



universität
wien

Diplomarbeit

Titel der Arbeit

Das AID 2 als Gruppen- oder Computerverfahren –
Eine Machbarkeitsstudie

Verfasserin

Angelika Eiter

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, im März 2011

Studienkennzahl: A - 298

Studienrichtung: Psychologie

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

Danksagung

Der erste Dank geht an Univ. Prof. Mag. Dr. Kubinger für die Vergabe dieses interessanten Themas an mich sowie für seine Unterstützung und seine Geduld. Jan Steinfeld danke ich für die Hilfe bei der Auswertung mit dem Programm R.

Vielen Dank an Mag. Johanna Böck, die mich in einer schwierigen Zeit unterstützt hat und ohne die die Erhebung nicht möglich gewesen wäre.

Den DirektorInnen, LehrerInnen und SchülerInnen der Schulen in Oberösterreich, die diese Untersuchung ermöglicht haben, danke ich ebenfalls sehr.

Der größte Dank gilt natürlich meiner Familie. Meinen Eltern, ohne die dieses Studium nicht möglich gewesen wäre, und die immer an mich geglaubt haben und hinter mir stehen, danke ich sehr. Auch René Schwenninger danke ich für seine Hilfe und seine Unterstützung.

Kurzzusammenfassung

Diese Arbeit beschreibt die Entwicklung neuer Items für die Untertests „Alltagswissen“, „Realitätssicherheit“ und „Angewandtes Rechnen“ des Adaptiven Intelligenz Diagnostikum 2 (Kubinger, 2009a) für ein Gruppen- oder Computerverfahren. Dafür wurde eine regelgeleitete Herangehensweise gewählt, mit dem Ziel einen Regelkatalog zu erstellen. Die so konstruierten Items der Untertests „Alltagswissen“ und „Angewandtes Rechnen“ wurden an 679 Kindern zwischen 8;00 und 15;11 Jahren in 4 Schulen in Oberösterreich auf Rasch-Modell-Konformität überprüft. Dabei konnte auch nach Ausschluss von Items anhand itemweiser, grafischer und qualitativer Methoden a posteriori für keine der vier Testformen des Untertests „Alltagswissen“ eine Konformität festgestellt werden. Hingegen waren drei der vier Testformen des Untertests „Angewandtes Rechnen“ dem Rasch-Modell konform. In zukünftigen Untersuchungen sollten deshalb für den Untertest „Alltagswissen“ mehr Regeln zur Erstellung von Fragen gefunden werden. Zusätzlich wäre eine größere Stichprobe nötig. Bei miteinander verbundenen Testformen wäre eine Überprüfung des Untertests mittels dem LLTM möglich, um die Passung der Regeln kontrollieren zu können, um einen verbindlichen Regelkatalog zur Itemkonstruktion erstellen zu können.

Schlüsselwörter: AID 2, Gruppentest, Computertest, Itemkonstruktion, Rasch-Modell

Abstract

Within this theses new items for the AID 2-subtests (Kubinger, 2009a) "everyday knowledge" ("Alltagswissen"), "competence in realism" ("Realitätssicherheit") and "applied mathematics" ("Angewandtes Rechnen") were developed for paper pencil and computer administration. For this reason a rule-based approach was chosen, to create a set of rules. The new items of "everyday knowledge" and "applied mathematics" have been administered in 679 8;00 to 15;11 year old students in Upper Austria. Even after eliminating 20% items a Rasch model fit wasn't given for the four test versions for the "everyday knowledge"-items. In contrast three out of four test versions of "applied mathematics" did fit the model. Further research should focus on finding more strategies for "everyday knowledge". To examine the fit of the rules and to create further obliging rules an LLTM should be calculated with a larger sample.

Keywords: AID 2, group setting, computer setting, item-generation, Rasch model

Inhaltsverzeichnis

I	Einleitung	1
II	Theoretischer Teil.....	2
1.	Adaptives Intelligenz Diagnostikum 2	2
2.	Adaptives Testen	3
2.1	Branched-testing	3
2.2	Tailored-testing	4
3	Itemgenerierung	4
3.1	Systematische Itemkonstruktion	5
3.2	Cognitive Design System Approach	5
3.3	Item-Modell-basierter Ansatz.....	5
3.4	Automatisierter Min-Max Ansatz.....	5
3.5	Fragemodus.....	6
3.6	Antwortmodus.....	6
III	Empirischer Teil.....	6
4	Alltagswissen	6
4.1	Untersuchung der Originalitems des AID 2.....	7
4.2	Ökologische Systemtheorie von U. Bronfenbrenner – kurzer Überblick	7
4.3	Weitere Untersuchungen	8
4.4	Gestaltung neuer Items	9
5	Realitätssicherheit	10
5.1	Untersuchung der Originalitems des AID 2.....	11
5.2	Gestaltung neuer Items	11
6	Angewandtes Rechnen	13
6.1	Untersuchung der Originalitems des AID 2.....	13
6.2	Weitere Untersuchungen	14
6.3	Gestaltung neuer Items	15
7	Durchführung der Untersuchung	18
7.1	Wahl der Analysestichprobe.....	18
7.2	Testvorgabe.....	18
7.3	Beobachtungen zur Testung	19
7.4	Verrechnung	20
7.5	Aufgabenanalyse	20
8	Deskriptive Statistik.....	21
8.1	Geschlecht.....	21

8.2	Region	22
8.3	Schultyp.....	23
8.4	Schulstufe	24
9	Ergebnisse	26
9.1	Alltagswissen, Testform A	26
9.1.1	<i>Erste Parameterschätzung</i>	26
9.1.2	<i>Itemselektion</i>	27
9.1.3	<i>Parameterschätzung und Modelltest nach Itemselektion</i>	32
9.2	Alltagswissen, Testform B	33
9.2.1	<i>Erste und zweite Parameterschätzung</i>	33
9.3	Alltagswissen, Testform C	34
9.3.1	<i>Erste Parameterschätzung</i>	34
9.3.2	<i>Itemselektion</i>	35
9.3.3	<i>Parameterschätzung und Modelltest nach Itemselektion</i>	40
9.4	Alltagswissen, Testform D	41
9.4.1	<i>Erste Parameterschätzung</i>	41
9.4.2	<i>Itemselektion</i>	41
9.5	Angewandtes Rechnen, Testform A	45
9.5.1	<i>Erste Parameterschätzung</i>	45
9.6	Angewandtes Rechnen, Testform B	48
9.6.1	<i>Erste und zweite Parameterschätzung</i>	48
9.7	Angewandtes Rechnen, Testform C	50
9.7.1	<i>Erste Parameterschätzung</i>	50
9.8	Angewandtes Rechnen, Testform D	51
9.8.1	<i>Erste Parameterschätzung</i>	51
9.9	Mittelwertvergleiche.....	53
9.9.1	<i>Ergebnisse Mittelwertvergleich</i>	55
10	Diskussion und Ausblick	57
11	Zusammenfassung	60
12	Literaturverzeichnis	62
13	Anhang.....	i
13.1	Anweisung für die Erstellung von Fragen für den Untertest „Alltagswissen“	i
13.2	Analyse der Items des Untertests 1 „Alltagswissen“	iii
13.3	Die gezeichneten Bilder des Untertests 2 „Realitätssicherheit“	vii
13.4	Analyse der Items des Untertests 3 „Angewandtes Rechnen“	xii

13.5	Anweisung für die Erstellung von Rechenbeispielen:.....	xvii
13.6	Anweisung für Untertest 1 „Alltagswissen“	xix
13.7	Fragen Untertest 1 „Alltagswissen“ Testform A	xix
13.8	Fragen Untertest 1 „Alltagswissen“ Testform B	xx
13.9	Fragen Untertest 1 „Alltagswissen“ Testform C	xxi
13.10	Fragen Untertest 1 „Alltagswissen“ Testform D	xxii
13.11	Anweisung für Untertest 3 „Angewandtes Rechnen“	xxiii
13.12	Fragen Untertest 3 „Angewandtes Rechnen“ Testform A	xxiii
13.13	Fragen Untertest 3 „Angewandtes Rechnen“ Testform B	xxv
13.14	Fragen Untertest 3 „Angewandtes Rechnen“ Testform C	xxvi
13.15	Fragen Untertest 3 „Angewandtes Rechnen“ Testform D	xxvii
13.16	Itemleichtigkeitsparameter	xxviii
13.17	Personenparameter	xxxi
13.18	Ansuchen an Landesschulrat.....	xxxiv
13.19	Beschreibung des Forschungsprojekts.....	xxxv
13.20	Elternbrief.....	xxxvi
13.21	Lebenslauf	xxxvii

Abbildungsverzeichnis

Abbildung 1: Schuh ohne Schnürsenkel.....	12
Abbildung 2: Violine ohne F-Löcher	12
Abbildung 3: Grafischer Modelltest, Alltagswissen A, TK: Höhe des Rohwertes ..	30
Abbildung 4: Grafischer Modelltest, Alltagswissen A, TK: Geschlecht	31
Abbildung 5: Grafischer Modelltest, Alltagswissen A, TK: Region	31
Abbildung 6: Grafischer Modelltest, Alltagswissen C, TK: Höhe des Rohwertes ..	38
Abbildung 7: Grafischer Modelltest, Alltagswissen C, TK: Geschlecht	38
Abbildung 8: Grafischer Modelltest, Alltagswissen C, TK: Region	39
Abbildung 9: Grafischer Modelltest, Alltagswissen D, TK: Höhe des Rohwertes ..	44
Abbildung 10: Grafischer Modelltest, Alltagswissen D, TK: Geschlecht	44
Abbildung 11: Grafischer Modelltest, Alltagswissen D, TK: Region	45
Abbildung 12: Grafischer Modelltest, Angewandtes Rechnen A, TK: Höhe des Rohwertes	46
Abbildung 13: Grafischer Modelltest, Angewandtes Rechnen A, TK: Geschlecht ..	47
Abbildung 14: Grafischer Modelltest, Angewandtes Rechnen A, TK: Region.....	47
Abbildung 15: Grafischer Modelltest, Angewandtes Rechnen B, TK: Höhe des Rohwertes	49
Abbildung 16: Grafischer Modelltest, Angewandtes Rechnen B, TK: Geschlecht ..	49
Abbildung 17: Grafischer Modelltest, Angewandtes Rechnen B, TK: Region.....	50
Abbildung 18: Grafischer Modelltest, Angewandtes Rechnen D, TK: Höhe des Rohwertes	52
Abbildung 19: Grafischer Modelltest, Angewandtes Rechnen D, TK: Geschlecht ..	52
Abbildung 20: Grafischer Modelltest, Angewandtes Rechnen D, TK: Region.....	53

Tabellenverzeichnis

Tabelle 1: Verteilung Geschlecht pro Untertest.....	21
Tabelle 2: Verteilung Geschlecht pro Testform.....	22
Tabelle 3: Verteilung Region pro Untertest	22
Tabelle 4: Verteilung Region pro Testform	23
Tabelle 5: Verteilung Schultyp pro Untertest.....	23
Tabelle 6: Verteilung Schultyp pro Testform.....	24
Tabelle 7: Verteilung Schulstufe pro Untertest	25
Tabelle 8: Verteilung Schulstufe pro Testform	25
Tabelle 9: Erste Parameterschätzung, Alltagswissen A, nicht Rasch-Modell- konform.....	27
Tabelle 10: Z-Test, Alltagswissen A, Teilungskriterien Intern, Geschlecht und Region.....	28
Tabelle 11: Itemfitstatistik, Alltagswissen A.....	29
Tabelle 12: Parameterschätzung nach Itemselektion, Alltagswissen A, nicht Rasch-Modell-konform	33
Tabelle 13: Zweite Parameterschätzung, Alltagswissen B, nicht Rasch-Modell- konform.....	34
Tabelle 14: Erste Parameterschätzung, Alltagswissen C, nicht Rasch-Modell- konform.....	35
Tabelle 15: Z-Test, Alltagswissen C, Teilungskriterien Intern, Geschlecht und Region.....	36
Tabelle 16: Itemfitstatistik, Alltagswissen C.....	37
Tabelle 17: Parameterschätzung nach Itemselektion, Alltagswissen C, nicht Rasch-Modell-konform	40
Tabelle 18: Erste Parameterschätzung, Alltagswissen D, nicht Rasch-Modell- konform.....	41
Tabelle 19: Z-Test, Alltagswissen D, Teilungskriterien Intern, Geschlecht und Region.....	42
Tabelle 20: Itemfitstatistik, Alltagswissen D	43
Tabelle 21 Erste Parameterschätzung, Angewandtes Rechnen A, Rasch-Modell- konform.....	46
Tabelle 22: Zweite Parameterschätzung, Angewandtes Rechnen B, Rasch-Modell- konform.....	48
Tabelle 23: Erste Parameterschätzung, Angewandtes Rechnen C, nicht Rasch- Modell-konform.....	51
Tabelle 24: Erste Parameterschätzung, Angewandtes Rechnen D, Rasch-Modell- konform.....	51
Tabelle 25: Dreifache Varianzanalyse, Alltagswissen, Testform A, C und D	55
Tabelle 26: Mittelwertvergleich, Alltagswissen, Testform B.....	55
Tabelle 27: Dreifache Varianzanalyse, Angewandtes Rechnen, Testform A & B..	55
Tabelle 28: Mittelwertvergleich, Angewandtes Rechnen, Testform C	56
Tabelle 29: Mittelwertvergleich, Angewandtes Rechnen, Testform D	56

I Einleitung

Seit dem Beginn des modernen Intelligenztests, eingeleitet durch Alfred Binet und mit Hilfe von Théodore Simon Anfang des 20. Jahrhunderts, wurden zahlreiche weitere Verfahren entwickelt. Diese unterschieden sich abgesehen von den zu Grunde liegenden Konzepten auch in der Art der Vorgabe. So können heutige Tests entweder mit einer Testperson oder mit mehreren auf einmal durchgeführt werden. Jede dieser Vorgabearten birgt verschiedene Vor- und Nachteile. So kann in der Einzeltestung Material eingesetzt werden, das in der Gruppentestung nicht eingesetzt werden kann. Als Beispiel für die Kinderdiagnostik seien hier z.B. die K-ABC (Melchers & Preuß, 2009) oder das AID 2 (Kubinger, 2009a) genannt.

Intelligenztests für Kinder und Jugendliche haben außerdem die große Aufgabe, zusätzlich den jeweiligen Entwicklungsstand miteinzubeziehen, um das Verfahren auch wirklich adäquat durchführen zu können. Einzeltestungen haben hier den Vorteil, dass der oder die TestleiterIn auf die Person eingehen kann. Dies ist grundsätzlich ein Vorteil, jedoch ist die Suche nach Verfahren ebenso gegeben, die Testleiterunabhängigkeit und damit größere Objektivität mit sich bringen. Unter manchen Umständen wird es außerdem notwendig und ökonomischer sein, viele Personen auf einmal testen zu können. Sei dies durch Gruppenverfahren unter Einsatz eines PP-Tests, oder mittels PC-Verfahrens. Im Zuge der immer wachsenden Möglichkeiten, die sich durch den Einsatz von Computer ergeben, wird auch eine Computertestung für Kinder und Jugendliche realistischer. Betrachtet man Konsolen wie Nintendo® Wii oder PlayStation 3, oder Handheld-Konsolen wie Nintendo® DS, PlayStation Portable und deren Nachfolger, so lassen sich auch für die Zukunft der Computertests immer mehr Möglichkeiten erahnen. Touchscreens & Co sind längst schon in den Alltag der Menschen gewandert und bei Lern- und Denkspielen der entsprechenden Anbieter werden diese Errungenschaften genutzt und eingesetzt. So könnte dies auch bei entsprechenden Programmierfertigkeiten für die psychologische Diagnostik von Nutzen sein.

In dieser Arbeit wurde deswegen versucht, die Machbarkeit eines Gruppen- oder PC-Tests auf Basis dreier Untertests des Adaptiven Intelligenz Diagnostikum 2 (Kubinger, 2009a) zu untersuchen. Die Diplomarbeit von Johanna Böck (2010)

behandelte dasselbe Thema zu drei anderen Untertests, die Erhebung der Daten erfolgte zusammen. Ähnlichkeiten der Arbeiten sind deswegen gegeben.

Im Folgenden erwartet den Leser nach einem kurzen Einblick in das AID 2 und Hintergründe zum Adaptiven Testen sowie zur Itemgenerierung die Vorstellung der drei Untertests „Alltagswissen“, „Realitätssicherheit“ und „Angewandtes Rechnen“, deren Untersuchung sowie die Prozedur der Gestaltung neuer Items. Danach wird die Durchführung der Erhebung der Daten in Schulklassen der 3. bis 9. Schulstufe vorgestellt mit anschließender deskriptiver Statistik. Zuletzt werden die Ergebnisse der Daten dargebracht.

II Theoretischer Teil

1. Adaptives Intelligenz Diagnostikum 2

An dieser Stelle sei ein Überblick des Adaptiven Intelligenz Diagnostikums 2, in Folge kurz AID 2, gegeben. Der interessierte Leser wird auf das Handbuch desselben von Kubinger (2009a) verwiesen.

Das AID 2 ist eine für Kinder und Jugendliche entwickelte Testbatterie, die komplexe und basale Kognitionen erfassen soll. An Wechsler thematisch angelehnt, soll eine Profilinterpretation von möglichst vielen Fähigkeiten erfolgen. Das AID 2 ist für den Altersbereich von 6;0 bis 15;11 Jahren konzipiert und wird als Individualverfahren vorgegeben. Für acht der elf Untertests (Alltagswissen, Angewandtes Rechnen, Synonyme Finden, Funktionen Abstrahieren, Soziales Erfassen und Sachliches Reflektieren, Soziale und Sachliche Folgerichtigkeit, Antizipieren und Kombinieren-figural, Analysieren und Synthetisieren-abstrakt) erfolgt die Vorgabe nach dem Prinzip des „branched-testings“, die restlichen drei Untertests (Realitätssicherheit, Unmittelbares Reproduzieren-numerisch, Kodieren und Assoziieren), sowie drei Zusatztests (Unmittelbares reproduzieren-figural/abstrakt), Merken und Einprägen, Strukturieren-visumotorisch) können konventionell vorgegeben werden.

2. Adaptives Testen

Um zu verstehen, was das Besondere am AID 2 mit seiner Vorgabeform ist, und warum diese bei einer Computertestung angestrebt wird und auch bei einer Gruppentestung von Nutzen sein kann, wird nun das Prinzip des adaptiven Testens erklärt. Ausführliche Beschreibungen findet der interessierte Leser bei Kubinger (1989 oder 2009b), Kubinger und Wild (1989), Wild (1989) oder Frey (2006).

Der Grundgedanke des adaptiven Testens ist, der Testperson abhängig von den Items die sie davor gelöst hat, die nächsten vorzugeben. Da jedes Item eine eigene Schwierigkeit aufweist werden dadurch Rückschlüsse auf die zugrunde liegende latente Eigenschaft gezogen. Ziel dieser Art der Vorgabe ist, die Testperson weder zu unter- noch zu überfordern, da sie ausschließlich die Items erhält, die der errechneten Fähigkeit entsprechen. Dadurch bekommt sie weder zu viele leichte, noch zu viele schwere Fragen. Durch diese Vorgangsweise bekommen Personen aber unterschiedliche Items vorgesetzt, was zu einem Problem der Vergleichbarkeit führen würde, wären die Items nicht nach einem Modell der Item-Response-Theorie überprüft worden. Dadurch erhalten die Items ihre jeweiligen Schwierigkeitsparameter.

Bei jeder Testung stellen sich dabei die Fragen „Mit welchem Item wird begonnen?“, „Mit welchem Item wird die Testung fortgesetzt?“ und „Wann ist der Test zu Ende?“ (Frey, 2006).

Diese Besonderheit des Vorgabemodus lässt zwei verschiedene Varianten zu, wobei die eine bei Papier-Bleistift-Einzeltestungen angewendet werden kann (branched-testing), die andere (tailored-testing) jedoch unbedingt einen Computer und dessen Rechenfähigkeit erfordert. Bei Papier-Bleistift-Gruppentestungen ist zwar auch das branched-testing nicht möglich, sind die Items jedoch nach der Item-Response-Theorie überprüft worden können Paralleltests konstruiert werden, die äquivalent sind.

2.1 Branched-testing

Bei dieser Form der Testung wird der Testperson eine Gruppe von Items vorgegeben. Je nachdem, wie viele davon richtig beantwortet werden, erhält die Person anschließend wieder eine Gruppe von Items die einen niedrigeren, fast

denselben oder einen höheren Schwierigkeitsparameter aufweisen. Das AID 2 verwendet diese Art der Testung. Dabei bekommen die Testpersonen in den entsprechenden Untertests 15 Items zu je 5-Item-Blöcken präsentiert. Dadurch wurde die Messgenauigkeit größtmöglich optimiert (Kubinger & Wild, 1989).

2.2 Tailored-testing

Diese computergestützte Art der Testung erlaubt eine noch genauere Anpassung des nächsten Items an die Testperson, da nach jeder Antwort der neue Personenparameter errechnet und das entsprechende Item ausgewählt wird. Erst bei Erfüllung eines bestimmten Abbruchkriteriums wird die Testung beendet. „Auf diese Weise kann z.B. ein aus 30 Items bestehender typischer konventioneller Test durch einen adaptiven der halben Länge ersetzt werden, ohne daß wesentliche Einbußen an Meßgenauigkeit in Kauf genommen werden müssen.“ (Wild, 1989, S. 163). Beim Adaptiven Matrizen Test (Hornke, Etzel & Rettig, 1999) oder beim Lexikon-Wissen-Test (Wagner-Menghin, 2004) wird diese Art der Testung eingesetzt. So ist sie auch der Gedanke an eine zukünftige Computerversion des AID 2.

Hornke (2001) schreibt in diesem Zusammenhang von der Unterscheidung zwischen einer *messorientierten* und *entscheidungsorientierten* Vorgehensweise. Dabei ergibt sich für die messorientierte Vorgehensweise mit einer Reliabilität von 0,85 im Durchschnitt eine Testlänge von 28 Items. Nach der entscheidungsorientierten Vorgehensweise kann nach etwa 9-10 Items eine Entscheidung gefällt werden. Hier steigt zwar der Messfehler, das Risiko einer Fehlentscheidung (z.B. $\alpha = 0,05$) wird jedoch eingehalten. Bei zusätzlichen Items würde zwar der Messfehler wieder sinken, die Entscheidung jedoch wäre dieselbe.

3 Itemgenerierung

Nach Irvine (2002) beginnt jede Form der Itemgenerierung mit der zugrunde liegenden Intelligenztheorie. Danach ist es wichtig, die Syntax und Grammatik der Items zu beachten und diese logisch zu konstruieren. Jedes Item wird weiter anhand von *radicals*, damit ist die Schwierigkeit der Items gemeint, und *incidentals*, also die inhaltlichen oberflächlichen Veränderungen, konstruiert.

Hornke et al. (2004) beschreiben vier Ansätze zur automatisierten Itemgenerierung:

3.1 Systematische Itemkonstruktion

In diesem nahezu top-down Ansatz wurden die einzelnen radicals frei variiert und abgestuft. Eine entsprechende Qualitätssicherungskomponente gab es jedoch nicht.

3.2 Cognitive Design System Approach

Dieser Ansatz bietet eine theoretische Fundierung, so baut er auf einem kognitiven Modell oder einem kognitiven Itemmodell auf. Daraus werden kognitive Prozesse, Wissensstrukturen und Bearbeitungsstrategien abgeleitet. Anschließend wird das kognitive Modell evaluiert und wenn notwendig erweitert. Hier wird auf klassisch konstruierte Tests zurückgegriffen sowie Item-Response-Theorie Modelle zu Hilfe genommen. Danach erfolgt die eigentliche semi- oder vollautomatische Itemgenerierung. Auch hier gibt es keine Qualitätssicherungskomponente.

3.3 Item-Modell-basierter Ansatz

Sind sich der Cognitive Design System Approach und der Item-Modell-basierte Ansatz sehr ähnlich (auch hier findet eine theoretische Fundierung statt), so unterscheiden sie sich doch in ihrer Zielsetzung. Geht es im Cognitive Design System Approach eher um eine Neukonstruktion von Items oder deren Rekonstruktion, so möchte der Item-Modell-basierte Ansatz die Erweiterung eines bestehenden Itempools durch Klone oder Isomorphe erreichen. Komponenten zur Qualitätssicherung wurden hier auch nicht realisiert.

3.4 Automatisierter Min-Max Ansatz

Dieser top-down Ansatz bezieht nicht nur theoretische Modelle mit ein, sondern auch die zugrunde liegende Definition des zu erfassenden Traits. Damit sollte bereits in der Konstruktionsphase durch die radicals eine höhere Konstruktrepräsentativität erreicht werden. So werden gleich vorweg Gedanken für die Qualitätssicherung verwendet. Die sogenannten *constraints* (Greeno,

Moore & Smith, 1993), also funktionale Einschränkungen, werden dadurch gebildet und diese sollen dazu beitragen dass nur die ausgewählten kognitiven Modelle verwendet werden. Diese Qualitätssicherungskomponente unterscheidet diesen Ansatz entscheidend von den restlichen.

