte die Position des Wortes innerhalb des Satzbaus variiert und das Fehlen des 1. bzw. 2. Wortteils gleichmäßig verteilt werden. Aus einer Vielzahl generierter Items wurden für die erstmalige Erprobung der Aufgaben letztendlich 63 Items ausgewählt, wovon zwei als Instruktionsbeispiele dienten. Für die späteren *Rasch*-Modell-Analysen

⁸ Nähere Erläuterungen finden sich in Abschnitt 9.3.2.

wurden die beiden Instruktionsbeispiele selbstverständlich ausgeschlossen. Nachfolgend findet sich eine Auflistung der Häufigkeiten der spezifischen Aufgabeneigenschaften.

Tabelle 3: Häufigkeiten der Charakteristika für 63 Items

Itemcharakteristika für 63 Items	Anzahl
<i>Schwierigkeit „einfach“</i>	13
<i>Schwierigkeit „einfach-mittel“</i>	4
<i>Schwierigkeit „mittel“</i>	16
<i>Schwierigkeit „mittel-schwierig“</i>	10
<i>Schwierigkeit „schwierig“</i>	20
<i>Selber Artikel der Wortteile</i>	40
<i>Unterschiedlicher Artikel der Wortteile</i>	23
<i>1. Wortteil gesucht</i>	31
<i>2. Wortteil gesucht</i>	32
<i>Position des Wortes am Satzanfang</i>	33
<i>Position des Wortes in der Satzmitte</i>	12
<i>Position des Wortes am Satzende</i>	18

9.6 Testdesign und Testformen (AID-Gruppe)

Für die Vorgabe der AID-Gruppenversion wurden acht verschiedene Aufgabenhefte zu je zwei parallelen Testformen zusammengestellt. Die jeweils parallelen Testformen beinhalteten dieselbe Aufgabenzusammenstellung, jedoch in unterschiedlicher Itemreihenfolge. Diese Vorgehensweise sollte es den Testpersonen erschweren, voneinander abzuschreiben.

Die Testformen wurden anhand ihrer *a priori* festgelegten Schwierigkeitsstufen zusammengestellt und für die Vorgabe innerhalb bestimmter Schulstufen vorgesehen. Pro Testheft wurden 20 Items des Untertests *Realitätssicherheit* vorgegeben.

Wie in Tabelle 4 ersichtlich, wurden für die jeweiligen Schulstufen mehrere Testhefte konstruiert. Verschiedene Zahlen kennzeichnen unterschiedliche Itemzusammenstellungen, die Bezeichnungen „a“ und „b“ markieren Paralleltestformen mit derselben Aufgabenzusammenstellung, jedoch in unterschiedlicher Reihenfolge. Die Abfolge der Testhefte in Tabelle 4 demonstriert die Vorgabe pro Sitzreihe in einem Klassenraum. Durch diese Vorgehensweise konnte das Risiko des Abschreibens auf ein Minimum reduziert werden.

Tabelle 4: Vorgabe der Testhefte je Schulform und Schulstufe

Schulstufe	Abfolge der Testhefte
3. VS-4.VS	1a - 1b - 2a - 2b - 1a - 1b - etc.
1. AHS-2.AHS 1.HS-2.HS	3a - 3b - 4a - 4b - 3a - 3b - etc.
3.AHS-5.AHS 3.HS-4.HS 1.BHS	5a - 5b - 6a - 6b - 5a - 5b - etc.
6.AHS-7.AHS 2.BHS-4.BHS	7a - 7b - 8a - 8b - 7a - 7b - etc.

Als Testdesign⁹ wurde ein verlinktes „*Booklet-Design*“ verwendet. Dieses Testdesign sieht vor, vereinzelte Items (sogenannte „*Linking-Items*“) in 2 bis 16 der bestehenden Testformen einzubauen. Durch diese Vorgehensweise werden die einzelnen Testformen durch wenige Items miteinander verknüpft, wodurch die Itemparameter im Zuge der *Rasch-Modell-Analysen* auf Basis der Gesamtstichprobe geschätzt werden können. Das ist durch diese Methode trotz

⁹ Das Schema zum Testdesign ist in Abschnitt V zu finden.

der Tatsache möglich, dass nicht alle Testpersonen alle Items bearbeitet haben (vgl. Wright & Stone, 1979).

9.7 Balancierter Versuchsplan

Abgesehen von der Erhebung einer Kalibrierungsstichprobe des Untertests *Realitätssicherheit* der Gruppenversion, sollte dieser auch anhand des gleichnamigen Untertests aus dem AID 3 validiert werden. Um die beiden Testarten direkt miteinander vergleichen zu können, wurde ein Teil der Stichprobe sowohl mit dem Individualverfahren als auch mit dem Gruppenverfahren getestet. Da die Items der beiden Testmodi nicht ident waren, musste zwischen den beiden Testungen pro Testperson kein zeitlicher Abstand eingehalten werden, weil die Wahrscheinlichkeit von Lerneffekten vernachlässigt werden konnte. Wie in Tabelle 5 ersichtlich, wurde pro Schulstufe jeweils die Hälfte der Teilstichprobe vor bzw. nach der Gruppenversion mit den entsprechenden Untertests des AID 3 getestet. Für Untersuchungen mit der Darbietung von identem Material durch unterschiedliche Vorgabemodi des AID 3, sei auf die beiden Diplomarbeiten von Scheffl (2011) und Goldmann (2011) verwiesen.

In Tabelle 5 wird der ursprüngliche *balancierte* Versuchsplan der Zielstichprobe für die zweifache Testung (AID 3, AID-Gruppe) skizziert.

Tabelle 5: Balancierter Versuchsplan für NÖ und OÖ gesamt

Schulstufe	Einzeltestung vor Gruppentestung – Anzahl der Testungen	Einzeltestung nach Gruppentestung – Anzahl der Testungen
3. VS	18	17
4. VS	18	17
1.AHS od. 1.HS	18	17
2.AHS od. 2.HS	18	17
3.AHS od. HS	12	12
4.AHS od. HS	12	12
5. AHS	12	12
Σ	108	104

Jene Testpersonen, die beide Testversionen bearbeiten sollten, wurden zufällig aus der akquirierten Gesamtstichprobe pro Bundesland gezogen. Die schlussendlich realisierte Stichprobengröße und deren Beschreibung sind in Abschnitt 11 skizziert.

9.8 Instruktion und Testvorgabe des AID-Gruppe

Für die Vorgabe des AID-Gruppe wurden den Testpersonen jeweils ein Aufgaben- und ein Testheft vorgelegt. Durch diese Vorgehensweise konnten die Produktionskosten des Testmaterials erheblich gesenkt werden, da die Aufgabenhefte mehrfach zum Einsatz kamen.

Die Untertests wurden konventionell vorgegeben. Jede/r Schüler/in musste 20-24 Items, entsprechend ihrer/seiner Altersstufe, bearbeiten. Die maximale Bearbeitungsdauer betrug zehn Minuten. Diese Dauer war für jeden der vorgegebenen Untertests gleich und bedeutete insbesondere für den Untertest *Realitätssicherheit* (jedenfalls für die Mehrheit der Testpersonen) einen Zeitüberschuss. Es kann daher angenommen werden, dass keine Vermengung der *Speed*- und *Power*-Komponente entstanden ist (siehe Abschnitt 6.2). Grundsätzlich war eine Zeitbegrenzung von Nöten, um willkürlichen Abbrüchen durch die Testleiter/innen entgegenzuwirken und somit alle Schüler/innen unter denselben zeitlichen Rahmenbedingungen zu testen.

Der/die Testleiter/in war angewiesen, nach einer exakt standardisierten Vorgehensweise zu instruieren. Für diesen Zweck gab es schriftliche Testleiteranweisungen, mit der genauen Abfolge der Vorgehensweise während der Testadministration.

Abbildung 6 zeigt einen Auszug aus der Testleiteranweisung für den Untertest *Realitätssicherheit*.

Abbildung 6: Testleiteranweisung – Realitätssicherheit des AID-Gruppe

Testleiter-Anweisung-Gruppentestung	
Bitte blättert nun auf <u>Seite 13</u> des Testhefts und auf <u>Seite 3</u> des Antwortbogens. Wir lesen nun wieder gemeinsam die Anleitung. Der/die Testleiter(in) liest folgendes vor, die Schüler(innen) können im Testheft mitlesen.	
Auf den nächsten Seiten des Testhefts findest du einige Sätze, die keinen Sinn ergeben, da der erste oder der zweite Teil eines zusammengesetzten Wortes fehlt. Ein zusammengesetztes Wort ist zum Beispiel Blumenstrauß. Dieses setzt sich aus den beiden Wörtern Blumen (erster Teil) und Strauß (zweiter Teil) zusammen. Beispiel 1: DIE SCHWESTER NIMMT IHM BLUT AB. ? Damit der Satz einen Sinn ergibt, musst du ein Wort durch ein zusammengesetztes Wort ersetzen. In Beispiel 1 ist KRANKENSCHWESTER die Lösung. Das Wort SCHWESTER wurde durch das zusammengesetzte Wort KRANKENSCHWESTER ersetzt. Nun hat der Satz einen Sinn! Daher ist in das Kästchen auf dem Antwortbogen unter Beispiel 1 bereits das Wort KRANKENSCHWESTER eingetragen. Bitte trage bei allen Aufgaben die Lösung in das Kästchen auf dem Antwortbogen ein. Achte dabei auf die Nummerierung der Aufgaben und darauf, nicht das Kästchen zu vertauschen! Beispiel 2: MIT DER LUFT KANN MAN AUF DEM WASSER LIEGEN. ? Versuche das Beispiel 2 zu lösen, indem du auf dem Antwortbogen in das Kästchen für Beispiel 2 die Lösung einträgst. Solltest du dir nicht sicher sein, wie man ein Wort richtig schreibt, schreibe es einfach so wie du denkst, dass es gehört. Es macht nichts, wenn du Rechtschreibfehler machst, diese werden nicht beurteilt!	

10. Datenerhebung

Nachfolgend werden die Rahmenbedingungen für die Datenerhebung erläutert. Darunter fallen die Akquirierung der Stichprobe, der Durchführungszeitraum und die Schulung für die Testleiter und Testleiterinnen.

10.1 Akquirierung der Stichprobe

Die beiden Intelligenz-Testbatterien sollten anhand einer repräsentativen Stichprobe im Erhebungszeitraum Mai/Juni 2012 erprobt werden. Dazu wurden Ansuchen an die Landesschulräte von Ober- und Niederösterreich, ausgehend von der Projektkoordinatorin Mag. Bettina Hagenmüller, gestellt. Da die Bewilligungen der Landesschulräte für die beiden Bundesländer erst kurz vor Beginn des geplanten Erhebungszeitraums zu erwarten waren, mussten vorläufige Zusagen der jeweiligen Schulen bereits früher eingeholt werden. Zu diesem Zweck wurden Anfragen bei verschiedenen Schulen, seitens der Autorin (NÖ) und Fr. Lorenz (OÖ), bereits ab April 2012 gestellt. Um die Anfragen möglichst effektiv und informativ zu handhaben, wurden Informationsschreiben zu der geplanten Untersuchung, breitgefächert an die verschiedenen Schultypen, per E-Mail versandt. Wenige Tage danach wurden die angeschriebenen Schulen telefonisch kontaktiert, um vorläufige Zusagen zu erhalten. Diese Vorgehensweise hat sich als zielführend erwiesen, da Schulleitung und Administration aufgrund der vorab versendeten Informationsschreiben bereits über die Durchführbarkeit und Planung der Testungen reflektieren konnten. Die Termine der Testungen wurden bereits während dieser Planungsphase grob fixiert. Nach Erhalt der Bewilligung durch die Landesschulräte (NÖ: 22.5.2012), wurde das Projekt den Direktoren und Direktorinnen nochmals persönlich vorgestellt. Im Zuge dessen wurden Projektbeschreibungen für die Lehrerschaft und Eltern sowie Einverständniserklärungen in den teilnehmenden Schulklassen verteilt. Um den Willen für eine rege Teilnahme zu steigern, wurden den Testpersonen individuelle Ergebnismeldungen zugesichert. In Oberösterreich konnten zwei Volksschulen, zwei Hauptschulen, eine AHS, eine höhere technische Lehranstalt (HTL) sowie eine polytechnische Schule zur Teilnahme bewegt werden. In Niederösterreich nahmen drei Volksschulen, eine AHS und ein Bundesoberstufenrealgymnasium (BORG) an den Untersuchungen teil. Nach

dem Erhalt des Rücklaufs der Teilnehmer und Teilnehmerinnen konnten genaue Testpläne zusammen mit den zuständigen Administratoren und Administratorinnen der jeweiligen Schulen erstellt werden. Sämtliche Materialien (Bewilligung des Landesschulrats, Eltern- und Lehrerbriefe, exemplarische Ergebnissrückmeldung etc.) zum Thema Organisation befinden sich in Anhang V.

