

## **Teil VI**

### **Analyse des Beobachtungsverfahrens**

Zur Analyse eines Datenerhebungsverfahrens gehören die Analyse der in ihm verwendeten Aufgaben und der Güteeigenschaften des Instrumentariums, gleich ob es sich um ein psychometrisches Testverfahren oder wie im vorliegenden Fall um ein Beobachtungsverfahren handelt. Bei einer Aufgabenanalyse, die sich hier nicht wie bei einem psychometrischen Test auf einzelne Items, sondern auf die anhand der Kriterien bei den verschiedenen Untersuchungsanordnungen zum Teil mehrfach bewerteten sprachlichen Untersuchungsbereiche bezieht, werden allgemein die Schwierigkeit und die Trennschärfe der einzelnen Aufgaben bestimmt. Die Ergebnisse dieser Analyse werden wiederum zur Ermittlung der Reliabilität benötigt, geben darüber hinaus aber auch Hinweise zu einer eventuell notwendigen Veränderung des Bewertungsmaßstabs oder der Überarbeitung der Untersuchungsanordnung.

Weniger gebräuchlich ist die Ermittlung der Spezifität, die bei der Entwicklung von Sprachdiagnoseverfahren dazu dient, allgemein und spezifisch geeignete Bereiche zur Beschreibung der Sprachkompetenz aufzufinden, und die unter anderem über die Faktorenanalyse berechnet werden kann. Mit der Faktorenanalyse lassen sich aus den Untersuchungsbereichen überdies diejenigen identifizieren, die wesentlich zur Aufklärung der beobachteten Varianz des hier interessierenden Sprachverhaltens beitragen.

Zu den wichtigsten Güteeigenschaften eines Verfahrens zählen die Objektivität, die Reliabilität und die Validität.<sup>244</sup> Aber auch ökonomische Aspekte, darunter vor allem die zur Durchführung und Auswertung notwendige Zeit, sollten nicht völlig unbeachtet bleiben, wenn ein Verfahren nicht nur vereinzelt zur Anwendung kommen soll. Daher wird im Anschluss an die Qualitätsprüfung der bei der Untersuchung verwendeten Pilotfassung eine kritische Musterung des sehr umfangreichen Beobachtungsverfahrens mit dem Ziel vorgenommen, durch die Selektion von Untersuchungsanordnungen zu sprachlichen Teilbereichen, die anhand der Analysedaten als Indikatoren geeignet erscheinen, ein Verfahren vorzuschlagen, das bei reduziertem Aufwand deutsche und italienische Sprachfähigkeiten gegen Ende eines 2. Schuljahres noch hinreichend genau zu erfassen vermag.

Bei der Analyse des Verfahrens sind allerdings bewährte statistische Methoden nur sehr begrenzt einzusetzen, da viele der bei psychometrischen Tests anwendbaren Rechenoperationen durch die andere Konstruktion des Beobachtungsverfahrens nicht ausgeführt werden können und auf einer Normalverteilung basierende Prüfmethode ebenfalls ausscheiden. Bei einer Untersuchung zur Zweisprachigkeit im Primarbereich liegen schließlich keine annähernd normal verteilten Bedingungen vor, sondern die Probanden weisen vielmehr äußerst heterogene Voraussetzungen auf.

---

<sup>244</sup> In der englischsprachigen Literatur zu Testverfahren wird das Kriterium Objektivität meistens als Element des Reliabilitätskonzepts dargestellt.

## **25 Qualitätsprüfung des Untersuchungsverfahrens**

### **25.1 Angaben zur Objektivität**

Zur Gewährleistung der Objektivität der Ergebnisse eines Untersuchungsverfahrens sollten die Daten weitgehend unabhängig von der Untersuchungssituation und vom Beobachter erhoben werden. Außerdem sollte die Auswertung durch standardisierte Vorgaben möglichst nicht subjektiv beeinflussbar sein. Hinsichtlich der Objektivität ist daher zwischen Durchführungsobjektivität und Auswertungsobjektivität des Verfahrens zu unterscheiden.