3.5 Fragemodus

Ist es bei herkömmlichen Papier-Bleistift-Tests so, dass die Items nacheinander auf einem Blatt dargebracht werden und natürlich auch die Reihenfolge bei selbständiger Bearbeitung geändert werden kann/könnte, können bei Computerverfahren die Items einzeln, der Reihe nach, vorgegeben werden. So fällt der Fokus immer nur auf ein Item. Da das AID 2 keinen Papier-Bleistift-Test darstellt, werden auch hier die Items in einer bestimmten Reihenfolge vorgegeben. Bei Paper-Bleistift-Gruppentestungen könnten mit Hilfe der Item-Response-Theorie Paralleltests entworfen werden, die äquivalent sind.

3.6 Antwortmodus

Bei Papier-Bleistift-Tests gibt es die Möglichkeit des Einsatzes von Multiple-Choice oder die des freien Antwortformates. Bei Computerverfahren besteht neben diesen Möglichkeiten noch die sequentielle Vorgabe der Antworten, bei der die Testperson bei jeder Antwort entscheiden muss, ob diese zutrifft oder nicht. Außerdem kann der Eingabemodus variieren „from standard computer keyboards to modified keyboards, to custom-made consoles (including joysticks), to different forms of mouse, or to analogues of paper and pencil involving touchscreens.“ (Irvine, 2002, S.8).

III Empirischer Teil

4 Alltagswissen

„Der Untertest 1 Alltagswissen soll die Fähigkeit prüfen, sich Sachkenntnisse über Inhalte anzueignen, die in der heutigen Gesellschaft alltäglich sind.“ (Kubinger & Wurst, 2000, S. 16).

Hierbei werden Fragen gestellt, auf die frei geantwortet werden kann.

4.1 Untersuchung der Originalitems des AID 2

Um neue Items für die Gruppen- bzw. Computertesting konstruieren zu können, wurden zunächst die des AID 2 analysiert. Dafür wurde jede Frage des Untertests nach folgenden Kriterien aufgeschlüsselt (siehe Anhang):

- Block Nr.,
- Item Nr.,
- Anzahl der Wörter,
- Sätze,
- Einleitungssatz (ja/nein),
- Zusatzinformation,
- Rubrik,
- Nähe zum Kind (zu Hause, nahe Umgebung, ferne Umgebung, gelernt) – in Anlehnung an die Ökologische Systemtheorie von Urie Bronfenbrenner (1986) sowie
- Antworten richtig (eine oder mehrere)

4.2 Ökologische Systemtheorie von U. Bronfenbrenner – kurzer Überblick

Diese geht davon aus, dass sich ein Kind über verschiedene Beziehungen entwickelt. Diese Beziehungen, auf unterschiedlichen Ebenen bestehend, bilden ein komplexes System, das das Kind in seiner Entwicklung beeinflussen.

Bronfenbrenner unterscheidet:

- Mikrosystem – beschreibt die innerste Schicht der Umwelt und setzt sich aus den Menschen zusammen, mit denen das Kind unmittelbar zu tun hat. (z.B. Familie, Kindergarten/Schule, Nachbarschaft)
- Mesosystem – hier spielen die Wechselwirkungen der verschiedenen Mikrosysteme eine Rolle. So wird die Stimmung zu Hause unter Umständen von den Erfahrungen in der Schule beeinflusst.
- Exosystem – das sind soziale Rahmenbedingungen, die mit dem Kind nicht direkt in Verbindung stehen.

- Makrosystem – in diesem System sind die kulturellen Wertvorstellungen, Gesetze, Sitten und Gebräuche einer Kultur zu finden. (vgl. Berk, 2005)

Auf die Entwicklung eines Kindes abgeändert kann man sagen, dass Kinder zuerst in ihrer nahen Umgebung lernen (zunächst überhaupt den eigenen Körper kennenlernen) und davon ausgehend ihre Umgebung erkunden. Mit fortschreitendem Alter des Kindes werden sein Wissen und sein Erfahrungsschatz erweitert. Aus diesem Grund können jüngere Kinder noch nicht über das Wissen von älteren verfügen (verallgemeinert gesprochen). Bei der Erstellung der Fragen soll in weiterer Folge genau das berücksichtigt werden.

4.3 Weitere Untersuchungen

Ebenso wurden andere Verfahren herangezogen, die einen Untertest ähnlich dem „1 Alltagswissen“ aufweisen, um auf mögliche Konstruktionsregeln zu stoßen.

- HAWIK IV: Hamburg-Wechsler-Intelligenztest für Kinder – IV (Petermann & Petermann, 2008) – Allgemeines Wissen

Im Handbuch des HAWIK IV wurde erwähnt, dass es versucht wurde, geläufige Wörter zu verwenden mit der Beifügung dieses Kriterium nicht zu eng zu sehen, da bei schwierigeren Aufgaben die Wörter ungebräuchlicher werden.

- IST 2000-R: Intelligenz-Struktur-Test 2000 R (Amthauer, Brocke, Liepmann & Beauducel, 2001) – Wissenstest (WI)

Das Manual des IST 2000-R listet zwar die Themengebiete und Kodierungen (verbal, numerisch, figural) auf, in Folge steht jedoch nur: „Für jede Kodierungsart (verbal, numerisch, figural) wurden mindestens vier und maximal fünf Fragen zu jedem Themenbereich konstruiert.“ (S. 49)

- K-ABC: Kaufman Assessment Battery for Children (Melchers & Preuß, 2009) - (14) Rätsel

Die K-ABC verweist in diesem Zusammenhang lediglich auf den „Conceptual Inference Test“ von Kagan und Klein (1973). Für den Untertest werden in der K-

ABC zwar neue Aufgaben verwendet, im Stil und Umfang jedoch wurde dem von Kagan und Klein benutzten Vorgehen entsprochen.

Wie sich herausstellte, wurde in keinem der Manuale erwähnt, wie die entsprechenden Fragen konstruiert wurden.

Lediglich das Handbuch des AID 2 enthält diesbezüglich Informationen über die Herangehensweise. Sie endete mit der Einschätzung durch Experten, weist aber auch keine Regelleitung auf. Besonders interessierte Leser seien an dieser Stelle an Kubinger und Wurst (2000, S.15) verwiesen.

4.4 Gestaltung neuer Items

Um die Konstruktion der Items mit mehr wissenschaftlichem Hintergrund zu gestalten, wurde nun versucht, Regeln für deren Erstellung herzuleiten. Dies stellte sich als relativ schwierig heraus. Die Grundregel jeder Frage war natürlich, dass kein Schulwissen abgefragt werden sollte. Es stellte sich jedoch die Frage, ob das Alltagswissen per Definition sich wirklich so einfach vom Schulwissen abgrenzen lässt. Schließlich wird z.B. in der Schule gelehrt, dass der menschliche Körper ein Herz besitzt, das für den Transport von Sauerstoff über das Blut verantwortlich ist. Diese Information würden viele Menschen sicherlich auch als alltäglich charakterisieren. Demnach wurde versucht Fragen zu finden, deren Inhalte zwar unter Umständen im Unterricht erworben werden können, jedoch keineswegs reines Schulwissen sind. Es wurde also nach Wissen gesucht, das im Alltag angewendet werden kann, unabhängig von der Schule erworben wird oder zumindest von vielen Menschen als selbstverständlich angesehen wird.

Nach Lienert und Raatz (1994) sollte man versuchen keine Begriffe zu verwenden die mehrere Bedeutungen haben. Außerdem sollten keine Begriffe verwendet werden, die unter Umständen nur einem Teil der Testpersonen bekannt sind.

Anhand der Aufschlüsselung der Items wurde beschlossen, nicht mehr als 2 Sätze pro Frage zu verwenden. Es wurden verschiedene Themenbereiche ausgewählt, die die incidentals bildeten (Irvine, 2002), und mit Hilfe von PsychologiestudentInnen altersentsprechend und unter der Zuhilfenahme von der ökologischen Systemtheorie als radical (Irvine, 2002) dazu Fragen erstellt. Als

Entscheidung, ob die Frage Alltagswissen verlangt, wurde überprüft, ob die Information auf Wikipedia (www.wikipedia.de) zu finden ist. Unter der Annahme, dass diese Seite von vielen Menschen geschaffen und erweitert wird, diente sie in dieser Diplomarbeit als Referenz und als quasi Qualitätssicherung.

Bei der Aufgabengestaltung wurde ein freies Antwortformat gewählt, um dem AID 2 bis zu einem gewissen Grad äquivalent zu sein. Es wurde darauf geachtet, dass die Fragen mit einem einzigen Wort zu beantworten sind. Die Testperson wird demnach nicht dazu verleitet, unnötig viel zu schreiben und sie muss sich ebenso keine Gedanken über eine entsprechende Formulierung machen.

Rechtschreibfehler wurden dabei nicht beachtet, bei einer möglichen Computertesting besteht die Option der Überprüfung durch einen Rechtschreibassistenten oder es werden mögliche Schreibformen eingetragen

Die Instruktion erfolgte im Rahmen einer Gruppentesting schriftlich. Die Fragen dieses Untertests wurden in vier Testformen aufgeteilt und je zusammen auf einem Blatt vorgegeben. Nachdem das selbständige Lesen, das Verständnis der Instruktion und das flüssige Schreiben-können von wesentlicher Bedeutung war (Bei jungen Kindern, die das Schreiben und Lesen erst erlernen ist dies noch nicht der Fall), wurde der ursprüngliche Altersbereich des AID 2 (6;0 – 15;11 Jahre) auf 8;0 bis 15;11 Jahre reduziert. Weiters wurde das Kriterium der Muttersprache „Deutsch“ bzw. das einwandfreie Beherrschen der Sprache in Wort und Schrift aufgestellt.

5 Realitätssicherheit

Vorweg ist gleich an dieser Stelle mitzuteilen, dass dieser Untertest bei der Erhebung der Daten aus Zeitgründen nicht miteinbezogen wurde, da die Testzeit von einer Schulstunde nicht überschritten werden durfte.

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

Hier werden der Testperson Bilder gezeigt, auf denen jeweils etwas Wichtiges fehlt. Dabei ist es wesentlich darauf hinzuweisen, dass es sich dabei um funktionale Lücken handelt.

5.1 Untersuchung der Originalitems des AID 2

Dieser Untertest besteht aus 12 Items exklusive einer warming-up-Aufgabe für alle Altersgruppen. Die Bilder sind stilisiert dargestellt, Unwichtiges (z.B. die Umgebung des Gegenstandes) wurde dabei weggelassen. Ebenso wurden keine Symmetrien verwendet. Thematisch konnte kein Hinweis auf die Schwierigkeit der Aufgaben ausgemacht werden. Bei der Analyse der fehlenden Details konnte ebenso keine Regel herausgefunden werden.

5.2 Gestaltung neuer Items

Bei der Suche nach neuen Aufgaben wurden zunächst Ideen jeglicher Art aus der realen Welt gesammelt. Wichtig dabei war, eine Darstellung zu finden, bei der etwas Funktionales fehlen kann. Anschließend wurde überlegt, ob die Bilder zu verwirklichen wären. Auch hinsichtlich der Schwierigkeit der Aufgaben wurden Überlegungen angestellt. Wie schon im AID 2 wurden hier ebenso Symmetrien vermieden (So hat es z.B. keinen Sinn, einen 4-eckigen Tisch mit drei Beinen vorzugeben.).

Um dem Namen des Untertests mehr gerecht zu werden, wurde zunächst versucht, ein mit der Kamera aufgenommenes Bild zu erhalten, das das fehlende Stück nicht enthält oder dieses Fotos mit einem Bildbearbeitungsprogramm zu verändern. Beispiele hierfür sind in Abbildung 1 und Abbildung 2 zu sehen.



Abbildung 1: Schuh ohne Schnürsenkel



Abbildung 2: Violine ohne F-Löcher

Diese Art der Bilder wurde jedoch verworfen, da zu hohe Fertigkeiten im entsprechenden Grafikprogramm notwendig gewesen wären sowie zu viele Details bei einigen der Items enthalten gewesen wären und dadurch die Übersicht verloren gegangen wäre.

Somit wurde der ursprüngliche Fragemodus der Items beibehalten – die Bilder wurden gezeichnet und sind im Anhang zu finden.

Die Frage des Antwortmodus ist hier nicht so einfach zu beantworten. Im AID 2 haben die Testpersonen die Möglichkeit verbal zu antworten oder an den Platz zu zeigen, an dem das fehlende Stück hingehören würde. Würde man hier nun die Möglichkeit geben die Lösung hinzuschreiben, fiel das Hinzeigen weg und der Untertest wäre nicht mehr sprach- bzw. in diesem Fall schreibfrei. Bei einer Papier-Bleistift-Testung könnte man der Testperson die Lösung zeichnen lassen, was jedoch eine gewisse Feinmotorik und Übung voraussetzt. Beim Einkreisen

des Areal in dem das Fehlende Teil zu finden ist, wäre der Kreis in manchen Fällen so groß zu ziehen, sodass fast das gesamte Bild eingeschlossen wäre (siehe z.B. beim Turnschuh in Abbildung 1). Somit könnte nicht gesagt werden, ob die Testperson die Lösung weiß oder einfach nur das Bild einkreist. Bei einer Computertestung bestehen natürlich ebenso die oben angeführten Möglichkeiten und Kritikpunkte. Zusätzlich könnte einfach auf den entsprechenden Bereich des Bildes geklickt werden, was jedoch ebenso den Rateeffekt bei den zurzeit existierenden Items nicht ausschließen würde. Diese müssten dann unter Umständen detailreicher gestaltet werden. Für die Bearbeitung der Items müsste in diesem Fall mehr Zeit anberaumt werden.

6 Angewandtes Rechnen

„Der Untertest 3 Angewandtes Rechnen soll weitgehend unabhängig von schulischen Rechenfertigkeiten prüfen, inwieweit die Testperson bei der Problemlösung alltäglicher Aufgabenstellungen durch entsprechende Schlußfolgerungen die passenden Rechenoperationen anzuwenden imstande ist.“ (Kubinger & Wurst, 2000, S. 16)

Die Testperson erhält Textaufgaben (die sie ab einer bestimmten Schwierigkeitsstufe auch mitlesen kann) die sie zu lösen hat. Zusätzlich können Nebenrechnungen auf einem Notizblatt gemacht werden.

6.1 Untersuchung der Originalitems des AID 2

Wie auch beim Untertest „Alltagswissen“ wurden auch hier zunächst die Items des AID 2 analysiert und aufgeschlüsselt:

- Block-Nummer
- Item-Nummer
- Anzahl der Sätze
- davon Hauptsätze und Nebensätze
- Vorkommende Namen (und deren Silben)
- Vokabeln in mathematische Metaregeln gepackt - Z.B. „gibt etwas her“ = „-“, oder „gleichmäßig aufgeteilt“ = „:“
- Anzahl der Operationen

- Richtung (der Rechenoperationen) – Hat das Kind die Möglichkeit beim Vorlesen gleich mitzurechnen oder erfährt es erst am Ende der Aufgabe die Ausgangszahl (Präsentationsreihenfolge)?
- Art der Operationen – Addition, Subtraktion, Multiplikation, Division
- Inhaltliche Komponenten – In welches Ereignis ist die Aufgabe integriert?
- Höchste angegebene Zahl
- Zusatzwissen – Z.B. Wissen um die Zeit, Umfangberechnung, etc. erforderlich

Im Anhang finden Sie die entsprechenden Daten, jedoch ohne „Anzahl der Operationen“ und „Vokabeln in mathematische Metaregeln gepackt“, da diese Punkte zu viel Information über die Items anzeigen würden.

6.2 Weitere Untersuchungen

Um auch hier Regeln für eine Itemerstellung ausfindig zu machen wurden andere Verfahren mit ähnlichen Untertests dazu herangezogen:

- WIE: Wechsler Intelligenztest für Erwachsene (Aster, Neubauer, & Horn, 2006). - 6) Rechnerisches Denken

Hier konnte kein Hinweis auf eine entsprechende Itemgenerierung gefunden werden.

- HAWIK IV: Hamburg-Wechsler-Intelligenztest für Kinder – IV (Petermann & Petermann, 2008) - (3) Rechnerisches Denken

Der einzige Hinweis, der hier gegeben wurde war, dass die Anweisungen möglich kurz formuliert wurden.

- INSBAT: Intelligenz-Struktur-Batterie (Hornke et al., 2004) – (AD) Algebraisches Denken

Der Untertest des INSBAT wurde mit Hilfe eines prüfenden (durch incidentals (Irvine, 2002) und generativen (durch radicals (Irvine, 2002)) Regelwerks

konstruiert (AGen: Arendasy & Sommer, 2004). Dieses kann im Manual des INSBAT (Hornke et al., 2004, S. 57) betrachtet werden.

Die Gestaltung neuer Items für das AID 2 wurde daran angelehnt.

6.3 Gestaltung neuer Items

Rollett (1997) postuliert, dass ohne „Kenntnis der Zahlenreihe und ihrer Eigenschaften ... Rechenoperationen nicht durchführbar [sind].“ (S. 55). Weiters beschreibt sie, dass höhere Malreihen für Kinder (in weiterer Folge auch für Erwachsene) oft schwieriger sind, weil im Unterricht auf das Vertauschen von Multiplikator und Multiplikand hingewiesen wird. So wird aus 3×8 auf einmal 8×3 . Sind aber beide Ziffern höher, so fällt die Multiplikation auch durch Tauschen der Faktoren nicht leichter (z.B. 8×7).

Greeno (1991) beschreibt das Erlernen der Struktur von Zahlen. „They [children] ... learn that numbers themselves have magnitudes and that the sequence of numbers is ordered by increasing magnitude-a number later in the sequence is greater than a number that is earlier.“ (S. 188). Er postuliert ebenso, dass Addition und Subtraktion früher gelernt werden als Multiplikation und Division.

Unter Berücksichtigung dieser Theorien und nach der Analyse der Items des AID 2 wurden diese der Übersicht halber in vier Schwierigkeitsstufen gegliedert. Dies vereinfachte die Erstellung neuer Textaufgaben. Als Inhalte dienten mögliche alltägliche Gegebenheiten bzw. dem jeweiligen Alter Verständliches. Weiters sollten die Items innerhalb einer gewissen Zeit lösbar sein (Stufe 1 und 2: 30 Sekunden, Stufe 3 und 4: 60 Sekunden). Kopfrechnen sollte überwiegend möglich sein, obwohl den Testpersonen Hilfsmittel (Papier und Stift) zur Verfügung stehen. Diese Bedingungen bilden mitunter die incidentals (Irvine, 2002).

Die weiteren incidentals und radicals (Irvine, 2002) der verschiedenen Stufen werden unten aufgelistet:

- Stufe 1 - Alter: 6-7 Jahre

Anzahl der Sätze: Bis zu 4 Hauptsätze möglich, keine Gliedsätze (eingeleitet z.B. durch: nachdem, obwohl, weil, dass, bevor, ob, damit, während, als...), generell sollen keine Beistriche verwendet werden (auch nicht zum Aufzählen).

Anzahl der Operationen: Überwiegend 1, maximal 2, alle Operationsarten möglich; Divisionen: ohne Rest.

Zahlenraum: bis 11.

Es soll kein Zusatzwissen erforderlich sein (Zeit, Umfang, Längen,...).

Überwiegend sollen Beispiele erstellt werden, bei denen schon beim Lesen mitgerechnet werden kann, die Rechenreihenfolge also der Lesereihenfolge entspricht.

- Stufe 2 - Alter: 8-9 Jahre

Anzahl der Sätze: 3-4 Hauptsätze, keine Gliedsätze wenn möglich, Beistriche zum Aufzählen möglich.

Anzahl der Operationen: Maximal 2, alle Operationsarten möglich; Divisionen: ohne Rest.

Zahlenraum: Großteils bis 30, bei höheren Werten sollten alle außer der 1. Stelle eine 0 sein.

Zusatzwissen: mit Zeit rechnen möglich (Stunden, Tage).

Die Rechenreihenfolge sollte überwiegend der Lesereihenfolge entsprechen.

- Stufe 3 - Alter: 10-11 Jahre

Anzahl der Sätze: bis zu 5 Hauptsätzen, bis zu 2 Nebensätze möglich, wenn eine geringe HS-Anzahl besteht.

Anzahl der Operationen: bis zu 4, alle Operationen möglich; bei Divisionen: Rest möglich.

Zahlenraum: Großteils bis 30, aber bis 70 möglich, bei höheren Werten sollten alle außer der 1. Stelle eine 0 sein.

Zusatzwissen: Zeit (Stunden, Tage, Monate), Umfang.

Die Reihenfolge kann beliebig sein.

- Stufe 4 - Alter: 12-15 Jahre

Anzahl der Sätze: bis zu 4 Hauptsätze, bis zu 2 Nebensätze möglich.

Zahlenraum: bis 50, darüber sind runde Zahlen zu verwenden, alle Operationen sind möglich; bei Divisionen: Rest möglich, bei höheren Werten sollten alle außer der 1. Stelle eine 0 sein. Kommazahlen vereinzelt verwendbar.

Zusatzwissen: Schlussrechnen, Zeit, Brüche, Flächen, Prozentrechnen, Algebra, Längenangaben.

Die Reihenfolge kann beliebig sein.

Die Anweisung für die Erstellung neuer Items für diesen Untertest finden Sie im Anhang.

Nach dieser Anweisung wurden im Rahmen einer Übung zur psychologischen Diagnostik von Psychologie-StudentInnen einige Textbeispiele erstellt. Diese wurden anschließend überprüft und gegebenenfalls verworfen oder verändert.

Auch hier wurde, wie beim Untertest Alltagswissen, ein freies Antwortformat gewählt. Der Antwortmodus wurde dabei so gewählt, dass kein Antwortsatz geschrieben werden musste, das Hinschreiben der Zahl war ausreichend. Bei einer möglichen Computertesting würde die Eingabe entweder über die Tastatur in ein Textfeld oder mit Hilfe eines Nummernfeldes am Bildschirm, auf dem die Ziffern mit der Maus angeklickt werden können, erfolgen. Bei Verwendung eines Touch-Screens könnte die Eingabe mit dem Finger direkt am Bildschirm erfolgen oder auch durch einen Eingabestift.

Die Instruktion erfolgte im Rahmen einer Gruppentesting ebenso schriftlich. Weitere Kriterien siehe Gestaltung neuer Items beim Alltagswissen.

7 Durchführung der Untersuchung

7.1 Wahl der Analysestichprobe

Die Testungen fanden in einer dritten und vierten Klasse Volksschule und zwei Hauptschulen aus dem Mühlviertel sowie einem Privatgymnasium aus Linz statt. Im Gymnasium wurde von der fünften bis neunten Schulstufe getestet. Die Schulstufen entsprachen dem Geltungsbereich des Alters von 8;00 bis 15;11 Jahren.

Die Stichprobe war dabei selektiv, da nur in jenen Schulen getestet wurde, in denen die Direktoren und der oberösterreichische Landesschulrat ihr Einverständnis gaben (Das Schreiben an den Landesschulrat und der Elternbrief befinden sich im Anhang).

Auch wenn damit der Geltungsbereich nicht im höchsten Maße repräsentativ ist, ist kein augenscheinlicher Unterschied festzustellen.

7.2 Testvorgabe

Zur Bearbeitung der Untertests stand den Kindern eine Unterrichtsstunde (50 Minuten) zur Verfügung. Neben den Untertests „Alltagswissen“ und „Angewandtes Rechnen“ erhielten die Testpersonen ebenso die Untertests „Wortschatz“ und „Funktionen abstrahieren“, die im Zuge einer anderen Diplomarbeit entwickelt wurden (Böck, 2010).

Mit der Aufteilung des Testmaterials in vier, optimaler Weise gleich schwere Testformen wurden die Aufgaben in einer Unterrichtseinheit gut bewältigt. Dies zeigte eine Voruntersuchung mit einem 8-, einem 10- und einem 13-jährigen Kind. Ein weiterer Grund für vier Testformen war auch der, dass damit das Abschreiben verhindert oder zumindest reduziert werden sollte.

Die Testung erfolgte in der Gruppe (Klasse) im jeweiligen Klassenzimmer. Anwesend waren ein oder zwei Testleiterinnen, manchmal die entsprechenden LehrerInnen. Zu Beginn wurden die SchülerInnen informiert, dass ihnen in dieser Stunde eine Reihe von Aufgaben vorgelegt werden würden, mit Hilfe derer herausgefunden werden soll, was sie schon alles können bzw. wissen. Außerdem wurde den Testpersonen mitgeteilt, dass sie sowohl leichte als auch schwerere

Aufgaben erhalten werden, da diese Fragen sowohl an jüngere als auch ältere Kinder gerichtet sind. Die Anonymität wurde versichert, ebenso wurde versichert, dass das Ergebnis keinen Einfluss auf eine Note haben wird. Sie wurden angewiesen alleine für sich zu arbeiten und es wurde ihnen mitgeteilt, dass Blicke zum Nachbarn aufgrund der verschiedenen Gruppen sinnlos seien. Danach wurde der erste Untertest „Funktionen abstrahieren“ ausgeteilt. Die Instruktion wurde leise gelesen, anschließend wurde das Übungsbeispiel mündlich gelöst. Sofern die Aufgabenstellung verstanden wurde, wurden die SchülerInnen gebeten, die eigentlichen Aufgaben zu bearbeiten und ein kurzes Handzeichen zu geben, sobald sie meinten fertig zu sein. War der Großteil der Klasse fertig, spätestens jedoch nach acht Minuten, wurden die SchülerInnen gebeten, den Testbogen beiseite zu legen. Der nächste Bogen „Angewandtes Rechnen“ wurde nun ausgeteilt und es wurde angemerkt, dass die nicht fertiggestellten Bögen, wenn Zeit bliebe, noch bearbeiten werden durften. Es wurde stets darauf geachtet, dass dieselbe Testform nicht nebeneinanderliegend ausgeteilt wurde. Auch hier wurden die SchülerInnen gebeten, die Instruktion leise zu lesen. Sie wurden darauf hingewiesen, dass kein Antwortsatz notwendig sei und die jeweilige Zahl genüge. Nach spätestens weiteren zehn Minuten wurde der Untertest „Wortschatz“ ausgeteilt. Als Abschluss, nach spätestens fünfzehn Minuten, bearbeiteten die SchülerInnen den Untertest „Alltagswissen“. Wurden alle Untertests fertig bearbeitet oder war die Unterrichtsstunde schon fast zu Ende, wurden alle Testbögen eingesammelt. Als kleines Dankeschön erhielten alle Kinder eine Süßigkeit, von der sie jedoch im Vorhinein nichts wussten.