10.2 Durchführungszeitraum und Testdauer

Der tatsächliche Durchführungszeitraum der Testungen in Niederösterreich erstreckte sich von 4.6.2012 bis 25.6.2012. Da sich die Terminplanung nach den organisatorischen Möglichkeiten (Zeitplanung, verfügbare Räumlichkeiten etc.) der Schulen richten musste, ergaben sich zeitgleiche Testungen an mehreren Lehranstalten. Die Testdauer pro Einzeltestung belief sich auf ca. 45 Minuten. Die Gruppentestungen pro Klasse konnten innerhalb von 80 Minuten administriert werden. Im Anschluss an die Testungen erhielten die Schüler/innen ein Schreiben ausgehändigt, in dem die Modalitäten zur Ergebnissrückmeldung angeführt waren (siehe Abschnitt V.). Durch den enormen zeitlichen Aufwand und dem vergleichsweise kurzen Durchführungszeitraum der Testungen, war der Einsatz einer größeren Anzahl geschulter Testleiter und Testleiterinnen von Nöten.

10.3 Testleiter/innen

Die Testungen der AID-Gruppenversion erfolgten je nach Rücklauf entweder klassenweise oder in zusammengelegten Gruppen mehrerer Klassen innerhalb einer Schulstufe. Für die ungestörte Durchführung der Individualtestungen wurden von den Schulen leer stehende Räume zur Verfügung gestellt.

Um zu gewährleisten, dass die Instruktionen je Untertest von allen teilnehmenden Personen verstanden wurden, wurde im Rahmen dieser vorliegenden Untersuchung, explizit innerhalb der Gruppe bzgl. des Verständnisses nachgefragt. Fehlinterpretationen die Bearbeitung der Aufgaben betreffend konnten somit weitestgehend vermieden werden. Um Störvariablen wie Lärm, Gespräche, Gelächter usw. möglichst klein zu halten, wurden die Testpersonen zum stillen Bearbeiten ermahnt und Interaktionen unterbunden. Die Gefahr des Abschreibens wurde durch die abwechselnde Vorgabe verschiedener Testformen auf ein Minimum reduziert (vgl. Amelang & Schmid-Atzert, 2006).

Für die Testungen der vorliegenden Arbeit wurden eine Vielzahl an Testleitern und Testleiterinnen eingesetzt, wobei eine standardisierte Vorgabe eine hohe Durchführungsobjektivität gewährleistete.

Als Testleiter/innen fungierten Diplomanden und Diplomandinnen des Arbeitsbereichs Psychologische Diagnostik. Diese konnten sich durch eine zweistündige Einschulung durch Projektkoordinatorin Mag. Bettina Hagenmüller für die Vorgabe der AID-Gruppenversion qualifizieren. Für die Individualtestungen des AID 3 wurden vorzugsweise Testleiter/innen gesucht, die bereits Testerfahrung im Rahmen des AID-Zertifizierungskurses gesammelt hatten. Der entstehende Arbeitsaufwand konnte entweder für den Praxisteil des AID-Zertifizierungskurses oder für die zu erbringenden Hilfestunden im Rahmen des „Forschungsseminars für Fortgeschrittene“ abgegolten werden. In Niederösterreich waren neben der Verfasserin selbst insgesamt 13 Testleiter/innen an den Erhebungen beteiligt. Die hauptverantwortliche Testleitung und Organisation erfolgte in Oberösterreich durch Fr. Lorenz und in Niederösterreich durch die Autorin dieser Arbeit.

11. Stichprobenbeschreibung

Insgesamt wurden 810 Testpersonen im Alter von 8 bis 19 Jahren mit dem AID-Gruppenverfahren getestet. Davon wurden 209 Testpersonen von der 3. bis zur 9. Schulstufe zusätzlich das AID 3 vorgegeben.

Nach Durchzählung der Rückläufe erklärten sich ursprünglich 836 Schüler/innen für die Untersuchung bereit. Während der Testungen fielen 26 Testpersonen aus, da sie z.B. aufgrund von Krankheit nicht anwesend waren. Während der Individualtestungen gab es zwei Testabbrüche aufgrund mangelnder Motivation seitens der Testpersonen. Die beiden Testungen wurden durch Zufallsbestimmungen zweier anderer Teilnehmer/innen ersetzt.

Die Gesamtstichprobe ($n = 810$) setzte sich aus 46 % Teilnehmern und 54 % Teilnehmerinnen im Alter von 8 bis 19 Jahren zusammen ($Md_{Alter} = 12$ Jahre). Die Daten wurden in Niederösterreich und in Oberösterreich gewonnen.

Im Folgenden werden die Verteilungen der Testpersonen der Gesamtstichprobe auf die Bundesländer (siehe Tabelle 6), auf die verschiedenen Schultypen (siehe Tabelle 7) und auf das Alter in Jahren (siehe Tabelle 8) dargestellt. In Tabelle 9 befindet sich außerdem der tatsächlich realisierte *balancierte* Versuchsplan für die Individualtestungen pro Schulstufe.

Tabelle 6: Die Verteilung der Testpersonen auf die Bundesländer Niederösterreich und Oberösterreich ($n = 810$).

Bundesland	Absolute Häufigkeit	%
NÖ	432	53,3
OÖ	378	46,6

Tabelle 7: Die Verteilung der Testpersonen auf die verschiedenen Schultypen ($n = 810$).

Schultyp	Absolute Häufigkeit	%
Volksschule	213	26,3
Hauptschule	91	11,2
AHS	432	53,3
Polytechnische L.	16	2
HTL	58	7,2

Tabelle 8: Die Verteilung der Testpersonen auf das Alter in Jahren ($n = 810$).

Alter in Jahren	Absolute Häufigkeit	%
8	16	2
9	83	10,2
10	103	12,7
11	102	12,6
12	119	14,7
13	83	10,2
14	81	10
15	72	8,9
16	81	10
17	67	8,3
18	2	0,2
19	1	0,1

Tabelle 9: Der balancierte Versuchsplan des AID 3 mit den Verteilungen auf die Schulstufen ($n = 209$).

Schulstufe	Absolute Häufigkeit vorher	Absolute Häufigkeit nachher
3	17	16
4	17	17
5	18	19
6	18	18
7	12	11
8	12	11
9	11	12

Für den Vergleich der geplanten Zielverteilung im Vorfeld sei auf Tabelle 5 in Abschnitt 9.7 verwiesen.

12. Ergebnisdarstellung

Die Durchführung der *Rasch*-Modell-Analysen erfolgte unter Verwendung des Programmpakets *eRm* (Mair, Hatzinger & Maier, 2010) der Statistik- Software R. Sämtliche anderen Auswertungen erfolgten mit Hilfe der Statistik-Software *PASW Statistics 18.0*.

12.1 *Rasch*-Modell-Analysen des Itempools

Die Prüfung der Modellgültigkeit der Items erfolgte unter Anwendung des *Likelihood-Ratio-Tests* nach Andersen (1973). Zur Veranschaulichung wurde der grafische Modelltest verwendet. Zur Überprüfung der Konformität einzelner Items bei signifikantem *LR*-Test wurde zusätzlich der *z-Test* nach Fischer und Scheiblechner (1970) berechnet (vgl. Abschnitt 7).

Das vergleichsbezogene Risiko 1. Art¹⁰ wurde mit $\alpha = 0,05$ festgelegt. Eine Adjustierung der einzelnen *Likelihood-Ratio-Tests* (LRT) erfolgte untersuchungsbezogen auf $\alpha = 0,01$, wodurch die Kumulierung des Risikos 1. Art bei der Anwendung von 3 Modelltests ein α von 0,029 ergibt. Nach Kubinger und Draxler (2007) sind die Modelltests so festzulegen, dass die insgesamt erzielte Wahrscheinlichkeit für einen Fehler 1. Art zwischen einem α von 0,01 und 0,05 liegt. Im Folgenden befindet sich die Einteilung der Teilungskriterien.

Internes Teilungskriterium:

- **Score** (Rohwert $\leq$ Median vs. Rohwert $>$ Median)

Externe Teilungskriterien:

- **Geschlecht** (männlich vs. weiblich)
- **Schulstufe** (Volksschule vs. höhere Schulen)

Nicht-Modell-konforme Items sollten schrittweise aus dem Itempool eliminiert werden, bis die Modellgültigkeit *a posteriori* erreicht wird. Üblicherweise wird Modellgültigkeit für mindestens zwei Drittel der Items angestrebt. Sollte dies nicht

¹⁰ Dies bezeichnet die Irrtumswahrscheinlichkeit, dass die H_0 : "Das *Rasch*-Modell gilt." fälschlicherweise verworfen wird.

der Fall sein, wird die zu untersuchende Fähigkeit anhand des Aufgabenuniversums nicht *eindimensional* erfasst.

Im grafischen Modelltest gelten jene Items als auffällig, deren Abweichungen „größer als 1/10 der Spannweite der Schwierigkeitsparameter im betroffenen Itempool“ sind (Kubinger & Draxler, 2007, S. 141).

Zusätzlich zu den bereits erwähnten Prüfverfahren (LRT, *z-Test*), sollte für den Ausschluss von einzelnen Items das Augenmerk auch auf inhaltliche Aspekte gelegt werden. Die inhaltlichen Ausschlussgründe der Items können aufgrund der Urheberrechte nicht im Detail angeführt werden. Der/die Leser/in muss sich mit dem allgemeinen Hinweis auf inhaltliche Gründe begnügen.

Die Ergebnisse der ersten *Rasch*-Modell-Analyse für 61 Items sind in Tabelle 10 aufgelistet. In allen drei Teilungskriterien kam es zu signifikanten Modellabweichungen. Für das Kriterium „Score“ waren vier Items nicht schätzbar (Anzahl schätzbarer Items = $df+1$). Dies ist dadurch begründet, dass in den jeweiligen Teilstichproben betreffende Items entweder von jeder oder von keiner Testperson gelöst wurden. Beim Teilungskriterium Volksschule vs. höhere Schulen (VS-HS) konnten nur jene Items geschätzt werden, die *Linking-Items* darstellten. Aufgrund der großen Anzahl an unterschiedlichen Testaufgaben für die Vorgabe in Volksschulen und höheren Schulen, ergibt sich eine deutlich reduzierte Anzahl schätzbarer Items.

Tabelle 10: Die Ergebnisse der LRTs vor dem Ausschluss eines Items.

Teilungskriterium	χ^2 (LRT)	df	$\chi^2_{\text{krit}} (\alpha=0.01)$	p
Score	111,771	56	83,51343	0,000
Geschlecht	134,075	60	88,37942	0,000
VS-HS	84,117	28	48,27824	0,000

Nachfolgend werden die grafischen Modelltests der gewählten Teilungskriterien vor dem Ausschluss eines Items wiedergegeben.

Abbildung 7: Der grafische Modelltest des Teilungskriteriums *Score* vor dem Ausschluss eines Items.

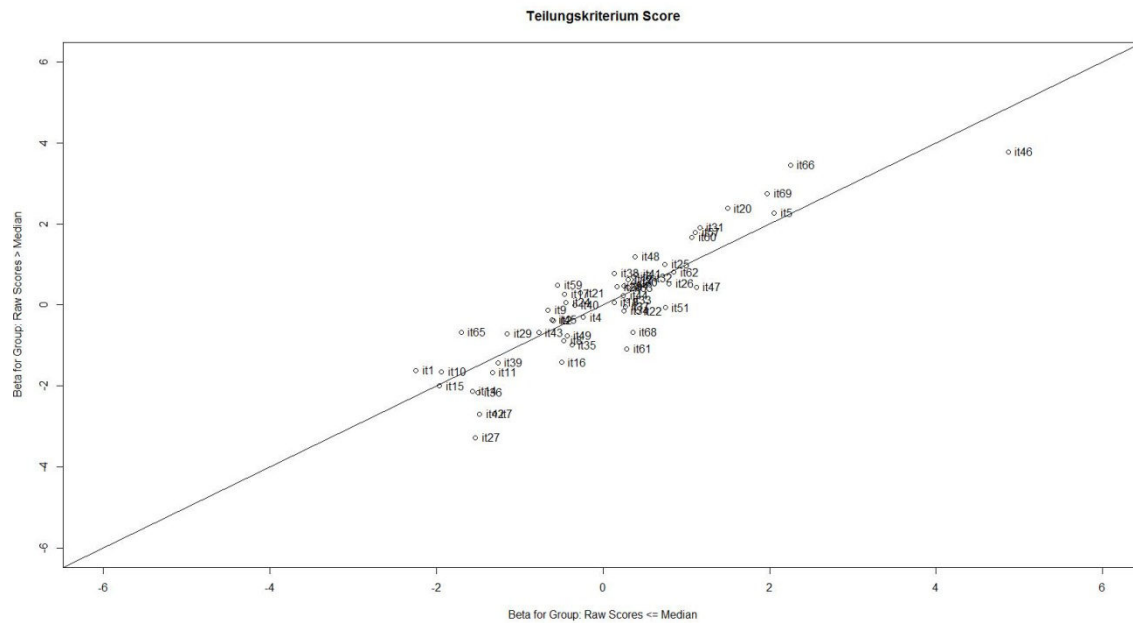


Abbildung 8: Der grafische Modelltest des Teilungskriteriums *Geschlecht* vor dem Ausschluss eines Items.

