

Sprachtest
Wissenschaftlicher Bericht



now you know

Sprachtest

Wissenschaftlicher Bericht

Tom Duindam, MA
Ömer Konak, MA
Frans Kamphuis, MSc

© Cito Deutschland GmbH, Butzbach (2010)

Alle Rechte vorbehalten. Sämtliche Inhalte dieser Wissenschaftlichen Publikation sind urheberrechtlich geschützt. Aus diesem Werk darf nichts ohne die vorherige schriftliche Zustimmung der Cito Deutschland GmbH mittels Druck, Fotokopie/Reprografie, Scannen, Computersoftware, Film oder Videokopie oder auf welche Weise auch immer, entnommen, veröffentlicht und/oder vervielfältigt oder sonst wie genutzt werden.

Inhaltsverzeichnis

Teil 1: Vorstudie Sprachtest 5

- 1 Einleitung 7**

- 2 Vom Test Zweisprachigkeit zum Sprachtest 9**
 - 2.1 Historischer Zusammenhang 9
 - 2.2 Organisatorischer Zusammenhang 9
 - 2.3 Komponenten des Sprachtests 10

- 3 Die Vorstudie 13**
 - 3.1 Die Forschungsfragen 13
 - 3.2 Stichprobeentnahme und Datenzusammenstellung 13
 - 3.3 Datenanalyse 14
 - 3.4 Zuverlässigkeit und Validität 15
 - 3.5 Korrelation: innerhalb und zwischen rezeptiven und produktiven Testkomponenten 16
 - 3.6 Vergleich der Sprachfähigkeiten der türkischen Kinder in der Testversion Türkisch und Deutsch 17
 - 3.7 Internationaler Vergleich (Deutschland, Türkei und Niederlande) 17
 - 3.8 Schlussfolgerung 18

Anlagen 21

- A Impressum 21
- B Durchführung des Pretests: Erfahrungen der Testleiter 22

Teil 2: Sprachtest 23

- 1 Einleitung 25**

- 2 Aufbau und Struktur des Sprachtests 27**
 - 2.1 Die Testkomponenten 27
 - 2.2 Vergleich deutsche und türkische Testversionen 27

- 3 Die Forschung: Fragestellung 29**
 - 3.1 Die Fragestellung 29
 - 3.2 Nähere Erklärung der 5 Gruppen 30

- 4 Testdaten-Analyse 31**
 - 4.1 Angaben auf dem Eingabeformular 31
 - 4.2 Zuverlässigkeit der Testversionen Deutsch und Türkisch 37
 - 4.3 Schlussfolgerungen 38

5	Profile der Sprachentwicklung	39
5.1	Ergebnisse aller Teilgruppen in der deutschen Testversion	39
5.2	Vergleich aller Sprachgruppen in der deutschen Testversion	40
5.3	Ergebnisse der beiden türkischen Teilgruppen in der türkischen Testversion	42
5.4	Ergebnisse der bilingualen, türkisch sprechenden Kinder in der deutschen und der türkischen Testversion	43
5.5	Internationaler Vergleich von Kindern in der Türkei und in den Niederlanden in den Altersgruppen 5 beziehungsweise 6 Jahre	44
5.6	Schlussfolgerungen	45

6	Testnormierung	47
6.1	Beschreibung des Normierungsprozesses	47
6.2	Darstellung der Normierung	48
6.3	Die Meßgenauigkeit	39
6.4	Abschließende Wahrnehmungen	53

7	Die im Sprachtest erzielten Ergebnisse	55
7.1	Ergebnisse in der deutschen Testversion	55
7.2	Ergebnisse in der türkischen Testversion	57
7.3	Übersicht der Ergebnisse in beiden Testversionen	59
7.4	Abschließende Erklärungen	60

8	Schlussfolgerungen und Empfehlungen	61
8.1	Sprachen-Atlas von Duisburg	61
8.2	Sprachprofile der Kinder	61
8.3	Empfehlungen für weitere Forschung	62
8.4	Empfehlungen für die Unterrichtspraxis	62

Literaturliste	63
-----------------------	-----------

Kurzfassung von 2 Artikeln zum Sprachtest	65
--	-----------

Anlagen	69	
A	Ergebnisse aller Teilgruppen in der deutschen Testversion	69
B	Ergebnisse der beiden türkischen Teilgruppen in der türkischen Testversion	71
C	Ergebnisse in der deutschen Testversion	73
D	Ergebnisse in der türkischen Testversion	75

Teil 1

Vorstudie Sprachtest

Eine Vorstudie bezüglich der Zuverlässigkeit, Gültigkeit und Konsistenz eines Sprachfähigkeitstests für Vorschulkinder in Duisburg

1 Einleitung

Dieser Artikel stellt einen Bericht über die Vorstudie des Sprachtests dar, die im Juli 2003 in Duisburg durchgeführt wurde.

Kapitel 2 geht auf den Bedarf der Stadt Duisburg an einem Test ein, mit dem sich Aussagen über die Sprachfähigkeit von Vorschulkindern in Deutsch treffen lassen, und das zu einem Zeitpunkt unmittelbar vor der Einschulung.

Es wird die Geschichte des Tests „Zweisprachigkeit“ dargestellt, der die Grundlage für den Sprachtest bildet. Dieses Kapitel enthält auch eine Beschreibung des Entwicklungsprozesses, wie der Test „Zweisprachigkeit“ zum Sprachtest weiterentwickelt wurde. Die Testkomponenten des Sprachtests werden aufgeführt und insbesondere die zwei mündlichen „Aktiver Wortschatz“ und „Satznachahmung“ erläutert. Kapitel 3 spiegelt den eigentlichen Bericht der Vorstudie wider. Hier gehen die Autoren auf die Fragen nach Validität, Zuverlässigkeit und interner sowie externer Konsistenz beider Testversionen ein. Zuerst stehen die Forschungsfragen im Mittelpunkt. Es wird eingegangen auf Alter und Einteilung der Stichprobenpopulation sowie die Anzahl der Teilnehmer je Teilgruppe. Sowohl von der deutschen als auch der türkischen Testversion sind unter anderem der Durchschnittswert und die Standarddeviation aufgeführt. Die Forschungsergebnisse werden im Bezug auf Zuverlässigkeit und Validität betrachtet und sowohl die interne Korrelation wie auch die Korrelation zwischen den digitalisierten und den mündlichen Testkomponenten dargestellt. Die Autoren bringen anschließend die in Duisburg gesammelten Daten in Verbindung mit bereits früher erhobenen Ergebnissen in den Niederlanden und in der Türkei. Die Schlussfolgerungen bilden den Abschluss des Kapitels 3.

2 Vom Test „Zweisprachigkeit“ zum Sprachtest

Der Sprachtest hat seinen Ursprung in dem bereits existierenden, in den Niederlanden entwickelten Test „Zweisprachigkeit“. Dieses Kapitel spiegelt die Charakteristiken dieses ursprünglichen Tests wider und beschreibt kurz die Testkomponenten. Außerdem wird der Kontext dargestellt, in dem die Entwicklung zum Sprachtest stattgefunden hat.

2.1 Historischer Zusammenhang

Die Universität Tilburg und CITO entwickelten zusammen den Test „Zweisprachigkeit“ in den Jahren 1989 bis 1992. Ab 1995 gehörte dieser Test zum Schüler-Monitoring-System von CITO. Ziel dieses so genannten "Bleistift- und Papier-Tests" ist es, Art und Ausmaß der Zweisprachigkeit von Kindern im Alter von vier bis sieben Jahren festzustellen. Verschiedene Studien¹ belegen, dass der Test „Zweisprachigkeit“ in genügendem Maße zuverlässig und valide ist.

2.2 Organisatorischer Zusammenhang

Bei der Stadt Duisburg bestand 2002 der Bedarf nach einem Sprachtest, der unmittelbar vor der Einschulung der Kinder ausweisen sollte, ob sie aufgrund ihrer Fähigkeiten in der deutschen Sprache den Anforderungen des Unterrichts gewachsen sein würden. Die Stadt stellte folgende Anforderungen an den Test, den sie benötigte: Er musste wissenschaftlich präzise, wissenschaftlich valide, zuverlässig, mehrsprachig und praktisch anzuwenden sein. Im Herbst 2002 machte das Landesinstitut für Schule in Soest/Nordrhein-Westfalen die Stadt Duisburg auf den niederländischen Test „Zweisprachigkeit“ aufmerksam. Dieser Test entsprach bis auf eine allen Anforderungen. Diese fehlende Anforderungserfüllung bezog sich auf die praktische Anwendung für große Gruppen von Vorschulkindern. Diese war jedoch enorm wichtig, da in kurzer Zeit ungefähr 5.000 Kinder diesen Test absolvieren sollten. Deshalb entschloss sich Cito, den Test „Zweisprachigkeit“ weiter zu entwickeln zu einem Computertest.

Damit der Test auch die Anforderungen des deutschen Bildungskontextes erfüllt, musste er "lokalisiert" werden; mit anderen Worten: Dort wo es erforderlich war, war er vom Inhalt her anzupassen. Letzten Endes sollte der Test Gegenstand einer Vorstudie bezüglich Zuverlässigkeit, Validität und Konsistenz werden. Duisburg zeigte auch Interesse daran, die türkischsprachige Testversion einzusetzen, da die zweisprachigen türkischsprachigen Kinder dann beide Testversionen absolvieren könnten.

Ogleich sich der ursprüngliche Test „Zweisprachigkeit“ in Bezug auf Validität, Zuverlässigkeit und interne Konsistenz bereits bewiesen hatte, war zum Einsatz dieses neuen und digitalisierten Tests in Deutschland eine weitere Studie erforderlich, die das erfolgreiche Funktionieren des Tests unter Beweis stellte.

Der Test „Zweisprachigkeit“ setzt sich zusammen aus den sechs nachstehend aufgeführten Testkomponenten. Zugunsten der Lesbarkeit sind auch die Abkürzungen hinzugefügt, die ab hier in diesem Bericht zur Verwendung kommen.

Passiver Wortschatz	PW
Kognitive Begriffe	KB
Phonologisches Bewusstsein	PB
Textverständnis	TV
Aktiver Wortschatz	AW
Satznachahmung	SN

¹ In den Niederlanden (L. Verhoeven u.A. 1992, G. Narain und L. Verhoeven 1994) und in der Türkei (Ö. Konak 1996).

Von folgenden Testkomponenten wurde angenommen, dass sie gut in einem Computertest funktionieren könnten: Passiver Wortschatz, Kognitive Begriffe, Phonologie und Textverständnis. Die Entwicklungsaktivitäten wurden von einer Arbeitsgruppe durchgeführt, die sich aus Experten und Fachleuten der Stadt Duisburg, dem Landesinstitut für Schule und der Universität Essen/Duisburg zusammensetzte (s. für Details das Impressum in der Anlage). Experten von CITO begleiteten die Arbeitsgruppe fachtechnisch. In Folge einer Diskussion über Eignung einiger Begriffe in der Testkomponente „Passiver Wortschatz“ der deutschen Testversion, wurden von sieben Testfragen zwei Varianten in den Test aufgenommen. In die gleiche Testkomponente der türkischen Version wurden aus dem gleichen Grund von einer Testfrage zwei Varianten aufgenommen. Letztendlich wurden die am besten geeigneten Testfragen selektiert. Für die Testkomponente KB schienen derartige Maßnahmen nicht erforderlich.

Für die Testkomponente PB genügte "Lokalisierung" nicht. Aufgrund der Satzung, die dieser Testkomponente zugrunde liegt, wurde für Deutschland eine ganz neue Testkomponente entwickelt. Diese basiert auf den distinktiven Phonemmerkmalen der deutschen Sprache. Bei Auswahl der Wörter für diese Testkomponente ist nicht nur an Lautoppositionen (*klankopposities*) und eventuelle Lautkombinationen (*klankcombinaties*) geachtet worden, sondern es wurde auch von dem deutschen Lexikon ausgegangen. Das heißt, dass in dieser Testkomponente, genauso wenig wie in den anderen Sprachversionen des Tests, keine Nonsenswörter aufgenommen worden sind. Für die ausführliche Beschreibung der anderen Testkomponenten ist die angegebene Literatur hinzuzuziehen (L. Verhoeven, u.A. 1992).

Die Texte der Testkomponente Textverständnis beider Testversionen unterscheiden sich nur in ein paar Einzelheiten vom Original. So stammen z.B. die Eigennamen von Personen, die in den Texten vorkommen, aus den jeweiligen Sprachumgebungen.

Da die Entwicklung des ursprünglichen "Bleistift und Papier-Tests" zum Computertest notwendigerweise auch Folgen hatte für die Testabnahmen, hat die Arbeitsgruppe auch in diesem Zusammenhang Lösungen beurteilt. Mehrere Anpassungen wurden deshalb zu dieser Beurteilung durchgeführt, z.B. die Art, in der der Computer die Testfragen mit Hilfe von Abbildungen und Lauten (*geluid*) präsentiert.

Der Beitrag der Arbeitsgruppe schlug sich außerdem in der Frage nieder, ob aufgrund eines Computertests Aussagen über die Sprachfähigkeit, vor allem die produktive Sprachfähigkeit, eines Kindes gemacht werden könnten. Daraufhin wurde beschlossen, in den Test zwei Komponenten des Tests „Zweisprachigkeit“, die nicht zum eigentlichen Sprachtest gehören würden, als mündliche Testkomponenten in die Vorstudie mit einzubeziehen. Von den Testkomponenten, um die es hier ging, „Aktiver Wortschatz“ und „Satznachahmung“, sollten nur deutschsprachige Versionen entwickelt werden. Die erste Testkomponente ließ sich auf vergleichbare Art und Weise wie die Testkomponenten PW und KB "lokalisieren". Die Testkomponente Satznachahmung musste aber für die deutsche Sprache ganz neu entwickelt werden, genau wie auch die Testkomponente Phonologie.

2.3 Komponenten des Sprachtests

Der Sprachtest setzt sich aus folgenden vier digitalen Testkomponenten zusammen: Passiver Wortschatz **PW**, Kognitive Begriffe **KB**, Phonologisches Bewusstsein **PB** und Textverständnis **TV**. Im Rahmen der Vorstudie wurden außerdem noch zwei nicht-digitale Testkomponenten zugefügt: Satznachahmung **SN** und Aktiver Wortschatz **AW**.

In diesem Bericht verwenden die Autoren nun ausschließlich die Abkürzungen, die hinter den Namen der Testkomponenten genannt sind. Die Testkomponenten werden in der Reihenfolge präsentiert, in der sie auch im Test vorkommen.

Passiver Wortschatz, **PW**, 67 Items (die türkische Version dieser Testkomponente hat 61 Items): Um die Kenntnisse in Wortbedeutungen zu messen, ist die Testkomponente "Passiver Wortschatz" entwickelt worden. Sprachbedeutung wird überwiegend mit Substantiven, Verben und Adjektiven ausgedrückt. Die Testkomponente "Passiver Wortschatz" umfasst 67 Testfragen, wobei jedes Mal von vier Abbildungen die Abbildung anzuklicken ist, die die Bedeutung des Wortes widerspiegelt.

Kognitive Begriffe, KB, 65 Items:

In der Testkomponente "Kognitive Begriffe" werden die Kenntnisse von Begriffen, die sich auf Farbe, Form, Maß, Räumlichkeit und Beziehungen zwischen Ereignissen beziehen, getestet. Die deutsche und die türkische Version testen die gleichen Begriffe. Bei jeder Testfrage werden mehrere Abbildungen gezeigt und das Kind wird gebeten, die Abbildung anzuklicken, die dem genannten Begriff entspricht.

Phonologisches Bewusstsein, PB, 30 Items:

In der Testkomponente "Phonologisches Bewusstsein" wird getestet, in wiefern die Kinder in Deutsch oder Türkisch Laute erkennen und voneinander unterscheiden können. Bei jeder Testfrage sagt die Teststimme zwei Wörter, bei denen der Unterschied vom Laut her minimal ist. Zum Beispiel können Wörter sich nur in einem Laut unterscheiden (Hall - hell), die Position des Lauts kann unterschiedlich sein (Kran - Karn) oder die Anzahl der Laute ist unterschiedlich (Miete - Mitte). In dieser Testkomponente werden ausschließlich Wörter mit sinngemäßer Bedeutung angeboten. Bei jedem Wortpaar wird das Kind gebeten anzugeben, ob die Wörter vom Klang her gleich oder unterschiedlich waren. Das Kind wählt dazu aus zwei Kreisen, einem grünen für "gleich lautend" und einem roten Kreis mit Kreuz, der "unterschiedlich vom Klang her" bedeutet.

Textverständnis, TV, 20 Items zu 4 Geschichten:

Hier wird getestet, in welchem Maß Kinder in der Lage sind, konzentriert einem kurzen gesprochenen Text zuzuhören und implizite und explizite Informationen zu erfassen. Dieser Testteil gibt Auskunft darüber, ob ein Kind den Text verstanden hat und Fragen dazu beantworten kann.

Die zwei nicht-digitalen Testkomponenten

Satznachahmung, SN, 20 Items:

Diese Testkomponente gibt es nur in der deutschsprachigen Version. Für jedes Item dieser mündlichen Testkomponente können aufgrund einer Korrekturvorschrift zwei Punkte erzielt werden. Zweck dieser Aufgabe ist es, Einsicht in die Wort- und Satzstrukturen der Kinder zu gewinnen. Das Kind muss dazu einige Sätze nachsprechen, die zu lang sind, um sie nur zu imitieren. Bei der Korrektur werden Aussprachefehler oder morphologische Fehler nicht in Betracht gezogen. Es geht vor allem um die richtige Realisation von syntaktischen Elementen und Mustern.

Aktiver Wortschatz, AW, 40 Items:

Diese Testkomponente gibt es nur in der deutschsprachigen Version. Die 40 Wörter dieser Komponente umfassen 30 Substantive und 10 Verben. Zu jedem Wort gibt es eine Abbildung, die das Kind benennen muss.

3 Die Vorstudie

Kapitel 3.1 geht kurz auf den neuen Test ein und formuliert die Forschungsfragen. Das Verfahren der Stichprobenentnahme wird in 3.2 erörtert, während Kapitel 3.3 die Datenanalyse darstellt. Einige vorsichtige Schlussfolgerungen stellen das Ergebnis von Kapitel 3.4 dar.

3.1 Die Forschungsfragen

Der Pretest sollte Antworten zu den folgenden Fragen liefern:

- 1 Weist die Duisburger Testversion genügend Zuverlässigkeit und Validität auf?
 - 2 Gibt es einen Zusammenhang zwischen den digitalen und den nicht-digitalen Subtests?
 - 3 Können Höchstgrenzeffekte bei einem oder mehreren Subtests festgestellt oder erkannt werden?
- Etwas getrennt von diesen Forschungsfragen war auch von Wichtigkeit, in Erfahrung zu bringen, wie die Kinder auf den Computertest reagieren würden. CITO hat in 2002 einige Lehrkräfte nach ihren Wahrnehmungen und Bewertungen beim Projekt der digitalisierten Kindergartentests gefragt (C.Sluijter, 2002). Die Anmerkungen in dem zugehörigen Bericht gaben Grund anzunehmen, dass es möglich sein müsste, auch sehr junge Kinder einen Computertest machen zu lassen.

3.2 Stichprobeentnahme und Datenzusammenstellung

Um diese Forschungsfragen beantworten zu können, wurden in Duisburg 448 Kinder getestet. Sie kamen aus unterschiedlichen Stadtteilen und sind aufgrund eines Fragebogens als Auskunftspersonen in drei Gruppen eingeteilt worden: Kinder, die zu Hause nur deutsch sprechen, Kinder, die zu Hause auch türkisch sprechen außer oder anstatt deutsch, und Kinder, die zu Hause eine andere Sprache sprechen (nicht türkisch) außer oder anstatt deutsch. In diesem Bericht wird die erste Gruppe „deutsche Kinder“ genannt, die zweite Gruppe „türkische Kinder“ und die dritte Gruppe wird „Rest“ genannt. Insgesamt wurden die Ergebnisse von 217 „türkischen Kindern“ analysiert. Der „Rest“ setzt sich aus dreißig Kindern zusammen.

Die Duisburger Testleiter sind von CITO in Zusammenarbeit mit der Arbeitsgruppe in Duisburg dazu ausgebildet worden, die Testabnahme begleiten zu können.

Die Begleitung bei den digitalen Subtests war eindeutig: Kinder beantworteten die Fragen am Computer und die Testleiter waren darauf vorbereitet, in Problemsituationen einzugreifen. Die zwei nicht-digitalen Subtests (Aktiver Wortschatz und Satznachahmung) wurden von muttersprachlichen Testleitern abgenommen.

Bei den türkischsprachigen Kindern nahmen diese Subtests solche Testleiter ab, die türkisch sprechen. Die letzten zwei Subtests sind nur auf Deutsch abgenommen worden. Die Testleiter haben die Fragebögen bei den Eltern der Kinder abgeholt.

Das Alter der Kinder lag zwischen fünf und acht Jahren. Die älteren Kinder waren größtenteils in der türkischen Gruppe vertreten.

Tabelle 3.1

	< 6	6 – 6,5	6,5 – 7	7 – 7,5	7,5 – 8	Gesamtzahl
türkische Sprache	5	79	93	30	10	217
deutsche Sprache und der Rest	15	102	94	16	4	231
Gesamtzahl	20	181	187	46	14	448

3.3 Datenanalyse

Der deutsche Test wurde im Ganzen und auf Gruppenniveau analysiert. Die Testergebnisse der Gruppe „Rest“ sind gesondert analysiert worden. Da die Teilnehmerzahl der Gruppe „Rest“ so gering war, wurden aus diesen Daten keine Schlussfolgerungen gezogen. Etwa 155 „deutsche Kinder“ haben den digitalen Test absolviert. Bei etwa 70 „türkischen Kindern“ ist die deutsche und die türkische Testversion abgenommen worden. 90 „türkische Kinder“ haben nur den türkischen Test gemacht, während ca. 80 „türkischen Kindern“ nur die deutsche Testversion angeboten worden ist. Genaue Teilnehmerzahlen stehen in den nachfolgenden Tabellen. Die Testergebnisse der „türkischen Kinder“, die beide Testversionen abgelegt haben, sind gesondert analysiert worden (Tabelle 3.2 und 3.3).

Für beide Testversionen wurde zuerst eine Ergebniseinteilung festgelegt. Dazu wurden die Durchschnittswerte und Standarddeviationen berechnet. Daraufhin wurden im Hinblick auf die Qualitätskontrolle, die interne Konsistenz und die Homogenität der Testaufgaben erörtert.

Zuerst erfolgte auf Itemniveau eine Berechnung der Item-Rest-Korrelationen (RIR). Als ungefähres Maß für den Test wurde Cronbachs Alfa bestimmt. Außerdem wurden die Frequenzen der Testergebnisse näher überprüft, um herauszufinden, ob Höchstgrenzeffekte eintraten.

Validitätsaspekte ließen sich beurteilen, indem man die Korrelation zwischen allen Subtests (digitalen und nicht-digitalen) in Betracht zog (Konkurrenz'-Validität). Schließlich sind die Testresultate in der Türkei und den Niederlanden zu Rate gezogen worden, um so Indizien auf die Frage zu haben, ob die Testresultate in verschiedenen Ländern oder bei verschiedenen Auskunftspersonen vergleichbar sind.

Tabelle 3.2 Durchschnittswerte und Standarddeviationen der „türkischen Kinder“ im Sprachtest Türkisch

Subtest	Anzahl Items	N	Durchschnittswerte	Standarddeviationen
PW	61	71	45.35	6.67
KB	65	67	44.97	7.44
PB	30	69	21.43	6.79
TV	20	66	13.54	2.95

Signifikant ist die weitläufige Verteilung der Ergebnisse beim Subtest KB.