#### **25.1.1 Durchführungsobjektivität**

Bei der Untersuchung zur Entwicklung der Sprachkompetenzen an der SESB wurde ein Beobachtungsverfahren eingesetzt, bei dem zwecks Annäherung an natürliche Kommunikationssituationen möglichst viele freie Sprachäußerungen evoziert werden sollten. Insofern war ein Einfluss interpersoneller Unterschiede bei der Durchführung der Untersuchungsanordnungen nicht zu vermeiden. Wenn freie Rede Gegenstand einer Untersuchung ist, würde die Reduzierung der sozialen Interaktion zwischen Beobachter und Proband dem Untersuchungsziel widersprechen.

Dennoch ist davon auszugehen, dass die Mitarbeiter bei dieser Untersuchung bei der Datenerhebung auf annähernd ähnliche Weise vorgegangen sind, da der Ablauf der Untersuchungsanordnungen durch das visuelle oder textuelle Material bis hin zu den Impulsen weitgehend vorstrukturiert war, vor Beginn der Untersuchung eine Schulung zur altersgemäßen Gesprächsführung stattfand und die ständige Präsenz eines zweiten Beobachters auch eine gewisse Kontrolle hinsichtlich des erwünschten und zulässigen Verhaltens seitens des Untersuchungsleiters ausübte. Aus den Protokollen geht hervor, dass es bei keinem Team zu gravierenden Abweichungen kam. Mitunter wurden zwar zunächst ausgelassene Fragen später nachgeholt oder völlig vergessen, was aber bei der Auswertung entsprechend der Protokollvermerke berücksichtigt werden konnte. Anhand der Tonaufnahmen ist festzustellen, dass manche Mitarbeiter sich förmlicher an die Instruktionen hielten als andere, die mit besonderem Geschick, Flexibilität und Einfühlungsvermögen eine entspannte Atmosphäre schufen, was des öfteren an fröhlichem Gelächter zu erkennen ist und sich sicher günstig auf die Gesprächssituation ausgewirkt hat. Durchwegs hielten sich aber alle Mitarbeiter an die Absprachen bezüglich einer geduldigen Vorgehensweise und eines freundlichen und stets ermunternden Umgangstons, so dass von Seiten der Interviewer her bei allen Schülern ähnliche Bedingungen vorlagen.

Dazu, dass die Sprachproben einem natürlichen Gespräch nahe kamen, hat sicherlich beigetragen, dass sich die jeweiligen Untersuchungsleiter voll auf ihre Aufgabe der Datenerhebung konzentrieren konnten und nicht nebenher mit der Aufnahmetechnik oder gar schon der Auswertung belastet wurden, wie das bei etlichen Sprachtests üblich ist. Abgesehen davon, dass sich ein natürlicher Dialog kaum ergeben kann, wenn der Interviewer laufend in sein Manual blickt oder auf dem Auswertungsbogen Kreuzchen einträgt, treten bei einer parallel zur Testdurchführung vorzunehmenden spontanen Feldbewertung bekanntlich gehäuft Durchführungsfehler auf. Zum Beispiel ist die geringe Durchführungsobjektivität des PI- Verfahrens vor allem auf die nicht zu vereinbarende Fülle von komplexen Anforderungen an den testdurchführenden Lehrer zurückzuführen.

*„Verlangt war die Kontrolle über das Testgeschehen und die richtige Vorlage der Aufgaben, die Wahrnehmung der Schüleräußerungen, ihre Beurteilung im Sinne der Kriterien des Tests sowie das Eintragen der Ergebnisse in den kompliziert gestalteten Auswertungsbogen.“<sup>245</sup>*

Da bei dieser Untersuchung die Bewertung der Schüleräußerungen grundsätzlich erst in einer späteren Phase erfolgte und der zweite Beobachter aus seinem größeren Abstand heraus neben dem Protokollieren der nicht vom Tonband erfassbaren Wahrnehmungen auch den Überblick über den Ablauf der Untersuchungsanordnung behalten und somit notfalls dezent an vergessene oder verkehrt dargebotene Teile erinnern konnte, sind die während der Durchführung der Aufgaben interpersonellen Unterschieden zuzuschreibenden Abweichungen vermutlich nur geringfügig.

Hingegen war bei dieser Untersuchung eine gewisse Beeinträchtigung der Durchführungsobjektivität durch schulorganisatorisch bedingte Abweichungen in Bezug auf die zur Verfügung stehenden Räume und den Zeitpunkt innerhalb des Schulvormittags leider nicht zu vermeiden, wie bereits ausführlich im Zusammenhang mit den Durchführungsbedingungen (Kapitel 12) dargestellt wurde.