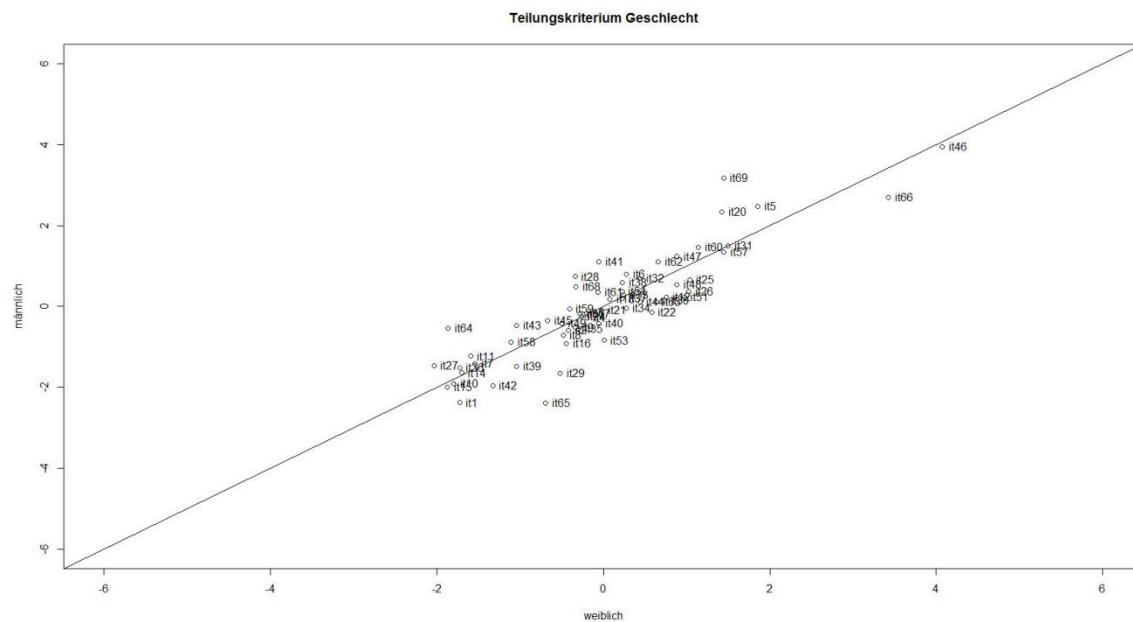
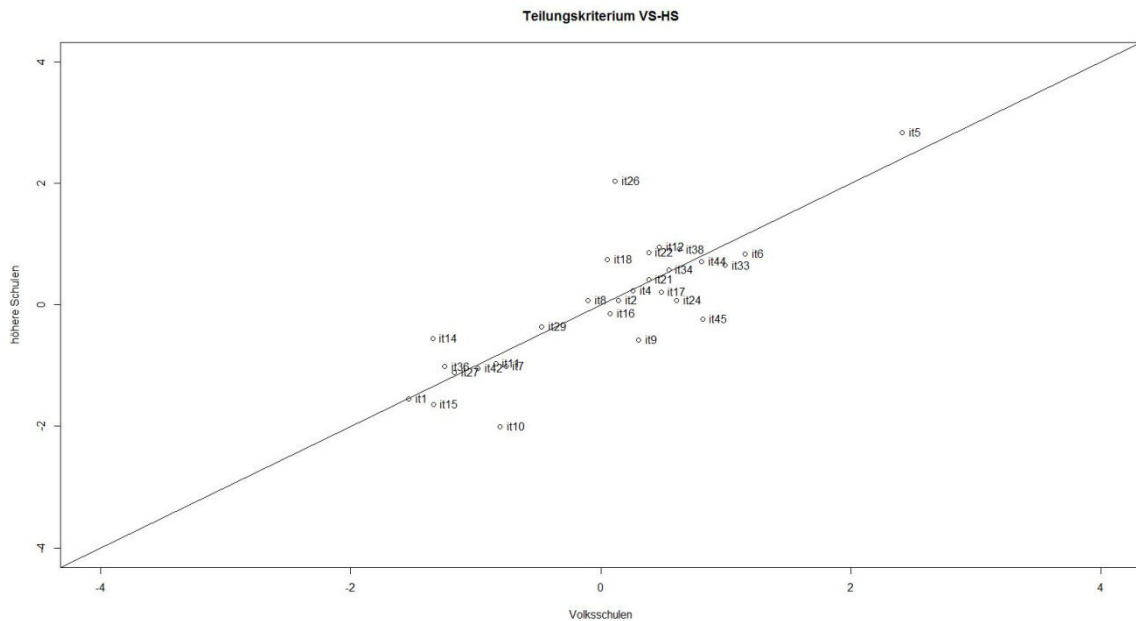


Abbildung 9: Der grafische Modelltest des Teilungskriteriums *Schulstufe* (VS-HS) vor dem Ausschluss eines Items.



Insgesamt flossen 61 der 69 Aufgaben in die Aufgabenhefte ein. Zwei Instruktionsbeispiele und sechs Items dieser Liste wurden nicht vorgegeben und somit auch nicht mittels *Rasch*-Modell-Analysen geprüft. Da die ursprüngliche Nummerierung der Aufgaben beibehalten wurde, können Itembezeichnungen bis zur Nummer 69 auftreten.

Im ersten Schritt wurde *Item 20* aus dem Itempool eliminiert, da es sowohl grafisch als auch im *z-Test* für die Teilungskriterien *Score* und *Geschlecht* auffällig wurde. Zudem konnte im Nachhinein noch eine weitere plausible Lösungsmöglichkeit für diese Aufgabe gefunden werden, die jedoch nicht im Lösungsschlüssel enthalten war. Demnach liegt bei dem Ausschluss dieses Items auch ein inhaltlicher Grund vor.

Nach dem Ausschluss des ersten Items gab es weiterhin in allen drei Teilungskriterien signifikante Modellabweichungen. Die nächste Elimination betraf das *Item 69*, welches ebenfalls besondere Auffälligkeiten in den Teilungskriterien *Score* und *Geschlecht* aufwies. Inhaltlich konnten zwei weitere sinnvolle Lösungsmöglichkeiten gefunden werden, die nicht im Lösungsschlüssel enthalten waren.

Auch nach dem Ausschluss des zweiten Items bestanden für alle drei Teilungskriterien Modellabweichungen, jedoch bereits in einer deutlich geringeren Ausprägung. Im nächsten Schritt wurde *Item 66* ausgeschlossen, da es grafisch in den Teilungskriterien *Score* und *Geschlecht* auffiel sowie im *z-Test* besonders im Teilungskriterium *Score* hervortrat. Inhaltlich konnten erneut weitere (zumindest umgangssprachlich) plausible Lösungsmöglichkeiten gefunden werden.

Eine weitere Auffälligkeit betraf *Item 21* im Teilungskriterium *Score*. Die Entscheidung für seinen Ausschluss wird zum einen dadurch begründet, dass es sich um ein *Linking-Item* jener Art handelt, welches sich in allen vorgegebenen Testformen befindet. Zum anderen erscheint mindestens eine weitere Lösungsmöglichkeit als plausibel und im Lösungsschlüssel fehlend. Nach dem Ausschluss des vierten Items ergab sich erstmals eine Modellanpassung für das Teilungskriterium *Score*.

Die nächste Elimination betraf das *Item 28*. Die Wahl fiel, neben grafischen Gesichtspunkten und deutlichen Ausprägungen im *z-Test* (das Teilungskriterium *Geschlecht* betreffend), auch aufgrund Erfahrungen während der Testvorgabe auf dieses Item. Während der Testvorgabe zeigte sich, dass ein Großteil der Testpersonen sämtlicher Altersstufen das in dem Satz enthaltene Verb nicht kannte. Dies führte vermutlich zu erheblichen Rateeffekten oder zur Unlösbarkeit dieser Aufgabe.

Im nächsten Schritt wurde *Item 41* aufgrund von deutlichen Auffälligkeiten im Teilungskriterium *Geschlecht* ausgeschlossen. Inhaltlich liegt die Vermutung nahe, dass dieses Item für männliche Testpersonen etwas leichter zu lösen ist als für weibliche (Anmerkung: Dieses Item ist ohne Kenntnisse des Regelwerks und der Fachsprache des Fußballsports nur schwer zu lösen.). Im grafischen Modelltest ist ein leichter geschlechtsbedingter Unterschied bzgl. der Lösbarkeit des Items beobachtbar. Nach dem Ausschluss dieses Items ergab sich Modellgültigkeit auch für das Teilungskriterium *Geschlecht*.

Im nächsten Schritt wurde *Item 26* aus dem Itempool ausgeschlossen, da es im Teilungskriterium *Schulstufe* sowohl grafisch als auch im *z-Test* deutlich abwich. Dieses Item war für Testpersonen der höheren Schulstufen leichter zu lösen.

Inhaltlich konnte eine begriffliche Uneindeutigkeit gefunden werden, die für die Missverständnisse bei den jüngeren Teilnehmer/innen verantwortlich scheint.

Die letzte eliminierte Aufgabe war *Item 10*, aufgrund seiner Auffälligkeit im *z-Test* sowie im grafischen Modelltest im Teilungskriterium *Schulstufe*. Inhaltlich konnten keine Gründe für die Untauglichkeit dieser Aufgabe gefunden werden. Nach dem Ausschluss dieses Items ergab sich die Modellgültigkeit auch für das Teilungskriterium *Schulstufe*.

Nach dem Ausschluss von insgesamt acht Items konnte *a posteriori* Modellgültigkeit für alle Teilungskriterien erreicht werden. Die Gesamtelimination entspricht einem Anteil von 13,1 % gegenüber dem ursprünglichen Itempool.

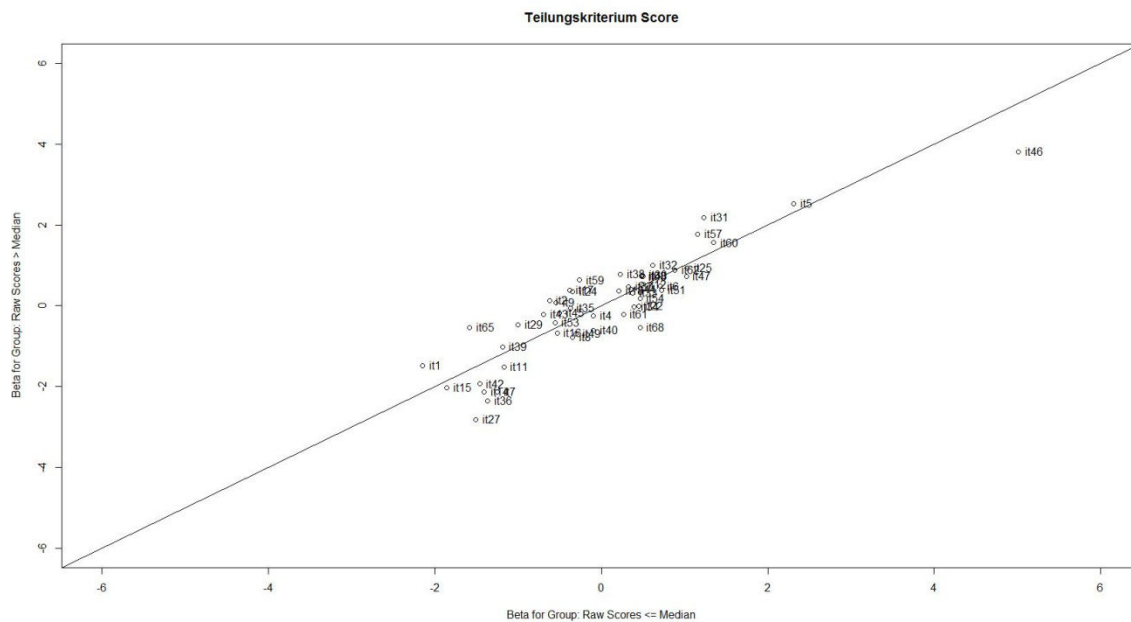
Die 53 verbliebenen Items des Untertests *Realitätssicherheit* messen somit die interessierende Eigenschaft *eindimensional*.

