



Institut für Schulqualität der Länder
Berlin und Brandenburg e.V.



Vorschulische Sprachstandsfeststellung ab dem Alter von 4 Jahren: Die Pilotierung von DEUTSCH PLUS 4.

Erste Ergebnisse.

Wolfgang Wendt, Thorsten Roick, Petra Stanat & Karin Tschackert
FU/Silberlaube, 25. April 2008

1. Funktion und Aufbau von DEUTSCH PLUS 4
2. Gütekriterien, Anlage der Untersuchung und Datenbasis
3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität
4. DEUTSCH PLUS 4: Ergebnisse zur Validität
5. DEUTSCH PLUS 4: Rückmeldungen der Erzieher/innen und Sonderpädagoginnen/innen
6. Zusammenhänge mit dem Bilanzbogen zum Sprachlerntagebuch
7. Erste Schlussfolgerungen

1. Funktion und Aufbau von DEUTSCH PLUS 4
2. Gütekriterien, Anlage der Untersuchung und Datenbasis
3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität
4. DEUTSCH PLUS 4: Ergebnisse zur Validität
5. DEUTSCH PLUS 4: Rückmeldungen der Erzieher/innen und Sonderpädagoginnen/innen
6. Zusammenhänge mit dem Bilanzbogen zum Sprachlerntagebuch
7. Erste Schlussfolgerungen

1. Funktion und Aufbau von DEUTSCH PLUS 4

Anlass und Ziel der Pilotierung

Ausgangslage

Bildungspolitisches Ziel, die vorschulische Sprachförderung auf ein ganzes Jahr auszudehnen.

DEUTSCH PLUS in seiner bisherigen Version hat die Gruppe der Kinder ab 4;6 Jahren als Zielgruppe.

Weiterentwicklung zu DEUTSCH PLUS 4

Eine neue Version, DEUTSCH PLUS 4, soll als standardisiertes Verfahren besonderen Sprachförderbedarf bei Kindern im Alter ab vier Jahren identifizieren.

Ziel der Pilotierung

Inwieweit erfüllt die neue Version von DEUTSCH PLUS seine Funktion?

Genügt DEUTSCH PLUS den grundlegenden Qualitätsanforderungen?

1. Funktion und Aufbau von DEUTSCH PLUS 4

Funktion

Funktion der Sprachstandsfeststellung

Unter Unsicherheit die Wahrscheinlichkeit falscher Entscheidungen reduzieren.

		Entscheidung	
		Förderbedarf	kein Förderbedarf
Tatsächliche Situation	Förderbedarf	richtige Entscheidung (wahr positiv)	falsche Entscheidung (falsch negativ)
	kein Förderbedarf	falsche Entscheidung (falsch positiv)	richtige Entscheidung (wahr negativ)

1. Funktion und Aufbau von DEUTSCH PLUS 4

Anlass und Ziel der Pilotierung

Zu Folie 5:

Es gibt zwei Typen falscher Entscheidungen, die nicht gleich schwerwiegend sind: Falsch positiv, also ein Kind, obwohl es das eigentlich nicht benötigt, sprachlich gezielt zu fördern, ist i.d.R. für das Kind nicht weiter schädlich, auch wenn unnötig Ressourcen gebunden werden. Ein Kind hingegen nicht zu fördern, das sprachliche Defizite aufweist (falsch negativ) kann gravierend sein.

Zu Folie 7 bis 9:

Damit nicht alle Kinder die gesamte Sprachstandsfeststellung durchlaufen müssen und somit allen Beteiligten Aufwand zu ersparen, wurde ein modularer Aufbau für DEUTSCH PLUS gewählt, der mehrere Ausstiegsmöglichkeiten kennt, bei denen versucht wurde, den unterschiedlich bedeutsamen Fehlertypen Rechnung zu tragen.

1. Funktion und Aufbau von DEUTSCH PLUS 4

Modularer Aufbau von DEUTSCH PLUS 4

Modul A: SPRACHUMFELD

Modul B: WIE GUT BEHERRSCHT DAS KIND DIE DEUTSCHE SPRACHE?

1. Entscheidung

Spezieller Sprachförderbedarf: Abbruch

Weiter mit Modul C

Modul C: PASSIVER WORTSCHATZ

Grundlage: Einfache Gegenstände. Fragen nach Nomen und Verben.

Beispiele: „Zeige mir den Schlüssel?“ „Womit kannst du essen?“ (Löffel)

2. Entscheidung

Abbruch: Spezieller Sprachförderbedarf

Abbruch: Kein spezieller Sprachförderbedarf

Weiter mit Modul D

Modul D: AKTIVER WORTSCHATZ

Ereignisbild, anhand dessen gezielt nach Situationen gefragt wird.

Beispiel *Restaurant*: Was gibt es zu hier zu essen? Was isst du gern?