### **25.1.2 Auswertungsobjektivität**

Dagegen weist die Untersuchung eine optimale Auswertungsobjektivität auf, obwohl die Auswertung des Pilotverfahrens teilweise recht kompliziert war. Die Beurteilung der sprachlichen Leistungen wurde durch die Kriterienangaben auf den Auswertungsbögen zwar erleichtert, aber bei der anschließenden Punktezuordnung wirkten die vielen verschiedenen Auswertungsschlüssel doch etwas verwirrend. Nicht nur dass ursprünglich bei drei Bereichen eine andere Bewertung galt, je nachdem ob die

---

<sup>245</sup> Boos- Nünning / Gogolin 1988, a.a.O., S.26

Aufgabe in der starken oder in der Partnersprache bewältigt wurde, auch der unterschiedliche Einbezug der Begriffsbildung bei der deutschen Sprachgruppe beim deutschen oder italienischen Wortschatz erforderte eine aufmerksame Beachtung. In der überarbeiteten Fassung wird die Auswertung durch das Entfallen der unterschiedlichen Bewertung und einiger Bereiche vereinfacht.

Doch zum Glück nahmen die meisten studentischen Mitarbeiter die Auswertung sehr gewissenhaft vor und brachten auch ausreichend Geduld für die äußerst zeitintensive Anfertigung der Transkriptionen auf, die mitunter durch die mangelhafte Tonqualität einiger Aufnahmen noch erschwert wurde. Überhaupt verdankt die Untersuchung ihre hohe Auswerterübereinstimmung sicherlich der günstigen Situation des begleitenden Seminars. Zwar wurde die eigentliche Auswertung stets von mehreren Personen unabhängig voneinander vorgenommen, aber bei den Sitzungen wurden vorher zu jedem Bereich anhand von Tonbandaufnahmen und Transkriptionen praktische Übungen durchgeführt, bei denen auch Unklarheiten beseitigt und in zweifelhaften Fällen gemeinsame Entscheidungskriterien vereinbart werden konnten. Es ist fraglich, ob eine ähnlich hohe Auswertungsobjektivität von nahezu 100% nur aufgrund der schriftlichen Anleitungen erreicht worden wäre.

**Auswerterübereinstimmung – (Ebene Gesamtpunkte<sup>246</sup>)**

Team	bei deutscher Sprachkompetenz	bei italienischer Sprachkompetenz
Team 1	0,99	0,98
Team 2	0,99	0,99
Team 3	0,99	0,95
Team 4	1,00	0,99
Team 5	0,99	0,98
Team 6	0,99	1,00

Wenn alle Teilergebnisse insgesamt in Beziehung gesetzt werden, ergeben sich bei allen an der SESB eingesetzten Teams hinsichtlich beider Sprachkompetenzen Produkt-Moment-Korrelationen von mindestens  $r = 0.95$ . Allerdings trägt zur starken Übereinstimmung bei einem Team sicherlich bei, dass nicht die von einem Studenten

---

<sup>246</sup> D.h. unter Berücksichtigung aller in den verschiedenen Bereichen vorgenommenen Bewertungen insgesamt (ausschließlich des Gesamtergebnisses).

angefertigte Auswertung, sondern stattdessen die von mir korrigierte Auswertung berücksichtigt wurde. Die Korrektur war in diesem Fall erforderlich, weil bei den stichprobenartigen Kontrollen im Gegensatz zu allen übrigen Beurteilungen eine fahrig und daher fehlerhafte Auswertung aufgefallen war. Dieser Mitarbeiter hatte die Auswertung offensichtlich sehr unkonzentriert vorgenommen. So waren auf den Auswertungsbögen mitunter gleiche Codebezeichnungen zum selben Bereich eingetragen worden, wodurch keine eindeutige Zuordnung erfolgen konnte. Auch waren die Sprachäußerungen durch einen Vergleich der Transkriptionen und der angekreuzten Kriterien zwar manchmal nachvollziehbar beobachtet, aber dann unter Anwendung eines nicht zutreffenden Auswertungsschlüssels falsch bewertet worden. Insofern musste die Auswertung hier gegebenenfalls berichtigt oder zum Teil erneut durchgeführt werden.