Nachfolgend werden die Ergebnisse der LR-Tests und die grafischen Modelltests der drei Teilungskriterien nach dem Ausschluss beschriebener Items dargestellt.

Tabelle 11: Die Ergebnisse der LRTs nach dem Ausschluss von 8 Items.

Teilungskriterium	$\chi^2 (LRT)$	<i>df</i>	$\chi^2_{krit} (\alpha=0.01)$	<i>p</i>
Score	65,583	49	74,91947	0,057
Geschlecht	68,993	52	78,61576	0,057
VS-HS	42,136	25	44,3141	0,017

Abbildung 10: Der grafische Modelltest für das Teilungskriterium *Score* nach dem Ausschluss von 8 Items.



Wie in Abbildung 10 ersichtlich, ergeben sich speziell für zwei Items deutliche grafische Auffälligkeiten (*Item 27*, *Item 46*). Für diese beiden Aufgaben sind jedoch keine Auffälligkeiten im *z-Test* erkennbar, speziell für *Item 27* auch keine inhaltlichen Gründe auffindbar. *Item 46* stellt die mit Abstand schwierigste Aufgabe im bestehenden Itempool dar. Da dies die einzig vorhandene Aufgabe mit dieser Schwierigkeitsausprägung ist, wäre ein zusätzlicher Ausschluss dieses Items für zukünftige Testungen zu überdenken.

Des Weiteren sei außerdem auf *Item 5* hingewiesen, da es in den Teilungskriterien *Geschlecht* und *Schulstufe* grafisch auffällt. Auch der *z-Test* für das Teilungskriterium *Geschlecht* zeigt deutliche Abweichungen diese Aufgabe betreffend. Inhaltliche Überlegungen legen nahe, dass das gesuchte zusammengesetzte Nomen an Aktualität verloren hat und nicht mehr dem heutigen Sprachgebrauch entspricht. Bei *Item 5* handelt es sich nach *Item 46*, um das zweitschwierigste Item (siehe Tabelle 12).

Abbildung 11: Der grafische Modelltest für das Teilungskriterium *Geschlecht* nach dem Ausschluss von 8 Items.

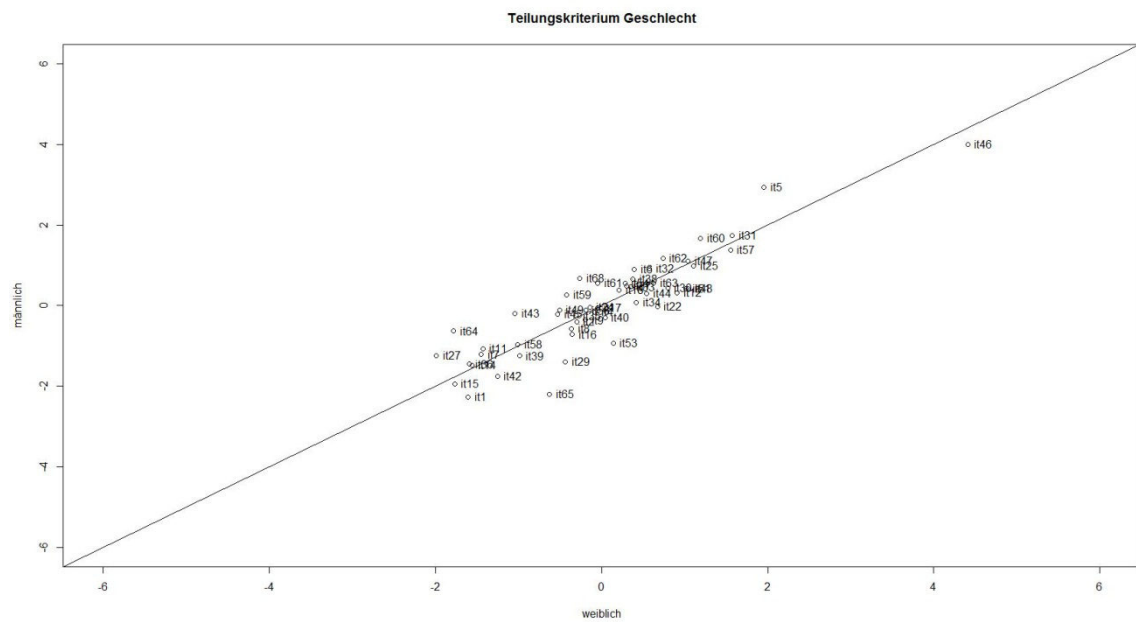
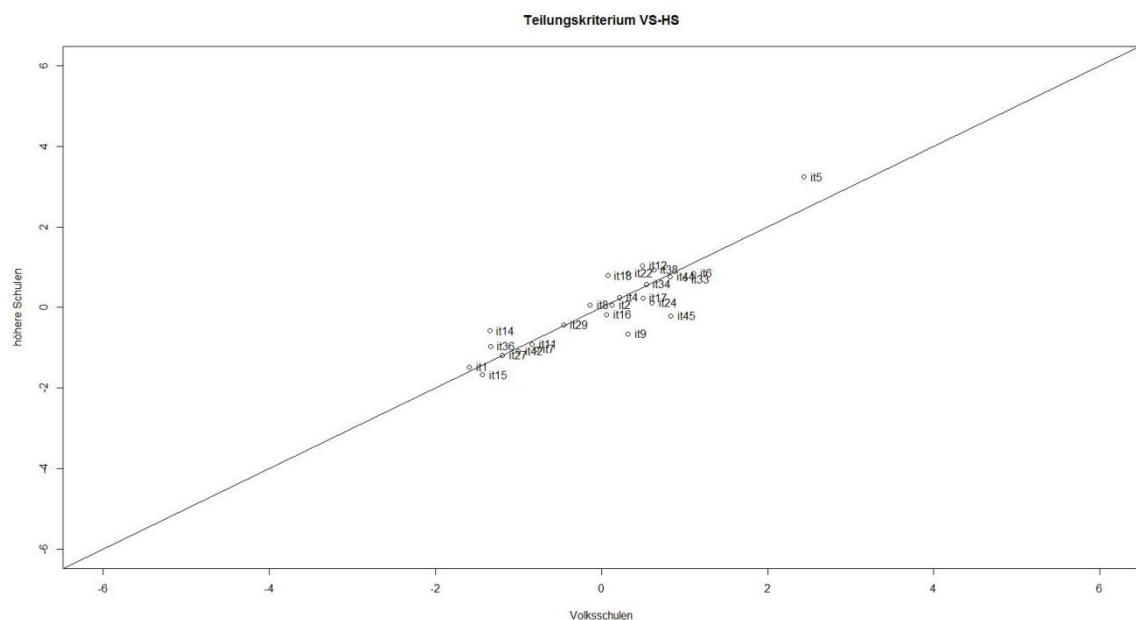


Abbildung 12: Der grafische Modelltest für das Teilungskriterium *Schulstufe (VS-HS)* nach dem Ausschluss von 8 Items.



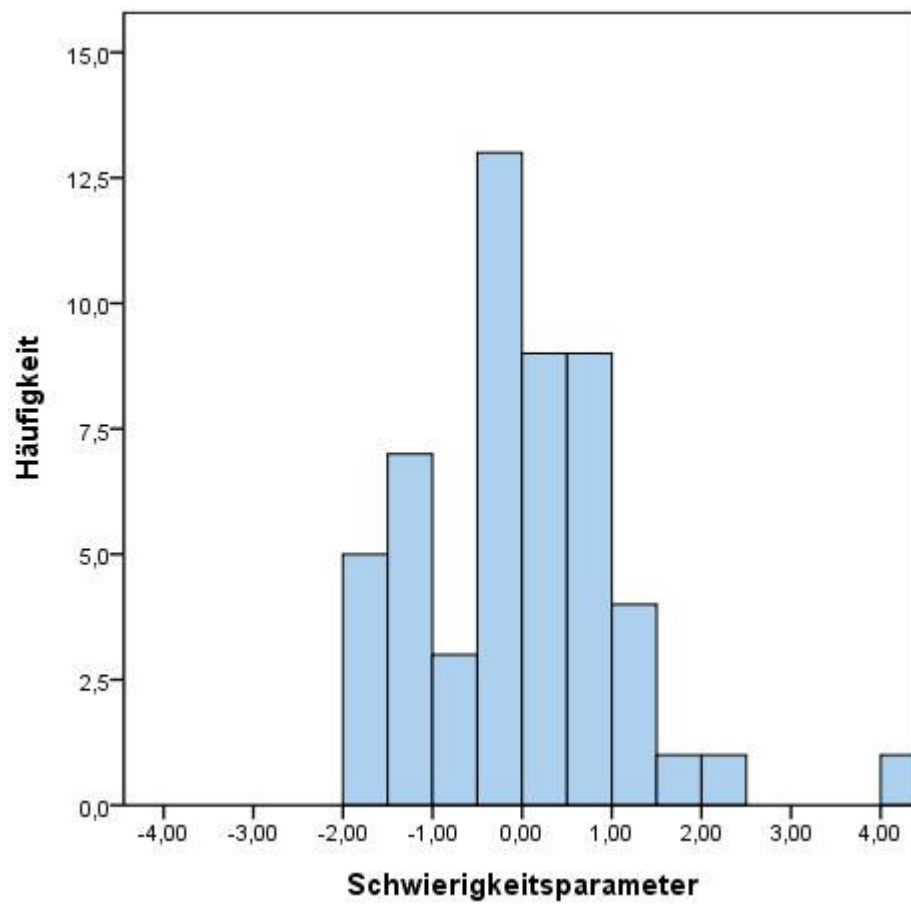
12.2 Itemschwierigkeitsparameter - *Realitätssicherheit*

In Tabelle 12 werden die Itemschwierigkeitsparameter der 53 *Rasch*-Modell-konformen Items dargestellt. Theoretisch kann der Itemschwierigkeitsparameter ξ Werte zwischen $-\infty$ und $+\infty$ annehmen, praktisch liegt er hingegen meist zwischen -5 und +5 (Kubinger, 1989). Die Schwierigkeitsparameter für den Untertest *Realitätssicherheit* sind, mit Ausnahme der oben genannten Items (*Item 46* und *Item 5*), in einem relativ schmalen Wertebereich verteilt (von -1,9 bis +1,6). Demnach vermögen die Aufgaben es nicht über einen weit gestreuten Fähigkeitsbereich zu differenzieren.

Tabelle 12: Itemschwierigkeitsparameter der *Rasch*-Modell-konformen Items des Untertests *Realitätssicherheit*; Rangreihung nach Schwierigkeit

Item	σ_i	Lower KI	Upper KI	Item	σ_i	Lower KI	Upper KI
it1	-1.941	-2.292	-1.590	it59	-0.098	-0.492	0.297
it15	-1.852	-2.249	-1.455	it17	-0.064	-0.328	0.201
it27	-1.594	-2.093	-1.095	it34	0.239	-0.118	0.596
it36	-1.513	-1.939	-1.088	it61	0.274	-0.130	0.677
it14	-1.509	-1.848	-1.170	it18	0.287	0.026	0.549
it42	-1.474	-1.867	-1.080	it22	0.337	0.047	0.628
it65	-1.345	-2.375	-0.315	it68	0.340	-0.251	0.932
it7	-1.339	-1.763	-0.916	it37	0.386	0.082	0.691
it11	-1.225	-1.542	-0.908	it33	0.394	0.205	0.583
it39	-1.116	-1.536	-0.697	it44	0.405	0.212	0.598
it58	-1.009	-1.950	-0.068	it54	0.425	-0.155	1.005
it64	-1.009	-1.950	-0.068	it38	0.510	0.203	0.816
it29	-0.844	-1.270	-0.418	it12	0.548	0.232	0.864
it43	-0.593	-0.976	-0.210	it63	0.574	0.080	1.069
it16	-0.522	-0.848	-0.195	it30	0.615	0.281	0.949
it8	-0.466	-0.770	-0.163	it6	0.628	0.277	0.979
it53	-0.427	-1.209	0.354	it48	0.651	0.060	1.243
it45	-0.378	-0.677	-0.079	it51	0.702	0.342	1.062
it2	-0.349	-0.724	0.027	it32	0.759	0.395	1.123
it49	-0.306	-0.758	0.147	it62	0.944	0.585	1.302
it35	-0.295	-0.703	0.113	it25	1.029	0.704	1.354
it9	-0.279	-0.579	0.021	it47	1.046	0.491	1.601
it56	-0.175	-0.676	0.326	it57	1.417	1.053	1.781
it40	-0.128	-0.450	0.194	it60	1.424	1.088	1.760
it4	-0.124	-0.420	0.171	it31	1.640	1.298	1.982
it24	-0.102	-0.439	0.235	it5	2.413	2.045	2.781
				it46	4.088	3.336	4.840

Abbildung 13: Die Verteilung der Schwierigkeitsparameter der 53 *Rasch*-Modell-konformen Items des Untertests *Realitätssicherheit*.