3. Entscheidung

Spezieller Sprachförderbedarf

Kein spezieller Sprachförderbedarf

1. Funktion und Aufbau von DEUTSCH PLUS 4

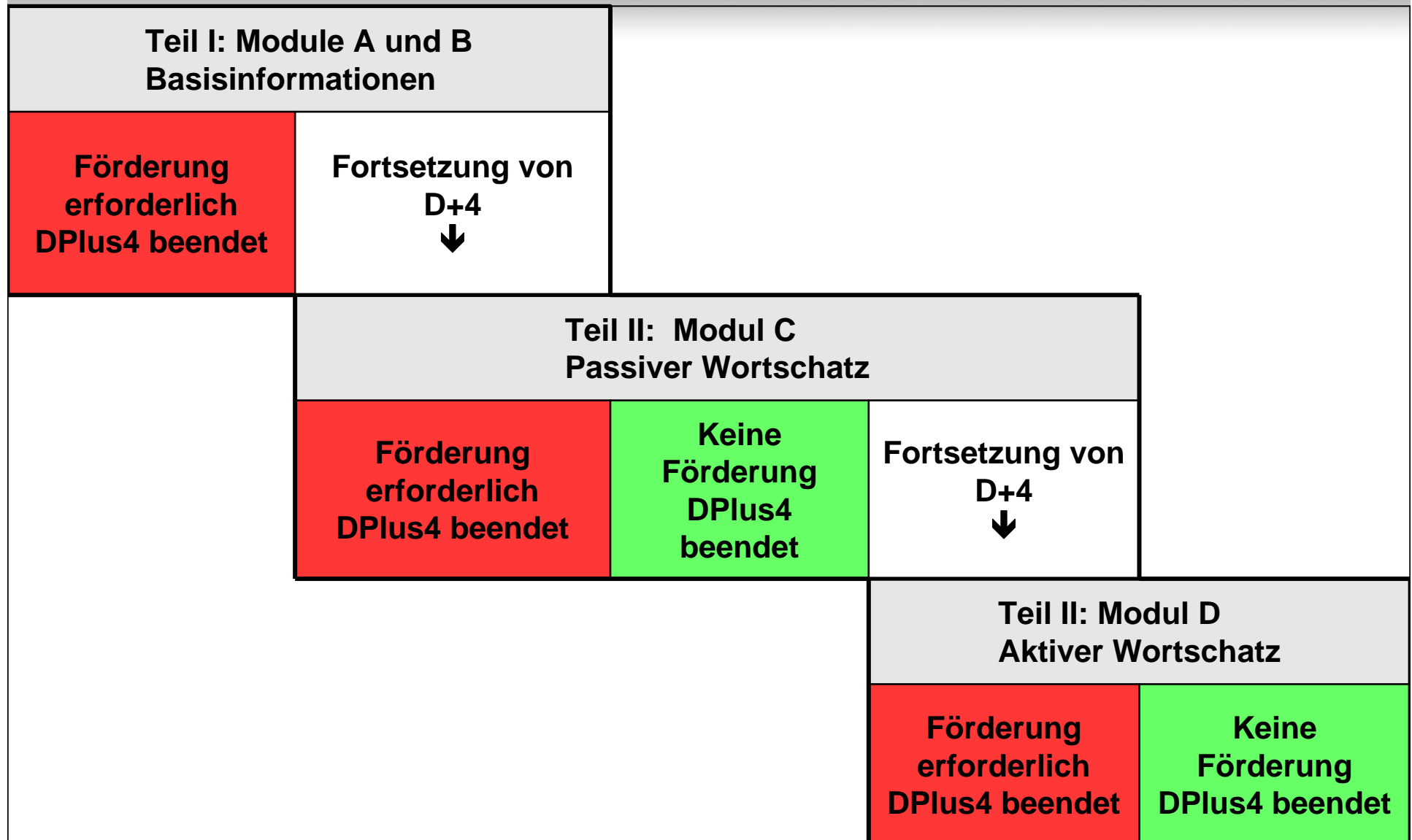
Modularer Aufbau von DEUTSCH PLUS 4: Modul D

Ereignisbild zum Modul D. Die Fragen zielen auf die markierten Situationen.



1. Funktion und Aufbau von DEUTSCH PLUS 4

Modularer Aufbau: Überblick



1. Funktion und Aufbau von DEUTSCH PLUS 4
2. Gütekriterien, Anlage der Untersuchung und Datenbasis
3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität
4. DEUTSCH PLUS 4: Ergebnisse zur Validität
5. DEUTSCH PLUS 4: Rückmeldungen der Erzieher/innen und Sonderpädagoginnen/innen
6. Zusammenhänge mit dem Bilanzbogen zum Sprachlerntagebuch
7. Erste Schlussfolgerungen

2. Gütekriterien, Anlage der Untersuchung und Datenbasis Gütekriterien: Eigenschaften standardisierter Tests

Was ist ein standardisierter Test?

- Testentwicklung basiert auf einer Testtheorie
- Systematische Auswahl von Aufgaben
- Hohes Maß an Standardisierung bzgl. Testdurchführung, Testauswertung, Interpretation
- Überprüfung der Testgütekriterien
- Häufig an großen Stichproben normiert

2. Gütekriterien, Anlage der Untersuchung und Datenbasis Gütekriterien: Eigenschaften standardisierter Tests

Zu Folie 11 – 14:

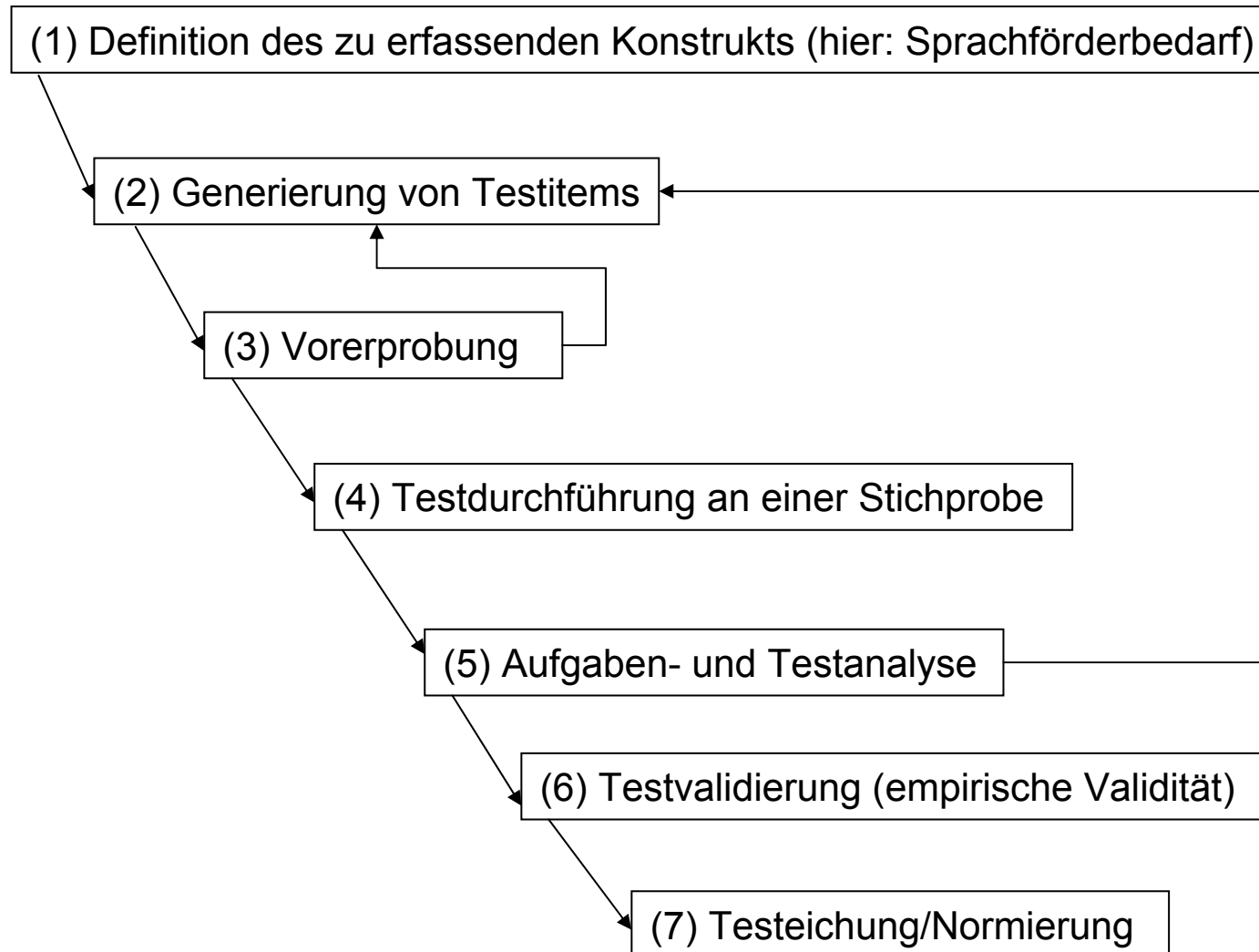
Der Aufwand, um ein aussagekräftiges Sprachstandsfeststellungsverfahren zu entwickeln, ist sehr hoch. Mehrere Etappen sind zu durchlaufen, die alle das Ziel eines Erhebungsinstrumentes haben,

- das aussagekräftige und zuverlässige Ergebnisse zum Sprachstand des Kindes macht
- unabhängig davon, wer das Verfahren durchführt.