Tabelle 3.3 Durchschnittswerte und Standarddeviationen der „türkischen Kinder“ im Sprachtest Deutsch

Subtest	Anzahl Items	N	Durchschnittswerte	Standarddeviationen
PW	67	71	44.37	7.64
KB	65	67	47.68	5.67
PB	30	69	21.16	6.70
TV	20	66	12.07	3.51
AW	40	29	22.68	6.16
SN	40	29	27.16	6.62

Die Tabellen 3.3 und 3.4 zeigen, dass die Resultate der „türkischen Kinder“ sich bei allen Subtests als niedrig herausstellen im Vergleich zu denen der „deutschen Kinder“.

Tabelle 3.4 Durchschnittswerte und Standarddeviationen der "deutschen Kinder" im CITO-Sprachtest Deutsch

Subtest	Anzahl Items	N	Durchschnittswerte	Standarddeviationen
PW (67)	67	154	55.54	6.64
KB (65)	65	154	54.65	6.61
PB (30)	30	160	23.77	6.46
TV (20)	20	149	15.31	3.95
AW(40)	40	79	32.32	5.55
SN (40)	40	80	34.32	7.99

3.4 Zuverlässigkeit und Validität

Interne Konsistenz

Im Hinblick auf die Qualitätskontrolle des Sprachtests wurde zuerst die interne Konsistenz jedes Subtests untersucht. Dazu berechnete man den Koeffizient Alpha und die durchschnittliche Item-Rest-Korrelation. Die Resultate dieser Berechnung für den türkischen Test stehen in Tabelle 3.5. In den einzelnen Subtests haben nur einige Aufgaben niedrige RIR-Werte. In all diesen Fällen trat gleichzeitig ein extrem niedriger P-Wert auf.

Tabelle 3.5 Cronbachs Alpha's und durchschnittliche Item-Rest-Korrelationen der „türkischen Kinder“ im Sprachtest Türkisch und im Sprachtest Deutsch

Test	Gruppe	Kinder	Items	P-Wert	Alpha	Durchschn. RIR
		(N)	(N)			
PW-d	T	71	67	65.90	.84	.26
PW-t	T	71	61	69.10	.66	.19
KB-d	T	69	65	72.90	.78	.21
KB-t	T	69	65	70.80	.83	.26
PB-d	T	67	30	69.80	.88	.42
PB-t	T	67	30	72.60	.91	.47
TV-d	T	66	20	59.00	.74	.31
TV-t	T	66	20	68.90	.75	.32

Die Resultate dieser Berechnung für den türkischen und deutschen Sprachtest stehen in Tabelle 3.6. In den einzelnen Subtests kommen nur einige Aufgaben vor mit niedrigem RIR-Wert in Kombination mit extrem niedrigen P-Werten.

Tabelle 3.6 Cronbachs Alpha's und durchschnittliche Item-Rest-Korrelationen für alle „türkischen Kinder“ und „deutschen Kinder“ im Sprachtest Türkisch und im Sprachtest Deutsch

Subtest	Gruppe	Kinder	Items	P-Wert	Alpha	Durchschn. RIR
		(N)	(N)			
PW-d	D	154	67	82.70	.84	.29
PW-d	T	163	67	63.20	.83	.26
PW-t	T	148	61	68.20	.69	.19
KB-d	D	154	65	83.80	.84	.28
KB-d	T	162	65	70.00	.80	.23
KB-t	T	145	65	68.80	.82	.25
PB-d	D	151	30	80.90	.90	.48
PB-d	T	160	30	64.80	.87	.40
PB-t	T	143	30	67.80	.87	.40
TV-d	D	149	20	77.40	.81	.38
TV-d	T	154	20	57.70	.66	.25
TV-t	T	140	20	64.10	.71	.28

3.5 Korrelation: innerhalb und zwischen rezeptiven und produktiven Testkomponenten

„Konkurrenz“-Validität

Um eine Sicht auf die ‚Konkurrenz‘-Validität zu erlangen sind zuerst die Korrelationen zwischen den Subtests bewertet worden.

Tabelle 3.7 Korrelationen zwischen den Subtests Türkisch

PW (N:71)	---	.401**	.793**
KB (N:69)	.934**	.484**	.821**
TV (N:66)	.793**	.677**	----
Subtest	PW	PB (N:67)	TV

In Tabelle 3.7 steht die Übersicht der Korrelationen zwischen den Subtests Türkisch. In allen Fällen zeigen sich signifikante Korrelationen in den Subtests Türkisch (P<01).

Tabelle 3.8 Korrelationen zwischen den Subtests Deutsch

PW (N:154)	----	.723**	.869**	.749**	.507**
KB (N:154)	.925**	.696**	.812**	.708**	.693**
TV (N:149)	.869**	.705**	----	.713**	.576**
AW (N:80)	.749**	.499**	.713**	----	.499**
SN (N:79) (mündlich)	.507**	.463**	.576**	.499**	----
Subtest	PW	PB (: 151)	TV	AW(mündlich)	ZN (mündlich)

Tabelle 3.8 zeigt eine Übersicht der Korrelationen zwischen den Subtests Deutsch.

In allen Fällen zeigen sich signifikante Korrelationen zwischen den Subtests Deutsch (P<01).

3.6 Vergleich der Sprachfähigkeiten der türkischen Kinder in Testversion Türkisch und Deutsch

Tabelle 3.9 Durchschnittliche Ergebnisse der „türkischen Kinder“ in allen digitalen Subtests auf Deutsch und Türkisch

Subtest	Anzahl Items	Testversion Türkisch		Testversion Deutsch		T	df	P-wert
		Durchschn.	Sd	Durchschn.	Sd			
PW	61 Türkisch 67 Deutsch	45.35	6.67	44.38	7.64	-776	70	n.s.
KB	65	44.96	7.44	47.68	5.67	3.172	66	.002
PB	30	21.43	6.79	21.16	6.70	-.325	68	n.s.
TV	20	13.54	2.95	12.07	3.51	-.874	65	.000

Tabelle 3.9 gibt die Durchschnittswerte der „türkischen Kinder“ in allen Subtests auf Türkisch und auf Deutsch sowie die T-Test-Statistik zum Vergleich zweier Durchschnittswerte.

Der Vergleich zeigt, dass die „türkischen Kinder“ im Subtest Kognitive Begriffe im deutschen Test höhere Werte erreichen als im türkischen Test. Im Subtest Textverständnis ist die Situation umgekehrt: Die Kinder erreichen im türkischen Test höhere Werte als im deutschen Test. Die Variante-Analyse zeigt, dass die Unterschiede in den Durchschnittswerten in diesen Gruppen signifikant sind (1% Niveau). In den Subtests PW und PB erreichen diese Kinder vergleichbare Werte im türkischen und im deutschen Test.

Die beobachteten Unterschiede sind hier nicht signifikant.

3.7 Internationaler Vergleich (Deutschland, Türkei und Niederlande)

Vergleich der Sprachfähigkeiten in Türkisch zwischen türkischen Kindern in Deutschland, in der Türkei und den Niederlanden

Tabelle 3.10 Durchschnittswerte der niederländischen Kinder in der Vorschule und der 1. Klasse in den niederländischen Subtests und der deutschen Kinder in vergleichbaren Subtests auf Deutsch.

Subtest	Vorschulklasse Niederlande		1. Klasse Niederlande		Deutschland	
	durchschnittlich	Sd	durchschnittlich	Sd	durchschnittlich	Sd
PW (67)	50.98	5.65	55.41 (60)	3.53	55.54	6.64
KB (65)	52.93	5.98	60.04	2.33	54.65	6.61
PB (30)	22.20	2.51	23.78	1,76	23.77	6.46
TV(20)	16.93	2.67	17.78	1.85	15.31	3.95
AW(40)	25.24	5.84	31.24	3.59	32.32	5.55
SN (40)	32.76	4.60	35.33	3.59	34.32	7.99

In der Annahme, dass die Subtests vergleichbar sind, kann vorsichtig der Schluss gezogen werden, dass sich die Durchschnittswerte in den Niederlanden und in Deutschland nicht wesentlich voneinander unterscheiden.

Absolute Schlüsse können nicht gezogen werden, da das Alter der Kinder in den Niederlanden nicht das gleiche ist wie das Alter der deutschen Kinder.

Außerdem ist die Schullaufbahn der Kinder unterschiedlich: In den Niederlanden werden Kinder im Alter von vier Jahren eingeschult, in Deutschland im Alter von sechs Jahren.

Tabelle 3.11 Durchschnittswerte der türkischen Kinder in den türkischen Subtests: in der Türkei, in der Vorschulklasse und 1. Klasse in den Niederlanden und in Deutschland.

	Kinder in der Türkei	Vorschulklasse Niederlande		1. Klasse Niederlande		Deutschland	
Subtest	durchschnittlich	Sd	Sd	durchschnittlich	Sd	durchschnittlich	Sd
PW (61)	50.70(60)	44.59 (60)	6.17	49.43 (60)	5.48	44.37	7.64
KB (65)	50.87	50.56	5.48	57.52	5.78	47.67	6.79
PB (30)	-	23.72 (25)	1.73	24.30 (25)	1.55	21.16	6.70
TV (20)	16.82	16.59	2.88	17.74	2.45	12.07	3.51
AW(40)	31.50	23.07	4.64	23.39	5.57	22.68	6.16
SN (40)	30.39	27.15	6.21	31.70	6.18	27.16	6.62

In der Annahme, dass die Subtests vergleichbar sind, kann vorsichtig der Schluss gezogen werden, dass sich die Durchschnittswerte in den Niederlanden, in Deutschland und in der Türkei nicht wesentlich voneinander unterscheiden.

Absolute Schlüsse können nicht gezogen werden, da das Alter der Kinder in der Türkei, in den Niederlanden und in Deutschland nicht gleich ist. Außerdem sind die Schullaufbahnen unterschiedlich: In den Niederlanden werden Kinder im Alter von vier Jahren eingeschult, in Deutschland im Alter von sechs Jahren und in der Türkei im Alter von fünf Jahren.

3.8 Schlussfolgerung

Aus der in Kapitel 3 durchgeführten Analyse dürfen wir schließen, dass die für Deutschland entwickelten Tests als Messinstrumente genügend Zuverlässigkeit und Validität aufweisen. Außerdem zeigt die Analyse, dass die zunächst genannten Subtests der türkischen Version (Verhoeven u.A.) auch in einem anderen Umfeld angemessene Qualitätskriterien erfüllen.

Was die Zuverlässigkeit angeht, zeigten alle entwickelten Testaufgaben ein hohes Maß an interner Konsistenz. Die Validität betreffend zeigt die Analyse, dass die gewählten Itemzusammensetzungen eindimensional und populationsunabhängig sind und sich von daher zur Verwendung bei Sprachfähigkeitstests für deutsche, aber auch für zweisprachige Kinder eignen.

Die hohen Korrelationen zwischen den Subtests wie auch die vergleichbaren Resultate der etwa Gleichaltrigen in der Türkei, Deutschland und den Niederlanden in Betracht gezogen, erscheint die ‚Konkurrenz‘-Validität der Tests ausreichend.

Es wurde eine signifikante, positive, gegenseitige Korrelation zwischen den digitalen Testkomponenten gefunden sowie zwischen den digitalen und den mündlichen Testkomponenten. Das heißt, dass der Sprachtest eine Aussage treffen kann sowohl über die rezeptive als auch die produktive Sprachfähigkeit der Kinder.

Bei diesem Pretest wurden vor allem die Zuverlässigkeit und die Validität der Tests erforscht. Im November/Dezember 2003 wurden in einer Großforschung über 5000 Kinder getestet. Außer der Zuverlässigkeit und der Validität wurde auch der Normfestsetzung des Sprachtests besondere Aufmerksamkeit gewidmet.

Der Vergleich (Tabelle 3.9) zwischen den Sprachfertigkeiten der „türkischen“ Kinder ist mit Vorsicht zu betrachten. Die Anzahl der Kinder ist nicht groß genug, um gravierende Schlüsse ziehen zu können. Es ist

möglich, dass es nach einer großen Testabnahme andere Ergebnisse gibt. Allerdings ist es schon ein Beispiel, wie die Daten zu analysieren sind.

Was die Frage nach dem Umgang der Kinder mit dem Computerprogramm betrifft, haben die Testleiter bereits von ihren Beobachtungen während der Testabnahmen berichtet. Wie aus einem schriftlichen Bewertungsbericht während einer Versammlung im Juli 2003 in Duisburg hervorgeht, urteilten die Testleiter nuanciert positiv über die Voraussetzungen der Kinder, um diesen Test mit dem Computer zu machen (siehe Anlage B).

Anlagen

A Impressum Sprachtest

Autoren

Ömer Ahmet Konak, Cito
Tom Duindam, Cito

Entwickelt in Zusammenarbeit mit

Barbara Bader	Schulleiterin GGS, Duisburg
Klaudia Katzer	Schulleiterin GGS Mevissen Str., Duisburg
Dagny Kuehlmann	Schulamt für die Stadt Duisburg, Fachberaterin Migranten
Elisabeth Pater	Stadt Duisburg, Regionale Arbeitsstelle zur Förderung von Kindern und Jugendlichen aus Zuwandererfamilien (RAA)
Wolfgang Wirtz	Schulleiter GGS Werthausen Str., Duisburg
İskender Yıldırım	Stadt Duisburg, RAA
Rita Rzycki	Stadt Duisburg, Dezernat für Umwelt, Schule und Gesundheit

Berater:

Jagoda Illner	Landesinstitut für Schule (Lfs) des Landes Nordrhein-Westfalen, Soest
Dr. Christoph Schroeder	Institut für Migrationsforschung und Interkulturelle Studien (IMIS), Universität Osnabrück

Für das Projekt Sprachtest sind wir Ludo Verhoeven, Goretti Nairan, Guus Extra, Ömer A. Konak und Rachid Zerrouk verbunden.

Sie sind die Autoren des Tests Zweisprachigkeit Cito (Toets Tweetaligheid), 1995, Arnheim.

Information:

Für Informationen über den Sprachtest oder über Cito können Sie sich wenden an:

Cito Deutschland GmbH
Schloßstraße 10
35510 Butzbach

Tel (06033) 74 63 00

Fax (06033) 74 63 01

deutschland@cito.com

www.de.cito.com

B Durchführung des Pre-Tests: Erfahrungen der Testleiter

Erfahrungen mit der Technik

- Kinder gehen sehr selbstverständlich und sicher mit dem PC und der Maus um.
- Kinder haben keine Schwierigkeiten, das „Touchpad“ eines Laptops zu benutzen.
- Das einwandfreie Funktionieren der Maus und des Programms ist eine absolute Grundvoraussetzung, um gültige Ergebnisse zu bekommen.
- Die Kopfhörer müssen sehr gut sitzen – also dürfen sie nicht rutschen, aber auch nicht drücken.
- Die Lautstärke der Kopfhörer muss durch die Kinder regulierbar sein.
- Die Zeiten zwischen den einzelnen Aufgaben, der Antworanforderung und der Weiterleitung zur nächsten Aufgabe sollten individuell gestaltet werden können.
- Eine Programmunterbrechung sollte auch während der Aufgabenerläuterung möglich sein.
- Eine Ergebniskontrolle nach Beendigung des Tests sollte möglich sein.

Erfahrungen mit den Eltern

- Der Informationsbedarf über den Testinhalt ist sehr groß.
- Das Interesse an den Ergebnissen ist sehr groß.
- Vorbehalte gegen die Testung am PC gab es sehr wenige.
- Der Einsatz eines standardisierten Testverfahrens wird überwiegend begrüßt.

Erfahrungen mit den Kindern

- Der Test hat den meisten Kindern großen Spaß gemacht.
- „Primo“ wird von den Kindern akzeptiert und gemocht.
- Die Kinder haben spontan die Kommunikation zu „Primo“ aufgenommen, sie haben ihm geantwortet, sich über sein Lob gefreut, haben ihm zum Abschied gewunken.
- Pausen während des Tests sind für die meisten Kinder notwendig.
- Der Test hat hohe Anforderungen an die Konzentrationsfähigkeit der Kinder.
- Die Kinder erleben den Test überwiegend als Spiel.
- Das Medium „PC“ gefällt den Kindern, sie schätzen es, damit arbeiten zu können.
- Die Kinder arbeiten nach der Einführung und Erklärung des Ablaufs selbstständig mit dem Test.

Erfahrungen der Testleiterinnen und Testleiter

- Gute Kenntnisse über den Test sind unabdingbar.
- Technisches Grundverständnis im Umgang mit dem PC ist erforderlich.
- Striktes Einhalten der Durchführungsanleitung ist wichtig.
- Die Kinder müssen während des Tests gut, aber unauffällig beobachtet werden.
- Ermüdungserscheinungen der Kinder müssen beachtet und aufgegriffen werden.
- Die Testsituation muss so gestaltet sein, dass die Kinder angst- und stressfrei den Test durchlaufen können.
- Die Kinder reagieren sehr unterschiedlich auf die Testsituation; individuelles Eingehen auf jedes Kind ist erforderlich.
- Die meisten Kinder arbeiten selbstständig – manche Kinder brauchen jedoch eine individuelle Begleitung.
- Der Testplatz/Arbeitsplatz muss gut gestaltet sein: möglichst wenig Störung von außen, gute Sitzposition, gute Raumtemperatur, gute Lichtverhältnisse.
- Je vertrauter die Testumgebung/der Testort/ der Testraum ist, desto besser können sich die Kinder auf den Test konzentrieren.
- Je weniger Erfolgsdruck seitens der Eltern vermittelt wurde, desto unbefangener konnten sich die Kinder auf die Testsituation einlassen.

Teil 2

Sprachtest

Bericht einer Untersuchung der Sprachprofile von Vorschulkindern in
Duisburg unter Einsatz eines Mehrsprachigen Tests als Messinstrument

1 Einleitung

Dieser Rapport berichtet über die Erforschung der Sprachfähigkeiten von Vorschulkindern in der Stadt Duisburg, die im Herbst 2003 stattgefunden hat. Das Instrument, das für diese Forschung eingesetzt wurde, nämlich der Sprachtest, war im Juli 2003 Thema einer Vorstudie.

Ein separater Artikel "Vorstudie Sprachtest" (Dezember 2004) beschreibt diese Vorstudie und ihre Schlussfolgerungen. In der Zusammenfassung wird außerdem kurz auf diese Vorstudie zurückgegriffen. Eine wichtige Zielsetzung dieses Berichts ist es, alle Stadien der Abnahme des Sprachtests im Herbst 2003 bei etwa 5.000 Vorschulkindern aus Duisburg zu beschreiben. Diese Beschreibung umfasst folgende Phasen: die Analyse der Testergebnisse, die Punktwerte der getesteten Kinder, das Normieren der zwei Testversionen des Sprachtests und die Berichterstattung über die Daten. Aus gesellschaftlicher Sicht, aber auch aus der Sicht des Schulwesens in Deutschland, ist es wesentlich, diese Informationen optimal erfassbar zu präsentieren.

Kapitel 2 beschreibt die vier Testkomponenten des Sprachtests, die in einer deutsch- und einer türkischsprachigen Version zur Verfügung stehen.

Kapitel 3 enthält das Kernthema dieses Rapports, nämlich die Forschungsfragen, die der Forschung, die im November/Dezember 2003 durchgeführt wurde, zugrunde liegen. Außerdem werden hier die fünf Gruppen, in die die Kinder unterteilt wurden, definiert.

Inhalt von Kapitel 4 ist die Analyse des endgültigen Tests, an dem aufgrund der Vorstudie einige Veränderungen vorgenommen wurden. Dazu ist hier das im Test integrierte Eingabeformular aufgenommen worden, mit dessen Hilfe relevante Informationen über die teilnehmenden Kinder gesammelt worden sind, unter anderem Anzahl und Aufteilung der Sprachen, die die Duisburger Kinder sprechen sowie ihr Alter und ihr Geburtsland. Zuletzt werden die Analysen sowohl der deutschen als auch der türkischen Testversion dargestellt. Dieses Kapitel behandelt die Forschungsfragen 1 und 2.

In Kapitel 5 wird pro Teilgruppe ein Profil der Sprachfähigkeit erstellt: Durchschnittswert, P-Werte, Alpha und Standarddeviationen von sowohl der deutschen als auch der türkischen Testversion werden angegeben. Anschließend sind in grafischer Darstellung die Ergebnisse pro Teilgruppe in allen Testkomponenten präsentiert. Schließlich erfolgt ein Vergleich der in Duisburg erzielten Resultate mit früheren Ergebnissen aus Holland und der Türkei. Die Forschungsfragen 3, 4 und 5 sind Gegenstand dieses Kapitels.

Kapitel 6 beschreibt, was unter Normierung zu verstehen ist und wie die Normierung des Sprachtests zustande gekommen ist. Die Normierung für die deutsche und für die türkische Version des Sprachtests wird in Normierungstabellen präsentiert. Dargelegt wird außerdem, wie die Normierungsregel anzuwenden ist. Für beide Testversionen wird sodann per Testkomponente überprüft, wie das Testergebnis sich zu dem wirklichen Ergebnis verhält. Die Forschungsfrage 6 ist Gegenstand dieses Kapitels.

In Kapitel 7 werden die Testergebnisse der Kinder pro Teilgruppe und Testkomponente dargestellt. Der Leser findet sowohl in absoluten Zahlen als auch in Prozentsätzen die Anzahl der Kinder, die für zusätzlichen Sprachunterricht in Frage kommen. Die Forschungsfrage 7 wird in diesem Kapitel behandelt.

Aufgrund der Gesamtübersicht, die die Daten und Schlussfolgerungen in den Kapiteln 4 bis einschließlich 6 bieten, werden in Kapitel 7 allgemeine Schlussfolgerungen gezogen. Schließlich folgen Empfehlungen zum Sprachtest und zum Unterrichtskontext, in dem der Test angewendet wird.

2 Aufbau und Struktur des Sprachtests

Der Sprachtest setzt sich aus folgenden vier Testkomponenten zusammen:

Passiver Wortschatz **PW**, Kognitive Begriffe **KB**, Phonologisches Bewusstsein **PB** und Textverständnis **TV**. Die hinter den jeweiligen Testkomponenten genannten Abkürzungen werden in diesem Rapport verwendet. Von jeder Testkomponente gibt es eine deutsche und eine türkische Version. Die Testkomponenten werden in der Reihenfolge präsentiert, in der sie auch im Test vorkommen.

2.1 Die Testkomponenten

Passiver Wortschatz, PW, 60 Testfragen: Um die Kenntnisse in Wortbedeutungen zu messen, ist die Testkomponente "Passiver Wortschatz" entwickelt worden. Sprachbedeutung wird überwiegend mit Substantiven, Verben und Adjektiven ausgedrückt. In der deutschen und der türkischen Version werden die gleichen Wörter getestet. Die Testkomponente "Passiver Wortschatz" umfasst 60 Testfragen, wobei jedes Mal von vier Abbildungen die Abbildung angeklickt werden muss, die die Bedeutung des Wortes angibt.

Kognitive Begriffe, KB, 65 Testfragen: In der Testkomponente "Kognitive Begriffe" werden die Kenntnisse über Begriffe getestet, die sich auf Farbe, Form, Maß, Räumlichkeit und Beziehungen zwischen Ereignissen beziehen. In der deutschen und der türkischen Version testet man die gleichen Begriffe. Bei jeder Testfrage werden mehrere Abbildungen gezeigt und das Kind wird gebeten, die Abbildung anzuklicken, die dem genannten Begriff entspricht.

Phonologisches Bewusstsein, PB, 30 Testfragen: In der Testkomponente "Phonologisches Bewusstsein" wird getestet, inwiefern die Kinder in Deutsch oder Türkisch Laute erkennen und voneinander unterscheiden können. Bei jeder Testfrage sagt die Teststimme zwei Wörter, bei denen der Unterschied vom Laut her minimal ist. Zum Beispiel können Wörter sich nur in einem Laut unterscheiden (Hall - hell), die Position des Lauts kann unterschiedlich sein (Kran - Karn) oder die Anzahl der Laute ist unterschiedlich (Miete - Mitte). In dieser Testkomponente werden ausschließlich Wörter mit sinngemäßer Bedeutung angeboten. Bei jedem Wortpaar wird das Kind gebeten anzugeben, ob die Wörter vom Klang her gleich oder unterschiedlich waren. Das Kind wählt dazu aus zwei Kreisen, einem grünen für "gleich lautend" und einem roten Kreis mit Kreuz, der "unterschiedlich vom Klang her" bedeutet.