7.3 Beobachtungen zur Testung

Nach dem oben beschriebenen Vorgabeplan verlief die Testung in allen Klassen reibungslos und ohne gravierende Vorfälle. Die Instruktion wurde durchgehend verstanden und die Übungsbeispiele konnten in sämtlichen Schulstufen gelöst werden. Die zur Verfügung stehende Zeit von 50 Minuten war für fast alle Kinder ausreichend.

Besonders wichtig erschien es einigen Kindern, dass ihren Lehrern keine Ergebnisse gezeigt werden und sie anonym bleiben.

Die Süßigkeit, die nach der Testung als Dankeschön ausgeteilt wurde, wurde von den Kindern freudig angenommen.

7.4 Verrechnung

Sowohl der Untertest „Alltagswissen“ als auch „Angewandtes Rechnen“ wurden dichotom verrechnet, die Antwort war entweder richtig oder falsch. Teilpunkte wären hier nicht sinnvoll gewesen, darum wurden auch keine vergeben. Die Rechtschreibung bei „Alltagswissen“ spielte keine Rolle. Auch Beantwortungen dem Wortlaut nach wurden gegebenenfalls als richtig gewertet. So fanden sich z.B. sehr unterschiedliche Schreibweisen für New York, Hiroshima oder Gandhi.

7.5 Aufgabenanalyse

Die Testformen wurden auf ihre Konformität mit dem dichotom logistischen Modell nach Rasch überprüft. Dies ist ein Ansatz aus der probabilistischen Testtheorie und analysiert die Testformen dahingehend, ob sie eindimensional messen. Der Testwert hätte dann Rationalskalenniveau.

Wer sich mit diesem Modell intensiv auseinandersetzen möchte, wird auf Fischer (1974) oder Kubinger (1989) verwiesen.

Die Konformität der Testformen mit dem Rasch-Modell wurde mit dem Likelihood-Quotienten-Test nach Andersen (1973) überprüft, der die Eigenschaft aufweist „stichprobenunabhängig“ zu sein. Die Stichprobe spielt keine Rolle bei der statistischen Inferenz der Parameter (Kubinger, 1989).

War eine Testform nicht Rasch-Modell-konform wurde sie anschließend itemweise überprüft. Dies erfolgte durch den z-Test nach Fischer und Scheiblechner, der Itemfitstatistik, einem graphischem Modelltest sowie einer qualitativen Analyse der Items.

Mit Hilfe dieser Überprüfung wurde entschieden, welches Item oder welche Items aus dem Pool zu entfernen sind um auf eine erneute Konformität zu testen. Die Ergebnisse finden sich im Unterpunkt „Ergebnisse“.

Die Auswertungen wurden mit Hilfe des Programmes R-2.12.2 für Windows sowie mit PASW Statistics 18, Version 18.0.0 durchgeführt.

8 Deskriptive Statistik

Die Verteilung der Daten in Bezug auf die Variablen „Geschlecht“, „Region“, „Schultyp“ und „Schulstufe“ wurde einerseits für die gesamten Untertests und zum anderen für die einzelnen Testformen überprüft.

Da die Untersuchung gemeinsam mit Böck (2010) durchgeführt wurde, und die Daten verglichen wurden, stellte sich heraus, dass in beiden Untertests Testbögen fehlten. Dadurch kam es zu einer Diskrepanz der Verteilung „Alltagswissen“ und „Angewandtes Rechnen“, in den Daten wurde diese als fehlende Werte kodiert. Natürlich wurde versucht, eine gleichmäßige Verteilung der Testformen zu erreichen.

8.1 Geschlecht

In den Tabelle 1 und 2 ist ersichtlich, dass die angestrebte gleichmäßige Verteilung der Variable Geschlecht sowohl in den Untertests, als auch in den verschiedenen Testformen annähernd eingetroffen ist (52,1% Buben, 47,9% Mädchen).

Geschlecht	Häufigkeit		Prozent	
	Alltags- wissen	Angewandtes Rechnen	Alltags- wissen	Angewandtes Rechnen
männlich	347	354	51,1	52,1
weiblich	325	322	47,9	47,4
fehlend	7	3	1,0	0,4
gesamt	679	679	100,0	100,0
gesamtgültig	672	676	99,0	99,6

Tabelle 1: Verteilung Geschlecht pro Untertest

Geschlecht	Alltagswissen				Angewandtes Rechnen			
	A	B	C	D	A	B	C	D
männlich	87	92	92	76	91	91	89	83
weiblich	86	79	75	85	83	80	78	81
fehlend	0	0	0	7	0	0	3	0
gesamt	173	171	167	168	174	171	170	164
gesamtgültig	173	171	167	161	174	171	167	164

Tabelle 2: Verteilung Geschlecht pro Testform

8.2 Region

Obwohl nur annähernd, kann man bei Werten von 47,4% Landkinder und 52,1% Stadtkinder (siehe Tabelle 3) doch auch hier von einer Gleichverteilung sprechen.

Die größte Differenz war zwischen Testform A und D beim „Alltagswissen“ zu finden und betrug 12 SchülerInnen (Tabelle 4).

Region	Häufigkeit		Prozent	
	Alltags- wissen	Angewandtes Rechnen	Alltags- wissen	Angewandtes Rechnen
Land	318	322	46,8	47,4
Stadt	354	354	52,1	52,1
fehlend	7	3	1,0	0,4
gesamt	679	679	100,0	100,0
gesamtgültig	672	676	99,0	99,6

Tabelle 3: Verteilung Region pro Untertest

Region	Alltagswissen				Angewandtes Rechnen			
	A	B	C	D	A	B	C	D
Land	83	79	79	77	83	82	79	78
Stadt	90	92	88	84	91	89	88	86
fehlend	0	0	0	7	0	0	3	0
gesamt	173	171	167	168	174	171	170	164
gesamtgültig	173	171	167	161	174	171	167	164

Tabelle 4: Verteilung Region pro Testform

8.3 Schultyp

Da in ländlichen Regionen die Volks- und Hauptschulen getestet wurden und in der Stadt das Gymnasium, sind sich die Variablen „Schultyp“ und „Region“ ähnlich. Es wurden insgesamt mehr GymnasiastInnen als HauptschülerInnen getestet (ersichtlich in Tabelle 5), verhältnismäßig wenige VolksschülerInnen. Doch auch hier kann man erkennen, dass die einzelnen Testformen selbst in der Volksschule gut verteilt wurden (Tabelle 6).

Schultyp	Häufigkeit		Prozent	
	Alltags- wissen	Angewandtes Rechnen	Alltags- wissen	Angewandtes Rechnen
Volksschule	34	34	5,0	5,0
Hauptschule	284	288	41,8	42,4
Gymnasium	354	354	52,1	52,1
fehlend	7	3	1,0	0,4
gesamt	679	679	100,0	100,0
gesamtgültig	672	676	99,0	99,6

Tabelle 5: Verteilung Schultyp pro Untertest

Schultyp	Alltagswissen				Angewandtes Rechnen			
	A	B	C	D	A	B	C	D
Volksschule	10	7	9	8	9	8	8	9
Hauptschule	73	72	70	69	74	74	71	69
Gymnasium	90	92	88	84	91	89	88	86
fehlend	0	0	0	7	0	0	3	0
gesamt	173	171	167	168	174	171	170	164
gesamtgültig	173	171	167	161	174	171	167	164

Tabelle 6: Verteilung Schultyp pro Testform

8.4 Schulstufe

Aus den beiden Tabellen die Schulstufe betreffend (Tabelle 7 und Tabelle 8) ist ersichtlich, dass zwar von der 5. Bis 8. fast gleich viele Kinder getestet wurden, in der 3. und 4. (den Volksschulklassen) sowie in der 9. die Anzahl jedoch wesentlich geringer ausfällt. (Diese Schulstufen wurden deshalb nicht in die Mittelwertvergleiche miteinbezogen.) Trotzdem genügt diese kleine Anzahl aber um festzustellen, dass die Vorgabe der Untertests „Alltagswissen“ und „Angewandtes Rechnen“ auch in diesem Altersbereich möglich und zumutbar ist.

Schulstufe	Häufigkeit		Prozent	
	Alltags- wissen	Angewandtes Rechnen	Alltags- wissen	Angewandtes Rechnen
3.	13	12	1,9	1,8
4.	21	22	3,1	3,2
5.	140	139	20,6	20,5
6.	162	162	23,9	23,9
7.	146	150	21,5	22,1
8.	139	139	20,5	20,5
9.	51	52	7,5	7,7
fehlend	7	3	1,0	0,4
gesamt	679	679	100,0	100,0
gesamtgültig	672	676	99,0	99,6

Tabelle 7: Verteilung Schulstufe pro Untertest

Schulstufe	Alltagswissen				Angewandtes Rechnen			
	A	B	C	D	A	B	C	D
3.	4	3	3	3	3	3	3	3
4.	6	4	6	5	6	5	5	6
5.	34	38	35	33	35	35	36	33
6.	40	40	42	40	40	40	41	41
7.	40	37	32	37	41	38	35	36
8.	36	36	36	31	36	37	34	32
9.	13	13	13	12	13	13	13	13
fehlend	0	0	0	7	0	0	3	0
gesamt	173	171	167	168	174	171	170	164
gesamtgültig	173	171	167	161	174	171	167	164

Tabelle 8: Verteilung Schulstufe pro Testform

9 Ergebnisse

Im Folgenden wurden die Testformen wie im Punkt 7.5 beschrieben ausgewertet. Die Item- und Personenparameter für die Rasch-Modell-konformen Testformen sind im Anhang zu finden.

Für sämtliche Testformen lautet die H_0 „Die Daten verhalten sich dem dichotom logistischen Modell von Rasch konform.“

Der Fehler 1. Art wurde adjustiert, da jeder Datensatz durch drei abhängige Tests, den Likelihood-Quotienten-Tests (LQTs), erfolgte. Somit wurde das α von 0,05 auf 0,013 herabgesetzt.

Die Gültigkeit aller acht Testformen wurde anhand folgender Teilungskriterien überprüft:

Internes Teilungskriterium:

- Höhe des Rohwertes – Testpersonen mit einem Rohwert niedriger oder gleich dem Median werden Personen mit einem Rohwert höher dem Median gegenübergestellt.

Externe Teilungskriterien:

- Geschlecht – Männliche Testpersonen werden weiblichen gegenübergestellt.
- Region – Testpersonen von der Stadt werden denen vom Land gegenübergestellt.

9.1 Alltagswissen, Testform A

9.1.1 Erste Parameterschätzung

Die erste Parameterschätzung des Untertests 1 „Alltagswissen“, Testform A erfolgte ohne Item 13, da sich herausstellte, dass mehrere Antworten richtig sind. So verbraucht die Flamme einer Kerze außer dem Wachs nicht nur Sauerstoff, sondern auch den Docht. Da man nicht mit Sicherheit sagen kann, dass die Antworten „Sauerstoff“ und „Docht“ demselben Niveau entspringen, wurde das Item ausgeschlossen.

Für den Itempool musste die H0 „Das Rasch-Modell gilt“ verworfen werden, da im Teilungskriterium „Geschlecht“ auf dem 1% Niveau ein signifikantes Ergebnis auftrat. Der Likelihood-Quotienten-Test nach Andersen umfasste beim Teilungskriterium Geschlecht und Region 14, beim internen nur 13 Items, da das Item 9 wegen ungünstiger Antwortmuster in den Subgruppen von der Berechnung ausgeschlossen wurde. Eine Übersicht der Ergebnisse ist in Tabelle 9 ersichtlich.

Andersen LQT	Intern	Kritischer Wert df=12	Geschlecht	Region	Kritischer Wert df=13
LQ-Wert	8,963	25,41	80,24	23,061	26,86
χ^2 df	12	12	13	13	13
p-Wert	0,706	0,013	0	0,041	0,013

Tabelle 9: Erste Parameterschätzung, Alltagswissen A, nicht Rasch-Modell-konform

9.1.2 Itemselektion

Z-Test

Beim Z-Test (Tabelle 10) mit dem Teilungskriterium „Geschlecht“ wurden die Items 1 (0,000), 3 (0,036), 5 (0,017), 6 (0,005), 7 (0,028) und 10 (0,019) signifikant. Die Items 5 (0,018), 14 (0,011) und 15 (0,040) unterschritten beim Teilungskriterium „Region“ den Wert. Item 9, welches beim Internen Teilungskriterium nicht geschätzt werden konnte, wurde sowohl beim Teilungskriterium „Geschlecht“, als auch bei „Region“ als Rasch-Modell-konform erachtet.

	z-Wert Intern	p-Wert Intern	z-Wert Geschl.	p-Wert Geschl.	z-Wert Region	p-Wert Region
Item1	1,176	0,240	-6,008	0,000	0,389	0,697
Item2	1,370	0,171	0,841	0,401	-0,461	0,645
Item3	0,068	0,946	2,092	0,036	0,619	0,536
Item4	-0,351	0,725	1,337	0,181	1,572	0,116
Item5	0,650	0,516	2,390	0,017	2,365	0,018
Item6	-0,736	0,462	2,790	0,005	0,953	0,341
Item7	-0,594	0,553	2,963	0,003	0,082	0,935
Item8	-1,246	0,213	2,200	0,028	0,656	0,512
Item9			0,130	0,896	0,962	0,336
Item10	0,889	0,374	2,350	0,019	1,285	0,199
Item11	0,381	0,703	0,716	0,474	-0,798	0,425
Item12	1,219	0,223	-0,552	0,581	0,835	0,404
Item14	0,409	0,682	0,289	0,773	2,558	0,011
Item15	0,628	0,530	2,058	0,040	1,740	0,082

Tabelle 10: Z-Test, Alltagswissen A, Teilungskriterien Intern, Geschlecht und Region

Itemfitstatistik

In der Itemfitstatistik wiesen die Items 1 und 3 signifikante Werte auf (Tabelle 11).

Item	χ^2	df	p-Wert
Item1	211,341	169	0,015
Item2	95,565	169	1,000
Item3	263,923	169	0,000
Item4	152,295	169	0,817
Item5	190,160	169	0,127
Item6	136,443	169	0,969
Item7	120,239	169	0,998
Item8	120,920	169	0,998
Item9	43,769	169	1,000
Item10	163,111	169	0,613
Item11	141,908	169	0,936
Item12	186,202	169	0,173
Item14	155,847	169	0,758
Item15	171,839	169	0,425

Tabelle 11: Itemfitstatistik, Alltagswissen A

Grafische Modelltests

Der grafische Modelltest mit dem Teilungskriterium „Höhe des Rohwertes“ (Abbildung 3) zeigte, dass sämtliche Konfidenz-Ellipsen die Gerade schneiden.

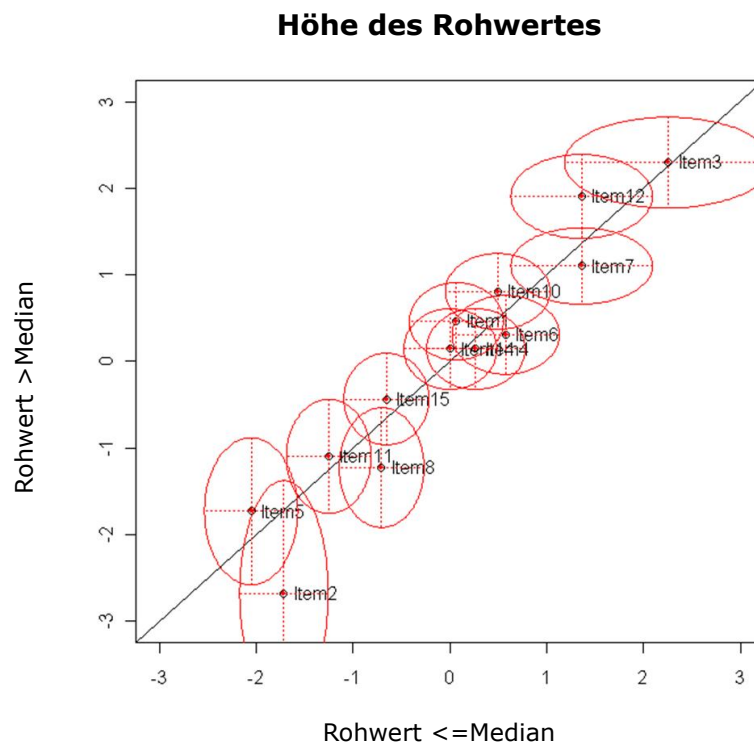


Abbildung 3: Grafischer Modelltest, Alltagswissen A, TK: Höhe des Rohwertes

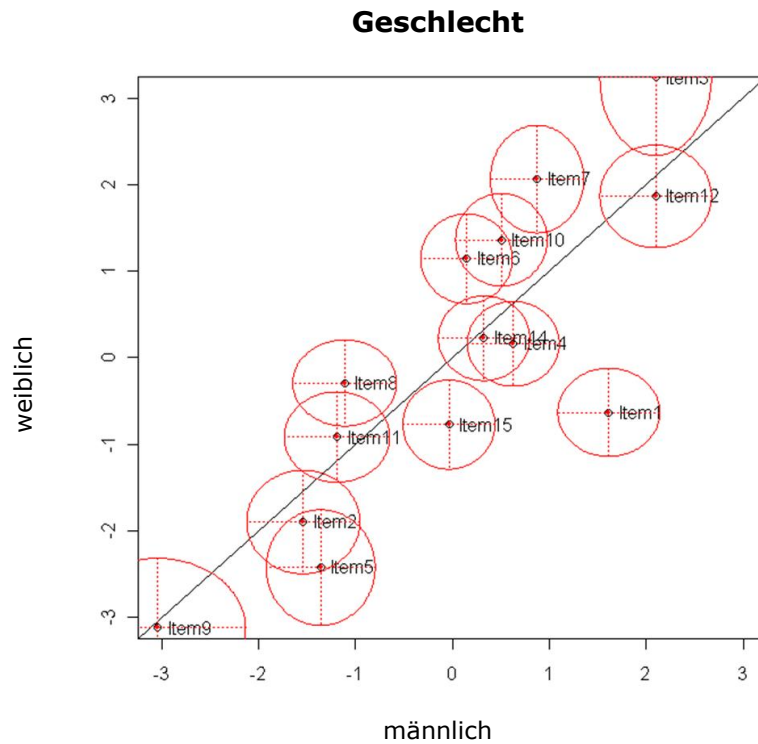


Abbildung 4: Grafischer Modelltest, Alltagswissen A, TK: Geschlecht

Im grafischen Modelltest mit dem Teilungskriterium „Geschlecht“ (Abbildung 4) schneiden die Konfidenz-Ellipsen der Items 1,3, 5, 6, 7, 8 und 10 die Gerade nicht.

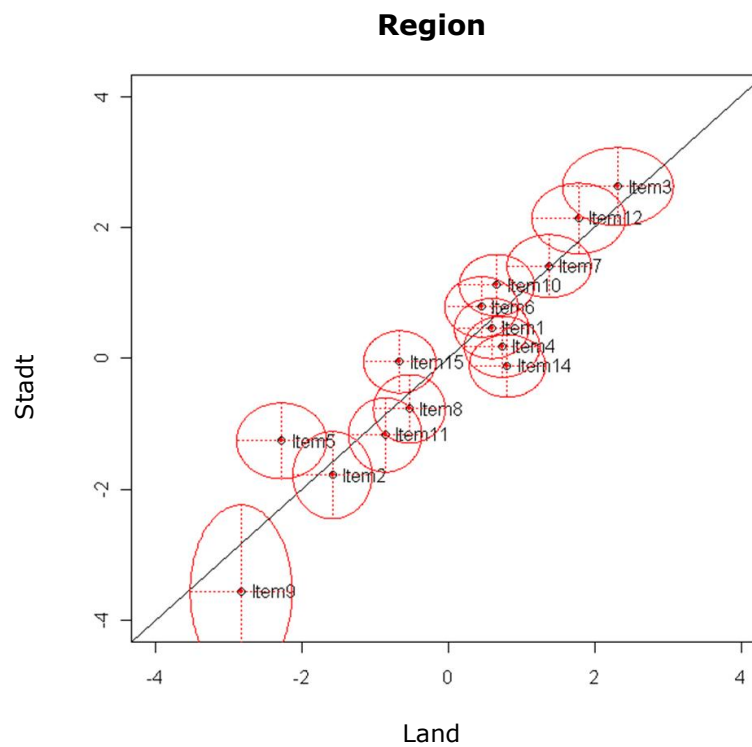


Abbildung 5: Grafischer Modelltest, Alltagswissen A, TK: Region

Der grafische Modelltest mit dem Teilungskriterium „Region“ (Abbildung 5) zeigt, dass die Konfidenz-Ellipsen der Items 5 und 14 die Gerade nicht schneiden.

Qualitative Analyse

- Item 1: „Wie nennt man die Nase bei Pferden?“

Dieses Item wies sowohl in der Itemfitstatistik, im grafischen Modelltest mit dem Teilungskriterium „Geschlecht“ als auch beim Z-Test mit dem Teilungskriterium „Geschlecht“ signifikante Werte auf. Es ist nachvollziehbar, dass Mädchen dieses Item eher lösen als Buben, da sie unter Umständen früher oder überhaupt mit Pferden in Kontakt kommen und sich für diese interessieren. Dieses Item ist demnach nicht fair dem Geschlecht gegenüber.

- Item 3: „Nenne ein Dateiformat, in dem Bilder am Computer abgespeichert werden können“

Bei diesem Item kristallisierte sich heraus, dass die SchülerInnen Schwierigkeiten mit dem Wort „Dateiformat“ hatten. Öfter wurde als Antwort „Eigene Dateien“, „Ordner“ oder ähnliches gegeben, was darauf schließen lässt, dass der Begriff „Format“ nicht gekannt wurde.

Zusätzlich konnte auch hier eine geringe, aber doch ungleiche Lösungsverteilung in Bezug auf das Geschlecht festgestellt werden. Hier könnte die Geschlechterrolle in der Kindheit und Jugend zu tragen kommen, da häufiger Männer an Computer interessiert sind.

So konnten sowohl beim Item 1 als auch beim Item 3 Schwachstellen gefunden werden. Deshalb wurde für die folgenden Parameterschätzungen zunächst Item 1, anschließend Item 3 und schlussendlich beide Items ausgeschlossen.

9.1.3 Parameterschätzung und Modelltest nach Itemselektion

Wie schon oben erwähnt wurden nun die Items 1, 3 und das ursprüngliche Item 13 aus dem Itempool ausgeschlossen. Da sowohl nach Ausschluss von Item 1, als auch 3 getrennt voneinander im Teilungskriterium „Geschlecht“ weiterhin ein signifikanter Wert auftrat, wurden nun beide gleichzeitig entfernt (einschließlich Item 13). Wie aus Tabelle 12 ersichtlich, konnte auch dadurch keine Rasch-Modell-Konformität in allen Teilungskriterien erreicht werden. Da nicht mehr als

32

20% der Items ausgeschieden werden sollen um weiterhin nach dem Rasch-Modell kalibrieren zu können, musste die H0 „Das Rasch-Modell gilt“ weiterhin verworfen werden.

Andersen LQT	Intern	Kritischer Wert df=10	Geschlecht	Region	Kritischer Wert df=11
LQ-Wert	9,896	22,44	37,449	23,223	23,94
χ^2 df	10	10	11	11	11
α	0,45	0,013	0	0,016	0,013

Tabelle 12: Parameterschätzung nach Itemselektion, Alltagswissen A, nicht Rasch-Modell-konform

Der Likelihood-Quotienten-Test nach Andersen mit dem Teilungskriterium „Intern“ verlief auch hier ohne Item 9, da es wegen ungünstiger Antwortmuster in den Subgruppen von der Berechnung ausgeschlossen wurde.

9.2 Alltagswissen, Testform B

9.2.1 Erste und zweite Parameterschätzung

Vor der ersten Parameterschätzung wurde nach erneuter Analyse der Items entschieden, aufgrund von Konstruktionsfehlern 3 Items vorweg auszuschließen.

Item 3: „Wo wurde die erste Atombombe abgeworfen?“ - erschien zwar auf den ersten Blick als eindeutig, die sehr unterschiedlich gestalteten Antworten zeigten jedoch, dass die Frage zu ungenau formuliert wurde. So wurden neben der eigentlich gewünschten Antwort „Hiroshima“ auch „Asien“, „Japan“ oder „Wüste“ gegeben (man denke hier an die Vorversuche in den USA). Da zwar alle diese Antworten richtig sind, es aber erstens nicht sinnvoll ist, ein so breites Spektrum an Antwortmöglichkeiten offen zu lassen und zweitens nicht mit Sicherheit gesagt werden kann, dass das Niveau der Antwortgeber bei den einzelnen Antworten gleich ist, wurde dieses Item aus dem Pool entfernt und sollte bei einer erneuten Testung genauer formuliert werden.