Die Folien 11 sowie 13 und 14 skizzieren die Etappen der Testentwicklung, zu denen auch gehört, das neue Instrument auszuprobieren, also zu pilotieren. Die Pilotierung, hier von DEUTSCH PLUS 4, umfasst dabei nur einen Teil der Testentwicklung. Das Instrument wurde von einer Entwicklergruppe unter Beteiligung des LISUMs erarbeitet und zum Pilotieren der FU und dem ISQ übergeben.

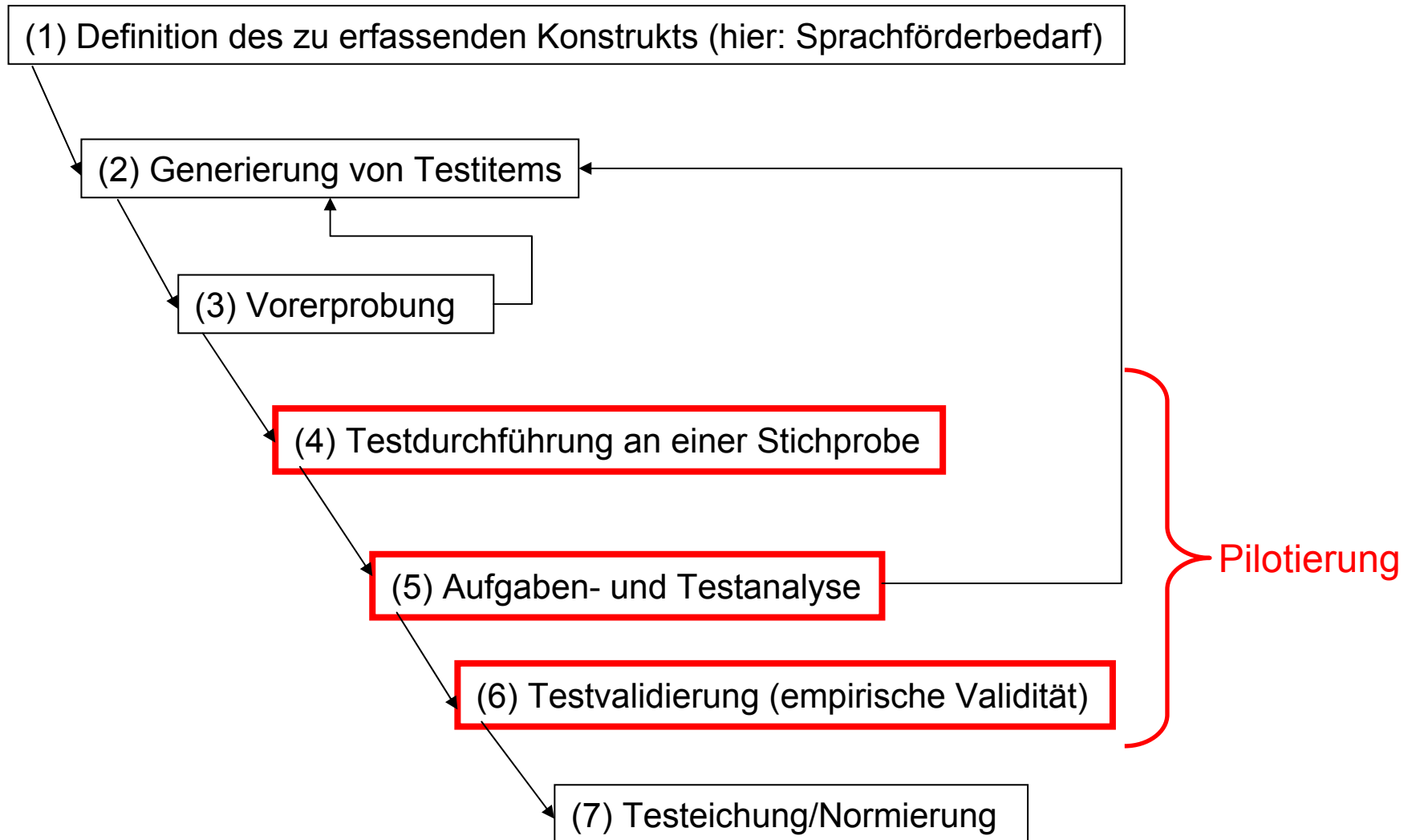
2. Gütekriterien, Anlage der Untersuchung und Datenbasis

Gütekriterien: Schritte zur Konstruktion eines standardisierten Tests



2. Gütekriterien, Anlage der Untersuchung und Datenbasis

Gütekriterien: Schritte zur Konstruktion eines standardisierten Tests



2. Gütekriterien, Anlage der Untersuchung und Datenbasis Gütekriterien und ihre Überprüfung

1. Objektivität

Sind die Ergebnisse unabhängig von der untersuchenden Person? Kommen zwei Untersucher bei einem Kind zu denselben Entscheidungen?

- Beurteilung der sprachlichen Äußerungen jedes Kindes in DEUTSCH PLUS durch Erzieherin und Sonderpädagogin. Überprüfung der Übereinstimmung zwischen den Urteilen

2. Reliabilität

Sind die Messungen präzise?

- Interne Konsistenz

3. Validität

Misst das Instrument das, was es messen soll, nämlich Förderbedarf?

- Überprüfung der Übereinstimmung zwischen den Ergebnissen von DEUTSCH PLUS und den Ergebnissen anderer Indikatoren für Sprachförderbedarf
 - AWST-R: Aktiver Wortschatztest
 - SETK 3-5: Verstehen von Sätzen
 - SETK 3-5: Enkodierung semantischer Relationen (Satzproduktion)

2. Gütekriterien, Anlage der Untersuchung und Datenbasis Durchführung der Untersuchung

Zu Folie 15:

Aufgeführt werden die drei zentralen klassischen Testgütekriterien.

Zur Reliabilität (Zuverlässigkeit): Hier beschränkt sich die Analyse darauf zu prüfen, wie weit die einzelnen Bestandteile des Testverfahrens „zusammen passen“ und wie präzise sie messen. Eine weitergehende Prüfung hätte darin bestanden, denselben Test, z.B. nach einem Monat, noch einmal durchzuführen. Das geht hier nicht wg. des Erinnerungs(Übungs)effekts und wg. des hohen damit verbundenen Aufwands für alle Beteiligten.