Textverständnis, TV 20, Testfragen zu 4 Geschichten: Hier wird getestet, in welchem Maß Kinder in der Lage sind, konzentriert einem kurzen gesprochenen Text zuzuhören und implizite sowie explizite Informationen zu erfassen. Dieser Testteil gibt Auskunft darüber, ob ein Kind den Text verstanden hat und Fragen zum Text beantworten kann.

2.2 Vergleich deutsche und türkische Testversion

Die Testkomponenten PW und KB der jeweiligen Version sind vom Inhalt her identisch. Die Testkomponente PB der beiden Testversionen beruht jeweils auf distinktiven Phonemmerkmalen der betreffenden Sprache. Bei Auswahl der Wörter für diese Testkomponente sind nicht nur der Klangunterschied und mögliche Klangkombinationen beachtet worden. Man berücksichtigte auch die von beiden Sprachen existierenden Lexika. Das heißt, dass in keine der beiden Testversionen in dieser Testkomponente Nonsenswörter aufgenommen worden sind. Gemäß der Zusammensetzung der Testkomponente PB sind die Wörter, die als Aufgaben präsentiert werden, in den verschiedenen Sprachversionen unterschiedlich.

Die Testkomponenten TV beider Testversionen unterscheiden sich nur in einigen Details. So stammen z.B. die Eigennamen von Personen, die in den Texten vorkommen, aus den jeweiligen Sprachbereichen.

3 Die Forschung

Der Sprachtest ist für Deutschland mit dem Ziel entwickelt worden, die Sprachfähigkeit von Vorschulkindern zu erfassen. Der Test wird etwa zehn Monate vor deren Einschulung abgenommen. Es bleibt dann noch genügend Zeit, um die Sprachfähigkeit der Kinder, die einen entsprechenden Bedarf haben, durch zusätzlichen Sprachunterricht zu verbessern. Mit Hilfe der Testresultate stellt man fest, welche Kinder für einen solchen Sprachunterricht in Frage kommen.

Die Frage nach der Reliabilität (Zuverlässigkeit), Validität (Gültigkeit) und Konsistenz des Sprachtests wurde nach Durchführung einer Vorstudie ("Vorstudie Sprachtest" Dezember 2004) positiv beantwortet. Diese Studie vermittelt Informationen über Kinder zu einem wichtigen Zeitpunkt ihrer schulischen Entwicklung. Die Ergebnisse der Studie sollen deshalb in erster Linie den Bedürfnissen des Schulwesens dienen. Da Duisburg, wie bekannt, in hohem Maße die Charakteristik einer mehrsprachigen Stadt widerspiegelt, war unser Ansicht nach von Anfang an die Unterteilung in verschiedene Sprachgruppen wesentlich. Die detaillierten Informationen über jede Sprachgruppe erweitern die Möglichkeiten des Schulwesens, um effektiv vorzugehen. Um einen wirkungsvollen Unterschied in Sprachgruppen realisieren zu können, wurden mit Hilfe eines Fragebogens (s. Abbildung 3.1) unter anderem Angaben über die „Heimsprache“ und das Geburtsland der Kinder und ihrer Eltern gesammelt. Wie in 2.1 dargestellt, erfolgte die Definition von fünf Sprachgruppen. Abgestimmt auf die Befunde aus der Literatur (Extra 2003) wurden Kriterien, bei denen es um Ethnizität oder Nationalität geht, nicht aufgeführt, da diese Kriterien nur unvollkommene Aussagen über die Sprachsituation der Kinder zulassen.

3.1 Die Fragestellung

- 1 Wie sieht die „taalkaart“ (Sprach-Atlas) der Kinder in Duisburg zu Beginn der Grundschule aus?
- 2 Misst der Sprachtest die Sprachfähigkeit auch in einer komplexen Situation mit 5 verschiedenen Sprachgruppen?
- 3 In welchem Ausmaß kontrastiert die Sprachfähigkeit Deutsch der folgenden 5 Sprachgruppen:
 - A monolinguale deutschsprachige Kinder,
 - B zweisprachige türkischsprachige Kinder, die angegeben haben, dass ihre Fähigkeit in Deutsch stärker ist,
 - C zweisprachige Kinder mit einer "anderen Sprache" (nicht Deutsch, nicht Türkisch) die angegeben haben, dass ihre Fähigkeit in Deutsch stärker ist,
 - D zweisprachige türkischsprachige Kinder, die angegeben haben, dass ihre Fähigkeit in Türkisch stärker ist,
 - E zweisprachige Kinder mit einer "anderen Sprache" (nicht Deutsch, nicht Türkisch) die angegeben haben, dass ihre Fähigkeit in der "anderen Sprache" stärker ist?
- 4 In welchem Ausmaß kontrastiert die Sprachfähigkeit Türkisch diese beiden Gruppen:
 - B zweisprachige türkischsprachige Kinder, die angegeben haben, dass ihre Fähigkeit in Deutsch stärker ist,
 - D zweisprachige türkischsprachige Kinder, die angegeben haben, dass ihre Fähigkeit in Türkisch stärker ist?
- 5 In wiefern sind die Testergebnisse in Duisburg zu vergleichen mit früher erzielten Ergebnissen in den Niederlanden und in der Türkei?
- 6 Welche Implikationen haben die Forschungsergebnisse zum Erstellen der Normierung?
- 7 Welche Implikationen haben die Forschungsergebnisse für die Unterrichtspraxis und für die Bildungspolitik in Duisburg und Deutschland?

3.2 Nähere Erklärung der fünf Gruppen

Zwecks Lesbarkeit des Berichts werden die bei Frage 3 und 4 unterschiedenen Gruppen aufgezeigt:

- A **Deutsch:** monolinguale deutschsprachige Kinder
- B **Deutsch-Türkisch** zweisprachige türkischsprachige Kinder, die angegeben haben, dass ihre Fähigkeit in Deutsch stärker ist
- C **Deutsch-Anders** zweisprachige Kinder mit einer "anderen Sprache" (nicht Deutsch, nicht Türkisch), die angegeben haben, dass ihre Fähigkeit in Deutsch stärker ist
- D **Türkisch** zweisprachige türkischsprachige Kinder, die angegeben haben, dass ihre Fähigkeit in Türkisch stärker ist
- E **Anders** zweisprachige Kinder mit einer "anderen Sprache" (nicht Deutsch, nicht Türkisch), die angegeben haben, dass ihre Fähigkeit in der "anderen Sprache" stärker ist

4 Testdaten-Analyse

In diesem Kapitel werden die Daten aus der Analyse der Testabnahme präsentiert und erläutert sowie die Ergebnisse aus der Vorstudie hinzugezogen. Am Anfang des Kapitels sind einige Änderungen im Test vermerkt, die aus der Vorstudie resultierten.

Einige relevante Eigenschaften der teilnehmenden Kinder, die aufgrund der Eingabeformulare deutlich wurden werden in Abschnitt 4.1 präsentiert. In den Abschnitten 4.2 und 4.3 sind die Analysen sowohl der deutschen als auch der türkischen Testversion dargestellt. Die Forschungsfragen 1 und 2, wie in Kapitel 3 aufgeführt, werden hier beantwortet. Das Kapitel schließt mit der Beschreibung ergänzender Wahrnehmungen in Abschnitt 4.4.

In dem Artikel über die Vorstudie wird beschrieben, dass in der Testkomponente Passiver Wortschatz der deutschen Version sieben Testfragen auf zwei unterschiedliche Arten aufgenommen sind. In der gleichen Testkomponente der türkischen Version ist eine Testfrage auf zwei Arten aufgenommen. Der Grund waren Zweifel über das Funktionieren der ursprünglichen Testfragen im deutschen Kontext. Beide Testfragenvarianten wurden in der Vorstudie getestet, bevor per Testfrage die am besten funktionierende Variante für den endgültigen Test herausgefiltert wurde. Die Abbildungen bei den betreffenden Testfragen wurden übrigens nicht geändert.

Nachfolgend stehen sechs dieser Testfragen der deutschen Version (die Wörter in Klammern wurden nicht für den endgültigen Test ausgewählt)

schützen	(verstecken)
die Krawatte	(der Schlips)
die Hacke	(die Ferse)
die Medaille	(der Orden)
die Grotte	(die Höhle)
der Ozean	(das Meer)

Nun folgt die betreffende Testfrage der türkischen Version (das Wort in Klammern wurde nicht für den endgültigen Test ausgewählt)

saklanmak	(sakinmak)
-----------	------------

Darüber hinaus weist die Testkomponente Passiver Wortschatz der deutschen Version noch einen Sonderfall aus, nämlich eine Testfrage mit der Abbildung eines Fingers. Diese wurde fälschlicherweise ein Mal mit dem Text "Klicke den Finger an" präsentiert und zusätzlich noch einmal mit dem richtigen Text "Klicke den Nagel an".

In der Vorstudie sind beide Testfragen von Kindern beantwortet worden. Im endgültigen Test ist nur das Original mit dem Text "Klicke den Nagel an" aufgenommen worden.

4.1 Angaben auf dem Eingabeformular

Im November/Dezember 2003 haben alle Kinder in Duisburg, die bei einer Grundschule angemeldet waren, den Sprachtest absolviert. Um die gewünschten Informationen über die Art und das Maß der Zweisprachigkeit zu erhalten, wurden den Kinder folgende Fragen vorgelegt (die Fragen wurden teilweise "G. Extra und P. Broeder, 2001" entnommen).

Abbildung 4.1 Eingabeformular

Sprachtest Deutsch / Türkisch	
EINSCHREIBUNG DES KINDES	
Teststandort	<input type="text"/>
Vorname	<input type="text"/>
Familienname	<input type="text"/>
Geburtsdatum	01 / 01 / 1998 ▾
Geschlecht	<input type="radio"/> männlich <input type="radio"/> weiblich
Kindergarten	<input type="radio"/> Ja <input type="radio"/> Nein
Schule	Unbekannt ▾
Testdatum	23 / 05 / 2005 ▾
Geburtsland Kind	<input type="text"/>
Geburtsland Mutter	<input type="text"/>
Geburtsland Vater	<input type="text"/>
Sprecht ihr zu Hause auch eine andere Sprache als Deutsch?	<input type="radio"/> Ja <input type="radio"/> Nein
War das Kind in einer Sprachförderungsgruppe?	<input type="radio"/> Ja <input type="radio"/> Nein
<input type="button" value="OK"/> <input type="button" value="Cancel"/>	

Aus der Analyse geht hervor, dass in Duisburg außer oder anstatt Deutsch noch 37 andere Sprachen als "Heimssprache" gesprochen werden. Nachstehend sind diese Sprachen und die Anzahl der Kinder, die sie sprechen, pro Sprache aufgeführt.

Während der ersten Analyse wurde das Eingabeformular im Ganzen für 5.330 Kinder ausgefüllt. 44,6 % der Kinder hat die Frage "Sprecht ihr zu Hause auch eine andere Sprache als Deutsch?" mit "Ja" beantwortet. Daraus geht hervor, dass 55,4 % der erforschten Kinder monolingual deutschsprachig sind. Die anderen Kinder, 44,6 % der gesamten Population, können als bilingual bezeichnet werden, wobei sich herausstellte, dass 26 % dieser Kinder türkisch sprechend sind und 15,4 % eine "andere" Sprache als "Heimssprache" hat. Mehr Angaben über die Sprachen der bilingualen Kinder, außer dem Deutschsprechen und über die Anzahl der Kinder, die die Sprache sprechen, werden in Tabelle 3.1 aufgeführt, in der die Antworten der Kinder auf die Frage "Welche Sprache sprichst Du am besten?" ausgewertet sind.

Tabelle 4.1 Die Sprachen, die von den Duisburger Kindern gesprochen werden, und die Anzahl der Kinder, die die betreffende Sprache sprechen

Sprache	Frequency	Percent	Valid Percent	Cumulative Percent
	2942	55.2	55.2	55.2
ab - Abchasisch	1	.0	0.	55.2
am - Amharisch	1	.0	0.	55.2
ar - Arabisch	43	.8	.8	56.0
as - Assamesisch	2	.0	.0	56.1
az - Aserbaischisch	2	.0	.0	56.1
da - Dänisch	2	.0	.0	56.2
de - Deutsch	846	15.9	15.9	72.0
el - Griechisch	10	.2	.2	72.2
en - Englisch	2	.0	.0	72.3
eo - Esperanto	2	.0	.0	72.3
es - Spanisch	1	.0	.0	72.3
fo - Färöisch	1	.0	.0	72.3
fr - Französisch	5	.1	.1	72.4
hi - Hindi	2	.0	.0	72.5
hr - Kroatisch	7	.1	.1	72.6
hy - Armenisch	4	.1	.1	72.7
it - Italienisch	20	.4	.4	73.1
ku - Kurdisch	27	.5	.5	73.6
mk - Mazedonisch	4	.1	.1	73.6
my - Burmesisch	2	.0	.0	73.7
na - Nauruisch	1	.0	.0	73.7
pa - Pundjabisch	2	.0	.0	73.7
pl - Polnisch	30	.6	.6	74.3
pt - Portugiesisch	2	.0	.0	74.3
ru - Russisch	30	.6	.6	74.9
sh - Serbokroatisch	4	.1	.1	75.0
sq - Albanisch	43	.8	.8	75.8
sr - Serbisch	3	.1	.1	75.8
ta - Tamilisch	6	.1	.1	75.9
tr - Türkisch	1247	23.4	23.4	99.4
ts - Tsongaisch	2	.0	.0	99.4
ur - Urdu	2	.0	.0	99.5
vi - Vietnamesisch	2	.0	.0	99.5
xx - [andere Sprachen]	26	.4	.4	99.9
zh - Chinesisch	4	.1	.1	100.0
SUMMEN	5330	100.0	100.0	

Tabelle 4.1 weist außer Deutsch noch 37 andere Sprachen auf. Die meisten werden weniger als zehn Mal genannt. Es geht hier um die Sprachen, von denen die Kinder gesagt haben, dass sie diese "am besten" sprechen.

Wenn hier auch die Antworten auf die anderen Fragen des Eingabeformulars hinzugezogen werden, ergibt sich eine Gesamtzahl von 42 Sprachen, die außer Deutsch gesprochen werden.

Auffallend ist, dass 842 bilinguale Kinder angeben, Deutsch als beste Sprache zu beherrschen. Die Zahl der Kinder, die sagt, dass sie Türkisch am besten beherrschen, liegt insgesamt bei 1240.

Tabelle 4.2 Die fünf Sprachgruppen: Frequenz und Prozentsatz

Gruppen	Frequenzen	Prozentsätze	kumulative Prozentsätze
Deutsch (mono)	2954	55.4	55.4
Deutsch-Türkisch	357	6.7	62.1
Deutsch-Anders	483	9.1	71.2
Anders	295	5.5	76.7
Türkisch	1241	23.3	100.0
Insgesamt	5330	100.0	

Wie in Kapitel 3 angegeben, werden die Auskunftspersonen in fünf Sprachgruppen unterteilt. Diese Einteilung wird hier noch einmal aufgeführt, wobei auch der Prozentsatz der Vertreter pro Sprachgruppe vermerkt ist.

- 55,4% Deutsch: monolinguale deutschsprachige Kinder
- 6,7% Deutsch-Türkisch: zweisprachige türkischsprachige Kinder, die angegeben haben, dass ihre Fähigkeit in Deutsch stärker ist
- 9,1% Deutsch-Anders: zweisprachige Kinder mit einer "anderen Sprache" (nicht Deutsch, nicht Türkisch), die angegeben haben, dass ihre Fähigkeit in Deutsch stärker ist
- 5,5% Türkisch-Deutsch: zweisprachige türkischsprachige Kinder, die angegeben haben, dass ihre Fähigkeit in Türkisch stärker ist
- 22,3% Anders-Deutsch: zweisprachige Kinder mit einer "anderen Sprache" (nicht Deutsch, nicht Türkisch), die angegeben haben, dass ihre Fähigkeit in der "anderen Sprache" stärker ist

Die hier aufgeführten Prozentsätze pro Sprachgruppe sind aufgrund der Zahl der ausgefüllten Eingabeformulare zustande gekommen. Aus weiteren Analysen geht jedoch hervor, dass für eine Gruppe von Kindern das Eingabeformular zwar ausgefüllt worden ist, sie aber keine einzige Testkomponente absolviert haben.

Außerdem gab es Kinder, die nur einige Testkomponenten absolviert haben. Da, wo es die bilingualen türkisch sprechenden Kinder betrifft, ist es so, dass manche Kinder zwar die türkische Testversion absolviert haben, aber nicht die deutsche, oder andersherum. Hierauf geht Kapitel 5 näher ein.

Für die Analyse der Testdaten wurden nur die Daten der Kinder in Betracht gezogen, die alle vier Testkomponenten der deutschen Testversion absolviert haben.

Wie in Tabelle 4.3 aufgeführt, setzte sich die Population, die danach übrig blieb, aus insgesamt 4.354 Kindern zusammen. Dies bildet die Basis für die Analyse der Testdaten. Die in Tabelle 4.1 und 4.2 genannten Frequenzen und Prozentsätze weichen daher ab von den Frequenzen und Prozentsätzen der Tabellen, die über die Analyse informieren, sowie Tabelle 4.3.

Tabelle 4.3 Die fünf Sprachgruppen: korrigierte Frequenzen und Prozentsätze

Gruppen	Frequenzen	Prozentsätze	kumulative Prozentsätze
Deutsch (mono)	2556	58.7	58.7
Deutsch-Türkisch	264	6.0	64.7
Deutsch-Anders	422	9.7	74.4
Anders	249	5.7	80.0
Türkisch	863	20.0	100.0
Insgesamt	4354	100.0	

Geburtsland des Kindes

Von den vier bilingualen Gruppen wurde das Geburtsland der Kinder ermittelt. In den folgenden Tabellen wird darauf näher eingegangen.

Tabelle 4.4 Geburtsland der Gruppe Deutsch-Türkisch

Gruppen	Frequenzen	Prozentsätze	Validierte Prozentsätze	kumulative Prozentsätze
DE	259	98.1	98.1	98.1
DK	2	.8	.8	98.9
DM	1	.4	.4	99.2
TR	2	.8	.8	100.0
Insgesamt	264	100.0	100.0	

Aus Tabelle 4.4 geht hervor, dass 98,1 % dieser Kinder in Deutschland geboren sind. Nur zwei Kinder, 0,8 %, sind in der Türkei geboren. Ferner zeigt sich, dass 1,2 % (3 Kinder) in einem anderen Land geboren sind.

Tabelle 4.5 Geburtsland der Gruppe Deutsch-Anders

	Frequenzen	Prozentsätze	Validierte Prozentsätze	kumulative Prozentsätze
DE	380	90	90	90
25 andere Länder	42	10	10	
Insgesamt	422	100.0	100.0	

Tabelle 4.5 zeigt, dass 90 % der Kinder aus der Gruppe Deutsch-Anders in Deutschland geboren sind. Die restlichen 10 % sind in 25 verschiedenen, anderen Ländern geboren, mit pro Land einem Prozentsatz, der niedriger ist als ein Prozent.

Tabelle 4.6 Geburtsland der Gruppe Anders

	Frequenzen	Prozent	Validierte Prozentsätze	kumulative Prozentsätze
DE	185	74.3	74.3	74.3
27 andere Länder	64	25.7	25.7	100
Insgesamt	249	100.0	100.0	

Tabelle 4.6 zeigt, dass 74,3 % der Kinder aus der Gruppe Anders in Deutschland geboren sind. Die restliche Gruppe, die sich aus 25,7 % zusammensetzt, ist in 27 verschiedenen anderen Ländern geboren.

Tabelle 4.7 Geburtsland der Gruppe Türkisch

	Frequenzen	Prozent	Validierte Prozentsätze	kumulative Prozentsätze
DE	832	96.4	96.4	96.4
TR	24	2.8	2.8	99.2
6 andere Länder	7	0.8	0.8	100
Insgesamt	863	100.0	100.0	

Tabelle 4.8 Geburtsland der Mutter

Gruppen	geboren in Deutschland	geboren in der Türkei	in einem anderen Land	Insgesamt
Deutsch	2364 (92.5%)	29 (1.1 %)	163 (6.4%)	2556
Deutsch-Türkisch	67 (25.4 %)	191 (72.3)	6 (2.3%)	264
Deutsch-Anders	94 (22.3 %)	46 (10.9)	282 (66.8)	422
Anders	16 (6.4 %)	21 (8.4)	212 (85.2%)	249
Türkisch	137 (15.9 %)	704 (81.6)	22 (2.5%)	863
Insgesamt	2678 (61.5)	991 (22.8)	685 (15.7)	4354

Tabelle 4.8 zeigt auffallende Angaben: 29 (1,1%) der insgesamt 2.364 Mütter, die zuhause keine andere Sprache sprechen als Deutsch, sind in der Türkei; 163 (6,4%) dieser Mütter in einem anderen Land geboren. Mit anderen Worten: 7,5% der Mütter, die mit ihren Kindern nur Deutsch sprechen, sind nicht in Deutschland geboren.

67 (25,4%) der Mütter der Gruppe Deutsch-Türkisch sind in Deutschland geboren und 191 (72,3%) in der Türkei.

94 (22,3%) der Mütter der Gruppe Deutsch-Anders sind in Deutschland geboren, 282 (69,9%) in einem anderen Land. Auffallend ist, dass 46 Mütter dieser Gruppe in der Türkei geboren worden sind. Mit anderen Worten: Diese Mütter sind in der Türkei geboren, aber sie sprechen zu Hause außer oder anstatt Deutsch eine andere Sprache als Türkisch. Das Gleiche gilt mehr oder weniger für die Gruppe Anders: 21 (8,4%) der Mütter sind in der Türkei geboren, aber sie sagen aus, nicht türkisch zu sprechen. Nur 6,4% dieser Gruppe wurde in Deutschland geboren. Mütter der Kinder der Gruppe Türkisch sind zum größten Teil in der Türkei geboren, insgesamt 704 (81,6%). Von dieser Mütter-Gruppe sind 137 (15,9%) in Deutschland und 22 (2,5%) in einem anderen Land geboren.

Tabelle 4.9 Geburtsland des Vaters

Gruppen	geboren in Deutschland	geboren in der Türkei	in einem anderen Land	Insgesamt
Deutsch	2303 (90.1%)	56 (2.2 %)	197 (7.7%)	2556
Deutsch-Türkisch	41 (15.5 %)	216 (81.8%)	7 (2.2%)	264
Deutsch-Anders	89 (21.1 %)	58 (13.7)	275 (65.2%)	422
Anders	20 (8.0 %)	20 (8.0)	209 (84.0%)	249
Türkisch	96 (11.1 %)	739 (85.6)	28 (3.7%)	863
Insgesamt	2549 (58.5%)	1089 (25.0)	716 (16.5%)	4354

Tabelle 4.9 zeigt auffallende Angaben: 56 (2.2 %) der insgesamt 2.556 Väter, die zu Hause keine andere Sprache sprechen als Deutsch, sind in der Türkei; 192 (7,7%) dieser Väter in einem anderen Land geboren. Mit anderen Worten: 7,7% der Väter, die mit ihren Kindern nur Deutsch sprechen, sind nicht in Deutschland geboren.

41 (15,5%) der Väter der Gruppe Deutsch-Türkisch sind in Deutschland geboren und 216 (81,8%) in der Türkei.