Item 7: „Was ist das Endprodukt der DNA?“ und Item 8: „Aus was besteht Salz überwiegend?“ – wurden nach erneuter Analyse als zu schwierig für SchülerInnen erachtet und fallen auch nicht in den Bereich des Alltagswissens.

So wurde die erste Parameterschätzung der Untertest 1 „Alltagswissen“, Testform B ohne die Items 3,7 und 8 durchgeführt.

Da der Likelihood-Quotienten-Test mit dem Teilungskriterium „Region“ wegen Item 5 nicht durchgeführt werden konnte, wurde auch dieses Item ausgeschlossen.

Für den Itempool musste die H_0 „Das Rasch-Modell gilt“ verworfen werden, da in allen Teilungskriterien auf dem 1% Niveau ein signifikantes Ergebnis auftrat. Der Likelihood-Quotienten-Test nach Andersen umfasste beim Teilungskriterium „Intern“ und „Region“ 9 Items, da Item 2 und 13 wegen ungünstiger Antwortmuster in den Subgruppen von der Berechnung ausgeschlossen wurde. Eine Übersicht der Ergebnisse ist in Tabelle 9 ersichtlich.

Andersen LQT	Intern	Krit. Wert df=8	Geschl.	Krit. Wert df=10	Region	Krit. Wert df=8
LQ-Wert	29,261	19,37	41,388	22,44	22,648	19,37
χ^2 df	8	8	10	10	8	8
α	0	0,013	0	0,013	0,004	0,013

Tabelle 13: Zweite Parameterschätzung, Alltagswissen B, nicht Rasch-Modellkonform

Eine weitere Itemselektion wäre hier nicht sinnvoll, da bereits zu viele Items ausgeschlossen wurden und die 20%-Grenze damit überschritten wurde.

9.3 Alltagswissen, Testform C

9.3.1 Erste Parameterschätzung

Auch hier wurde bereits vor der ersten Parameterschätzung nach erneuter Analyse der Items entschieden, aufgrund von Konstruktionsfehlern 2 Items vorweg auszuschließen.

Item 13: „Von wem stammt das Theaterstück „Maria Stuart“? (nur Nachname)“ und Item 15: „Ein stärkeres und länger anhaltendes Steigen der Kurse an der Börse wird auch wie genannt?“ - wurden nach erneuter Analyse als zu schwierig für SchülerInnen erachtet und fallen auch nicht in den Bereich des Alltagswissens.

So wurde die erste Parameterschätzung der Untertest 1 „Alltagswissen“, Testform C ohne die Items 13 und 15 durchgeführt.

Für den Itempool musste die H_0 „Das Rasch-Modell gilt“ verworfen werden, da in allen Teilungskriterien auf dem 1% Niveau ein signifikantes Ergebnis auftrat. Der Likelihood-Quotienten-Test nach Andersen umfasste beim Teilungskriterium „Intern“ nur 11 Items, da Item 8 wegen ungünstiger Antwortmuster in den Subgruppen von der Berechnung ausgeschlossen wurde. Eine Übersicht der Ergebnisse ist in Tabelle 9 ersichtlich.

Andersen LQT	Intern	Kritischer Wert df=11	Geschlecht	Region	Kritischer Wert df=12
LQ-Wert	88,593	23,94	65,123	42,867	25,41
χ^2 df	11	11	12	12	12
α	0	0,013	0	0	0,013

Tabelle 14: Erste Parameterschätzung, Alltagswissen C, nicht Rasch-Modell-konform

9.3.2 Itemselektion

Z-Test

Beim Z-Test (Tabelle 15) mit dem Teilungskriterium „Intern“ wurden die Items 1 (0,007), 2 (0,049), 3 (0,000), 4 (0,000), 11 (0,023) und 14 (0,000) signifikant. Die Items 1 (0,001), 2 (0,009), 3 (0,015), 4 (0,001), 5 (0,000), 6 (0,031) und 11 (0,020) unterschritten beim Teilungskriterium „Geschlecht“ den Wert. Item 3 (0,000), Item 6 (0,023) und Item 11 (0,005) wurden beim Teilungskriterium „Region“ signifikant.

Item 8, welches beim Internen Teilungskriterium nicht geschätzt werden konnte, wurde sowohl beim Teilungskriterium „Geschlecht“, als auch bei „Region“ als Rasch-Modell-konform erachtet.

	z-Wert Intern	p-Wert Intern	z-Wert Geschl.	p-Wert Geschl.	z-Wert Region	p-Wert Region
Item1	2,704	0,007	-3,329	0,001	0,515	0,607
Item2	-1,971	0,049	2,618	0,009	-0,689	0,491
Item3	6,304	0,000	-2,421	0,015	4,414	0,000
Item4	4,030	0,000	-3,319	0,001	1,573	0,116
Item5	0,428	0,668	3,785	0,000	1,483	0,138
Item6	0,209	0,835	-2,154	0,031	2,270	0,023
Item7	0,842	0,400	-1,528	0,127	0,641	0,521
Item8			1,229	0,219	-1,237	0,216
Item9	-1,807	0,071	0,933	0,351	-1,684	0,092
Item10	-1,346	0,178	-1,533	0,125	-0,274	0,784
Item11	-2,269	0,023	2,332	0,020	-2,791	0,005
Item12	-0,104	0,917	0,345	0,730	-0,295	0,768
Item14	-3,835	0,000	0,966	0,334	-1,216	0,224

Tabelle 15: Z-Test, Alltagswissen C, Teilungskriterien Intern, Geschlecht und Region

Itemfitstatistik

In der Itemfitstatistik wiesen die Items 1, 3 und 4 signifikante Werte auf (Tabelle 16).

Item	χ^2	df	p-Wert
Item1	207,474	164	0,012
Item2	92,902	164	1,000
Item3	330,955	164	0,000
Item4	245,147	164	0,000
Item5	167,775	164	0,404
Item6	166,932	164	0,422
Item7	158,186	164	0,613
Item8	68,951	164	1,000
Item9	130,687	164	0,974
Item10	134,856	164	0,953
Item11	114,228	164	0,999
Item12	175,045	164	0,263
Item14	86,394	164	1,0000

Tabelle 16: Itemfitstatistik, Alltagswissen C

Grafische Modelltests

Der grafische Modelltest mit dem Teilungskriterium „Höhe des Rohwertes“ (Abbildung 6) zeigte, dass die Konfidenz-Ellipsen der Items 1, 3, 4, 11 und 14 die Gerade nicht schneiden.

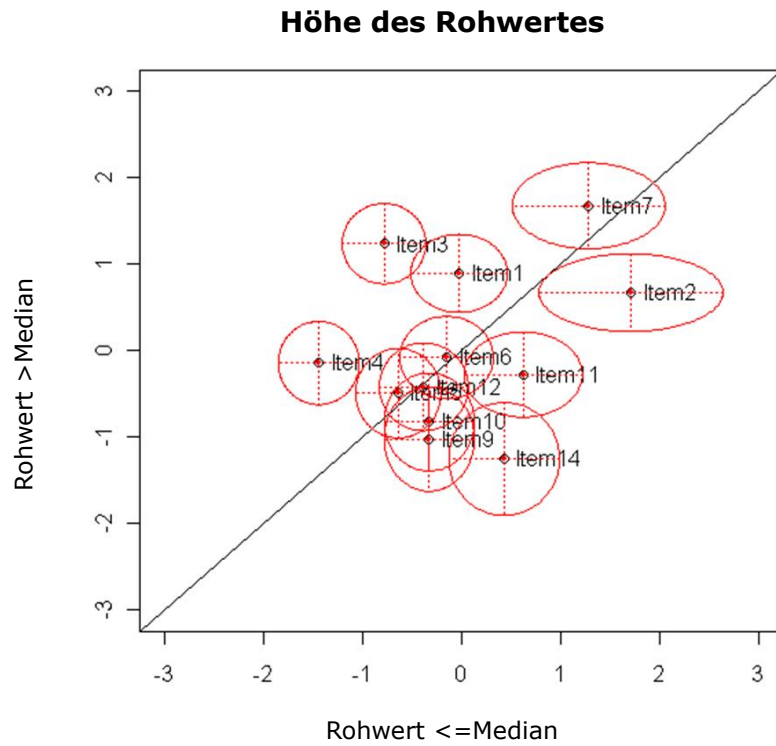


Abbildung 6: Grafischer Modelltest, Alltagswissen C, TK: Höhe des Rohwertes

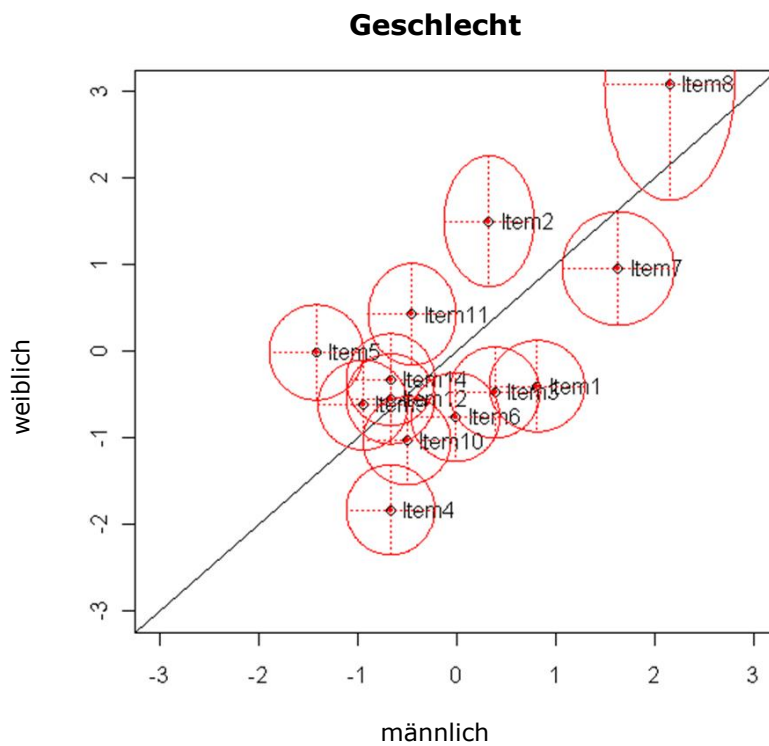


Abbildung 7: Grafischer Modelltest, Alltagswissen C, TK: Geschlecht

Im grafischen Modelltest mit dem Teilungskriterium „Geschlecht“ (Abbildung 7) schneiden die Konfidenz-Ellipsen der Items 1, 2, 3, 4, 5, 6 und 11 die Gerade nicht.

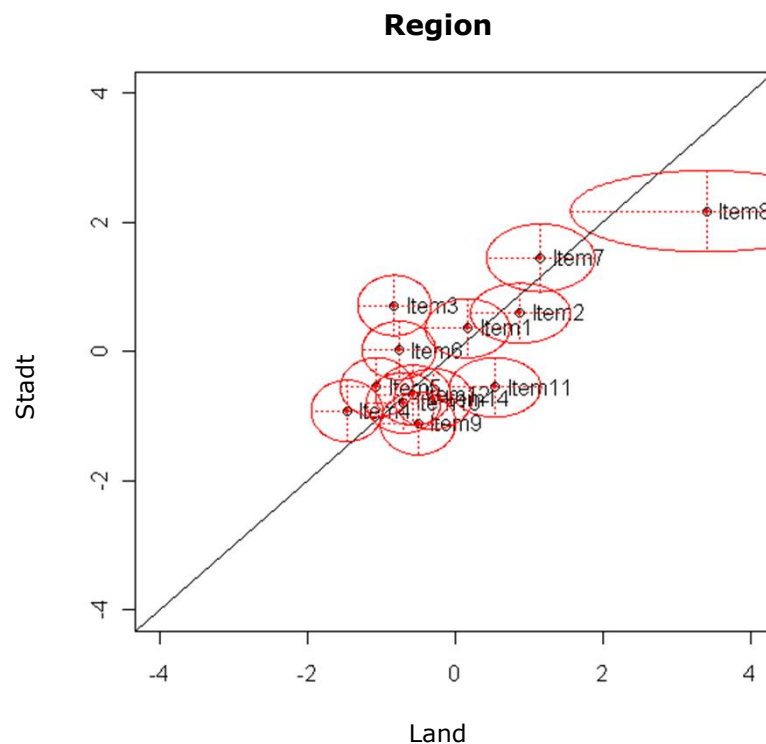


Abbildung 8: Grafischer Modelltest, Alltagswissen C, TK: Region

Der grafische Modelltest mit dem Teilungskriterium „Region“ (Abbildung 8) zeigt, dass die Konfidenz-Ellipsen der Items 3, 6 und 11 die Gerade nicht schneiden.

Qualitative Analyse

- Item 3: „Welche Farbe entsteht, wenn man gelb und grün mischt?“

Dieses Item wurde eher von Mädchen als von Buben gelöst und war für Landkinder einfacher als für Kinder aus der Stadt. Kann man in Bezug auf das Geschlecht argumentieren, dass Mädchen sich vielleicht eher dem Malen hingeben, und sie deswegen eher die Erfahrung des Mischens gemacht haben, so bleiben Spekulationen über den Unterschied zwischen Stadt und Land aus.

- Item 11: „Von welchem Interpreten stammt das Lied „We are the Champions“?“

Hier lösten eher Stadtkinder das Item, und eher Buben als Mädchen konnten die richtige Antwort finden. Ein Grund dafür könnte in der Verwendung des Liedes liegen. So wird es z.B. gerne nach wichtigen Fußballspielen gesendet und deshalb kann es sein, dass sich Buben eher mit dem Lied beschäftigt haben. Den Unterschied zwischen Stadt- und Landkindern könnte man in der Menge und im guten Empfang von Radiosendern begründen. Während in der Stadt mehrere Sender gut empfangen werden können, die unterschiedliche Musikrichtungen abdecken, empfängt man in ländlicheren Regionen nicht mehr alle Sender und der Empfang ist auch nicht mehr so gut. Ebenso besteht ein Unterschied der gehörten Musik zwischen Stadt und Land.

9.3.3 Parameterschätzung und Modelltest nach Itemselektion

Wie schon oben erwähnt wurden nun die Items 3 und 11, sowie die ursprünglichen Items 13 und 15 aus dem Itempool ausgeschlossen. Da sowohl nach Ausschluss von Item 3, als auch 11 getrennt voneinander allen Teilungskriterien weiterhin ein signifikanter Wert auftrat, wurden nun beide gleichzeitig entfernt (einschließlich der Items 13 und 15). Wie aus Tabelle 17 ersichtlich, konnte auch dadurch keine Rasch-Modell-Konformität in allen Teilungskriterien erreicht werden. Lediglich im Teilungskriterium „Region“ wurde der Wert nicht mehr signifikant. Da nicht mehr als 20% der Items ausgeschieden werden dürfen um weiterhin nach dem Rasch-Modell kalibrieren zu können, musste die H0 „Das Rasch-Modell gilt“ weiterhin verworfen werden.

Andersen LQT	Intern	Kritischer Wert df=9	Geschlecht	Region	Kritischer Wert df=10
LQ-Wert	44,683	20,92	56,65	17,024	22,44
χ^2 df	9	9	10	10	10
α	0	0,013	0	0,074	0,013

Tabelle 17: Parameterschätzung nach Itemselektion, Alltagswissen C, nicht Rasch-Modell-konform

Der Likelihood-Quotienten-Test nach Andersen mit dem Teilungskriterium „Intern“ verlief auch hier ohne Item 8, da es wegen ungünstiger Antwortmuster in den Subgruppen von der Berechnung ausgeschlossen wurde.

9.4 Alltagswissen, Testform D

9.4.1 Erste Parameterschätzung

Für den Itempool der ersten Parameterschätzung des Untertest 1 „Alltagswissen“, Testform D musste die H0 „Das Rasch-Modell gilt“ verworfen werden, da in allen Teilungskriterien auf dem 1% Niveau ein signifikantes Ergebnis auftrat. Der Likelihood-Quotienten-Test nach Andersen umfasste beim Teilungskriterium „Intern“ nur 13 Items (ohne Item 12 und 15) und beim Teilungskriterium „Region“ nur 14 Items (ohne Item 3) da sie wegen ungünstiger Antwortmuster in den Subgruppen von der Berechnung ausgeschlossen wurde. Eine Übersicht der Ergebnisse ist in Tabelle 18 ersichtlich.

Andersen	Intern	Krit.	Geschl.	Krit.	Region	Krit.
LQT		Wert		Wert		Wert
		df=12		df=14		df=13
LQ-Wert	39,117	25,41	34,428	28,29	42,171	26,86
χ^2 df	12	12	14	14	13	13
α	0	0,013	0,002	0,013	0	0,013

Tabelle 18: Erste Parameterschätzung, Alltagswissen D, nicht Rasch-Modell-konform

9.4.2 Itemselektion

Z-Test

Beim Z-Test (Tabelle 15) mit dem Teilungskriterium „Intern“ wurden die Items 7 (0,003), 8 (0,027) und 11 (0,001). Die Items 6 (0,000) und 13 (0,004) unterschritten beim Teilungskriterium „Geschlecht“ den Wert. Item 1 (0,009), Item 4 (0,004) und Item 9 (0,000) wurden beim Teilungskriterium „Region“ signifikant.

Item 12 und 15, welche beim Internen Teilungskriterium nicht geschätzt werden konnte, wurden sowohl beim Teilungskriterium „Geschlecht“, als auch bei „Region“ als Rasch-Modell-konform erachtet. Item 3 konnte zwar nicht beim

Teilungskriterium „Region“ nicht geschätzt werden, war jedoch bei den anderen zwei Teilungskriterien Rasch-Modell-konform.

	z-Wert Intern	p-Wert Intern	z-Wert Geschl.	p-Wert Geschl.	z-Wert Region	p-Wert Region
Item1	1.531	0.126	0.462	0.644	2.596	0.009
Item2	0.707	0.480	-1.602	0.109	0.764	0.445
Item3	1.834	0.067	-0.101	0.919		
Item4	-1.281	0.200	0.000	1.000	-2.859	0.004
Item5	-0.566	0.571	-0.162	0.872	-0.414	0.679
Item6	1.338	0.181	3.944	0.000	1.908	0.056
Item7	-2.920	0.003	0.275	0.783	-0.754	0.451
Item8	-2.211	0.027	0.173	0.863	-0.563	0.573
Item9	-1.940	0.052	0.772	0.440	-4.404	0.000
Item10	-0.546	0.585	-1.324	0.185	1.813	0.070
Item11	-3.416	0.001	-0.330	0.742	-1.711	0.087
Item12			1.314	0.189	0.734	0.463
Item13	1.690	0.091	2.857	0.004	1.668	0.095
Item14	1.779	0.075	-0.176	0.860	-1.070	0.284
Item15			-1.896	0.058	0.946	0.344

Tabelle 19: Z-Test, Alltagswissen D, Teilungskriterien Intern, Geschlecht und Region

Itemfitstatistik

In der Itemfitstatistik wiesen die Items 1, 3, 6, 13 und 14 signifikante Werte auf (Tabelle 20).

Item	χ^2	df	p-Wert
Item1	872.141	160	0.000
Item2	176.684	160	0.174
Item3	335.304	160	0.000
Item4	153.380	160	0.632
Item5	125.282	160	0.980
Item6	392.121	160	0.000
Item7	107.703	160	0.999
Item8	118.978	160	0.994
Item9	105.798	160	1.000
Item10	67.493	160	1.000
Item11	94.965	160	1.000
Item12	59.950	160	1.000
Item13	216.867	160	0.002
Item14	219.521	160	0.001
Item15	39.048	160	1.000

Tabelle 20: Itemfitstatistik, Alltagswissen D

Grafische Modelltests

Der grafische Modelltest mit dem Teilungskriterium „Höhe des Rohwertes“ (Abbildung 9) zeigte, dass die Konfidenz-Ellipsen der Items 7, 8 und 11 die Gerade nicht schneiden.

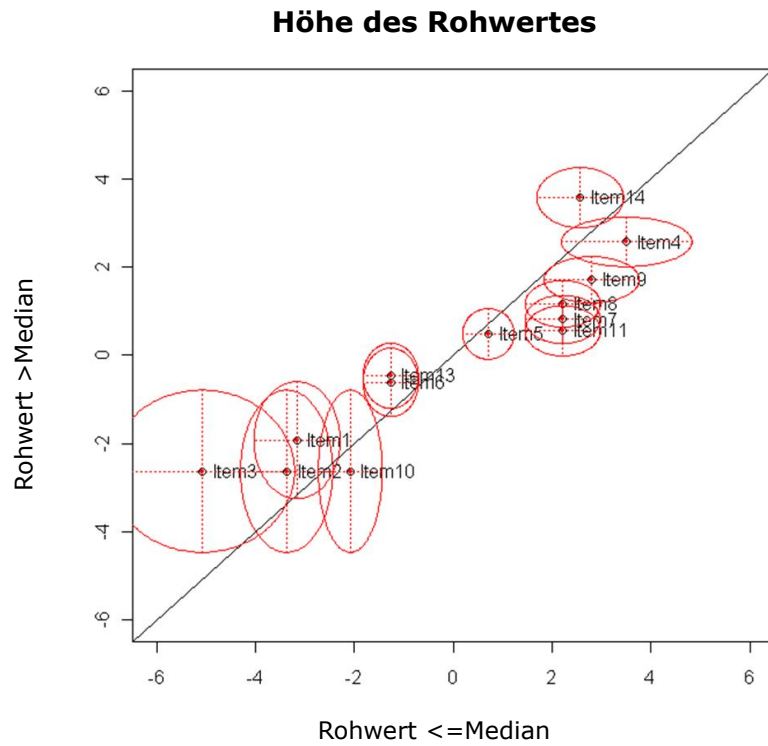


Abbildung 9: Grafischer Modelltest, Alltagswissen D, TK: Höhe des Rohwertes

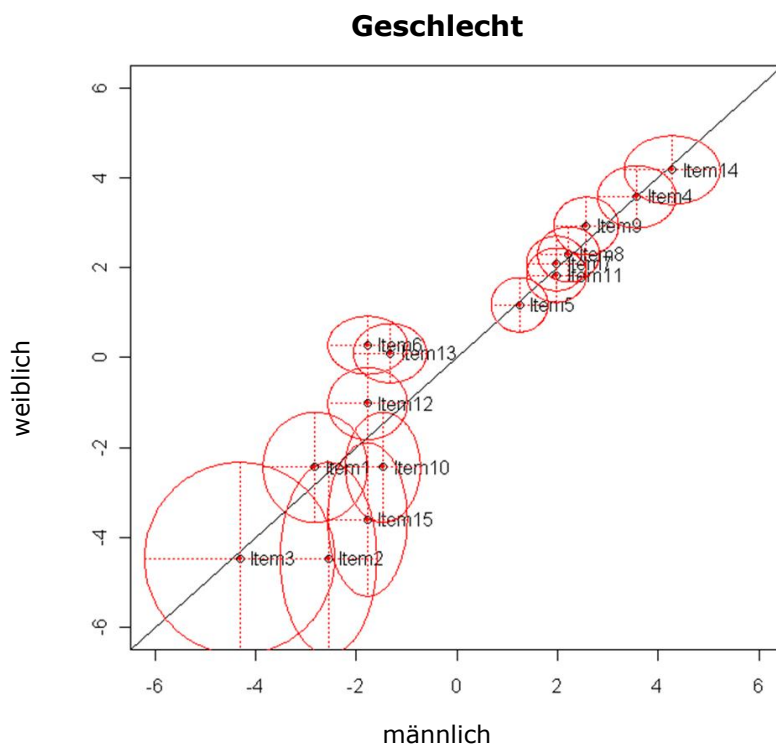


Abbildung 10: Grafischer Modelltest, Alltagswissen D, TK: Geschlecht

Im grafischen Modelltest mit dem Teilungskriterium „Geschlecht“ (Abbildung 10) schneiden die Konfidenz-Ellipsen der Items 6 und 13 die Gerade nicht.

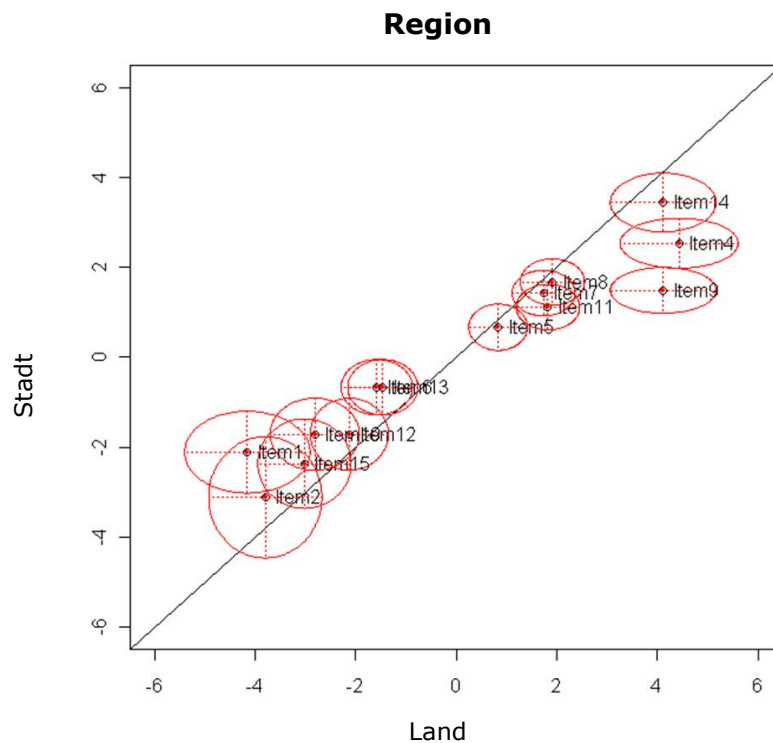


Abbildung 11: Grafischer Modelltest, Alltagswissen D, TK: Region

Der grafische Modelltest mit dem Teilungskriterium „Region“ (Abbildung 11) zeigt, dass die Konfidenz-Ellipsen der Items 1, 4 und 9 die Gerade nicht schneiden.