### Auswerterübereinstimmung – (Ebene Untersuchungsbereiche)

starke Sprache Italienisch		Partnersprache Deutsch	
<u>Bild A</u>		<u>Bild B</u>	
Hörverständnis	0,88	Hörverständnis	1,00
Phonematik/Prosodie	0,68	Phonematik/Prosodie	0,97
Wortschatz	0,86	Wortschatz	0,14
Morphosyntax	0,83	Morphosyntax	0,98
mündl. Sprachgebrauch	0,92	mündl. Sprachgebrauch	0,89
<u>I- 1 Lesen</u>		<u>PS-D</u>	
Phonematik/Prosodie	1,00	Phonematik/Prosodie	0,90
Lesefertigkeit	0,89	Lesefertigkeit	0,98
<u>I - 2 Textv.</u>		<u>Spiel 1</u>	
Textverständnis	0,93	Hörverständnis	1,00
Wortschatz	0,98	mündl. Sprachgebrauch	1,00
Morphosyntax	0,98		
mündl. Sprachgebrauch	0,99		
<u>Schreibfertigkeit</u>	0,84		
I - 3 Abschreiben	0,63		
I - 4 Diktat	1,00		
concetti – Rohwerte	0,91		
kommun. Sprachverhalten	0,89	kommun. Sprachverhalten	0,87
Gesamtergebnis	0,99	Gesamtergebnis	0,98
bil. kom. Sprachverhalten			1,00

Da eine Gesamtübereinstimmung auch in Bezug auf Differenzen täuschen kann, weil sich eventuelle Unterschiede bei einer Vielzahl von zu bewertenden Bereichen auch gegenseitig aufheben können<sup>247</sup>, wurde die Auswerterübereinstimmung bei einem Team zusätzlich anhand der Bewertungen bei den Untersuchungsbereichen einzeln überprüft. In der Tat scheint beim Team 6 zwischen den Bewertungen der beiden Auswerter trotz grundsätzlicher Bestätigung der Tendenz einer ähnlichen Beurteilung nicht in allen Bereichen eine hohe Übereinstimmung vorzuliegen. Besonders auffallend ist die sehr geringe Korrelation von 0,14 bei der Bewertung des Wortschatzes beim Bildimpuls-gesteuerten Interview in der Partnersprache.

Betrachtet man die Bewertungen im Einzelnen, ist jedoch festzustellen, dass die Auswerter bei 4 von den 6 zu beurteilenden lexikalischen Fähigkeiten durchaus zum gleichen Urteil kamen und beide die max. Punktzahl vergaben. Nur bei 2 Schülern treten Differenzen auf. Da die Bewertungen aber nicht in beiden Fällen bei einem Auswerter strenger ausfielen als beim andern, sondern in umgekehrtem Verhältnis voneinander abweichen (C6: 3/ 3,5 - C16: 4/ 2,5) und Korrelationsberechnungen bei wenigen Daten sehr empfindlich auf geringfügige Veränderungen von Verhältnissen reagieren, widerspricht auch dieses vereinzelt Phänomen, das durch die geringe Stichprobengröße bei Betrachtung der Bewertung der einzelnen Untersuchungsbereiche von nur 6 Probanden zu erklären ist, nicht der allgemein sehr guten Auswertungsobjektivität der Untersuchung.

## **25.2           Angaben zur Reliabilität**

### **25.2.1        Aufgabenanalyse**

Bei dem zur Untersuchung an der SESB verwendeten Verfahren handelt es sich nicht um einen Test im herkömmlichen Sinne, schon gar nicht um einen homogenen Test, für dessen Analyse statistische Handbücher vor allem nützliche Hinweise bieten. Bei der Konstruktion des Verfahrens wurde bewusst auf eine Aufgliederung komplexer Sprachleistungen in minutiöse Items sowie auf eine sonst wegen der leichteren Auswertung beliebte Dichotomisierung verzichtet. Anliegen des hier vorgestellten Verfahrens ist keine Bewertung von Kenntnissen, wie es beim Fremdsprachenunterricht mitunter durchaus angebracht sein kann, sondern die Beobachtung eines langwierigen Spracherwerbsprozesses, bei dem es eher um „schon mehr“ oder „noch weniger“ richtigen Sprachgebrauch geht als um die eindeutige Zuordnung zu Kategorien wie Falsch und Richtig. Selbst bei dem am ehesten einem Testverfahren entsprechenden Fragenkatalog zur Begriffsbildung wurden deshalb produktive

---

<sup>247</sup> Siehe in Bezug auf das PI-Verfahren: Boos-Nünning / Gogolin 1988, a.a.O., S.24.