89 (21,1%) der Väter der Gruppe Deutsch-Anders sind in Deutschland geboren, 275 (65,2%) in einem anderen Land. Auffallend ist, dass 58 (13,7%) der Väter dieser Gruppe in der Türkei geboren sind. Mit anderen Worten: Diese Väter sind in der Türkei geboren, aber sie sprechen zu Hause außer oder anstatt Deutsch eine andere Sprache als Türkisch. Das Gleiche gilt mehr oder weniger für die Gruppe Anders: 20 (8,0%) der Väter sind in der Türkei geboren, aber sie sagen aus, nicht türkisch zu sprechen. Auch 20 (8%) dieser Gruppe wurde in Deutschland geboren. Väter der Kinder der Gruppe Türkisch sind zum größten Teil in der Türkei geboren, insgesamt 739 (85,6%). Von dieser Gruppe Väter sind 96 (11,1%) in Deutschland und 28 (3,3%) in einem anderen Land geboren.

Tabelle 4.10 Übersicht der Auskunftsgebenden nach Altersgruppe (Intervall ist 6 Monate)

Gruppen	<5	5 – 5.5	5.5 – 6	6 – 6.5	6.5 – 7	7 >	unbekannt	insgesamt
Deutsch	26	424	1132	795	71	16	92	2556
Deutsch-Türkisch	0	29	115	99	14	2	0	264
Deutsch-Anders	9	60	170	145	16	5	17	422
Anders	1	31	105	84	18	5	5	249
Türkisch	7	96	404	282	49	9	16	863
Insgesamt	43	640	1926	1402	168	37	135	4354

Das Alter der Auskunftsgebenden bewegt sich zwischen 4+ und 6+. Auffallend ist, dass die Gruppen <5 und >7 sehr klein sind.

4.2 Zuverlässigkeit der Testversionen Deutsch und Türkisch

In der Vorstudie ist die psychometrische Qualität der Testaufgaben untersucht worden. Für eine detaillierte Behandlung der Zuverlässigkeit und Validität wird auf Konak, Duindam und Kamphuis (2004) verwiesen. In den Tabellen 3.11 und 3.12 werden von beiden Testversionen (Deutsch und Türkisch) der Durchschnittswert, der P-Wert, der Alpha-Wert und die Standarddeviation aufgeführt.

Tabelle 4.11 Durchschnittswert, P-Wert, Alpha-Wert und die Standarddeviation aller Kinder im Test Deutsch

Aufgabe	Aufgaben insg.	N	Durchschnitt	P-Wert	alpha	Sd
PW	(60)	4660	42.47	70.80	.95	12.60
KB	(65)	4527	46.70	71.80	.91	10.28
PB	(30)	4428	19.83	66.10	.90	7.03
TV	(20)	4355	12.80	64.00	.86	4.49

Tabelle 4.12 Durchschnittswert, P-Wert, Alpha-Wert und die Standarddeviation aller türkischsprachigen Kinder im Test Türkisch

Aufgabe	Aufgaben insg.	N	Durchschnitt	P-Wert	alpha	Sd
PW	(60)	967	39.66	66.10	.85	7.52
KB	(65)	958	40.81	62.80	.84	8.55
PB	(30)	946	17.38	57.90	.84	6.13
TV	(20)	936	10.94	54.7	.69	3.62

In Tabelle 4.12 werden nur die Kinder erfasst, die einen türkischsprachigen Migrationshintergrund haben und die den Test etwa zwei Wochen nach der Testabnahme auf Deutsch auch in der türkischen Sprache absolviert haben.

Die Reliabilitätskoeffizienten in den Tabellen 3.11 und 3.12 zeigen, dass die Messgenauigkeit beider Testversionen gut bis sehr gut ist. Die Wahrnehmungen während der Vorstudie bestätigen diese Angabe.

4.3 Schlussfolgerungen

Mit über 40% Schulkindern, die zu Hause eine andere Sprache als Deutsch sprechen, ist eindeutig, dass Duisburg eine multilinguale Stadt ist, wie viele andere europäische Städte.

Fast 30% der bilingualen Kinder geben vor, dass sie die eigene Sprache besser beherrschen als die deutsche Sprache. Der gleiche Prozentsatz spricht diese Sprache auch mit den Eltern. Das Türkische nimmt zwischen den anderen Sprachen, die in Duisburg gesprochen werden, einen wichtigen Platz ein. Über 25% der gesamten Population gibt an, zu Hause türkisch zu sprechen und 20% der gesamten Population sagt, diese Sprache besser zu beherrschen als die deutsche Sprache.

Auffallend ist, dass die Frage nach dem Geburtsland der Kinder und ihrer Eltern keine relevanten Daten hervorbringt. 98% der Kinder aus der Gruppe Deutsch-Türkisch ist in Deutschland geboren.

Ein Prozentsatz, der sich nur wenig unterscheidet von der Gruppe Türkisch: 96 Prozent. Von den Kindern der Gruppe Anders ist der Prozentsatz der in Deutschland geborenen am niedrigsten, nämlich 74,3%.

In der Gruppe Deutsch-Anders gibt 90 % an, dass sie in Deutschland geboren sind.

5 Profile der Sprachentwicklung

In diesem Kapitel werden die durchschnittlichen Ergebnisse und die Standarddeviationen in der deutschen Testversion angegeben (Tabelle 5.1). Diese Werte werden auch von der türkischen Testversion präsentiert die nur von den türkischsprachigen Auskunftsgebenden absolviert worden ist.

Die Zahl der Kinder beider türkischsprachigen Gruppen unterscheidet sich übrigens von der Zahl die (unter anderem) in Tabelle 5.10 vermerkt ist. Der Grund besteht darin, dass einige Kinder zwar die deutsche, aber nicht die (vollständige) türkische Testversion absolviert haben.

In diesem Kapitel wird ferner je Sprachgruppe das Profil der Sprachfähigkeit aufgeführt. Mit Hilfe grafischer Darstellungen in den Anlagen werden die Ergebnisse je Sprachgruppe in allen Testkomponenten präsentiert. Zuletzt folgt ein Vergleich aller in Duisburg ermittelten Resultate mit früheren Ergebnissen in den Niederlanden und in der Türkei. Dieses Kapitel gibt Antwort auf die Forschungsfragen 3, 4 und 5.

5.1 Ergebnisse aller Sprachgruppen in der deutschen Testversion

Tabelle 5.1 Durchschnittliche Ergebnisse und Standarddeviationen aller Sprachgruppen in der deutschen Testversion

Testkomponente Gruppe	PW (60)	KB (65)	PB (30)	TV (20)
Deutsch (N)	2556	2556	2556	2556
D.schn.	49.10	51.33	21.81	14.57
Sd	8.00	7.34	6.82	3.92
Deutsch-Türkisch (N)	264	264	264	264
D.schn.	38.70	43.50	17.86	10.89
Sd	8.86	7.34	6.51	3.88
Deutsch-Anders(N)	422	422	422	422
D.schn.	43.68	47.14	19.86	12.73
Sd	8.95	8.50	6.60	4.13
Anders (N)	249	249	249	249
D.schn.	34.90	40.73	17.15	10.32
Sd	10.42	9.28	5.71	4.06
Türkisch (N)	863	863	863	863
D.schn.	31.26	38.42	16.32	8.91
Sd	9.10	8.15	5.30	3.35

Tabelle 5.1 zeigt die durchschnittlichen Ergebnisse und die Standarddeviationen aller fünf Gruppen, so wie diese in Abschnitt 3.2 aufgeführt wurden.

Die Gruppe Deutsch schneidet am besten ab, die Gruppe Türkisch am schlechtesten. Auffallend ist, dass die Gruppe Anders eine relativ hohe Verstreuung aufweist, ausgenommen die Testkomponente Phonologisches Bewusstsein. Die größte Streuung ist in der Testkomponente Passiver Wortschatz sichtbar, die niedrigste in der Testkomponente Textverständnis. Letztes gilt für alle Gruppen.

5.2 Vergleich aller Sprachgruppen in der deutschen Testversion

Tabelle 5.2 Vergleich der durchschnittlichen Ergebnisse der Gruppen Deutsch und Deutsch-Türkisch in der deutschen Testversion

Aufgabe	durchschn. Deutsch	durchschn Deutsch-Türkisch	F	t	df	P
PW	49.10	38.70	10.01	19.870	2818	.000
KB	51.33	43.50	7.32	16.40	2818	.000
PB	21.81	17.86	7.68	9.02	2818	.000
TV	14.57	10.89	.01	14.51	2818	.000

Tabelle 5.2 zeigt die durchschnittlichen Ergebnisse der Gruppen Deutsch und Deutsch-Türkisch für alle Testkomponenten der Testversion Deutsch sowie die T-Statistiken zum Vergleich der beiden Durchschnitte. Die Varianzanalyse zeigt auf, dass die Unterschiede zwischen beiden Gruppen in allen Testkomponenten signifikant sind.

Tabelle 5.3 Vergleich der durchschnittlichen Ergebnisse der Gruppen Deutsch und Deutsch-Anders in der deutschen Testversion

Aufgabe	durchschn. Deutsch	durchschn Deutsch-Anders	F	t	df	P
PW	49.10	43.69	19.21	12.61	2976	.000
KB	51.33	47.14	11.62	10.74	2976	.000
PB	21.81	19.86	1.81	5.50	2976	.000
TV	14.57	12.73	4.73	8.40	2976	.000

Tabelle 5.3 zeigt die durchschnittlichen Ergebnisse der Gruppen Deutsch und Deutsch-Anders für alle Testkomponenten der Testversion Deutsch sowie die T-Statistiken zum Vergleich der beiden Durchschnitte. Die Varianzanalyse zeigt auf, dass die Unterschiede zwischen beiden Gruppen in allen Testkomponenten signifikant sind.

Tabelle 5.4 Vergleich der durchschnittlichen Ergebnisse der Gruppen Deutsch und Anders in der deutschen Testversion

Aufgabe	durchschn. Deutsch	Durchschn. Anders	F	T	df	P
PW	49.10	34.90	43.154	25.92	2803	.000
KB	51.33	40.73	27.28	21.30	2803	.000
PB	21.81	17.15	44.28	10.44	2803	.000
TV	14.57	10.33	0.74	16.25	2803	.000

Tabelle 5.4 zeigt die durchschnittlichen Ergebnisse der Gruppen Deutsch und Anders für alle Testkomponenten der Testversion Deutsch sowie die T-Statistiken zum Vergleich der beiden Durchschnitte. Die Varianzanalyse zeigt auf, dass die Unterschiede zwischen beiden Gruppen in allen Testkomponenten signifikant sind.

Tabelle 5.5 Vergleich der durchschnittlichen Ergebnisse der Gruppen Deutsch und Türkisch in der deutschen Testversion

Aufgabe	durchschn. Deutsch	Durchschn. Türkisch	F	t	df	P
PW	49.10	31.26	38.08	54.58	3417	.000
KB	51.33	38.42	12.51	43.59	3417	.000
PB	21.81	16.32	186.55	21.57	3417	.000
TV	14.57	8.91	36.42	37.94	3417	.000

Tabelle 5.5 zeigt die durchschnittlichen Ergebnisse der Gruppen Deutsch und Türkisch für alle Testkomponenten der Testversion Deutsch sowie die T-Statistiken zum Vergleich der beiden Durchschnitte. Die Varianzanalyse zeigt auf, dass die Unterschiede zwischen beiden Gruppen in allen Testkomponenten signifikant sind.

Tabelle 5.6 Vergleich der durchschnittlichen Ergebnisse der Gruppen Deutsch-Türkisch und Deutsch-Anders in der deutschen Testversion

Aufgabe	durchschn. Deutsch-Türkisch	durchschn. Deutsch-Anders	F	t	df	P
PW	38.70	43.69	0.11	7.13	684	.000
KB	43.50	47.14	0.01	5.51	684	.000
PB	17.86	19.86	1.86	3.89	684	.000
TV	10.89	12.73	2.39	5.81	684	.000

Tabelle 5.6 zeigt die durchschnittlichen Ergebnisse der Gruppen Deutsch-Türkisch und Deutsch-Anders für alle Testkomponenten der Testversion Deutsch sowie die T-Statistiken zum Vergleich der beiden Durchschnitte. Die Variante-Analyse zeigt auf, dass die Unterschiede zwischen beiden Gruppen in allen Testkomponenten signifikant sind.

Tabelle 5.7 Vergleich der durchschnittlichen Ergebnisse der Gruppen Anders und Deutsch-Türkisch in der deutschen Testversion

Aufgabe	durchschn. Anders	durchschn. Deutsch-Türkisch	F	t	df	P
PW	34.90	38.70	6.24	4.45	511	.000
KB	40.73	43.50	3.02	3.57	511	.000
PB	17.15	17.86	7.56	1.31	511	n.s
TV	10.33	10.89	0.51	1.62	511	n.s

Tabelle 5.7 zeigt die durchschnittlichen Ergebnisse der Gruppen Anders und Deutsch-Türkisch für alle Testkomponenten der Testversion Deutsch sowie die T-Statistiken zum Vergleich der beiden Durchschnitte. Die Varianzanalyse zeigt auf, dass die Unterschiede zwischen beiden Gruppen in PW und KB signifikant sind, aber nicht in PB und TV.

Tabelle 5.8 Vergleich der durchschnittlichen Ergebnisse der Gruppen Anders und Türkisch in der deutschen Testversion

Aufgabe	durchschn. Anders	durchschn. Türkisch	F	t	df	P
PW	34.90	31.26	6.81	5.37	1110	.000
KB	40.73	38.42	6.85	3.82	1110	.000
PB	17.15	16.32	1.58	2.13	1110	.n.s
TV	10.33	8.91	18.77	5.68	1110	.000

Tabelle 5.8 zeigt die durchschnittlichen Ergebnisse der Gruppen Anders und Türkisch für alle Testkomponenten der Testversion Deutsch sowie die T-Statistiken zum Vergleich der beiden Durchschnitte. Die Varianzanalyse zeigt auf, dass die Unterschiede zwischen beiden Gruppen in PW, KB und TV signifikant sind, aber nicht in PB.

Tabelle 5.9 Vergleich der durchschnittlichen Ergebnisse der Gruppen Deutsch-Türkisch und Türkisch in der deutschen Testversion

Aufgabe	durchschn. Deutsch-Türkisch	durchschn. Türkisch	F	t	df	P
PW	38.70	31.26	0.31	11.67	1125	.000
KB	43.50	38.42	0.19	8.23	1125	.000
PB	17.86	16.32	24.98	3.89	1125	.000
TV	10.89	8.91	12.26	11.25	1125	.000

Tabelle 5.9 zeigt die durchschnittlichen Ergebnisse der Gruppen Deutsch-Türkisch und Türkisch für alle Testkomponenten der Testversion Deutsch sowie die T-Statistiken zum Vergleich der beiden Durchschnitte. Die Varianzanalyse zeigt, dass die Unterschiede zwischen allen Gruppen signifikant sind.

5.3 Ergebnisse der beiden türkischen Sprachgruppen in der türkischen Testversion

Wie zu Anfang des Kapitels bereits erwähnt, unterscheidet sich die Zahl der beiden türkischsprachigen Gruppen in untenstehender Tabelle etwas von den Zahlen in Tabelle 5.1. Dies begründet sich darin, dass nicht alle türkischsprachigen Kinder beide Testversionen vollständig absolviert haben.

Tabelle 5.10 Durchschnittswerte und Standarddeviationen in der türkischen Testversion der Gruppen Deutsch-Türkisch und Türkisch

Testkomponente Gruppe	PW (60)	KB (65)	PB (30)	TV (20)
Deutsch-Türkisch (N)	204	201	200	198
D.schn.	40.39	42.80	19.03	11.88
Sd	5.75	8.03	6.77	3.49
Türkisch (N)	701	697	689	682
D.schn.	39.72	40.36	16.77	10.54
Sd	7.28	8.35	5.75	3.52

Alle in Tabelle 5.10 aufgeführten Kinder haben auch die deutsche Testversion absolviert.

Tabelle 5.11 Durchschnittliche Ergebnisse der Gruppen Deutsch-Türkisch und Türkisch in der türkischen Testversion

Aufgabe	durchschn. Deutsch-Türkisch	durchschn. Türkisch	F	T	df	P
PW	40.39	39.72	.70	2.95	903	n.s
KB	42.80	40.36	.202	2.23	896	.000
PB	19.03	16.77	18.74	2.52	887	.000
TV	11.88	10.54	0.19	4.25	878	.000

Tabelle 5.11 zeigt die durchschnittlichen Ergebnisse der Gruppen Deutsch-Türkisch und Türkisch für alle Testkomponenten der Testversion Türkisch sowie die T-Statistiken zum Vergleich der beiden Durchschnitte. Die Varianzanalyse zeigt auf, dass die Unterschiede zwischen beiden Gruppen in allen Testkomponenten signifikant sind, ausser in der Testkomponente PW.

5.4 Ergebnisse der bilingualen, türkisch sprechenden Kinder in der deutschen und der türkischen Testversion

Tabelle 5.12 Durchschnittliche Ergebnisse der Gruppe Deutsch-Türkisch in der deutschen und der türkischen Testversion

	Test Deutsch	Test Türkisch	sd	T	df	P
PW	37.80	40.45	6.49	-3.10	122	.003
KB	43.61	42.60	8.39	- 1.33	120	n.s
PB	18.38	18.93	9.61	-.92	119	n.s
TV	11.14	12.17	2.79	-4.10	117	.000

Tabelle 5.12 zeigt die durchschnittlichen Ergebnisse der Gruppe Deutsch-Türkisch für alle Testkomponenten der Testversion Deutsch und der Testversion Türkisch sowie die T-Statistiken zum Vergleich der beiden Durchschnitte. Die Analyse zeigt auf, dass die Kinder der Gruppe Deutsch-Türkisch in der türkischen Testversion besser abschneiden als in der deutschen Testversion. Außerdem weist die Analyse nach, dass die Unterschiede der erreichten Werte in PW und TV signifikant sind, während die Unterschiede der erreichten Werte in KB und PB nicht signifikant sind.

Tabelle 5.13 Durchschnittliche Ergebnisse der Gruppe Türkisch in der deutschen und der türkischen Testversion

	Test Deutsch	Test Türkisch	sd	T	df	P
PW	30.40	40.14	9.56	-20.03	382	.000
KB	37.93	40.73	7.76	-7.11	384	.000
PB	16.25	17.10	6.41	-2.60	385	.010 (n.s)
TV	8.73	10.64	2.97	-12.628	381	.000

Tabelle 5.13 zeigt die durchschnittlichen Ergebnisse der Gruppe Türkisch für alle Testkomponenten der Testversion Deutsch und der Testversion Türkisch sowie die T-Statistiken zum Vergleich der beiden Durchschnitte.

Die Analyse zeigt auf, dass die Kinder der Gruppe Türkisch in der türkischen Testversion besser abschneiden als in der deutschen Testversion. Außerdem weist die Analyse nach, dass die Unterschiede der erreichten Werte in KB, PW und TV signifikant sind, während die Unterschiede der erreichten Werte in PB nicht signifikant sind.

5.5 Internationaler Vergleich von Kindern in der Türkei und in den Niederlanden in den Altersgruppen 5 beziehungsweise 6 Jahre

Tabelle 5.14 Vergleich der Testergebnisse der deutschen und niederländischen monolingualen Kinder

Lokation Klasse Älter Aufgabe	Duisburg gem 5-6	Duisburg 6-7	Niederlanden eind gr 1 5-6	Niederlanden eind gr 2 6-7
PW	48.4222	50.2506	50.98	55.41
KB	50.8413	52.3141	52.93	60.04
TV	12.32	12	15.00	16.93

Tabelle 5.14 zeigt alle Ergebnisse in drei Testkomponenten in zwei Sprachen (Deutsch und Niederländisch). Die Ergebnisse sind an sich nicht ohne weiteres zu vergleichen. Erstens ist der Test in Holland ein "Papier und Bleistift-Test", während der deutsche Test von den Kindern am Computer gemacht wurde. Außerdem ist in der deutschen Testversion der Inhalt der Testkomponente TV etwas abgeändert worden. Und schließlich ist die niederländische Testversion bereits 1989 und 1992 absolviert worden, die deutsche Version in 2003.

Trotzdem sagen die Angaben in Tabelle 5.14 einiges aus. Auffallend ist, dass sich Unterschiede im Bezug auf die Entwicklung innerhalb der Gruppen zeigen. Niederländische Kinder waren im Allgemeinen am Ende der Basisschulklasse 1 über fünf Jahre alt und hatten ein Jahr Unterricht. Die Kinder waren am Ende der Basisschulklasse 2 etwa sechs Jahre alt und hatten zwei Jahre Unterricht erhalten. Außerdem ist es interessant zu sehen, dass die Unterschiede zwischen den einsprachig niederländischen Kindern größer sind als zwischen den deutschen einsprachigen Kindern.

Tabelle 5.15 Vergleich der Testergebnisse in der türkischen Testversion, die in Deutschland, in der Türkei und in den Niederlanden erlangt wurden

Lokation Klasse Älter Aufgabe	Duisburg gem 5-6	Duisburg 6-7	Niederlanden 5-6	Niederlanden gr 2 6-7	Türkei 5-6
PW	39.78	41.60	44.59	49.43	50.69
KB	40.22	42.92	50.56	57.52	50.87
TV	10.51	11.33	16.59	17.74	16.60

Tabelle 5.15 zeigt alle Ergebnisse in drei Testkomponenten in der türkischen Testversion. Die Ergebnisse sind an sich nicht ohne weiteres zu vergleichen. Erstens ist der Test in Holland ein "Papier und Bleistift-Test", während der deutsche Test von den Kindern am Computer gemacht wurde. Außerdem ist in der deutschen Testversion der Inhalt der Testkomponente TV etwas abgeändert worden. Und schließlich ist die niederländische Testversion bereits 1989 und 1992 absolviert worden, die deutsche Version in 2003.

Trotzdem sagen die Angaben in Tabelle 5.15 einiges aus. Signifikant ist, dass sich Unterschiede im Bezug auf die Entwicklung innerhalb der Gruppen zeigen. Niederländische Kinder waren im Allgemeinen am Ende der Basisschulklasse 1 über 5 Jahre alt und hatten ein Jahr Unterricht. Die Kinder waren am Ende der Basisschulklasse 2 in etwa 6 Jahre alt und hatten zwei Jahre Unterricht erhalten. Die türkischen Kinder im Alter von ca. fünf Jahren hatten keinen formellen Unterricht erhalten, die Kinder im Alter von etwa 6 Jahren hatten ein Jahr Unterricht.

Im Allgemeinen kann man sagen, dass die Entwicklung innerhalb der Gruppen in den Niederlanden und in der Türkei (außer in TV) sichtbarer ist als die Entwicklung innerhalb der Gruppen in Deutschland. Dies gilt übrigens nicht für die Teilgruppe Türkisch in der Testkomponente TV.

Dies ist vorsichtig zu formulieren, da Kinder in Deutschland zum Testzeitpunkt noch keinen formellen Unterricht erhalten haben.

5.6 Schlussfolgerungen

Vergleiche in der deutschen Testversion

Die monolingualen deutschen Kinder schneiden besser ab als alle anderen Gruppen in allen Testkomponenten auf Deutsch. Die Unterschiede in allen Testkomponenten sind signifikant. Auffallend ist, dass die jüngeren Kinder besser abschneiden als die älteren (6.5 bis 7 und 7+). Eine relativ kleine Gruppe von etwa 25 Kindern, die jünger als 5 Jahre sind, schneidet sogar bedeutend besser ab.

Zwischen den Gruppen zeigen sich noch einige Unterschiede: Die Kinder der Gruppe Deutsch-Anders (die also nach eigenem Urteil Deutsch als dominante Sprache beherrschen) schneiden nach der Gruppe Deutsch am besten ab. Sie schneiden besser ab als die Gruppen Deutsch-Türkisch, Anders und Türkisch in allen Testkomponenten. Die Unterschiede sind in allen Testkomponenten signifikant.

Vergleichen wir die Gruppe Deutsch-Türkisch mit der Gruppe Anders, dann sehen wir, dass die Gruppe Deutsch-Türkisch in allen Testkomponenten besser abschneidet. Die Unterschiede sind aber nicht immer signifikant. In den Testkomponenten PB und TV unterscheiden sie sich nicht signifikant. Aber in PW und KB sind signifikante Unterschiede zwischen beiden Gruppen.