Qualitative Analyse

Durch die qualitative Analyse konnten keine Erklärungen für die unterschiedlichen Ergebnisse in den Teilungskriterien gefunden werden. Daher musste die H0 „Das Rasch-Modell gilt“ verworfen werden.

9.5 Angewandtes Rechnen, Testform A

9.5.1 Erste Parameterschätzung

Für den Itempool des Untertest 3 „Angewandtes Rechnen“ Testform A konnte die H0 „Das Rasch-Modell gilt“ beibehalten werden. In allen Teilungskriterien war der Wert über dem 1% Niveau. Der Likelihood-Quotienten-Test nach Andersen umfasste beim internen Teilungskriterium nur 6 Items, da die Items 1 bis 8,

sowie Item 12 wegen ungünstiger Antwortmuster in den Subgruppen von der Berechnung ausgeschlossen wurde. Eine Übersicht der Ergebnisse sind in Tabelle 21 und den Abbildungen 12, 13 und 14 ersichtlich.

Andersen LQT	Intern	Kritischer Wert df=5	Geschlecht	Region	Kritischer Wert df=14
LQ-Wert	4,057	14,45	16,522	15,905	28,29
χ^2 df	5	5	14	14	14
α	0,541	0,013	0,283	0,319	0,013

Tabelle 21 Erste Parameterschätzung, Angewandtes Rechnen A, Rasch-Modell-konform

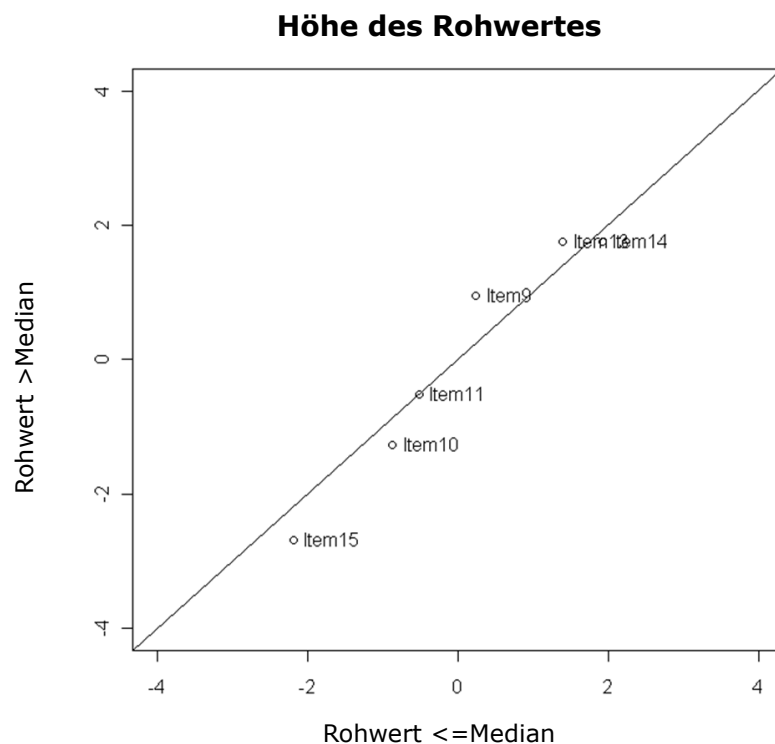


Abbildung 12: Grafischer Modelltest, Angewandtes Rechnen A, TK: Höhe des Rohwertes

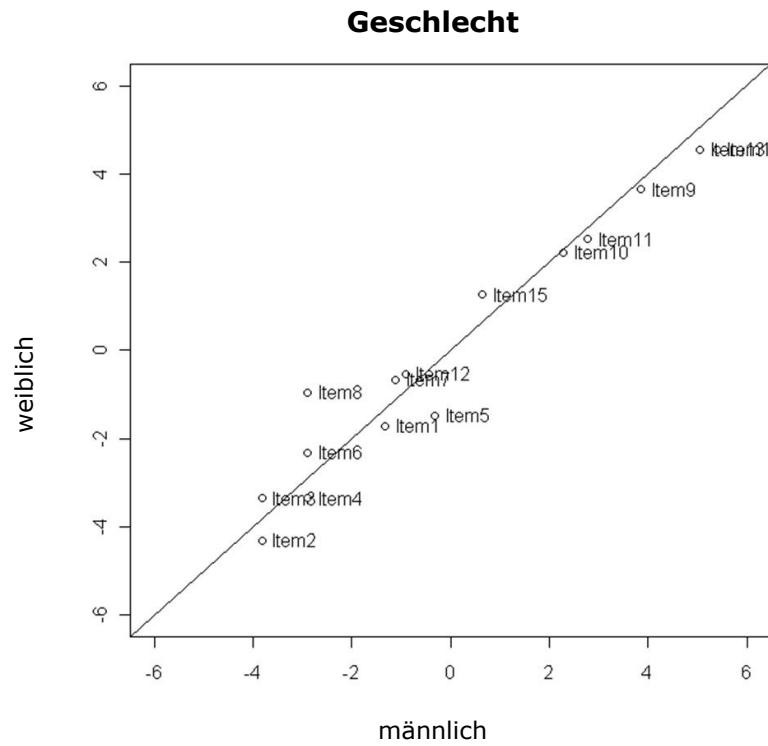


Abbildung 13: Grafischer Modelltest, Angewandtes Rechnen A, TK: Geschlecht

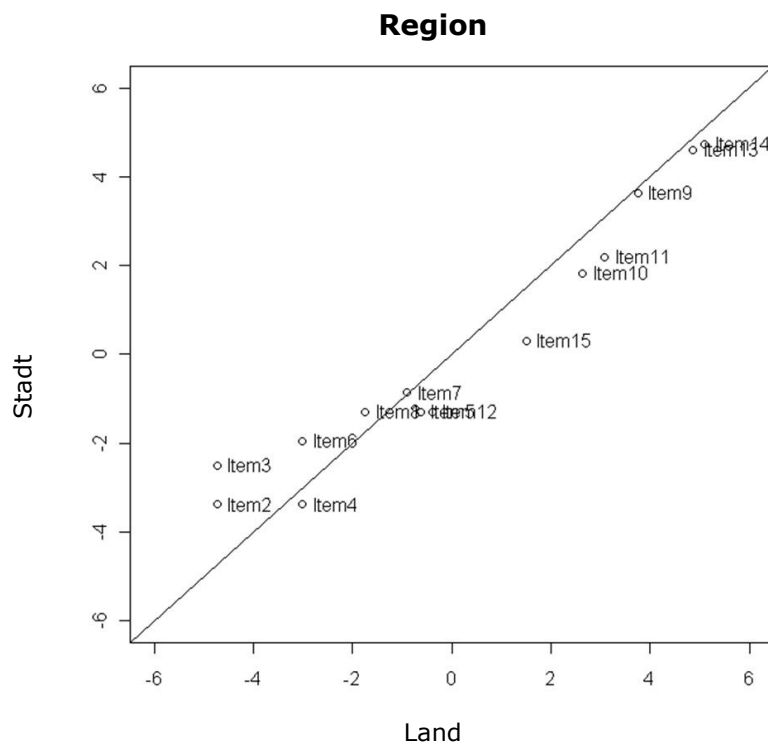


Abbildung 14: Grafischer Modelltest, Angewandtes Rechnen A, TK: Region

9.6 Angewandtes Rechnen, Testform B

9.6.1 Erste und zweite Parameterschätzung

Vor der ersten Parameterschätzung wurde Item 12 aus dem Pool entfernt, da sich nach weiterer Analyse herausstellte, dass es missverständlich formuliert wurde.

Item 12: „Ein Haus wird gebaut und soll mehrere Stöcke haben. Die Zimmerhöhe muss 2 Meter betragen, und die Decke muss 75cm dick sein. Wie viele Stöcke kann man maximal bauen, wenn die Höhe des Hauses ohne Dach maximal 8 Meter betragen darf?“

Es ist hier nicht klar, ob das Erdgeschoß als Stock gerechnet werden soll oder nicht. Da dadurch natürlich unterschiedliche Ergebnisse zustande kommen, wurde das Item für die weitere Überprüfung entfernt.

Für den Itempool des Untertest 3 „Angewandtes Rechnen“ Testform B konnte die H_0 „Das Rasch-Modell gilt“ beibehalten werden, da in allen Teilungskriterien der Wert nicht signifikant wurde. Der Likelihood-Quotienten-Test nach Andersen umfasste beim internen Teilungskriterium nur 6 Items, da die Items 1, 2, 4, 5, 6, 8, 9 und 15 wegen ungünstiger Antwortmuster in den Subgruppen von der Berechnung ausgeschlossen wurde. Bei den Teilungskriterien „Geschlecht“ und „Region“ wurde ebenfalls Item 1 ausgeschlossen, deshalb wurde dieses Item auch aus dem Pool entfernt. Die zweite Parameterschätzung erfolgte deshalb ohne Item 1 und 12. Eine Übersicht der Ergebnisse sind in Tabelle 22 und den Abbildungen 15,16 und 17 ersichtlich.

Andersen LQT	Intern	Kritischer Wert df=5	Geschlecht	Region	Kritischer Wert df=12
LQ-Wert	11,192	14,45	6,736	19,162	25,41
χ^2 df	5	5	12	12	12
α	0,048	0,013	0,875	0,085	0,013

Tabelle 22: Zweite Parameterschätzung, Angewandtes Rechen B, Rasch-Modell-konform

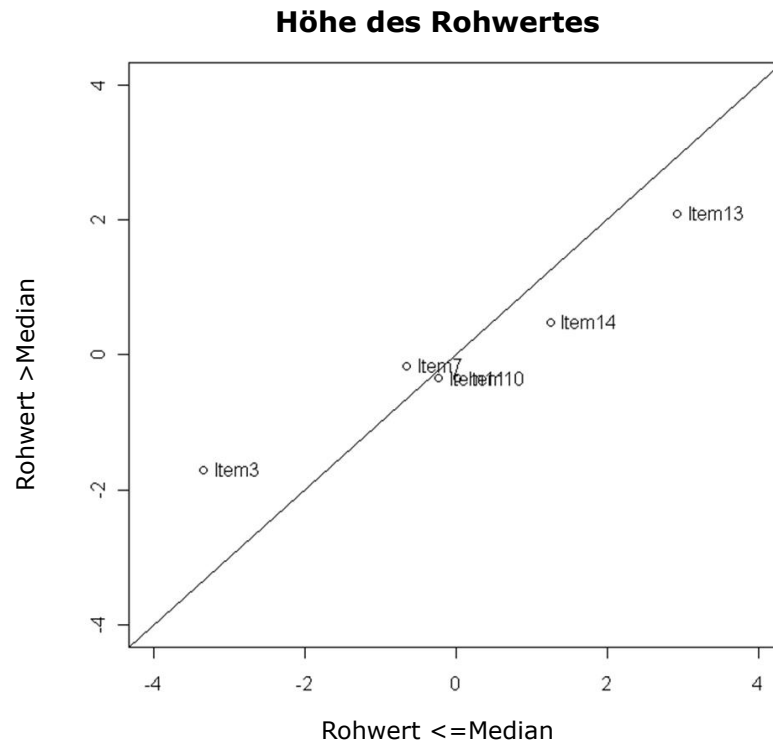


Abbildung 15: Grafischer Modelltest, Angewandtes Rechnen B, TK: Höhe des Rohwertes

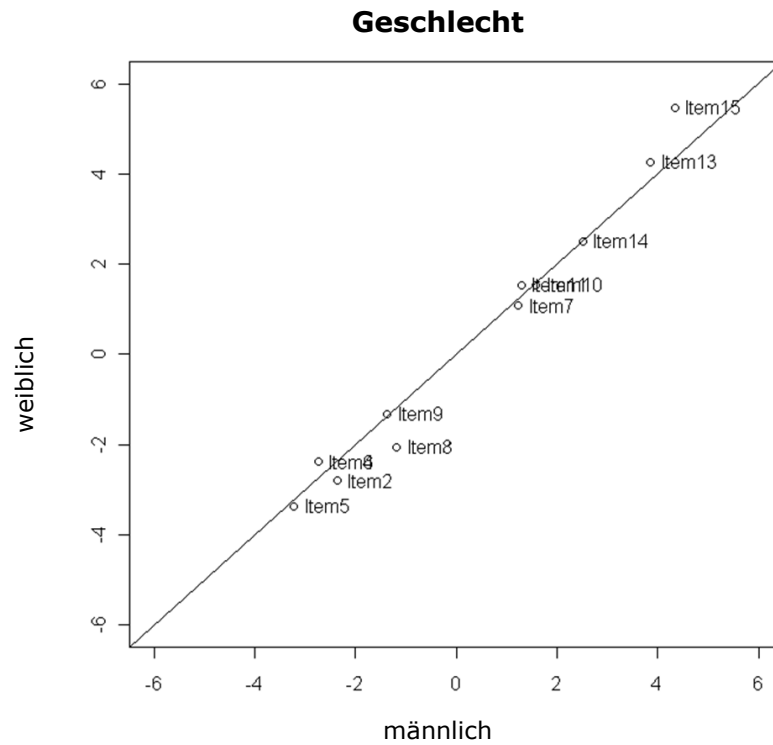


Abbildung 16: Grafischer Modelltest, Angewandtes Rechnen B, TK: Geschlecht

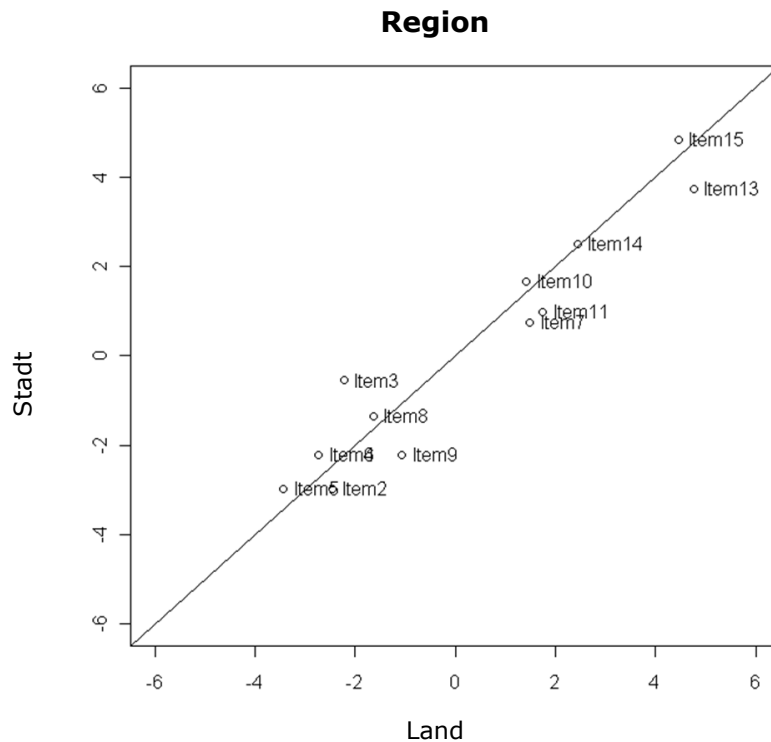


Abbildung 17: Grafischer Modelltest, Angewandtes Rechnen B, TK: Region

9.7 Angewandtes Rechnen, Testform C

9.7.1 Erste Parameterschätzung

Für den Itempool des Untertest 3 „Angewandtes Rechnen“ Testform C musste die H_0 „Das Rasch-Modell gilt“ verworfen werden, da signifikante Ergebnisse in allen Teilungskriterien auftraten. Der Likelihood-Quotienten-Test nach Andersen umfasste beim internen Teilungskriterium nur 11 Items, da die Items 1, 4, 7, und 14 wegen ungünstiger Antwortmuster in den Subgruppen von der Berechnung ausgeschlossen wurde. Bei den Teilungskriterien „Geschlecht“ und „Region“ wurde ebenfalls die Items 1, 4 und 14 ausgeschlossen. Da damit bereits 3 Items wegfielen, musste die H_0 weiter verworfen bleiben. Eine Übersicht der Ergebnisse ist in Tabelle 23 ersichtlich.

Andersen LQT	Intern	Kritischer Wert df=10	Geschlecht	Region	Kritischer Wert df=11
LQ-Wert	33,07	22,44	38,141	30,906	23,94
χ^2 df	10	10	11	11	11
α	0	0,013	0	0,001	0,013

Tabelle 23: Erste Parameterschätzung, Angewandtes Rechen C, nicht Rasch-Modell-konform

9.8 Angewandtes Rechnen, Testform D

9.8.1 Erste Parameterschätzung

Für den Itempool des Untertest 3 „Angewandtes Rechnen“ Testform D konnte die H_0 „Das Rasch-Modell gilt“ beibehalten werden. In allen Teilungskriterien war der Wert über dem 1% Niveau. Der Likelihood-Quotienten-Test nach Andersen umfasste beim internen Teilungskriterium nur 11 Items, da die Items 2, 6, 9 und 10 wegen ungünstiger Antwortmuster in den Subgruppen von der Berechnung ausgeschlossen wurde. Beim Teilungskriterium „Geschlecht“ war dies bei Item 6 der Fall, beim Teilungskriterium „Region“ bei den Items 2, 5, 6 und 9. Da Item 6 in keinem Test miteinbezogen wurde, wurde es aus dem Pool entfernt. Die restlichen Items wiesen nicht signifikante Werte in den übrigen Teilungskriterien auf. Eine Übersicht der Ergebnisse sind in Tabelle 24 und den Abbildungen 18, 19 und 20 ersichtlich.

Andersen LQT	Intern	Krit. Wert df=10	Geschl.	Krit. Wert df=13	Region	Krit. Wert df=10
LQ-Wert	7,153	22,44	20,71	26,86	16,611	22,44
χ^2 df	10	10	13	13	10	10
α	0,711	0,013	0,079	0,013	0,083	0,013

Tabelle 24: Erste Parameterschätzung, Angewandtes Rechen D, Rasch-Modell-konform

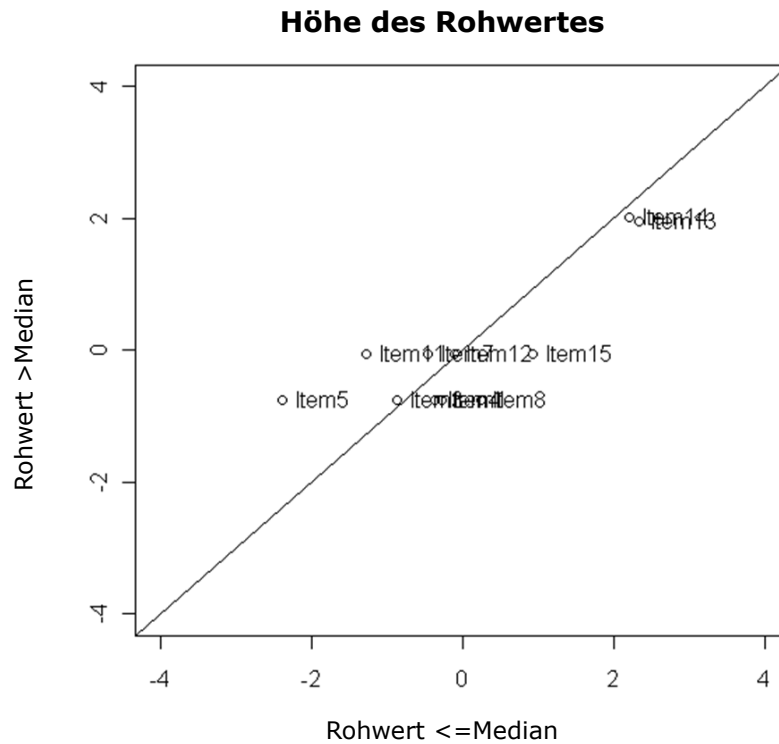


Abbildung 18: Grafischer Modelltest, Angewandtes Rechnen D, TK: Höhe des Rohwertes

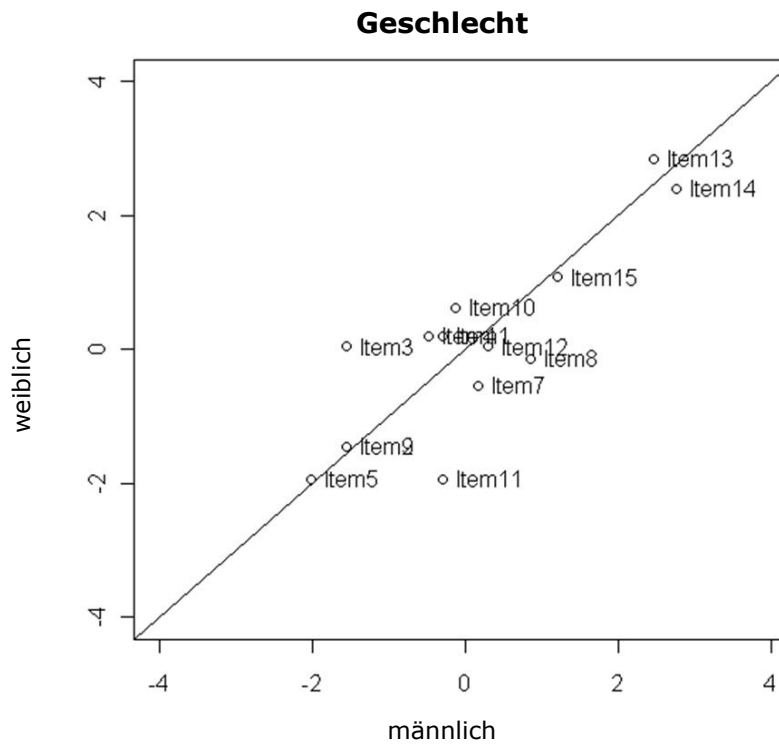


Abbildung 19: Grafischer Modelltest, Angewandtes Rechnen D, TK: Geschlecht

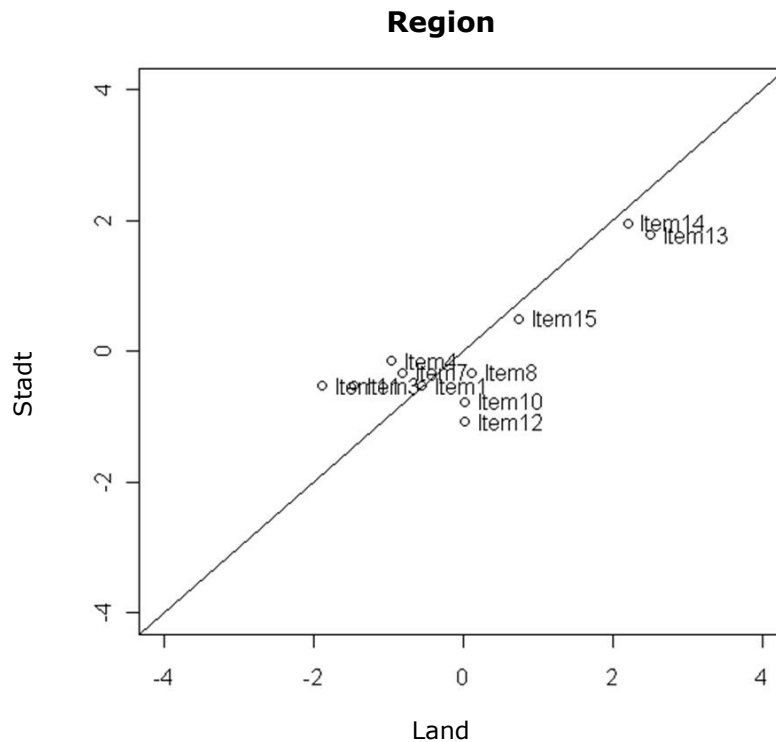


Abbildung 20: Grafischer Modelltest, Angewandtes Rechnen D, TK: Region

9.9 Mittelwertvergleiche

Jetzt sollte noch untersucht werden, ob es einen Einfluss von Geschlecht, Schultyp und Schulstufe auf die Höhe des Rohwertes gibt. Sind die entsprechenden Voraussetzungen dafür gegeben, wird eine dreifache Varianzanalyse durchgeführt. Dabei wurden die SchülerInnen der 3., 4. und 9. Schulstufe von der Analyse ausgeschlossen, da hier zu wenige Daten pro Testform vorhanden waren. Der Levene Test für die Überprüfung der Homogenität der Varianzen wurde jedoch in einigen Testformen signifikant - Bei „Alltagswissen“, Testform B und bei „Angewandtes Rechnen“, Testform C und D. Hier wurde je nach Signifikanz bei Überprüfung der Homogenität der Varianzen auf andere parametrische oder nichtparametrische Verfahren zurückgegriffen.