Sprachleistungen gefordert und keine Auswahlantworten angeboten, was hinsichtlich der statistischen Analyse einen gewissen Vorteil hat, weil ein Raten entfiel und sich daher immerhin mühselig zu berechnende Zufallskorrekturen erübrigen.

Da das Beobilingua-2dit-Verfahren gewisse Bedingungen einer Standardisierung erfüllt und bei der Konstruktion entsprechend der komplexen Sprachkompetenz möglichst viele verschiedene Bereiche der Sprachfähigkeit berücksichtigt wurden, kann das Verfahren analog einer heterogenen Testbatterie einigen der dafür geeigneten statistischen Verfahren unterworfen werden. Den Untertests einer Testbatterie entsprechen bei diesem Beobachtungsverfahren jedoch die Untersuchungsbereiche und nicht etwa die Untersuchungsanordnungen. Diese bilden lediglich kindgemäße Darbietungen der jeweils dabei interessierenden Untersuchungsbereiche, deren jeweilige Beobachtungskriterien jedoch nicht wie bei einem psychometrischen Test als einzelne Items, sondern eher wie ein Bündel von Items, die dasselbe Merkmal erfassen, zu betrachten sind. Die Zuordnung der Rohwertpunkte erfolgte nicht pro erfülltem Teilkriterium, sondern unter Berücksichtigung des Ausprägungsgrades einer komplexen Sprachleistung anhand einer für diesen Sprachbereich definierten Skala. Sollen statistische Verfahren angewendet werden, ist bei diesem Beobachtungsverfahren die maximal zu erreichende Punktzahl der mehrstufigen Bewertung in einem Untersuchungsbereich mit der Anzahl von Einzelaufgaben gleichzusetzen. Dann können die Trennschärfe und der Schwierigkeitsgrad der Untersuchungsbereiche wie bei einer Aufgabenanalyse mit den dafür geeigneten Formeln bestimmt werden.

#### **25.2.1.1 Schwierigkeit**

Die Berechnung von Schwierigkeitsindizes ist eine der verbreitetsten und einfachsten Methoden der Aufgabenanalyse. Bei psychometrischen Tests ist dabei die Prozentzahl derjenigen Probanden einer Stichprobe zu berechnen, die dieses Item richtig gelöst haben. Dabei ist ein hoher Schwierigkeitsindex  $P$  ein Zeichen für eine leichte Aufgabe, ein niedriger Index hingegen weist auf eine schwierige Aufgabe hin. Allgemein sollen Datenerhebungsverfahren möglichst vorwiegend aus Aufgaben mittlerer Schwierigkeit, also  $P = 50\%$  bestehen. Liegen dichotome Items vor, braucht lediglich die Anzahl der richtigen Lösungen abgezählt und diese durch die Anzahl der Probanden dividiert zu werden.

Bei einer mehrstufigen Bewertung wie bei der vorliegenden Untersuchung kann jedoch nicht etwa nur das Erreichen der maximalen Punktzahl in einem Bereich als Lösung der Aufgabe angesehen werden. Stattdessen sind bei der mehr oder weniger erfolgreichen Performanz einer Sprachleistung unbedingt unterschiedliche Niveaus zu berücksichtigen.



sichtigen. Auch schwächere Sprachfähigkeiten sollten zwar nicht als optimale, aber als hinreichende Lösung gelten. Daher wurde zur Berechnung der Schwierigkeit in dieser Untersuchung bei jedem Untersuchungsbereich sowohl ermittelt, wie viel Schüler die maximale Punktzahl erreichten, als auch wie viel Schüler wenigstens mit 50% des Maximalwerts beurteilt wurden. Als Schwierigkeitsindex für diesen Bereich wurde dann der Mittelwert der beiden Ergebnisse definiert.