Die Gruppe Anders schneidet in drei Testkomponenten signifikant besser ab als die Gruppe Türkisch. Nur die Unterschiede in der Testkomponente PB ist nicht signifikant.

Die Gruppe Deutsch-Türkisch schneidet in allen Testkomponenten signifikant besser ab als die Gruppe Türkisch.

Vergleiche in der türkischen Testversion

Die unterschiedlichen Ergebnisse zwischen beiden türkischen Gruppen (Deutsch-Türkisch und Türkisch) sind auffallend. Die Gruppe Deutsch-Türkisch schneidet auch, außer in PW, in allen Testkomponenten signifikant besser ab als die Gruppe Türkisch. Eine vorsichtige Schlussfolgerung wäre, dass die Gruppe Deutsch-Türkisch balanciert zweisprachig ist.

Vergleiche in beiden Testversionen

Was die zweisprachigen Kinder angeht, so weist auch die Gruppe Deutsch-Türkisch auffallende Ergebnisse auf. Aus weiteren Analysen der Resultate der Kinder dieser Gruppe, die alle Testkomponenten in beiden Sprachen absolviert haben, geht hervor, dass sie in den Testkomponenten PW und TV der türkischen Version signifikant besser abschneiden als in den gleichen Testkomponenten der deutschen Version. Die beiden anderen Testkomponenten beider Sprachen (KB und PB) weisen keine signifikanten Unterschiede auf. Bei der Gruppe Türkisch, die also ausgesagt hat, dass ihr Türkisch besser ist als ihr

Deutsch, wurden signifikante Unterschiede in den drei Testkomponenten PW, KB und TV wahrgenommen. In der Testkomponente PB zeigt sich kein signifikanter Unterschied.

Internationale Vergleiche

Beim internationalen Vergleich monolingualer und bilingualer Kinder ist wahrzunehmen, dass die Ergebnisse in den Testkomponenten KB und PW vergleichbar sind, während die Kinder in der Türkei und in den Niederlanden in der Testkomponente TV besser abschneiden. Eine eventuelle Begründung könnte sein, dass die Testkomponente TV der digitalen Version etwas schwieriger ist. Die Testkomponente PB der deutschen Testversion kann nicht verglichen werden, da die Testfragen vollkommen verändert sind. Was sich sowohl in der Türkei als auch in den Niederlanden zeigt, ist, dass mit zunehmendem Alter die Rede ist von sich steigernden Ergebnissen, möglicherweise weil die Kinder zur Schule gehen. Bei den in Deutschland lebenden Kindern, die noch nicht zur Schule gehen, konnte solch eine Steigerung nämlich nicht festgestellt werden.

6 Testnormierung

In Kapitel 6 ist die Normierung des Sprachtests das zentrale Thema.

Abschnitt 6.1 beschreibt, was unter Normieren verstanden wird und wie die Normierung des Sprachtests zustande gekommen ist. Angesichts der Tatsache, dass in Deutschland (Duisburg) keine Beispielnormierungen existierten, war für den Sprachtest eine Normierung zu erstellen. Dabei wurde, wo es möglich war, von universellen Standpunkten ("wie sprachfähig muss ein Kind im Alter von etwa 6 Jahren sein") ausgegangen. In diesen Prozess sind auch die Ergebnisse der fünf unterschiedlichen Gruppen einbezogen worden. Außerdem sind durch das Hinzuziehen von relativ vielen Fachleuten, zur Erstellung der Normierung die Meinungen und Gutachten dieser Spezialisten der deutschen Bildungsstruktur (deutsche Lehrer und Fachexperten) eingeflossen. Ferner sind die Ergebnisse des Sprachtests auf internationalem Niveau hinzugezogen worden.

In Abschnitt 6.2 werden die Normierungstabellen aufgeführt und in Abschnitt 6.3 die Zuverlässigkeitstabellen präsentiert, sowohl von der deutschen wie von der türkischen Version des Sprachtests.

In Abschnitt 6.4 werden abschließend einige Schlussfolgerungen formuliert.

6.1 Beschreibung des Normierungsprozesses

Ziel des Testes ist zu entscheiden, welche Kinder einen besonderen Sprachförderbedarf haben. Es geht darum herauszufinden, ob die getesteten Kinder über ein ausreichendes Sprachvermögen verfügen, um dem Unterricht in der Grundschule folgen zu können. Der Test dient nicht der "AusSORTIERUNG" von Kindern, sondern der Ermittlung eines besonderen Förderbedarfs. Die sprachliche Entwicklung der hier getesteten Kinder befindet sich in einer ausreichend frühen Entwicklungsphase, die es erlaubt, ermittelte sprachliche Defizite bei geeigneter Förderung in aller Regel gut aufzuarbeiten.

Das Prädikat "Sprachförderbedürftig" beschreibt nicht grundsätzlich die sprachliche Kompetenz eines Kindes, sondern bildet den aktuellen Stand ab.

Die Normierung soll möglichst genau angeben, wie viele Aufgaben (Items) ein Kind mindestens richtig beantwortet haben sollte, damit ein zusätzlicher Sprachförderbedarf weitestgehend ausgeschlossen werden kann.

Unter Normierung ist das Ergebnis eines Beratungs- und Entscheidungsprozesses zu verstehen, das erzielt wurde auf Grundlage von:

- Fachliteratur
- Erfahrungen im Umgang mit bisher erworbenen Erkenntnissen im Test „Zweisprachigkeit“
- Analyse der erzielten Testergebnisse
- Fachmeinungen

Die hier festgesetzte Normierung ist wissenschaftlich belastbar, muss sich jedoch durch die dauerhafte Beobachtung sowohl der Kinder, die Sprachförderung erhalten haben, als auch der Kinder, die aufgrund des Testergebnisses keine zusätzliche Sprachförderung in Anspruch nahmen, beweisen. Diese Vorgehensweise entspricht der für die Entwicklung von wissenschaftlich belastbaren Sprachtests standardisierten Verfahrensweise.

Literatur

Der Umfang des passiven Wortschatzes kann aufgrund der empirisch äußerst schwierigen Forschungslage nur näherungsweise beschrieben werden. Die Autoren beziehen sich vorzugsweise auf die Erkenntnisse, die *Verhoeven* 1994 veröffentlicht hat:

Der passive Wortschatz sollte bei 4-jährigen Kindern mindestens ca. 2000 Wörter umfassen. Bezogen auf den Test hieße das, dass minimal 19 von 60 Items richtig beantwortet sein müssten.

Für den Bereich kognitive Begriffe gilt, auch wieder bezogen auf ein 4-jähriges Kind, dass minimal 18 von 65 Items richtig beantwortet sein müssten. Entsprechend erhöht sich die Zahl der richtig beantworteten Items für 5-, 6- und 7-jährige Kinder.

Erfahrungen

Der Test „Zweisprachigkeit“ wird in den Niederlanden seit 15 Jahren zur Testung von Kindern im Alter zwischen vier und sieben Jahren angewendet.

Dort wurden für verschiedene ethnische Gruppen Normierungswerte festgelegt, die die Grenze beschreiben, ab wann ein Kind keinen zusätzlichen Sprachförderbedarf hat. Diese Normierung wurde durch eine Langzeitstudie erhärtet: Kinder, die auf Grund der Normierung ohne zusätzliche Sprachförderung in den Unterricht der Grundschule gingen, konnten diesem ohne Schwierigkeiten folgen.

Die festgelegten Normierungswerte haben sich damit als tauglich erwiesen und konnten somit in die Betrachtung einbezogen werden.

Untersuchung

Die in Duisburg im Rahmen der Durchführung des Tests im November und Dezember 2003 erworbenen Daten standen im Zentrum des Normierungsprozesses.

Die vorliegenden Ergebnisse der Durchschnittswerte, der Standardabweichungen, der Alpha-Werte und der Werte für die jeweiligen Subtests in den definierten Subgruppen erhärteten die zuvor aufgestellten Thesen, die auf der Grundlage von Literatur und Testforschungserfahrungen aufgestellt wurden.

Fachmeinungen

Die Fachmeinung ist das vierte Element im Prozess der Normierung. Experten, d.h. Sprachwissenschaftler, Psychometriker, Testentwicklungsforscher seitens Cito, des Landesinstituts für Schule und der Universität Duisburg/Essen beraten die örtlichen Fachleute aus Verwaltung und Praxis auf den oben dargestellten Grundlagen, wo die Normierungsgrenzen liegen könnten oder sollten. Dieses Gremium aus Experten und Fachleuten hat sich am Ende des Beratungs- und Entscheidungsprozesses konsensual für die im Folgenden dargestellten Normierungsgrenzen entschieden. Dieser konsensuale Prozess ist ein wesentliches Element in dieser Phase der Testentwicklung. Die wissenschaftliche Erhärtung der Normierungsgrenzen ist daran gebunden, um die für wissenschaftliche Anerkennung erforderliche Komponente der Dimension "gesellschaftlicher Kontext und Konsens" abzubilden.

6.2 Darstellung der Normierung

Als Ergebnis des in Abschnitt 6.1 beschriebenen Prozesses wurden die folgenden Normierungstabellen festgelegt. Außerdem wurde im Bezug auf Anwendung der Normierung bestimmt, dass ein Kind für zusätzlichen Sprachunterricht in Frage kommt, wenn es in einer Testkomponente "förderbedürftig" abschneidet.

Tabelle 6.1 Die Normierungstabelle der deutschen Testversion

	Förderbedürftig	Befriedigend	Gut
Passiver Wortschatz 60	0-33	34-40	41-60
Kognitive Begriffe 65	0-39	40-45	46-65
Phonologisches Bewusstsein 30	0-13	14-19	20-30
Textverständnis 20	0-8	9-12	13-20

Tabelle 6.2 Die Normierungstabelle der türkischer Testversion

	Förderbedürftig	Befriedigend	Gut
Passiver Wortschatz 60	0-36	37-40	41-60
Kognitive Begriffe 65	0-35	36-42	43-65
Phonologisches Bewusstsein 30	0-12	13-19	20-30
Textverständnis 20	0-9	10-12	13-20

Wie die sprachliche Förderung für ein förderbedürftiges Kind konkret aussieht, hängt davon ab, in welchen und in wie vielen Subtests das Kind „förderbedürftig“ abgeschnitten hat. So wird eine dem individuellen Bedarf entsprechende Förderung sichergestellt.

6.3 Die Messgenauigkeit

Die Zuverlässigkeitskoeffizienten in Kapitel 4 weisen auf, dass die Messgenauigkeit gut bis sehr gut ist. In diesem Abschnitt wird die Messgenauigkeit im Zusammenhang mit der gewählten Normierung betrachtet. Diesbezügliche Ergebnisse sind in den so genannten Reliabilitätstabellen aufgeführt.

In den ersten acht Tabellen wird die Beziehung zwischen den Ergebnisgruppen aufgrund des Testergebnisses und den übereinstimmenden Gruppen aufgrund des wirklichen Ergebnisses dargestellt. Die Angaben basieren auf allen Kindern, die an dem bestimmten Test teilgenommen haben.

So zeigt Tabelle 6.4 zum Beispiel, dass 83,2% der Kinder, die aufgrund ihrer Punktwerte förderbedürftig abschneiden, auch mit ihrem wirklichen Ergebnis in diese Normierungskategorie fallen. Anders ausgedrückt: Die Wahrscheinlichkeit, dass das Testergebnis eines Kindes mit Recht als förderbedürftig klassifiziert wird, liegt bei 83,2%. Von den Kindern mit einem "förderbedürftigen" Testresultat sind 16,8% nicht korrekt klassifiziert worden (16,4% befriedigend, 0,4% in Wirklichkeit gut). In diesen Tabellen kann der Prozentsatz der Missklassifikationen problemlos abgelesen werden.

Tabelle 6.3 Reliabilitätstabelle PW Test Deutsch

Ergebnisgruppe (Testergebnis)	Punktwerte (wirkliche Werte)		
	Fb	Bf	Gt
Fb	87.8	12.2	.0
Bf	11.5	71.8	16.7
Gt	.0	6.3	93.7

Tabelle 6.4 Reliabilitätstabelle KB Test Deutsch

Ergebnisgruppe (Testergebnis)	Punktwerte (wirkliche Werte)		
	Fb	Bf	Gt
Fb	83.2	16.4	.4
Bf	15.4	62.5	22.1
Gt	.1	7.1	92.8

Tabelle 6.5 Reliabilitätstabelle PB Test Deutsch

Ergebnisgruppe	Punktwerte (wirkliche Werte)		
(Testergebnis)	Fb	Bf	Gt
Fb	79.0	21.0	.0
Bf	9.5	72.2	18.3
Gt	.0	8.5	91.5

Tabelle 6.6 Reliabilitätstabelle TV Test Deutsch

Ergebnisgruppe	Punktwerte (wirkliche Werte)		
(Testergebnis)	Fb	Bf	Gt
Fb	70.1	29.4	0.5
Bf	10.6	66.3	23.1
Gt	.1	10.3	89.6

Die Tabellen 6.3, 6.4, 6.5 und 6.6 geben die Messgenauigkeit der Subtests der deutschen Testversion im Bezug auf die gewählte Normierung wieder. Auffallend ist, dass bei PW der Prozentsatz an richtigen Klassifikationen am höchsten ist. 87,8% der Kinder, die aufgrund ihrer Punktwerte als "förderbedürftig" eingestuft werden, befinden sich auch mit ihrem wirklichen Testergebnis innerhalb dieser Normierungskategorie. Anders ausgedrückt: Die Wahrscheinlichkeit, dass das Testergebnis eines Kindes zu Recht als "förderbedürftig" eingestuft wird, liegt bei 87,8%.

12,2% der Kinder sind inkorrekt eingestuft worden (in Wirklichkeit 12,2% befriedigend und 0.0% gut).

Tabelle 6.7 Reliabilitätstabelle PW Test Türkisch

Ergebnisgruppe	Punktwerte (wirkliche Werte)		
(Testergebnis)	Fb	Bf	Gt
Fb	81.2	16.9	1.9
Bf	23.4	50.0	26.6
Gt	.8	11.0	88.1

Tabelle 6.8 Reliabilitätstabelle KB Test Türkisch

Ergebnisgruppe	Punktwerte (wirkliche Werte)		
(Testergebnis)	Fb	Bf	Gt
Fb	78.4	21.1	.6
Bf	14.4	66.3	19.3
Gt	.1	13.6	86.3

Tabelle 6.9 Reliabilitätstabelle PB Test Türkisch

Ergebnisgruppe	Punktwerte (wirkliche Werte)		
(Testergebnis)	Fb	Bf	Gt
Fb	77.9	21.8	.3
Bf	11.6	72.0	16.4
Gt	.1	15.6	84.3

Tabelle 6.10 Reliabilitätstabelle TV Test Türkisch

Ergebnisgruppe (Testergebnis)	Punktwerte (wirkliche Werte)		
	Fb	Bf	Gt
Fb	73.4	24.7	1.9
Bf	20.3	56.7	23.0
Gt	1.1	20.1	78.8

Die Tabellen 6.7, 6.8, 6.9 und 6.10 geben die Messgenauigkeit der Subtests der türkischen Testversion im Bezug auf die gewählte Normierung wieder. Auffallend ist, dass bei PW der Prozentsatz an richtigen Klassifikationen am höchsten ist. 81,2% der Kinder, die aufgrund ihrer Punktwerte als "förderbedürftig" eingestuft werden, befinden sich auch mit ihrem wirklichen Testergebnis innerhalb dieser Normierungskategorie. Anders ausgedrückt: Die Wahrscheinlichkeit, dass das Testergebnis eines Kindes zu Recht als "förderbedürftig" eingestuft wird, liegt bei 81,2%.

18,8% der Kinder sind inkorrekt eingestuft worden (in Wirklichkeit 16,9% befriedigend und 1.9% gut).

In den folgenden acht Tabellen wird die Beziehung zwischen den Ergebnisgruppen aufgrund des wirklichen Testergebnisses und die übereinstimmende Ergebnisgruppe aufgrund der Punktwerte dargestellt. Die Angaben basieren auf allen Kindern, die die betreffende Testkomponente absolviert haben.

So zeigt Tabelle 6.12 zum Beispiel, dass 85,9% der Kinder, die aufgrund ihrer wirklichen Punktwerte als "förderbedürftig" eingestuft wurden, auch mit ihrem Testergebnis in diese Normierungskategorie fallen. Anders ausgedrückt: Die Wahrscheinlichkeit, dass ein Kind mit förderbedürftigen Fähigkeiten auch wirklich nach dem Testen als förderbedürftig eingestuft wird, liegt bei 85.9%.

Tabelle 6.11 Reliabilitätstabelle PW Test Deutsch

Punktwerte (wirkliche Werte)	Ergebnisgruppe(Testergebnis)		
	Fb	Bf	Gt
Fb	88.6	11.4	.0
Bf	12.6	72.2	15.3
Gt	.0	6.9	93.1

Tabelle 6.12 Reliabilitätstabelle KB Test Deutsch

Punktwerte (wirkliche Werte)	Ergebnisgruppe(Testergebnis)		
	Fb	Bf	Gt
Fb	85.9	13.8	.3
Bf	18.7	61.8	19.5
Gt	.2	7.9	91.9

Tabelle 6.13 Reliabilitätstabelle PB Test Deutsch

Punktwerte (wirkliche Werte)	Ergebnisgruppe(Testergebnis)		
	Fb	Bf	Gt
Fb	83.7	16.2	.1
Bf	12.9	71.0	16.1
Gt	.0	9.4	90.6

Tabelle 6.14 Reliabilitätstabelle TV Test Deutsch

Punktwerte	Ergebnisgruppe(Testergebnis)		
(wirkliche Werte)	Fb	Bf	Gt
Fb	78.8	20.7	.6
Bf	16.2	63.5	20.2
Gt	.1	11.1	88.7

Tabelle 6.15 Reliabilitätstabelle KB Test Türkisch

Punktwerte	Ergebnisgruppe(Testergebnis)		
(wirkliche Werte)	Fb	Bf	Gt
Fb	83.0	16.8	.2
Bf	17.9	62.3	19.7
Gt	.3	12.6	87.1

Tabelle 6.16 Reliabilitätstabelle PB Test Türkisch

Punktwerte	Ergebnisgruppe(Testergebnis)		
(wirkliche Werte)	Fb	Bf	Gt
Fb	82.9	16.9	.2
Bf	15.0	67.8	17.2
Gt	.2	14.2	85.6

Tabelle 6.17 Reliabilitätstabelle PW Test Türkisch

Punktwerte	Ergebnisgruppe(Testergebnis)		
(wirkliche Werte)	Fb	Bf	Gt
Fb	84.4	14.3	1.3
Bf	27.0	46.9	26.1
Gt	1.3	10.6	88.1

Tabelle 6.18 Reliabilitätstabelle TV Test Türkisch

Punktwerte	Ergebnisgruppe(Testergebnis)		
(wirkliche Werte)	Fb	Bf	Gt
Fb	79.4	19.1	1.6
Bf	24.9	49.8	25.3
Gt	1.6	16.6	81.8

6.4 Abschließende Wahrnehmungen

Wie in diesem Kapitel beschrieben, folgt man einem sorgfältigen Verfahren, um eine Normierung für den Sprachtest zu erstellen. Auch die direkt Beteiligten in Duisburg haben in wichtigem Maße zu diesem Resultat beigetragen.

Im Allgemeinen bestimmen Normierungen letztlich, dass etwa 20% der am schwächsten leistenden Testkandidaten das Prädikat förderbedürftig bekommen. Festzustellen ist, dass dies auch auf den Sprachtest zutrifft.

Im Bezug auf das Funktionieren der Normierung sind auch folgende Bemerkungen wichtig: Zwar basiert die gewählte Normierung u.a. auf der systematischen Sammlung der Daten der Kinder und geht aus der Normierungsforschung (siehe Abschnitt 6.3) hervor, dass die Beziehung zwischen den wirklichen Punktwerten und den Testergebnissen sehr gut ist. Gleichzeitig dienen aber immer auch die Wahrnehmung und Beurteilung der Lehrer als Ausgangspunkte für die Normierung. Es ist also nachdrücklich in Betracht gezogen worden, dass, wenn Grund dazu besteht, ein bestimmtes Ergebnis im Nachhinein angeglichen werden kann.

7 Die im Sprachtest erzielten Ergebnisse

In diesem Kapitel werden zunächst die Resultate in den deutschen Testkomponenten aufgeführt (7.1). Aufgrund der Normierung der deutschen und türkischen Testversion wurde erfasst, wie viele Kinder je Gruppe das Prädikat förderbedürftig, befriedigend und gut erlangt haben. Neben der Frequenz werden auch die Prozentsätze gegeben. Diese zeigen zum Beispiel, wie viel Prozent der Gruppe Deutsch in der Testkomponente KB gut abschneidet, aber auch wie viel Prozent der gesamten Population mit gut abgeschnitten hat. In diesen Angaben sind alle Kinder, die den Test absolviert haben, mitgezählt. In dem Sinne weichen die Frequenzen etwas ab von denen, die in Kapitel 4 und 5 aufgeführt sind. In diesen Kapiteln ging man von den "Netto"-Frequenzen aus, das heißt von Kindern, die alle vier Testkomponenten des Tests absolviert haben, weil in den betreffenden Kapiteln vor allem auf eine gute Analyse geachtet worden ist (siehe auch Erklärung bei Tabelle 6.2).

In diesem Kapitel geht es um alle Kinder, die ein Ergebniszeugnis von Cito bekommen haben. In Abschnitt 7.2 werden die Ergebnisse in den türkischen Testkomponenten aufgeführt. Danach zeigt Abschnitt 7.3 die Ergebnisübersichten in beiden Testversionen. In Abschnitt 7.4 werden einige abschließende Erklärungen formuliert.

7.1 Ergebnisse in der deutschen Testversion

In diesem Abschnitt werden die Ergebnisse in der deutschen Testversion je Testkomponente und je Gruppe aufgeführt.

Tabelle 7.1 Ergebnisse in PW Deutsch

	Deutsch	Deutsch-Türkisch	Deutsch-Anders	Anders	Türkisch	Total
Förderbedürftig	207	76	76	130	558	1047
	7,6%	27,7%	16,9%	46,9%	60,3%	22,5%
Befriedigend	252	90	99	69	247	757
	9,2%	32,8%	22,0%	24,9%	26,7%	16,2%
Gut	2274	108	275	78	120	2855
	83,2%	39,4%	61,1%	28,2%	13,0%	61,3%
Insgesamt	2733	274	450	277	925	4659
	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Tabelle 7.1 zeigt, dass insgesamt 4.659 Kinder die Testkomponente PW absolviert haben; 1.047 davon (22,5%) sind als "förderbedürftig" eingestuft worden.

Dies ist die Rangordnung nach Teilgruppen: die Gruppe Türkisch (558 Kinder); Deutsch (207 Kinder); Anders (130 Kinder); Deutsch-Anders (76 Kinder); Deutsch-Türkisch (76 Kinder). Die Reihenfolge verändert sich, wenn man die Prozentsätze der Sprachgruppen in Betracht zieht:

60,3% der Kinder der Gruppe Türkisch; 46,9% der Kinder der Gruppe Anders, 27,7 % der Kinder der Gruppe Deutsch-Türkisch, 16,9% der Kinder der Gruppe Deutsch-Anders und 7,6% der Kinder der Gruppe Deutsch sind als "förderbedürftig" eingestuft worden.

Tabelle 7.2 Ergebnisse in KB Deutsch

	Deutsch	Deutsch-Türkisch	Deutsch-Anders	Anders	Türkisch	Total
Förderbedürftig	228	76	80	109	492	985
	8.6%	28.4%	18.2%	40.8%	54.4%	21.8%
Befriedigend	296	80	83	74	242	775
	11.2%	29.9%	18.9%	27.7%	26.7%	17.1%
Gut	2123	112	277	84	171	2767
	80.2%	41.8%	63.0%	31.5%	18.9%	61.1%
Insgesamt	2647	268	440	267	905	4527
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Tabelle 7.2 zeigt, dass insgesamt 4.527 Kinder die Testkomponente KB absolviert haben; 985 davon (21,8%) sind als "förderbedürftig" eingestuft worden.