Zu Folie 17 und 18:

Im Hinblick auf die verschiedenen Gütekriterien wurde eine Vielzahl von Erhebungsinstrumenten eingesetzt, um die Ergebnisse von DEUTSCH PLUS 4 mit den Werten anderer Instrumente vergleichen zu können. Darüber hinaus erlaubt das breite Spektrum an Instrumenten, die Ergebnisse in größere Zusammenhänge (z.B. durch den Elternfragebogen) einbetten und somit besser verstehen und interpretieren zu können.

Durchführung:

- Erprobung von DEUTSCH PLUS 4 zunächst für die größte Gruppe von Kindern, die mit DEUTSCH PLUS 4 getestet werden sollte (also Kinder in Kitas)
- Auswahl einer Zufallsstichprobe von Kitas
- Auswahl einer Zufallsstichprobe von etwa 10 Kindern in jeder Kita
- Durchführung von DEUTSCH PLUS 4 durch Erzieher/innen unter Anwesenheit von Sonderpädagogen/innen
- Durchführung der Parallelinstrumente durch die Sonderpädagogen/innen an einem zweiten Testtag

2. Gütekriterien, Anlage der Untersuchung und Datenbasis

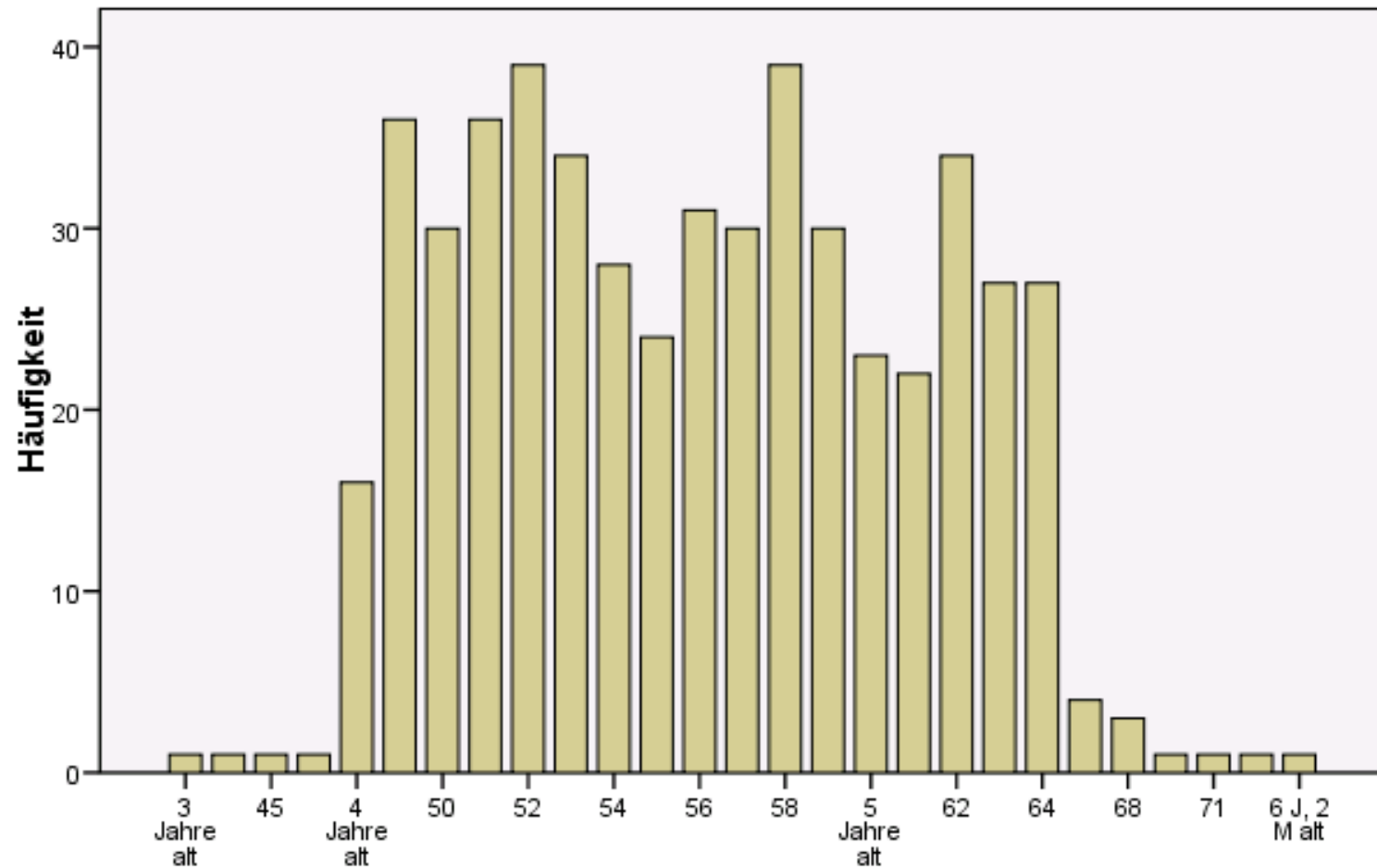
Anlage der Untersuchung: Erhebungsinstrumente

Erhebungsinstrument	Erz	SP
Fragebogen zum sprachbezogenen Verhalten	X	
DEUTSCH PLUS 4	X	X (C und D)
Testablaufprotokoll DEUTSCH PLUS 4	X	X
Parallelinstrumentarium (AWST, SETK3-5/SV und SB)		X
Testablaufprotokoll Parallelinstrumentarium		X
Bilanzbogen zum Sprachlerntagebuch	X	
Fazitbogen zur Sprachstandsfeststellung DEUTSCH PLUS 4	X	X

2. Gütekriterien, Anlage der Untersuchung und Datenbasis Datenbasis

N = 53 Kitas
N = 631 Kinder

Erzieher/innen-Angabe: Kindesalter in Monaten



Erzieher/innen-Angabe: Kindesalter in Monaten

2. Gütekriterien, Anlage der Untersuchung und Datenbasis Durchführung der Untersuchung

Zu Folie 19 und 21:

Die Stichprobe fokussiert, wie gewünscht, die Zielgruppe von DEUTSCH PLUS 4, nämlich die Kinder im Alter von 4 Jahren bis 5½-Jahren. Folie 19 zeigt die Verteilung der Kinder nach Alter: Der „Schwerpunkt“ der Verteilung liegt bei 56 Monaten: Der Mittelwert beträgt 56 Monate, wie Folie 21 zeigt.

Folie 21 listet einige Charakteristika der Stichprobe auf. Neben dem Mittelwert (MW) wird dort ebenfalls die Standardabweichung (s) aufgeführt, die die „Breite“ der Verteilung quantifiziert, wie weit also die Verteilung sich auf den Mittelwert konzentriert (hohe Homogenität, niedriges s) oder sich von ihm gewissermaßen entfernt (hohe Heterogenität, hohes s).