Angesichts der heterogenen Sprachfähigkeiten der Probanden sehr unterschiedlicher Spracherwerbsstufen ist eine spezifisch auf jede Zielgruppe bezogene Analyse angebracht, denn *“a given item will have differing information functions at different levels of ability.”*<sup>248</sup> Um die Schwierigkeit für Lerngruppen unterschiedlicher Voraussetzungen zu ermitteln, wurden bei gleicher Aufgabenstellung und Bewertung in der starken Sprache und der Partnersprache die Schwierigkeitsindizes für die einzelnen Gruppen und insgesamt berechnet.

Beispiel: Berechnung des Schwierigkeitsindex im Bereich Hörverständnis – Bild A (it.)

Hörverständnis (alle Gruppen)	N	ab 2,0		max. 4		p
SESB (it. Sprachgr.)	18	18	43/ 45 = 0,96	15	24/ 45	0,74
it. Kontrollgruppe	9	9		7	= 0,53	
SESB (it. Partnersp.)	18	16		2		P = 74%
Summe	45	43		24		

Der Untersuchungsbereich Hörverständnis beim italienischen Interview erweist sich insgesamt als ziemlich einfach. Die Schüler der italienischen Sprachgruppe konnten sogar fast alle das Maximum erreichen (P = 92%), bei der Gruppe der Partnersprache liegt eine mittlere Schwierigkeit von P = 50% vor.

Bei Betrachtung aller Untersuchungsbereiche aller in die Untersuchung einbezogener Schüler der unterschiedlichsten Spracherwerbsstadien ergibt sich insgesamt eine recht befriedigende mittlere Schwierigkeit von P = 62%. Die Untersuchungsbereiche Kommunikatives Sprachverhalten, Wortschatz und Begriffsbildung (Konzepte) sind mit Schwierigkeitsindizes um P = 50% besonders zufriedenstellend. Durch eine strengere Bewertung bei den offensichtlich sehr leichten Bereichen Hör- und Textverständnis

<sup>248</sup> Bachmann 1990, a.a.O., S.207

sowie der Schreibfertigkeit in der starken Sprache und der Lesefertigkeit in der starken und in der Partnersprache könnte die mittlere Schwierigkeit jedoch noch stärker dem Idealwert von 50% angenähert werden. Bei den Bereichen Phonematik/Prosodie und Morphosyntax, die insgesamt ebenfalls als zu leicht wirken, ist hingegen zu beachten, dass unter Anwendung derselben Auswertungskriterien Kinder mit sehr verschiedenen sprachlichen Voraussetzungen zu beurteilen sind. Bei einem Verfahren zur Erfassung bilingualer Sprachfähigkeiten muss auch beachtet werden, wie leicht oder schwer bestimmte Aufgaben jeweils für die Gruppen mit allgemein starker bzw. schwächerer Kompetenz ausfallen.