Dies ist die Rangordnung nach Teilgruppen: die Gruppe Türkisch (492 Kinder); Deutsch (228 Kinder); Anders (109 Kinder); Deutsch-Anders (80 Kinder); Deutsch-Türkisch (76 Kinder). Die Reihenfolge verändert sich, wenn man die Prozentsätze der Sprachgruppen in Betracht zieht:

54.4% der Kinder der Gruppe Türkisch; 40.8% der Kinder der Gruppe Anders, 28.4 % der Kinder der Gruppe Deutsch-Türkisch, 18.2% der Kinder der Gruppe Deutsch-Anders und 8.6% der Kinder der Gruppe Deutsch sind als "förderbedürftig" eingestuft worden.

Tabelle 7.3 Ergebnisse in PB Deutsch

	Deutsch	Deutsch-Türkisch	Deutsch-Anders	Anders	Türkisch	Total
Förderbedürftig	403	81	88	75	288	935
	15.6%	30.3%	20.5%	29.1%	32.6%	21.1%
Befriedigend	599	88	140	105	358	1290
	23.1%	33.0%	32.6%	40.7%	40.5%	29.1%
Gut	1588	98	202	78	237	2203
	61.3%	36.7%	47.0%	30.2%	26.8%	49.8%
Insgesamt	2590	267	430	258	883	4428
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Tabelle 7.3 zeigt, dass insgesamt 4.428 Kinder die Testkomponente PB absolviert haben; 935 davon (21,1%) sind als "förderbedürftig" eingestuft worden.

Dies ist die Rangordnung nach Teilgruppen: die Gruppe Deutsch (403 Kinder); Türkisch (288 Kinder); Deutsch-Türkisch (81 Kinder); Deutsch-Anders (88 Kinder); Anders (75 Kinder). Die Reihenfolge verändert sich, wenn man die Prozentsätze der Sprachgruppen in Betracht zieht:

32,6% der Kinder der Gruppe Türkisch; 30,3% der Kinder der Gruppe Deutsch-Türkisch, 29,1% der Kinder der Gruppe Anders, 20,5% der Kinder der Gruppe Deutsch-Anders und 15,6% der Kinder der Gruppe Deutsch sind als "förderbedürftig" eingestuft worden.

Auffallend ist, dass der Prozentsatz "förderbedürftig" der Gruppe Deutsch in dieser Testkomponente höher ist als in den anderen Testkomponenten. Bei der Gruppe Türkisch ist das genau umgekehrt; diese Gruppe

schneidet mit dem niedrigsten Prozentsatz in dieser Testkomponente im Vergleich zu den anderen Testkomponenten ab.

Tabelle 7.4 Ergebnisse in Deutsch TV

	Deutsch	Deutsch-Türkisch	Deutsch-Anders	Anders	Türkisch	Total
Förderbedürftig	206	75	73	87	402	843
	8.1%	28.4%	17.3%	34.9%	46.6%	19.4%
Befriedigend	472	99	125	85	339	1119
	18.4%	37.5%	29.6%	34.1%	39.3%	25.7%
Gut	1879	90	224	77	122	2392
	73.5%	34.1%	53.1%	30.9%	14.1%	54.9%
Insgesamt	2556	264	422	249	863	4354
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Tabelle 7.4 zeigt, dass insgesamt 4.354 Kinder die Testkomponente TV absolviert haben; 843 davon (19,4%) sind als "förderbedürftig" eingestuft worden.

Dies ist die Rangordnung nach Teilgruppen: die Gruppe Türkisch (402 Kinder); Deutsch (206 Kinder); Anders (87 Kinder); Deutsch-Türkisch (75 Kinder); Deutsch-Anders (73 Kinder). Die Reihenfolge verändert sich, wenn man die Prozentsätze der Sprachgruppen in Betracht zieht:

46,6% der Kinder der Gruppe Türkisch; 34,9% der Kinder der Gruppe Anders, 28,4% der Kinder der Gruppe Deutsch-Türkisch, 17,3% der Kinder der Gruppe Deutsch-Anders und 8,16% der Kinder der Gruppe Deutsch sind als "förderbedürftig" eingestuft worden.

7.2 Ergebnisse in der türkischen Testversion

In diesem Abschnitt werden die Ergebnisse in der türkischen Testversion je Testkomponente und je Gruppe aufgeführt.

Tabelle 7.5 Ergebnisse in PW Türkisch

	Deutsch-Türkisch	Türkisch	Total
Förderbedürftig	48	165	213
	23.5%	23.5%	23.5%
Befriedigend	49	169	218
	24.0%	24.1%	24.1%
Gut	107	367	474
	52.5%	52.4%	52.4%
Insgesamt	204	701	905
	100.0%	100.0%	100.0%

Tabelle 7.5 zeigt, dass insgesamt 905 Kinder die Testkomponente PW absolviert haben. 213 von ihnen (23,5%) sind als "förderbedürftig" eingestuft worden. In der Gruppe Türkisch hatten 165 Kinder das Ergebnis "förderbedürftig", also mehr als in der Gruppe Deutsch-Türkisch (48 Kinder). Die Prozentsätze der förderbedürftigen Ergebnisse waren für beide Gruppen gleich: 23%.

Tabelle 7.6 Ergebnisse in KB Türkisch

	Deutsch-Türkisch	Türkisch	Total
Förderbedürftig	38	163	201
	18.9%	23.4%	22.4%
Befriedigend	62	245	307
	30.8%	35.2%	34.2%
Gut	101	289	390
	50.2%	41.5%	43.4%
Insgesamt	201	697	898
	100.0%	100.0%	100.0%

Tabelle 7.6 zeigt, dass insgesamt 898 Kinder die Testkomponente KB absolviert haben. 201 von ihnen (22,4%) sind als "förderbedürftig" eingestuft worden. In der Gruppe Türkisch waren mehr "förderbedürftige" Ergebnisse (163 Kinder, 23,4%) als in der Gruppe Deutsch-Türkisch (38 Kinder, 18,9%).

Tabelle 7.7 Ergebnisse in PB Türkisch

	Deutsch-Türkisch	Türkisch	Total
Förderbedürftig	42	158	200
	21.0%	22.9%	22.5%
Befriedigend	60	331	391
	30.0%	48.0%	44.0%
Gut	98	200	298
	49.0%	29.0%	33.5%
Insgesamt	200	689	889
	100.0%	100.0%	100.0%

Tabelle 7.7 zeigt, dass insgesamt 889 Kinder die Testkomponente PB absolviert haben. 200 von ihnen (22,5%) sind als "förderbedürftig" eingestuft worden. In der Gruppe Türkisch waren mehr "förderbedürftige" Ergebnisse (158 Kinder, 22,9%) als in der Gruppe Deutsch-Türkisch (42 Kinder, 21,0%).

Tabelle 7.8 Ergebnisse in TV Türkisch

	Deutsch-Türkisch	Türkisch	Total
Förderbedürftig	34	195	229
	17.2%	28.6%	26.0%
Befriedigend	82	296	378
	41.4%	43.4%	43.0%
Gut	82	191	273
	41.4%	28.0%	31.0%
Insgesamt	198	682	880
	100.0%	100.0%	100.0%

Tabelle 7.8 zeigt, dass insgesamt 880 Kinder die Testkomponente TV absolviert haben. 229 von ihnen (26,0%) sind als "förderbedürftig" eingestuft worden. In der Gruppe Türkisch waren mehr "förderbedürftige" Ergebnisse (195 Kinder, 28,6%) als in der Gruppe Deutsch-Türkisch (34 Kinder, 17,2%).

7.3 Übersicht der Ergebnisse in beiden Testversionen

Dieser Abschnitt gibt die Ergebnisse in der deutschen Testversion je Testkomponente und je Gruppe wieder. Daneben werden auch die Ergebnisse aller Kinder angegeben, die die türkische Testversion absolviert haben, nämlich die Kinder der zwei türkischsprachigen Gruppen Deutsch-Türkisch und Türkisch. Ziel dieses Abschnittes ist es zu verdeutlichen, wie die Kinder in den Testkomponenten abgeschnitten haben. Faktisch ist in der Reihe die mit 0 beginnt sichtbar, wie viele Kinder in alle Testkomponenten mit "gut" oder "befriedigend" abgeschnitten haben; in der Reihe die mit 1 beginnt, wie viele Kinder in einer Testkomponente mit "förderbedürftig" abgeschnitten haben; in der Reihe die mit 2 beginnt, wie viele Kinder in zwei Testkomponenten ... usw.

Anders als in den Abschnitten 6.1 und 6.2 wird hier also kein Unterschied pro Testkomponente gemacht. Zuletzt ist zu beachten, dass in diesen Tabellen nur die Kinder aufgeführt sind, die alle vier Testkomponenten absolviert haben.

Tabelle 7.9 Ergebnisse aller Gruppen in allen Testkomponenten der deutschen Testversion

	Deutsch	Deutsch/ Türkisch	Deutsch/ Anders	Anders	Türkisch/ Deutsch	Total
Absolute Zahl	1951	113	271	101	197	2633
% in der Gruppe	76.3%	42.8%	64.2%	40.6%	22.8%	60.5%
Absolute Zahl	453	81	86	53	237	910
% in der Gruppe	17.7%	30.7%	20.4%	21.3%	27.5%	20.9%
Absolute Zahl	100	41	42	50	190	423
% in der Gruppe	3.9%	15.5%	10.0%	20.1%	22.0%	9.7%
Absolute Zahl	41	24	20	29	167	281
% in der Gruppe	1.6%	9.1%	4.7%	11.6%	19.4%	6.5%
Absolute Zahl	11	5	3	16	72	107
% in der Gruppe	.4%	1.9%	.7%	6.4%	8.3%	2.5%
Absolute Zahl	2556	264	422	249	863	4354
Insgesamt	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Die Tabelle 7.9 liefert eine Übersicht der Ergebnisse aller Gruppen in der deutschen Testversion.

Die Angaben in Reihe 0 zeigen, dass 76,3% der Gruppe Deutsch, 42,8% der Gruppe Deutsch-Türkisch, 64,2% der Gruppe Deutsch-Anders, 40,6% der Gruppe Anders und 22,8% der Gruppe Türkisch in allen Testkomponenten mit "gut" oder "befriedigend" abgeschnitten haben. Diese Kinder kommen nicht für zusätzlichen Sprachunterricht in Frage, aufgrund der Vorschrift, die mit der Normierung festgelegt worden ist (wenn in einer Testkomponente mit "förderbedürftig" abgeschnitten wird, bekommt das Kind zusätzlichen Sprachunterricht).

Insgesamt betrifft das 2.633 Kinder, die 60,5% der totalen Population ausmachen.

Auffallend ist, dass beide türkischsprachigen Gruppen in Reihe 1 (in einer Testkomponente das Ergebnis "förderbedürftig") im Vergleich zu den anderen Gruppen einen höheren Prozentsatz aufweisen.

Die Prozentsätze der Kinder, die in allen Testkomponenten mit Ergebnis "förderbedürftig" abgeschnitten haben (Reihe 4) sind auch auffallend: Insgesamt geht es um 2,5% der totalen Population, die folgendermaßen unterteilt ist: die Gruppe Türkisch 72 Kinder (8.3 %); die Gruppe Anders 16 Kinder (6.4%), die Gruppe Deutsch-Türkisch 5 Kinder (1,9%), Deutsch-Anders 3 Kinder (0.7 %) und die Gruppe Deutsch 11 Kinder (0.4 %).

7.4 Abschließende Erklärungen

Wie in diesem Kapitel dargestellt, geht aus den Übersichten der Leistungen der Kinder hervor, dass die Rangordnung der fünf Gruppen die gleiche ist wie die Rangordnung, die die Ergebnisse aufweisen (siehe Kapitel 5):

- 1 die Gruppe Deutsch
- 2 die Gruppe Deutsch-Anders
- 3 die Gruppe Deutsch-Türkisch
- 4 die Gruppe Anders
- 5 die Gruppe Türkisch.

Was die Testkomponente PB der deutschen Testversion angeht, gibt es noch etwas zu erklären.

Anderssprachige Kinder beherrschen die Fähigkeit deutsche Laute zu unterscheiden offensichtlich besser als die anderen Sprachfähigkeiten. Dieser Befund stimmt überein mit der Theorie, dass Vorschulkinder die Laute problemlos erlernen können. Der Befund illustriert die Tatsache, dass Zweisprachigkeit positive Effekte haben kann, auf den Prozess des Erlernens einer Sprache.

Aus den Leistungen in der türkischen Testversion geht hervor, dass die Gruppe Deutsch-Türkisch bessere Ergebnisse aufweist als die Gruppe Türkisch.

Es ist interessant feststellen zu können, dass Kinder in der Lage sind, ihre eigenen Fähigkeiten in Deutsch zu beurteilen: Kinder die aussagten, Deutsch besser zu beherrschen als die eigene Sprache, haben tatsächlich mehr "gute" oder "befriedigende" Ergebnisse erzielt als Kinder, die sagten, dass sie die eigene Sprache besser beherrschen als die deutsche.

Zusammengefasst: Die Ergebnisse bieten eine klare Übersicht der Fähigkeiten der Kinder. Diese Tatsache bestätigt die Zuverlässigkeit des Tests und der sorgfältigen Normierung.

8 Schlussfolgerungen und Empfehlungen

In diesem Kapitel werden die Ergebnisse zusammengefasst und erörtert, die sich im Anschluss an die Forschung aufgrund der Fragen in Kapitel 3 ergeben haben. In Abschnitt 8.1 werden die Befunde zu der „taalkaart“ (Sprachatlas) von Duisburg aufgeführt. Der Abschnitt 8.2 widmet seine Aufmerksamkeit den Duisburger Kindern in der Vorschulphase. Dann werden in Abschnitt 8.3 Empfehlungen gegeben für weitere Forschungen und in Abschnitt 8.4 für die Unterrichtspraxis.

8.1 Sprachatlas (taalkaart) von Duisburg

Die Forschung im Bezug auf den Sprachhintergrund der Kinder, worüber hauptsächlich in Kapitel 4 berichtet worden ist, bestätigt noch einmal, dass Duisburg eine mehrsprachige Stadt ist. Die Befunde aus Kapitel 5, in dem die Ergebnisse je Sprachgruppe dargestellt werden, zeigen, dass die Unterverteilung in fünf Sprachgruppen ein Mehr an Informationen gebracht hat.

Es ist interessant, dass die Kinder aufgrund des so genannten Heim-Sprachen-Kriteriums in diese Sprachgruppen eingeteilt werden konnten. Wie in 4.3. festgestellt, hat das Geburtsland-Kriterium hierfür keine aussagekräftigen Informationen erbracht.

Dennoch gab nur die Feststellung, dass es sich um mehrsprachige Kinder handelt, noch unzulängliche Angaben über die Sprachfähigkeit dieser Kinder. Insbesondere die Frage, welche Sprache die Kinder am besten sprechen, muss demnach noch beantwortet werden. Diesbezüglich wurden, sowohl was die Heimsprache betrifft als auch was die zweite Sprache betrifft, Unterschiede in der Sprachfähigkeit innerhalb der zweisprachigen Gruppen wahrgenommen.

8.2 Sprachprofile der Kinder

Aus der Forschung geht hervor, dass Kinder in der Lage sind, ihre eigene Sprachfähigkeit zu beurteilen. Die zweisprachigen Kinder der Gruppe Deutsch-Anders, die ausgesagt hatten, dass sie besser seien in Deutsch als in ihrer eigenen Sprache, schnitten in allen Testkomponenten auch wirklich besser ab als die Kinder der Gruppe Anders.

Das Gleiche sieht man bei der Gruppe Deutsch-Türkisch, die in allen Testkomponenten auf Deutsch signifikant besser abschnitt als die Gruppe Türkisch.

Über die Sprachfähigkeit in der Heimsprache der zweisprachigen Kinder aus den Gruppen Deutsch-Anders und Anders gibt es keine Angaben. Es kann also keine Aussage über Art und Ausmaß der Zweisprachigkeit dieser Gruppen getroffen werden.

Weil die Ergebnisse der beiden türkischsprachigen Gruppen (Deutsch-Türkisch und Türkisch) sowohl in der deutschen als auch der türkischen Testversion zur Verfügung stehen, kann über Art und Ausmaß der Zweisprachigkeit dieser Gruppen dahingegen schon eine Aussage gemacht werden.

Kinder der Gruppe Türkisch schnitten besser in der türkischen Testversion als in der deutschen Testversion ab. Auffallend ist aber, dass die Gruppe Deutsch-Türkisch auch in der türkischen Testversion besser abschnitt als die Gruppe Türkisch. Die Gruppe Deutsch-Türkisch ist also nicht nur befähigt in Deutsch, sondern auch in Türkisch.

Und es ist noch eine Bemerkung zum Sprachprofil der Gruppe Deutsch-Türkisch zu machen. Obwohl die Gruppe vorgibt, besser zu sein in Deutsch als in Türkisch, stimmt das nicht ganz überein mit den tatsächlichen Ergebnissen: In zwei Testkomponenten beider Testversionen sind keine signifikanten Unterschiede zu erkennen und in zwei anderen Testkomponenten schneiden sie in Türkisch sogar signifikant besser ab. In dem Sinne können diese Kinder balanciert zweisprachige Kinder genannt werden. Die Bemerkung über die balancierte Zweisprachigkeit der Gruppe Deutsch-Türkisch kann möglicherweise in Zusammenhang gebracht werden mit einer Wahrnehmung aus früherer Forschung (Cummins, Hakuta, Narain und Verhoeven), dass ausgeglichenen zweisprachigen Kindern ein hohes metalinguistisches Bewusstsein zur Verfügung steht.

In diesem Zusammenhang sind die Ergebnisse in der Testkomponente PB interessant. Die Kinder der monolingualen Gruppe Deutsch schneiden absolut gesehen zwar am besten in dieser Testkomponente ab, aber im Vergleich zu anderen Testkomponenten relativ niedrig in der Testkomponente PB. Interessant ist, dass die zweisprachigen Kinder aller vier Gruppen in dieser Testkomponente gut abschneiden.

Betrachtet man die Ergebnisse aus internationaler Perspektive (im Konkreten aus türkischer und niederländischer Perspektive), dann zeigt sich, dass der Abstand der Sprachentwicklung von Kindern im Alter von vier bis sechs Jahren in diesen Ländern größer wird gegenüber den Duisburger Kindern. Dies ist mit der Tatsache zu erklären, dass die Kinder in den Niederlanden und in der Türkei schon im genannten Alter unterrichtet werden, während Kinder in Deutschland erst ab sechs Jahren am Unterricht teilnehmen. Was Zuverlässigkeit betrifft, können die Tests als gut bis sehr gut bezeichnet werden. Das zeigt sich auch bei der Genauigkeitsmessung im Bezug auf die Normierung.

8.3 Empfehlungen für weitere Forschung

In weiterer Forschung können mehr Testfragen in den Testkomponenten aufgenommen werden, um den Test gegenwartsnahe bleiben zu lassen. Außerdem kann man an das adaptive Umgestalten der Tests denken.

Ferner wird empfohlen, die Gültigkeit der Normierung auch für andere Situationen zu prüfen. Der Sprachtest lässt sich zum Beispiel nach zehn Monaten (die Kinder gehen dann im Prinzip in die Grundschule) nochmals absolvieren. Das Analysieren der Daten einer solchen Abnahme ist zu empfehlen, ebenso wie das Durchführen einer Validierungsstudie der Normierung, indem die Ergebnisse des Sprachtests mit dem Befund der Lehrer in der Klasse verglichen werden.

8.4 Empfehlungen für die Unterrichtspraxis

Die Forschung zeigt, dass die Gruppen (mit ein- und mehrsprachigen Kindern) Unterschiede in ihren Sprachprofilen aufweisen.

Um der Entwicklung von Kindern beim Schulanfang entgegenkommen zu können, brauchen Lehrkräfte Einschätzungsmethoden. Differenzierung und Individualisierung innerhalb des Sprachunterrichts sind nur wirklich möglich, wenn den Lehrern Instrumente zur Verfügung stehen, die konkrete Indizien bezüglich Sprachentwicklung beziehungsweise Art und Ausmaß von Zweisprachigkeit liefern. Wenn man den Sprachtest zwei Mal im Jahr absolvieren lässt, kann man die Sprachentwicklung von Kindern nachvollziehen. Außerdem lässt sich dieses Instrument mit anderen in Deutschland vertretenen Sprachgruppen ergänzen.

Literaturliste

Broeder, P & G. Extra (2005), *Minderheidsgroepen en Minderheidstalen*, Vereniging van Nederlandse Gemeenten, 1995.

Cor Sluijter (2002), *De ontwikkeling van computer gestuurde adaptieve toetsen binnen het Leerlingvolgsysteem van het Cito*. In: Proceedings van de 29ste Onderwijs Research Dagen 2002 te Antwerpen. Samenstellers: Fr. Daems e.a., Antwerpen: Universiteit Antwerpen 2002.

Cummins J. & M. Swain (1986), *Bilingualism in education*. Londen: Longman.

Extra, G & K. Yağmur (2003), *Urban Multilingualism in Europe, Immigrant Minority Languages at Home and School*, Tilburg University.

Hakuta R. (1986) *Mirror of language: the debate on bilingualism*. Newyork: Basic Boks

Konak, Ö.A, (1996), *Taalontwikkeling van Turkse kinderen in Turkije en in Nederland*, Doctoraalscriptie, Tilburg University.

Kuyk, J. van & L. Verhoeven (1991), Conceptuele en metalinguïstische kennis bij jongste kleuters. In: B. van Hout & L.de Klerk (eds), *Onderwijsleerprocessen; cognitie en motivatie*. Amsterdam SCO

Narain, G & L.Verhoeven (1994), *Ontwikkeling van tweetaligheid bij allochtone kleuters. Studies in Meertaligheid 5*. Tilburg University Press.

Verhoeven, L. G. Extra, Ö.A. Konak, Goretti Narain, R. Zerrouk (1992), *Tweetaligheid bij allochtone kleuters. Studies in Meertaligheid 2*. Tilburg University Press.

Kurzfassung von 2 Artikeln zum Sprachtest

1 Kurzfassung Teil 1 der Vorstudie Sprachtest (Dezember 2004)

Obgleich der Rapport der Vorstudie als separater Artikel herausgegeben worden ist, fanden die Autoren es geeignet und zweckmäßig, in dieser Kurzfassung auf die Vorstudie einzugehen. Diese Fassung beschreibt den ganzen Prozess der Vorstudie, die von Juli 2003 bis einschließlich der Hauptforschung im November/Dezember 2003 stattgefunden hat.

Die Stadt Duisburg war bei ihrer Suche (Herbst 2002) nach einem Sprachtest, der bei der Einschulung nachweisen sollte, ob Kinder zusätzlichen Unterricht in der deutschen Sprache benötigten, vom Landesinstitut für Schule in Soest (NRW) auf einen niederländischen Test (Test Zweisprachigkeit) aufmerksam gemacht worden.

Obwohl dieser Test Zweisprachigkeit im niederländischen Kontext vor allem angewandt wurde zur Messung der Sprachfähigkeit von zweisprachigen Kindern, hat sich der Test auch als geeignet erwiesen, um die Sprachfähigkeit von monolingual niederländischsprachigen Kindern in Niederländisch zu messen.

In Duisburg setzt man den Sprachtest ein, um den Sprachstand in Deutsch von sowohl monolingualen deutschsprachigen Kindern als auch anderssprachigen Kindern zu messen. Der Stadt war es wichtig, auch die türkischsprachige Testversion einzusetzen. Die zweisprachigen türkischsprachigen Kinder können sowohl die deutschsprachige wie auch die türkischsprachige Testversion absolvieren. Die Stadt Duisburg stellte bestimmte Anforderungen an einen Sprachtest (u.a. Zuverlässigkeit, Präzision, Zweisprachigkeit), die man im Test Zweisprachigkeit zum größten Teil realisiert sah. Die Anforderung eines umfangreichen, praktischen Einsatzes konnte man jedoch nicht erfüllen. Aus diesem Grund wurde der Test zu einem Computertest weiterentwickelt.