Folie 21 zeigt das immer vorhandene Phänomen bei empirischen Erhebungen, dass nämlich nicht von allen Kindern alle Daten vollständig vorliegen; beachte die unterschiedlichen Fallzahlen N.

2. Gütekriterien, Anlage der Untersuchung und Datenbasis Datenbasis

Zusammensetzung der Stichprobe

Alter	N=521, MW = 56 Monaten, s = 5 Monate
Geschlecht	N=533, 50% Mädchen, 50% Jungen
Herkunftssprache	N=535, 69% dH
Verkehrssprache zu Hause	N=507, 75% dV, 10% tV, 4% rV, 11% aV
Verweildauer in der Kita	N=436, MW = 30 Monate, s = 12 Monate
Dem/der Erzieher/in bekannt seit	N=451, MW = 23 Monate, s = 14 Monate

1. Funktion und Aufbau von DEUTSCH PLUS 4
2. Gütekriterien, Anlage der Untersuchung und Datenbasis
3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität
4. DEUTSCH PLUS 4: Ergebnisse zur Validität
5. DEUTSCH PLUS 4: Rückmeldungen der Erzieher/innen und Sonderpädagoginnen/innen
6. Zusammenhänge mit dem Bilanzbogen zum Sprachlerntagebuch
7. Erste Schlussfolgerungen

3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität

a) Ergebnisse zur Objektivität: Gesamtergebnis

	Am Ende des ersten Teils (Modul A und B) (N=535)	Am Ende des Moduls C (N=530)	Am Ende des Moduls D (N=525)
Förderung erforderlich	7%	15%	21%
Keine Entscheidung: Weitertesten	93% → 8% ↗ → 14% → → 71% ↘	14% → 6% ↗ → 8% ↘	-----
Förderung nicht erforderlich	-----	71%	79%

3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität

a) Ergebnisse zur Objektivität: Gesamtergebnis

Zu Folie 23:

Zum Gesamtergebnis: Werden die von DEUTSCH PLUS 4 vorgegebenen Regeln eingehalten, so kann bei 7% der Kinder bereits nach Teil I (Modul A und B) mit dem Befund *Sprachförderbedarf* abgebrochen werden. Nach dem Modul C gibt es drei Entscheidungsmöglichkeiten. Für weitere 8% der Kinder kann hier mit dem Befund *Sprachförderbedarf*, für 71% mit dem Ergebnis *Kein Sprachförderbedarf* abgebrochen werden. 14% der Kinder durchlaufen das Modul D: Weitere 6% erhalten die Diagnose *Sprachförderbedarf*, 8% *Kein Sprachförderbedarf*, so dass die Sprachförderquote insgesamt 21% beträgt, demnach identisch mit dem Ergebnis der letzten Sprachstandsfeststellung in Berlin ist.

Zu Folie 25:

Auswertungsobjektivität: Die Sprachstandsfeststellung DEUTSCH PLUS 4 kennt eine Reihe von Auswertungsschritten, die mehr oder weniger fehleranfällig sein können. Vergleicht man nunmehr das Gesamtergebnis so, wie die Erzieher/innen es angaben, mit dem Gesamtergebnis, das sich ausgehend von den Einzelangaben der Erzieher/innen nachrechnen lässt, so ergibt sich insgesamt eine sehr hohe Übereinstimmung von 95%. Hier nicht dokumentiert sind die hohen Übereinstimmungen, die die Auswertungen der Erzieher/innen und die der Sonderpädagogen/innen aufweisen.

3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität

a) Ergebnisse zur Objektivität: Gesamtergebnis

N = 375 Errechnet → Erz-Angaben ↓	Sprachförderung nicht erforderlich	Sprachförderung erforderlich
Sprachförderung nicht erforderlich	75%	2%
Sprachförderung erforderlich	3%	20%

3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität

b) Ergebnisse zur Reliabilität: Konsistenzkoeffizient α

Modul	n (Items)	N (Fälle)	α
Teil I: Sprachkenntnisse/Basisinformationen	12	535	.83
A: Sprachumfeld	2	535	.92
B: Sprachbeherrschung	10	535	.83
C: Passiver Wortschatz	15	532	.70
C-1: Richtig benannte Nomen	10	532	.56
C-2: Richtig benannte Verben	5	532	.77
D-11: Anzahl aller gebildeten Sätze	7	481	.86
D-12: Anzahl der Sätze mit 4 Wörtern u.m.	7	496	.86
D-21: Anzahl aller genannten Nomen	7	499	.84
D-22: Anzahl aller korrekten Nomen	7	499	.85
D-31: Anzahl aller genannten Verben	7	499	.90
D-32: Anzahl aller korrekten Verben	7	499	.90

3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität

a) Ergebnisse zur Objektivität: Gesamtergebnis

Zu Folie 26:

Interne Konsistenz: Ein wichtiger Aspekt der Reliabilität ist die interne Konsistenz, die sich durch den Koeffizienten α quantifizieren lässt. Die für die Komponenten von DEUTSCH PLUS 4 ermittelten Konsistenzkoeffizienten sind bis auf jenen des Moduls C-1 hoch bis sehr hoch.

Zu Folie 28:

Interne Konsistenz: Widersprechen sich die Ergebnisse der einzelnen Module?

Um diese Frage überprüfen zu können, durchliefen in der Pilotierung unabhängig von den eigentlichen Regeln von DEUTSCH PLUS 4 alle Kinder die gesamte Sprachstandsfeststellung. Dabei stellte sich heraus, dass insbesondere Modul C und Modul D in beachtlichem Ausmaße und in gravierender Richtung voneinander abweichen: Bei 16% der getesteten Kinder wäre an sich nach dem Modul C mit dem Befund *Kein Sprachförderbedarf* abgebrochen worden, wohingegen Modul D auf einen *Sprachförderbedarf* erkannte.