### Schwierigkeitsindizes aller Untersuchungsbereiche – SESB u. Kontrollgruppen

Italiano		Deutsch		Bild A/B (it.+ d)	
<u>Bild A (it.)</u>		<u>Bild B (d)</u>		<u>Bild A/B (it.+ d)</u>	
Hörverständnis	0,74	Hörverständnis	0,84	Hörverständnis	0,79
Phonematik/Prosodie	0,53	Phonematik/Prosodie	0,82	Phonematik/Prosodie	0,67
Wortschatz	0,47	Wortschatz	0,69	Wortschatz	0,58
Morphosyntax	0,59	Morphosyntax	0,80	Morphosyntax	0,70
mündl. Sprachgebrauch	0,33	mündl. Sprachgebrauch	0,65	mündl. Sprachgebrauch	0,49
<u>I - 2 Textv.</u>		<u>D - 2 Textv.</u>		<u>I/D - 2 Textv.</u>	
Textverständnis	0,74	Textverständnis	0,89	Textverständnis	0,82
Wortschatz	0,43	Wortschatz	0,50	Wortschatz	0,46
Morphosyntax	0,74	Morphosyntax	0,84	Morphosyntax	0,79
mündl. Sprachgebrauch	0,44	mündl. Sprachgebrauch	0,52	mündl. Sprachgebrauch	0,48
<u>Spiel 1 (it.)</u>		<u>Spiel 1 (d)</u>		<u>Spiel 1</u>	
Hörverständnis	0,56	Hörverständnis	0,89	Hörverständnis	0,72
mündl. Sprachgebrauch	0,45	mündl. Sprachgebrauch	0,78	mündl. Sprachgebrauch	0,61
<u>Schreibfertigkeit (it.)</u>		<u>Schreibfertigkeit (d)</u>		<u>Schreibfertigkeit (it. + d)</u>	
I - 3 Abschreiben	0,70	D - 3 Abschreiben	0,64	D - 3 Abschreiben	0,67
I - 4 Diktat	0,68	D - 4 Diktat	0,68	D - 4 Diktat	0,68
Schreibfertigkeit G	0,63	Schreibfertigkeit G	0,64	Schreibfertigkeit G	0,64
<u>concetti</u>		<u>Konzepte</u>		<u>concetti / Konzepte</u>	
Punkte	0,43	Punkte	0,53	Punkte	0,48
kom. Sprachverhalten I	0,51	kom. Sprachverhalten D	0,54	kom. Sprachverh. I + D	0,53
<u>I - 1 Lesen</u>		<u>D - 1 Lesen</u>		<u>I/D - 1 Lesen</u>	
Phonematik/Prosodie	0,64	Phonematik/Prosodie	0,66	Phonematik/Prosodie	0,65
Lesefertigkeit	0,80	Lesefertigkeit	0,70	Lesefertigkeit	0,75
<u>PS-I - Lesen</u>		<u>PS-D - Lesen</u>		<u>PS-I/D - Lesen</u>	
Phonematik/Prosodie	0,31	Phonematik/Prosodie	0,39	Phonematik/Prosodie	0,35
Lesefertigkeit	0,56	Lesefertigkeit	0,61	Lesefertigkeit	0,59
<div style="background-color: #cccccc; padding: 2px; display: inline-block;">unterlegte Bereiche ergeben sich aus der Addition der Punktwerte der zugehörigen Komponenten</div>					
mittlere Schwierigkeit	0,56	mittlere Schwierigkeit	0,68	mittlere Schwierigkeit	0,62

## Schwierigkeitsindizes der Untersuchungsbereiche - SESB/ starke Sprache

### starke Sprache Italienisch

Schreibfertigkeit (it.)	
I - 3 Abschreiben	0,72
I - 4 Diktat	0,61
Schreibfertigkeit G	0,61

concezzi	
Punkte	0,44

kom. Sprachverhalten I	0,61
------------------------	------

I - 1 Lesen	
Phonematik/ Prosodie	0,69
Leseferdigkeit	0,83

Bild A (it.)	
Hörverständnis	0,92
Phonematik/ Prosodie	0,81
Wortschatz	0,47
Morphosyntax	0,59
mündl. Sprachgebrauch	0,45

I - 2 Textv.	
Textverständnis	0,75
Wortschatz	0,47
Morphosyntax	0,78
mündl. Sprachgebrauch	0,47

mittlere Schwierigkeit	0,66
------------------------	------

### starke Sprache Deutsch

Schreibfertigkeit (d)	
D - 3 Abschreiben	0,73
D - 4 Diktat	0,75
Schreibfertigkeit G	0,69

Konzepte	
Punkte	0,50

kom. Sprachverhalten D	0,56
------------------------	------

D - 1 Lesen	
Phonematik/ Prosodie	0,75
Leseferdigkeit	0,80

Bild B (d)	
Hörverständnis	0,89
Phonematik/ Prosodie	0,83
Wortschatz	0,69
Morphosyntax	0,80
mündl. Sprachgebrauch	0,50

D - 2 Textv.	
Textverständnis	0,92
Wortschatz	0,61
Morphosyntax	0,80
mündl. Sprachgebrauch	0,61

mittlere Schwierigkeit	0,75
------------------------	------

### starke Sprachen It. + D

Schreibfertigkeit (it. + d)	
D - 3 Abschreiben	0,73
D - 4 Diktat	0,68
Schreibfertigkeit G	0,65

concezzi/ Konzepte	
Punkte	0,47

kom. Sprachverh. I + D	0,58
------------------------	------

I / D - 1 Lesen	
Phonematik/ Prosodie	0,72
Leseferdigkeit	0,82

Bild A/B (it.+ d)	
Hörverständnis	0,91
Phonematik/ Prosodie	0,82
Wortschatz	0,58
Morphosyntax	0,70
mündl. Sprachgebrauch	0,49

I / D - 2 Textv.	
Textverständnis	0,83
Wortschatz	0,54
Morphosyntax	0,79
mündl. Sprachgebrauch	0,54

mittlere Schwierigkeit	0,70
------------------------	------

Zur Diagnose der Sprachfähigkeiten in der starken Sprache erscheint das Verfahren insgesamt als etwas zu leicht. Sollten bei der Überarbeitung Aufgaben zur Lese- und Schreibfertigkeit und dem Textverständnis in der starken Sprache beibehalten werden, wären die Anforderungen entweder durch schwierigere Aufgaben oder eine strengere Bewertung zu erhöhen.