Der Test Zweisprachigkeit wurde zwischen 1989 und 1992 in Zusammenarbeit mit der Universität Tilburg und CITO entwickelt. Ziel dieses so genannten "Bleistift und Papier-Tests" ist es, Art und Ausmaß der Zweisprachigkeit von Kindern im Alter von vier bis sieben Jahren festzustellen.

Um den Test Zweisprachigkeit zum Sprachtest weiterentwickeln zu können, mit anderen Worten zu einem Test, der wie gewünscht im deutschen Bildungskontext einsetzbar ist, waren eine Reihe von Maßnahmen vorzunehmen, u.a. den Test vom Inhalt her anzupassen, ihn zu digitalisieren (damit Großabnahme möglich wurde) und eine Vorstudie durchzuführen.

Ausgehend von der niederländischen und türkischen Version, wurde der Test für die deutsche Situation von einer Arbeitsgruppe weiterentwickelt, die sich aus Experten und Fachleuten der Stadt Duisburg, des Landesinstituts und der Universität Essen/Duisburg, zusammensetzte; sie wurden fachtechnisch von Cito-Experten begleitet. Im Rahmen der Vorstudie sind in der deutschen Testkomponente PW (Passiver Wortschatz) sieben Testfragen mit einer extra Alternative als Antwort aufgenommen worden, während in der türkischen Testversion in derselben Testkomponente eine Testfrage mit einer extra Alternative als Antwort enthalten ist. In diesem Zusammenhang sind aufgrund der Daten der Analyse die besten Testfragen für die endgültige Testversion ausgewählt worden. Für die Testkomponente KB hielt man derartige Maßnahmen nicht für notwendig.

Für die Testkomponente PB genügte eine kritische Übersetzung nicht. Aufgrund der Satzung, die dieser Testkomponente zugrunde liegt, wurde für Deutschland eine neue Testkomponente entwickelt. Die Texte der Testkomponente Textverständnis sind geringfügig bearbeitet worden, aber die Themen dieser kurzen Texte wurden nicht geändert.

Wie bereits vermerkt, wurde als Lösung des Problems, dass viele Kinder in kurzer Zeit getestet werden mussten, der Test außerdem zu einem Computertest weiterentwickelt. Da diese Entwicklung auch Folgen für die Art der Testpräsentation hatte, bewertete die genannte Arbeitsgruppe auch die Resultate dieser Entwicklung. Aufgrund dieser Bewertungsergebnisse sind mehrere Anpassungen durchgeführt worden, insbesondere im Hinblick auf die Darbietung der Testaufgaben mit Hilfe von Abbildungen und die Art der Interaktion zwischen Kind und Computer. Die Arbeitsgruppe war nicht ohne weiteres davon überzeugt, dass ein Computertest für die Messung der Sprachfähigkeit von Kindern geeignet ist. Aus dem Grund zog man zwei Bestandteile des Tests Zweisprachigkeit, die nicht digitalisiert worden waren, in die Vorstudie mit ein.

Es beteiligten sich gut 400 Kinder an der Vorstudie, die im Juli 2003 stattfand. Aufgrund eines Fragebogens unterschied man folgende Gruppen:

- monolinguale deutschsprachige Kinder;
- bilinguale Kinder mit Türkisch als erster Sprache;
- bilinguale Kinder mit "einer anderen" Sprache als erster Sprache.

Etwa 160 monolingual deutschsprachige Kinder nahmen an der Studie teil; von den ca. 217 teilnehmenden türkischsprachigen Kindern haben 70 sowohl die deutsche als auch die türkische Testversion absolviert. 30 Kinder gehörten zu der Gruppe mit einer "anderen Sprache".

Die Fragen, auf die diese Vorstudie Antwort geben sollte, waren:

- Ist der Sprachtest in ausreichendem Maß valide und zuverlässig?
- Treten keine Höchstgrenzeffekte auf?
- Existiert eine Korrelation zwischen den vier rezeptiven (digitalen) und den zwei produktiven Testkomponenten?

Etwas getrennt von diesen Forschungsfragen war auch wichtig in Erfahrung zu bringen, wie die Kinder auf den Computertest reagieren würden. Anlässlich dieser Fragen haben die Testleiter ihre Observationen der Kinder während der Testabnahmen protokolliert und darüber referiert. Sie bewerteten nuanciert positiv, dass die Kinder in der Lage sind, diesen Test am Computer zu machen.

Mit Hilfe verschiedener statistischer Bearbeitungen hat man während der Ergebnisanalyse die Durchschnittswerte und die Standarddeviation festgestellt. Es wurde die interne Konsistenz und Homogenität der Testfragen kontrolliert. Aufgrund der damit festgestellten Angaben, konnten die Fragen nach Validität, Reliabilität und Korrelation zwischen den digitalen und den mündlichen Testkomponenten positiv beantwortet werden. Diese Angaben waren ein zusätzliches Zeichen dafür, dass die Kinder während der Computertests keine großen Probleme zeigten.

Zuletzt wurde festgestellt, dass keine Höchstgrenzeffekte auftraten.

2 Kurzfassung Teil 2 des Forschungsberichtes Sprachtest (Dezember 2004)

In diesem Bericht werden die vier Testkomponenten des Sprachtests, der eine deutschsprachige und eine türkischsprachige Version umfasst, die übrigens nahezu identisch sind, in ihrer endgültigen Form präsentiert.

Ende 2003 wurde eine Forschung durchgeführt, die auf Forschungsfragen über folgende Themen basierte:

- *der Sprachatlas der Duisburger Kinder*
- *die Zuverlässigkeit des Tests*
- *die Sprachprofile in Deutsch der folgenden fünf zu unterscheidenden Sprachgruppen*

Deutsch	monolinguale deutschsprachige Kinder
Deutsch-Türkisch	zweisprachige türkischsprachige Kinder, die angegeben haben, dass ihre Fähigkeit in Deutsch stärker ist
Deutsch-Anders	zweisprachige Kinder mit einer "anderen Sprache" (nicht Deutsch, nicht Türkisch), die angegeben haben, dass ihre Fähigkeit in Deutsch stärker ist
Anders	zweisprachige Kinder mit einer "anderen Sprache" (nicht Deutsch, nicht Türkisch), die angegeben haben, dass ihre Fähigkeit in der "anderen Sprache" stärker ist
Türkisch	zweisprachige türkischsprachige Kinder, die angegeben haben, dass ihre Fähigkeit in Türkisch stärker ist.

- *die Sprachprofile der türkischsprachigen Kinder*
- *ein Vergleich der Ergebnisse in Duisburg mit früheren Ergebnissen aus der Türkei und aus den Niederlanden*
- *die Normierung*
- *die Implikationen für weitere Forschung und für die Unterrichtspraxis.*

Die Antworten auf die Fragen, die 5.330 Kinder im Rahmen der Forschung beantwortet haben, zeigen, dass 55,4% der Kinder monolingual deutsch sind und 44,6% zweisprachig. 26% der Auskunftgebenden hat Türkisch als Heimsprache angegeben. Neben oder anstelle von Deutsch werden von dieser Population noch 42 andere Sprachen gesprochen. Von diesen Sprachen werden 37 als "best gesprochene Sprache" bezeichnet. Die Frage nach dem Geburtsland der Kinder und ihrer Eltern hat keine relevante Informationen über die tatsächliche Sprachsituation hervorgebracht. So sagten gut 95% der türkischsprachigen Kinder, dass sie in Deutschland geboren sind.

Aus der Analyse ging hervor, dass die Messgenauigkeit der deutschen und der türkischen Testversion gut bis sehr gut war. Die Wahrnehmungen aus der Vorstudie wiesen in die gleiche Richtung.

Die Rangordnung der Fähigkeiten in Deutsch basiert auf den erzielten Ergebnissen, und ist in der deutschen Testversion je kodierte Sprachgruppe folgendermaßen:

- 1 Deutsch
- 2 Deutsch-Anders
- 3 Deutsch-Türkisch
- 4 Anders
- 5 Türkisch.

Beim Vergleich der Teilgruppen wurde ermittelt, zwischen welchen Teilgruppen und in welchen Testkomponenten die Ergebnisse signifikant auseinander gehen. Hierbei stellte sich heraus, dass die Gruppe Deutsch signifikant besser abschnitt als alle anderen Gruppen in allen Testkomponenten. Nach dieser Gruppe schnitt die Gruppe Deutsch-Anders in allen Testkomponenten signifikant besser ab als die Gruppen Deutsch-Türkisch, Anders und Türkisch. Die Gruppe Deutsch-Türkisch schnitt im Vergleich zur Gruppe Türkisch in allen Testkomponenten signifikant besser ab, aber diese Gruppe schnitt im Vergleich zur Gruppe Anders nur in den Testkomponenten PW und KB signifikant besser ab. Schliesslich schnitt die Gruppe Anders im Vergleich zur Gruppe Türkisch in drei Testkomponenten – PW, KB und TV – signifikant besser ab.

In der türkischen Testversion schnitt die Gruppe Deutsch-Türkisch in drei Testkomponenten signifikant besser ab als die Gruppe Türkisch ab. Nur der Unterschied in der Testkomponente PW war nicht signifikant.

Ein Vergleich der beiden türkischsprachigen Gruppen in beiden Testversionen wies nach, dass die Gruppe Deutsch-Türkisch in den Testkomponenten PW und TV der türkischen Testversion signifikant besser abschneidet als in denselben deutschen Testkomponenten. Die Gruppe Türkisch schnitt in drei Testkomponenten der türkischen Testversion signifikant besser ab als in denselben deutschen Testkomponenten.

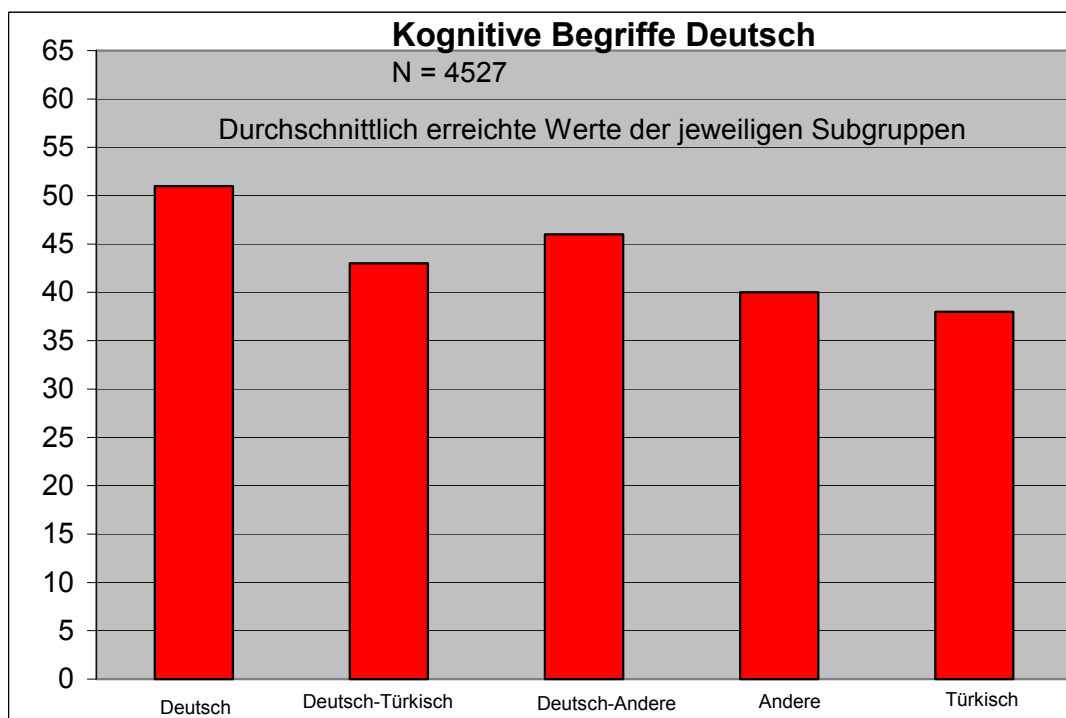
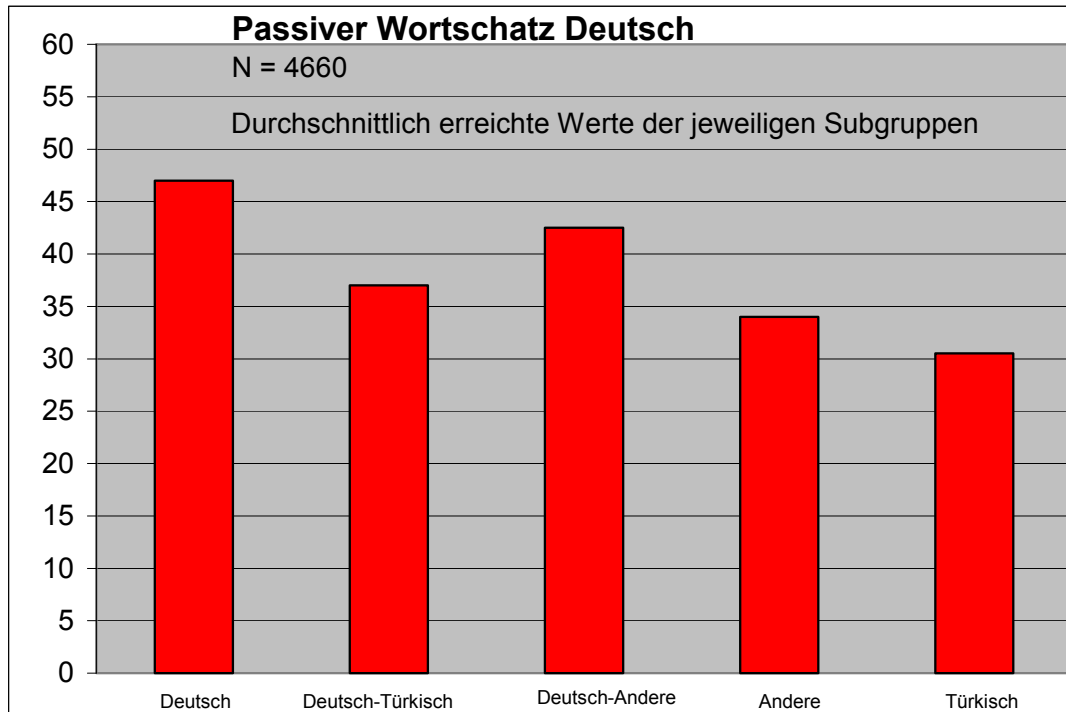
Die Normierung wurde während einer Sondersitzung in Duisburg mit Spezialisten und Fachleuten des Landesinstitutes, der Universität Essen/Duisburg, des Bildungswesens und dem politischen Vorstand der Stadt Duisburg festgestellt.

Die Cito-Stellungnahme, die dem Normierungsentschluss zugrunde lag, kam u.a. aufgrund von Befunden aus der Literatur im Bezug auf Fragen wie "Wie groß muss der Wortschatz eines Kindes sein, um dem Unterricht gewachsen zu sein?" zustande. Darüber hinaus waren die Normierung des Tests Zweisprachigkeit und die Erfahrungen damit in Betracht gezogen worden. Die Normierung wurde in Duisburg mit Berücksichtigung des Ausgangspunktes eingesetzt, dass ein Kind Anspruch auf zusätzlichen Sprachunterricht hat, wenn es in einer Testkomponente als "förderbedürftig" eingestuft wird. Aus genauer Analyse der Normierung ging hervor, dass die Relation zwischen den wirklichen Ergebnissen und den Testergebnissen sehr gut ist. Aus den Leistungen der Kinder nach der Normierung war ersichtlich, dass sich die Rangordnung nicht geändert hat, also so war, wie in dieser Kurzfassung dargestellt. Man kann vorsichtig den Schluss ziehen, dass die zweisprachigen Kinder verhältnismäßig gut in der Lage zu sein scheinen, ihre eigene Sprachfähigkeit einzuschätzen. Am deutlichsten wurde dies bei den Kindern der Gruppe Türkisch sichtbar.

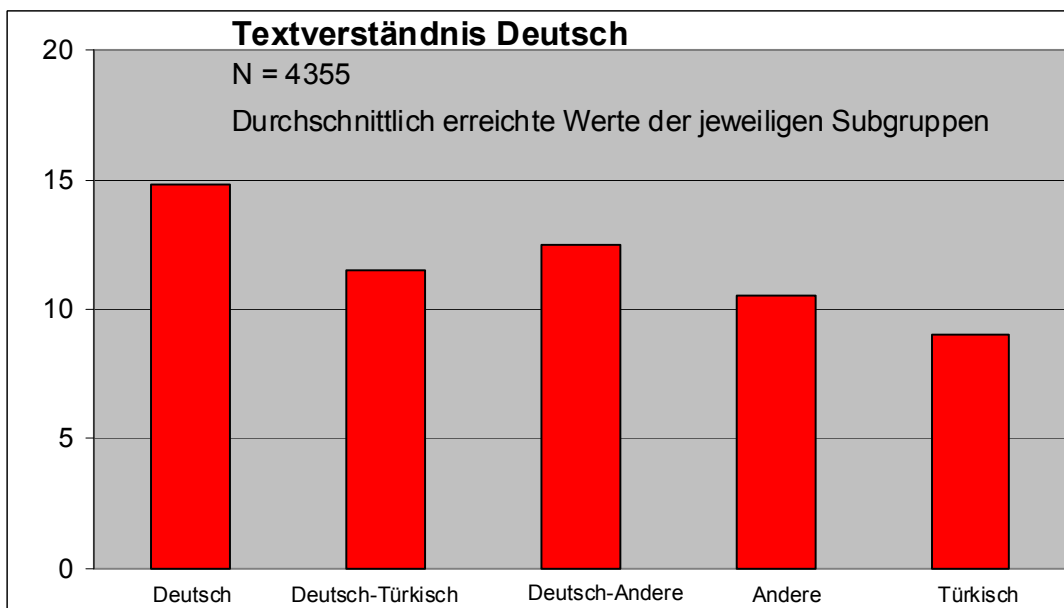
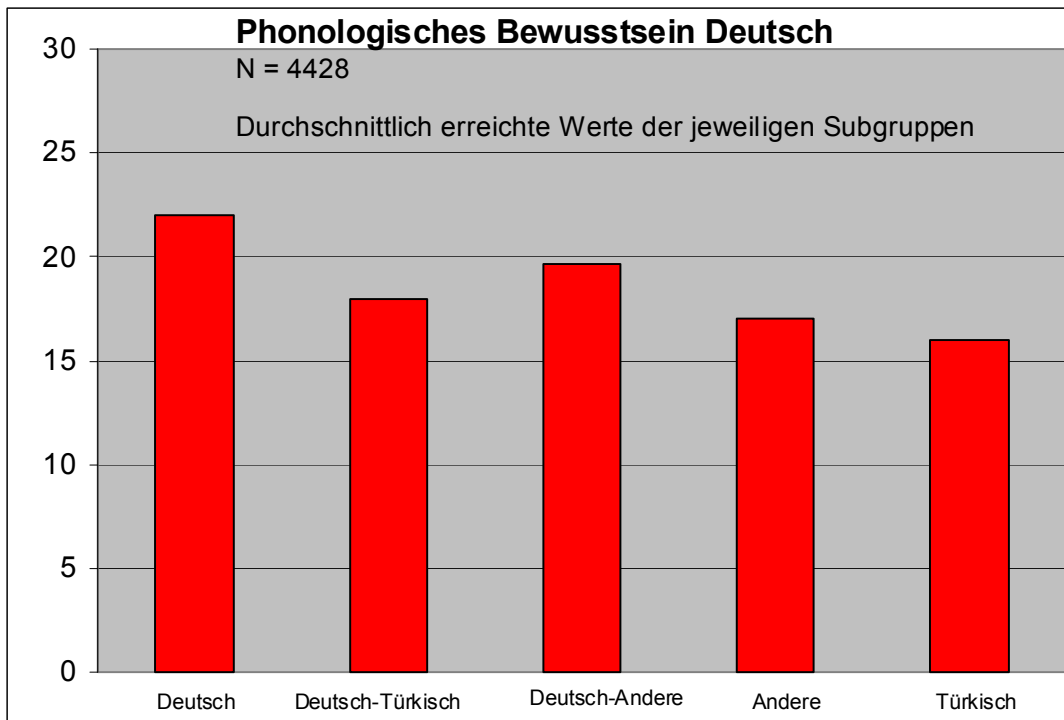
Bei der Gruppe Deutsch-Türkisch war das Erkennen der eigenen Fähigkeiten etwas weniger deutlich sichtbar. Sie schnitt zwar bezüglich Deutsch höher ab als die Gruppe Türkisch, aber insgesamt schnitt sie in Türkisch besser ab, als sie selbst angenommen hatte.