3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität

b) Ergebnisse zur Reliabilität: Widersprüchliche Modulergebnisse

	C/Abbruch: Förderung nicht erforderlich	D Förderung nicht erforderlich	D Förderung erforderlich
Teil I (Modul A und B) Abbruch: Förderung erforderlich	4% (N=530)	2% (N=479)	-----
Modul C Abbruch: Förderung erforderlich	-----	5% (N=478)	-----
Modul C Abbruch: Förderung nicht erforderlich	-----	-----	16% (N=478)

1. Funktion und Aufbau von DEUTSCH PLUS 4
2. Gütekriterien, Anlage der Untersuchung und Datenbasis
3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität
4. DEUTSCH PLUS 4: Ergebnisse zur Validität
5. DEUTSCH PLUS 4: Rückmeldungen der Erzieher/innen und Sonderpädagoginnen/innen
6. Zusammenhänge mit dem Bilanzbogen zum Sprachlerntagebuch
7. Erste Schlussfolgerungen

4. DEUTSCH PLUS 4: Ergebnisse zur Validität

a) Zusammenhänge mit dem Parallelinstrumentarium

Deutsch Plus 4	Alter (in Monaten)	Geschlecht	Wortschatz (AWST-R)	Satzverstehen (SETK 3-5)	Satzbildung (SETK 3-5)
Modul A Sprachumfeld	-.05	-.03	.65* (.69)	.50* (.56)	.51* (.57)
Modul B Sprachbeherrschung	.09	-.05	.52* (.58)	.51* (.61)	.49* (.59)
Modul C Passiver Wortschatz	.23*	.05	.51* (.61)	.53* (.67)	.46* (.60)
Modul D Aktiver Wortschatz	.14*	-.01	.44* (.53)	.36* (.45)	.41* (.52)
Gesamt-Rohwert	.17*	-.03	.71* (.79)	.65* (.75)	.65* (.77)

Anmerkungen. * $p < .01$, reliabilitätskorrigierte Koeffizienten in Klammern.

4. DEUTSCH PLUS 4: Ergebnisse zur Validität

a) Zusammenhänge

Zu Folie 30:

- es gibt einen geringen Alterszusammenhang mit den „Leistungs“-Modulen C und D von DEUTSCH PLUS 4
- DEUTSCH PLUS 4 scheint geschlechtsunspezifisch zu sein
- erkennbar ist, dass sich die höchsten Zusammenhänge (mit und ohne Korrektur) für den Gesamtrohwert von DEUTSCH PLUS 4 zeigen
- bemerkenswert ist, dass auch die „Fremdangaben“ (Sprachumfeld, Sprachbeherrschung) bereits substantielle Zusammenhänge zu den Außenkriterien zeigen

Zu Folie 32:

- es zeigen sich insgesamt moderate Zusammenhänge zwischen DEUTSCH PLUS 4 und dem Kontextbogen zum sprachbezogenen Verhalten
- Schüchternheit des Kindes wirkt sich insbesondere auf die im Modul B erfassten Erzieherinnenangaben aus, d.h. in diesem Fremdrating der Sprachbeherrschung schneiden schüchterne Kinder schlechter ab; möglicherweise werden im Modul B auch Aspekte des Sozialverhaltens erfasst.

4. DEUTSCH PLUS 4: Ergebnisse zur Validität

a) Zusammenhänge mit dem Kontextfragebogen

Kontextfragebogen zum sprachbezogenen Verhalten

Deutsch Plus 4	Schüchternheit	Sprechfreude	Sprachliche Kompetenz
Modul A Sprachumfeld	-.10 (-.12)	.27* (.30)	.42* (.46)
Modul B Sprachbeherrschung	-.40* (-.50)	.60* (.71)	.66* (.78)
Modul C Passiver Wortschatz	-.15* (-.20)	.28* (.35)	.42* (.52)
Modul D Aktiver Wortschatz	-.20* (-.27)	.30* (.37)	.31* (.39)
Gesamtrohwert	-.35* (-.44)	.55* (.64)	.64* (.74)

Anmerkungen. * $p < .01$, reliabilitätskorrigierte Koeffizienten in Klammern.

4. DEUTSCH PLUS 4: Ergebnisse zur Validität

b) Klassifikatorische Vorhersage

Frage: Wie weit stimmen die Entscheidungen von DEUTSCH PLUS 4 mit Entscheidungen anhand von anderen Verfahren überein?

Möglich sind Abweichungen in zwei Richtungen.

		Entscheidung auf der Grundlage der DEUTSCH PLUS 4 Ergebnisse	
		Sprachförderung erforderlich	Sprachförderung nicht erforderlich
Entscheidung aufgrund eines anderen Kriteriums (Außenkriterium)	sprachliche Leistung auffällig	wahr positiv (WP) Zugehörigkeit zur Sprachfördergruppe zutreffend angenommen	falsch negativ (FN) Zugehörigkeit zur Sprachfördergruppe fälschlich abgelehnt
	sprachliche Leistung nicht auffällig	falsch positiv (FP) Zugehörigkeit zur Sprachfördergruppe fälschlich angenommen	wahr negative (WN) Zugehörigkeit zur Sprachfördergruppe zutreffend abgelehnt

4. DEUTSCH PLUS 4: Ergebnisse zur Validität

b) Klassifikatorische Vorhersage

Zu Folie 33 bis 37:

- während sich die Korrelationskoeffizienten auf den gesamten Variationsbereich der Merkmale beziehen, sagen die klassifikatorischen Analysen etwas über die Qualität getroffener Entscheidungen aus
- bei einer einfachen Entscheidung mit DEUTSCH PLUS 4 lassen sich zwei grundlegende Konsequenzen unterscheiden: eine Sprachförderung ist erforderlich oder sie ist nicht erforderlich
- ein zur Validierung der Entscheidung hinzugezogenes Außenkriterium (welches aber nur im idealen Fall etwas über den realen Zustand der Kinder aussagt) lässt ebenfalls zwei Fälle entscheiden: Sprachförderung erforderlich oder nicht erforderlich
- somit ergeben sich vier Entscheidungssituationen mit ihren eigenen Fehlerwahrscheinlichkeiten: zwei dieser Entscheidungen sind richtig (wahr positiv und wahr negativ), zwei sind fehlerhaft (falsch negativ und falsch positiv)
- zur Beurteilung der Zuordnungsqualität und somit auch der Qualität des testdiagnostischen Verfahrens lassen sich verschiedene Koeffizienten bestimmen, welche die Zellhäufigkeiten und die Randsummen zueinander in Beziehung setzen:
 - Sensitivität: Anteil der positiven Sprachförderurteile (erforderlich) an der Gesamtzahl identifizierter Sprachförderfälle des Außenkriteriums; ein „gutes“ Screening-Instrument zeichnet sich auch durch eine hohe Sensitivität aus
 - Spezifität: Anteil der negativen Sprachförderurteile (nicht erforderlich) an der Gesamtzahl identifizierter Sprachförderfälle des Außenkriteriums
 - RATZ-Index: kombiniertes Maß; relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote unter Berücksichtigung der maximalen Trefferquote (gute Klassifikation: $RATZ > .33$, sehr gute Klassifikation: $RATZ > .66$)

4. DEUTSCH PLUS 4: Ergebnisse zur Validität

b) Klassifikatorische Vorhersage

		Entscheidung auf der Grundlage des 1. Abbruchkriteriums von DEUTSCH PLUS 4	
		Sprachförderung erforderlich	Sprachförderung nicht erforderlich (bzw. noch unklar)
Entscheidung aufgrund des AWST-R Gesamt- rohwerthes	sprachliche Leistung auffällig	wahr positiv (WP) 3%	falsch negativ (FN) 4%
	sprachliche Leistung nicht auffällig	falsch positiv (FP) 4%	wahr negativ (WN) 90%