12.3 Bestimmung der *konvergenten Validität*

Die Bestimmung der *konvergenten Validität* als Teil der Konstruktvalidität sollte durch die Berechnung des Korrelationskoeffizienten der beiden AID-Testversionen erfolgen. Für diesen Zweck wurden die Personenparameter beider Testarten von 207 Testpersonen herangezogen. Die statistische Prüfung der Normalverteilung der Daten ergab für beide Variablen ($p_1 = 0,005$, $p_2 = 0,003$) keine Normalverteilung. Daher wurde das nicht-parametrische Rangkorrelationsverfahren nach Spearman angewendet. Die Rangkorrelation wurde mit der Statistik-Software *PASW Statistics 18.0* berechnet. Die Ergebnisse sind in Tabelle 13 angeführt.

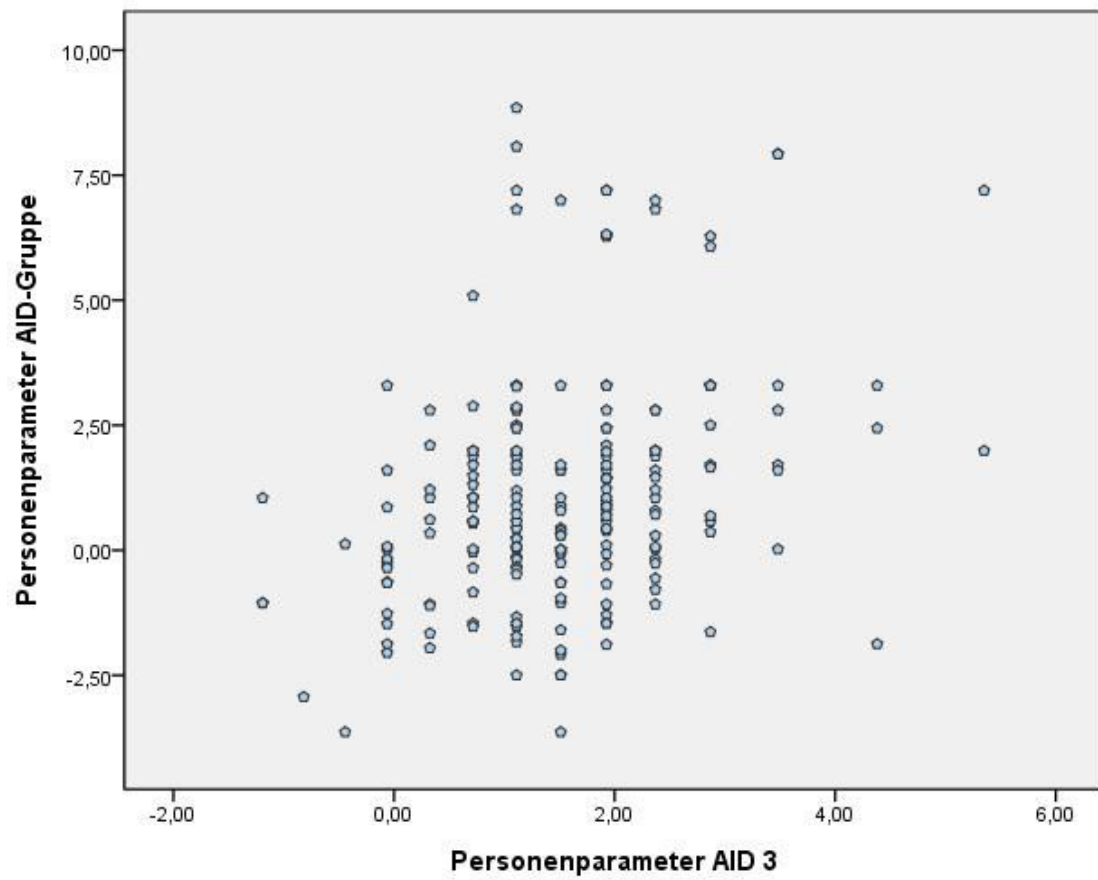
Tabelle 13: Die Rangkorrelation der Personenparameter für die Untertests *Realitätssicherheit*, des AID 3 und des AID-Gruppe.

Rangkorrelation nach Spearman		Personenparameter AID 3	Personenparameter AID-Gruppe
Personenparameter AID 3	Korrelationskoeffizient	1.000	.317
	p-Wert		.000
	Stichprobenanzahl	207	207
Personenparameter AID-Gruppe	Korrelationskoeffizient	.317	1.000
	p-Wert	.000	
	Stichprobenanzahl	207	207

Es besteht ein signifikanter Zusammenhang für die Untertests *Realitätssicherheit* der beiden AID-Versionen. Der Korrelationskoeffizient von $r = 0,317$ beschreibt jedoch einen niedrigen Zusammenhang. 10 % der Varianz des einen Tests werden durch den anderen Test erklärt. Demnach wird nur ein kleiner Teil desselben Konstrukts durch beide Testarten erfasst. 90 % der Varianz der Variabilität je Testart hängen von anderen nicht bekannten Einflussfaktoren ab (vgl. Rasch & Kubinger, 2006).

In Abbildung 14 wird der lineare Zusammenhang der beiden Variablen grafisch dargestellt. Für die Begründungen und Überlegungen zu den möglichen Unterschieden zwischen den beiden Versionen des Untertests *Realitätssicherheit* sei auf Abschnitt 13 verwiesen.

Abbildung 14: Streudiagramm der Personenparameter für den Untertest *Realitätssicherheit* des AID 3 und des AID-Gruppe



13. Diskussion und Ausblick

Das Ziel der vorliegenden Arbeit war, eine verbale Version des Untertests *Realitätssicherheit* zu konstruieren, um diesen in der AID-Gruppenversion vorgeben zu können. Es konnte ein Itempool von 53 *Rasch*-Modell-konformen Aufgaben entwickelt werden. Von der ursprünglich entwickelten Aufgabenmenge von 61 Items mussten acht Items ausgeschieden werden, um *a posteriori Rasch*-Modell-Konformität zu erreichen. Das entspricht einem Ausschluss von 13,1 %.

Durch die *Rasch*-Modell-Konformität wird außerdem dem Gütekriterium *Skalierung* Rechnung getragen, da die Anzahl der gelösten Aufgaben ein adäquates Maß für die zu messende Fähigkeit darstellt (vgl. Kubinger, 2009b).

Bei der inhaltlichen Betrachtung der auszuschließenden Items fiel auf, dass besonders jene ungeeignet waren, die mehrere plausible Lösungsmöglichkeiten beinhalteten, jedoch im Lösungsschlüssel nicht berücksichtigt wurden. Wie geläufig gewisse Vokabeln im allgemeinen Sprachgebrauch sind, stellte ein weiteres relevantes Merkmal für die Tauglichkeit von Items dar.

Eine weitere Auffälligkeit betraf die Schwierigkeit der entwickelten Testaufgaben. Mit Ausnahme von zwei Items (siehe Tabelle 12), befanden sich alle Aufgaben innerhalb eines sehr schmalen Schwierigkeitsspektrums. Ein normorientierter Test sollte mittlere Schwierigkeit im Verhältnis zu dem Kollektiv auf das er angewendet werden soll aufweisen, da die Differenzierungsfähigkeit auf diese Weise am größten ist (Lienert & Raatz, 1998). Das Schwierigkeitsspektrum (mit Ausnahme von *Item 5* und *Item 46*) befindet sich im Fähigkeitsbereich von -1,9 und +1,6. Kritisch anzumerken ist jedoch, dass Items mit anderen Schwierigkeitsausprägungen gänzlich fehlen. Die Konstruktion von schwierigen Items stellte sich bei der Aufgabenfindung als große Herausforderung dar und könnte zum Ziel zukünftiger Itementwicklungen - diesen Untertest betreffend -, gemacht werden. Für zukünftige Untersuchungen wären außerdem Testungen in allen neun österreichischen Bundesländern sowie in Deutschland und der Schweiz zweckmäßig.

Die Untersuchung zur *konvergenten Validität* mittels der beiden Untertests *Realitätssicherheit* aus dem AID 3 und AID-Gruppe ergab nur einen niedrigen Zusammenhang der beiden Versionen. Ein ähnliches Ergebnis war im Grunde

auch zu erwarten. Die Erfassung des Konstrukts *Realitätssicherheit* im AID 3 erfolgt durch bildhaftes Material und kann sprachfrei erfasst werden. Die Fähigkeit in der Gruppenversion hingegen verlangt vermutlich tiefgreifende sprachliche Kenntnisse und Fähigkeiten. Das Auffinden der richtigen zusammengesetzten Nomen erfordert Kenntnisse der deutschen Grammatik, einen umfassenden Wortschatz sowie ein erhebliches Maß an Sprachlogik und Sprachverständnis. Der Untertest *Realitätssicherheit* in der AID-Gruppenversion unterscheidet sich durch seine sprachliche Konzeption deutlich von jener der Version des AID 3. Der geringe Zusammenhang der beiden Versionen wird vermutlich hauptsächlich dadurch begründet, wobei auch der Vorgabemodus (Einzel- vs. Gruppentestung) an sich seinen Beitrag dazu leistet.

Das Potenzial dieses konstruierten Untertests liegt vermutlich in erster Linie in der Erfassung oben genannter sprachlichen Fähigkeiten. Weiterführende Untersuchungen dieses Konstrukt betreffend, sollten sich mit der Validierung im Zusammenhang mit sprachlichem Erhebungsmaterial beschäftigen, um umfassendere Erkenntnisse über die durch den Itempool zu erfassende Fähigkeit zu erlangen. Für die Erfassung des Konstrukts *Realitätssicherheit* in Anlehnung an den AID 3 wäre eine Itemkonstruktion mit bildhaftem Aufgabenmaterial anzudenken.

Die Gruppenvorgabe war durch exakt strukturierte Anweisungen für die Testleiter/innen standardisiert. Das Testgütekriterium *Objektivität* kann somit als ausreichend erfüllt angesehen werden. Inwieweit sich jedoch trotzdem Abweichungen (im Sinne der *Testleiterabhängigkeit*) in der Art der Vorgabe zwischen den verschiedenen Testleitern und Testleiterinnen ergaben, kann nicht beurteilt werden. Es ist jedoch denkbar, dass bestimmte Vokabeln von der Testleitung anhand von Synonymen klassenweise erklärt wurden und somit Testeffekte auftraten. Obwohl pro Schulklasse meist nur eine Testleiterin oder ein Testleiter anwesend war, konnten Absprachen unter den Testpersonen weitestgehend unterbunden werden. Eine zweite Aufsichtsperson während der Testungen wäre für eine reibungslosere Handhabung bei zukünftigen Erhebungen trotzdem zu empfehlen. Durch die genaue Festlegung eines Lösungsschlüssels kann die *Auswertungsobjektivität* und die *Interpretationseindeutigkeit* als erfüllt betrachtet werden.

Da im Teilungskriterium *Geschlecht* keine Auffälligkeiten zugunsten oder zum Nachteil von männlichen oder weiblichen Testpersonen auftraten, kann dem verbliebenen Itempool *Fairness* für deutschsprachige Testpersonen zugesprochen werden.

Vor- und Nachteile zum Thema der *Ökonomie* wurden in Abschnitt 6.2 ausführlich behandelt. Der vorhandenen Testversion können in jedem Fall die Hauptaspekte der Ökonomie von psychologischen Tests zugesprochen werden (Wirtschaftlichkeit und Aufwandsminimierung).

Die Testgütekriterien der *Zumutbarkeit* und der *Unverfälschbarkeit* sind sowohl durch die Testdauer als auch durch die Art der Vorgabe als erfüllt zu bewerten.