So ergaben sich die unten stehenden Hypothesen für beide Untertests, wobei die H01 bis H07 für die Testformen A, C und D bei „Alltagswissen“ und die Testformen A und B für „Angewandtes Rechnen“ geprüft werden konnten. Für die Testform B bei „Alltagswissen“ und Testformen C und D bei „Angewandtes Rechnen“ konnten nur die H1 bis H3 untersucht werden.

Bei der Auflistung der 0-Hypothesen steht das X für „Alltagswissen“ oder „Angewandtes Rechnen“, das Y kann für A, B, C oder D stehen.

H01: „Die Testwerte im Untertest X Testform Y unterscheiden sich nicht zwischen den einzelnen Stufen des Faktors Geschlecht“

H02: „Die Testwerte im Untertest X Testform Y unterscheiden sich nicht zwischen den einzelnen Stufen des Faktors Schultyp“

H03: „Die Testwerte im Untertest X Testform Y unterscheiden sich nicht zwischen den einzelnen Stufen des Faktors Schulstufe“

H04: „Es gibt keine Wechselwirkungen zwischen Geschlecht und Schultyp in Bezug auf die Testwerte im Untertest X Testform Y“

H05: „Es gibt keine Wechselwirkungen zwischen Geschlecht und Schulstufe in Bezug auf die Testwerte im Untertest X Testform Y“

H06: „Es gibt keine Wechselwirkungen zwischen Schulstufe und Schultyp in Bezug auf die Testwerte im Untertest X Testform Y“

H07: „Es gibt keine Wechselwirkungen zwischen Geschlecht, Schulstufe und Schultyp in Bezug auf die Testwerte im Untertest X Testform Y“

Da angenommen werden kann, dass mit der Höhe der Schulstufe auch mehr Fragen beantwortet werden können, wird bei der 3. Hypothese die H1 erwartet (H13).

9.9.1 Ergebnisse Mittelwertvergleich

Untertest Alltagswissen

Testform	Quelle	F	Sig	η^2
Testform A	Schultyp	10,940	0,001	0,075
	Schulstufe	8,628	0,000	0,162
Testform C	Geschlecht	33,187	0,000	0,205
	Schultyp	31,049	0,000	0,194
	Schulstufe	29,281	0,000	0,405
Testform D	Geschlecht	33,187	0,000	0,205
	Schultyp	31,049	0,000	0,194
	Schulstufe	29,281	0,000	0,405

Tabelle 25: Dreifache Varianzanalyse, Alltagswissen, Testform A, C und D

Testform	Quelle/Test	Z	sig	ϕ
Testform B	Schultyp/ Mann- Whitney-U	-4,242	0,000	0,345
	Schulstufe/ Kruskal-Wallis- Test	χ^2/df 53,584/3	sig 0,000	ω 0,596

Tabelle 26: Mittelwertvergleich, Alltagswissen, Testform B

Untertest Angewandtes Rechnen

Testform	Quelle	F	Sig	η^2
Testform A	Geschlecht	5,408	0,022	0,038
	Schultyp	7,963	0,005	0,055
	Schulstufe	4,430	0,005	0,089
Testform B	Schultyp	28,215	0,000	0,174
	Schulstufe	12,543	0,000	0,219

Tabelle 27: Dreifache Varianzanalyse, Angewandtes Rechnen, Testform A & B

Testform	Quelle/Test	T/df	Sig.	d
Testform C	Schultyp/T-Test	-2,858/144	0,005	0,472
		F	Sig.	η^2
	Schulstufe/ Varianzanalyse	11,763	0,000	0,199

Tabelle 28: Mittelwertvergleich, Angewandtes Rechnen, Testform C

Testform	Quelle/Test	Z	sig	ϕ
Testform D	Schultyp/ Mann-Whitney-U	-2,647	0,008	0,222
		χ^2/df	sig	ω
	Schulstufe/ Kruskal-Wallis-Test	17,309/3	0,001	0,349

Tabelle 29: Mittelwertvergleich, Angewandtes Rechnen, Testform D

Die signifikanten Ergebnisse der Tests sind in den Tabellen 25 bis 29 zu betrachten. In der letzten Spalte der Tabellen sind jeweils die verschiedenen Effektstärken angegeben.

Die Faktoren Schultyp und Schulstufe wurden in sämtlichen Testformen signifikant, weshalb die H02 und H03 verworfen werden müssen. Die Erwartung, dass die H13 gilt, konnte also bestätigt werden. Der Faktor Geschlecht wurde in Testform C und D des Untertests Alltagswissen, sowie in Testform A des Untertests Angewandtes Rechnen signifikant. In diesen Testformen musste die H0 also verworfen werden. Die Wechselwirkungen zwischen zwei bzw. allen drei Faktoren ergab in keinen der Untertests und deren Testformen, in denen die dreifache Varianzanalyse durchgeführt werden konnte, ein signifikantes Ergebnis. Somit konnten die H04 bis H07 bei den entsprechenden Testformen beibehalten werden.

10 Diskussion und Ausblick

Die Machbarkeitsstudie des AID 2 als Gruppen- oder Computerverfahren kann zumindest für eine Gruppentestung für die Untertests „Alltagswissen“ und „Angewandtes Rechnen“ positiv abgeschlossen werden. Die Testung konnte in sämtlichen Schulstufen problemlos durchgeführt werden, die Anweisungen wurden durchwegs verstanden. Für eine entsprechende Computertestung wurden zwar Ideen gesammelt, diese konnte jedoch aufgrund zu geringer Programmierfertigkeiten nicht umgesetzt werden.

Sämtliche Testformen des Untertest „Alltagswissen“ verhielten sich dem Rasch-Modell nicht konform. Dies könnte an der zu kurzen Anleitung zur Itemkonstruktion gelegen haben. Weitere Recherchen wären hier notwendig und zumindest bei diesem Untertest wäre es unter Umständen sinnvoll, die Items weiterhin von Experten beurteilen zu lassen. Da viele Items gerade im Teilungskriterium „Geschlecht“ auffällig waren, sollten sie gerade in Bezug darauf hinterfragt werden. Die Überprüfung der Inhalte der Items durch Wikipedia (www.wikipedia.de) hat sich als durchaus passend herausgestellt. Man könnte diesen Schritt zwar als unwissenschaftlich betrachten, so erfüllte er jedoch den Zweck der Überprüfung.

Eine Ursache der bei der Auswertung bemerkten generell eher geringen Antwortquote könnte auch die Vorgabereihenfolge gewesen sein. So wurde dieser Untertest an vierter Stelle vorgegeben, nahezu am Ende jeder Unterrichtsstunde. Hier könnte die Motivation schon etwas gesunken sein oder es trat Erschöpfung auf, nachdem bereits drei anspruchsvolle Untertests vorgegeben wurden. So wurde bemerkt, dass z.B. Fragen wie „Welcher Buchstabe ist der letzte des Alphabetes?“ nicht beantwortete wurden.

Auch ein anderes Antwortformat wäre denkbar. So wurde zwar ein Modus dem LEWITE (Wagner-Menghin, 2004) ähnlich in Betracht gezogen, dieser wurde jedoch verworfen. Das freie Reproduzieren aus dem Gedächtnis erschien in diesem Fall sinnvoller, jedoch könnte es durch ein Multiple-Choice-Format zu einer größeren Passung kommen. So war es in diesem Fall interessant zu erfahren, dass manche Kinder z.B. nicht wussten, was man unter einem Kontinent versteht. Die Frage „In welchem Kontinent fließt der Nil?“ wurde häufig mit „Ägypten“ beantwortet. Bei der Auswahl passender Antwortmöglichkeiten

müsste dies beachtet werden, anstatt z.B. die Namen anderer Kontinente anzuführen.

Die theorie- und regelgeleitete Itemkonstruktion im Untertest „Angewandtes Rechnen“ erwies sich als relativ passend, wenn auch Verbesserungen angestrebt werden können. So könnten vor allem bei Stufe 1 und 2 Items konstruiert werden, die den angegebenen Zahlenraum besser ausnutzen. Dann würde es möglicherweise zu weniger Ausfällen bei der Modellüberprüfung durch ungünstige Antwortmuster in den Subgruppen kommen.

Die Mittelwertvergleiche ergaben bei einigen Testformen Unterschiede in Bezug auf das Geschlecht, die Effektstärke ist dabei aber gering. Unterschiede des Testwertes in sämtlichen Testformen gab es zwischen den Schultypen und Schulstufen. Dabei fiel die Effektstärke der Schulstufe bis auf eine Ausnahme (Angewandtes Rechnen, Testform C) immer höher aus als die des Schultyps. Die Höhe des Effektes variierte dabei von gering (0,055 bei Angewandtes Rechnen, Testform A) bis stark (0,596 bei Alltagswissen, Testform B).

Nach diesen Verbesserungen kann bei einer erneuten Untersuchung eine größere Stichprobe und mehr Teilungskriterien verwendet werden (vgl. Kubinger & Draxler, 2007). So stellen die ca. 170 Personen pro Testform eine eher geringe Anzahl für eine Überprüfung auf Rasch-Modell-Konformität dar. Das Alter als weiteres Teilungskriterium wäre erforderlich, zusätzlich könnte ebenso das Teilungskriterium „Lösung vs. Nicht-Lösung“ bei einem willkürlich gewählten Item durch den Andersen Likelihood-Ratio-Test angewendet werden oder eine Überprüfung durch Martin-Löfs Likelihood-Ratio-Test erfolgen.

Ein weiteres Manko der Untersuchung stellte die nicht randomisierte Stichprobe dar. So fanden sämtliche Testungen in Oberösterreich statt, wobei dabei in der Stadt das Gymnasium, und am Land die Volksschule und Hauptschulen getestet wurden. Unterschiede zwischen Stadt und Land könnten sich daher ebenso auf Unterschiede zwischen Gymnasium und Hauptschule beziehen. Daher sollten bei einer weiteren Untersuchung diesbezüglich Vorkehrungen getroffen werden.

In weiterer Folge sollte auch dazu übergegangen werden das AID 2 in eine erwünschte Computerform zu bringen.

Weiters könnte man die vier Testformen auch miteinander verbinden um festzustellen, ob sich diese auch zueinander Rasch-Modell-konform verhalten.

Dadurch wäre es möglich, mit den Daten mittels dem Linearen Logistischen Testmodell von Fischer (LLTM) die Passung der erstellten Regeln zu überprüfen.

Ein verbindlicher Regelkatalog für die Erstellung von Items kann in Zukunft dazu beitragen, dass die Konstruktion von Items erleichtert wird und eine gewisse Sicherheit dabei gegeben ist. So ist es denkbar, den automatisierten Min-Max-Ansatz (Hornke et al., 2004) mit Hilfe eines Itemgenerators auch auf die Untertests des AID 2 anzuwenden.

11 Zusammenfassung

Diese Arbeit hatte das Ziel die Untertests „Alltagswissen“, „Realitätssicherheit“ und „Angewandtes Rechnen“ des Adaptiven Intelligenz Diagnostikum 2 (Version 2.2) von Kubinger (2009a) zu einer Gruppentestversion bzw. Computertestversion zu entwickeln.

Nach intensiver Literaturrecherche und der Analyse der Originalitems des AID 2 erfolgte die Generierung neuer Items.

Für den Untertest „Alltagswissen“ wurde die Theorie aufgestellt, dass sich die Schwierigkeitsstufen ähnlich der ökologischen Systemtheorie von Urie Bronfenbrenner verhalten (Berk, 2005). Also dass je früher ein Kind eine Erfahrung macht, diese leichter zugänglich ist.

Der Untertest „Realitätssicherheit“ erwies sich als recht schwierig für die Konstruktion neuer Items. So wurden Ideen für neue Items gesammelt und diese umgesetzt. Wichtig dabei war, der Definition des Untertests gerecht zu werden, also darauf zu achten dass etwas Funktionales fehlte. Symmetrien sollten dabei nicht verwendet werden.

Beim „Angewandten Rechnen“ wurden die Annahmen getroffen, dass addieren und subtrahieren früher erlernt wird als multiplizieren und dividieren (Greeno, 1991) und dass höhere Malreihen schwieriger sind (Rollett, 1997).

Die Vorgabe der neu entwickelten Items erfolgte in vier oberösterreichischen Schulen an 679 Kindern zwischen 8;00 und 15;11 Jahren. Die Überprüfung des Itempools erfolgte mit dem Likelihood-Quotienten-Test nach Andersen (1973) auf Konformität mit dem dichotom-logistischen Modell nach Rasch. Die Teilungskriterien dabei waren Median („Intern“), Geschlecht und Region (Stadt – Land).

Zur Konformität mit dem Rasch-Modell ist zu sagen, dass sämtliche Testformen des Untertests „Alltagswissen“ nicht dem Modell entsprechen. Lediglich Testform A zeigte in den Teilungskriterien „Intern“ und „Region“ eine Passung an das Modell. Auch nach Ausschluss von 20% der Items konnte diese in keine der Testformen erreicht werden. Da eine größere Prozentzahl hier nicht sinnvoll ist, nachdem die einzelnen Testformen von Haus aus nur je 15 Items aufwiesen und

bei einer höheren Prozentzahl die Kalibrierung gelitten hätte, mussten die H0 weiterhin verworfen bleiben.

Anders schnitt der Untertest „Angewandtes Rechnen“ ab. Hier konnte für drei der vier Testformen eine Rasch-Modell-Konformität erreicht werden.

Bei Entwicklung neuer Items des Untertests „Alltagswissen“ sollten diese vor allem in Bezug auf Geschlechtsneutralität untersucht werden. Beim „Angewandten Rechnen“ könnte der jeweilige Zahlenraum der Stufen 1 und 2 besser ausgeschöpft werden.

12 Literaturverzeichnis

- Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (2001). *I-S-T 2000 R. Intelligenz-Struktur-Test 2000 R*. Göttingen: Hogrefe.
- Andersen, E.B. (1973). A goodness of fit test for the Rasch model. *Psychometrika*, 38(1), 123-140.
- Arendasy, M. & Sommer, M. (2004). *AGen: Ein Itemgenerator zur Herstellung algebraischer Textaufgaben*. Wien: Eigenverlag.
- Aster, M.v., Neubauer, A. & Horn, R. (Hrsg.). (2006). *WIE. Wechsler Intelligenztest für Erwachsene*. Deutschsprachige Bearbeitung und Adaptation des WAIS-III von David Wechsler. Frankfurt/M.: Harcourt.
- Berk, L. E. (2005). *Entwicklungspsychologie*. München [u.a.]: Pearson Studium.
- Böck, J. (2010). *AID 2 als Gruppentest?: Eine Machbarkeitsstudie*. Dipl. Arb. Univ. Wien, Wien.
- Bronfenbrenner, U. (1986). Ecology of the family as a context for human development: research perspectives. *Developmental Psychology*. 22(6). 723-742.
- Frey, A. (2006). *Validitätssteigerung durch adaptives Testen*. Frankfurt (Main): Peter Lang GmbH.
- Greeno, J.G. (1991). Number sense as situated knowing in a conceptual domain. *Journal for Research in Mathematics Education*, 22, 170-218.
- Greeno, J.G., Moore, J.L. & Smith, D.R. (1993). Transfer of situated learning. In Dettermann, D.K. & Sternberg, R.J. (Eds.). *Transfer on trial: Intelligence, cognition, and instruction*. (pp. 99-167). Wetsport, CT, US: Ablex Publishing.
- Hornke, L.F. (2001). Benötigte Itemanzahlen beim meß- und entscheidungsorientierten adaptiven Testen. *Zeitschrift für Differentielle und Diagnostische Psychologie*. 22(3), 185-193.
- Hornke, L.F., Arendasy, M., Sommer, M., Häusler, J., Wagner-Menghin, M., Gittler, G., Bognar, B., & Wenzl, M. (2004). *Intelligenz-Struktur-Batterie (INSBAT): Eine Testbatterie zur Messung von Intelligenz [Intelligence Structure Battery (INSBAT): A test battery for the measurement of intelligence]*. Mödling: Dr. G. Schuhfried GmbH.
- Hornke, L.F., Etzel, S. & Rettig, K. (1999) *Adaptiver Matrizen Test. AMT [Computerprogramm mit Manual]*. Mödling: Schuhfried.
- Irvine, S.H. (2002). The foundation of Item Generation for Mass Testing. In S.H. Irvine & P.C. Kyllonen (Eds.). *Item Generation for Test Development* (p. 3-34). Mahwah, NJ: Lawrence Erlbaum Associates.

- Kagan, J. & Klein, R.E. (1973). Cross-Cultural Perspectives on Early Development. *American Psychologist*. 28,947-961.
- Kubinger, K.D. (1989). Aktueller Stand und kritische Würdigung der Probabilistischen Testtheorie. In K.D. Kubinger (Hrsg.), *Moderne Testtheorie - Ein Abriß samt neuesten Beiträgen* (2. Aufl., S. 19-83). Weinheim: Beltz.
- Kubinger, K.D. (2009a). *Adaptives Intelligenz Diagnostikum – Version 2.2 (AID 2) samt AID 2-Türkisch*. Göttingen: Beltz.
- Kubinger, K.D. (2009b). *Psychologische Diagnostik – Theorie und Praxis psychologischen Diagnostizierens* (2., überarb. u. erw. Aufl.). Göttingen: Hogrefe.
- Kubinger, K.D. & Draxler, C. (2007). Probleme bei der Testkonstruktion nach dem Rasch-Modell. *Diagnostica*, 53 (3), 131-143.
- Kubinger, K.D. & Wild, B. (1989). Die Optimierung der Meßgenauigkeit beim „branched“-adaptiven Testen. In K.D. Kubinger (Hrsg.), *Moderne Testtheorie - Ein Abriß samt neuesten Beiträgen* (2. Aufl., S. 187-218). Weinheim: Beltz.
- Kubinger, K.D. & Wurst, B. (2000). *AID 2. Adaptives Intelligenz Diagnostikum 2*. Göttingen: Beltz.
- Lienert, G.A. & Raatz, U. (1994). *Testaufbau und Testanalyse* (5., vollst. überarb. u. erw. Auflage). Weinheim: Beltz.
- Melchers, P. & Preuß, U. (2009). *Kaufman Assessment Battery for Children* (deutsche Version) (8., unveränd. Aufl.). Frankfurt/M.: Pearson Assessment.
- Petermann, F. & Petermann, U. (2008). *HAWIK-IV. Hamburg-Wechsler-Intelligenztest für Kinder - IV*. Übersetzung und Adaptation der WISC-IV von David Wechsler (2., ergänzte Auflage). Bern: Huber.
- Rollett, B. (1997). *Lernen und Lehren: eine Einführung in die Pädagogische Psychologie und ihre entwicklungspsychologischen Grundlagen*. Wien: WUV-Universitätsverlag
- Wagner-Menghin, M. (2004). *Manual Lexikon-Wissen Test (LEWITE)* [Manual Lexical Knowledge Test (LEWITE)]. Mödling: Dr. G. Schuhfried GmbH.
- Wild, B. (1989). Neue Simulationsstudien zur Effizienz verschiedener Parameterschätz- und Itemauswahl-Strategien beim „tailored-testing“. In K.D. Kubinger (Hrsg.), *Moderne Testtheorie - Ein Abriß samt neuesten Beiträgen* (2. Aufl., S. 19-83). Weinheim: Beltz.

13 Anhang

13.1 Anweisung für die Erstellung von Fragen für den Untertest „Alltagswissen“

Bitte erstellen Sie Fragen, die das Alltagswissen von Kindern im Alter zwischen 6 und 15 Jahren abfragen sollen. Mit Alltagswissen sind Sachkenntnisse über Inhalte gemeint, die in der heutigen Gesellschaft alltäglich sind.

Denken Sie dabei an verschiedene Genres wie:

Biologie (Mensch/Tier/Pflanzen, Natur,...)

Geographie

Geschichte

Politik

Wirtschaft

Verkehr

Physik

Musik

Kunst

Sport

Literatur

Technologie

Berühmte Persönlichkeiten (reale & fiktive)

u.s.w.

Ihrer Ideenvielfalt sind dabei keine Grenzen gesetzt, bedenken Sie aber das Alter der Kinder und dass mit den Fragen **nicht** das reine angeeignete Schulwissen abgefragt werden soll, sondern die alltäglichen Kenntnisse aus der Gesellschaft.

Bedenken Sie auch die verschiedenen Systeme (Mikro, Meso, Exo, Makro) bei Ihrer Formulierung. Das heißt, entspricht die Frage dem System, in dem sich das Kind gerade befindet? (Mikrosystem: Schule, Familie, Mesosystem: Freunde, weite Familie, Exosystem: Systeme der Schule, Familie, Markosystem: Kulturelles System, Politisches System)

Dabei sollen **keine** Daten abgefragt werden.

Als erste Überprüfung der Gültigkeit der Frage stellen Sie bitte fest, ob die entsprechende Information bei Wikipedia enthalten ist.

Zusätzlich dient Google als Überprüfung der Schwierigkeit der Items – nur nach deutschen Seiten suchen:

Dazu geben Sie bitte alle erforderlichen Stichwörter ein, notieren diese und die Anzahl der Stichwörter.

z.B. „männliche Stimmlage“ – 656.000

Als Hilfe sei Ihnen hier eine Gruppierung der Altersgruppen gegeben:

1. Stufe - Alter: 6-7 Jahre
2. Stufe - Alter: 8-9 Jahre
3. Stufe - Alter: 10-11 Jahre
4. Stufe - Alter: 12-15 Jahre

Verwenden Sie nicht mehr als zwei Sätze, Zusatzinformationen können gegeben werden.