N = 528

Sensitivität = .43, Spezifität = .96, RATZ-Index = .39

4. DEUTSCH PLUS 4: Ergebnisse zur Validität

b) Klassifikatorische Vorhersage

		Entscheidung auf der Grundlage des 2. Abbruchkriteriums von DEUTSCH PLUS 4	
		Sprachförderung erforderlich	Sprachförderung nicht erforderlich (bzw. noch unklar)
Entscheidung aufgrund des AWST-R Gesamt- rohwerthes	sprachliche Leistung auffällig	wahr positiv (WP) 6%	falsch negativ (FN) 7%
	sprachliche Leistung nicht auffällig	falsch positiv (FP) 5%	wahr negativ (WN) 82%

N = 524

Sensitivität = .45, Spezifität = .94, RATZ-Index = .44

4. DEUTSCH PLUS 4: Ergebnisse zur Validität

b) Klassifikatorische Vorhersage

		Entscheidung auf der Grundlage des Endergebnisses von DEUTSCH PLUS 4	
		Sprachförderung erforderlich	Sprachförderung nicht erforderlich (bzw. noch unklar)
Entscheidung aufgrund des AWST-R Gesamtrohwertes	sprachliche Leistung auffällig	wahr positiv (WP) 13%	falsch negativ (FN) 8%
	sprachliche Leistung nicht auffällig	falsch positiv (FP) 7%	wahr negativ (WN) 72%

N = 519

Sensitivität = .62, Spezifität = .91, RATZ-Index = .54

1. Funktion und Aufbau von DEUTSCH PLUS 4
2. Gütekriterien, Anlage der Untersuchung und Datenbasis
3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität
4. DEUTSCH PLUS 4: Ergebnisse zur Validität
5. DEUTSCH PLUS 4: Rückmeldungen der Erzieher/innen und Sonderpädagoginnen/innen
6. Zusammenhänge mit dem Bilanzbogen zum Sprachlerntagebuch
7. Erste Schlussfolgerungen

5. DEUTSCH PLUS 4: Rückmeldungen der Erzieher/innen und der Sonderpädagogen/innen (Fazitbogen)

	Mittelwert
FAZIT_01 Die Durchführung ist verständlich erklärt	2,6
FAZIT_02 Die Auswertung ist verständlich erklärt	2,1
FAZIT_03 Die Materialien sind geeignet, den passiven Wortschatz zu erfassen	2,0
FAZIT_04 Das Ereignisbild ist geeignet, den aktiven Wortschatz zu erfassen	1,5
FAZIT_05 Das erste Abbruchkriterium (nach Teil IB) ist an dieser Stelle sinnvoll platziert	1,7
FAZIT_06 Das erste Abbruchkriterium ist verständlich erklärt	2,6
FAZIT_07 Das zweite Abbruchkriterium (nach Teil IIC) ist an dieser Stelle sinnvoll platziert	1,6
FAZIT_08 Das zweite Abbruchkriterium ist verständlich erklärt	2,5
FAZIT_09 Die Anleitung zum Instrument DEUTSCH PLUS 4 (im Manual) ist hilfreich	2,3
FAZIT_10 DEUTSCH PLUS 4 macht den Kindern Freude	1,7
FAZIT_11 Fazit: DEUTSCH PLUS 4 ist insgesamt geeignet, den Sprachstand der Kinder zu erfassen	1,4

5. DEUTSCH PLUS 4: Rückmeldungen der Erzieher/innen und der Sonderpädagogen/innen (Fazitbogen)

Zwei Ergebnisse

1. Fast jeder der einzelnen Aspekte wird überdurchschnittlich bewertet, aber insgesamt liegt die Bewertung genau auf dem mittleren Niveau.
2. Die Einschätzungen der Sonderpädagoginnen sind generell skeptischer als die der Erzieherinnen, aber keine der Mittelwertsunterschiede sind signifikant bis auf zwei Ausnahmen:

5. DEUTSCH PLUS 4: Rückmeldungen der Erzieher/innen und der Sonderpädagogen/innen (Fazitbogen)

	MW: Erz	MW: SP
01 Die Durchführung ist verständlich erklärt	2,6	2,5
02 Die Auswertung ist verständlich erklärt	2,2	1,8
03 Die Materialien sind geeignet, den passiven Wortschatz zu erfassen	2,1	1,9
04 Das Ereignisbild ist geeignet, den aktiven Wortschatz zu erfassen	1,5	1,4
05 Das erste Abbruchkriterium (nach Teil IB) ist an dieser Stelle sinnvoll platziert	1,7	1,7
06 Das erste Abbruchkriterium ist verständlich erklärt	2,5	2,6
07 Das zweite Abbruchkriterium (nach Teil II-C) ist an dieser Stelle sinnvoll platziert	1,6	1,5
08 Das zweite Abbruchkriterium ist verständlich erklärt	2,5	2,6
09 Die Anleitung zum Instrument DEUTSCH PLUS 4 (im Manual) ist hilfreich	2,4	2,2
10 DEUTSCH PLUS 4 macht den Kindern Freude	1,8	1,7
11 Fazit: DEUTSCH PLUS 4 ist insgesamt geeignet, den Sprachstand der Kinder zu erfassen	1,5	1,2

1. Funktion und Aufbau von DEUTSCH PLUS 4
2. Gütekriterien, Anlage der Untersuchung und Datenbasis
3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität
4. DEUTSCH PLUS 4: Ergebnisse zur Validität
5. DEUTSCH PLUS 4: Rückmeldungen der Erzieher/innen und Sonderpädagoginnen/innen
6. Zusammenhänge mit dem Bilanzbogen zum Sprachlerntagebuch
7. Erste Schlussfolgerungen

6. Zusammenhänge mit dem Bilanzbogen zum Sprachlerntagebuch

Bilanzbogen und Parallelinstrumentarium

Bilanzbogen zum Sprachlerntagebuch	Alter (in Monaten)	Geschlecht	Wortschatz (AWST-R)	Satzverstehen (SETK 3-5)	Satzbildung (SETK 3-5)
Basales	.25*	-.15*	.33* (.36)	.38* (.44)	.32* (.37)
Phonologische Bewusstheit	.29*	-.08	.57* (.61)	.52* (.59)	.51* (.58)
Sprachhandeln	.13*	-.10	.51* (.53)	.53* (.58)	.51* (.56)
Erste Erfahrungen	.19*	-.04	.52* (.55)	.54* (.59)	.49* (.55)
Sprachstrukturen	.11	-.02	.55* (.57)	.51* (.56)	.53* (.59)
Förderurteil der Erzieherinnen ¹	.07	-.08	.50* (.51)	.51* (.54)	.47* (.51)

Anmerkungen. * $p < .01$, 1) einfach, ansonsten doppelt reliabilitätskorrigierte Koeffizienten in Klammern.