Anlagen

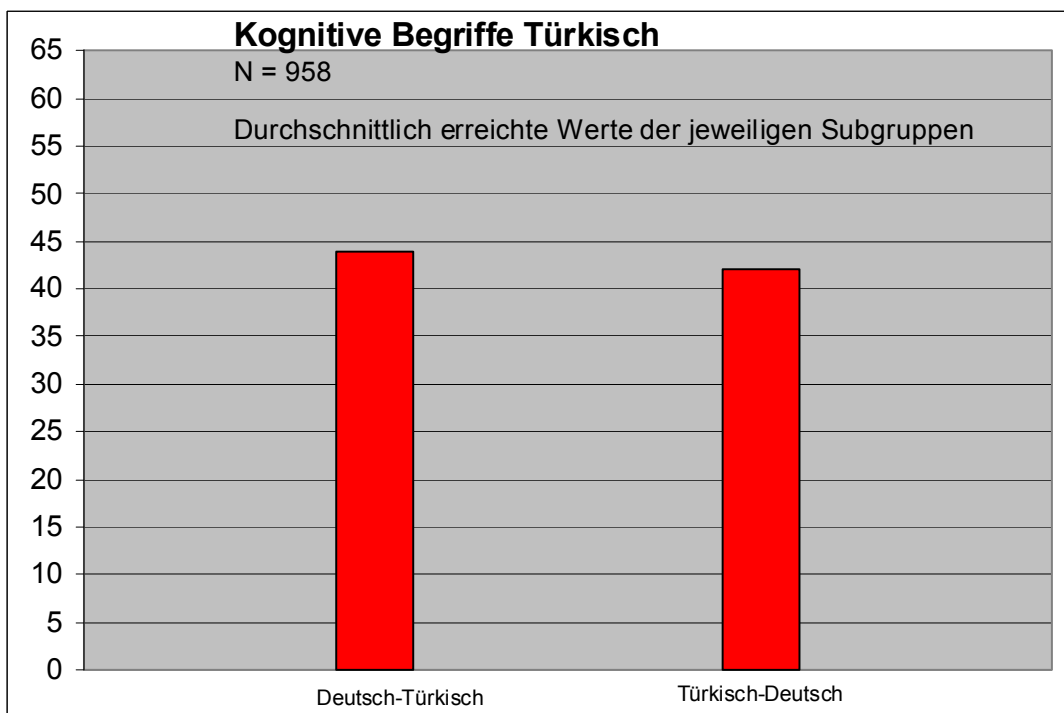
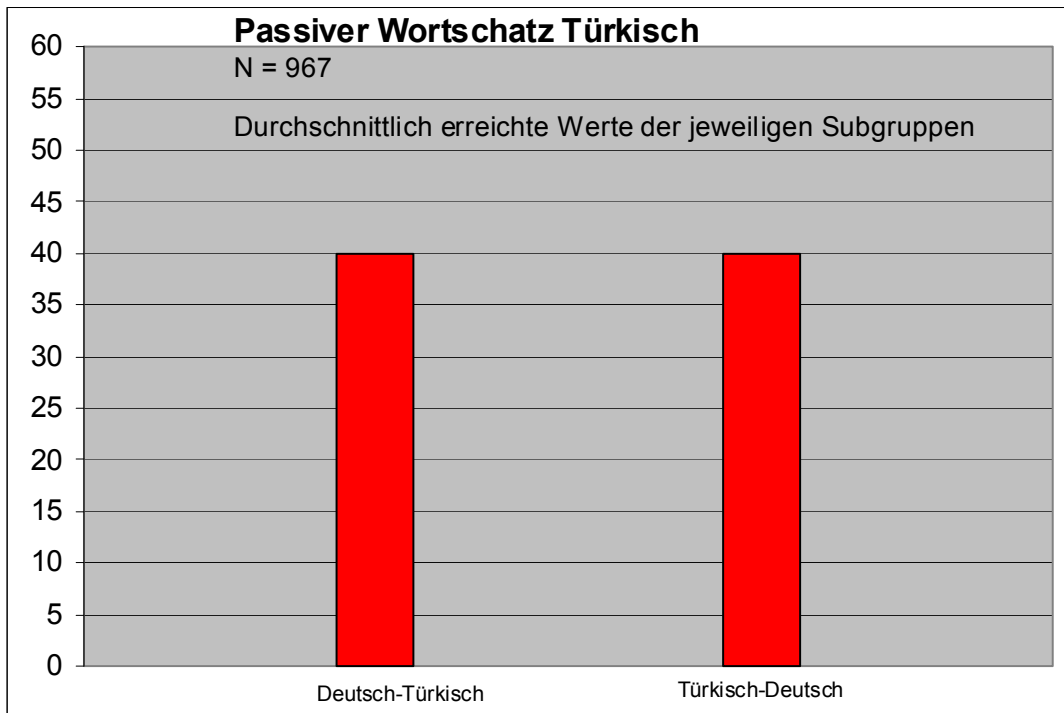
A Ergebnisse aller Teilgruppen in der deutschen Testversion



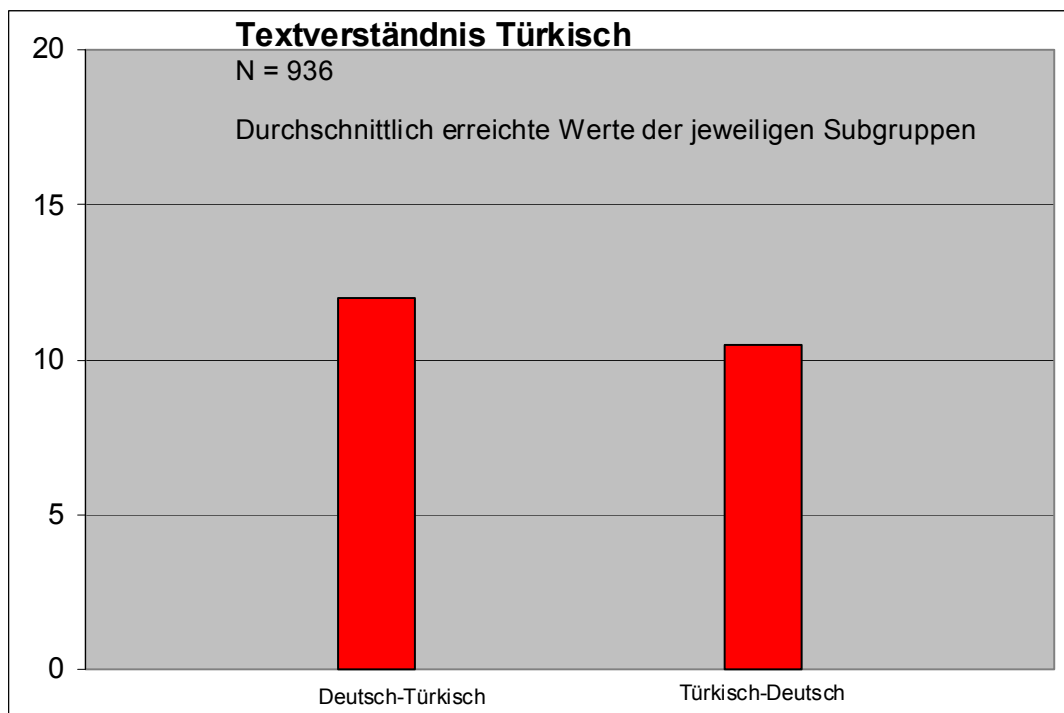
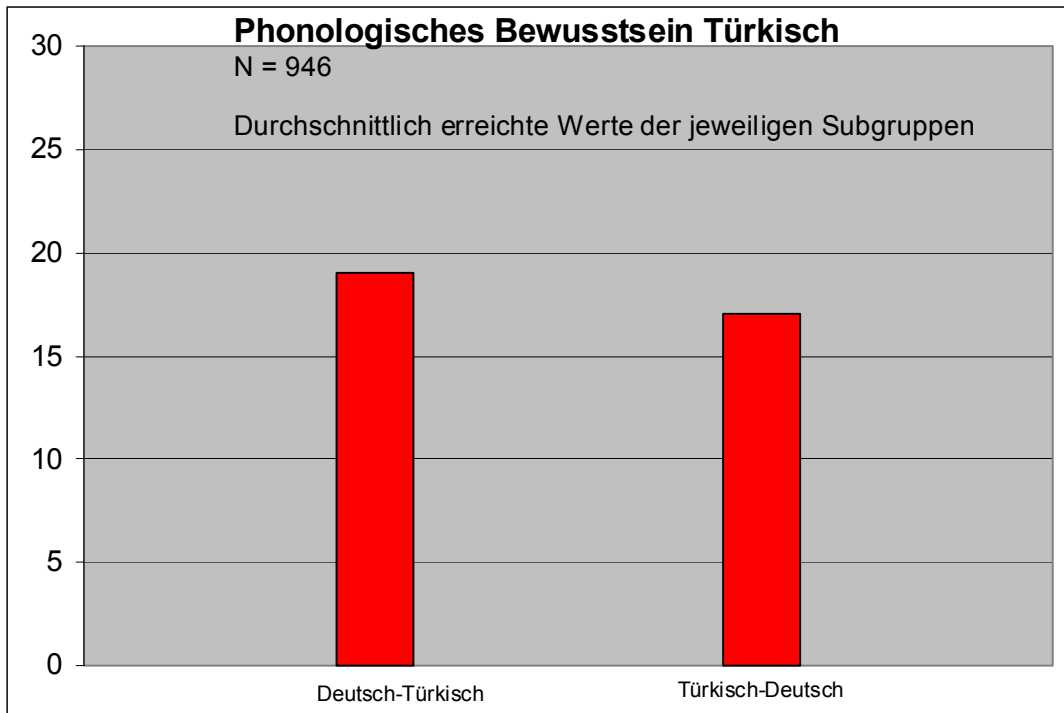
A Ergebnisse aller Teilgruppen in der deutschen Testversion



B Ergebnisse der beiden türkischen Teilgruppen in der türkischen Testversion

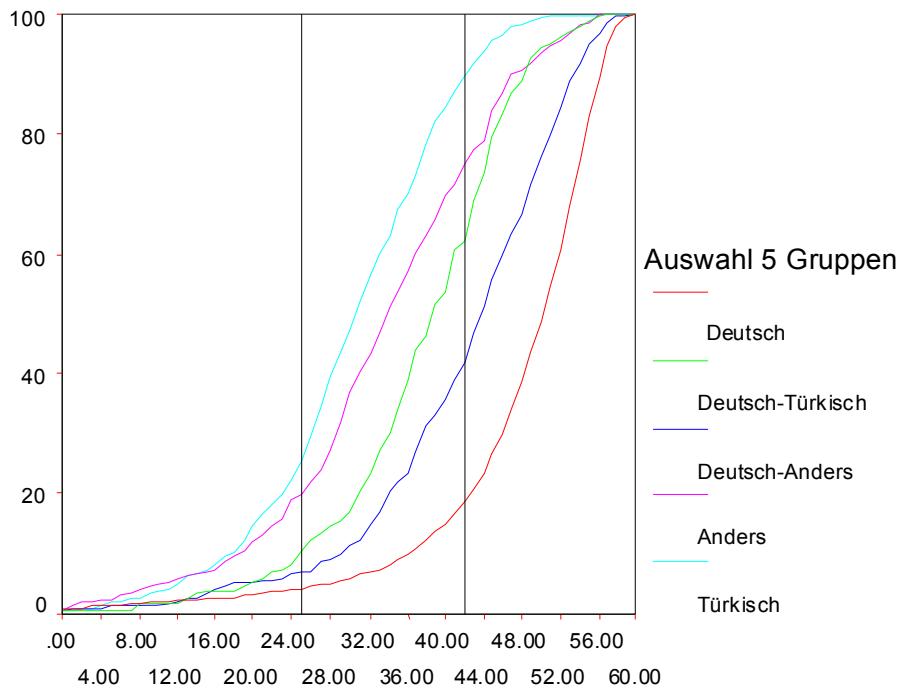


B Ergebnisse der beiden türkischen Teilgruppen in der türkischen Testversion



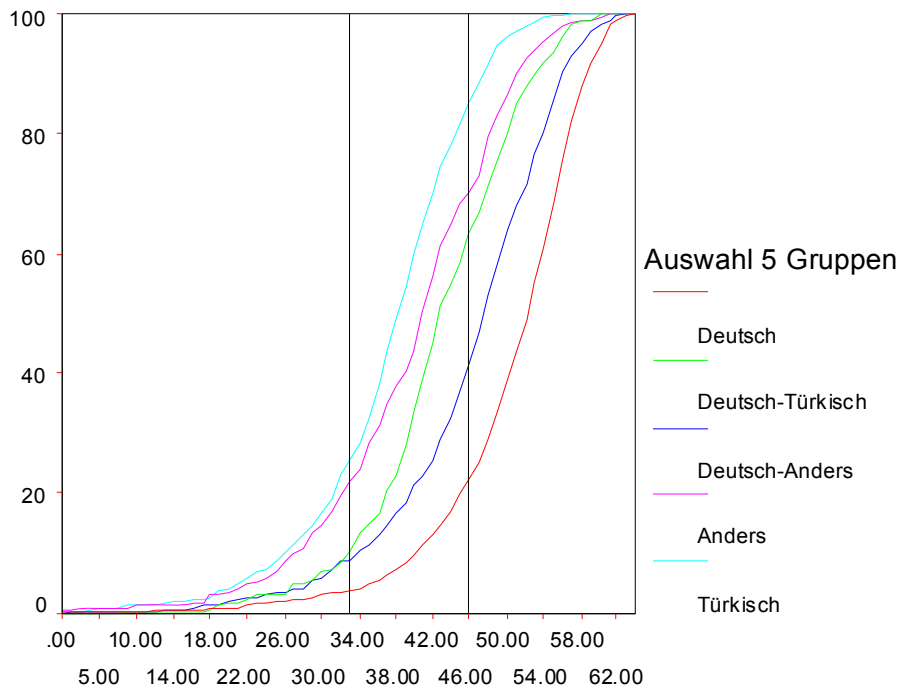
C Ergebnisse in der deutschen Testversion

Cumulative Percent



PW

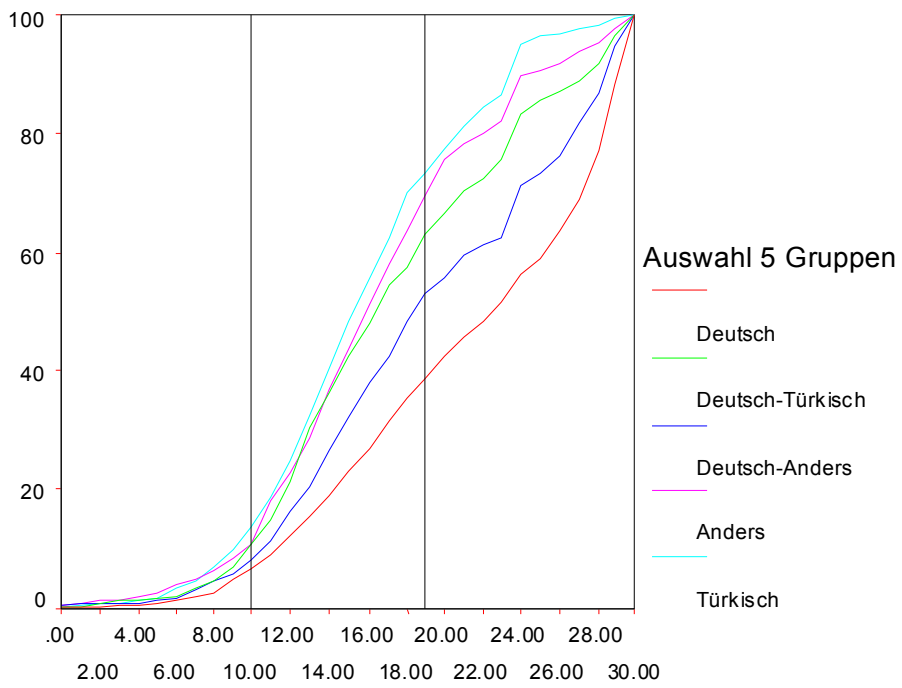
Cumulative Percent



KB

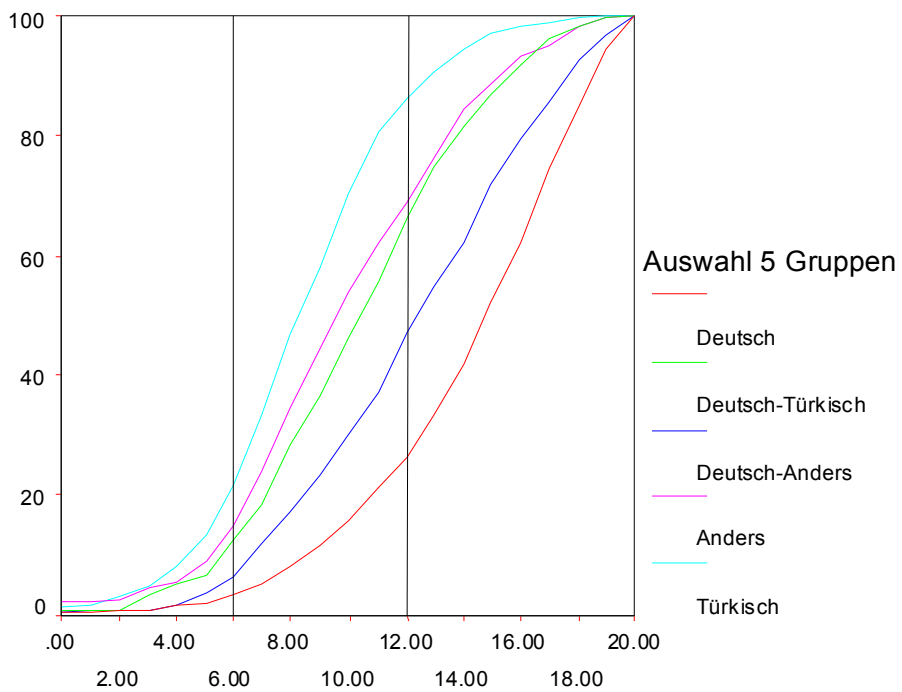
C Ergebnisse in der deutschen Testversion

Cumulative Percent



Phonologisches Bewußtsein

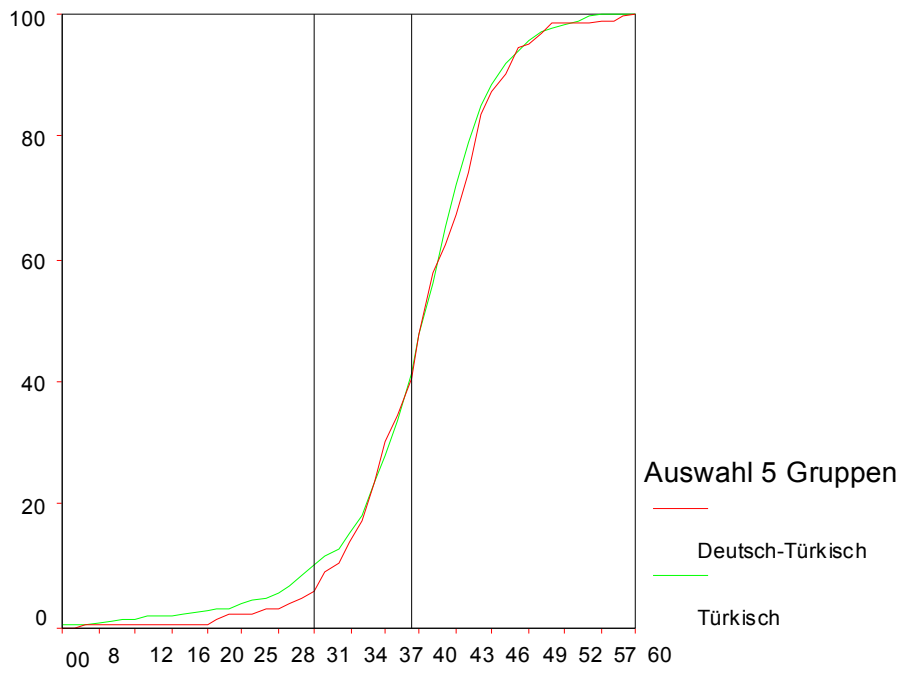
Cumulative Percent



Text-Verständnis

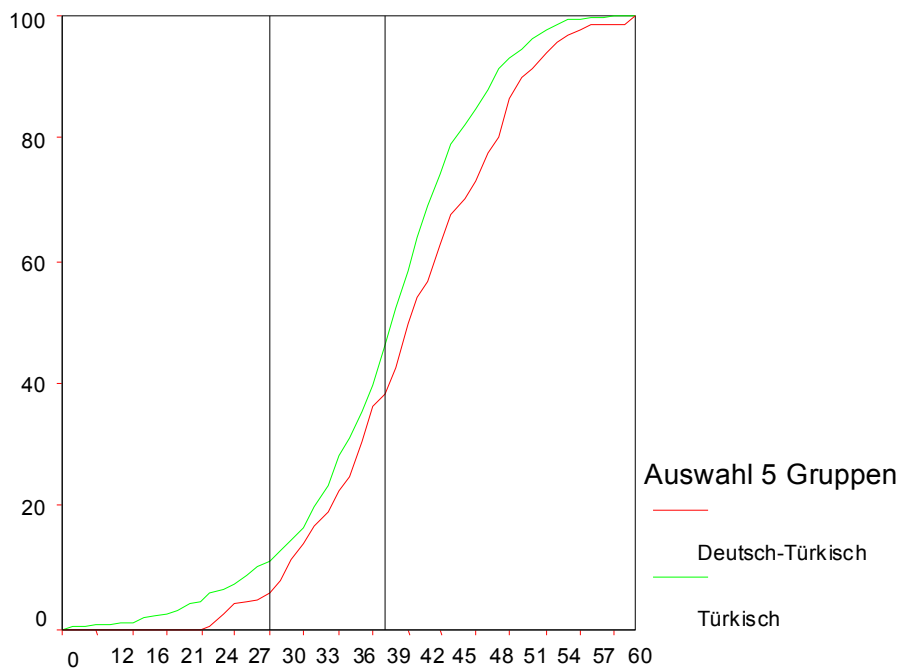
D Ergebnisse in der türkischen Testversion

Cumulative Percent



Passiver Wortschatz

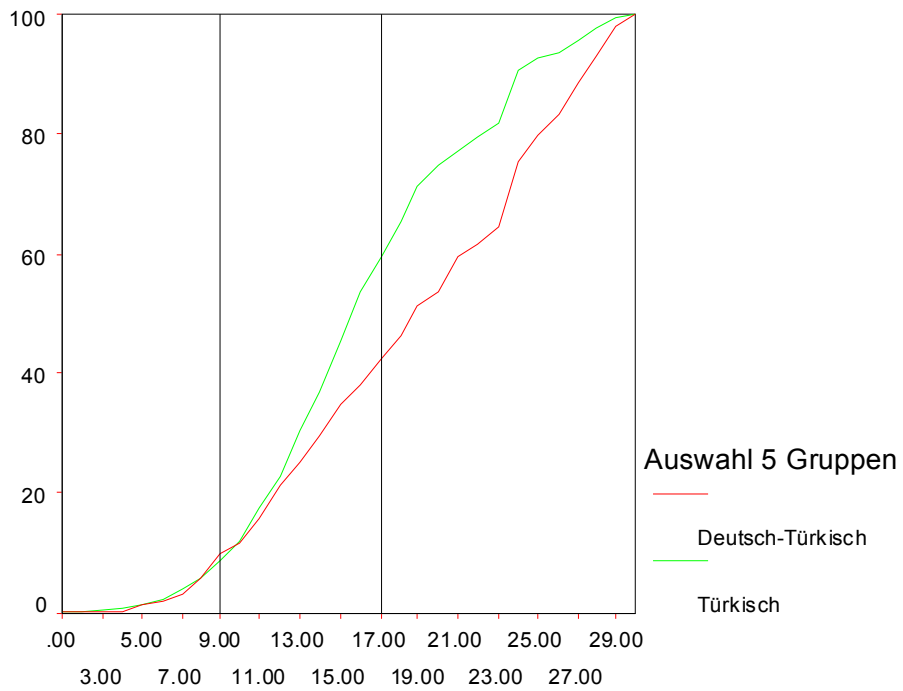
Cumulative Percent



Kognitive Begriffe

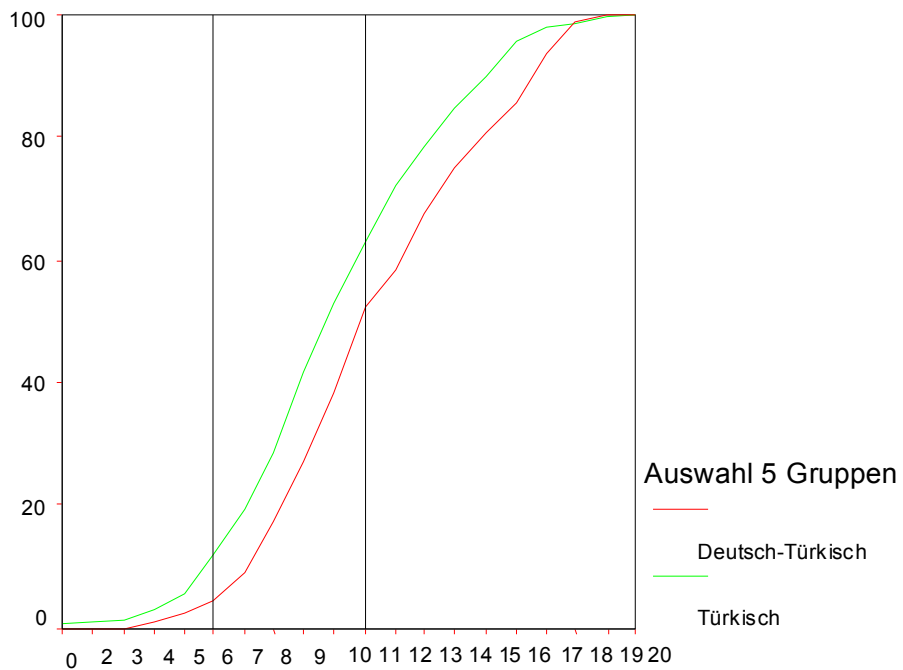
D Ergebnisse in der türkischen Testversion

Cumulative Percent



Phonologisches Bewußtsein

Cumulative Percent



Text-Verständnis

Cito setzt sich weltweit für gute und ehrliche Prüfungen und Beurteilungen ein. Mit den Mess- und Folgemethoden von Cito bekommt man ein objektives Bild von Kenntnis, Allgemeinwissen und Kompetenzen. Hierdurch sind verantwortliche Wahlen auf dem Gebiet der persönlichen und professionellen Entwicklung möglich. Unsere Expertise setzen wir nicht nur für unsere eigene Arbeit ein, sondern auch um Anderen Beratung, Unterstützung und Untersuchung zu bieten.

Cito Deutschland
Schlossstraße 10
35510 Butzbach
T (06033) 74 63 00
F (06033) 74 63 01
deutschland@cito.com
www.de.cito.com

Fotografie: Ron Steemers
© Cito Deutschland GmbH (2010)