Da sich der Bereich Morphosyntax nicht nur in der starken Sprache, sondern auch bei der Partnersprache als relativ leicht erweist, sollte der Bewertungsschlüssel bei einer Überarbeitung des Verfahrens wie auch bei den Bereichen Hörverständnis und Leseferdigkeit leicht angehoben werden. Indessen wäre bei den Bereichen Wortschatz und Phonematik/Prosodie eine Erhöhung der Schwierigkeit nicht angebracht, obwohl diese Bereiche auch in der deutschen Partnersprache als sehr leicht auffallen. Aber der Erwerb des Deutschen als Umgebungssprache wird bei einem in Deutschland stattfindenden bilingualen Spracherwerb stets schneller voranschreiten als der Erwerb

einer im außerschulischen Umfeld ziemlich funktionslosen Zweitsprache. Die unterschiedlichen Chancen dürfen nicht durch verschiedene Anforderungen nivelliert werden, damit die Sprachfähigkeiten vergleichbar bleiben.

### Schwierigkeitsindizes bei Untersuchungsbereichen der Partnersprache

Partnersprache Italienisch		Partnersprache Deutsch		Partnersprachen I + D	
<u>Spiel 1 (it.)</u>		<u>Spiel 1 (d)</u>		<u>Spiel 1</u>	
Hörverständnis	0,56	Hörverständnis	0,89	Hörverständnis	
mündl. Sprachgebrauch	0,45	mündl. Sprachgebrauch	0,78	mündl. Sprachgebrauch	
<u>kom. Sprachverhalten I</u>	<u>0,22</u>	<u>kom. Sprachverh. D</u>	<u>0,52</u>	<u>kom. Sprachverh. I + D</u>	
<u>PS-I - Lesen</u>		<u>PS-D - Lesen</u>		<u>PS-I/D - Lesen</u>	
Phonematik/Prosodie	0,31	Phonematik/Prosodie	0,39	Phonematik/Prosodie	
Lesefertigkeit	0,56	Lesefertigkeit	0,61	Lesefertigkeit	
<u>Bild A (it.)</u>		<u>Bild B (d)</u>		<u>Bild A/B (it.+ d)</u>	
Hörverständnis	0,50	Hörverständnis	0,89	Hörverständnis	
Phonematik/Prosodie	0,14	Phonematik/Prosodie	0,83	Phonematik/Prosodie	
Wortschatz	0,08	Wortschatz	0,83	Wortschatz	
Morphosyntax	0,56	Morphosyntax	0,80	Morphosyntax	
mündl. Sprachgebrauch	0,14	mündl. Sprachgebrauch	0,75	mündl. Sprachgebrauch	
mittlere Schwierigkeit	<u>0,35</u>	mittlere Schwierigkeit	<u>0,73</u>	mittlere Schwierigkeit	

Bei den Aufgaben, die gleichfalls von Schülern in der Partnersprache bearbeitet werden, ist die geringe Schwierigkeit, die sie Schülern in der starken Sprache bereiten, nicht vermeidbar. Ein für die bilinguale Sprachdiagnose geeignetes Verfahren muss schließlich die Bandbreite aller beim Spracherwerb erreichten Stadien erfassen und darf andererseits Sprachanfänger nicht völlig überfordern. Bezüglich der Schwierigkeit ist daher ein Kompromiss nicht zu umgehen, der gleichzeitig eine ausreichende Trennschärfe gewährleistet. Als ideal wäre bei diesem Verfahren sicherlich anzusehen, wenn die Schwierigkeit für Kinder in der Partnersprache/Zweitsprache zwischen 40% bis 50% und in der starken Sprache etwa bei 60%, höchstens jedoch bei 70% liegen würde.