Zu Folie 44:

- die einzelnen Bereiche des Bilanzbogens zum Sprachlerntagebuch zeigen geringe Zusammenhänge mit dem Alter und sind weitgehend geschlechtsunspezifisch
- der Bilanzbogen zeigt moderate korrelative Zusammenhänge zu den gewählten Außenkriterien AWST-R und den beiden Subskalen des SETK 3-5

Zu Folie 46:

- es zeigen sich insgesamt moderate Zusammenhänge zwischen den Subbereichen des Bilanzbogens und den drei Bereichen des Kontextbogens zum sprachbezogenen Verhalten
- erkennbar ist, dass eine Schüchternheit des Kindes offenbar mit der Einschätzung in allen Subbereichen des Bilanzbogens in Zusammenhang steht (dies ist ähnlich, aber sehr viel deutlicher als zum Modul B von DEUTSCH PLUS 4); möglicherweise werden im Bilanzbogen auch sehr viele Aspekte des kindlichen Sozialverhaltens erfasst.

6. Zusammenhänge mit dem Bilanzbogen zum Sprachlerntagebuch Bilanzbogen und sprachbezogenes Verhalten

Kontextfragebogen zum sprachbezogenen Verhalten			
Bilanzbogen zum Sprachlerntagebuch	Schüchternheit	Sprechfreude	Sprachliche Kompetenz
Basales	-.30* (-.37)	.33* (.37)	.43* (.48)
Phonologische Bewusstheit	-.32* (-.39)	.42* (.47)	.53* (.59)
Sprachhandeln	-.42* (-.49)	.62* (.67)	.72* (.78)
Erste Erfahrungen	-.35* (-.41)	.51* (.56)	.59* (.65)
Sprachstrukturen	-.29* (-.34)	.45* (.50)	.68* (.75)
Förderurteil der Erzieherinnen ¹	-.26* (-.30)	.43* (.46)	.63* (.67)

Anmerkungen. * $p < .01$, 1) einfach, ansonsten doppelt reliabilitätskorrigierte Koeffizienten in Klammern.

6. Zusammenhänge mit dem Bilanzbogen zum Sprachlerntagebuch

Gesamtergebnis DEUTSCH PLUS 4 und Erzieher/innenurteil

N = 484 DEUTSCH PLUS 4 → Erz-Angaben ↓	Sprachförderung erforderlich	Sprachförderung nicht erforderlich
Sprachförderung erforderlich	15%	18%
Sprachförderung nicht erforderlich	5%	63%

Zu Folie 47:

- die klassifikatorische Analyse zum Zusammenhang zwischen einer Sprachförder-Entscheidung durch DEUTSCH PLUS 4 und dem Votum der Erzieherinnen als Außenkriterium zeigt insbesondere im Bereich der falsch negativen Entscheidungen eine erhöhte Quote
- ca. 45% der Fälle, die die Erzieherinnen mit Sprachförderbedarf identifizieren, werden auch von DEUTSCH PLUS 4 entsprechend klassifiziert
- insgesamt identifizieren die Erzieherinnen eine höhere Quote an Kindern mit Sprachförderbedarf, sie sind im Sinne einer klassifikatorischen Analyse sensitiver, während die Vorteile von DEUTSCH PLUS 4 in der Spezifität des Verfahrens liegen

1. Funktion und Aufbau von DEUTSCH PLUS 4
2. Gütekriterien, Anlage der Untersuchung und Datenbasis
3. DEUTSCH PLUS 4: Ergebnisse zur Objektivität und Reliabilität
4. DEUTSCH PLUS 4: Ergebnisse zur Validität
5. DEUTSCH PLUS 4: Rückmeldungen der Erzieher/innen und Sonderpädagoginnen/innen
6. Zusammenhänge mit dem Bilanzbogen zum Sprachlerntagebuch
7. Erste Schlussfolgerungen

7. Erste Schlussfolgerungen

1. Zur Sprachförderquote: Sie liegt mit 21% auf dem bisherigen Niveau. Offenbar wurde mit der Änderung des Instruments die Schwelle entsprechend mitverschoben. Die Sprachförderquote ist in zu erwartender Weise abhängig von Herkunfts- und Verkehrssprache.
2. Zur Objektivität: Die verschiedenen Modi zum Gesamtergebnis zu gelangen (Erzieher/innen und Sonderpädagogen/innen-Angaben direkt und nachgerechnet), stimmen in hohem Maße überein.
3. Zur Reliabilität: Konsistenzkoeffizient α reicht bei den Teilskalen von niedrigen bis hohen Werten. Es gibt beachtenswerte Inkonsistenzen in den Ergebnissen der Module C und D.
4. Zur Validität: DEUTSCH PLUS 4 weist eine ausreichend gute empirische Validität auf. Bemerkenswert ist, dass auch die "Fremdangaben" (Sprachumfeld, Sprachbeherrschung) bereits substantielle Zusammenhänge zu den Parallelinstrumenten zeigen.

7. Erste Schlussfolgerungen

5. Zur Sensitivität: Die klassifikatorischen Analysen deuten darauf hin, dass DEUTSCH PLUS 4 weniger ein sensitives Instrument ist (was für ein Screeningverfahren wichtig wäre), als vielmehr eine gute Spezifität aufweist.
6. Ansätze für Weiterentwicklung: Die bisherigen Analysen deuten an, dass sich die Sensitivität verbessert, wenn man nicht die Abbruchkriterien berücksichtigt, sondern das Verfahren komplett durchführt.
Alternativ oder zugleich sind die Items des Moduls C kritisch zu sichten, zumal es im Modul C einen Deckeneffekt gibt. Die Schwelle für das Urteil "Keine Förderung erforderlich" müsste erhöht werden, um die Sensitivität zu verbessern.
7. Zum Bilanzbogen des Sprachlerntagebuchs: Wie DEUTSCH PLUS 4 weist der Bilanzbogen substantielle Zusammenhänge mit dem Parallelinstrumentarium auf.
Die Zusammenhänge des Bilanzbogens mit den Merkmalen aus dem Kontextfragebogen zum sprachbezogenen Verhalten sind etwas enger als dies bei DEUTSCH PLUS 4 der Fall ist. Dies ist ein Hinweis darauf, dass in das Erzieherinnenurteil in höherem Maße Sozialverhalten eingeht als dies bei dem reaktiven Verfahren DEUTSCH PLUS 4 der Fall ist.

ISQ

Otto-von-Simson-Str. 15
14195 Berlin

Tel (030) 844 166 80
Fax (030) 844 166 10

info@isq-bb.de

www.isq-bb.de

FU/Empirische Bildungsforschung

Habelschwerdter Allee 45
14195 Berlin

(030) 838 55 981
(030) 838 55 986

empbild@zedat.fu-berlin.de

www.ewi-psy.fu-berlin.de