Bsp.: Der Orang-Utan ist ein Menschenaffe. Welcher andere Affe ist auch ein Menschenaffe? (mit Zusatzinfo)

oder: Nenne einen Menschenaffen? (ohne Zusatzinfo)

13.2 Analyse der Items des Untertests 1 „Alltagswissen“

Block Nr.	Item Nr.	Anzahl der Wörter	Sätze	Einleitungssatz (ja/nein)	Zusatzinfo	Rubrik	Nähe zum Kind (zu Hause - nahe Umgebung - ferne Umgebung - gelernt)	Antworten richtig (1 oder mehrere)
1	1	4	1	n	n	Mensch/Biologie	zh	1
1	2	4	1	n	n	Mensch/Biologie	zh	1
1	3	5	1	n	n	Tier/Biologie	zh/nu	1
1	4	4	1	n	n	Tier/Biologie	zh/nu	1
1	5	6	1	n	n	Alltagswissen/Schule	nu	mehrere
2	1	3	1	n	n	Mensch/Biologie	zh	1
2	2	7	1	n	j (Märchen)	Alltagswissen/Deutsch/Märchen/Bücher	zh/nu	1
2	3	5	1	n	n	Alltagswissen/Wirtschaft	nu	mehrere
2	4	5	1	n	j (Tier)	Tier/Biologie	nu/fu	1
2	5	5	1	n	n	Alltagswissen/Biologie	zh/nu/fu	mehrere
3	1	8	1	n	j (Tier)	Tier/Biologie/Alltagswissen	nu/fu	mehrere
3	2	5	1	n	n	Sachunterricht/Physik?	fu/g	mehrere
3	3	8	1	n	j (Nagetier)	Tier/Biologie	fu/g	1
3	4	11	2	j	j (Orang-Utan)	Tier/Biologie	fu/g	mehrere
3	5	5	1	n	n	Alltagswissen	zh/g	1
4	1	3	1	n	n	Natur/Biologie	nu/g	mehrere
4	2	6	1	n	n	Alltagswissen/Religion	zh/g	1
4	3	12	2	j	n	Geschichte	g	1
4	4	12	2	j	j (Laufen, Radfahren)	Alltagswissen/Sport	g	1
4	5	10	2	j	j (Geige, Kontrabass)	Musik	zh/nu/g	mehrere

5	1	14	2	j	j (Deutschland, Schweiz)	Geografie	fu/g	mehrere
5	2	5	1	n	n	Geografie/Wirtschaft?	g	1
5	3	9	1	n	n	Geschichte	g	1
5	4	5	1	n	n	Geschichte	g	mehrere
5	5	9	2	j	j (Sopran, Alt)	Musik	nu/fu/g	mehrere
6	1	7	1	n	j (Land)	Geografie	fu/g	1
6	2	6	1	n	j (Amt)	Alltagswissen/Geschichte/Wirtschaft?	fu/g	1
6	3	5	1	n	n	Bildnerische Erziehung/Geschichte	g	mehrere
6	4	6	1	n	n	Geschichte	g	mehrere
6	5	3	1	n	n	Geschichte/Religion?	zh/nu/g	mehrere
7	1	6	1	n	n	Geschichte/Alltagswissen/Politik	fu/g	1
7	2	9	1	n	n	Geschichte	g	1
7	3	4	1	n	n	Geschichte/Alltagswissen/Politik	g	1
7	4	7	1	n	n	Geografie/Alltagswissen/Sport	fu/g	1
7	5	3	1	n	n	Geschichte/Politik	g	mehrere
8	1	8	1	n	j (Meer)	Geografie	fu/g	1
8	2	5	1	n	j (Insekt)	Biologie	g	1
8	3	15	2	j	j	Geschichte	g	1
					(Gegenreformation, Absolutismus)			
8	4	12	1	n	n	Geschichte/Alltagswissen/Politik	fu/g	1
8	5	6	1	n	j (Roman)	Deutsch/Literatur	zh/nu/g	1
9	1					= 1-4		
9	2					= 1-5		
9	3	4	1	n	n	Biologie/Alltagswissen	nu/fu/g	mehrere
9	4	6	1	n	n	Sachunterricht/Alltagswissen	zh	1
9	5					= 1/1		
10	1	9	2	j	j (Monat)	Sachunterricht	zh	1
10	2	13	1	n	n	Alltagswissen/Technik	nu/g	mehrere
10	3	6	1	n	j (Tier)	Alltagswissen/Biologie	fu/g	mehrere
10	4	7	1	n	n	Physik/Sachunterricht	g	1
10	5	6	1	n	n	Sachunterricht/Alltagswissen	zh	1
11	1				n	= 9/1		

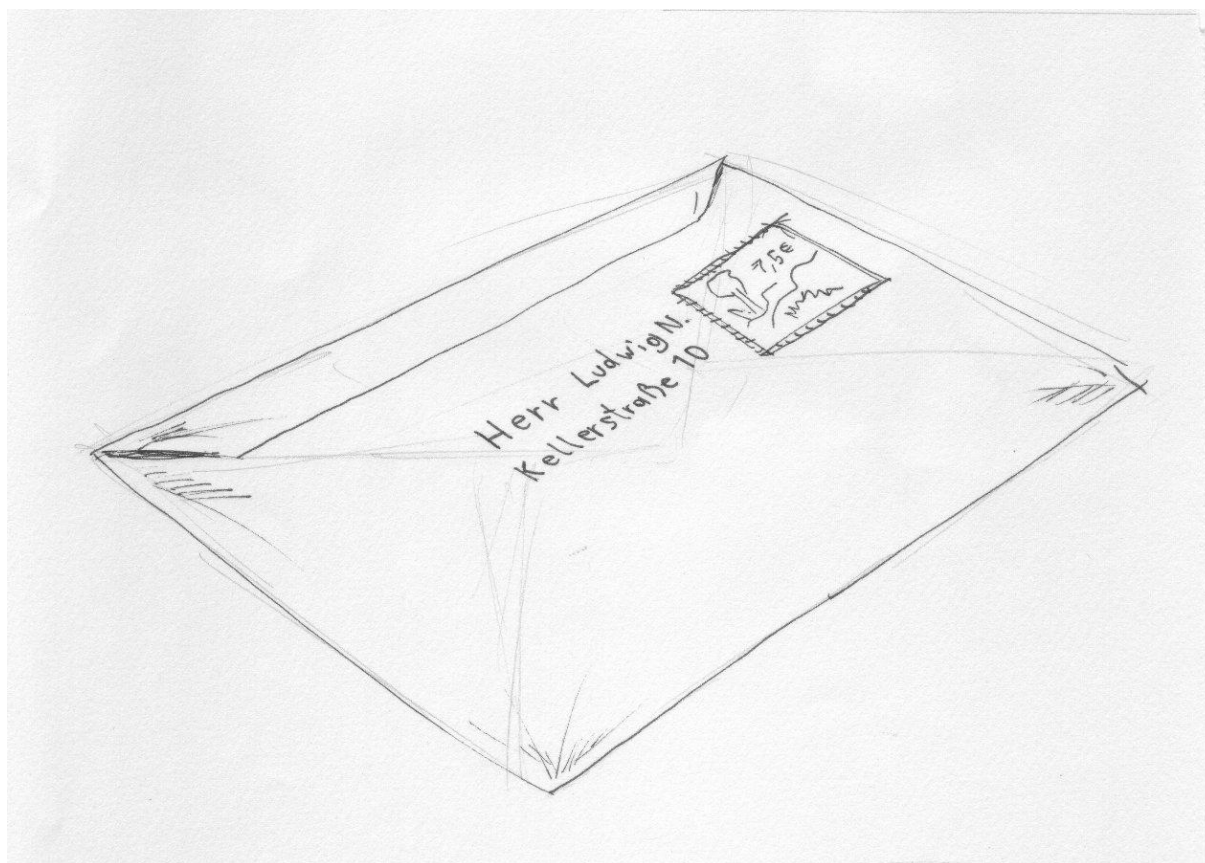
11	2				n	= 9/4		
11	3	7	1	n	n	Biologie	g?	mehrere
11	4	6	1	n	n	Geografie	fu/g	1
11	5	5	1	n	n	Bildnerische Erziehung/Geschichte	g	mehrere
12	1	9	2	j	j (Europa, Amerika)	Geografie	fu/g	mehrere
12	2	9	1	n	j (Land)	Geschichte	g	1
12	3	8	2	j	j (Schmuck)	Alltagswissen/?	g	mehrere
12	4	5	1	n	j (Tier)	Biologie/Geografie/Alltagswissen	g	1
12	5	9	1	n	j (menschlicher Körper)	Biologie	g	mehrere
13	1	5	1	n	j (Drama)	Deutsch/Literatur	zh/nu/fu/g	1
13	2	9	2	j	j (Romanik, Gotik)	Geschichte	g	mehrere
13	3	7	1	n	n	Geschichte	g	mehrere
13	4	4	1	n	n	Geografie	fu/g	mehrere
13	5	6	1	n	n	Religion/Geschichte/Alltagswissen	nu/fu/g	mehrere
14	1					= 12/3		
14	2					= 12/5		
14	3					= 8/1		
14	4					= 8/3		
14	5					= 8/2		
15	1					= 2/5		
15	2					= 2/1		
15	3					= 2/2		
15	4					= 11/4		
15	5					= 9/3		
16	1					= 10/5		
16	2					= 10/1		
16	3					= 12/1		
16	4					= 12/4		
16	5					= 10/4		
17	1					= 11/2		
17	2					= 11/5		

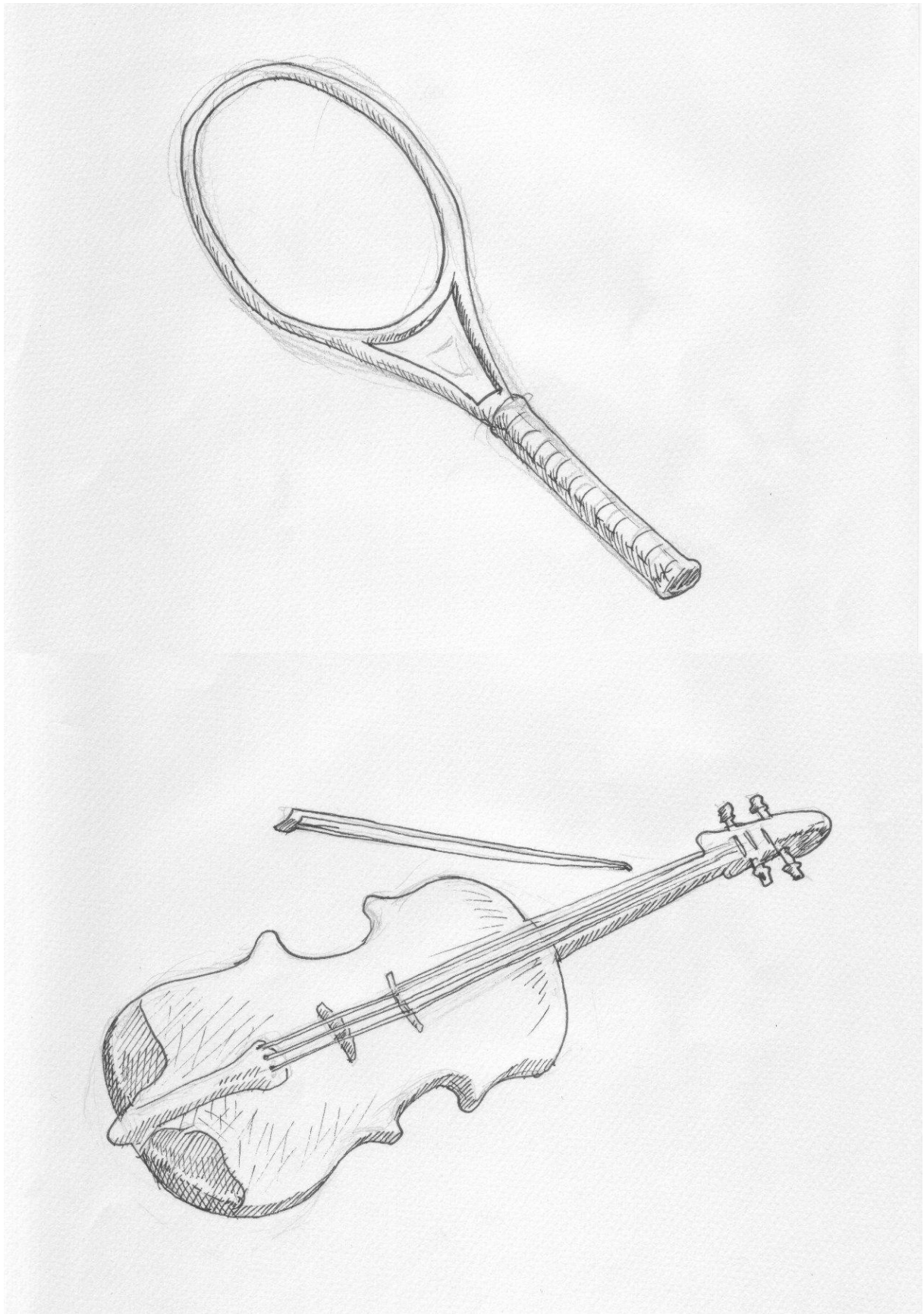
17	3	= 11/4
17	4	= 13/5
17	5	= 13/2
18	1	= 12/3
18	2	= 12/4
18	3	= 12/5
18	4	= 7/1
18	5	= 7/2

13.3 Die gezeichneten Bilder des Untertest 2 „Realitätssicherheit“











13.4 Analyse der Items des Untertest 3 „Angewandtes Rechnen“

Block Nr.	Item Nr.	Anzahl der Sätze	HS/NS	Vorkommende Namen (und deren Silben)	Richtung	Art der Operationen	Inhaltliche Komponenten	Höchste angegebene Zahl	Zusatzwissen (Zeit, Umfang,...)
1	1	2	3HS		-->	Addition	Knopf von Jacke abgerissen	1	
1	2	3	0		-->	Subtraktion	Kerzen ausblasen	4	
1	3	2	0	Mutter (2)	-->	Multiplikation	Birne durchschneiden	1	
1	4	3	0		-->	Subtraktion	Marienkäfer machen Ausflug	5	
1	5	3	0		-->	Subtraktion	Banane von Teller nehmen	1	
2	1	3	0	Peter (2)	-->	Subtraktion	Schokolade herschenken	3	
2	2	3	0		-->	Subtraktion	Mäuse werden von Katze gefressen	7	
2	3	3	0	Daniel (3), Martina (3)		1m = 1m	Größenvergleich	1	
2	4	2	3HS	Ulrich (2), Thomas (2)	-->	Multiplikation	Weitsprung	1	
2	5	3	4HS	Inge (2), Anna (2)	-->	Subtraktion	Stofftiersammlung	5	
3	1	2	3HS		-->	Addition	Bücher vom Regal nehmen	2	
3	2	3	?		-->	Subtraktion	Rosen verschenken	3	
3	3	2	3HS	Lotte (2), Karl (1)	<--	Subtraktion	Bücher vergleichen	11	
3	4	4	0		-->	Subtraktion, Division	Autoreifen werden	1000	

							ausgefahren		
3	5	3	4HS	Bernadette (3), Edith (2)	<--	Subtraktion	Gehäkelte Schlangen	1m 80cm	Längen
4	1	3	0		-->	Division	Torte wird aufgeteilt	12	
4	2	4	3HS, 1 Satzglied	Petra (2), Karl (1)	-->	Multiplikation	Über Wiese laufen	10	
4	3	3	4HS		-->	Subtraktion	Blumen zählen	30	
4	4	3	4HS	Maximilian (5)	-->	Subtraktion, Multiplikation	Sparen	5000	
4	5	4	0		<--	Division	Tabletten schlucken	60	
5	1	3	0		-->	Division	Lastwagen - Sportwagen vergleich	40	Zeit
5	2	4	0	Gerda (2)	-->, <--	Multiplikation, Subtraktion	Lose kaufen	20	
5	3	4	0	Karin (2), Nina (2)	-->	Addition	Karten kaufen	3	
5	4	3	0		-->		Gips Brei anrühren	3	
5	5	3	2		-->	Addition, Multiplikation	Schnur um Beet spannen	2	Umfang
6	1	3	0			Division, (Multiplikation)	Butter für Jause	500	Schlussrechnen
6	2	2	2		-->	Addition	Zeit für Hausaufgaben	25 bzw. halbe Stunde	Zeit umrechnen
6	3	3	1	Peter (2)	-->	Division, Multiplikation	Taschengeld	60	mit Brüchen rechnen

6	4	3	3?		-->	Multiplikation, Subtraktion	Papierschlange basteln	50	Vorstellungsvermög en
6	5	2	3HS		-->	Multiplikation, Addition	Zahlungen für Strom und Wasser	200	Zeitwissen
7	1	3	2		gm	Multiplikation, (Division), Addition	Flaschen mit Saft füllen	3	Bruchwissen
7	2	4	0		gm	Multiplikation, Division	Bretter für Boden	30	Fläche ausrechnen, Schlussrechnen
7	3	4	0	Stefan (2)	-->				
7	4	3	0		gm	Division, Multiplikation	Gehaltserhöhung	600	Schlussrechnen, Prozentrechnen
7	5	2	0		gm	Subtraktion, Division	Aufteilung der Klasse	30	Algebra
8	1	4	0		gm	Division, Multiplikation, (Subtraktion)	Läufer laufen um die Wette	100	Längenangaben umrechnen
8	2	3	1		gm	Multiplikation, Division	Metallplatte herstellen	20,14	Fläche ausrechnen, Prozentrechnen
8	3	2	0		-->	Multiplikation, Addition	Alkoholgehalt mischen	50	Prozentrechnen,
8	4	2	3HS	Elisabeth (4)	gm	Subtraktion	Alter von Vater und Tochter	38	
8	5	3	1		-->	Division	Blumenfrau bindet Sträuße	50	
9	1	$= 1/4$							
9	2	3	0		-->	Subtraktion	Banane essen	1	
9	3	3	1	Monika (3)	?		Apfel essen	1	
9	4	2	3HS, 1NS		-->	Addition, Division	Dinosaurier teilen	5	
9	5	3	5HS		-->	Addition	Kinder auf Spielplatz	2	
10	1	3	0		-->	Subtraktion	Kegel	5	

10	2	3	0		-->	Division	umschmeißen Orangen teilen	6	
10	3	= 5/1							
10	4	3	4HS		-->	Subtraktion, Addition	Häuser werden abgerissen und gebaut	9	
10	5	3	0		-->	Division	Wasser von Aquarium wird ausgeschüttet	20	
11	1	=9/4							
11	2	2	0		-->	Multiplikation	Uhr geht nach	3	Wie viel Tage hat die Woche?
11	3	1	1		gm	Addition, umrechnen	Zeit vergangen	100	Umrechnen der Zeit
11	4	2	3HS		gm	Addition	Thermometer anzeige	3	Minusbereich,
11	5	3	2		gm	Division, (Multiplikation , Subtraktion	Kastanien für Männchen	30	
12	1	2	3HS	Rainer (2), Lukas (2)	?	Subtraktion	Altersvergleich	8	Wie viele Monate hat ein Jahr?
12	2	3	5HS		gm	Addition, Division, Multiplikation	Kastanien sammeln	3	
12	3	4	0		gm	Multiplikation, Subtraktion	Lehrer teilt Hefte aus	70	
12	4	3	1		gm nahe - ->	Division, Multiplikation, Addition	Platz für Personen im Zug	50	
12	5	3	4HS		gm nahe - ->	Multiplikation, Subtraktion, Addition	Plätze für Theater	20	
13	1	2	0		<--	Division	Preis für Waschmittel	50	

13	2	2	1	Renate (3)	-->	Addition, Subtraktion	Kilo abnehmen	4	
13	3	3	0	Margit (2), Heinz (1)	gm	Division, Multiplikation	Taschengeld ausrechnen	30	Algebra ($2x+x=30$)
13	4	=8/1							
13	5	2			-->	Multiplikation, Addition	Geld zählen in Kasse	1000	Geldscheinwissen
14	1	= 12/5							
14	2	= 12/4							
14	3	= 8/5							
14	4	= 8/2							
14	5	= 8/4							
15	1	= 9/5							
15	2	= 2/5							
15	3	= 2/4							
15	4	= 11/5							
15	5	= 11/2							
16	1	= 10/4							
16	2	= 10/5							
16	3	= 5/1							
16	4	= 12/2							
16	5	= 12/3							
17	1	= 3/1							
17	2	= 11/4							
17	3	= 13/1							
17	4	= 13/3							
17	5	= 13/5							
18	1	= 12/1							
18	2	= 12/5							
18	3	= 7/1							
18	4	= 12/4							
18	5	= 7/3							

13.5 Anweisung für die Erstellung von Rechenbeispielen:

Grundsätzliches:

Verwenden Sie Inhalte, die dem Alltag der Kinder des jeweiligen Alters entstammen könnten bzw. für dieses Alter verständlich sind.

Die angegebenen Zahlenräume können beliebig verwendet werden. Ein hoher Zahlenraum meint nicht, dass keine Rechenaufgabe mit niedrigen Zahlen gegeben werden darf.

Auch 3 und 4stellige Zahlen dürfen ab der 2. Stufe vereinzelt verwendet werden, jedoch unter Berücksichtigung der anderen Vorgaben.

Zusatzwissen kann, muss aber nicht in die Aufgabe eingebaut werden.

Bedenken Sie, dass für die Lösung der Aufgabe für die ersten zwei Stufen nur 30Sek., für die anderen zwei 60 Sekunden zur Verfügung stehen.

Hilfsmittel wie Papier und Bleistift stehen den Kindern zur Verfügung, sie werden die Items dann selbst lesen und die Antwort aufschreiben. Trotzdem sollte es größtenteils möglich sein, im Kopf zu rechnen.

1. Stufe - Alter: 6-7 Jahre

Anzahl der Sätze: Bis zu 4 Hauptsätze möglich, keine Gliedsätze (eingeleitet z.B. durch: nachdem, obwohl, weil, dass, bevor, ob, damit, während, als...), generell keine Beistriche verwenden (auch nicht zum Aufzählen)

Anzahl der Operationen: Überwiegend 1, maximal 2, alle Operationsarten möglich; Divisionen: ohne Rest

Zahlenraum: bis 11

Es soll kein Zusatzwissen erforderlich sein (Zeit, Umfang, Längen,...)

Überwiegend Beispiele erstellen, bei denen schon beim Lesen mitgerechnet werden kann, die Rechenreihenfolge also der Lesereihenfolge entspricht.

2. Stufe - Alter: 8-9 Jahre

Anzahl der Sätze: 3-4 Hauptsätze, keine Gliedsätze wenn möglich, Beistriche zum Aufzählen möglich

Anzahl der Operationen: Maximal 2, alle Operationsarten möglich; Divisionen: ohne Rest

Zahlenraum: Großteils bis 30, bei höheren Werten sollten alle außer der 1. Stelle eine 0 sein.

Zusatzwissen: mit Zeit rechnen möglich (Stunden, Tage)

Rechenreihenfolge sollte überwiegend der Lesereihenfolge entsprechen

3. Stufe - Alter: 10-11 Jahre

Anzahl der Sätze: bis zu 5 Hauptsätzen, bis zu 2 Nebensätze möglich, wenn geringe HS-Anzahl

Anzahl der Operationen: bis zu 4, alle Operationen möglich; bei Divisionen: Rest möglich

Zahlenraum: Großteils bis 30, aber bis 70 möglich, höheren Werten sollten alle außer der 1. Stelle eine 0 sein.

Zusatzwissen: Zeit (Stunden, Tage, Monate), Umfang

Rechenreihenfolge kann beliebig sein

4. Stufe - Alter: 12-15 Jahre

Anzahl der Sätze: bis zu 4 Hauptsätze, bis zu 2 Nebensätze möglich

Zahlenraum: bis 50, darüber runde Zahlen verwenden, alle Operationen möglich; bei Divisionen: Rest möglich, höheren Werten sollten alle außer der 1. Stelle eine 0 sein. Kommazahlen vereinzelt verwendbar

Zusatzwissen: Schlussrechnen, Zeit, Brüche, Flächen, Prozentrechnen, Algebra, Längenangaben

Rechenreihenfolge kann beliebig sein

13.6 Anweisung für Untertest 1 „Alltagswissen“

Gleich wirst du einige Fragen zu verschiedenen Themen gestellt bekommen. Als Antwort reicht **ein einziges** Wort oder eine Zahl aus. Wenn du dir bei der Rechtschreibung nicht sicher bist, schreibe das Wort einfach so hin, wie du es für richtig hältst.

Falls du etwas nicht weißt, kannst du die Frage überspringen und zur nächsten gehen.

Nun geht's los!

13.7 Fragen Untertest 1 „Alltagswissen“ Testform A

- 1 Wie nennt man die Nase bei Pferden?
- 2 Wie viele Grundrechenarten gibt es?
- 3 Nenne ein Dateiformat, in dem Bilder am Computer abgespeichert werden können.
- 4 Wie heißt die Zahl über dem Bruchstrich?
- 5 Wie heißt der beste Freund der Biene Maja?
- 6 Welche Einheit misst die elektrische Spannung?
- 7 Welche Maßeinheit bezeichnet den Reinheitsgehalt von Gold?
- 8 In welchem Kontinent fließt der Nil?
- 9 Wie viele Weltkriege hat es bis jetzt gegeben?
- 10 Wer gilt allgemein als der Erfinder der Glühbirne? (nur Nachname)
- 11 Zu welcher Tierklasse (Tierart) gehört der Wal?
- 12 Den höchsten Titel, den eine Person im Mittelalter überhaupt haben konnte, war...?
- 13 Außer Wachs verbraucht die Flamme einer Kerze noch?
- 14 Welches Musikinstrument ist das größte Streichinstrument?
- 15 In welchem Monat beginnt (astronomisch) der Frühling?

13.8 Fragen Untertest 1 „Alltagswissen“ Testform B

- 1 Wie viele Beine haben Spinnentiere?
- 2 Wie viele Monate hat ein Jahr?
- 3 Wo wurde die erste Atombombe abgeworfen?
- 4 Welcher Ton wird gespielt, wenn alle Löcher der (Sopran)Blockflöte zugehalten werden?
- 5 Welches Nahrungsmittel erzeugen Bienen?
- 6 Wie wurde Ludwig XIV, König von Frankreich, noch bezeichnet?
- 7 Was ist das Endprodukt der DNA?
- 8 Aus was besteht Salz überwiegend?
- 9 Welche Ladung weisen Protonen auf?
- 10 Wofür steht in der Abkürzung „mfg“ das „f“?
- 11 Wofür steht die Abkürzung „PKW“?
- 12 Was ist das wichtigste religiöse Schriftstück im Islam?
- 13 Wie viele Chromosomen (-paare) besitzt der Mensch?
- 14 Wie heißt die Substanz, die für den Antrieb von Düsenflugzeugen verwendet wird?
- 15 Welchem antiken Schriftsteller wird das Epos „Odyssee“ zugeschrieben?

13.9 Fragen Untertest 1 „Alltagswissen“ Testform C

- 1 Wie viele Planeten hat unser Sonnensystem?
- 2 Welche Firma hat Bill Gates gegründet?
- 3 Welche Farbe entsteht, wenn man gelb und grün misch?
- 4 Wie nennt man das Ergebnis einer Addition?
- 5 Man sagt, Lucky Luke ziehe schneller als sein...?
- 6 Nenne eine Primärfarbe.
- 7 Was wird mit kg streng physikalisch gesehen bezeichnet?
- 8 Wie wird Software genannt, die frei erhältlich und voll einsatzfähig ist?
- 9 Welche Bezeichnung trägt James Bond?
- 10 Wie hieß der Rockband, bei der John Lennon und Paul McCartney Mitglieder waren?
- 11 Von welchem Interpreten stammt das Lied „We are the Champions“?
- 12 Welcher Strom fließt durch Wien?
- 13 Von wem stammt das Theaterstück „Maria Stuart“? (nur Nachname)
- 14 In welcher Stadt standen früher die „Twin Towers“ (oder auch World Trade Center genannt)?
- 15 Ein stärker und länger anhaltendes Seigen der Kurse an der Börse wird auch wie genannt?

13.10 Fragen Untertest 1 „Alltagswissen“ Testform D

- 1 Wer wird mit der Nummer „133“ in Österreich angerufen?
- 2 Bei welcher Farbe der Fußgängerampel ist das Gehen erlaubt?
- 3 Von welchem Tier stammt der Hauptanteil der Milch, die verkauft wird?
- 4 Wer war der berühmte Führer der indischen Unabhängigkeitsbewegung und Begründer des „gewaltlosen Widerstandes“ (nur Nachname)?
- 5 Auf welchem Kontinent sind Koalas beheimatet?
- 6 Was wird mit Hilfe von Windkraftanlagen erzeugt?
- 7 Wie nennt man die Krankheit, die in Folge einer Infektion mit dem HIV entstehen kann?
- 8 Wie wird ein Film genannt, in dem man nicht hören kann, was die SchauspielerInnen sprechen?
- 9 Wie wird Schmetterlingschwimmen noch genannt?
- 10 Welche Blume wird als Symbol der Liebe verwendet?
- 11 Welches Organ pumpt Blut durch den (menschlichen) Körper?
- 12 Wie wird ein Ensemble, in dem vier Musiker spielen, genannt?
- 13 Auf welcher Seite der Straße ist das Fahren mit dem Auto in Österreich erlaubt?
- 14 Mit wie viel °C wird Kochwäsche gewaschen, wenn die höchstmögliche Temperatur eingestellt wird?
- 15 Welcher Buchstabe ist der letzte des Alphabetes?