Das Kriterium der *Nützlichkeit* muss hingegen zum gegebenen Zeitpunkt kritisch betrachtet werden. Der vorliegende konstruierte Untertest *Realitätssicherheit* eignet sich für die Erfassung des benannten Konstrukts nur sehr eingeschränkt. Wie oben beschrieben, erscheint die Möglichkeit plausibel, dass sich dieser Untertest für die Erfassung tiefgreifender, sprachlogischer Fähigkeiten besser eignet. Um diese Behauptung zu prüfen, wären entsprechende Untersuchungen zur Bestimmung der Konstruktvalidität von Nöten.

Für die Testvorgaben des AID 3 mussten entsprechende Räumlichkeiten aufgesucht werden, in denen eine ungestörte Einzeltestung möglich war. Durch die räumliche Knappheit gewisser Schulen, war es nicht immer möglich, vollkommen lärmbefreite Orte aufsuchen zu können. Es wurde jedoch darauf geachtet, den entstehenden Lärm in den Schulpausen abzuwarten.

Den teilnehmenden Schülern und Schülerinnen wurden Rückmeldungen der Testergebnisse in anonymisierter Form zugesichert. Für diesen Zweck bekamen die Testpersonen ein Informationsschreiben unmittelbar nach den Testungen ausgehändigt (siehe Punkt V. Anhang). In diesem Schreiben war das Passwort zum Öffnen der zukünftig zugesandten Pdf-Dateien der Rückmeldungen angeführt. Es stellte sich jedoch heraus, dass die meisten Informationsschreiben bis zum Zeitpunkt der Rückmeldung verloren gegangen waren und der Großteil der Teilnehmer und Teilnehmerinnen somit erneut über das Passwort informiert werden musste. Um diesen Arbeitsaufwand zu minimieren, wäre ein Informationsschreiben über elektronischem Wege vermutlich sinnvoller. Dadurch

hätten die Testpersonen das Passwort bereits im entsprechenden Posteingang ihrer E-Mailadresse gespeichert. Außerdem würde sich daraus der Vorteil ergeben, dass die Verifizierung der Gültigkeit der angegebenen E-Mailadressen bereits unmittelbar nach den Testungen erfolgen kann.

Zusammenfassend kann zwar von einem gelungenen *Rasch*-Modell-konformen Untertest gesprochen werden, dessen Items dieselbe Fähigkeit messen. Jedoch bleibt die Frage nach dem zugrunde liegenden Konstrukts offen, welches eben durch diesen Untertest erfasst wird. Das Konstrukt *Realitätssicherheit*, in seiner Darstellung innerhalb des AID 3, kann nur in sehr geringer Ausprägung dem konstruierten Untertest der Gruppenvorgabe zugesprochen werden.

14. Zusammenfassung

Das Ziel der vorliegenden Arbeit war, eine verbale Umsetzung des Untertests *Realitätssicherheit* des AID 3 für die sich in Entwicklung befindende AID-Gruppenversion zu konstruieren. Der Geltungsbereich für die Gruppenvorgabe dieses Untertests reichte von der 3. bis zur 12. Schulstufe.

Die Entwicklung der Items sollte regel- und erfahrungsgeleitet vonstattengehen sowie Richtlinien für den sprachlichen Aufbau von Testaufgaben berücksichtigen. Die Erfassung der interessierenden Fähigkeit erfolgte durch unvollständige Sätze, in denen jeweils ein fehlender Wortteil eines zusammengesetzten Nomens aufzufinden und zu produzieren war. Jeder Testperson wurden ihrer Schulstufe entsprechend 20 Aufgaben in schriftlicher Form vorgelegt. Die Lösungswörter wurden durch ein freies Antwortformat, das durch die Beschränkung auf die im Satz befindlichen Nomen gebunden war, erfasst. Die Verrechnung der Lösungen erfolgte durch einen *a-priori* festgelegten Lösungsschlüssel.

Für die altersadäquate Vorgabe der Testaufgaben wurden die konstruierten Items im Vorfeld unterschiedlichen Schwierigkeitsstufen zugeteilt. Die daraus entstandenen acht verschiedenen altersbezogenen Testformen ergaben sich aus einer Kombination von 61 verschiedenen Items, die durch ein verlinktes Testdesign miteinander verknüpft waren. Für jede Testform gab es noch eine weitere parallele Testform, in der dieselben Aufgaben in unterschiedlicher Reihenfolge dargeboten waren. Insgesamt fanden 16 Testformen zu acht verschiedenen Formen ihre Anwendung. Der konstruierte Untertest wurde einer Stichprobe von 810 Testpersonen im Alter von 8 bis 19 Jahren in Oberösterreich und in Niederösterreich vorgegeben.

Neben der Gruppenvorgabe erhielten 209 zufällig ausgewählte Testpersonen zusätzlich das AID 3 vorgegeben. Die Vorgabe von AID 3 und AID-Gruppe erfolgte dabei in einem balancierten Testdesign.

Die gewonnenen Daten des Gruppenverfahrens wurden mittels *Rasch*-Modell-Analysen untersucht und anhand des *Likelihood-Ratio*-Tests nach Andersen (1973) auf Modellgültigkeit überprüft. Die dafür verwendeten Teilungskriterien waren *Score*, *Geschlecht* und *Schulstufe* (Volksschule vs. höhere Schulen). Nach dem schrittweisen Ausschluss acht nicht *Rasch*-Modell-konformer Items konnte a

posteriori Modellgültigkeit erzielt werden. Dies entspricht einem Itemausschluss von 13,1 %. Das zugrunde liegende Konstrukt der verbliebenen 53 Items wird *eindimensional* erfasst.

In weiterer Folge wurde die *konvergente* Validität des Untertests *Realitätssicherheit* in Bezug auf den entsprechenden Untertest im AID 3 bestimmt. Für diesen Zweck wurde der Korrelationskoeffizient der Personenparameter beider Vorgabemodi herangezogen. Der signifikante Zusammenhang zwischen den Untertests der beiden Testversionen war gering ($r = 0,317$). Das zu erfassende Konstrukt des Untertests *Realitätssicherheit* in der Testversion des AID-Gruppe kann durch 10% der Varianz des Untertests im AID 3 erklärt werden.

Gründe für den niedrigen Zusammenhang zwischen den beiden Erfassungsmethoden des Konstrukts *Realitätssicherheit* werden in den verschiedenen Verarbeitungsarten vermutet. Die Erfassung der interessierenden Fähigkeit erfolgt im AID 3 durch sprachfreies, bildhaftes Material, im AID-Gruppe hingegen durch sprachlich-schriftliche Erfassung. Für die Bearbeitung der Testaufgaben des konstruierten Untertests sind tiefgreifende sprachlogische Fähigkeiten von Nöten sowie erhebliche Wortschatz- und Grammatikkenntnisse der deutschen Sprache.

Zukünftige Validierungsuntersuchungen im Zusammenhang mit verbaler Leistungsfähigkeit wären zur Klärung des mit dem konstruierten Untertest erfassbaren Fähigkeitskonstrukts zu empfehlen. Für die Umsetzung eines Untertests zur Erfassung des Konstrukts *Realitätssicherheit* für die Intelligenz-Testbatterie AID-Gruppe in Anlehnung an den AID 3, könnte eine Entwicklung von Items mit Bildmaterialien angedacht werden.

Literaturverzeichnis

Amelang, M., Bartussek, D., Stemmler, G. & Hagemann, D. (2006). *Differentielle Psychologie und Persönlichkeitsforschung*. (6., überarb. Aufl.). Stuttgart: Kohlhammer.

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. akt. und erw. Aufl.). München: Pearson Studium.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: a critical experiment. *Journal of Educational Psychology*, 54, 1-22.

Fischer, G. H. (1974). *Einführung in die Theorie psychologischer Tests. Grundlagen und Anwendungen*. Bern: Hans Huber.

Glas, C.A.W., Verhelst, N.D. (1995). Testing the Rasch Model. In G.H. Fischer & I.W. Molenaar (Eds.). *Rasch Models – Foundations, Recent Developments, and Applications* (p. 69-95). New York: Springer.

Goldmann, J. (2011). *Einzel- versus Gruppenvorgabe der Untertests Antonyme, Synonyme Finden und Funktionen Abstrahieren des AID 3*. Unveröffentlichte Diplomarbeit, Universität Wien, Wien.

Herle, M. (2003). Individualtestung. . In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 205-206). Weinheim: Beltz.

Horn, J. L. & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, 57, 253-270.

- Jäger, R. S. (1988). Psychologische Diagnostik: ein Lehrbuch. München-Weinheim: Psychologie Verlags Union.
- John, O. P., Angleitner, A. & Ostendorf, F. (1988). The lexical approach to personality: a historical review of trait taxonomic research. *European Journal of Personality*, 2(3), 171-203.
- 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. & Wurst, E. (2000). AID 2: *Adaptives Intelligenz Diagnostikum 2*. Göttingen: Beltz Test.
- Kubinger, K. D. (2003). Adaptives Testen. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 1-9). Weinheim: Beltz.
- Kubinger, K. D. (2003). Gütekriterien. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 195-204). Weinheim: Beltz.
- Kubinger, K. D. (2003). Probabilistische Testtheorie. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 415-423). Weinheim: Beltz.
- Kubinger, K. D. & Draxler, C. (2007). Probleme bei der Testkonstruktion nach dem Rasch-Modell. *Diagnostica*, 53, 131-143.
- 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. überarbeitete und erweiterte Aufl.). Göttingen: Hogrefe.
- Kubinger, K. D. & Holocher-Ertl, S. (2013). *Adaptives Intelligenz Diagnostikum – Version 3.1 (AID 3)*. [in Druck]
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse*. (6. Aufl.). Weinheim: Beltz.
- Lorenz, R. (2012). *Machbarkeitsstudie zum Untertest „Alltagswissen“ des AID-Gruppe: erste Befunde zur psychometrischen Güte und konvergenten Validität*. Unveröffentlichte Diplomarbeit, Universität Wien, Wien
- 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> [05.01.2013]
- Rasch, D. & Kubinger, K. D. (2006). *Statistik für das Psychologiestudium*. München: Elsevier.
- Reindl, A. (1998). *Einzel- vs. Gruppenvorgabe: Niveauunterschiede im Leistungsbereich*. Unveröffentlichte Diplomarbeit, Universität Wien, Wien.
- Rost, J. (2004). *Lehrbuch Testtheorie-Testkonstruktion* (2. überarbeitete und erweiterte Aufl.). Bern: Hans Huber.
- Weber, B. (2011). *Konstruktion des sprachlichen Untertests „Antonyme finden“ für die Intelligenztestbatterie AID 3*. Unveröffentlichte Diplomarbeit, Universität Wien, Wien.
- Wright, B. D. & Stone, M. H. (1979). *Best test design. Rasch measurement*. Chicago: Mesa.

Scheffl, M. (2011). *Einzel- vs. Gruppenvorgabe der AID 3 Untertests*
Alltagswissen, Angewandtes Rechnen, Synonyme Finden, Funktionen
Abstrahieren und Soziales Erfassen und Sachliches Reflektieren.
Unveröffentlichte Diplomarbeit, Universität Wien, Wien.

V. Anhang

Testdesign AID-Gruppe

[illegible]

Anmerkung: Testdesign für die 8 verschiedenen Testformen zu je 2 Paralleltestformen mit derselben Aufgabenliste in unterschiedlicher Reihenfolge (Bezeichnungen „a“ oder „b“).