13.11 Anweisung für Untertest 3 „Angewandtes Rechnen“

Gleich wirst du einige Textbeispiele gestellt bekommen. Du kannst auf dem Blatt Nebenrechnungen rechnen, aber verwende keinen Taschenrechner, falls du einen hast. Wenn du etwas nicht weißt, kannst du das Beispiel überspringen und zum nächsten gehen.

Nun geht's los!

13.12 Fragen Untertest 3 „Angewandtes Rechnen“ Testform A

- 1 Paul hat 9 Äpfel geschenkt bekommen. Er möchte nun diese mit seinem Bruder und seiner Schwester teilen. Wie viele Äpfel bekommt jeder von ihnen?
- 2 Eine Hündin bekommt 7 Hundebabys. 4 davon werden an Familien verschenkt. Wie viele Hundebabys bleiben bei der Mutter?
- 3 Auf einem Parkplatz stehen drei rote und sieben blaue Autos. Wie viele Autos stehen insgesamt auf diesem Parkplatz?
- 4 Peter bekommt 23 Gummibärchen geschenkt. 3 davon isst er gleich auf. 10 verschenkt er an seine Schwester. Wie viele Gummibärchen hat Peter dann noch?
- 5 Martin kauft 18 Orangen und 4 Birnen. Die Hälfte der Orangen und der Birnen soll Martin zur Oma bringen. Wie viele Obststücke bleiben ihm über?
- 6 Jakob hat 10 Computerspiele. Thomas hat 22 Computerspiele. Um wie viele Spiele hat Jakob weniger als Thomas?
- 7 Hans darf nach der Schule für 1 Stunde Fußball spielen gehen. Er verspätet sich um 12 Minuten. Wie viel Zeit hat er noch zum Fußball spielen übrig?
- 8 Eine Fabrik stellt pro Monat 20 Busse her. Wie viele Monate dauert es, bis diese Fabrik 60 Busse hergestellt hat?
- 9 Eine Flasche mit 1 Liter Sirup soll im Verhältnis 1:6 (1 Liter Sirup, 6 Liter Wasser) verdünnt werden. Für wie viele Flaschen zu je 1 Liter reicht das?
- 10 Anna kann in einer Stunde 16 Rechenbeispiele lösen. Wie viele Rechenbeispiele löst sie in einer Stunde und 15 Minuten?
- 11 Ein Busfahrer beginnt mit seinem leeren Bus die vorgesehene Strecke abzufahren. Es steigen an der ersten Haltestelle 11 Menschen ein, an der zweiten acht und an der dritten 13. An der dritten Haltestelle sind aber auch vier Menschen ausgestiegen. Wie viele Menschen, einschließlich des Busfahrers, sind im Bus, wenn dieser von der dritten Haltestelle abfährt?
- 12 Martin wohnt 20km von der Schule entfernt und benötigt dafür 15 Minuten mit dem Schulbus. Alex wohnt 40km von der Schule entfernt. Wie lange benötigt er mit dem Schulbus?
- 13 Ein Fahrzeug fährt mit 80km/h 50 Minuten lang und kommt 65km weit. Ein anderes Fahrzeug fährt mit 130km/h dieselbe Strecke. Um wie viel Minuten ist dieses Fahrzeug schneller?
- 14 Sandra kauft sich einen neuen Pullover. Der ursprüngliche Preis war 48 Euro, im Schlussverkauf kostet der Pullover nur mehr 33 Euro und 60 Cent. Um wie viel Prozent war der Pullover verbilligt?

- 15 Mario hat zu Weihnachten 600 Euro von den Verwandten bekommen. Zu Ostern hat er noch $\frac{1}{3}$ des Weihnachtsgeldes übrig. Jetzt gibt er nochmals die Hälfte davon aus. Wie viel Geld hat er jetzt noch übrig?

13.13 Fragen Untertest 3 „Angewandtes Rechnen“ Testform B

- 1 Anne hat sechs Eier in einer Schachtel gekauft. Allerdings stolpert sie und fällt hin. Dabei gehen zwei Eier kaputt. Wie viele ganze Eier hat Anne jetzt noch in der Schachtel?
- 2 Johanna hat zu ihrem Geburtstag ihre drei besten Freundinnen eingeladen. Jede Freundin bringt zwei kleine Geschenke für Johanna mit. Wie viele Geschenke bekommt Johanna somit insgesamt von ihren Freundinnen?
- 3 In einem Wohnzimmer befinden sich zwei Sessel für jeweils eine Person sowie ein großes Sofa. Au dem Sofa können höchstens vier Menschen sitzen. Wie viele Menschen können insgesamt in diesem Wohnzimmer sitzen?
- 4 Ein Bauer hat vier Schweine. Vor einem Jahr hatte er noch doppelt so viele. Wie viele Schweine hatte er vor einem Jahr?
- 5 Im Kühlschrank sind zwei Packungen Eier. In jeder Packung sind sechs Eier. Der Vater nimmt von diesen Eiern neun Stück. Wie viele Eier sind jetzt noch im Kühlschrank?
- 6 Eine Tafel Schokolade besteht aus 15 Schokostücken. Diese soll zwischen 3 Kindern gleich aufgeteilt werden. Wie viele Stückchen bekommt jedes Kind?
- 7 Josef kann 2m weit springen. Er möchte eine 10m lange Strecke zurücklegen, hin und zurück. Wie oft muss er springen?
- 8 Karl hat zu Hause 30 Bücher. Nur die Hälfte dieser Bücher ist spannend. Von diesen verleiht er eines an seinen besten Freund. Wie viele spannende Bücher hat Karl jetzt noch zu Hause?
- 9 Eine sehr gefräßige Katze benötigt am Tag den Inhalt von zwei Dosen Katzenfutter. Den Inhalt wie vieler Dosen Katzenfutter frisst diese Katze in 14 Tagen?
- 10 Eine große Torte wird in der Mitte durchgeschnitten. Jede der beiden Hälften wird dann noch einmal jeweils in zwei Teile geteilt. Jedes der so entstandenen Stücke wird dann noch einmal geteilt. Dann noch ein weiteres Mal. Wie viele Stück Torte hat man nun?
- 11 Maria ist um drei Jahre jünger als ihr Bruder Peter. Ihre Mutter ist 48 Jahre alt. Sie ist dreimal so alt wie Maria. Wie alt ist Peter?
- 12 Ein Haus wird gebaut und soll mehrere Stöcke haben. Die Zimmerhöhe muss 2 Meter betragen, und die Decke muss 75cm dick sein. Wie viele Stöcke kann man maximal bauen, wenn die Höhe des Hauses ohne Dach maximal 8 Meter betragen darf?
- 13 Kathi liest in einer Stunde 9 Seiten eines Buches mehr als Nina. Zusammengezählt lesen sie 107 Seiten. Wie viele Seiten liest Kathi in einer Stunde?
- 14 Auf einem rechteckigen Grundstück mit einer Länge von 15 Meter und einer Breite von 8 Meter soll ein Haus gebaut werden. Es dürfen aber nur $\frac{2}{3}$ der Fläche für den Bau des Hauses verwendet werden. Wie groß darf die Grundfläche des Hauses maximal sein?
- 15 Du bekommst in ersten Lehrjahr 300€ pro Monat. Am Ende jeden Monats musst du 13% Steuern zahlen. Im zweiten Lehrjahr bekommst du eine Gehaltserhöhung von 20%. Wie viel Geld bekommst du im zweiten Lehrjahr nach Abzug der Steuern?

13.14 Fragen Untertest 3 „Angewandtes Rechnen“ Testform C

- 1 In einem Korb liegen 3 Äpfel. DU nimmst dir einen und isst ihn. Wie viele Äpfel sind jetzt noch im Korb?
- 2 Susi hat 3 Birnen. Sie isst eine Birne. Wie viele hat sie noch?
- 3 Eine Frau hat neun gleiche Blumen und möchte diese mit der Gießkanne gießen. In diese Gießkanne passt aber immer nur das Wasser für drei Blumen. Wie viele Mal muss die Frau die Kanne mit Wasser füllen um alle Blumen zu gießen?
- 4 Max hat 7 Murmeln. Seine Mutter schenkt ihm noch 2 Murmeln. Wie viele hat Max jetzt?
- 5 Ein Mann hat einen neuen Teich in seinen Garten angelegt. Er hat acht Goldfische gekauft und gibt diese in den Teich. Von seiner Nachbarin erhält er drei weitere geschenkt und gibt diese ebenfalls in den Teich. Wie viele Goldfische sind jetzt im Teich?
- 6 Thomas ist 1,80m groß. Judith ist 15cm kleiner. Wie groß ist Judith?
- 7 In einem Regal sind 10 Hefte und doppelt so viele Bücher. Wie viele Gegenstände sind insgesamt in diesem Regal?
- 8 Franz hat 17 Kugeln, Susi hat um 7 Kugeln mehr als Franz, und Lisa hat um 6 Kugeln mehr als Susi. Wie viele Kugeln hat Lisa?
- 9 Martina möchte ein Fahrrad kaufen und spart monatlich 14,- In wie vielen Monaten hat sie genug Geld gespart, wenn das Rad 70,- kostet?
- 10 Susi bekommt 25 Tafeln Schokolade. 4 davon schenkt sie ihrem Bruder, eine isst sie gleich auf. Jeden zweiten Tag will Susi eine Tafel Schokolade essen. Wie viele Tage wird sie benötigen, bis sie die übrigen Tafeln aufgegessen hat?
- 11 Anna hat ein neues Buch bekommen. Es hat 345 Seiten. In wie vielen Tagen kann sie das Buch fertiglesen, wenn sie 50 Seiten pro Tag liest?
- 12 Sarah hat 50,- bekommen. 60% gibt sie für neue Kleidung aus, 20% für Essen. Wie viel Geld bleibt ihr übrig?
- 13 Die 25 SchülerInnen einer Schulklasse sowie zwei LehrerInnen besuchen gemeinsam eine Theatervorstellung, für die sie nun die Eintrittskarten kaufen müssen. Der reguläre Preis für SchülerInnen beträgt 3,- und für Erwachsene 12,50. Da die SchülerInnen und die LehrerInnen aber zusammen als eine Gruppe die Karten kaufen, erhalten sie eine Ermäßigung von 20% auf den Gesamtpreis. Welcher Gesamtpreis ergibt sich?
- 14 Ein Parkettboden soll in zwei Räumen verlegt werden. Der erste Raum hat eine Länge von 5m und eine Breite von 3m, der zweite Raum hat eine Fläche von 10qm. Ein Brett des Parkettbodens ist 1,3m lang und 20cm breit. Wie viele Bretter werden für beide Zimmer benötigt?
- 15 Wenn eine Spielkonsole, wie beispielsweise ein Play Station 300 Euro kostet und um 25 Prozent verbilligt wird, wie viel kostet die Konsole dann?

13.15 Fragen Untertest 3 „Angewandtes Rechnen“ Testform D

- 1 In einer Straße gibt es 11 Häuser. In acht davon wohnen Menschen. In jedem der restlichen Häuser ist jeweils ein großes Kaufhaus. Wie viele große Kaufhäuser gibt es in dieser Straße?
- 2 Oma schenkt dir 4 Tafeln Schokolade. Du nimmst sie mit in die Schule und schenkst deinem besten Freund eine Tafel. Maria schenkst du auch eine. Wie viele bleiben übrig?
- 3 Mark hat sechs Stück Schokolade. ER möchte diese mit Martin und Stefan teilen. Alle drei Buben sollen dabei dieselbe Menge erhalten. Wie viel Stück bekommt jeder?
- 4 Eine kleine Fabrik stellt an jedem Tag sechs Autos her. Wie viele Tage dauert es daher, bis die Fabrik 30 Autos hergestellt hat?
- 5 An einem kleinen Apfelbaum hängen neuen Äpfel. Bei einem Sturm fallen zwei hinunter. Wie viele Äpfel hängen nun noch daran?
- 6 Thomas hat 10,- Taschengeld bekommen. Er kauft sich um 5,- Kaufgummis. Wie viel Geld bleibt ihm noch übrig?
- 7 Auf einem See fahren acht Ruderboote, 16 Tretboote und drei Segelboote. Wie viele Boote fahren insgesamt auf diesem See?
- 8 Lilli hat 2 Geschwister, eine kleine Schwester und einen großen Bruder. Kathrin hat einen kleinen Bruder. Sie wollen ins Kino gehen. Wie viele Kinokarten müssen sie für alle gemeinsam kaufen?
- 9 Jeder Tag besteht aus 24 Stunden. Aus wie vielen Stunden besteht ein halber Tag?
- 10 Um wie viel °C ist die Tagestemperatur gestiegen, wenn es um 5:00 morgens -10 Grad hatte und jetzt um 12:00 mittags +5 Grad?
- 11 Tom kauft sich ein Computerspiel. Es kostet 15Euro. Niklas kauft sich 2 Computerspiele jeweils um denselben Preis. Wie viel muss Niklas bezahlen?
- 12 Wie viele Minuten sind vergangen, wenn du seit 13:45 auf deine Freundin (deinen Freund) wartest und es jetzt 14:10 ist?
- 13 Herr Huber hat einen Garten, der 2 Meter breit und 3 Meter lang ist. Er kauft den angrenzenden Nachbargarten dazu, der 3 Meter breit und 4 Meter lang ist. Wie viele qm (m^2) hat sein Garten jetzt?
- 14 Sebastian kauft in einem Elektronikmarkt 3 Stück DVD's um ja 8,-, eine DVD-Box um 18,- und 2 Druckerfarbbänder zu je 21,-. Wie viel Geld erhält er zurück, wenn er mit einem 100,- Schein bezahlt?
- 15 Ein quadratisches Grundstück, bei dem die Seitenlänge 4 Meter beträgt, soll eingezäunt werden. Auf einer Seite des Grundstückes befindet sich eine Mauer. Wie viel Meter Zaun wird benötigt, um das Grundstück vollständig einzuzäunen?

13.16 Itemleichtigkeitsparameter

Angewandtes Rechnen, Testform A

Item	Item- parameter	Standard- schätzfehler	Konfidenzintervall (p = 0,95)	
beta Item1	1.572	0.345	0.895	2.249
beta Item 2	4.066	0.774	2.550	5.583
beta Item 3	3.546	0.656	2.261	4.831
beta Item 4	3.146	0.575	2.019	4.273
beta Item 5	0.896	0.288	0.331	1.461
beta Item 6	2.566	0.473	1.638	3.493
beta Item 7	0.896	0.288	0.331	1.461
beta Item 8	1.572	0.346	0.895	2.249
beta Item 9	-3.686	0.252	-4.181	-3.191
beta Item 10	-2.203	0.212	-2.619	-1.787
beta Item 11	-2.589	0.217	-3.015	-2.163
beta Item 12	0.745	0.278	0.200	1.290
beta Item 13	-4.721	0.320	-5.348	-4.093
beta Item 14	-4.899	0.337	-5.558	-4.239
beta Item 15	-0.908	0.216	-1.331	-0.485

Item	Item- parameter	Standard- schätzfehler	Konfidenzintervall (p = 0,95)	
beta Item 1	3.714	0.706	2.329	5.099
beta Item 2	2.222	0.422	1.394	3.050
beta Item 3	1.244	0.324	0.609	1.879
beta Item 4	2.222	0.422	1.394	3.050
beta Item 5	2.924	0.528	1.889	3.960
beta Item 6	2.222	0.422	1.394	3.050
beta Item 7	-1.396	0.214	-1.816	-0.975
beta Item 8	1.244	0.324	0.609	1.879
beta Item 9	1.041	0.308	0.436	1.645
beta Item 10	-1.795	0.212	-2.212	-1.379
beta Item 11	-1.632	0.213	-2.050	-1.215
beta Item 13	-4.268	0.280	-4.816	-3.720
beta Item 14	-2.740	0.220	-3.170	-2.310
beta Item 15	-5.002	0.339	-5.667	-4.338

Item	Item- parameter	Standard- schätzfehler	Konfidenzintervall (p = 0,95)	
beta Item 1	0.084	0.277	-0.459	0.628
beta Item 2	1.520	0.423	0.690	2.349
beta Item 3	0.649	0.322	0.018	1.280
beta Item 4	0.167	0.283	-0.388	0.721
beta Item 5	1.986	0.502	1.002	2.970
beta Item 7	0.167	0.283	-0.388	0.721
beta Item 8	-0.410	0.248	-0.896	0.076
beta Item 9	1.520	0.423	0.690	2.349
beta Item 10	-0.214	0.259	-0.721	0.294
beta Item 11	0.889	0.345	0.212	1.565
beta Item 12	-0.143	0.263	-0.659	0.372
beta Item 13	-2.573	0.204	-2.973	-2.174
beta Item 14	-2.541	0.204	-2.940	-2.142
beta Item 15	-1.099	0.220	-1.531	-0.668

13.17 Personenparameter

Angewandtes Rechnen, Testform A

Rohwert	Personen- parameter	Standard- schätzfehler
0	-5.75904778	NA
1	-4.72376731	1.1297484
2	-3.74166647	0.8950406
3	-2.99747169	NA
4	-2.40440237	0.7687435
5	-1.83050670	0.7494586
6	-1.27136826	0.7494694
7	-0.69555575	0.7723125
8	-0.06412435	0.8208204
9	0.66304633	0.8844551
10	1.49069933	0.9287622
11	2.36829098	0.9420247
12	3.26715696	0.9568960
13	4.21928335	1.0035487
14	5.37414014	1.1899168

Rohwert	Personen- parameter	Standard- schätzfehler
0	-4.7834080	NA
1	-4.0857809	NA
2	-3.3881537	0.8721795
3	-2.7135602	0.7826842
4	-2.1322753	0.7485986
5	-1.5759622	0.7480222
6	-0.9990195	0.7755629
7	-0.3594855	0.8263087
8	0.3670384	0.8737887
9	1.1457438	0.8860959
10	1.9369544	0.8973878
11	2.7824053	0.9495383
12	3.7726581	1.0466403
13	5.0422732	1.2415098
14	6.3773714	NA

Rohwert	Personen- parameter	Standard- schätzfehler
0	-3.14898027	NA
1	-2.67057502	NA
2	-2.19216978	NA
3	-1.71376453	0.7118444
4	-1.24801726	0.6580557
5	-0.83593858	0.6288110
6	-0.45033408	0.6153764
7	-0.07330273	0.6147168
8	0.31055077	0.6265433
9	0.71793441	0.6525273
10	1.17069952	0.6964365
11	1.70161000	0.7654353
12	2.36937069	0.8781863
13	3.33769998	1.1295571
14	4.37803988	NA

13.18 Ansuchen an Landesschulrat

Johanna Böck
Uferstraße 2
4191 Vorderweißenbach

Landesschulrat Linz
Sonnensternstraße 20
4020 Linz

Ansuchen um Unterstützung eines universitären Forschungsprojekts

Sehr geehrte Frau Wagner!

Mit diesem Schreiben möchten wir Sie um die Unterstützung bei der Durchführung eines Forschungsprojekts im Rahmen unserer Diplomarbeit (Machbarkeitsstudie zum Thema AID 2 als PC Version) an der Fakultät für Psychologie an der Universität Wien bitten.

Wir bräuchten dazu Ihre Genehmigung eine Testung im Umfang von einer Schulstunde
an folgenden Schulen durchzuführen:

- Volksschule Vorderweißenbach (3. und 4. Schulstufe)
- Hauptschule Vorderweißenbach (5.-8. Schulstufe)
- Hauptschule Bad Leonfelden (5.-8. Schulstufe)
- Petrinum Linz (5.-9. Schulstufe)

Die Beschreibung des Forschungsprojekts und des eingesetzten Tests sowie ein Elternbrief liegen bei.

Mit freundlichen Grüßen

Böck Johanna

13.19 Beschreibung des Forschungsprojekts

Ziel dieses Forschungsprojekts ist eine Abklärung der Machbarkeit einer Computerversion des AID 2.

Im Zuge unserer Diplomarbeit, welche als eine Voruntersuchung dient, soll für vier der elf Untertests abgeklärt werden, ob eine Vorgabe dieser in Form einer Gruppentestung ohne wesentliche Qualitätseinbußen z.B. Aufrechterhaltung der Rasch-Modell-Konformität machbar ist.

Bearbeitet wurden dabei folgende Untertests:

UT 1 Alltagswissen

UT 3 Angewandtes Rechnen

UT 6 Synonyme finden

UT 9 Funktionen abstrahieren

Die erstellten Items, welche jenen des Original AID 2 nachempfunden sind, würden den Testpersonen in Form einer Papier-Bleistift-Testung vorgegeben werden. Um möglichst generalisierbare Aussagen treffen zu können, benötigen wir für unsere Testung SchülerInnen der 3. und 4. VS, sowie der 1.-4. HS bzw. AHS und der 5. AHS bzw. BHS. Der Stichprobenumfang sollte dabei mindestens 800 betragen.

Sinn dieses Forschungsprojektes und einer Computerversion des AID 2 ist es, ökonomische (Vor)-untersuchungen zu ermöglichen. Durch die Vorgabe der PC-Version in Gruppen wäre es auch möglich, im Bedarfsfall ganze Schulklassen zu testen.

Eingesetzte Verfahren

Bei dem eingesetzten Verfahren handelt es sich um eine Gruppentestung im Papier-Bleistift-Format.

Wobei die vorgegebenen Aufgaben jenen des AID 2 nachempfunden, jedoch nicht mit diesem identisch, sind.

Den Testpersonen werden vier verschiedene Aufgabenstellungen vorgegeben:

- 1) Fragen zum Thema Alltagswissen werden im Multiple-Choice-Format vorgegeben und sollen von den Testpersonen beantwortet werden
- 2) Vorgabe von Textrechnungen, freies Antwortformat
- 3) Vorgabe eines Wortes von dem die richtige Definition erarbeitet werden muss
- 4) Vorgabe eines Wortes und fünf Antwortmöglichkeiten, die Testperson soll angeben, welche zwei der fünf möglichen Antworten eine zentrale Gemeinsamkeit mit dem vorgegeben Wort teilen

Zu Beginn der Testung erfolgt eine Instruktion durch die Testleiterin. Anschließend sollen die Kinder selbstständig arbeiten. Falls sich Fragen ergeben, ist die Testleiterin natürlich zu Stelle und verhält sich je nach Art der Frage entsprechend. Die Testung wird eine Schulstunde in Anspruch nehmen.

13.20 Elternbrief

Sehr geehrte Damen und Herren, liebe Eltern!

Wir wenden uns durch diesen Brief mit der Bitte an Sie, uns bei der Umsetzung eines Forschungsprojektes im Zuge unserer Diplomarbeit zu unterstützen.

Thema unserer Arbeit ist die Umsetzung des Adaptiven Intelligenzdiagnostikum 2 (AID 2 – eine in der Schulleistungsdiagnostik häufig eingesetzte Testbatterie zur Erfassung kognitiver Teilfertigkeiten) in eine Computerversion desselbigen. Eine PC-Version des AID 2 würde eine zeitsparende und ökonomische Voruntersuchung bei Kindern und die Durchführung von Gruppenuntersuchungen in Klassen ermöglichen. Außerdem eröffnet eine Vorgabe mittels Computer neue Möglichkeiten der Aufgabengestaltung.

Um die Machbarkeit einer Computerversion des AID 2 zu überprüfen, ist es notwendig durch Testungen in verschiedenen Schulen Daten zu erhalten und diese auszuwerten, wobei die Testungen im Zuge einer Schuleinheit durchgeführt werden. Die Testung beinhaltet verschiedene Aufgaben, wie das Durchführen von Rechnungen, das Definieren von Wörtern etc.

Sie können uns und die Forschungsarbeit an der Universität Wien unterstützen, wenn sie sich als bereit erklären, Ihr Kind an dieser Testung teilnehmen zu lassen. Vielen Dank im Voraus.

Mit freundlichen Grüßen

Böck Johanna und Eiter Angelika

Abschnitt zum Unterzeichnen

Ich erkläre mich bereit, dass mein/e Sohn/Tochter _____ im Zuge einer Schuleinheit an der von Studentinnen der Universität Wien durchgeführten Testung teilnimmt.

Unterschrift

13.21 Lebenslauf

Persönliche Daten

Angelika Eiter

Geb. am 11. November 1983 in Wien

Bildungsweg

seit 10/02 Diplomstudium der Psychologie an der Universität Wien

9/97 – 6/02 BAKIP Patrizigasse, Diplom- und Reifeprüfung mit
ausgezeichnetem Erfolg

9/93 – 6/97 AHS am Schulschiff „Bertha von Suttner“

9/89 – 6/93 VS in Deutsch Wagram

Fachbezogene

Berufserfahrung

7/06 – 9/ 06 12-Wochen-Pflichtpraktikum in der
gerontopsychiatrischen Abteilung des SMZ-Ost

seit 03/08 Studienassistentin an der Fakultät für Psychologie, am
Institut für Entwicklungspsychologie und Psychologische
Diagnostik