Itemleichtigkeitsparameter vor einem Itemausschluss

Item easiness Parameters (beta) with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
beta it1	2.062	0.177	1.716	2.409
beta it2	0.511	0.186	0.146	0.876
beta it4	0.248	0.148	-0.042	0.538
beta it5	-2.146	0.178	-2.494	-1.797
beta it6	-0.516	0.176	-0.861	-0.171
beta it7	1.480	0.213	1.063	1.897
beta it8	0.599	0.151	0.302	0.895
beta it9	0.438	0.152	0.140	0.735
beta it10	1.854	0.137	1.585	2.124
beta it11	1.393	0.160	1.079	1.707
beta it12	-0.443	0.162	-0.760	-0.126
beta it14	1.667	0.170	1.333	2.001
beta it15	1.940	0.198	1.552	2.328
beta it16	0.669	0.163	0.349	0.989
beta it17	0.201	0.133	-0.060	0.462
beta it18	-0.118	0.131	-0.375	0.139
beta it20	-1.931	0.140	-2.206	-1.656
beta it21	0.073	0.098	-0.119	0.265
beta it22	-0.223	0.146	-0.508	0.063
beta it24	0.268	0.168	-0.061	0.597
beta it25	-0.837	0.160	-1.151	-0.523
beta it26	-0.674	0.184	-1.034	-0.314
beta it27	1.735	0.253	1.238	2.232
beta it28	-0.263	0.171	-0.599	0.073
beta it29	1.008	0.217	0.583	1.434
beta it30	-0.437	0.166	-0.761	-0.112
beta it31	-1.477	0.168	-1.807	-1.147
beta it32	-0.566	0.180	-0.918	-0.214
beta it33	-0.247	0.095	-0.433	-0.061
beta it34	-0.103	0.180	-0.455	0.249
beta it35	0.441	0.203	0.043	0.839
beta it36	1.609	0.214	1.189	2.029
beta it37	-0.197	0.151	-0.492	0.099
beta it38	-0.406	0.156	-0.711	-0.101
beta it39	1.260	0.210	0.848	1.672
beta it40	0.235	0.164	-0.086	0.556
beta it41	-0.526	0.138	-0.796	-0.255
beta it42	1.609	0.198	1.221	1.998
beta it43	0.734	0.192	0.356	1.111
beta it44	-0.254	0.097	-0.443	-0.065
beta it45	0.519	0.151	0.223	0.816
beta it46	-3.941	0.370	-4.666	-3.216
beta it47	-1.067	0.282	-1.619	-0.514
beta it48	-0.670	0.300	-1.258	-0.081
beta it49	0.461	0.227	0.017	0.905
beta it51	-0.573	0.180	-0.926	-0.219
beta it53	0.416	0.398	-0.363	1.196
beta it54	-0.295	0.295	-0.873	0.283
beta it56	0.237	0.255	-0.264	0.737
beta it57	-1.359	0.185	-1.721	-0.998
beta it58	0.996	0.480	0.056	1.936
beta it59	0.247	0.196	-0.136	0.631
beta it60	-1.286	0.167	-1.613	-0.958
beta it61	-0.152	0.201	-0.546	0.242
beta it62	-0.868	0.180	-1.221	-0.514
beta it63	-0.358	0.241	-0.830	0.115
beta it64	0.996	0.480	0.056	1.936
beta it65	1.481	0.527	0.448	2.514
beta it66	-2.900	0.191	-3.275	-2.525
beta it68	-0.210	0.301	-0.800	0.380
beta it69	-2.348	0.171	-2.683	-2.013

Itemleichtigkeitsparameter nach Ausschluss von 8 Items

Item easiness Parameters (beta) with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
beta it1	1.941	0.179	1.590	2.292
beta it2	0.349	0.192	-0.027	0.724
beta it4	0.124	0.151	-0.171	0.420
beta it5	-2.413	0.188	-2.781	-2.045
beta it6	-0.628	0.179	-0.979	-0.277
beta it7	1.339	0.216	0.916	1.763
beta it8	0.466	0.155	0.163	0.770
beta it9	0.279	0.153	-0.021	0.579
beta it11	1.225	0.162	0.908	1.542
beta it12	-0.548	0.161	-0.864	-0.232
beta it14	1.509	0.173	1.170	1.848
beta it15	1.852	0.203	1.455	2.249
beta it16	0.522	0.167	0.195	0.848
beta it17	0.064	0.135	-0.201	0.328
beta it18	-0.287	0.134	-0.549	-0.026
beta it22	-0.337	0.148	-0.628	-0.047
beta it24	0.102	0.172	-0.235	0.439
beta it25	-1.029	0.166	-1.354	-0.704
beta it27	1.594	0.255	1.095	2.093
beta it29	0.844	0.217	0.418	1.270
beta it30	-0.615	0.171	-0.949	-0.281
beta it31	-1.640	0.174	-1.982	-1.298
beta it32	-0.759	0.186	-1.123	-0.395
beta it33	-0.394	0.097	-0.583	-0.205
beta it34	-0.239	0.182	-0.596	0.118
beta it35	0.295	0.208	-0.113	0.703
beta it36	1.513	0.217	1.088	1.939
beta it37	-0.386	0.156	-0.691	-0.082
beta it38	-0.510	0.156	-0.816	-0.203
beta it39	1.116	0.214	0.697	1.536
beta it40	0.128	0.164	-0.194	0.450
beta it42	1.474	0.201	1.080	1.867
beta it43	0.593	0.195	0.210	0.976
beta it44	-0.405	0.098	-0.598	-0.212
beta it45	0.378	0.153	0.079	0.677
beta it46	-4.088	0.384	-4.840	-3.336
beta it47	-1.046	0.283	-1.601	-0.491
beta it48	-0.651	0.302	-1.243	-0.060
beta it49	0.306	0.231	-0.147	0.758
beta it51	-0.702	0.184	-1.062	-0.342
beta it53	0.427	0.399	-0.354	1.209
beta it54	-0.425	0.296	-1.005	0.155
beta it56	0.175	0.256	-0.326	0.676
beta it57	-1.417	0.186	-1.781	-1.053
beta it58	1.009	0.480	0.068	1.950
beta it59	0.098	0.201	-0.297	0.492
beta it60	-1.424	0.172	-1.760	-1.088
beta it61	-0.274	0.206	-0.677	0.130
beta it62	-0.944	0.183	-1.302	-0.585
beta it63	-0.574	0.252	-1.069	-0.080
beta it64	1.009	0.480	0.068	1.950
beta it65	1.345	0.525	0.315	2.375
beta it68	-0.340	0.302	-0.932	0.251



Sehr geehrte Lehrerinnen und Lehrer!

Viele Kinder werden im Laufe ihrer schulischen Karriere aus den unterschiedlichsten Gründen mit psychologischen Tests untersucht. Bei schulpyschologischen Fragestellungen wie z.B. schulische Unter- bzw. Überforderung, Verhaltensprobleme im schulischen Kontext, Aufmerksamkeitsprobleme, Abklärung einer möglichen Legasthenie/Dyskalkulie, Schullaufbahnberatungen etc. kommt dabei meist ein Intelligenztest zum Einsatz, um die intellektuellen Stärken und Schwächen des Kindes abschätzen zu können.

Das *Adaptive Intelligenz Diagnostikum AID 2 – aktuelle Version 2.2 (AID 2.2, Kubinger, 2009)* – ist ein im deutschen Sprachraum sehr etabliertes Verfahren zur Erfassung der intellektuellen Fähigkeiten von Kindern und Jugendlichen zwischen 6 und 15 Jahren.

Im Zuge eines Forschungsprojekts der Universität Wien (Leitung: Univ.-Prof. Dr. Mag. Klaus D. Kubinger) findet nun eine Schüler(innen)-Testung mit einer gruppentauglichen Version des AID 3 (Kubinger & Holocher-Ertl, in Vorb.) – dem AID-Gruppe – statt.

Zusätzlich werden einige teilnehmende Kinder aus einigen Schulklassen ausgewählt, die neben der AID-Gruppen-Testung, noch mit einigen Untertests des AID 3 und psychologischen Verfahren zur Erfassung der Lern- und Leistungsmotivation, des Selbstkonzepts und leistungsbezogener Persönlichkeitseigenschaften getestet werden. Ziel ist es, neben der Entwicklung des Intelligenztests AID-Gruppe, zu untersuchen, wodurch Leistungsunterschiede in Einzel- und Gruppentests bedingt sind. Zudem soll untersucht werden, ob es möglich ist mit den beiden Verfahren AID-Gruppe und AID 3 dieselben Fähigkeitsbereiche bei Kindern und Jugendlichen zu erfassen, was eine wesentliche Information für praktisch tätige Psycholog(innen) bzw. Schulpyscholog(innen), die diese Verfahren in Zukunft einsetzen können, darstellt.

Wir wenden uns daher mit der Bitte an die Eltern Ihrer Schüler(innen), diese an der Untersuchung teilnehmen zu lassen, vorausgesetzt natürlich, dass das Kind/der Jugendliche damit einverstanden ist. Die Untersuchung findet während der Schulzeit in der Gruppe bzw. bei einem Teil der Kinder vor oder nach der Gruppentestung zusätzlich einzeln statt und dauert ca. ein bis zwei Stunden (je nach Zeitbedarf des Kindes). Durchgeführt werden die Testungen von speziell dafür ausgebildeten Testleitern(innen). Erfahrungsgemäß macht den Kindern die Mitarbeit an den Aufgaben viel Spaß. Möchte das Kind jedoch einmal eine Pause einlegen oder die Untersuchung aus irgendeinem Grund frühzeitig abbrechen, ist das natürlich jederzeit möglich.

Auf Wunsch werden wir den Eltern als kleines Dankeschön einen kurzen schriftlichen Ergebnisbericht über die intellektuellen Stärken und Schwächen des Kindes zuschicken.

Die gewonnenen Daten werden im Sinne des Datenschutzes ausschließlich für wissenschaftliche Zwecke genutzt. Sämtliche Ergebnisse der Schüler(innen) werden (noch während der Testung) bzw. nach Erstellung des schriftlichen Ergebnisberichts von den Testleitern(innen) anonymisiert. Es können keine Ergebnisse an Sie oder die Schuldirektion weitergegeben werden.

Wir möchten Sie bitten, die Elternbriefe an Ihre Schüler(innen) zu übergeben und die Rückmeldungen der Eltern gesammelt in der Direktion Ihrer Schule abzugeben. Wenn sich Eltern Ihrer Schüler(innen) mit der

Mag. Bettina Hagenmüller
Psychologin
Fakultät f. Psychologie der Universität Wien
Inst. f. Angewandte Psychologie: Gesundheit, Entwicklung und
Förderung
Arbeitsbereich Psychologische Diagnostik
Liebiggasse 5
A-1010 Wien
email: bettina.hagenmueller@univie.ac.at
Tel: 01 – 4277 47274



universität
wien



Test- und
Beratungsstelle

Wien, Mai bis Juni 2012

Liebe Eltern!

Viele Kinder werden im Laufe ihrer schulischen Karriere aus den unterschiedlichsten Gründen mit psychologischen Tests untersucht. Bei schulpsychologischen Fragestellungen wie z.B. schulische Unter- bzw. Überforderung, Schullaufbahnberatungen, Verhaltensprobleme im schulischen Kontext, Aufmerksamkeitsprobleme, Abklärung einer möglichen Legasthenie/Dyskalkulie etc. kommt dabei meist ein Intelligenztest zum Einsatz, um die intellektuellen Stärken und Schwächen des Kindes abschätzen zu können. Das *Adaptive Intelligenz Diagnostikum AID 2 – aktuelle Version 2.2 (AID 2.2, Kubinger, 2009)* - ist ein im deutschen Sprachraum sehr etabliertes Verfahren zur Erfassung der intellektuellen Fähigkeiten von Kindern und Jugendlichen zwischen 6 und 15 Jahren.

Im Zuge eines Forschungsprojekts der Universität Wien (Leitung: Univ.-Prof. Dr. Mag. Klaus D. Kubinger) findet nun eine Schüler(innen)-Testung mit einer gruppentauglichen Version des AID 3 (Kubinger & Holocher-Ertl, in Vorb.) – dem AID-Gruppe – statt.

Zusätzlich werden einige teilnehmende Kinder ausgewählt, die neben der AID-Gruppen-Testung, zusätzlich noch mit einigen Untertests des AID 3 und psychologischen Verfahren zur Erfassung der Lern- und Leistungsmotivation, des Selbstkonzepts und leistungsbezogener Persönlichkeitseigenschaften getestet werden.

Wir würden es sehr begrüßen, in dieses Forschungsprojekt auch Ihr Kind einbeziehen zu können.

Wir wenden uns daher mit der Bitte an Sie, Ihr Kind an dieser Untersuchung teilnehmen zu lassen, vorausgesetzt natürlich, Ihr Kind ist einverstanden. Die Untersuchung findet während der Schulzeit in der Gruppe bzw. bei einem Teil der Kinder zusätzlich vor oder nach der Gruppentestung einzeln statt und dauert ca. ein bis zwei Stunden. Durchgeführt werden die Testungen von speziell dafür ausgebildeten Testleitern(innen). Erfahrungsgemäß macht den Kindern die Mitarbeit an den Aufgaben viel Spaß (natürlich kann Ihr Kind dabei eine kleine Pause einlegen).

Auf Wunsch ist es auch möglich, Ihnen als kleines Dankeschön einen kurzen schriftlichen Ergebnisbericht über die intellektuellen Stärken und relativen Schwächen Ihres Kindes per E-Mail zuzuschicken. Das PDF können Sie mit dem Passwort „aidgruppe2012“ öffnen.

Für eventuelle Rückfragen steht Ihnen als Ansprechperson Frau Mag. Bettina Hagenmüller, Projektkoordinatorin (Tel: +43-1-4277 47274, email: bettina.hagenmueller@univie.ac.at), jederzeit gerne zur Verfügung.

Die gewonnenen Daten werden im Sinne des Datenschutzes ausschließlich für wissenschaftliche Zwecke genutzt. Sämtliche Ergebnisse der Schüler(innen) werden noch während der Testung von den Testleitern(innen) bzw. nach dem Versand des Ergebnisberichts anonymisiert. Direktion bzw. Lehrer(innen) der Schule werden selbstverständlich nicht über die Ergebnisse informiert.

Wir bitten Sie, mit Ihrer Unterschrift auf dem beiliegenden Formular, Ihr Einverständnis zur Teilnahme Ihres Kindes an der oben beschriebenen Untersuchung zu erteilen.

Mit freundlichen Grüßen und herzlichen Dank im Voraus!

Mag. Bettina Hagenmüller (im Auftrag der Projektleitung)

Ich erkläre mich mit der Teilnahme meiner Tochter/meines Sohnes

_____, geboren am _____, an der

Name des Kindes

Schüler(innen)-Erhebung zum AID-Gruppe einverstanden.

- ☐ Ich bitte um die Zusendung eines kurzen schriftlichen Ergebnisberichts in Form eines PDF an folgende E-Mail-Adresse:

Sollten Sie über keine E-Mail-Adresse verfügen und dennoch an dem Ergebnisbericht interessiert sein, melden Sie sich bitte telefonisch unter +43-1-4277 47274 bei Frau Mag. Bettina Hagenmüller.

- ☐ Ich wünsche keinen Ergebnisbericht.

Datum

Unterschrift des/der Erziehungsberechtigten

Mag. Bettina Hagenmüller
Psychologin
Fakultät f. Psychologie der Universität Wien
Inst. f. Angewandte Psychologie: Gesundheit, Entwicklung und
Förderung
Arbeitsbereich Psychologische Diagnostik
Liebiggasse 5
A-1010 Wien
email: bettina.hagenmueller@univie.ac.at
Tel: 01 – 4277 47274



universität
wien



Test- und
Beratungsstelle

Wien, Mai bis Juni 2012

Liebe Eltern!

Vielen herzlichen Dank, dass Ihr Kind an der Schüler(innen)-Testung mit dem Intelligenztest AID-Gruppe und/oder AID 3 teilnehmen konnte!

Wenn Sie einen Ergebnisbericht angefordert haben, schicken wir Ihnen diesen voraussichtlich bis Ende Oktober, an die von Ihnen angegebene E-Mail-Adresse als PDF. Das Passwort zum Öffnen des PDF lautet „aidgruppe2012“.

Nochmals möchten wir darauf hinweisen, dass nach Versenden des Ergebnisberichtes die Daten anonymisiert werden.

Mit freundlichen Grüßen und herzlichem Dank!

Mag. Bettina Hagenmüller (im Auftrag der Projektleitung)

Landesschulrat für Niederösterreich

Rennbahnstraße 29
3109 St. Pölten

Isr
noe

Frau
Mag. Bettina Hagenmüller

Per E-Mail: bettina.hagenmueller@univie.ac.at

Sachbearbeiterin:
Mag. Christina Unterberger

t: +43 2742 280 5370
f: +43 2742 280 1111
e: christina.unterberger@lsr-noe.gv.at

Beilage(n):

Präs.-420/861-2012

Datum: 22.05.2012

Betrifft:

Genehmigung der Durchführung einer empirischen Untersuchung

Der Landesschulrat für Niederösterreich genehmigt die Durchführung der vorgelegten empirischen Untersuchung im Rahmen eines Forschungsprojektes zum Thema „AID-Gruppe und AID 3 (Adaptives Intelligenz Diagnostikum)“ durch Angehörige der Fakultät für Psychologie an der Universität Wien (Institut für Angewandte Psychologie: Gesundheit, Entwicklung und Förderung/Arbeitsbereich Psychologische Diagnostik).

Die Untersuchung, die sich an SchülerInnen richtet (Gruppen- und Einzeltestung), darf in dem vorliegenden Umfang antragsgemäß an folgenden Schulen in Niederösterreich durchgeführt werden:

Allgemeinbildende Pflichtschulen:

Allgemeinbildende höhere Schulen:



Auf die Einhaltung der Datenschutzbestimmungen darf hingewiesen werden, außerdem ist vor Beginn der Erhebungen (Gruppen- und Einzeltestung) das Einverständnis der Eltern bzw. Erziehungsberechtigten und die Zustimmung der jeweiligen Direktion einzuholen.

Parteienverkehr
Dienstag 8-12 Uhr

<http://www.isr-noe.gv.at>
office@lsr-noe.gv.at
DVR: 0064394

Amtsstunden
Mo.-Fr. 8-16 Uhr

Die an dieser Untersuchung teilnehmenden SchülerInnen sind vor Beginn der Erhebung ausdrücklich auf die Freiwilligkeit ihrer Teilnahme hinzuweisen, außerdem ist deren Anonymität – sofern nicht die Erziehungsberechtigten eine Rückmeldung über die Ergebnisse ihres eigenen Kindes wünschen - zu wahren. Es ist darauf zu achten, dass die Durchführung der Untersuchung nur die unbedingt erforderliche Zeit in Anspruch nimmt.

Für den Amtsführenden Präsidenten
Dr. F r e u d e n s p r u n g
Wirkl. Hofrat

Mag. Bettina Hagenmüller
Fakultät f. Psychologie der Universität Wien
Inst. f. Angewandte Psychologie:
Gesundheit, Entwicklung und Förderung
Liebiggasse 5
A-1010 Wien
email: bettina.hagenmueller@univie.ac.at
Tel: 01 – 4277 47274



universität
wien



Test- und
Beratungsstelle

Wien, im November/Dezember 2012

Liebe Eltern!

Vielen Dank für Ihr Einverständnis zur Teilnahme Ihres Kindes an der Schüler(innen)erhebung zur Intelligenz-Testbatterie AID-Gruppe im Rahmen eines Forschungsprojekts der Universität Wien (Leitung: Univ.-Prof. Dr. Mag. Klaus D. Kubinger, Projektkoordination: Mag. Bettina Hagenmüller).

Das Adaptive Intelligenz Diagnostikum AID – aktuelle Version 2.2 (Kubinger, 2009) – ist ein im deutschen Sprachraum sehr etabliertes Verfahren zur Erfassung der intellektuellen Fähigkeiten von Kindern und Jugendlichen zwischen 6 und 15 Jahren.

Einzelne Untertests dieser Intelligenz-Testbatterie wurden nun im Rahmen dieses Forschungsprojekts derart umgestaltet, dass man Kinder bzw. Jugendliche im Alter von 8 bis 19 Jahren nun in der Gruppe (im Klassenverband) testen kann. Dazu wurden gänzlich neue Aufgaben entwickelt, um die intellektuellen Fähigkeiten der Kinder bzw. Jugendlichen zu erfassen, und diese nun erstmals an Schüler/innen in Niederösterreich und Oberösterreich erprobt.

Wir wollen Ihnen nun über die Testergebnisse Ihres Kindes berichten. Für eine anonymisierte Verarbeitung und Speicherung der Daten haben wir Ihrem Kind folgenden Probandencode zugewiesen:

ABXXX

Die Testung Ihres Kindes mit einigen Untertests des AID-Gruppe fand während der Unterrichtszeit statt und umfasste die Dauer von in etwa einer Stunde. Die Testung wurde dabei von einer/einem speziell dafür ausgebildeten Testleiter/in durchgeführt.

Testergebnisse:

Die Leistungen in den einzelnen Untertests wurden jeweils mit einer altersspezifischen Stichprobe aus Niederösterreich und Oberösterreich verglichen.

Die Testergebnisse werden in Prozenträngen (PR) angegeben, wobei ein Prozentrangwert (PR) von 25 bis 75 als durchschnittlich (alterstypisch) gilt. Werte darüber gelten als überdurchschnittlich, Werte darunter als unterdurchschnittlich.

Der PR gibt an, wie viel Prozent der Gleichaltrigen eine gleich gute oder niedrigere Leistung erbringen. So bedeutet ein PR = 40, dass 40 Prozent der Gleichaltrigen eine schlechtere oder gleich gute Leistung erbringen. 60 Prozent der Gleichaltrigen schneiden hingegen besser ab.

Untertest	Interpretation
Alltagswissen PR = 95	Es wird die Fähigkeit gemessen, sich Sachkenntnisse über Inhalte anzueignen, die in der heutigen Gesellschaft alltäglich sind (Wissen zu den Themen: Geschichte, Geografie, Verkehr, Sport, Kunst, Musik, Biologie, Chemie, Physik, Mathematik, Sprache, Wirtschaft, neue Medien). Die Leistung des/der Schülers/in liegt hier über dem Altersdurchschnitt.
Antonyme Finden PR = 95	Es wird die Fähigkeit gemessen, die Bedeutung sprachgebundener Begriffe zu erfassen, indem das Gegenteil zu vorgegebenen Begriffen gefunden werden soll. Es wird hierbei die Größe des Wortschatzes bzw. das Sprachvermögen erfasst. Die Leistung des/der Schülers/in liegt hier über dem Altersdurchschnitt.
Realitätssicherheit PR = 91	Es wird die Fähigkeit gemessen, die Dinge des Alltags sprachgebunden auf deren Korrektheit zu prüfen. Dabei sollen sinnfreie Sätze durch das Ersetzen eines Wortes korrigiert werden. Die Leistung des/der Schülers/in liegt hier über dem Altersdurchschnitt.
Angewandtes Rechnen PR = 80	Es wird die Fähigkeit gemessen, (weitestgehend) unabhängig von schulischen Rechenfertigkeiten, alltägliche Problemstellungen durch Anwendung passender Rechenoperationen zu lösen. Die Leistung des/der Schülers/in liegt hier über dem Altersdurchschnitt.
Funktionen abstrahieren PR = 91	Es wird die Fähigkeit gemessen, durch schlussfolgerndes Denken im sprachlichen Bereich, die gemeinsame Funktion vorgegebener Objekte zu abstrahieren (bzw. zu erkennen). Es wird hierbei das sprachliche Abstraktionsvermögen erfasst. Die Leistung des/der Schülers/in liegt hier über dem Altersdurchschnitt.

Die Leistungen Ihres Kindes wurden im Rahmen eines Forschungsprojektes gewonnen und sind daher nur eingeschränkt aussagekräftig. Sollten Sie genauere Informationen zu der Leistungsfähigkeit Ihres Kindes wünschen, so empfehlen wir Ihnen eine nochmalige Testung mit ausführlicher Beratung bei einem/einer niedergelassenen Kinderpsychologen/in (www.psychologie.at).

Nochmals vielen Dank für Ihre Teilnahme und alles Gute für Sie und für die Zukunft Ihres Kindes.

Mag. Bettina Hagenmüller

Lebenslauf

Persönliche Daten

Name: Angelika Erika Bichler
Geburtsdaten: St. Pölten, 05.05.1980
Staatbürgerschaft: Österreich
E-Mail: angelika.bichler80@gmail.com

Schulische Ausbildung

1998-lfd. Studium Psychologie, 1010 Wien
09/1998 Reifeprüfung
1990-1998 BG/BRG Josefstrasse, 3100 St. Pölten

Berufliche Tätigkeiten

2011-lfd. Trainerin im Rahmen des Caritas-Microsoft-Computer-ABCs für Frauen, Landesstelle St. Pölten
2007-lfd. Gastronomische Tätigkeit in der „Aquacity“, St. Pölten
2005-lfd. Mitarbeiterin bei Cimedia-Print
Medienvertrieb bzw. Feibra Medienvertrieb
1998-2011 Projektmitarbeiterin Kuratorium für Verkehrssicherheit, Landesstelle St. Pölten
2009 Angestellte des Kuratorium für Verkehrssicherheit, zuständig für Verkehrssicherheitsaktionen, Administration und Gestaltung von Informationsmaterial
2005-2007 Gastronomische Tätigkeit im Buffet der NÖGKK, St. Pölten
2000-2004 Gastronomische Tätigkeit im „Cafe Amadeus“, St. Pölten
Juli-August 2001 Mitarbeit bei der Volkszählung der Stadt St. Pölten
1999 Gastronomische Tätigkeit, Internetcafe-Tor zum Landhaus, St. Pölten

Publikationen/Veröffentlichungen

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

Praktika

02/08/2011-23/09/2011 Klinisch-psychologisches Praktikum im Rahmen des Studiums
im Landeskrankenhaus St. Pölten, Neurologische Abteilung.

Fremdsprachenkenntnisse

Englisch	fließend (in Wort und Schrift)
Russisch	fortgeschritten

EDV-Kenntnisse

MS Office (Word, Excel, Power Point)
SPSS, R